Sio-Iong Ao · Haeng Kon Kim
Mahyar A. Amouzegar *Editors*

Transactions on Engineering Technologies

World Congress on Engineering and Computer Science 2017

Springer

Transactions on Engineering Technologies

Sio-Iong Ao · Haeng Kon Kim
Mahyar A. Amouzegar
Editors

Transactions on Engineering Technologies

World Congress on Engineering and Computer Science 2017

Springer

Editors
Sio-Iong Ao
International Association of Engineers
Hong Kong, Hong Kong

Mahyar A. Amouzegar
University of New Orleans
New Orleans, LA, USA

Haeng Kon Kim
Department of Computer and Communication
Daegu Catholic University
Daegu, Korea (Republic of)

ISBN 978-981-13-2190-0 ISBN 978-981-13-2191-7 (eBook)
https://doi.org/10.1007/978-981-13-2191-7

Library of Congress Control Number: 2018943704

© Springer Nature Singapore Pte Ltd. 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

A large international conference on Advances in Engineering Technologies and Physical Science was held in San Francisco, California, USA, October 25–27, 2017, under the auspices of the World Congress on Engineering and Computer Science (WCECS 2017). The WCECS 2017 is organized by the International Association of Engineers (IAENG). IAENG, originally founded in 1968, is a nonprofit international association for the engineers and the computer scientists. The WCECS Congress serves as an excellent platform for the members of the engineering community to meet and exchange ideas. The Congress in its long history has found a right balance between theoretical and application development, which has attracted a diverse group of researchers, leading its rapid expansion. The conference committees have been formed with over 200 members including research center heads, deans, department heads/chairs, professors, and research scientists from over 30 countries. The full committee list is available at the Congress' website: www.iaeng.org/WCECS2017/committee.html. WCECS conference is truly an international meeting with a high level of participation from many countries. The response to WCECS 2017 conference call for papers was outstanding, with more than five hundred manuscript submissions. All papers went through rigorous peer review process, and the overall acceptance rate was 50.39%.

This volume contains 27 revised and extended research articles, written by prominent researchers, participating in the Congress. Topics include engineering mathematics, electrical engineering, communications systems, computer science, chemical engineering, systems engineering, manufacture engineering, and industrial applications. This book offers the state of the art of tremendous advances in engineering technologies and physical science and applications; it also serves as an exceptional source of reference for researchers and graduate students working with/on engineering technologies and physical science and applications.

Hong Kong, Hong Kong — Sio-Iong Ao
Daegu, Korea (Republic of) — Haeng Kon Kim
New Orleans, USA — Mahyar A. Amouzegar

Contents

Chapter 1
Classical Young Measures Generated by Oscillating Sequences with Uniform Representation

Andrzej Z. Grzybowski and Piotr Puchała

Abstract The paper is devoted to the theory of classical Young measures. It focuses on the situation where a sequence of rapidly oscillating functions has uniform representation in a sense that is defined in this article. There is stated a proposision characterizing the Young measures generated by such a class of sequences. This characterization enables one to find an explicit formulae for the density functions of these generated measures as well as the computations of the values of the related Young functionals. Examples of possible applications of the new results are presented as well.

Keywords Classical Young measures · Non-convex optimization · Numerical computations · Oscillating sequences · Uniform distribution · Uniform representation

1.1 Introduction

Non-convex optimization problems are at the core of various contemporary engineering applications. They arise e.g. in optimal control, nonlinear evolution equations, variational calculus, micromagnetic phenomena in ferromagnetic materials as well as in microstructures theory in continuum mechanics. It appears hovewer that the optimization problems may not possess clasical minimizers especially when elements of the minimizing sequences oscillate rapidly. The following example, attributed to

A. Z. Grzybowski (✉) · P. Puchała
Faculty of Mechanical Engineering and Computer Science, Institue of Mathematics, Czestochowa University of Technology, al. Armii Krajowej 21, 42-201 Częstochowa, Poland
e-mail: andrzej.grzybowski@im.pcz.pl

P. Puchała
e-mail: piotr.puchala@im.pcz.pl

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_1

Oscar Bolza and Laurence Chisholm Young, illustrates the nature of the problem, see [7].

Example 1 We are minimizing the integral functional $\mathscr{J}$ acting on a Sobolev space $W^{1,4}(0,1)$ of real functions defined on the unit interval in $\mathbb{R}$ and such that their first derivative belongs to the space $L^4(0,1)$. The functional $\mathscr{J}$ is of the form

$$\mathscr{J}(u) = \int_0^1 \left[u^2 + \left(\left(\frac{du}{dx}\right)^2 - 1\right)^2\right] dx.$$

We impose the boundary conditions $u(0) = 0 = u(1)$.

It can be proved that $\inf \mathscr{J} = 0$.

Consider the function

$$u(x) = \begin{cases} x, & \text{for } x \in \left[0, \frac{1}{4}\right) \\ \frac{1}{2} - x, & \text{for } x \in \left[\frac{1}{4}, \frac{3}{4}\right) \\ x - 1, & \text{for } x \in \left[\frac{3}{4}, 1\right). \end{cases}$$

Then the sequence $u_n(x) := \frac{1}{n} u(nx)$ (divergent strongly but convergent weakly in $W^{1,4}(0,1)$) is the minimizing sequence for $\mathscr{J}$, that is it satisfies the condition $\mathscr{J}(u_n) \to \inf \mathscr{J}$. However, we have $\inf \mathscr{J} \neq \mathscr{J}(\lim u_n)$: if the limit u_0 of (u_n) were the function satisfying the equality $\inf \mathscr{J} = \mathscr{J}(u_0)$, it would have to satisfy the conditions: $u_0 \equiv 0$ and $\frac{du_0}{dx} = \pm 1$ a.e. (with respect to the Lebesgue measure on $[0,1]$), which is imposible. This means that $\mathscr{J}$ does not attain its infimum. It is seen, that the number of 'teeth' of u_n (with slope ± 1) grows with n, that is the elements of the minimizing sequence oscillate more and more 'wildly' around zero.

Such a behavior of the sequences requires a generalization of the notion of a solution for such problems. It often can be achieved by means of Young measures.

Young measures theory has a long history. It starts with the seminal work [12] of L. C. Young who introduced the notion (called by himself "generalized curves") to provide extended solutions for some non-convex problems in variational calculus. He developed these pioneering ideas in [13].

Nowadays we are provided with vast literature where the Young measures are defined under different assumptions about underlying spaces and analysed from different standpoints. However this paper focuses on the clasical Young measures related to sequences of rapidly oscilating functions.

This article is a revised and enlarged version of the talk [5] given by the first author during the World Congress on Engineering and Computer Science 2017 in San Francisco, USA. It is organized as follows. In the next section we introduce some preliminary definitions and results. In Sect. 1.3 we define some classes of fast-oscillating sequences and state new proposition that allows us to find explicit forms of the density functions of related classical Young measures in various situations. Section 1.4 presents some examples that illustrate the possible applications of the main result stated in Sect. 1.3. Finally we make some remarks about possible further extensions and applications.

1.2 Preliminary Definitions and Results

We now introduce basic notions of the Young measure theory from the point of view of nonlinear elasticity. Our presentation follows the approach taken in [8], where the reader is referred to for detailed information along with necessary notions from functional analysis and further bibliography. Another book treating Young measures thoroughly in the context the optimization theory and variational calculus is [11].

Let Ω be a nonempty, open and bounded subset of $\mathbf{R}^d$ with smooth boundary. Denote by $L^\infty(\Omega)$ the Banach space of essentially bounded functions defined on Ω with values in a compact set $K \subset \mathbf{R}^l$. Let (f_n) be a sequence of functions converging to some function f_0 weakly* in L^∞ and denote by φ a continuous real valued function with domain $\mathbf{R}^l$. By the continuity of φ the sequence $\varphi(f_n)$ is uniformly bounded in L^∞ norm and Banach-Alaoglu theorem yields the existence of the (not relabeled) subsequence such that $\varphi(f_n) \to g$ weakly* in L^∞. However, in general g is not $\varphi(f_0)$, moreover, it is not even a function with domain in $\mathbf{R}^l$. To quote from [8]:

> The Young measure associated with (f_n) furnishes the link among (f_n), f_0, g and φ.

Recall that a function $H : \mathbf{R}^d \times \mathbf{R}^l \to \mathbf{R} \cup \{\infty\}$ is called a *Carathéodory function* if it is measurable with respect to the first and continuous with respect to the second variable.

We now state the basic existence theorem for Young measures in its full generality.

Theorem 1 (see Theorem 2.2 in [8]) *Let $\Omega \subset \boldsymbol{R}^d$ be a measurable set and let $z_n : \Omega \to \boldsymbol{R}^l$ be measurable functions such that*

$$\sup_n \int_\Omega h(|z_n|)dx < \infty,$$

where $h : [0, \infty) \to [0, \infty)$ is a continuous, nondecreasing function such that $\lim_{t\to\infty} h(t) = \infty$. There exist a subsequence, not relabeled, and a family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ (the associated Young measure) depending measurably on x with the property that whenever the sequence $\big(H(x, z_n(x))\big)$ is weakly convergent in $L^1(\Omega)$ for any Carathéodory function $H(x, \lambda) : \Omega \times \boldsymbol{R}^l \to \boldsymbol{R} \cup \{\infty\}$, the weak limit is the function

$$\overline{H}(x) = \int_{\boldsymbol{R}^l} H(x, \lambda)d\nu_x(\lambda).$$

Definition 1 (i) the family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ of the above Theorem is called the *Young measure associated with the sequence* (z_n) (or *Young measure generated by the sequence* (z_n));

(ii) if the Young measure $\nu = \{\nu_x\}_{x\in\Omega}$ does not depend on $x \in \Omega$, then it is called *homogeneous Young measure* and is denoted merely by ν.

Remark 1 The homogeneous Young measures, although being the 'simplest examples of the species', are very important both in theory and in applications. For instance, the Young measures calculated in [7] as examples in micromagnetism or elasticity, are homogeneous ones.

One may also look at the Young measure as at object associated with *any* measurable function defined on a nonempty, open, bounded subset Ω of $\mathbf{R}^d$ with values in a compact subset K of $\mathbf{R}^l$. Such a conclusion can be derived from the Theorem 3.6.1 in [11]. Due to this theorem it can be proved that the Young measure associated with a simple function is homogeneous and is the convex combination of Dirac measures (such measures are called *discrete*). These Dirac measures are concentrated at the values of the simple function under consideration while coefficients of the convex combination are proportional to the Lebesgue measure of the sets on which the respective values are taken on by the function; see [9] for details and more general results concerning simple method of obtaining explicit form of Young measures associated with oscillating functions (similar, although mathematically more complicated situation, is met in elasticity when the deformed body has a laminate structure; see e.g. Sect. 4.6 in [7]).

This result gives rise to the use of computer tools in performing one of the main tasks when dealing with Young measures: calculating the explicit form of the Young measure. Generally this is very difficult or sometimes even impossible to perform. It has however turned out, that using the approach presented in [11] and the notion of a *quasi-Young measure* introduced in [9], the Monte Carlo simulation can be succesfully used to partial treatment of the mentioned task (apparently, quasi-Young measures associated with Borel functions are precisely *the* homogeneous Young measures associated with them, see [10] for the proof). The word 'partial' refers to the fact that the simulation has worked well for the particularly simple Young measure: the one associated with the continuous piecewise affine real function. In the example presented in [2] proposed computer routine passes the χ^2 Pearson test. Moreover, in the article [3] two more routines for Monte Carlo simulating of the Young measures associated with such functions has been presented and compared. The analysis of those algorithms has led to the much more general *theoretical* result: the characterisation of a Young measure associated with *any* Borel function. The main result stated there provides direct link between the Young measure concepts and the probability theory. Namely, the following theorem holds:

Theorem 2 (see [6]) *Let $f : \boldsymbol{R}^d \supset \Omega \to K \subset \boldsymbol{R}^l$ be a Borel function with Young measure ν. Then ν is the probability distribution of the random variable $Y = f(U)$, where U has a uniform distribution on Ω.*

The above theorem opens the possibility not only for broadening the class of functions whose Young measures can be calculated via Monte Carlo simulation but also for calculating the values of the classical (i.e. generated by sequences of rapidly oscillating functions) Young measures. The Monte Carlo method in the second opportunity has been done in [4]. The main topic of this article is to present yet another way.

Now, let us recall the notion of classical Young measure associated with a sequence of oscillating functions $\{f_k\}$, see e.g. [11].

Definition 2 The classical Young measure generated by the sequence $\{f_k\}$ is a family of probability measures $\nu = \{\nu_x\}_{x\in\Omega}$ satisfying the condition:

For any Carathéodory function H

$$\int_{\Omega} H(x, f_k(x))dx \overset{k\to\infty}{\longrightarrow} \int_{\Omega}\int_{K} H(x, y)d\nu_x(y)dx \tag{1.1}$$

The application that assignes to any Carathéodory function H the integral given on the right-hand side of the above equation is called *Young functional*. Its values on H will be denoted here sa YF(h), while the integrals on the left-hand side of (1.1) will be denoted as $C(f_k, H)$.

Basically, the above definition presents the original understanding of the Young measure, as introduced in his work [13]. These measures are of our main concern in this paper.

1.3 Main Results

This section contains main results of the article. We first recall the notion of the sequence whose elements are violently oscillating functions and then consider classical Young measures generated by such sequences. Finally we formulate proposition characterizing those Young measures.

1.3.1 Rapidly Oscillating Sequences with Uniform Representation

Let function $f : [a, b) \to K \subset \mathbf{R}$ be a Borel function defined on the interval $[a, b)$, $b > a$, and let $f^e : \mathbf{R} \to K$ be the periodic extension of f (with the period equal to $T = b - a$).

Let Ω be a given interval. A sequence $\{f_k\}$ of functions $f_k : \Omega \to K,\ k = 1, 2, ...$ defined by the formula

$$f_k(x) = f^e(kx),\ x \in \Omega \tag{1.2}$$

will be called a *Rapidly Oscillating Sequence with Uniform representation* f, and denoted as ROSU(f). In such a case we will also say that f generates rapidly oscillating sequence $\{f_k\}$. Note, that the interval Ω—the domain of elements of ROSU(f)—does not have to be the same as $[a, b)$, i.e. the domain of f.

Example 2 In this example we present illustrative plots of some elements of ROSU(f), where

$$f(x) = 2 - 2\sin(x),\ x \in [0, 3\pi/2) \tag{1.3}$$

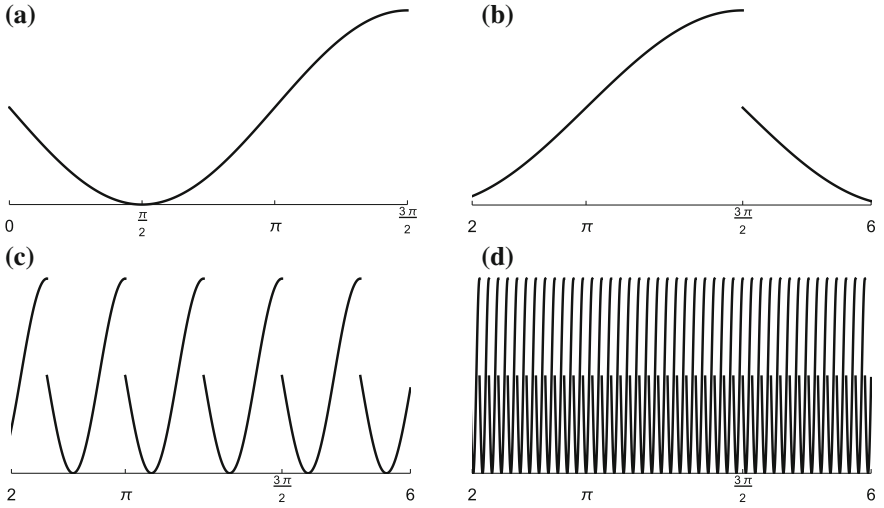


Fig. 1.1 A function f given by Eq. (1.3) and functions f_1, f_5 and f_{50} belonging to ROSU(f) with the domain $\Omega = [2, 6)$. Plots of the functions f, f_1, f_5 and f_{50} are labeled as **a**, **b**, **c** and **d**, respectively

Its purpose is not only to illustrate the behaviuor of ROSU (which is quite obvious, in fact) but also to illustrate the concept of classical Young measure associated with the sequence $\{f_k\}$. The plots of the function f given by 1.3 as well as of examplary elements of ROSU(f) with the domaine $\Omega = [2, 6)$ are presented in Fig.1.1. Namely it shows f, f_1, f_5 and f_{50}.

It can be easilly seen that the graphs of f_k are getting denser when k tends to infinity. Unfortunately, a conventional weak* cluster point of $\{f_k\}$ loses most of the information about the fast oscilations in $\{f_k\}$ because, in some sense, it takes into account only the mean values of $\{f_k\}$—as integrals do. That is why we need a new concept of the limit and here the theory of Young measures helps us. If ν_x is the Young measure associated with $\{f_k\}$ then, rougly speaking, for any measurable set $A \subset K$ the intuitive meaning of $\nu_x(A)$ is the probability that for an *infinitesimally* small neighbourhood S of $x \in \Omega$ and sufficiently large k's we can "find" $f_k(s)$ in A, when s changes within S. The "density" of the values in K can be observed in the plot (d) (for f_{50}) where the "more probable values" create darker straps in the figure.

1.3.2 Classical Young Measures Generated by the ROSU(f)

It results directly from the definition of ROSU(f) that its behaviour, as k tends to infinity, is exactly the same in every neighbourhood of any $x \in \Omega$. In other words, its asymptotic behaviour in an arbitrarily small interval $I \subset \Omega$ does not depend on where the interval is placed within the domain. Consequently, it is obvious that the

classical Young measure generated by the ROSU(f) is the homogeneous Young measure (i.e. it does not depend on $x \in \Omega$).

Now, let U_S denote a random variable uniformly distributed on S. Let us consider two functions $g_1 : \Omega_1 \to K$ and $g_2 : \Omega_2 \to K$. We will say that the two functions *identically transform a uniform distribution* if the distributions of the random variables $g_1(U_{\Omega_1})$ and $g_2(U_{\Omega_2})$, are the same. This fact will be denoted by $g_1 \approx g_2$. Obviously the "$\approx$" is the equivalence relation.

Note that if the ROSU(f) is defined on the same interval as the generating function f, then *any* of its elements transform a uniform distribution identically as the function f, i.e. $f_k \approx f$ for any $k = 1, 2, \ldots$. Now let us consider the case where the ROSU(f) is defined on interval Ω that is different than the domain $[a, b)$ of the generating function f. In such a case it can be seen that $f_k(U_\Omega) \xrightarrow{D} f(U_{[a,b)})$, where $\xrightarrow{D}$ denotes the *convergence in distribution*, see [1]. In other words the distribution of $f(U_{[a,b)})$ is a vague limit of the sequence of distributions of $f_k(U_\Omega)$. Indeed, for k's large enough so the length of interval Ω is greater than $(b - a)/k$ we have for *any* measurable subset $A \subset K$:

$$|P(f_k(U_\Omega) \in A) - P(f(U_{[a,b)}) \in A)| < 1/k$$

Finally, by Theorem 2 we know, that the distribution of $Y = f(U_{[a,b)})$ is the Young measure associated with the function f. Thus we can formulate the following result concerning classical Young measures.

Proposition 1 *Classical Young measure generated by ROSU(f) is the homogeneous Young measure. This measure is identical with the distribution of the random variable* $Y = f(U_{[a,b)})$.

1.4 Illustrative Examples

On the basis of the above general description of the classical Young measure generated by ROSU(f) we can obtain a number of rules which allow one to find an explict form of the classical Young measure in various specific cases. For example, let us consider the following situation.

Let $[a, b)$ be the interval—domain of the function f. Let us consider an open partition of $[a, b)$ into a number of open intervals $I_1, I_2, \ldots, I_n$ such that the intervals are pairwise disjoint and $\bigcup_i^n \overline{I}_i = [a, b]$, where $\overline{A}$ denotes the closure of the set A.

Let function f be continuously differentiable on each interval of the partition and let $f'(x) \neq 0$ for all $x \in \bigcup_{l=1}^n I_l$.

Recall that for any set A the symbol $\mathbf{1}_A$ stands for the characteristic function of this set, i.e. $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ otherwise.

Using the well-known probabilistic result concernig the dostributions of such functions of random variables we can obtain the following corollary of the Proposition 1.

Corollary 1 *The classical Young measure generated by any ROSU(f) is a homogeneous one and its density g with respect to the Lebesgue measure on K is of the following form*

$$g(y) = \frac{1}{b-a} \sum_{l=1}^{n} |h_l'(y)| \mathbf{1}_{D_l}(y) \tag{1.4}$$

where h_l is the inverse of f on the interval I_l, while $D_l = f(I_l)$ is the domain of h_l.

To show how Proposition 1 works in practice, let us consider a specific function f and an exemplary Carathéodory function H for which both sides of the Eq. (1.1) can be computed precisely.

Example 3 Let us consider a function $f(x) = \sin x$ defined on the interval $[0,\ 2\pi)$. Now, let the ROSU(f) be defined on the interval $[2, 4)$.

Let us compute the integrals $C(f_k, H)$ that appears on the left-hand side of the Eq. 1.1 for the exemplary Carathéodory function $H(x, y) = xy^2$. We get

$$C(f_k, H) = \int_2^4 H(x, f_k(x))dx = \int_2^4 x \sin^2(kx)dx$$
$$= \frac{1}{8k^2}[\cos(4k) - \cos(8k) + 4k(6k + \sin(4k) - 2\sin(8k))]$$

The left hand side of the Eq. 1.1 is a limit of the above expression when k tends to infinity, so it equals 3.

In order to compute the value of the Young functional on H, i.e. the integral given by right-hand side of the Eq. 1.1 we need to know the classical Young measure generated by the ROSU(f). For this purpose we make use of the Corollary 1 and receive the following formula:

$$g(y) = \frac{1}{\pi\sqrt{1-y^2}} \mathbf{1}_{(-1,1)}(y)$$

Thus the value of the Young functional in the considered case is the following (recall that the Young measure ν_x is homogeneous in this case, so the subscript 'x' is omitted):

$$YF(H) = \int_\Omega \int_K H(x, y)d\nu(y)dx = \int_{[2,4)} \Big(\int_{(-1,1)} H(x, y)g(y)dy \Big)dx$$
$$= \int_2^4 \int_{-1}^1 \frac{xy^2}{\pi\sqrt{1-y^2}} dydx = 3$$

As we can see, the "whole information" about the rapid oscillations in this case is contained in the clasical Young measure and—due to the Proposition 1—can be

revealed with the help of the Corollary 1. In this example the considered generating function f has continuously differentiable extension f^e and the integrations needed to compute the $C(f_k, H)$ were easy to perform. But although this example is intentionally simple, it perfectly shows the benefits resulting from our proposition. It is quite clear that even in this case, where the oscillation have such smooth nature, the computation of the limit of $\{C(f_k, H)\}$ for more complex Carathéodory functions could be a much more difficult task than the calculation of the value of the Young functional YF(H). Moreover, the problem is getting even more difficult if the extension f^e of the generating function f is piecewise continuous with countably many discontinuity points. The next example deals with such a case.

Example 4 Let us consider the generating function f given by Eq. 1.3 and ROSU(f) introduced in Example 1. Although we use again sinus function as the "basis" for the definition of f, it appears that due to the discontinuity of its extension f^e the general symbolic form for the integrals $C(f_k, H)$ cannot be computed, as long as k is unspecified. They can be only computed for given values of k (and not for large values) or approximated numerically, but even this can be very challenging task for such a functions. Moreover it is certainly insufficient for calculation of the limiting value, which is our aim. With the help of Wolfram Mathematica 10.4 software we computed numerically the integrals in the considered case assuming the Carathéodory functions $H(x, y) = x^2y^3$. The exemplary computed values are as follows (recall that the domain of ROSU(f) is the interval $[2, 6)$): $C(f_5, H) = 941.71, C(f_{10}, H) = 891.05, C(f_{50}, H) = 942.32, C(f_{100}, H) = 956.47$, $C(f_{500}, H) = 953.92, C(f_{600}, H) = 955.75$.

As we can see it is not easy to guess what is the limit value of the sequence $\{C(f_k, H)\}$.

Now to compute this limit let us make use of the Young concept. For this purpose we need the density function g of the classical Young measure genereted by the ROSU(f). With the help of Corollary 1 in this case we obtain:

$$g(y) = \frac{4}{3\pi\sqrt{4-(y-2)^2}}\mathbf{1}_{(0,2)}(y) + \frac{2}{3\pi\sqrt{4-(y-2)^2}}\mathbf{1}_{[2,4)}(y)$$

Thus the value of the Young functional on H is the following:

$$\begin{aligned}\mathrm{YF}(H) &= \int_{\Omega}\int_{K} H(x,y)dv(y)dx \\ &= \int_{2}^{6}\int_{0}^{2} \frac{4x^2y^3}{3\pi\sqrt{4-(y-2)^2}}dydx \\ &+ \int_{2}^{6}\int_{2}^{4} \frac{2x^2y^3}{3\pi\sqrt{4-(y-2)^2}}dydx \\ &= \frac{832(45\pi-44)}{27\pi}\end{aligned}$$

The above limit value could hardly be guessed on the basis of the approximately computed values of the elemnents of $\{C(f_k, H)\}$. The decimal form of the limit, which is 955.09, differs from the computed numerically value of $C(f_{600}, H) = 955.92$. It is worth mentioning at this point that the numerical integration of $C(f_k, H)$ for $k > 600$ failed to converge due to highly oscillatory integrand.

1.5 Final Remarks

The probability theory provides us with a number of different versions of the theorems concerning the shapes of distributions of functions of random variables/vectors. Consequently various other rules for computing explicit formulae for the density functions of classical Young measures generated by ROSU(f) can also be developed on the basis of the new result stated in Proposition 1. We are also sure that the same approach enables development of analogous results related to rapidly oscillating sequences with uniform representation which are defined on open and bounded subsets of $\mathbf{R}^d$. It is promising direction of future research.

The possibility of derivation of explicit formulae for the density functions of classical Young measures is not the only benefit resulting from our Proposition 1. In many interesting cases explicit formulae for these densities cannot be found. For instance, if one wants to make use of Corollary 1 they have to obtain the inverses of f on particular subintervals of its domain, but it is not always possible. However in all such cases thanks to Proposition 1 one may use directly Monte Carlo simulations in order to compute values of the Young functionals. That fact significantly broaden the range of possible applications of our main result.

References

1. P. Billingsley, *Convergence of Probability Measures*. Willey (1968)
2. A.Z. Grzybowski, P. Puchała, Remarks about discrete Young measures and their Monte Carlo simulation. J. Appl. Math. Comput. Mech. **14**, 195–199 (2015)
3. A.Z. Grzybowski, P. Puchała, Monte Carlo simulation in the evaluation of the young measures—comparison of random-number generators, in *Proceedings of 2015 IEEE 13th International Scientific Conference on Informatics*, ed. by V. Novitzká, Š. Korečko, A. Szakál (2015)
4. A.Z. Grzybowski, P. Puchała, Monte Carlo simulation in the evaluation of the young functional values, in *Proceedings of 2017 IEEE 14th International Scientific Conference on Informatics*, ed. by V. Novitzká, Š. Korečko, A. Szakál, pp. 221–225 (2017)
5. A.Z. Grzybowski, P. Puchała, On classical Young measures generated by certain rapidly oscillating sequences, in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2017*, San Francisco, USA, 25–27 Oct, 2017, pp. 889–892
6. A.Z. Grzybowski, P. Puchała, On general characterization of Young measures associated with Borel functions. arXiv: 1601.00206v2 (2017)
7. S. Müller, Variational models for microstructure and phase transitions, in *Calculus of Variations and Geometric Evolution Problems*, ed. by S. Hildebrandt, M. Struwe (Springer, Berlin, Heidelberg, 1999), pp. 85–210
8. P. Pedregal, *Variational Methods in Nonlinear Elasticity* (SIAM, Philadelphia, 2000)
9. P. Puchała, An elementary method of calculating Young measures in some special cases. Optimization **63**, 1419–1430 (2014)
10. P. Puchała, A simple characterization of homogeneous Young measures and weak L^1 convergence of their densities. Optimization **66**, 197–203 (2017)
11. T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus* (Walter de Gruyter, Berlin, New York, 1997)
12. L.C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations. Comptes Rendus de la Société des Sciences et des Lettres de Varsovie classe II **I**(30), 212–234 (1937)
13. L.C. Young, Generalized surfaces in the calculus of variations. Ann. Math. **43**, part I: 84–103, part II: 530–544 (1942)

Chapter 2
Energy Requirements Estimation Models for Iron and Steel Industry Applied to Electric Steelworks

Lorenzo Damiani, Roberto Revetria, Pietro Giribone and Maurizio Schenone

Abstract The price of electric energy depends on additional factors since the introduction of renewable energy sources, which has changed the basics of electricity production and the determination of its price. Iron and steel industries strongly require forecasting procedures for the energy amount of their production cycles: today production planning is performed without taking into account that the difference in electricity price between night and day can overcome 500%. The aim of this work is to create a model allowing to estimate energy requirements for steel industry; the model correctness is assessed, for both energy and power analysis, by comparison with real data. A planning tool is employed to provide data to a computer platform able to assess, on the basis of required energy, the best market on which power can be purchased ensuring money saving for the steelworks.

Keywords Computer simulation · Consumption forecasting · Decision support system · Electric power · Energy market · Industrial process optimization Steelworks · Production planning · Production process · Software tool

L. Damiani · R. Revetria (✉) · P. Giribone
Genoa University, Via Opera Pia 15, 16145 Genova, Italy
e-mail: Roberto.Revetria@unige.it

L. Damiani
e-mail: lorenzo.damiani@unige.it

P. Giribone
e-mail: pietro.giribone@unige.it

M. Schenone
Politecnico di Torino, Corso Duca degli Abruzzi, Torino, Italy
e-mail: maurizio.schenone@polito.it

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_2

2.1 Introduction

Nowadays, electric grid users are very heterogeneous: from domestic user to second homes, from city lighting to industrial plants of remarkable consumption. Such a differentiated panorama implies the necessity to produce electric energy in the moment in which the same is required; therefore the installed power must be sufficient to satisfy a demand which is not constant with time. In steelworks industry [1, 2], not long time ago, the Companies focus was mainly turned to production process itself, to its control and automation. The success of this focusing is evident, and the quality standards offered by the market, which nowadays are taken for granted, are extremely high, so as applications, which are more and more advanced. In order not to risk the falling of the Occidental industries competitiveness compared to the big Asiatic Companies, it is necessary to become competitive not only on a technological and quality horizon, but also, and above all, to be able to excel in the customer satisfaction field, being able to suit in a repeatable manner, customers requirements in terms of delivery time, building a stable process of Planning, Insertion and Management of the orders. Such system requires firstly a production order and a definition of the status quo; after that, an optimized production planning needs to be implemented to satisfy the orders. It must be taken into account that during production problems may occur that can require an immediate production re-programming. The introduction of a virtual planning, in all the possible problems that can require a production re-programming, allows to assess in a virtual time the plant status; the system is able to process and show the new organization of the production line. At the same time, the system is able to interact in a continuative manner with the energy market basing on plant necessities. In this paper, the Authors propose a planning method for a complex steelworks plant. The proposed method performance is assessed by comparison with real data obtained by measurements on the plant. The analysis involved fully productive days, days with up and down power ramps, non productive days and anomalous days. To provide thorough terms of comparison, the MAPE (mean absolute percentage error), the Least Squares and the ANOVA (analysis of variance) methods were employed.

2.2 The Complexity of the Electric System

The complexity of the inter-relation between demand and offer, imposes an equally complex management of the electric system [3, 4], which can be divided into: (1) Power generation plants: these are characterized by a mix of technologies suitable for satisfying the time-varying behavior of the demand, which appears not only within 24 hours, but also within weeks and in general during the whole year. In order to assure power supply to users, the installed power needs to be 10–15% higher than peak demand power. In fact, other than maximum yearly demand, installed power needs to satisfy the power for programmed or extraordinary events, e.g. maintenance or

renovation, failures or natural causes such as the incomplete filling of hydro-electric basins due to rain scarcity. For this reason, a certain number of power plants are kept in operation H24 (e.g. nuclear plants, flowing water plants and modern coal plants, which supply for the base-load), while the plants characterized by a short startup time (e.g. gas turbines, combined cycle plants, reciprocating engines and water basin hydroelectric plants) are employed to satisfy the peak demand.

(2) Transmission grid: this connects power generation plants to the distribution grid and represents the place in which electric energy demand and offer meet. Along the whole chain going from generation to final use, energy losses inevitably occur by Joule effect. Generally electric power is transmitted at high voltage to reduce losses, which increase with increasing distance between generation and use locations.

A transmission grid is composed by:

- three lines, usually overhead, which transport power connecting two electric stations, or an energy immission/withdrawal point;
- electric stations, used both to distribute the power between the lines of a network and for transferring electricity between different voltage networks;
- load rephasing systems;
- switches;
- regulation and control systems.

Electric power transmission grids are structured in a knit way, in order to make available alternative routes in case of breakdown or to share the load on the network.

(3) distribution grid: this brings electricity to the final consumer. The national transmission grid is connected to the distribution grid, which supplies the electricity to the final user with a voltage dependent on the type of use. It is divided into:

- High voltage (HV) grid, ensuring the primary distribution of electric power. This grid includes lines with voltage values between 30 and 150 kV;
- primary cabins, transforming electric power from high voltage to medium voltage;
- Medium voltage (MV) lines, from 1 kV up to 30 kV;
- secondary cabins, transforming electric power from medium to low voltage;
- Low voltage (LV) lines, characterized by voltage values lower than 1 kV;
- Low voltage grid, feeding all domestic users.

The energy produced by the electric system is subdivided into six classes, to account for internal uses and losses:

- Gross production: it represents the sum of the energy produced by all the plants;
- Net production: this represents the gross production minus the generation plants requirements;
- Energy for pumping: this represents the energy required in water basin pumping systems to pump the water from the lower to the higher basin;
- Production intended for consumption: this is net production minus the pumping energy;

- Electric energy required by the grid: this represents the energy production intended for consumption minus the exported energy plus the imported energy;
- Electrical consumption: this represents the energy required by the grid minus the transmission losses.

2.3 The Large Energy Consumers: Steelworks

The production size of electric steelworks is between 1 and 2 million tons per year; the furnaces casting capacity is limited, which imposes a high frequency of casting in modern systems to maintain a high cycle production. Electric-powered steelworks, smaller than the integral cycle ones, allow the construction of plants with an acceptable relationships between investment and production capacity.

Electric ovens are small ovens with a maximum production capacity of 12 tons approximately. Their development and use has in many cases been held back by the cost of energy, which significantly affects the cost of steelmaking.

The melting furnaces with electric heating can be divided into two large categories:

- electric arc furnaces (EAF);
- Induction furnaces.

Induction furnaces with magnetic cores are made of a metallic bath contained in an annular crucible of small cross-section forming the secondary of a transformer whose primary is wound on a frame-shaped iron core, located in a vertical plane. Alternated current is delivered in the primary, so that a low voltage but high intensity current is induced in the metallic bath, warming it by Joule effect. Their application in the iron and steel industry is very limited, while it is more extensive in the field of special metallurgy such as copper and nickel.

Induction furnaces without magnetic core have had a remarkable development for the production of high quality steel, or foundry for the production of cast iron.

In ovens without core, the primary consists of a copper spiral with a tubular section traveled internally by a stream of cooling water and wrapped around a refractory crucible containing the metal bath.

The arc furnace operates through an electrical discharge that melts the metal; to reduce the consumption of the electrode, steelworks have begun to introduce preheated scrap. This type of furnace is formed by a cylindrical crucible with a rounded bottom and whose walls form the vat. The vault covers the vat and, rotating on a peripheral vertical axis, opens the oven allowing to quickly load it from the top through the charging baskets.

The power involved may range from 500 to over 100 000 kVA; the potentiality of the oven is measured by the diameter of the basin. Consumption is in the order of 500–700 kWh per ton of product. Once produced in semi-finished steel, the final desired shape must be impressed. This part of the steelmaking process is common to the two oven types; there are two ways to proceed: casting into ingots and following rolling, or continuous casting.

The competitiveness of such plants is based on the ability to achieve very important economies of scale in terms of production volumes; the optimal production threshold for large steelworks ranges between 5 and 10 million tons per year.

2.4 Steel Manufacturing Plant

The first phase of the Electric Arc Furnace is the loading of the scrap, which usually comes from the demolition of steel structures or process residuals. The scrap is transported above the electric oven through a basket which, by overturning, allows the scrap to be poured inside the vat; here the scrap begins the melting process through the electric arcs that appear between the electrodes and the scrap itself. Once liquid state is reached, pressurized oxygen is blown in the reaction volume, enabling to withdraw undesired chemical elements, such as nitrogen and phosphorus. In this phase, slags are formed, collecting on the top of the metallic bath. When the steel has reached the prescribed value of alloying elements and the ideal temperature for the subsequent treatments, it is poured into the ladle; this is then taken to the Ladle Furnace, in which molten steel is kept in temperature by means of electrodes and is further added with alloying elements. In Fig. 2.1 is represented the Electric Arc Furnace, completed by the various flows of materials and energy and the main auxiliary equipment.

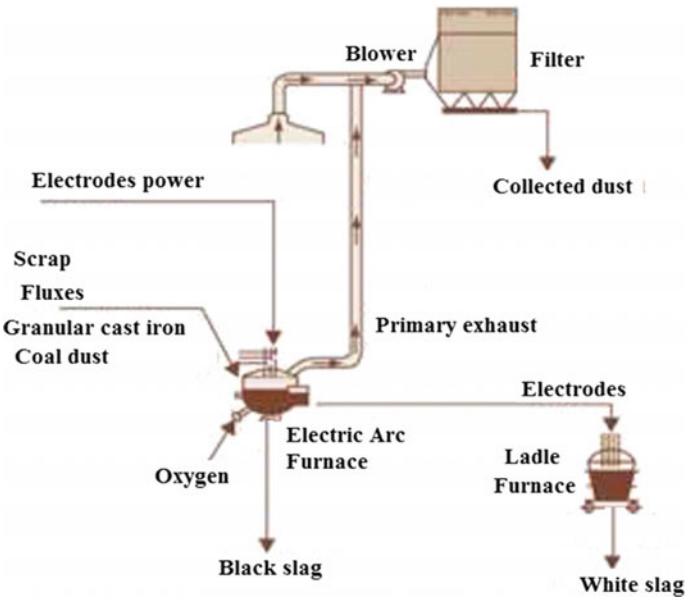


Fig. 2.1 Process of the electric arc furnace

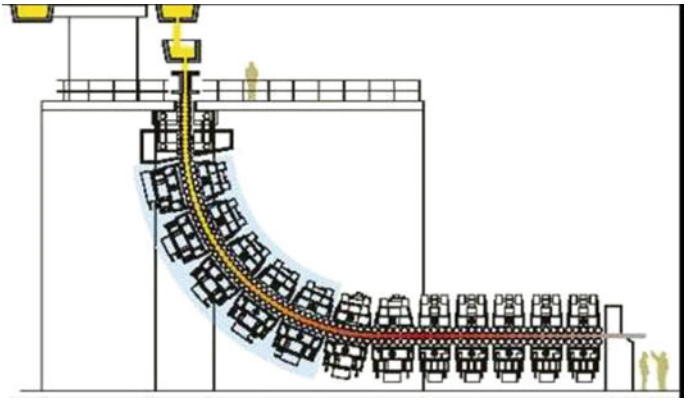


Fig. 2.2 Process of the electric arc furnace

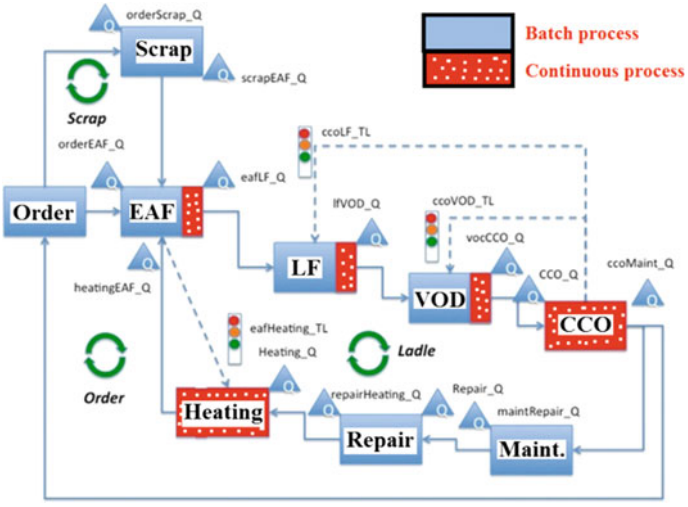


Fig. 2.3 Process scheme of the analyzed steelworks

According to the plant type, steel can then pass directly to a Continuous Casting Machine, assuming different shapes, or be placed in molds, solidified in ingots and then milled. Continuous Casting Machine is an oscillating mold, in which liquid steel is poured assuming the shape impressed by the casting channel. Usually, the mold is constituted by water-cooled copper walls.

The longer is the path made by steel, the longer will be the cooling phase, therefore modern machines are developed along a curved line, as visible in Fig. 2.2.

Figure 2.3 represents the scheme of the operations of the steelworks analyzed in this paper. The processes represented in figure are the following:

- *Order*: arriving orders;
- *Scrap*: scrap warehouse;
- *EAF*: Electric Arc Furnace;
- *LF*: Ladle Furnace;
- *VOD*: degassing treatment;
- *CCO*: Continuous casting;
- *Maint and Repair*: processes of repair and maintenance of the ladles;
- *Heating*: ladles heating process.

Rectangular blocks correspond to the processes, ie the treatments that the initial product undergoes to get the finished product. Triangles represent the queues in input and output of the processes.

Operations in full color (CCO and Heating) are continuous processes, those in dotted color are batch, those with both the colors (EAF, LF, VOD) present both a discrete and a continuous component. Continuous arrows indicate the processes temporal dependencies and order. Traffic lights and dotted arrows, with direction opposite to continue arrows, indicate the functional dependency between processes: the process from which arrow starts commands the process under the traffic light (for example, the EAF regulates the arrival of a ladle from the Heating process). The three green circular arrows identify three types of cycle:

- scrap cycle (Scrap) following the Scrap, EAF, LF, VOD, CCO and Order path;
- order cycle (Order) that follows the route Order, EAF, LF, VOD, CCO;
- ladles cycle (Ladle) that follows the path Heating, EAF, LF, VOD, CCO, Maint. and Repair.

The following Fig. 2.4 provides an overview of the plant studied in this work, highlighting for the mentioned components the hourly capacity of each individual machine.

In Fig. 2.4 is visible the flow scheme of the whole process, including the material flow rates processed by each component. The examined plant provides two production lines; in particular:

- EAF: Electrical Arc Furnace that can work 7 days a week; after that it requires maintenance (these operations take between 8 and 12 hours)
  - EAF A: processing every 75 min with 110t/h capacity;
  - EAF B: processing every 60 min with 150t/h capacity;
- LF: Ladle Furnaces with the same capacity of the EAF of its line; they consume 20 MW;
- CASTER: it is the station where casting starts. The plant has three casters:
  - NNS (Near Net Caster): capacity 100t/h, it mainly supplies billets to Large Section Mill (LSM);
  - B CASTER: capacity between 100 and 160t/h, it supplies billets to Medium Section Mill (MSM);

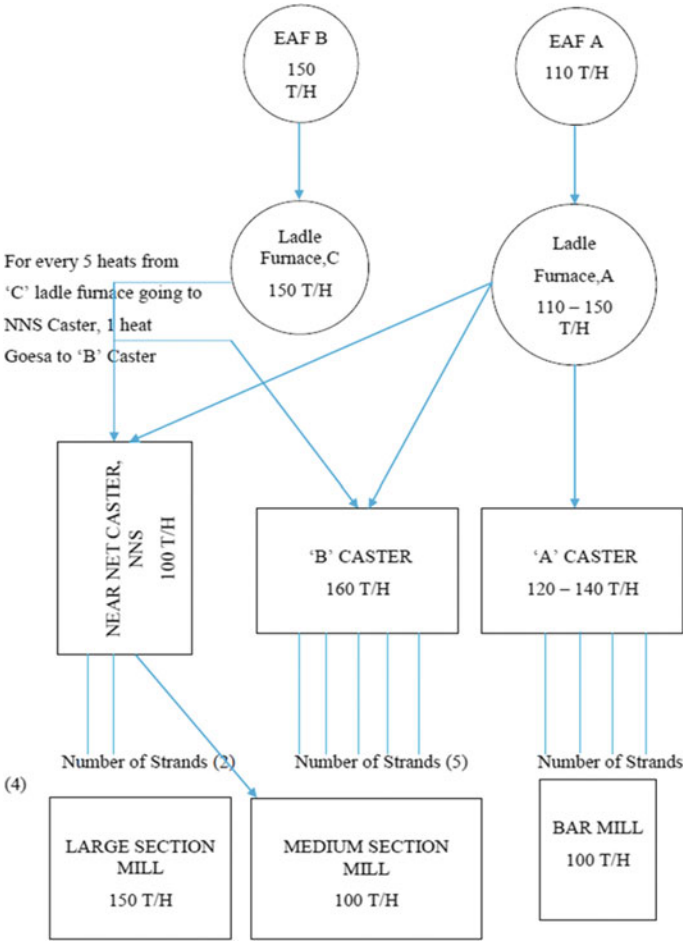


Fig. 2.4 Process flow diagram

  - A CASTER: capacity 100t/h, it supplies billets to Bar Section Mill (BSM).
- MILLER: three rolling mills.

2.5 Consumption Data

The steelworks consumptions have been monitored and collected in a data-base for the period between 1/1/2010 and 30/9/2015, excluding the year 2012 for which it was not possible to obtain an exhaustive documentation.

Each day has been divided into 3 shifts (from 9 p.m. of the day before to 4.00 a.m. of the current day, from 5.00 a.m. to 12.00 a.m. of the current day, from 1.00 p.m. to 8.00 p.m. of the current day). The average of hourly required power has been calculated, such as the average of all day. Based on this, can be identified:

- Days of production: if the average of the absorbed power is 21 MW at least;
- Days of ramp down production if: (i) The second half of the first shift, the second shift and the first half of the third are productive; (ii) The day hasnt been classified as productive; (iii) The first half of the third shift, the second half of itself and the first half of the first shift (the one related to the day after) have to be characterized by a ramp down;
- We have a ramp down if:

1. The second half of the previous shift is productive;
2. The first half of the next shift is no-productive;
3. The average of the half under exam is higher than the average of the next half;
4. The first value of the half under exam is higher than the value associated to the last hour of the same half;
5. The next half is no-productive.

- Days of ramp up production if

1. The second half of the first shift, the first half of the second or the second half of the same must be characterized by a rising ramp;
2. The second half of the second shift have to be productive just like all the third shift and the first half of the first shift (the one connected to the next day);
3. The day must not have already been classified as fully productive.

- We have a ramp up if:

1. the second half of the previous shift is non-productive;
2. the first half of the next shift is productive;
3. the average of the half in question is less than average of the later half;
4. the first value of the half under examination, corresponds to the first hour of the same, is less than the value associated with the last hour of the same half;
5. the half later is productive.

- Days are classified as no-production if:

1. the first half and the second half of the first shift are not productive;
2. the second and the third shift are not productive;
3. the day hasn't been categorized as productive with one of the two ramps.

- Abnormal days: any day does not fit into any of the classes above described.

Each of the 1732 days analyzed has been assigned to the category with the following result:

- 336 whole days production;
- 523 no-productive days;
- 81 productive days characterized by ramp down;
- 88 productive days characterized by ramp;
- 704 abnormal days.

2.6 Construction of Production Profiles Based on Planning

In this section will be described the steps to create a simplified ideal production profile basing on which it is possible to organize and manage the purchase of electric energy, avoiding to buy useless energy in those days in which the plant is stopped or only partially operative.

An optimized production plan, based on times that each machine uses for fulfilling its task, was implemented. In particular, in the following are described the processing times for a complete production process.

EAF: 60 min; Transport between EAF and LF: 5 min; LF: 40 min; Transport between LF and CCO/VD: 5 min; Degassing (VOD): 20 min; Continuous casting (CCO): 40 min. The last two processes were considered as one for simplicity, creating a unique process CCO/VOD lasting 60 min.

The last two processes were considered as one for simplicity, creating a unique process CCO/VOD lasting 60 min.

For each of the three equipments, a chart was built (Fig. 2.5), showing the absorbed MW (y axis) in function of time (x axis); the equipment consumptions are estimated in 50 MW for EAF, 20 MW for LF and 1 MW per both degassing and CCO.

Planning begins from zero-day and the first process of each equipment doesnt consider the time interval needed to reach the processing temperature.

To optimize working time, in both the lines of the plant the second casting starts when the first semi-finished piece is exiting from the ladle furnace (LF) (gap time of 5 min) in order to avoid delays which would turn into costs for missed production.

The treatment in ladle furnaces LF-A and LF-B lasts for 40 min, after which billets are transported, in a 5 min time, to degassing machine. Degassing treatment lasts 20 min, downstream of which there is the continuous casting plant which terminates its work after 40 min.

Reporting on a single diagram the power required by all the three components (EAF, LF and CCO/VOD) in order to complete a production phase, in the ideal case of considering the temperature ramps instantaneous (zero startup time), to transform a whole steel batch into finished product, and overlapping of three of these optimized power diagrams, opportunely shifted in time one each other, the diagram in Fig. 2.5 is obtained.

Summing in function of time the power values and considering a base-load power for auxiliaries equal to 5 MW, the diagram of the total absorbed power is obtained, as visible in Fig. 2.6.

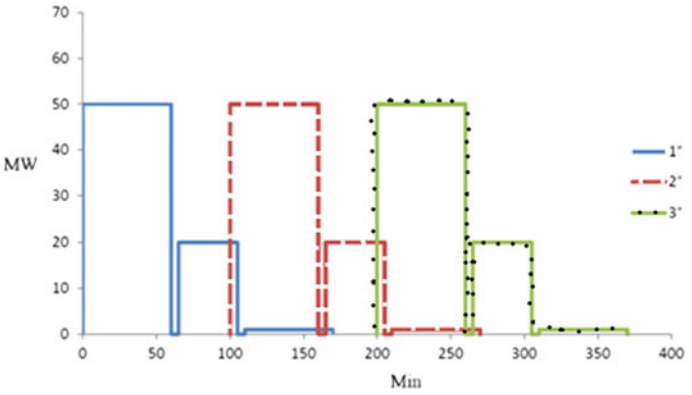


Fig. 2.5 Three cycles in series

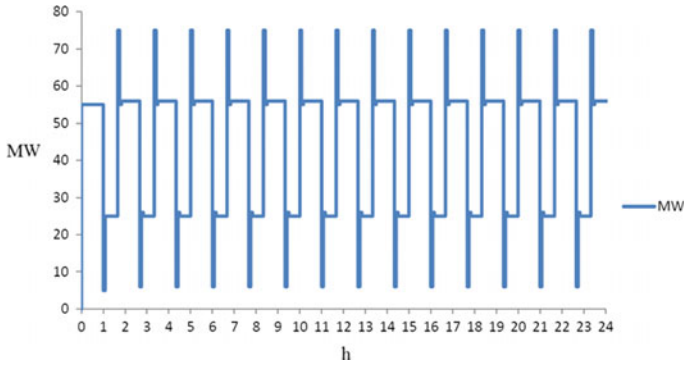


Fig. 2.6 Total absorbed power

From this diagram it is possible to estimate the energy amount required for each hour of plant operation, as visible in Fig. 2.7.

2.6.1 MAPE

The MAPE (Mean Absolute Percentage Error) is defined as: (Expected Value-Real Value)/Real Value. Such value was calculated for all the days typologies above described, both for energy and for power values, obtaining the following results:

- production days: 27,2% (energy) and 32% (power);
- no-production days: 76,7% (energy) and 84% (power);
- production days with ramp up: 12,7% (energy) and 13% (power);
- production days with ramp down: 80,4% (energy) and 81% (power);
- anomalous days: 38% (energy) and 76% (power).

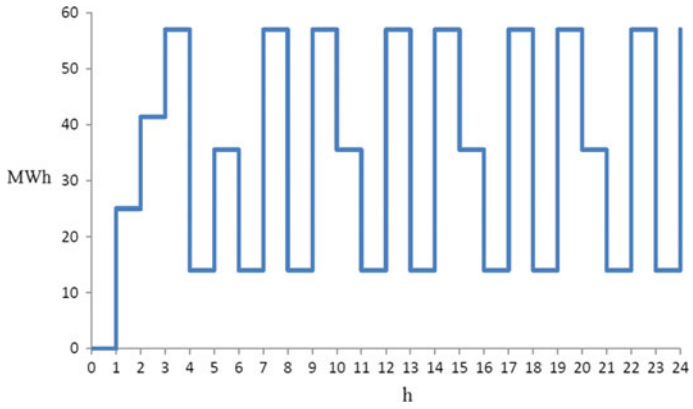


Fig. 2.7 Energy consumed in each hour of plant operation

It is an optimization technique that allows to determine the linear function that minimizes the sum of squares of distances between data. The general formula of the straight line is: $y = mx + q$, where y in this case is the actual plant data, x is the planning data. The parameters m and q are found by fitting the data by using least squares. Also in this case the analysis is carried out for all the five day typologies.

It was first calculated the hourly average of relative power errors for the days from which were determined:

- Average Error: the average of the power errors;
- Error^2: square of all the related errors;
- Average Error^2: the average of the values just above;
- Average Hours: the average of the twenty four hours per day;
- AverageHours^2: the average of the squares of hours;
- Error*Hour: the product of the errors and their hours;
- Average (Err*Hour): the average of the 24 values calculated above.

The coefficient m is given by:

$$m = \frac{(x \cdot y)^* - x^* \cdot y^*}{(x^2)^* - (x^*)^2} \tag{2.1}$$

where

$(x \cdot y)^*$: it is the average of the product of the actual data for the ideal ones;
x^* : it is the average oh the ideal data;
y^* : the average of real data;
$(x^2)^*$: it is the average of the square of the ideal data;
$(x^*)^2$: it is the square of the average of the real data.

The known term q is given by $q = ymx$. The calculation of the correlation coefficient (CP), also called Pearson coefficient, gives us an indication on the goodness of

our approximation: its range is [−1, +1], more its value tends to 1, better was the approximation.

$$CP = \frac{(x \cdot y)^* - x^* \cdot y^*}{\sqrt{[(x^2)^* - (x^*)^2] - [(y^2)^* - (y^*)^2]}} \tag{2.2}$$

As for the power analysis:

- For productive days CP is 0,1167
- For no-productive days CP = 0,066
- For abnormal days CP = 0,7670
- For productive days with ramp CP is 0,739
- For productive days with ramp down CP = 0,33.

As regards the energy analysis is obtained:

- For productive days CP = −4,5E−02
- For no-productive days CP = 6,9E−02
- For abnormal days CP = 7E−02
- For productive days with ramp CP = 1E−02
- For productive days with ramp down CP = 3E−02.

2.6.2 Analysis of Variance (ANOVA)

Analysis of Variance (ANOVA) is a collection of statistical models used to analyze the differences among group means and their associated procedures. It is based on the test null hypothesis: H0 = the average of different populations are the same; indicating with 1, 2, 3 the average dimensions of populations, the null hypothesis can be written as H0: 1 = 2 = 3 = = k. Once we have gathered data, the solidity of the null hypothesis can be gauged. The variability measured between k average (on groups) are compared to variability on each population (in groups). The comparison between variance on groups and variance in groups gives F-value: low F-value means that H0 is true, high F-value means that H0 is false. In this context, the P-value is defined as the probability that the observed data come from the null hypothesis or from the alternative hypothesis. In particular, high P-values favor the null hypothesis, while low P-values are against the null hypothesis. The P-values calculated for the different day typologies are presented in the following:

For productive days: 6,85E−73

- For no-productive days: 1,05E−50
- For abnormal days: 1,24E−271
- For productive days with ramp: 1,55E−06
- For productive days with ramp down: 1,91E−34

Such small values are an evidence against the null hypothesis, but relative errors are not classifiable, thus it is not possible to establish their nature by means of this analysis [5].

2.7 Electric Market

2.7.1 Market Structure

Since 2005, the spot energy market has been divided into Day-ahead market (MGP), intraday market (MI), adjustment market (MA) and the market for ancillary services (MSD).

1. The day-ahead market presents an auction system where both bidders and buyers take part; bids are characterized by quantity and unit price for energy, the purpose of this market is to point out the possibility to sell and/or to buy energy not at a lower price than the one that has been proposed. GME (Electric Market Manager) arranges bids and purchase offers and it draws two graphs: (i) The sale curve: bids are ordered by descending price. (ii) The purchase curve: purchases offers are ordered by descending price.
   The intersection of the two curves (point P*) defines how much energy can be ex-changed, the reached price, the approved offers and injection and withdrawal pro-grams. The selling price of the accepted offers is not higher than P* and the purchase price is not lower than P*.
2. The intraday markets (MI). Also these markets are managed by GME. The price calculation and the method of acceptance are the same as MGP. MI is divided into 4 submarkets (MI 1 to MI 4) according to the opening hour.
3. Adjustment market (MA). It opens at 10.30 after communication of the results of MGP, it have to allow operators to modify programs that have been determined after results of MGP; they can make new bids and it closes at 14.00.
4. The market for ancillary services (MSD). It opens at 14.30 after results of MA and it closes at 16.00.

2.7.2 Software Tool for Energy Purchase on the Market

To complete the optimization tool for steelworks, a software able to forecast energy price in the market based on energy price historical data was implemented. It provides a support system to decisions (Decision Support System) for managers and all people who must make strategic decisions and has the purpose of allowing to extract in a short time and in a flexible way, by large size data, the information needed to support and effectively improve the process decision.

The developed software platform uses two types of values: on the one hand the forecasts of market prices, on the other the simulation data of the plant. Both of these data must be estimated on purpose: in particular the expected consumption of the plants are generated with the aid of the simulator, energy prices are processed using statistical methods starting from previous market outcomes. The platform incorporates the simulator and after receiving the production orders, generates the simulation.

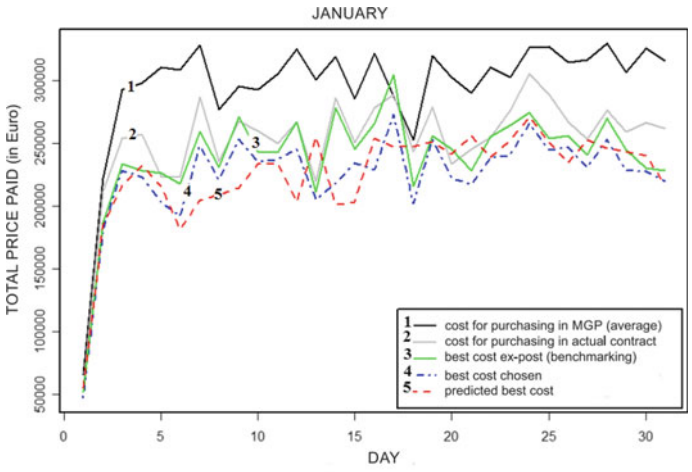


Fig. 2.8 Comparison diagram for energy purchase decision making

Once the hypothetical hourly consumption is acquired, these are crossed with the price forecast generated by a SARIMA model: it is a combinatorial calculation, so there will be a huge number of possible scenarios; therefore, they are filtered, and only the most significant ones will be kept. The first will be to buy at the price of contract, although in reality each contract provides a variable unitary price in function of the power delivered in the hour and based on the number of power peaks that overcome the limit. Five other scenarios are those of purchasing the whole energy required in the corresponding five markets of the Electricity Exchange. The most important scenario is the one with the best expected prices; for each hour the tool selects the most advantageous market, indicating the offer price and on which market the purchase should be made.

Starting from the production Gantt diagram, the total consumption in MWh is determined for each hour, and such data are crossed with the forecasted market price. The result of the forecast of the purchase scenarios is given in the form of bar chart (see Fig. 2.8), in which the total purchase price is shown so as the total energy price and the different mix of purchase markets. In Fig. 2.8 is visible the graphical interface that helps the user to make decisions about the possible power buying options, this model was implemented using Systems Dynamics formalism based approach tested by Authors in many applications [6]. Each bar of the diagram shows the total price of energy if we act on markets in different way; referring to the graph in Fig. 2.8:

- Plot 1: all the energy is purchased at MGP.
- Plot 2: all the energy is purchased at the current contract.
- Plot 3: all the energy is at the best market price (MGP, MI 1, MI 2, MI 3, MI 4),

i.e. benchmarking.
- Plot 4: purchase price suggested by the tool.
- Plot 5: predicted best cost.

2.8 Conclusions

This paper presented a planning methodology for a steelworks production. The methodology efficacy was tested by comparison with real power and energy data collected from an operative plant in a long observation time. The comparison between real data and planning was carried out by three mathematical methods, whose main results are summarized in the following.

1. For the MAPE method is:
   - Acceptable for productive day and productive days with ramp up (for both energetic and power analysis)
   - Not acceptable for no-productive days and productive days with ramp down both for energetic and power analysis.
2. For least squares:
   - Acceptable for productive days with ramp and abnormal days for power analysis
   - Non acceptable for no-productive days for power analysis and for all 5 types for energetic analysis.
3. Analysis of Variance:
   - Errors are determined by events, so it is not possible to classify with this analysis.

Finally, using the proposed methodology with a market analysis tool is was possible to obtain significant savings on energy purchase.

References

1. E. Briano, C. Caballini, R. Revetria, A. Testa, M. De Leo, F. Belgrano, A. Bertolotto, Anticipation models for on-line control in steel industry: methodologies and case study, CP1303, in *CASYS09 9th International Conference on Computing Anticipatory Systems* (2009)
2. G. Fiorani, L. Damiani, R. Revetria, P. Giribone, M. Schenone, Models to estimate energy requirements for iron and steel industry: application case for electric steelworks, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 Oct, San Francisco, USA, pp. 920–924 (2017)
3. L. Damiani, P. Giribone, R. Revetria, Simulink study of a smart node for domestic applications equipped with PV panel, Energy Storage and Home Automation, in *IAENG Transactions on*

Engineering Sciences, Special Issue for the International Association of Engineers Conferences 2016, vol. II (2016). ISBN 978-9813230-76-7

4. L. Damiani, P. Giribone, R. Revetria, A. Testa, An innovative model for supporting energy-based cost reduction in steel manufacturing industry using online real-time simulation, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science 2014*, San Francisco, USA, 22–24 October, 2014, pp. 1–7
5. E. Briano, C. Caballini, P. Giribone, R. Revetria, Using system dynamics for short life cycle supply chains evaluation. Proc. Winter Simul. Conf. Art. **5678887**, 1820–1832 (2010)
6. L. Cassettari, R. Mosca, R. Revetria, Monte Carlo simulation models evolving in replicated runs: a methodology to choose the optimal experimental sample size. Math. Prob. Eng. (2012); in *The Technical Writers Handbook*, ed. by M. Young (University Science, Mill Valley, CA, 1989)

Chapter 3
Technology Landscape 4.0

Yübo Wang, Thilo Towara and Reiner Anderl

Abstract Companies are facing manifold challenges while trying to implement Industrie 4.0, which are in great parts rooted in the complexity of Industrie 4.0 and its associated fields of research. Initiatives have developed abstract reference architectural models to mitigate these challenges and structure Industrie 4.0. The research on hand uses the reference architectural model Industrie 4.0 (RAMI 4.0), which is developed by the German Platform Industrie 4.0. This work aims to create a concrete, yet universal, application-oriented model that fosters the widespread of RAMI 4.0 in practice. It supports further research and amendments, and hence, facilitates the implementation of Industrie 4.0 in small and medium-sized enterprises. An information and technological navigation tool is developed for mapping technologies in RAMI 4.0. Finally, the foundation for a subsequent inclusion of IT security in RAMI 4.0 is laid.

Keywords Application-oriented RAMI 4.0 · Cyber-physical system · Industrie 4.0 · Internet of things · IT security · Technology landscape 4.0

Y. Wang (✉) · R. Anderl
Department of Computer Integrated Design, Technische Universität Darmstadt,
Otto-Berdt-Str. 2, 64287 Darmstadt, Germany
e-mail: y.wang@dik.tu-darmstadt.de

R. Anderl
e-mail: anderl@dik.tu-darmstadt.de

T. Towara
Campana & Schott GmbH, Graefstrasse 99, 60487 Frankfurt am Main, Germany
e-mail: Thilo.Towara@campana-schott.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_3

3.1 Introduction

Industrie 4.0 is increasingly gaining importance in the global manufacturing industry. While it is usually regarded as a prospective for large global enterprises, small and medium-sized enterprises (SMEs) are also recognizing the impact of Industrie 4.0 on the industrial future.

Platform Industrie 4.0, which is the leading institution of German Industrie 4.0, is dedicated to promote research in Industrie 4.0 for the German industry and facilitating the realization of this vision in manufacturing enterprises. In 2013 Platform Industrie 4.0 announced 17 technology development areas covering all the Industrie 4.0-related aspects in "Recommendations for implementing the strategic initiative INDUSTRIE 4.0". An implementation roadmap for Industrie 4.0 is described in this recommendation [1].

As the extension of the first three industrial revolutions, the technical innovation for Industrie 4.0 is based on the vertical and horizontal integration of manufacturing systems, a continuous digital engineering throughout the product lifecycle, and finally the decentralization of computing resources. Enabling such an intelligent network, new technologies like modern internet technologies and flexible hardware and software interfaces are needed [2].

Equipping products and manufacturing environments with these new technologies, Industrie 4.0 offers manifold opportunities for enterprises to enhance efficiency and flexibilize production processes. Therefore, current products will increase in value and the development of new business models is facilitated [1].

According to the existing experiences of the Department of Computer Integrated Design (DiK), the Industrie 4.0 competences regarding communication technologies, digitalization, and IT infrastructure in enterprises, which are considered as key technologies, have to be promoted to bolster Industrie 4.0 implementation and finally, to fully realize the potential benefits [3].

In order to support research, norming processes, as well as practitioners, Platform Industrie 4.0 developed the Reference Architectural Model Industrie 4.0 (RAMI 4.0) . The model helps to classify and identify areas of Industrie 4.0 and creates a solid foundation for the further development of technologies. However, the model itself is rather abstract and its application is complicated for the practical implementation, especially for SMEs. Consequently, the use of RAMI 4.0 is limited to research institutions and first individual use-cases. Due to the abstract design of the model its generic applicability is provided, however, enterprises have to commit substantial resources to populate the model with specific technologies in order to use it as a knowledge foundation [4]. Developing their individual technological roadmaps for the next years, enterprises are facing questions like, which technical solutions would facilitate the development of their enterprise best, how these technologies interact, and how to implement Industrie 4.0 most efficiently.

This work takes these questions into account and proposes an interactive approach to support enterprises in realizing Industrie 4.0 based on RAMI 4.0. The aim is to provide a comprehensive overview of technologies, as well as their topological rela-

tionships, and therefore, a knowledge foundation for the implementation of Industrie 4.0 [5].

3.2 Global Digital Transformation

3.2.1 Global Leading Initiatives

The global digitalization of industries and values chains, along with the corresponding demand for structured research and norming, has brought forward four major initiatives in China, Japan, USA, and Germany. These initiatives address potentials and challenges of digitalization.

The "Made in China 2025" strategy is part of a larger modernization campaign of China. The core elements of the Chinese strategy are enhanced creative ability, improved quality and efficiency, as well as green development. Until 2025 industry and information technology are supposed to be integrated and China's capacity to innovate and manufacturing productivity are planned to have been improved.

The Japanese Industrial Value Chain Initiative (IVI) is a platform to combine manufacturing and information technologies and facilitate collaboration between companies. It was founded in 2015 and aims to discuss potentials of human-centric manufacturing processes and to build a mutually connected system architecture.

The Industrial Internet Consortium (IIC), founded in 2014 by major telecommunications and technology companies in the US, aims to accelerate the development and deployment of interconnected machines and devices, as well as intelligent analytic tools, not just in manufacturing environments but also in areas like energy, transportation and smart cities. Its focus is the identification and promotion of best practices, while bringing together practitioners, researchers and government institutions.

Plattform Industrie 4.0 is dedicated to research in Industrie 4.0 for German industry and supporting the realization of this vision in manufacturing enterprises. The platform aims to identify all relevant technologies and developments in the manufacturing industry and to foster a common understanding of Industrie 4.0.

3.2.2 Global Leading Reference Architectures

Pivotal for all these initiatives is the development and adoption of common standards and norms to enable technology independent connectivity and interoperability across systems and industries. The IVI, the IIC, as well as the Plattform Industrie 4.0 have developed reference architectures to provide fundamental definitions.

3.2.3 Industrial Value Chain Reference Architecture (IVRA)

The Japanese IVI published the Industrial Value Chain Reference Architecture (IVRA) in 2016 to foster the widespread of so called Smart Manufacturing Units (SMUs), which are defined as individual units within industrial systems that interact with each other autonomously through mutual communication and thereby improve productivity and efficiency. The model analyzes SMUs from three different perspectives: asset view, activity view, and management view. While the asset view shows assets valuable to the enterprise (Personnel, Process, Product, and Plant), the activity view addresses the question how smart manufacturing creates values as the outcome of activities. The latter therefore deploys the common Deming cycle (Plan, Do, Check, Act). The management view shows purposes and indices relevant for management like quality, cost, delivery, and environment, which are used for steering assets and activities [6].

3.2.4 Industrial Internet Reference Architecture (IIRA)

The IIRA was first published in 2015 by the IIC and seeks to allow practitioners and researchers to develop and analyze systems based on common frameworks and concepts put together in a standards-based architectural template. It uses a multi-dimensional approach based on four different viewpoints of stakeholders at individual levels ranging from business topics over usage and functional considerations, to the final implementation. Each viewpoint is detailed in further levels and domains, while being intersected by topics like interoperability, security, and machine-learning [7].

3.2.5 Reference Architectural Model Industrie 4.0 (RAMI4.0)

The Reference Architectural Model Industrie 4.0 (RAMI 4.0) was developed by the Platform 4.0 in 2015. It consists of several layers, hierarchical levels, and the product lifecycle representing the value stream [1].

The fundamental structure is shown in Fig. 3.1. The **hierarchical levels** represent the functional characteristics of the components and are oriented on IEC 62264 and IEC 61512. These levels consist of seven individual levels: *product*, *field device*, *control device*, *station*, *work centers*, *enterprise* and *connected world*, from bottom to top. The lowest level, called p*roduct*, includes products that, due to their ability to communicate, are active elements within the production system. They provide information on their individual properties and necessary production steps. The f*ield device* level includes intelligent field devices such as sensors and actuators, the level c*ontrol device* in turn control devices, controllers and embedded controllers. Production machines, robots, or intelligent logistics vehicles are situated on the *station*

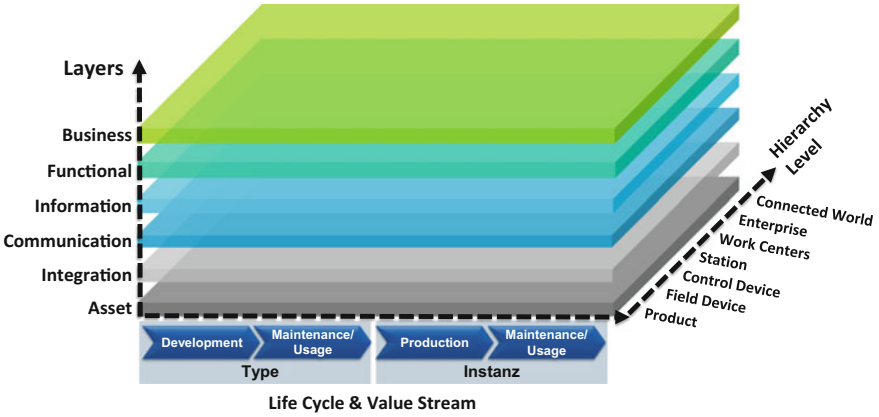


Fig. 3.1 RAMI 4.0

level. Both production plants and departments within a company are assigned to the *work centers* level. The e*nterprise* level considers the company as a whole and the *connected world* level represents its outer networks, e.g. collaboration with business partners and customers [4].

Each component consists of **six layers**. Starting with the lowest layer, the structure consists of *asset*, *integration*, *communication*, *information*, *functional* and *business*. While each layer is distinctively different from the others, elements within one layer are supposed to be homogenous regarding their attributes [8, 9]. The lowest layer is the *asset* layer, i.e. the representation of the physical reality. It contains all real objects such as machines, sensors and documents, as well as humans. In addition, intangible objects such as models, ideas, or patents are also attributed to this layer. Each element or function of the overlying layers must be attributable to an object of the asset layer. The *integration* layer supports the provision of computer-usable information related to the physical assets, such as geometry, hardware, and software, to the overlying layers. It contains all elements associated with the IT, including HMIs, and generates events based on the acquired information. The *integration* layer performs the final control of the technical processes. The purpose of the *communication* layer is to enable communication between the different elements of the network on the basis of uniform communication protocols and data formats. It also provides services to control the *integration* layer. Within the *information* layer the rule-based (pre-) processing of events takes place. For this purpose, data are checked for integrity, summarized into new, higher-quality data, and made available to the superordinate layers via interfaces. Events are received from the communication layer, transformed and forwarded accordingly. The *functional* layer represents the runtime environment for services and applications. It is the platform for the horizontal integration of the various functions and generates rules and application logic. Remote access and the integration of applications and functions take place only here, without interfering with the underlying layers. This ensures the integrity of the information, as well as

the technical level. The *business* layer covers the abstract business models and the resulting process logic. It provides legal and regulatory frameworks and ensures the integrity of the functions along the value chain.

The third axis describes the **life cycle and the value chain** of an Industrie 4.0 component. The structure of this axis is based on IEC 62890 and assumes a basic division into product *type* and *instance*. While a type already exists with the basic product idea and covers the phases from order intake over product development to prototype production, an instance stems from the transition to production after the successful completion of all tests. The manufactured product then represents the instantiation of the type. Types, as well as instances, are subdivided into two phases: development and maintenance/usage for types and production and maintenance/usage for instances respectively [1, 4].

3.3 Challenges of German Small and Medium-Sized Enterprises

While implementing Industrie 4.0 companies have to overcome manifold issues. As small and medium-sized enterprises often lack the needed capabilities, for them Industrie 4.0 is especially challenging. The vertical and horizontal integration along the value chain, a continuous digital engineering, as well as decentralized production processes, inhibit great potentials, however, require a profound restructuring of IT and work organization and finally, bring along severe security threats.

3.3.1 Increase of Process Efficiency

Vertical integration of industrial value chains enables a dynamic adaptation to change environments. Accessibility of data is improved which can be used to optimize process efficiency and performance. Decentralized intelligent production systems can be altered in real-time, making even small batches, last-minute changes, or completely customized products economically feasible. Through modern engineering paradigms like collaborative, smart, and digital engineering, development time is further reduced and product supply is ultimately matching product demand [1].

3.3.2 Flexibility of Work Organization

Collaboration of employees and machines is enhanced by the implementation of multimodal Human-Machine Interfaces (HMIs). This symbiosis utilizes human attributes like creativity and flexibility together with the fatigue-free precision of machines.

Work not only becomes more efficient but also more humane. Challenges like the demographic change, the inclusion of employees with special needs, or other barriers can be encountered by intelligent workers' assistance systems. Modern Information and Communication Systems (ICS) allow remote telework, reduce transportation and travel expenses, and flexibilize site selection [1, 10].

3.3.3 Enabling New Business Models

The digitalization of processes along the entire life cycle produces vast amounts of data. The ability to analyze these data employing intelligent algorithms in combination with powerful hardware systems such as in-memory databases is an unprecedented source of information. All potentially relevant information can be recorded and taken into account in operational or strategic planning. Problems with products can be detected at an early stage on the basis of indirect customer feedback via social media. Furthermore, target groups can be analyzed more profoundly and product development adapted accordingly. New data-driven business models are emerging both in B2B (business-to-business) and B2C (business-to-costumer) areas. Services are increasingly important and partly replacing conventional products [10].

Along with these new opportunities, however, companies also face new challenges. Different platforms of various manufacturers must be networked and reliably communicate with each other on a uniform basis. Therefore, cross-platform interfaces and protocols must be defined and used extensively. In order to successfully transition to Industrie 4.0 not only technological solutions have to be implemented within the scope of production, but the entire structure of the company has to be transformed. Individualized product service systems partly displace conventional products. Data-based services and platform solutions require the development of new business models [11].

3.3.4 IT Security in Industrie 4.0

By connecting the formerly separated production and business networks, as well as exchanging data with external partners, new security risks arise. However, due to the requirements of real-time capabilities and limited computing resources, conventional security measures are not fully feasible [10]. To protect sensitive know-how and allow for trust-based collaboration confidentiality and integrity of transferred data must be ensured at all times. In the event of breakdowns, the expected damage is not only limited to monetary aspects, but may also affect the operational safety of machines and systems [12]. In order to implement effective safety concepts, systems must be secured comprehensively and, if possible, already in the development stage (Security by Design) and on several levels (Defense in Depth) [13, 14]. The realization of IT security measures therefore, takes a central position within the framework of

Industries 4.0. It is at the same time the cornerstone for successful networking and one of the key challenges for companies on the road to Industrie 4.0 [2]. To unfold the full potential of Industrie 4.0 applications, resilient and trusted components have to be developed on the asset level (Security by Design) and topics of secure networks, secure processes, secure services, and secure data addressed on the system level following a Defense in Depth approach [14–16]. The approach on hand does not incorporate security aspects yet, but can be extended by further dimensions to regard vulnerabilities and corresponding security technologies.

3.4 Topological Approach

Our topological approach populates RAMI 4.0 with relevant technologies and thereby, describes absolute topologies within the Industrie 4.0 landscape and interrelations between technologies. *Relevant* refers to technologies that are particularly important in the context of Industrie 4.0. Major technologies have hereby been identified based on core Industrie 4.0 paradigms such as *Engineering 4.0*, *Smart Factory 4.0*, and the development towards *Smart Products and Services*.

3.4.1 Requirements to the Topological Approach

Aiming at supporting the introduction of Industrie 4.0 solutions in SME, any concept has to fulfill certain criteria to enable a flexible and effective usage in practice. The requirements of the topological approach are shown in Fig. 3.2. Firstly (**R1**), every relevant technology in the context of Industrie 4.0 has to be considered and presented in a structured manner. As manifold industries are involved, generic technological categories rather than specific manifestations are to be preferred [1]. Secondly (**R2**), the overall concept along with the selection of technologies has to be implemented dynamically. Adjustments and adaptations of the concept to the individual, company-specific, circumstances have to be possible [2]. Thirdly (**R3**), due to the substantial importance of IT security, any concept addressing Industrie 4.0 should either already cover security aspects or provide a foundation for subsequent amendments [16].

3.4.2 Global Structure of the Topological Approach

Technologies are assigned to the corresponding levels based on their area of usage. While sensors and actuators can be clearly assigned to the *field device level*, computer and terminals may be used as discrete stations to control production facilities or on a higher level within departments for engineering or administrative purposes. Servers can equally be used on the *enterprise level* or, if provided by external partners, on

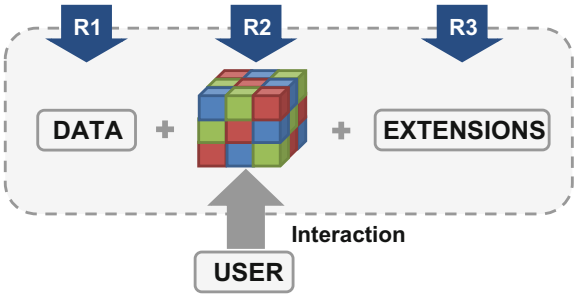


Fig. 3.2 Requirements for topological approach

the *connected world level*. The assignment of technologies to the hierarchy levels of RAMI 4.0 is therefore not definite but depends, within boundaries, on the concrete use case.

The positioning of technologies on the second axis of RAMI 4.0 (layers) depends on their core functionalities. Technologies that are providing means for communication like transmission technologies or communication protocols are assigned to the *communication layer*. Technologies providing specific functions like software products for engineering purposes or control software are assigned to the *functional layer*.

Technologies can be used throughout different product lifecycle phases. While there are certain technologies that are used regardless of the currently viewed lifecycle phase like computers and servers, technologies, such as x-in-the-loop, will be primarily used throughout the *type* lifecycle phase. The global view of the topological approach is shown in Fig. 3.3.

The three-dimensional design along with its dynamic usability concept allow users to view the model from six different perspectives, change the currently displayed level, layer, and lifecycle phase and actively interact with its content using intuitive multi-touch gestures. By selecting any technology, additional information regarding specific manifestations and related use cases are provided. This concept allows for adding further informational layers to the model, e.g. addressing vulnerabilities and corresponding security mechanisms. The detailed view of Instance, Production is shown in Fig. 3.4.

3.4.3 Mapping Technologies in RAMI 4.0

Cloud services represent a model for providing scalable, ubiquitous and on-demand retrievable network resources that can be flexibly adapted to current demand. Platforms in the context of industry 4.0 describe open application platforms for service orchestration [11].

Web services are services provided by applications or machines that can be used by other entities within a network. Offering independent, frequently standardized and modularized services, they play a central role within M2M communication within a service-oriented software architectures (SOA) [16].

Business Intelligence (BI) describes all concepts, models and information systems that are designed to support, carry out and control operational activities. Information systems (IS) can, in turn, be described as socio-technical systems, which aim to provide information and communication skills including applications and systems to support collaboration [17].

The integration of cyber-physical systems within industrial applications offers numerous possibilities for process monitoring, as well as remote services. **Smart** sensors, as well as smart material systems provide information about material stresses and loads. The data obtained is used for developing new business models, products or **services** [18].

Before physical production, manufacturing, assembly and testing processes can be simulated, the components' behavior and material stresses must be analyzed. Various product configurations and interactions third-party components can be simulated via powerful and interoperable software systems like x-in-the-loop methods in an **application or control software** [10].

To exchange model and simulation data, **software interfaces** are required, that enable cross-platform data exchange. When it comes to cross-company collaboration, integrated data security functionalities and protection of intellectual property gains importance [19].

Virtual and Augmented Reality (VR/AR) technologies are increasingly being used for industrial purposes. VR describes a computer-generated virtual environment that

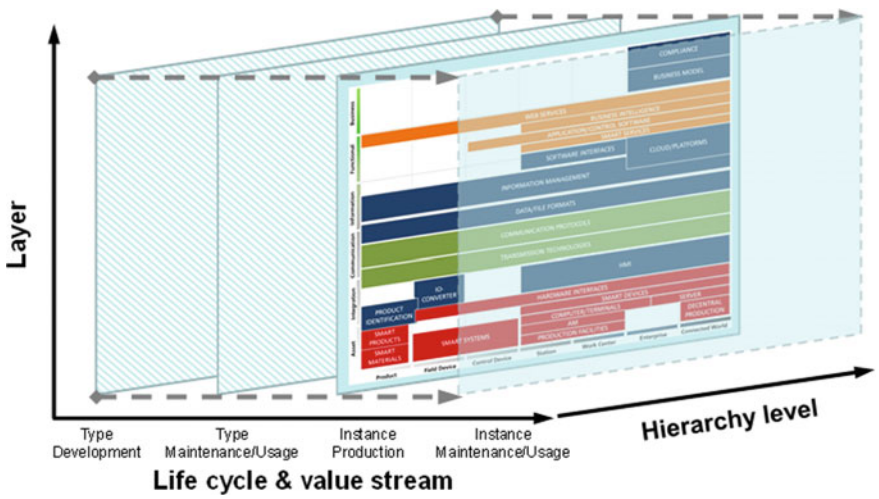


Fig. 3.3 Global view of topological approach to map technologies in RAMI 4.0

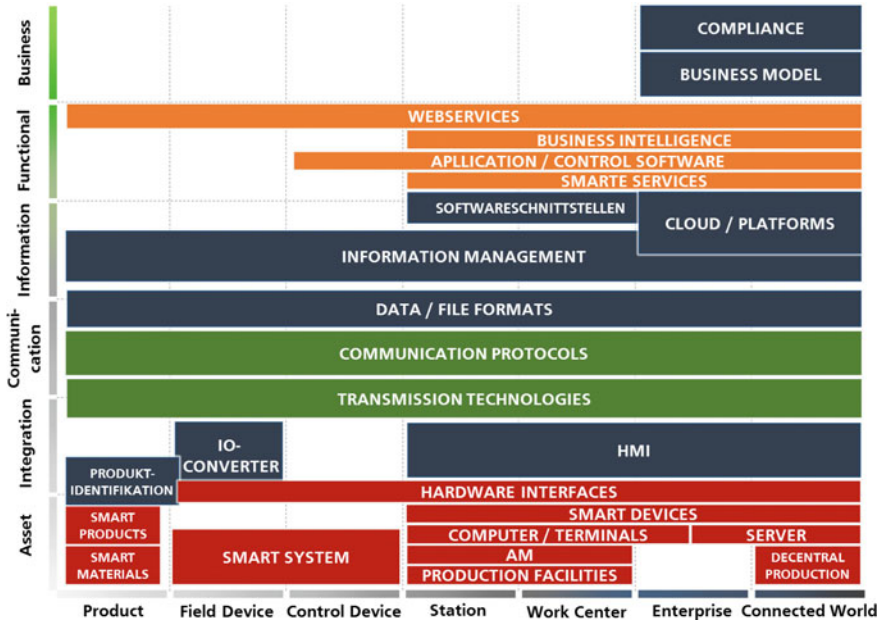


Fig. 3.4 Technology Landscape 4.0

serves as a HMI. The **human-machine interaction** can hereby take place in a multimodal way—addressing tactile or kinesthetic senses in addition to an audiovisual stimulation. AR refers to an expanded reality, in which the physically perceived reality is enriched by computer-generated information. Apart from VR/AR applications, **smart devices** providing context sensitive information can be used for alerts or tracking purposes and displace static, often paper based, information tools [10, 20].

Through digitization of value-added processes, large quantities of data are created. The efficient management and utilization of this data is a core task in Industrie 4.0. Only by deploying structured recording, storage, analyses and data provision processes in an **information and data management** system inefficient media breaches can be avoided and insights generated [21].

A prerequisite for cross-platform data exchange is the definition and use of uniform standards for **data formats**. In order to avoid redundant data generation and data losses, data formats must be fully transferable [22].

In order to ensure the collaboration of elements involved in the value-added process, their ability to communicate via platforms, interfaces and various transmission technologies must be ensured. The core technology cross-company **communication protocols**, and therefore for vertical and horizontal integration, is the Internet technology.

For wireless **data transmission** in industrial applications various open and proprietary technologies are available. In addition to WLAN, Infrared, NFC and Bluetooth,

or further technologies based thereupon (e.g. ZigBee) are used. Mobile network technologies such as HSPA (High Speed Packet Access) or LTE (Long Term Evolution) can also be used for mobile networking of employees and systems. The individual technologies differ in frequencies used, transmission rates, range, as well as power consumption [23].

Manufacturing needs to be dynamically adaptable to changing demands and requirements. Therefore, modular production systems with interchangeable components need to be deployed. The usage of such components requires standardized, platform independent **hardware interfaces** to assure a sound communication among the different components. These hardware interfaces can either be wireless or tethered [10].

The **identification of products** and objects along the value chain is a core element of Industrie 4.0. Besides active technologies based on embedded systems, in practice, passive identification technologies via RFID, barcodes or data matrix codes are used. In contrast to active technologies, passive technologies allow only a temporary intelligence of the products, associated with the demand for additional reading devices. Alternatively, objects can also be identified and addressed directly through their assigned IP addresses [24].

Additive Manufacturing (AM) is a primary forming process in which the object geometry is produced by laminating volume elements of the same thickness. This allows a flexible production of any geometry based on three-dimensional CAD models. The potential for this lays above all in the freedom regarding component design, the elimination of special tools and machines, the possibility for a flexible individual parts production and the further reduction of development times [25].

Modern sensor-actuator systems are distinguished into passive, adaptive and active systems. While the characteristics of passive systems are firmly defined due to their construction and the materials used, adaptive systems can be adapted to altered environments in discrete stages. Semi-active systems even allow for continuous adaptation. Both adaptive and semi-active systems are often based on passive components such as springs and dampers and therefore limited in their solution space. This limitation is overcome by the combination of sensor and actuator elements together with controls to active, respectively smart, systems. Hereby, sensors and actuators are not additionally attached to the structure but directly integrated into it. Core components of **smart systems** are **smart materials**. These are often integrated with auxiliary materials into a multifunctional material system and fulfill requirements to actuators and sensors, as well as load-bearing tasks [26].

3.5 Benefits of the Topological Approach

Figure 3.4 presents a two-dimensional projection of the topological approach. In order to improve clarity within this paper, a projection of the hierarchy levels, as well as the layers, for the product lifecycle phase *instance—production* is performed. Its elements represent relevant technologies for the respective product lifecycle phase.

Through the population of RAMI 4.0 with generic technologies, the approach enables users to get a profound, yet swift, overview of the Industrie 4.0 landscape. Core elements are displayed and their topological relationships exposed. The dynamic interaction with the three-dimensional model further reveals the spread of technologies throughout the different product lifecycles. Users can navigate through the dimensions using intuitive controls that are based on current smartphone applications and support the future deployment of VR technologies.

The model is used in many projects as a knowledge foundation along the different stages of Industrie 4.0 implementation as described by the Generic Procedure Model to introduce Industrie 4.0 in SME (GPMI 4.0) [15]. It supports the preparation, analyzation, idea generation and valuation of solutions and concepts for the implementation of Industrie 4.0 in small and medium enterprises. The approach thereby fosters the implementation of Industrie 4.0 and enables practitioners to develop and validate their technological roadmap.

Regarding the future development and widespread of RAMI 4.0 the topological approach helps to reify the abstract RAMI 4.0 without omitting its generic applicability to the diverse industries and company specific backgrounds. It is the first model to actively promote the usage of RAMI 4.0 in practice and therefore, will create valuable feedback from practice and support the future development of RAMI 4.0. Especially regarding IT security, the topological approach sets the foundation for adding another dimension to RAMI 4.0 that specifically addresses security issues in Industrie 4.0.

The approach addresses the prior set requirements as every currently relevant technology is covered and due to the generic and dynamic design of the model newly emerging technologies can be assigned to an already existing category or simply added to the model in the future (R1). The three-dimensional and VR-ready design allows for manifold user interaction to display information in a profound, yet understandable and clear manner (R2). Due to its design, further informational layers can be added and thereby, the original model amended. Based on the technologies described associated vulnerabilities and corresponding security mechanism could be added in the future. The approach, therefore, already lays the foundation for a later integration of IT security aspects with the security-by-design concept (R3).

3.6 Conclusion

Industrie 4.0 requires a horizontal and vertical integration of the industrial landscape towards a fully digitalized enterprise. This integration brings along manifold security issues, which to address is one of the key prerequisites for Industrie 4.0. The diversity of technologies employed in industrial IT, the complexity of networks and new business models, as well as the challenges of securing those, are hurdles companies need to overcome.

The aim of RAMI 4.0 is, among others, to support the implementation of Industrie 4.0 and to help companies to overcome these hurdles. The approach on hand

improves its direct applicability as the abstract model design is populated with more concrete technologies and their relationships among one another are exposed. Therefore, the transmission of RAMI 4.0 from research to practice is facilitated. Despite its importance for Industrie 4.0, IT security is so far not covered in RAMI 4.0. The topological approach described lays the foundation for an expansion of RAMI 4.0 to add another dimension addressing security aspects. Further research is needed to develop such an expansion and analyze vulnerabilities of the relevant technologies described, as well as corresponding security mechanisms.

References

1. H. Kagermann, W. Wahlster, J. Helbig, Recommendations for implementing the strategic initiative Industrie 4.0—Securing the future of German manufacturing industry (2013)
2. BITKOM, VDMA, and ZVEI, Umsetzungsstrategie Industrie 4.0—Ergebnisbericht der Plattform Industrie 4.0 (2015)
3. R. Anderl, A. Picard, Y. Wang, J. Fleischer, S. Dosch, B. Klee, J. Bauer, Guideline Industrie 4.0—Guiding principles for the implementation of Industrie 4.0 in small and medium sized businesses (2015)
4. P. Adolphs, H. Bedenbender, D. Dirzus, M. Ehlich, U. Epple, M. Hankel et al., Referenzarchitekturmodell Industrie 4.0 (RAMI 4.0) (2015)
5. Y. Wang, T. Towara, R. Anderl, Topological approach for mapping technologies in reference architectural model Industrie 4.0 (RAMI 4.0), in *Proceedings of the World Congress on Engineering and Computer Science 2017*. Lecture Notes in Engineering and Computer Science, 25–27 October, 2017, San Francisco, USA, pp. 982–990
6. IVI, Industrial Value Chain Reference Architecture (2016), https://iv-i.org/docs/IndustrialValueChainReferenceArchitecture.pdf
7. IIC, The Industrial Internet Reference Architecture v1.8 (2017), http://www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf
8. VDI, Industrie 4.0 Service Architecture—Basic Concepts for Interoperability (2016)
9. VDI, Industrie 4.0 Components—Modeling Examples (2016)
10. T. Bauernhansl, M. ten Hompel, B. Vogel-Heuser, Industrie 4.0 in Produktion, Automatisierung und Logistik (Springer Vieweg, Wiesbaden, 2014)
11. M. Barbian, I. Gräßler, F. Piller, C. Gülpen, P. Welp, H. Kamal et al., Digitale Chancen und Bedrohungen—Geschäftsmodell für Industrie 4.0 (2016)
12. R. Anderl, Industrie 4.0—Technological approaches, use cases, and implementation. Automatisierungstechnik **63**, 753–765 (2015)
13. M. Jochem, W. Klasen, L. Linke, L. Jaenicke, T. Garner, M. Stolz et al., Security in RAMI 4.0 (2016)
14. DHS, Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies (2016)
15. Y. Wang, G. Wang, R. Anderl, Generic procedure model to introduce Industrie 4.0 in small and medium-sized enterprises, in *Proceedings of the World Congress on Engineering and Computer Science 2016, WCECS 2016*. Lecture Notes in Engineering and Computer Science, 19–21 October, 2016, San Francisco, USA, pp. 971–976
16. C. Eckert, IT-Sicherheit. Konzepte-Verfahren-Protokolle, 8th edn (Oldenbourg, 2013)
17. T. Clark, M. Jones, C. Armstrong, the dynamic structure of management support systems: theory development, research focus, and direction. MIS Q. **31**(3), 579–615 (2007)
18. R. Anderl, Industrie 4.0—advanced engineering of smart products and smart production, in *19th International Seminar on High Technology*, Piracicaba, Brasil (2014)

19. M. Otter, T. Blochwitz, M. Arnold, Functional Mock-up Interface for Model Exchange and Co-simulation (Vers. 2.0) (2014)
20. R. Blach, Virtual reality technology: an overview, in *Product Engineering: Tools and Methods Based on Virtual Reality,* ed. by D. Talaba, A. Amditis, (Springer, Dordrecht, NL, 2008), pp. 21–64
21. O. Drumm, Durchgängiges Engineering in Industrie 4.0-Wertschöpfungsketten (2016)
22. M. Pratt, Introduction to ISO 10303—the STEP standard for product data exchange. J. Comput. Inf. Sci. Eng. (Technical Note) **1**(1), 102–103 (2001)
23. N. Chhabra, Comparative analysis of different wireless technologies. Int. J. Sci. Res. Netw. Secur. Commun. (Research Paper) **1**(5), 13–17 (2013)
24. R. Anderl, O. Anokhin, A. Arndt, Effiziente Fabrik 4.0 Darmstadt—Industrie 4.0 Implementierung für die mittelständige Industrie, in *Industrie 4.0 grenzenlos,* ed. by U. Sendler (Springer Vieweg, Berlin, Germany, 2016), pp. 121–136
25. A. Gebhardt, *Additive Fertigungsverfahren, Additive Manufacturing und 3D-Drucken für Prototyping - Tooling - Produktion*, 5th edn. (Hanser, Munich, Germany, 2016)
26. H. Janocha, *Adaptronics and Smart Structures. Basics, Materials, Design and Applications*. 2nd edn (Springer, Berlin, Heidelberg, 2007)

Chapter 4
Modeling of Vehicle-Cargo Interaction Under Different Environments

Frank Otremba and José A. Romero

Abstract The safety of any transport system depends on a multitude of conditions, parameters and circumstances. In this regard, the interaction of the carried cargo with the carrying vehicle represents a factor influencing the overall safety of any transport. The effects of cargo on the vehicle have to do with the vibration or shifting of the cargo, affecting the lateral stability of the vehicles and the braking performance. Such interaction has been associated to road crashes and maritime vehicles capsizing. Simulation of cargo-vehicle interaction thus represents an interesting topic when a reduction in transport accidents is pursued. In this paper, the fundamentals principles for simulating the interaction of the liquid cargo and the carrying vehicle, is presented. In the case of a road transportation, the proposed simplified simulation methodologies, show good agreement with a full-scale test.

Keywords Braking performance · Experimental approach · Newton approach
Ship stability · Sloshing · Transition matrix approach

4.1 Introduction

The level of safety of a transport depends in great extent, of factors associated to the dynamic forces that are developed due to the motion of the cargo on the vehicle. The mobility of the cargo within the vehicle is exhibited when the cargo is liquid or is hanging from the vehicle chassis. However, the most common situation of cargo vibration on the vehicle, occurs in the case of liquid cargo. Liquid cargo can include a variety of products, of different densities and levels of danger. Physical properties

F. Otremba
Federal Institute of Materials Research and Testing (BAM), Unter den Eichen 87, 12205 Berlin, Germany
e-mail: frank.otremba@bam.de

J. A. Romero (✉)
Queretaro Autonomous University, Rio Moctezuma 249, San Juan del Rìo, Queretaro, Mexico
e-mail: jaromero@uaq.mx; jromero1@hotmail.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_4

affecting the vehicle dynamics, however, include the density and viscosity of the carried cargo.

The density of the cargo influences the inertia of the liquid, while its viscosity affects its damping characteristics. The density plays a crucial role in the context of the regulations involving trucks' weight and dimensions, because in order to observe those regulations, the vehicles have to travel partially filled.

Partially filled tankers pose particular stability risk to the transport, regardless of the transportation mode. The motion of the liquid cargo in the context of transportation, is called sloshing.

Sloshing of the cargo implies at least three negative effects on the vehicle dynamics: (i) lateral cargo shifting that affects the longitudinal and/or lateral position of the vehicle's center of gravity; (ii) a vertical shifting of the center of gravity of the vehicle, further affecting the longitudinal/lateral stability of the cargo; and (iii) the potential dynamic coupling of the liquid and vehicle vibration.

Figure 4.1 illustrates cargo shifting situation, in case of a cylindrical road tanker, travelling at several fill levels while negotiating a turn.

Figure 4.2 illustrates the calculation of the roll moments for a sloshing and non-sloshing cargoes. It is assumed in this figure, that the first mode of sloshing vibration is occurring, which is acceptably realistic. However, the lateral acceleration, defining the lateral slope of the fluid's free surface, is exaggerated, to better visualize the geometric relationships so created.

From the isosceles triangle in Fig. 4.2, the following expression for a can be deduced:

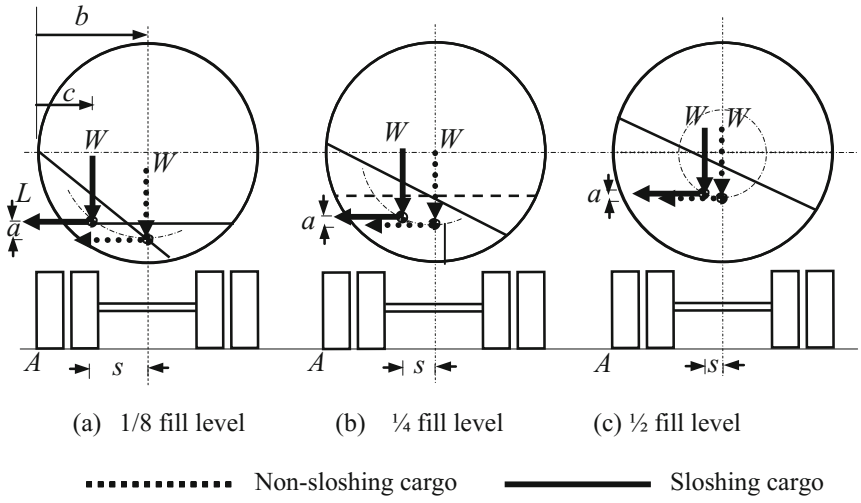


Fig. 4.1 Vertical and lateral shifting of the cargo, when a partially loaded road tanker is subjected to lateral accelerations

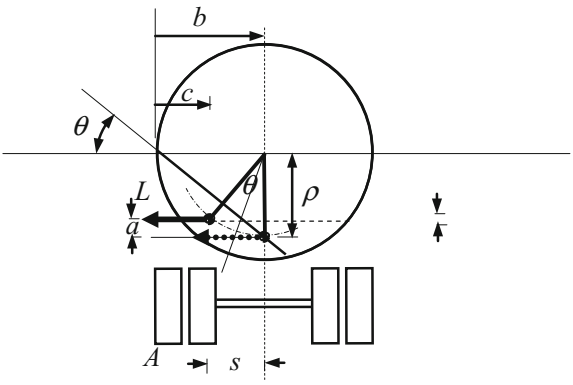


Fig. 4.2 Representation of the vertical and lateral shifting of the center of gravity of the cargo (one-quarter fill level, maximum lateral acceleration)

$$a = \rho \sin(\theta/2)(2)(\sin(\theta/2) \tag{4.1}$$

When negotiating a turn, there are two moments acting on the vehicle: one to destabilize it, and the other to stabilize it. The destabilizing moment is represented by the inertia force derived from the turning maneuver, while the stabilizing moment derives from the vertical acceleration (gravity). When the cargo shifts both vertically and laterally, the stabilizing moment is affected. Figure 4.3 illustrates the effect of lateral acceleration and fill level, on the resulting overall moment (stabilizing minus destabilizing moments). According to these results the lateral acceleration influences in a significant way the increase of the moments to destabilize the vehicle. For the lateral acceleration considered (0.068 g), the increase of roll moment varies from 5 to 20%, as a function of the fill level. However, it should be noted that in the case of the minimum fill level considered, the inertia associated to this volume of cargo, would represent a very small amount with respect to the total inertia of the vehicle (chassis and body). Consequently, while the increase of the roll moment is the greater for this low fill level, the consequences on the lateral stability of the vehicle, would be minimal.

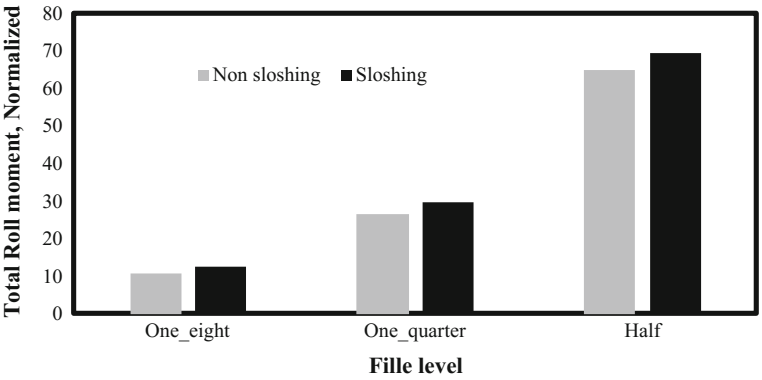
On the other hand, the longitudinal dynamics of the sloshing cargo keeps a close relationship with the lateral effect described above. Figure 4.4 describes this longitudinal acceleration situation, where the resulting pressure on any point on the front head or surface of the tank, is the result of a vertical pressure p_v, plus a longitudinal pressure p_L. Both of these pressures depend on the longitudinal acceleration, as such acceleration defines the angle of the liquid's free surface, as follows:

$$p_v = \rho\, g y = \gamma\, y \tag{4.2}$$

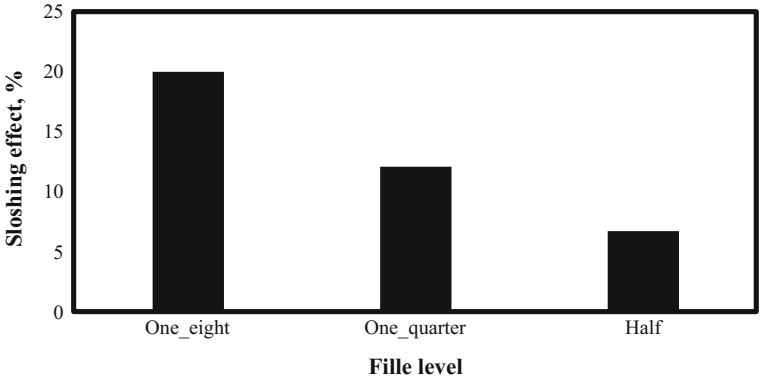
$$p_L = \rho\, a_L\, L \tag{4.3}$$

where ρ_L is the mass density of the liquid (kg/m^3); a_L is the longitudinal acceleration of the vehicle; and g is the acceleration due to gravity.

This paper is based upon a WCECS 2017 Conference paper [1].



(a) Effect of the cargo shifting on the level of roll moment.



(b) Cargo shifting effect on the increase of the roll moment.

Fig. 4.3 Cargo shifting effect on the roll moments (lateral acceleration of 0.667 m/s^2)

4.2 Modelling Principles

While the representations described above correspond a steady state condition, for both the lateral and the longitudinal behavior of the liquid in the tank, there is the need to model the oscillation characteristics of the fluid within the tank. For that, a simplified model has been proposed and used in a multitude of cases to simulate such dynamics. Until recently, however, the dynamic properties of the pendulum representing the sloshing cargo, has been based upon experimental data [2]. Such pendulum property is its length. The basis for such approach derives from the free vibration of a liquid in a rectangular tank, called "gravity waves", according to the following equation [3]:

$$c = \lambda\, f = \sqrt{\left(\frac{g}{\kappa} \tanh\, \kappa h\right)} \tag{4.4}$$

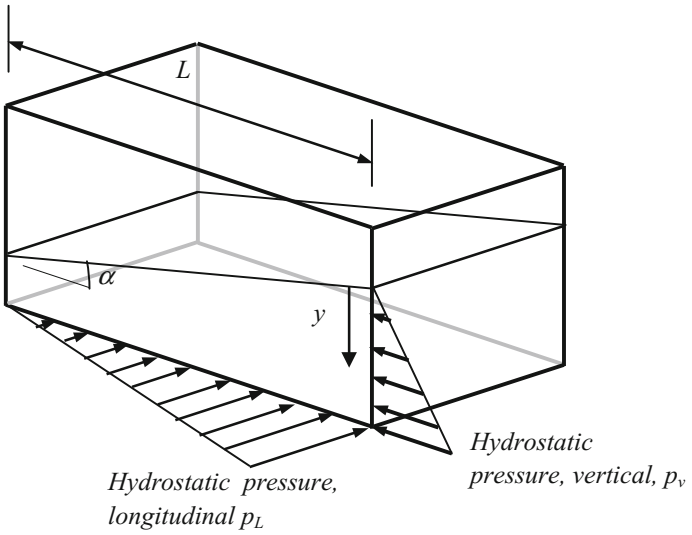


Fig. 4.4 Superposition of hydrostatic pressures (vertical and longitudinal)

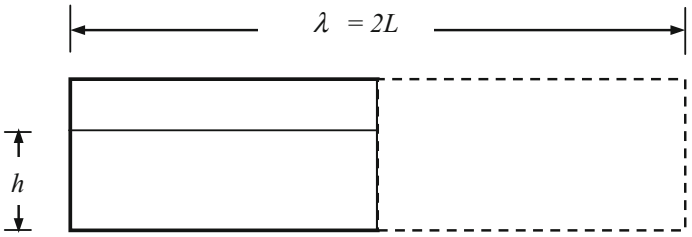


Fig. 4.5 Natural free sloshing frequency in a rectangular container

where κ is the wave number, equal to $2\pi/\lambda$. This is illustrated in Fig. 4.5.

Consequently, the pendulum representing the sloshing cargo, has a length as a function of the following equation:

$$f = \frac{1}{2\pi}\sqrt{\frac{g}{l_p}} \tag{4.5}$$

According to this approach, it if feasible to simulate the transient and steady state of the sloshing of the cargo within the tankers.

The mechanical analogy of the sloshing cargo with the pendulum, involves a mechanical system that is able to vibrate, and that is supported by the chassis of the vehicle. That part of the vehicle, further represents a mechanical spring-mass-damper system. Figure 4.6 illustrates the resultant model, in which the simple pendulum rests on an equivalent mechanical system representing the sprung mass of the vehicle.

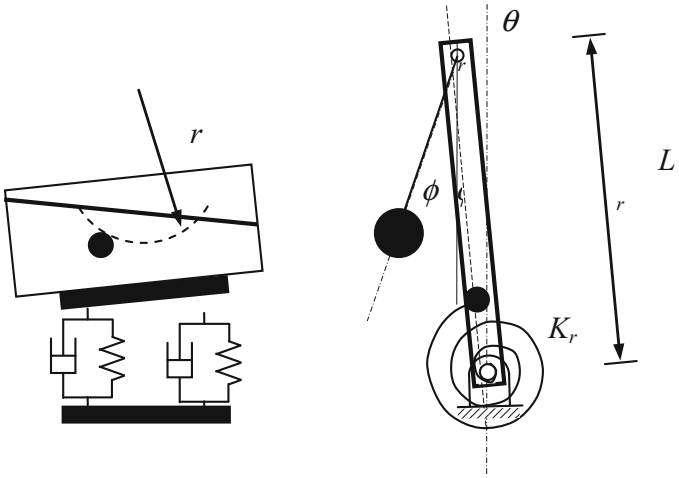


Fig. 4.6 Model of sloshing cargo, supported on the sprung mass of the vehicle

The resulting mechanical system representing the sloshing cargo and the sprung mass of the vehicle, consists of a two degree-of-freedom system, namely, the sloshing angle of the first mode of vibration of the fluid, and the pitch or roll motion of the sprung mass of the vehicle. Derivation of the equations of motion can be done through the use of Newton's Second Law.

On the other hand, the linear mathematical model resulting from the combination of the different mathematical approaches just described, are solved on the basis of the solution of a first order system. For this, the second-order equations are expressed as vector state variables, as follow [4]:

$$\{\dot{y}(t)\} = [A]\{y(t)\} + [B]\{Y(t)\} \tag{4.6}$$

where the matrices $[A]$ and $[B]$ are coefficient matrices, expressed in terms of the mechanical properties of the system (stiffness, damping, inertial and dimensional characteristics). The time-discrete exact solution of this system of equations, results as follows [4]:

$$\{y(t + \Delta t)\} = [\Phi]\{y(t)\} + [\Gamma]\{Y(t)\} \tag{4.7}$$

where $[\Phi]$ is the transition matrix; $[\Gamma]$ is the particular response matrix; and $\{Y(t)\}$ is the perturbation vector. The size of the different vectors and matrices of this method, thus depends on the number of degrees of freedom of the resulting system. The state vector for a two-degree-of-freedom system, is four, as it involves the state variable and its derivative. On the other hand, a number of particular response matrix $[\Gamma]$ can be generated, as a function of the number of perturbations, $\{Y(t)\}$.

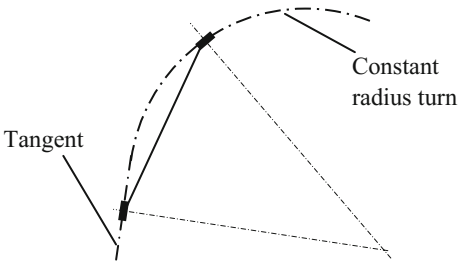


Fig. 4.7 Methodology two calculate the instantaneous lateral acceleration of a vehicle while incorporating into a curved infrastructure

To simulate the circulation of the vehicles along the infrastructure, there is still another modelling principle that must be considered for the simulation of the tanker vehicles dynamic response to the geometrical characteristics of the infrastructure. This principle considers the instantaneous lateral acceleration on the vehicle when it incorporates to the curved part of the infrastructure, as a function of the geometrical characteristics of the road, rail or maritime infrastructure. Figure 4.7 describes the situation of the vehicle when entering a curve portion of the infrastructure. The respective formulation is described in [5], to calculate the instantaneous radius of curvature when the vehicle passes from a tangent to a constant radius turn.

4.3 Examples

The models described in the previous section, as well as the equation solving methodology, are used to simulate two common situations in liquid cargo transportation. On the one hand, an emergency braking maneuver is simulated, of a trailer tanker combination. For that, the model consists of a longitudinal pendulum, and the vehicle is equipped with Antilock Brake System. Figure 4.8 schematically illustrates the trailer tanker [1]. Two of the seven load compartments, were instrumented to sense the hydrostatic pressure on the compartment's wall.

The braking algorithm consisted of a bang-bang control, set to an optimal wheel slip of 0.2, leading to a maximum braking coefficient of 0.7. The role of sloshing on the braking efficiency of the vehicle, is defined on the basis of the stopping distance. Figure 4.9 illustrates the measured braking acceleration, with maximum decelerations on the order of 0.65 g [6].

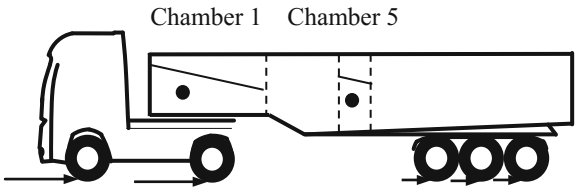


Fig. 4.8 Schematic representation of a tractor-trailer combination for carrying liquid cargo

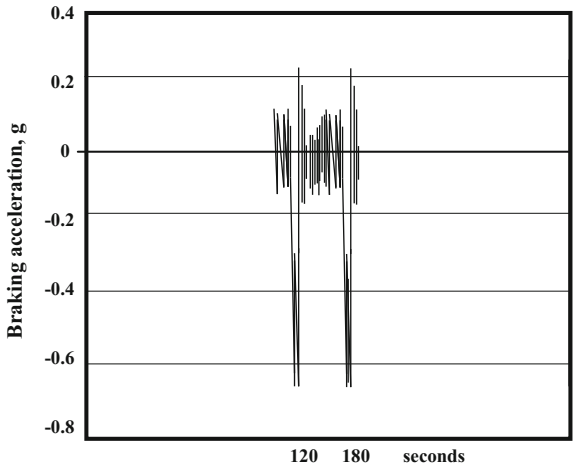


Fig. 4.9 Experimental data of the braking acceleration [1]

Figure 4.10 illustrates the measured and simulation data, of the hydrostatic pressure on the pressure sensors of the two instrumented chambers. According to these data, there is a close correlation between both data. Furthermore, Fig. 4.11 illustrates a detailed view of the hydrostatic pressure once the vehicle was in rest, where the natural frequency of the sloshing was appropriately reproduced through the simulation.

According to these results, the longitudinal dynamics of the sloshing cargo—vehicle system, can be closely reproduced through the use of simplified models, for both, the vehicle and the sloshing cargo.

The simulation methodologies described above are now used to simulate the roll behavior of a ship when travelling along a river infrastructure.

Figure 4.12 describes the roll plane of a ship when subjected to a lateral acceleration. As it was described in the previous section, the sloshing cargo is simulated through a simple pendulum, whose characteristics are defined on the basis of a validated methodology to calculate the natural sloshing frequency of the fluid within the vessel.

A series of simplifications are assumed for this model, including the bottom of the river tanker, which is semicircular instead of flat [7].

Figure 4.13 illustrates the simulation results from this model. It should be noted that the non-sloshing condition was obtained by artificially increasing the damping of the sloshing cargo within the tanker. According to these results, the sloshing cargo represents a dangerous situation, as the maximum roll angle increases in a certain percentage.

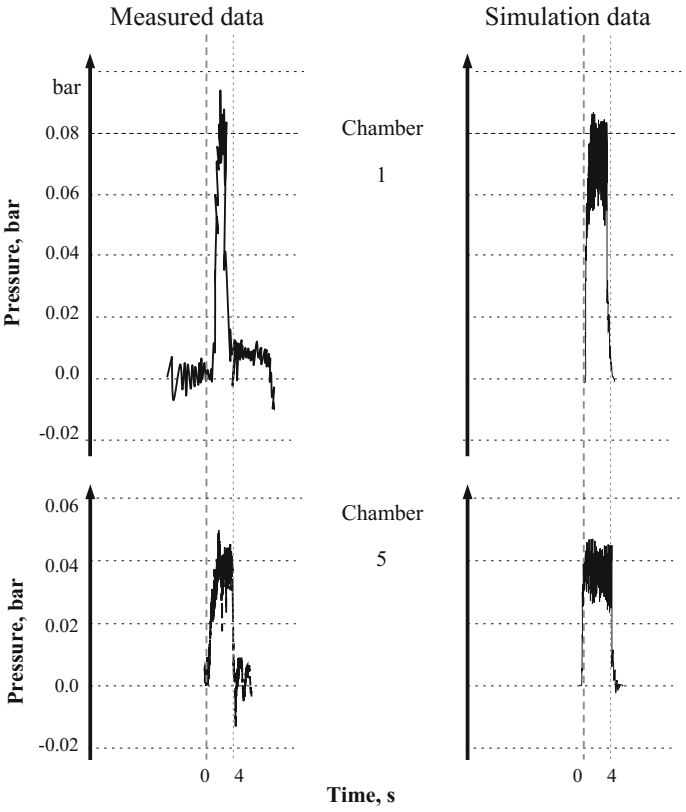


Fig. 4.10 Simulation and measure data for the pressure on the vehicle's chambers

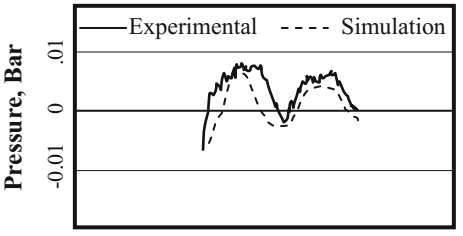


Fig. 4.11 Simulation and experimental data of the residual vibration of the sloshing cargo

4.4 Conclusions

A simplified physical modelling and an equation-solving algorithm, have been presented to simulate the behavior of a vehicle (tanker) under road and maritime infrastructure conditions. For the road infrastructure, the emergency braking of a semi-trailer tanker, was presented, and for the maritime environment, the travelling of a river ship was modelled. In the case of the braking maneuver, experimental results

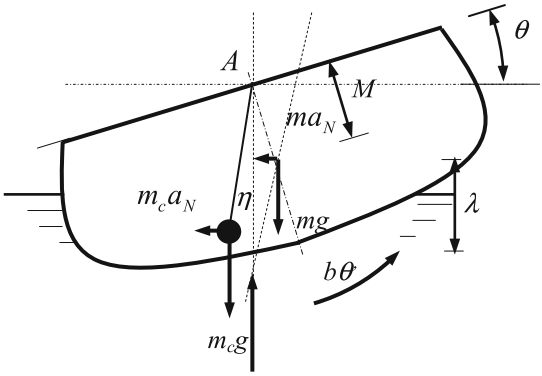


Fig. 4.12 Roll model for a ship when subjected to lateral accelerations due to infrastructure's geometric design

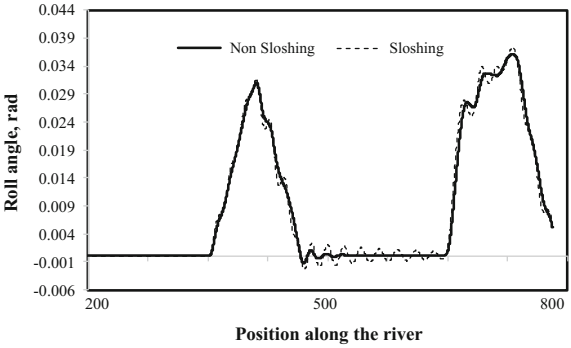


Fig. 4.13 Time response history of the roll angle for a sloshing and non-sloshing cargo when the ship travels along a winding river

are compared with theoretical results, finding a good agreement between both series of data. The simulation of a ship along a river infrastructure containing several turns, reveals that the roll stability of the ship can be affected by the sloshing cargo. The simulation principles can be considered for other environments and conditions. For example, the performance of a braking maneuver under the conditions of changing, simultaneously, the speed of the vehicle, and the direction of the vehicle.

References

1. F. Otremba, J.A. Romero, Modelling a partly filled road tanker during an emergency braking, in *Proceedings of the World Congress on Engineering and Computer Science 2017*. Lecture Notes in Engineering and Computer Science, 25–27 October, 2017, San Francisco, USA, pp. 610–614
2. J.A. Romero, R. Hildebrand, W. Ortiz, J.C. Gomez, Natural sloshing frequencies of liquid cargo road tankers. Int. J. Heavy Veh. Syst. **12**(2), 121–138 (2005)
3. C.W. Elmore, M.A. Heald, *Physics of Waves* (Dover Publications Inc., New York, 1989)
4. L. Meirovitch, *Elements of Vibration Analysis*, 2nd edn (Mc Graw Hill Int., 1986)

5. J.A. Romero, A.A. Lozano-Guzmán, G.M. Arroyo-Contreras, Study of the effect of variations in bogie centre distance on rail car roll performance, in *Proceedings, 2015 IFToMM World Congress*, Taipei, Taiwan (2015)
6. BAM, Sloshing tests on a tanker semi-trailer, Federal Institute for Materials Research and Testing (BAM), Hazardous goods tanks and accident mechanics (In German, 2004)
7. F. Otremba, J.A. Romero, Analysis of the capsizing of a tanker in the Rhine river. ASME paper IMECE2017-70488 (2017)

Chapter 5
Temperature Optimized Hydrolysis of Acetic Acid Catalyzed Magnesium Hydride for Hydrogen Generation in a Batch System Hydrogen Reactor

Joshua Adeniyi Adeniran, Romeo Sephyrin Fono-Tamo, Esther Titilayo Akinlabi and Tien-Chien Jen

Abstract Interests in hydrogen storage in magnesium hydride (MgH_2) is on growing due to its high gravimetric and volumetric hydrogen concentration and less harmful environmental impact of its reaction. This study reports the usage of acetic acid as an accelerator in hydrogen generation experiment via hydrolysis, using a batch system hydrogen reactor. The study was conducted in two phases: investigation of role of substrate weight and catalyst concentration on hydrogen yield and optimization of temperature for hydrogen generation. The study investigated the roles of the organic acids in the reduction of the poor reaction kinetics limitation in MgH_2 through the catalytic characteristic of acetic acid. Three acetic acid concentrations were used in the study namely 40, 50 and 60 wt%. Similarly, three substrate weights (0.4, 0.8 and 1.2 g) were investigated for the role of substrate weight on hydrogen yield. Application of the acetic acid accelerated the hydrogen yield across board. Similarly, the hydrogen yield increased with MgH_2 weight in the study. The highest hydrogen yield of 0.0189 L was recorded from 1.2 g MgH_2 at 30 °C and 50 °C respectively and 50 wt% acetic acid. The lowest hydrogen yield in the study was 0.0048 L hydrogen gas from 0.4 g MgH_2 at 50 wt% acetic acid. The experiment conducted at ambient temperature (25 °C) revealed comparable hydrogen yield with higher temperatures.

Keywords Acetic acid · Hydrogen generation · Hydrolysis · Reaction kinetics Temperature optimization · Thermodynamics

J. A. Adeniran (✉) · R. S. Fono-Tamo · E. T. Akinlabi · T.-C. Jen
Department of Mechanical Engineering Science, University of Johannesburg, Auckland Park Kingsway Campus, Johannesburg 2006, South Africa
e-mail: nikjocrown2000@yahoo.com; joshuaa@uj.ac.za

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_5

5.1 Introduction

Hydrogen storage could be achieved through various means such liquid or cold hydrogen storage (cryogenic), pressure hydrogen storage, steam reform, biohydrogen storage such as hydrogen storage in wastewater, carbon nanotubes, solid state hydrogen storage in metal and complex metal hydrides [1–3]. Solid state hydrogen storage particularly metal hydrides have attracted attention due to its relative safe storage thus making it an alluring candidate for an on-board vehicular fuel applications. Furthermore, most metal hydrides have high specific hydrogen volumetric and gravimetric volumes [4]. Magnesium hydride (MgH_2) is one of the leading candidates among metal hydrides for solid state hydrogen storage purpose. The development is due to some of the endearing characteristics of the MgH_2. For example, MgH_2 is relatively cheap compared to majority of the metal hydrides. In the same vein, magnesium (Mg) a major component of MgH_2 is abundant in earth crust. It is regarded as the eighth most abundant element in earth crust [5, 6]. Furthermore, other advantages of MgH_2 as a hydrogen storage substrate include its non-corrosive nature which makes it an environment benign hydrogen generation and storage material [7, 8]. Similarly, MgH_2 has high volumetric hydrogen storage capacity of about 7.6 wt%.

However, hydrogen storage in MgH_2 like most metal and complex metal hydrides is hampered by slow reaction kinetics and poor thermodynamics. The thermodynamic challenge being high reaction enthalpy and entropy. Consequently, high temperature is required to desorb the hydrogen embedded in the matrix of the MgH_2. This application of heat leads to increase in reaction cost thereby limiting scale up of the hydrogen storage technology. Furthermore, different researchers have explored use of various materials or compounds as catalysts in a bid to lessen the poor reaction kinetics and unfavorable thermodynamics limitations to the utilization of MgH_2 as hydrogen storage medium. Noble metal catalysts such as platinum have been used and reported to catalyze hydrogen desorption or release in metal hydrides, but the major limitation of the intervention is the prohibitive cost of platinum metal to large scale hydrogen storage and generation [9]. Moreover, some of the reaction products of noble metal catalysis of hydrogen storage in metal and complex metal hydrides are toxic to the environment [8]. The use of ball milling technique with its attendant increase in reaction surface area due to reduction in crystalline sizes of the substrates have been reported to increase reaction kinetics and improve thermodynamic properties of hydrogen storage experiments. Moreover, fracturing of the metal hydrides crystals due to the milling process improves nucleation reaction which further enhances hydrogen desorption [10]. However, despite the successes recorded in hydrogen generation using ball milling, the quantity of hydrogen released is not enough for the use of the technology in large scale production. Organic acids have also been used as accelerators in for hydrogen release from metal hydrides. The advantages of this technique include relative cheap low cost of the organic acids and their nontoxic/environmental friendly nature. For example, in our laboratory we have reported highest hydrogen generation of 0.018 L from a reaction of 1.2 g MgH_2 powder with 50 wt% acetic acid [7]. In another study from the same research

group, it was observed that hydrogen generation increased with substrate weight and acetic acid concentration accelerated hydrogen release when the reaction of MgH_2 pills was catalyzed with acetic acid [11]. The study reported maximum hydrogen generation of about 1.285 L. Sustainable hydrogen generation is crucial to the scale up of the technology and cost reduction is an integral component of it. To ameliorate the challenges of unfavorable thermodynamic and slow reaction kinetics that have limited the scale up of hydrogen storage in metal hydrides, interventions of increased reaction temperature are among the interventionist approach employed. However, the application of heat to the experiments further add to the reaction cost because of the cost of powering the temperature regulators/water baths used for such purposes. Consequently, a successful elimination of need to raise reaction temperature above room temperature would reduce reaction cost, thus driving the hydrogen storage technique towards sustainability. Thus, optimization of reaction temperature will provide insight into hydrogen yield with regards to reaction temperature. In this study, optimum reaction temperature for hydrogen generation in the designed reactor was investigated. Furthermore, the role of substrate weight and acetic acid concentration on hydrogen yield were examined.

5.2 Experimental Design and Method

5.2.1 Hydrogen Generation Reactor Operation

Figure 5.1 indicates the experimental design for the hydrogen generation reactor employed in the study. The reactor is made up of a three-neck round bottom flask which serves as the reaction vessel, thermostatic regulated water bath [Julabo TW20, Julabo GmbH (Germany)], the moisture absorbent unit, the flowmeter [T1000, Fujikin (Japan)] and the data logger.

The hydrogen generation experiment is essentially a hydrolysis reaction. The substrate (MgH_2 powder) was poured inside round bottom flask wherein acetic acid of various concentrations was released through the soxhlet apparatus attached to the middle neck of the round bottom flask. The outer left and right neck of the reaction vessel was attached to the thermometer and the tube for harvesting the hydrogen generated from the experiment respectively. The moisture absorbent in the design trapped the moisture in hydrogen thus ensuring only hydrogen is recorded by the flowmeter. The hydrogen generation was recorded using the data logger connected to the flowmeter [11].

An organic acid catalyzed the experiment—acetic acid [99.8%, Labchem, South Africa (SA)]. Magnesium hydride powder 99.8% purity (Rockwood Lithium, Germany) employed throughout the course of experiment as hydrogen storage media (substrate) was used as received (AR) from the supplier without further treatment with average particle size of 50 μm. The equation for the reaction of magnesium hydride with acetic acid catalyzed hydrolysis reaction can be seen in Eq. (5.1):

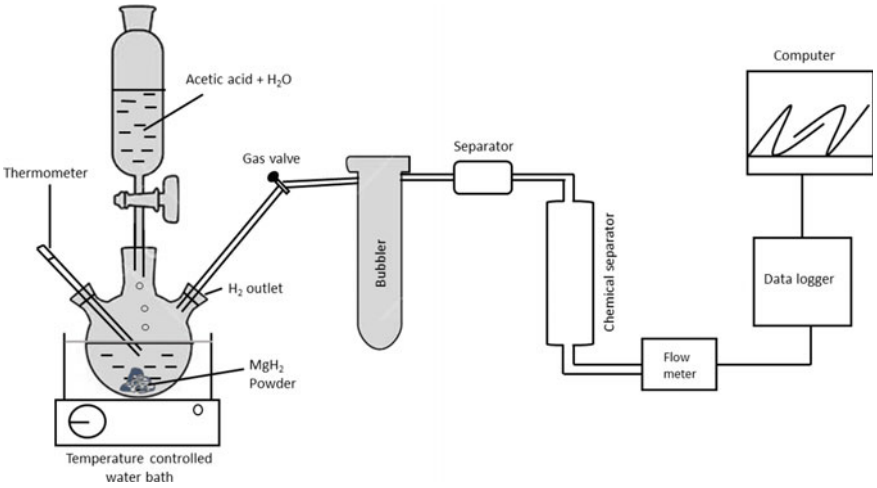


Fig. 5.1 Batch system hydrogen reactor experimental set up

$$2MgH_2 + 2H_2O + 2CH_3COOH \rightarrow Mg(OH)_2 + Mg(CH_3COO)_2 + 4H_2 + \Delta H_{r\times n} \quad (5.1)$$

The heat of reaction ($\Delta H_{r\times n}$) is approximately -277 kJ/mole.

The investigation of the impact of substrate weight on hydrogen yield was carried with three different MgH_2 powder weights (0.4, 0.8 and 1.2 g), except in 40 wt% acetic acid concentrations where only 0.4 and 0.8 g were utilized. Weighing of the substrate samples were carried out using BM-200 analytical balance (United Kingdom) with 0.0001 g repeatability to enhance weighing uniformity. The experiment was carried out at 25, 30, 50 and 60 °C external temperature. Furthermore, the experiment was conducted using different catalyst concentrations (40, 50, 60 and 70%).

5.2.2 Material Characterization of Reaction Substrate (MgH_2) Powder

Scanning electron microscopy (SEM), energy dispersive X-ray spectroscopy (EDS) analysis was conducted on the MgH_2 powder for adequate material characterization. The SEM analysis of the MgH_2 powder surface morphology was conducted using JSM 7600F Jeol ultra-high-resolution field emission gun scanning electron microscope [(FEG-SEM), (USA)] equipped with EDS was utilized for the EDS analysis.

5.3 Results and Discussion

5.3.1 Role of Substrate Weight and Catalyst Concentration on Hydrogen Yield

In this study, the roles of substrate (MgH_2 powder) concentration and acetic acid concentration on hydrogen yield were examined to find the optimum parameters that can enhance hydrogen generation. Investigation of role of MgH_2 powder concentration on hydrogen generation was conducted using 0.4 g, 0.8 g and 1.2 g MgH_2 powders at various acetic acid concentrations namely 40 wt%, 50 wt%, 60 wt% and 70 wt% respectively. The results of the study are presented in Figs. 5.2, 5.3, 5.4 and 5.5.

Figure 5.2, Present the result of hydrogen generation experiment at 40 wt% acetic acid concentration for 0.4, 0.8 and 1.2 g MgH_2 powder. From the results, it can be observed that highest hydrogen generation of about 0.016 L was recorded from 1.2 g substrate while by 0.0098 L hydrogen yield from 0.8 g MgH_2 compared to maximum hydrogen generation of 0.005 L obtained in the 0.4 g MgH_2 experiment.

From Fig. 5.3, 0.4 g MgH_2 powder recorded the least hydrogen yield of about 0.0056 L, followed by maximum hydrogen generation of about 0.013 L at 0.8 g while the highest hydrogen yield of 0.018 L obtained when 1.2 g substrate was hydrolyzed in 50 wt% acetic acid.

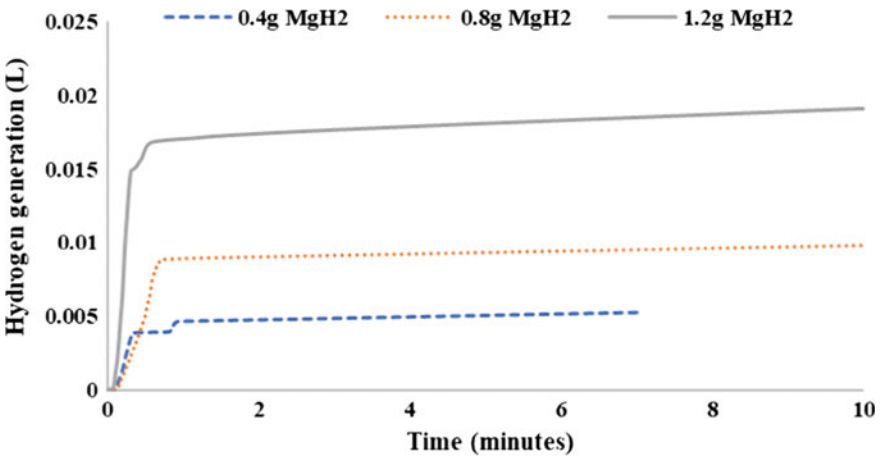


Fig. 5.2 Hydrogen generation at 40 wt% acetic acid concentration

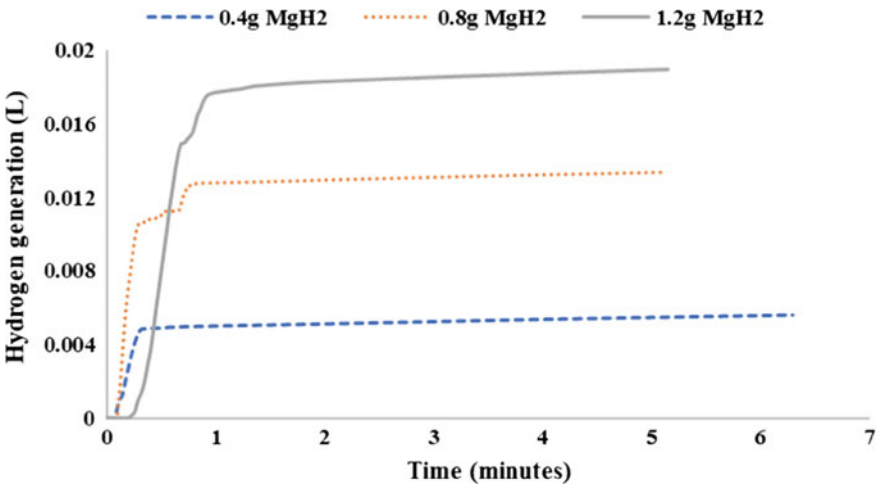


Fig. 5.3 Hydrogen generation at 50 wt% acetic acid concentration

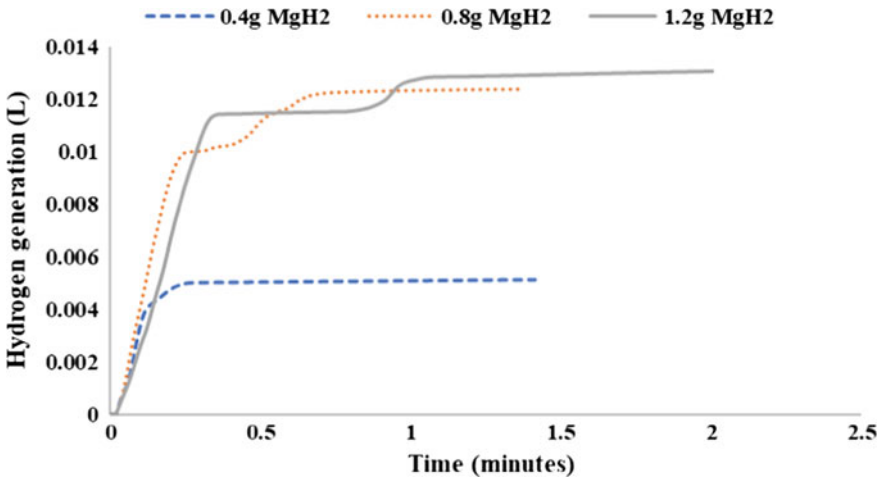


Fig. 5.4 Hydrogen generation at 60 wt% acetic acid concentration

Furthermore, the result of hydrogen generation experiment with 60 wt% acetic acid (Fig. 5.4) indicated the least hydrogen yield of 0.005 L (0.4 g MgH_2 powder), followed by 0.012 L and 0.013 L hydrogen at 0.8 g and 1.2 g MgH_2 powder respectively. Similarly, at 70 wt% acetic acid concentration (Fig. 5.5), hydrogen generation increased with weight of substrate with 0.0048 L, 0.009 L, 0.013 L obtained from 0.4 g, 0.8 g and 1.2 g substrate weight respectively.

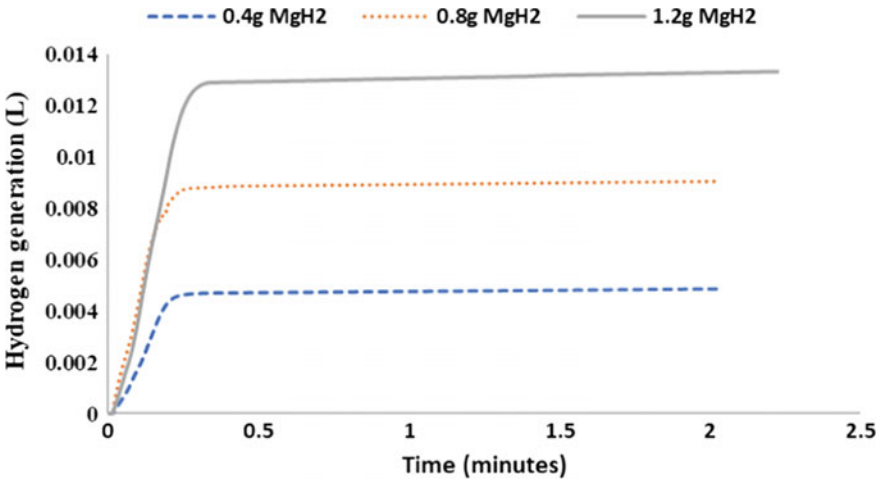


Fig. 5.5 Hydrogen generation at 70 wt% acetic acid concentration

5.3.2 Scanning Electron Micrograph Characterization of Substrate

The SEM micrographs of the MgH_2 powder at different magnifications are indicated in Fig. 5.6a, b. From the micrographs, it can be observed that the particles are of different orientations, some are flake like in nature, while some are rod like and debris particles could also be observed. This also reveals the hydrogen generation sites on the particle of the MgH_2 powder.

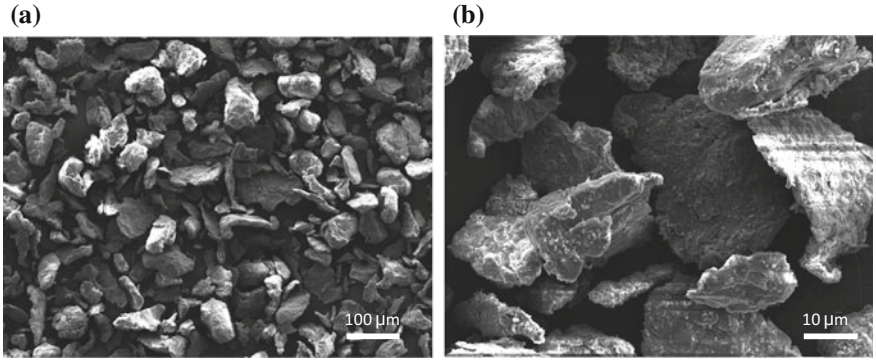


Fig. 5.6 SEM micrograph of the MgH_2 powder as received from the supplier

5.3.3 Substrate Elemental Composition Investigation

Composition/purity of the substrate is important to hydrogen yield. To ascertain the elemental composition of the MgH_2 sample, EDS analysis was conducted. From the EDS result in Table in 5.1 and Fig. 5.7, it can be revealed that three elements were observed in the MgH_2 powder namely magnesium (Mg), oxygen (O), and iridium (Ir). The Mg represent the major constituent as expected in the sample with weight and atomic compositions of 88.88 and 90.35% respectively. This composition is expected because Mg is the major composition of MgH_2. The oxygen in the result could be attributed to oxidation process in the substrate while Ir is obtained from the coating material used in the preparation of the substrate for EDS analysis.

5.3.4 Temperature Optimization Effect on Hydrogen Yield

Temperature is believed to enhance reaction kinetics and thermodynamics in hydrogen storage in light weight metals. However, increasing reaction temperature also involves cost in terms of supply of heat throughout the duration of reaction which

Table 5.1 Elemental composition of MgH_2 powder

Element	Weight (%)	Atomic (%)
0	5.81	8.97
Mg	88.88	9.35
Ir	5.31	0.68
Total	100	100

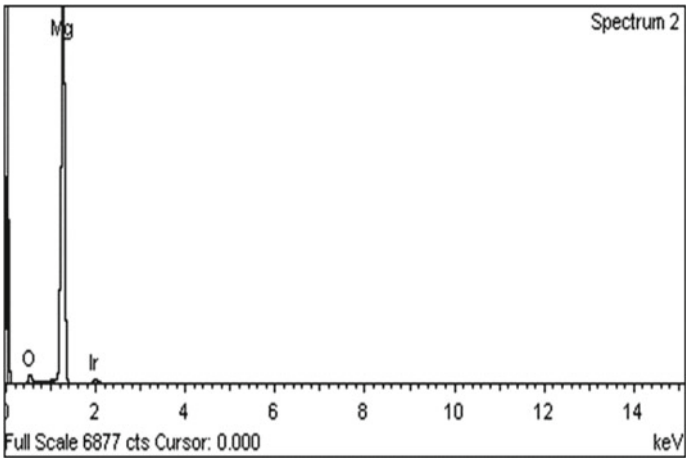


Fig. 5.7 EDS spectrum of MgH_2 powder

could be prohibitive to the sustainable scale of the technique of hydrogen storage in real time. It is important to examine the impact of external temperature supply on hydrogen yield to quantify its role. In this section three temperature parameters were examined in the addition to 50 °C wherein the first phase of experiments was conducted namely 25 °C, 30 °C and 60 °C respectively.

From the temperature optimization experiments (Figs. 5.3, 5.8, 5.9 and 5.10) it was observed that temperature change didn't have a significant impact on hydrogen yield. Interestingly, from Fig. 5.8, hydrogen yield at ambient temperature of 25 °C is comparable to what was obtained at higher temperatures in this study (Figs. 5.3, 5.9 and 5.10), this is a significant discovery as it proves that hydrogen release/desorption can take place at room temperature. Conversely, the test carried out at 25 °C recorded marginal difference in hydrogen generation from 0.8 g (0.139 L) and 1.2 g (0.015 L) MgH_2 respectively. While 0.0058 L hydrogen yield observed from the reaction of 0.4 g substrate was the third highest for all the 0.4 g MgH_2 reaction conducted in this study.

From all the experiments conducted, hydrogen generation increased as a function of substrate concentration. Moreover, the least hydrogen yield was recorded from 0.4 g MgH_2 at 70 wt% acetic acid (Fig. 5.5) with a value of 0.0048 L while the highest hydrogen generation of 0.018 L was recorded at 1.2 g MgH_2 (50 wt% acetic acid concentration). The results also followed the similar pattern to what was obtained in our previous study where MgH_2 pill was utilized as reaction substrate [11]. Thus, laying credence to the important role of substrate concentration to hydrogen yield in MgH_2 based hydrolysis experiment for hydrogen storage.

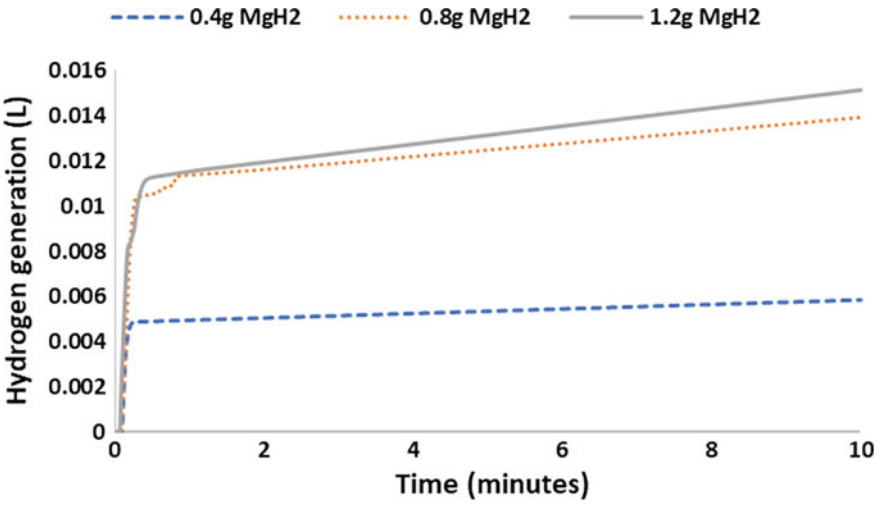


Fig. 5.8 Hydrogen generation at 25 °C, 50 wt% acetic acid concentration

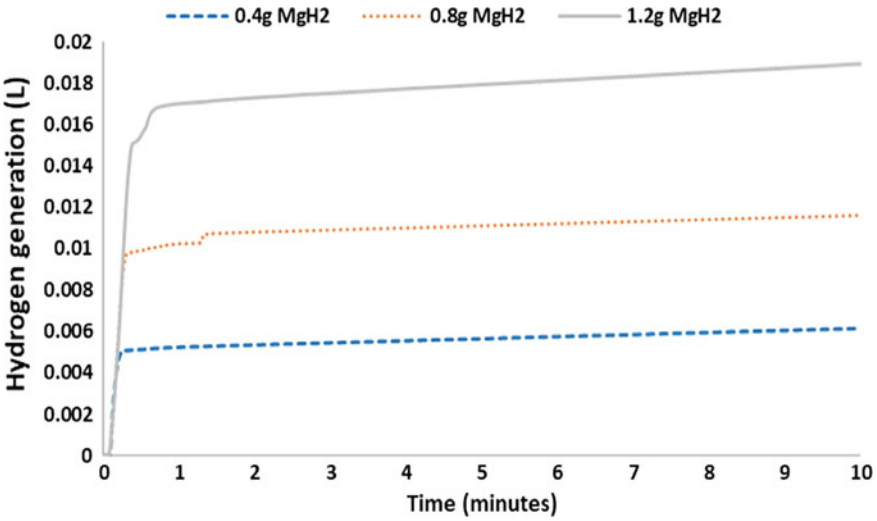


Fig. 5.9 Hydrogen generation at 30 °C, 50 wt% acetic acid concentration

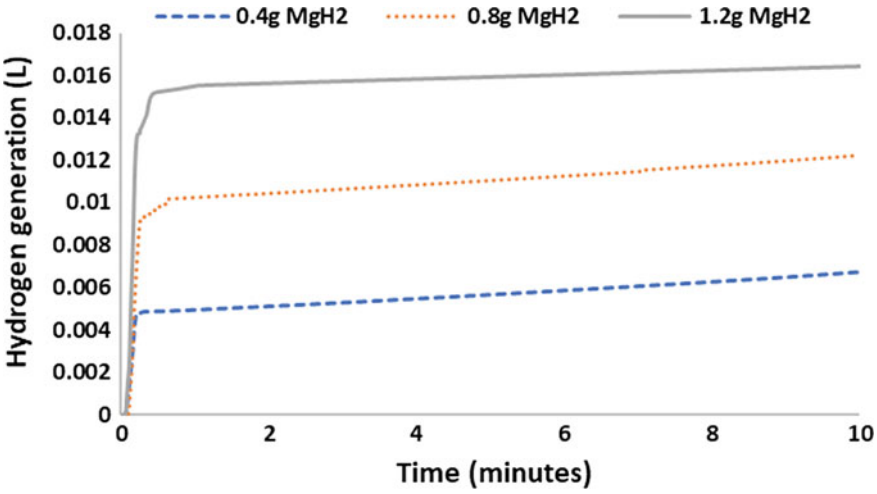


Fig. 5.10 Hydrogen generation at 60 °C, 50 wt% acetic acid concentration

5.4 Conclusion and Future Work

Hydrolysis of MgH_2 powder was catalyzed by an organic acid (acetic acid) using batch mode hydrogen reactor in this study. Acetic acid being a cheap catalyst employed also reduce the reaction cost thereby potentially enhancing the scale up potential of the technology. The observation that appreciable hydrogen can be generated at room temperature means experiment can be carried out without increasing

the reaction temperature thus eliminating the use of energy to supply heat which in turn reduce reaction cost. The cost reduction will further set the study in the path of sustainability drive necessary for the scale up of solid state hydrogen storage for on board vehicular application.

Acknowledgements National research foundation (NRF), South Africa is appreciated for funding the study of the first author. The authors also gratefully acknowledge the financial support from TIA grant of South Africa and University of Johannesburg.

References

1. S. Zhao et al., Carbon nanotubes covalent combined with graphitic carbon nitride for photocatalytic hydrogen peroxide production under visible light. Appl. Catal. B **224**, 725–732 (2018)
2. A. Chinnappan et al., Hydrogen generation from the hydrolysis of sodium borohydride using chemically modified multiwalled carbon nanotubes with pyridinium based ionic liquid and decorated with highly dispersed Mn nanoparticles. J. Power Sources **293**, 429–436 (2015)
3. A. Züttel, Hydrogen storage methods. Naturwissenschaften **91**(4), 157–172 (2004)
4. M.Y. Song et al., Hydrogen-storage properties of MgH_2–10Ni–2$NaAlH_4$–2Ti prepared by reactive mechanical grinding. J. Ind. Eng. Chem. **20**(4), 1591–1595 (2014)
5. Y. Kojima, Platform science and technology for advanced magnesium alloys, in *Materials Science Forum* (Trans Tech Publ, 2000)
6. T.M. Pollock, Weight loss with magnesium alloys. Science **328**(5981), 986–987 (2010)
7. J.A. Adeniran, E.T. Akinlabi, H.S. Chen, R. Fono-Tamo, T.C. Jen, Organic acid-catalyzed hydrolysis of magnesium hydride for generation of hydrogen in a batch system hydrogen reactor, in *Proceedings of the World Congress on Engineering and Computer Science*, 25–27 October, 2017, San Francisco, USA. Lecture Note in Engineering and Computer Science (2017), pp. 615–619
8. J.A. Adeniran et al., Sustainable hydrogen generation substrates, catalysts and methods: an overview, in *International Conference on Mechanical, Industrial and Manufacturing Technologies* (Institute of Electrical and Electronic Engineers, CapeTown, South Africa, 2017), pp. 21–26
9. J.Y. Uan, C.Y. Cho, K.T. Liu, Generation of hydrogen from magnesium alloy scraps catalyzed by platinum-coated titanium net in NaCl aqueous solution. Int. J. Hydrog. Energy **32**(13), 2337–2343 (2007)
10. C.J. Webb, A review of catalyst-enhanced magnesium hydride as a hydrogen storage material. J. Phys. Chem. Solids **84**, 96–106 (2015)
11. T.C. Jen et al., Hydrogen generation from acetic acid catalyzed magnesium hydride using an on-demand hydrogen reactor, in *International Mechanical Engineering Congress and Exposition* (American Society of Mechanical Engineers, Phoenix, Arizona, USA, 2016)

Chapter 6
Investigation into the Electrical Conductivity of Carbon Nanosphere-Based Green Nanofluids

Gloria Adedayo Adewumi, Freddie Inambao, Andrew Eloka-Eboka, Mohsen Sharifpur and Josua Meyer

Abstract Electrical conductivity measurements of green nanofluids prepared from carbon nanospheres dispersed in 60:40 ethylene glycol and water (60:40 EG/W) based nanofluids have been studied. In order to investigate the effect of temperature and volume concentration on the electrical conductivity of the nanofluids, the temperature was varied from 15 to 60 °C and volume fractions of 0.04, 0.1, 0.12, and 0.2 vol% were used. The results show that the electrical conductivity is greatly enhanced with an increase in temperature and volume fraction. The highest enhancement is seen at 0.2 vol% with 1470% increase in electrical conductivity. The high conductivity enhancement indicates a potential for cooling applications.

Keywords Activated carbon · Carbon nanospheres · Coconut fibre · Electrical conductivity · Green nanofluids · Green nanoparticles · Temperature · Volume concentration

G. A. Adewumi (✉)
University of KwaZulu Natal, Room 138B Mechanical Engineering Building, Durban 4041, South Africa
e-mail: adewumigloria@gmail.com

F. Inambao · A. Eloka-Eboka
Department of Mechanical Engineering, University of KwaZulu Natal, Durban 4041, South Africa
e-mail: inambaof@ukzn.ac.za

A. Eloka-Eboka
e-mail: fatherfounder@yahoo.com

M. Sharifpur · J. Meyer
Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria 002, South Africa
e-mail: mohsen.sharifpur@up.ac.za

J. Meyer
e-mail: josua.meyer@up.ac.za

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_6

6.1 Introduction

Synthesis of nanoparticles from "green" bio-precursors has environmental advantages over other conventional methods because of their low-costs and low-toxicity [1, 2]. Nanomaterials synthesized from plants already have their surfaces functionalized; an advantage which is absent in synthetic nanomaterials and this enables them to have better dispersability and stability in base-fluids and subsequently enhanced properties [3]. A detailed review by Buzea et al. [4] outlines the numerous effects of nanoparticle exposure on the human body and environment. These effects can be reduced by investing more into "green" nanotechnology.

A study by Kumar et al. [5] investigated the use of natural sources such as fossil hydrocarbons, waste natural products and botanical hydrocarbon precursors as a means of synthesizing carbon nanomaterials and graphene. It was revealed that those based on fossil hydrocarbon are mostly expensive and not readily accessible. In addition, a lot of the liquid and gaseous fossil hydrocarbons are explosive or toxic in nature and are not acceptable due to atmospheric pollution and its negative effect on human health. On the other hand, carbon nanomaterials derived from carbon based natural precursors have the advantage of producing scalable amounts, they are safe to use in the environment, cheap and allows for fast production techniques.

Thermophysical properties of fluids are those properties that vary with temperature and yet do not have any impact on the chemical structure of the fluid. Just like other thermo-physical properties, electrical conductivity of fluids needs to be enhanced as this can improve the overall working efficiency of their applications relating to cooling, such as in proton electron membranes (PEM) [6]. They can also improve the life span of the electrodes used in metallic cathodes and nozzles [7]. To deal with emission issues, PEM fuel cells are starting to be used in place of batteries as they are light-weight, quiet, have exceptional storage density and possess high fuel energy efficiency [8]. The life span of electrodes is influenced by the electrical conductivity of the liquids flowing through them.

The concept of electrical conductivity arises from the movement of ions is a medium. When an electric field is applied to the medium or nanofluids, the central ion will be attracted and the result is an asymmetric field from a formerly symmetric field [9].

From Stokes law, the hydrodynamic force needed for the movement of a spherical body with radius r and a velocity ν through a fluid with viscosity μ is given as [9]:

$$F_s = 6\pi\mu r\nu \tag{6.1}$$

For microscopic particles with relatively same size as the base fluid, the hydrodynamic drag force on an ion moving under the influence of a chemical potential gradient is [9]:

$$-F_d = 6\pi\mu r_i \upsilon_i \tag{6.2}$$

where r_i and υ_i is the radius and velocity of i ion respectively. The Stokes equation is given as [9]:

$$D_i = \frac{kT}{6\pi\mu r_i} \tag{6.3}$$

where k and T are the thermal conductivities and temperatures respectively. When there is movement of ions in a solution, the local solvent molecules are dragged along. For electric current to be created in a fluid, the positive and negative ions move in opposite directions under the influence of an electromagnetic force (e.m.f). This results in an electrophoretic effect with a countercurrent velocity υ_e [9]. If the number of positive ions is denoted as n_{+ve} and the number of negative ions as n_{-ve} both having drift velocities of v_{+ve} and v_{-ve} respectively with charge q_{+ve} and q_{-ve} respectively, then the current density will be [10]:

$$J = n_{+ve}q_{+ve}v_{+ve} - n_{-ve}q_{-ve}v_{-ve} \tag{6.4}$$

Equation (6.4) can also be written as:

$$J = \frac{nq^2\tau_f E}{m} \tag{6.5}$$

where τ_f, E, and m are the mean free time, electric field and mass respectively. From Ohms law, the electrical conductivity σ is:

$$\sigma = \frac{J}{E} \tag{6.6}$$

Combining Eqs. (6.5) and (6.6),

$$\sigma = \frac{nq^2\tau_f}{m} \tag{6.7}$$

The electrical conductivity in nanofluids can be assumed to be governed by the number of ions present per unit of volume and their respective drift velocities [10].

From past research [7, 11–14], conclusions can be made that the electrical conductivity of nanofluids depend on temperature, volumetric concentration and particle diameter. Increasing the volumetric concentration increases the conducting pathway hence, increasing the electrical conductivity. Similarly, increasing the temperature of the nanofluids will result in a decrease in viscosity thereby enhancing the mobility of ions in the nanofluids. In addition, the numbers of ions are increased due to molecular dissociation, resulting in an enhanced electrical conductivity [11]. The interface effect was considered by [15] as a factor that could enhance the electrical conductivity in carbon nanotubes as the electrical conductivity of complex CNTs in their study decreased rapidly with increasing interfacial layer thickness depending on the percolation threshold.

Coconut fibre (CF) is obtained from the pericarp of coconut fruit and a coconut is made up of 33–35% of husk. Presently, CF are being used as fuel in the processing of coconut-based products, as a fibre source for manufacturing ropes and mats and as a fuel for domestic application [16]. To benefit from this plentiful and low-cost agricultural waste, coconut fibre has been transformed into a carbon nanomaterial in the present study. The conversion of CF into carbon nanomaterials will function as an important raw material obtained from agricultural waste as they are bacteria and fungi resistant. Carbon is a group 14 element that is distributed very widely in nature. It also has a remarkable ability to bond with several other elements as a result of its polyvalent properties creating structures with distinct properties. Recent research has tuned carbon into nanoscale, resulting into carbon nanospheres, carbon nanotubes and carbon nanofibres. Heptagonal and pentagonal pairing of carbon atoms can result in the creation of carbon nanospheres. The graphite sheets in nanospheres occur as waving flakes instead of closed shells which takes the form of the sphere, hence having several open ends at the surface and establishing carbon nanospheres (CNS) as suitable materials for catalytic and adsorption application [17].

Laboratory fabrication of carbon nanospheres involve methods such as chemical vapour deposition (CVD) [18–21], hydrothermal treatment [22], pyrolysis of polymers [23], ultrasonic treatments and chlorination of cobaltocene [24]. Generally, chemical vapour deposition (CVD) occurs when rare earth metal oxides or metal oxides are used as catalysts which results in a need for purification of the synthesized carbon spheres in order to get rid of the catalyst. This makes the process limited to a small scale [25]. Various sources of biomass are being used for the production of nanomaterials derived from carbon [26, 27] because of their low toxicity and availability. These biomasses are often first carbonized, and then activated before being converted to carbon nanomaterials using different methods [28–33]. The method of activation used can either be physical (thermal), chemical or a combination of both physical and chemical processes [16]. The physical process involves the use of CO_2, steam or air and takes place at higher temperatures while the chemical process which is one-step, takes place at lower temperatures and involves the co-carbonization of a parent feedstock with a suitable chemical compound.

To the best of the authors' knowledge, there has been no study carried out on the behavior of electrical conductivity of green nanofluids from coconut fibre based nanospheres. In this study, the authors present the results from the dispersion of already synthesized carbon nanospheres from coconut fibre in 60:40 ethylene glycol/water (EG/W) and finally electrical conductivity measurements. The effect of volume concentration and temperature on the electrical conductivity was also reported.

6.2 Experimental Method

6.2.1 Materials and Stable Nanofluid Preparation

Nanoparticles used in this study were synthesized from coconut fibre. They were first carbonized followed by physical activation in the presence of CO_2 and then treated with ethanol vapour at 800 °C. The procedure is outlined in our previous study of the synthesis procedure in [34]. A known weight of the nanoparticles were measured using a digital weight balance (RADWAG model: AS 220. R2 Max: 220 g Min: 10 mg) and dispersed in 60:40 EG/W (Merck (Pty) Ltd.), gum arabic (GA) acquired from Fluka Analytical and deionized water was used for the nanofluid preparation. The nanotubes were dispersed in a base fluid consisting of 60:40% EG/W in the volume fractions of 0.04, 0.1, 0.12 and 0.2% respectively. The mixture of the nanoparticles, GA and 60:40 ethylene glycol/water (EG/W) was magnetically stirred using a hotplate stirrer (Lasec from Benchmark Scientific Inc., model-H4000-HSE) and sonicated with a 20 kHz, 700 W, QSonica ultrasonic processor. The nanofluid was kept in a programmable temperature bath (LAUDA ECO RE1225 Silver temperature bath) the whole time of sonication and the temperature maintained at 15 °C.

The nanospheres are in the size range of 10–150 nm and the size with the highest occurrence is 60–89 nm. From Fig. 6.1, nanospheres of diameter in the range 60–89 nm constituted almost 50% (48.5%) while nanospheres of diameter 30–59 nm constituted 41%. This shows uniformity in the size range of the nanospheres. The particle size distribution and morphology of the green nanoparticles is given in Fig. 6.1. The morphology indicates smooth round nanospheres with some degree of aggregation due to strong Van der Waals forces. These aggregates can lead to reduced stability and electrical conductivity, hence the need to break the agglomerates.

6.2.2 Measurement of Electrical Conductivity

CON700 conductivity meter was used to measure electrical conductivity (EUTECH instruments). The CON700 conductivity meter is comprised of an electrode with a nominal cell constant of $k = 1.0$, built-in temperature sensor, and 1 m cable. The electrode design offers fast temperature response and reduces air entrapment, ensuring accurate, repeatable, and stable readings. Measurements were taken at temperatures ranging from 15–60 °C at 5 °C intervals. This temperature control was done using a temperature control bath. The conductivity meter was initially calibrated at room temperature with a 1413 μS standard fluid from the supplier.

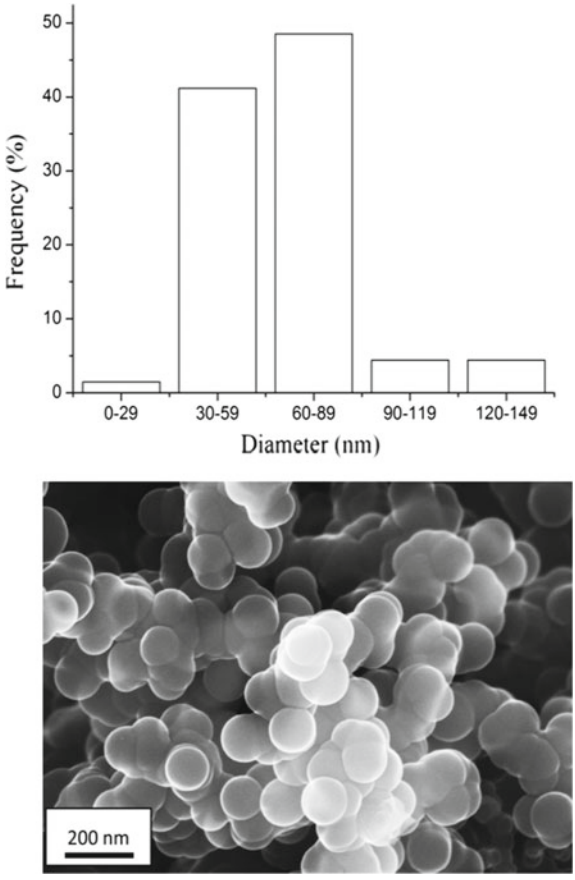


Fig. 6.1 Nanosphere particle size distribution

6.3 Results and Discussion

6.3.1 XRD and Stability of Nanofluid

The graphitization and crystallinity of the synthesized nanospheres is studied through XRD analysis. Figure 6.2 is the XRD results which shows two Bragg diffraction peaks at 30.34° and 50.44°. These peaks can be assigned to typical graphite (003) and (101) planes [35]. The d-spacing calculated is 0.342 nm which is close to graphite 3R given as 0.340 nm. The broadening peaks suggest a low graphitization degree and the possibility of the presence of amorphous carbon. These results fall in the range of values from authors [36] and [37] with d-spacing of 0.33 nm and 0.36 nm respectively. No other peaks are visible in the XRD pattern, which could be due to the high purity of the product.

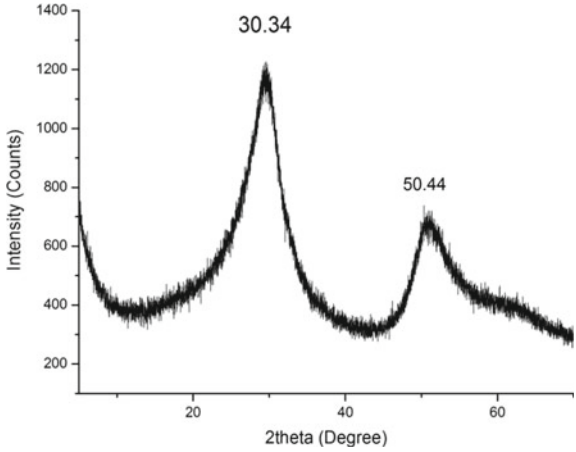


Fig. 6.2 XRD pattern of carbon nanospheres synthesized at 1100 °C

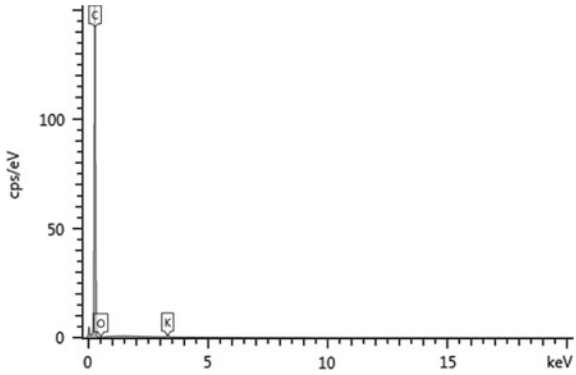


Fig. 6.3 EDX of CNS

Figure 6.3 confirms the purity of the material with a well-defined presence of carbon (98.59%) with a trace of oxygen and potassium.

The stability of the green nanofluids was determined by observing the viscosity at a constant temperature (20 °C) for 720 min. Figure 6.4 is the result of the observation from stability of nanofluids which shows that the viscosity values were hardly changed for the whole duration. This is an indication that the fluids were stable for more than 720 min which is more than sufficient time to take the electrical conductivity measurements. It was therefore determined that a ratio of 1:3.5 CNS/GA gave a good stability and this ratio was maintained throughout this study.

6.3.2 Electrical Conductivity Evaluation

Temperature effects on the electrical conductivity of prepared green nanofluids from coconut fibre nanosphere are shown in Fig. 6.5. The results indicate the dependence

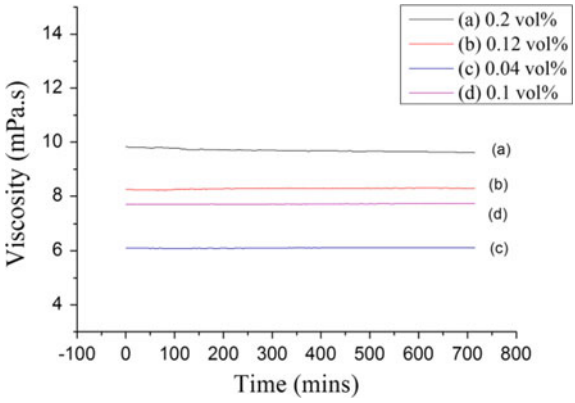


Fig. 6.4 Stability test using viscosity at a constant temperature for 720 min

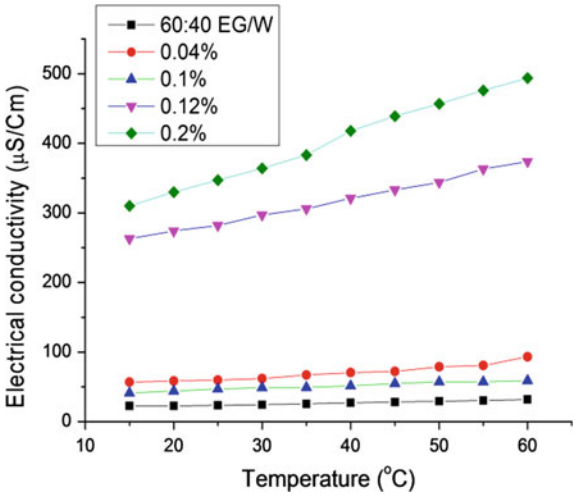


Fig. 6.5 Electrical conductivity at different temperatures for different volume concentrations

of electrical conductivity on temperature. From Fig. 6.6 the effect of the nanosphere volume concentration on the electrical conductivity is reported. The result reveals that an increase in the nanosphere loading impacted greatly on the electrical conductivity as this led to a change in the ionic configuration [38].

At 0.2% volume fraction, the highest electrical conductivity was observed at all temperatures. This is in line with results from recent studies on nanofluids electrical conductivity [38]. With an increase in particle volume concentration, the availability of conducting pathways is increased in the nanofluids consequently leading to a rise in electrical conductivity.

The electrophoretic mobility in nanofluids is also increased when it is stable thereby reducing equivalent particulate masses resulting in an increased electrical conductivity [12]. It is therefore safe to say increasing the temperature of nanofluids

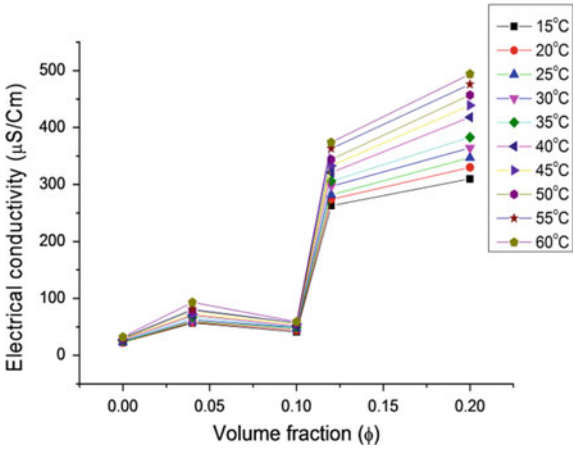


Fig. 6.6 Electric conductivity at various mass fractions

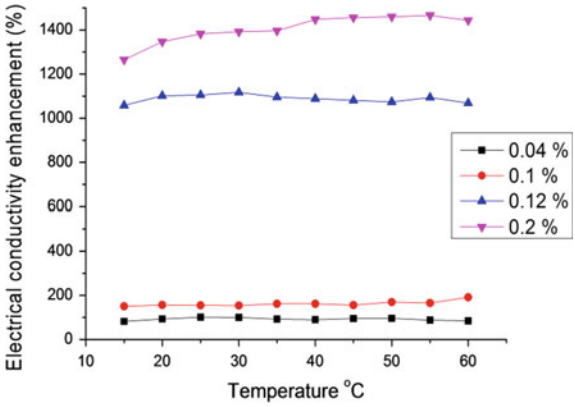


Fig. 6.7 Electrical conductivity enhancements at various temperatures for different volume concentrations

reduces the equivalent particulate masses and increases the electrophoretic mobility. The enhancement in electrical conductivity was calculated based on Eq. 6.3 [13]:

$$\frac{\sigma_{nf} - \sigma_{bf}}{\sigma_{bf}} \times 100 \tag{6.8}$$

where σ_{nf} is the nanofluids electrical conductivity and σ_{bf} is the base fluid electrical conductivity. From Figs. 6.7 and 6.8, the percentage enhancement in electrical conductivity increases with an increase in temperature and volume fraction. At 0.04 vol% an enhancement of about 100% is seen while an enhancement of 1470% is observed at 55 °C for 0.2 vol%.

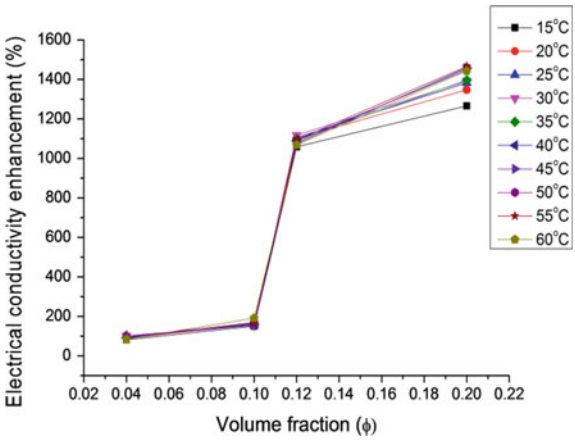


Fig. 6.8 Electrical conductivity enhancement at different volume fractions

6.4 Conclusion

This study presents the results on the electrical conductivity behavior of green nanospheres synthesized from coconut fibre dispersed in 60:40 EG/W and 1:3.5 nanosphere/GA. The measurement of electrical conductivity was varied with volume concentration and temperature. The results obtained show an improved electrical conductivity with increase in temperature and volume concentration. At 0.2 vol%, a maximum electrical conductivity of 1470% was achieved at a temperature of 55 °C. The results presented show the potentials of carbon nanosphere-based green nanofluids as heat transfer fluid for cooling applications. Due to its less toxic nature, it can also be used in applications where human contact is required.

Acknowledgements The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged.

References

1. B. Orhevba, M. Umaru, I.A. Garba, B. Suleiman, M.U. Garba, N. Ernest, Synthesis of composite biomass briquettes as alternative household fuel for domestic application, in *Proceedings of the World Congress on Engineering and Computer Science 2016, WCECS 2016*, 19–21 October 2016, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2016), pp. 696–700
2. J.D. Bayani, J.R.M.D. Pena, E.R. Magdaluyo Jr., Infrared spectra and mechanical properties of corn oil-based polyurethane reinforced with silica nanoparticles from rice husk ash, in *Proceedings of the World Congress on Engineering 2016, WCE 2016*, 29 June–1 July 2016, London, U.K. Lecture Notes in Engineering and Computer Science (2016), pp. 998–1002
3. V. Makarov et al., "Green" nanotechnologies: synthesis of metal nanoparticles using plants. Acta Naturae (англоязычная версия) **6**(1)(20) (2014)

4. C. Buzea, I.I. Pacheco, K. Robbie, Nanomaterials and nanoparticles: sources and toxicity. Biointerphases **2**(4), MR17-MR71 (2007)
5. R. Kumar, R.K. Singh, D.P. Singh, Natural and waste hydrocarbon precursors for the synthesis of carbon based nanomaterials: graphene and CNTs. Renew. Sustain. Energy Rev. **58**, 976–1006 (2016)
6. M.R. Islam, B. Shabani, G. Rosengarten, Electrical and thermal conductivities of 50/50 water-ethylene glycol based TiO_2 nanofluids to be used as coolants in PEM Fuel cells. Energy Procedia **110**, 101–108 (2017)
7. H. Konakanchi, R. Vajjha, D. Misra, D. Das, Electrical conductivity measurements of nanofluids and development of new correlations. J. Nanosci. Nanotechnol. **11**(8), 6788–6795 (2011)
8. J.W. Pratt, L.E. Klebanoff, K. Munoz-Ramos, A.A. Akhil, D.B. Curgus, B.L. Schenkman, Proton exchange membrane fuel cells for electrical power generation on-board commercial airplanes. Appl. Energy **101**, 776–796 (2013)
9. S.I. Smedley, *The Interpretation of Ionic Conductivity in Liquids* (Springer Science & Business Media, 2012)
10. H. Semat, R. Katz, *Physics, Chapter 28: Electrical Conduction in Liquids and Solids* (1958)
11. I. Nurdin, Satriananda, Investigation on electrical conductivity enhancement of water based maghemite (γ-Fe_2O_3) nanofluids. Int. J. Mater. Sci. Appl. J. **6**, 32–36, Art. no. 1 (2017)
12. M. Zawrah, R. Khattab, L. Girgis, H. El Daidamony, R.E.A. Aziz, Stability and electrical conductivity of water-base Al_2O_3 nanofluids for different applications. HBRC J. **12**(3), 227–234 (2016)
13. T.T. Baby, S. Ramaprabhu, Investigation of thermal and electrical conductivity of graphene based nanofluids. J. Appl. Phys. **108**(12), 124308 (2010)
14. M. Kole, T. Dey, Investigation of thermal conductivity, viscosity, and electrical conductivity of graphene based nanofluids. J. Appl. Phys. **113**(8), 084307 (2013)
15. K. Yan, Q. Xue, Q. Zheng, L. Hao, The interface effect of the effective electrical conductivity of carbon nanotube composites. Nanotechnology **18**(25), 255705 (2007)
16. I. Tan, A. Ahmad, B. Hameed, Preparation of activated carbon from coconut husk: optimization study on removal of 2,4,6-trichlorophenol using response surface methodology. J. Hazard. Mater. **153**(1), 709–717 (2008)
17. A. Nieto-Márquez, R. Romero, A. Romero, J.L. Valverde, Carbon nanospheres: synthesis, physicochemical properties and applications. J. Mater. Chem. **21**(6), 1664–1672 (2011)
18. H.-S. Qian, F.-M. Han, B. Zhang, Y.-C. Guo, J. Yue, B.-X. Peng, Non-catalytic CVD preparation of carbon spheres with a specific size. Carbon **42**(4), 761–766 (2004)
19. Y.Z. Jin et al., Large-scale synthesis and characterization of carbon spheres prepared by direct pyrolysis of hydrocarbons. Carbon **43**(9), 1944–1953 (2005)
20. J.-Y. Miao et al., Synthesis and properties of carbon nanospheres grown by CVD using Kaolin supported transition metal catalysts. Carbon **42**(4), 813–822 (2004)
21. M. Ibrahim Mohammed, R. Ismaeel Ibrahim, L.H. Mahmoud, M.A. Zablouk, N. Manweel, A. Mahmoud, Characteristics of carbon nanospheres prepared from locally deoiled asphalt, in *Advances in Materials Science and Engineering*, vol. 2013 (2013)
22. X. Yu et al., Synthesis of activated carbon nanospheres with hierarchical porous structure for high volumetric performance supercapacitors. Electrochim. Acta **182**, 908–916 (2015)
23. Y. Wang, F. Su, C.D. Wood, J.Y. Lee, X.S. Zhao, Preparation and characterization of carbon nanospheres as anode materials in lithium-ion secondary batteries. Ind. Eng. Chem. Res. **47**(7), 2294–2300 (2008)
24. N. Katcho et al., Structure of carbon nanospheres prepared by chlorination of cobaltocene: experiment and modeling. Phys. Rev. B **77**(19), 195402 (2008)
25. P. Zhang, Z.-A. Qiao, S. Dai, Recent advances in carbon nanospheres: synthetic routes and applications. Chem. Commun. **51**(45), 9246–9256 (2015)
26. A.A. Arie, H. Kristianto, M. Halim, J.-K. Lee, Biomass based carbon nanospheres as electrode materials in lithium ion batteries. ECS Trans. **66**(11), 13–19 (2015)
27. H. Kristianto, C.D. Putra, A.A. Arie, M. Halim, J.K. Lee, Synthesis and characterization of carbon nanospheres using cooking palm oil as natural precursors onto activated carbon support. Procedia Chem. **16**, 328–333 (2015)

28. X.-W. Chen, O. Timpe, S.B. Hamid, R. Schlögl, D.S. Su, Direct synthesis of carbon nanofibers on modified biomass-derived activated carbon. Carbon **47**(1), 340–343 (2009)
29. J. Zhu, J. Jia, F.L. Kwong, D.H.L. Ng, S.C. Tjong, Synthesis of multiwalled carbon nanotubes from bamboo charcoal and the roles of minerals on their growth. Biomass Bioenerg. **36**, 12–19 (2012)
30. H. Li, X. He, Y. Liu, H. Yu, Z. Kang, S.-T. Lee, Synthesis of fluorescent carbon nanoparticles directly from active carbon via a one-step ultrasonic treatment. Mater. Res. Bull. **46**(1), 147–151 (2011)
31. J.O. Alves, C. Zhuo, Y.A. Levendis, J.A. Tenório, Catalytic conversion of wastes from the bioethanol production into carbon nanomaterials. Appl. Catal. B **106**(3), 433–444 (2011)
32. K. Shi, J. Yan, E. Lester, T. Wu, Catalyst-free synthesis of multiwalled carbon nanotubes via microwave-induced processing of biomass. Ind. Eng. Chem. Res. **53**(39), 15012–15019 (2014)
33. A. Melati, E. Hidayati, Synthesis and characterization of carbon nanotube from coconut shells activated carbon. J. Phys. Conf. Ser. (IOP Publishing) **694**(1), 012073 (2016)
34. G.A. Adewumi, N. Revaprasadu, A.C. Eloka-Eboka, F. Inambao, C. Gervas, A facile low-cost synthesis of carbon nanosphere from coconut fibre, in *Proceedings of the World Congress on Engineering and Computer Science 2017, WCECS 2017*, 25–27 October 2017, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 577–582
35. P. Debye, P. Scherrer, Interference on inordinate orientated particles in X-ray light. III. Physikalische Zeitschrift **18**, 291–301 (1917)
36. A. Nath, D.D. Purkayastha, M. Sharon, C.R. Bhattacharjee, Catalyst free low temperature synthesis and antioxidant activity of multiwalled carbon nanotubes accessed from ghee, clarified butter of cow' s milk. Mater. Lett. **152**, 36–39 (2015)
37. A.D. Faisal, A.A. Aljubouri, *Synthesis and Production of Carbon Nanospheres Using Noncatalytic CVD Method*
38. S.A. Adio, M. Sharifpur, J.P. Meyer, Investigation into effective viscosity, electrical conductivity, and pH of γ-Al2O3-glycerol nanofluids in Einstein concentration regime. Heat Transf. Eng. **36**(14–15), 1241–1251 (2015)

Chapter 7
Convective Drying of Ginger Rhizomes

Gbasouzor Austin Ikechukwu and Sam Nna Omenyi

Abstract This paper presents the results of convective drying of ginger rhizomes under blanched, unblanched, peeled and unpeeled conditions using the ars-0680 environmental chambers for the drying process and td1002a—linear heat conduction experimental equipment to measure the thermal conductivities of the ginger at six temperature levels ranging from 10 to 60 °C and drying times of 2 and 24 h. The drying curves were drawn using the moisture and conductivity data. The drying rate at higher drying times (24 h) was 0.889/°C and 0.4437/°C for 2 h-drying, giving 50% in moisture reduction rate. Whereas the initial moisture content was 95.12%, it reduced to 59.33% for the 24 h-drying time. The result of this study shows that the lowest moisture content (5.98%) was obtained for unpeeled ginger while the highest was the blanched (9.04%) all for 24 h-drying and at 60 °C. The average moisture content for 2 h drying at 60 °C was 70.6% while for 24 h drying; it was an average of 7.55%. Which is close to the target of 4–7% desired for this research. Though our results made our target, they are in line with the literature results that recommend moisture content of 7–12%. These show the superiority of higher temperature drying and the use of the convective drying method. The thermal conductivity for 24 h-dried ginger at 60 °C approximates to the thermal conductivity of dried ginger and it is 0.050 W/mK on the average. The unpeeled ginger gave the lowest value of 0.046 W/mK while the unblanched ginger gave the highest value of 0.055 W/mK. For 2 h of drying, the average value was 0.079 W/mK while the unblancehd ginger gave the lowest (0.076 W/mK) while the blanched the highest (0.084 W/mK).

Keyword Blanched · Convective drying · Environmental chamber · Ginger rhizomes · Moisture content · Peeled · Temperature drying · Thermal conductivity · Unblanched · Unpeeled

G. A. Ikechukwu (✉)
Department of Mechanical Engineering, Chukwuemeka Odumegwu Ojukwu University, P. M. B. 02, Uli, Nigeria
e-mail: unconditionaldivineventure@yahoo.com; ai.gbsouzor@coou.edu.ng

S. N. Omenyi
Department of Mechanical Engineering, Nnamdi Azikiwe University, P. M. B. 5025, Awka, Anambra State, Nigeria

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_7

7.1 Introduction

7.1.1 Preamble

Convective drying can be employed to remove volatile liquid from porous materials such as food stuffs, ceramic products, clay products, wood and so on. Porous materials have microscopic capillaries and pores which cause a mixture of transfer mechanisms to occur simultaneously when subjected to heating or cooling. The drying of moist porous solids involves simultaneous heat and mass transfer. Moisture is removed by evaporation into an unsaturated gas phase. Drying is essentially important for preservation of agricultural crops for future use. Crops are preserved by removing enough moisture from them to avoid decay and spoilage. For example, the principle of the drying process of ginger rhizomes involves decreasing the water content of the product to a lower level so that micro-organisms cannot decompose and multiply in the product. The drying process unfortunately can cause the enzymes present in ginger rhizomes to be killed.

Ginger is the rhizome of the plant *Zingiber officinale*. It is one of the most important and most widely used spices worldwide, consumed whole as a delicacy and medicine. It lends its name to its genus and family *zingiber aceae*. Other notable members of this plant family are turmeric, cardamom, and galangal. Ginger is distributed in tropical and subtropical Asia, Far East Asia and Africa (Fig. 7.1).

Ginger is not known to occur in the truly wild state. It is believed to have originated from Southeast Asia, but was under cultivation from ancient times in India as well as in China. There is no definite information on the primary center of domestication. Because of the easiness with which ginger rhizomes can be transported long distances, it has spread throughout the tropical and subtropical regions in both hemispheres. Ginger is indeed, the most wildly cultivated spice. India with over 30% of the global share, now leads in the global production of ginger.

Fig. 7.1 Fresh ginger rhizomes

7.1.2 Ginger Rhizomes in Nigeria

In Nigeria, large-scale cultivation of ginger began in 1927 in southern Zaria, especially within Jemima's federated districts as well as in the adjoining parts of the Plateau. Nigeria has tried to widen the genetic base of the crop through introduction of ginger cultivars, mainly from India. Currently, Nigeria is one of the largest producers and exporters of split-dried ginger. Ginger is readily available in the local Nigerian markets and is inexpensive. It is obtained in numerous forms in the market: fresh, dry and powdered ginger rhizomes [14]. Kaduna State is adjudged the largest producer of ginger whereas other states like Nassarawa, Gombe, Benue, Sokoto, Zamfara, Akwa Ibom, Oyo, Abia, Lagos and Bauchi are among the main producers of the farm produce. However Southern Kaduna still remains the largest producer of fresh ginger in Nigeria [4, 9, 10].

7.1.3 Nigeria Ginger Rhizomes Research

Studies conducted in the 1980s were focused on sun-drying and solar drying methods. Several studies conducted on ginger rhizomes were centered on effects of pricking, sun-drying and sieving on Ginger (*Zingiber officinale Roscoe*) colour and powder [13]; composition of volatile oil [6]; Bio-chemical changes in ginger during storage [17]; Development of ginger processing machines [1, 3, 5, 12, 15, 16]; efficiency of ginger production in selected local government areas of Kaduna state, Nigeria [11]; isolation and characterization studies of Ginger root starch as a potential industrial biomaterial [2] etc. In those periods, commercial ginger was exploited. The major difficulties encountered were on pests, diseases and pollutants. Extensive studies were done in the area of post-harvest chemical dips, improved and controlled air storage, spraying of fungicide, hot water treatment, cool storage, etc.

The moisture content of Ginger rhizomes has a major influence on the difficulties encountered in processing ginger rhizomes produced in Nigeria. Other difficulties include vulnerability to fungal rots and quality of dried ginger using open sun drying and/or solar drying. Ginger experiences moisture content loss either vigorously as a segment of the drying process or flaccidly under controlled storage of the farm produce which will not guarantee its' freshness thereafter and therefore will have to be dried to assured moisture content of about 20–35% (Fig. 7.2).

Although several studies have been conducted on drying of ginger rhizomes there are no published work on the convective drying of ginger rhizomes (*Zingiber officinale*) to the knowledge of the authors. This work therefore is centered on the convective drying of ginger rhizomes [8]. The hitherto assumed principal processing of ginger rhizomes involves sorting, washing, soaking, splitting or peeling and drying it to moisture content 7–12%. In this work, the target using the convective drying methodology would be 4–7% from initial moisture content of 87–90% (wb).

Fig. 7.2 Dried split ginger

7.2 Theoretical Aspects

Drying is a very complex process which involves simultaneous heat and mass transfer. Drying is one of the least understood processes at the microscopic level, because of the complexities and deficiencies in mathematical formulations. It is a form of unit operation that converts a liquid, solid or semi-solid feed material into a solid product of very low moisture content. Ginger drying is very complicated because of the differential structure of products. The mechanisms used for drying are surface diffusion or liquid diffusion on the pore surfaces, liquid or vapor diffusion due to moisture concentration differences, and capillary action in granular and porous foods due to surface forces.

Drying processes are categorized into two major models:

Distributed model: This model considers simultaneous heat and mass transfer. It takes into account both the internal and external heat and mass transfers. It predicts the temperature and moisture gradients in the product better. The distributed model depends on the Luikov equations that were derived from Fick's second law of diffusion as shown in Eq. 7.1.

$$\frac{\partial M}{\partial t} = \nabla^2 K_{11} M + \nabla^2 K_{12} T + \nabla^2 K_{13} P$$
$$\frac{\partial T}{\partial t} = \nabla^2 K_{21} M + \nabla^2 K_{22} T + \nabla^2 K_{23} P$$
$$\frac{\partial P}{\partial t} = \nabla^2 K_{31} M + \nabla^2 K_{32} T + \nabla^2 K_{33} P \tag{7.1}$$

where K_{11}, K_{22}, K_{33} are the phenomenological coefficients while $K_{12}, K_{13}, K_{21}, K_{23}$, K_{31}, K_{32} are the coupling coefficients.

In most of the drying processes, the effects of pressure are negligible compared with the temperature and moisture effect. Hence, Luikov equations reduce to:

$$\frac{\partial M}{\partial t} = \nabla^2 K_{11} M + \nabla^2 K_{12} T$$
$$\frac{\partial T}{\partial t} = \nabla^2 K_{21} M + \nabla^2 K_{22} T \tag{7.2}$$

Equation 7.2 is the modified form of Luikov equations and may not be solved using analytical methods due to the complexities of real drying mechanisms. However, the modified form can be solved with the Finite Element Method.

Lumped parameter model: This model does not consider the temperature gradient in the product but assumes a uniform temperature distribution that is equal to the drying air temperature in the product. This assumption reduces the Luikov equation to:

$$\frac{\partial M}{\partial t} = \nabla^2 K_{11} M \tag{7.3}$$
$$\frac{\partial T}{\partial t} = \nabla^2 K_{22} T \tag{7.4}$$

The phenomenological coefficient K_{11} is known as effective moisture diffusivity (D_{eff}) and K_{22} is known as thermal diffusivity (α). For constant values of D_{eff} and α, Eqs. 7.3 and 7.4 can be rearranged as:

$$\frac{\partial M}{\partial t} = D_{eff}\left[\frac{\partial^2 M}{\partial x^2} + \frac{a_1}{x}\frac{\partial M}{\partial x}\right] \tag{7.5}$$
$$\frac{\partial T}{\partial t} = \alpha\left[\frac{\partial^2 T}{\partial x^2} + \frac{a_1}{x}\frac{\partial T}{\partial x}\right] \tag{7.6}$$

where parameter $a_1 = 0$ for planar geometries, $a_1 = 1$ for cylindrical shapes and $a_1 = 2$ for spherical geometries.

Assumptions resembling the uniform temperature distribution and temperature equivalent of the ambient air and product were found to cause errors [7]. Henderson and Pabis reported that this error can be reduced to acceptable values with reduction

in the thickness of the product. This necessitates the derivation of the thin layer drying equations. In this report, the mathematical expressions were not solved but have been presented to alert on the existence of such equations. The work presented here is purely experimental.

7.3 Methodology

7.3.1 Materials

The Ginger rhizomes used in this study were purchased from the popular Eke Awka market in Awka, Anambra State and stored at room temperature before being used for the experimentations. The drying experiments were carried out at the Electronic Manufacturing Engineering Laboratory (ERMERG) Hawkes building, University of Greenwich.

The ginger rhizomes used for the experiment were classified under: (a) Blanched (b) Unblanched (c) Peeled (d) Unpeeled.

7.3.2 Sample Preparation

(a) *Blanched*

- Fill a large pot with water until half full. Put the pot on a stovetop, and turn the burner to high heat. Add several shakes of salt to the water.
- Strip the ginger of its outer peel by running a knife vertically and horizontally.
- Put the ginger into the boiling pot of water. Set the stove timer for 3 min.
- Remove blanched ginger and drop it into ice cold water. This will suddenly put a stop to the cooking process.
- Wait for another 3 min for the ginger to complete the blanching process. Remove the ginger and place on a paper towel linen plate to dry.

(b) *Unblanched*

- Freshly unwashed ginger with water

(c) *Peeled*

- Wash the ginger with water and peel
- Hold a piece of ginger and scrap the edge of ginger with a spoon to peel off the skin

(d) *Unpeeled*

- Wash ginger with water and then keep unpeeled

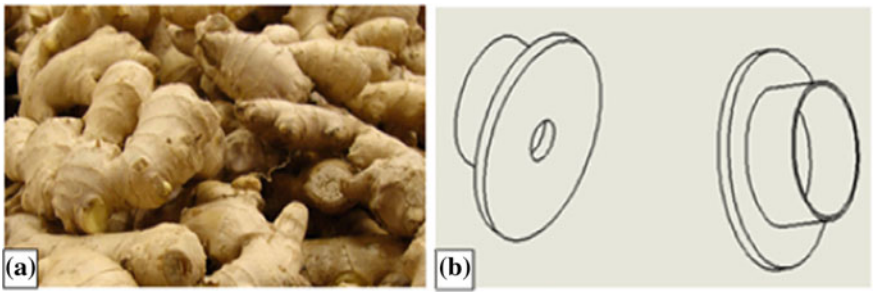


Fig. 7.3 **a** Raw materials for the experiments (Ginger Rhizomes) **b** device designed for the chopping of ginger rhizomes to the required sizes (18 × 30 mm diameter) for the drying and heat conduction experiments

Raw ginger and the device used to chop the ginger rhizomes to size are shown in Fig. 7.3.

7.4 Methods

ESPEC's ARS-0680 Environmental Humidity and Temperature Chamber as shown in Fig. 7.4 was used for heating the specimen at low or high temperature with controlled humidity. The ESPEC's ARS-0680 Environmental Humidity and Temperature chamber has the following features:

- Internal dimension of W850 × H1000 × D800 and an External dimension of W1050 × H1955 × D1805
- Operating temperature ranging from −73 °C to + 180 °C(−103 °F to + 356 °F)
- Temperature function of 0.3 K
- Temperature deviation in space of ±1.5 K
- Temperature gradient of 3.0 K
- Rate of temperature change 6.0 K/min or more while heating and
- Rate of temperature of change 4.2/min or more while cooling.

ESPEC's Environmental Stress Chambers can withstand heat loads produced by the specimen, improve temperature change rates, and provide expanded ranges for temperature and humidity. Each chamber is also equipped with a specimen temperature control function to meet stringent testing demands typically required for automotive parts and mobile products.

The ginger used was cut into slices of 30 mm diameter and 18 mm thickness by scoopers designed for this purpose and prepared as Blanched, Unblanched, Peeled and Unpeeled as previously described at temperatures of 10–60 °C for drying times of 2 and 24 h and the Linear Heat Conduction Experiment was used to measure the thermal conductivity of the sample.

Fig. 7.4 ARS-0680 temperature and humidity chamber

The temperature and humidity chamber installed at the Hawke building, University of Greenwich was used for the drying of the ginger rhizomes at a minimum temperature of 10 °C; maximum temperature of 60 °C and resident time of 10 min starting at a room temperature (RT) of 24 °C in the environmental chamber. A total of 16 samples were placed in the environmental chamber which was programmed to run for 2 or 24 h initially. However, at the end of every cycle, a sample would be retrieved from the environmental chamber for analysis and measurement to evaluate the percentage moisture content and its thermal conductivity using the TD1002A—Linear Heat Conduction Experiment Unit shown in Fig. 7.5a. Humidity test was totally ignored in this research, as it was not one of the objectives to meet in this study.

The equipment for Linear Heat Conduction Experiment shown in Fig. 7.5—TD1002A—has a wooden bar of circular cross-section made up of two sections with an interchangeable middle section. It is mounted on a base platen with a clear schematic of the experiment layout. The first brass section includes three thermocouples and the electric heater (heat source). The second brass section includes a small water-cooled chamber (heat sink) and three more thermocouples. The interchangeable middle section was modified with wood for this experiment to prevent heat loss during the experiment. Each middle section has a thermocouple. The electric heater and thermocouples connect to sockets on the Heat Transfer experiments base unit, which also supplies the cold water feed and drain for the heat sink. The cooling water flow and the heater power were turned on until the materials attained temperature equilibrium; the temperatures were then recorded along the

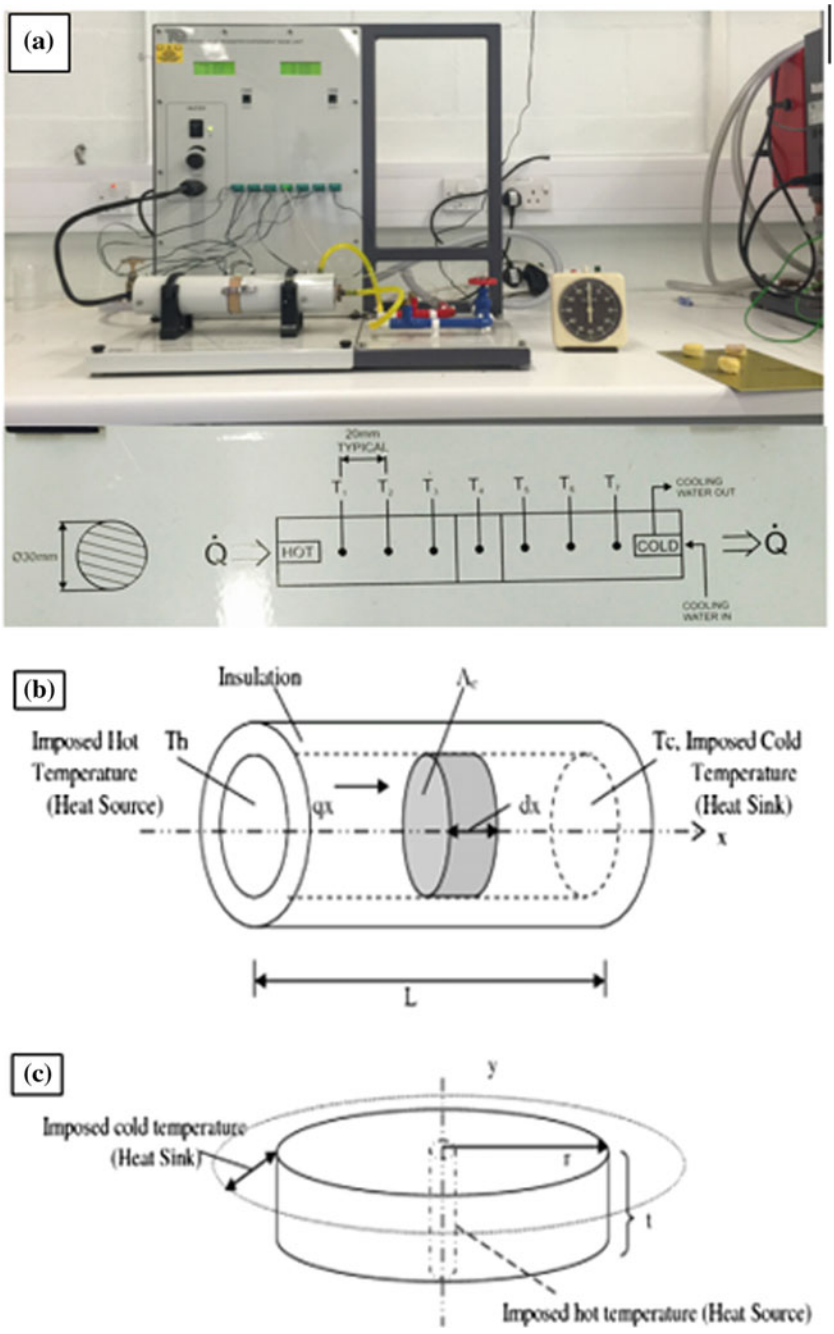


Fig. 7.5 **a** TD1002A—linear heat conduction experiment unit (LHTEU) with TD1002 heat transfer experiments base unit **b**, **c** diagram for heat conduction along a well-insulated cylindrical rod

bar. Insulation around the bar reduced heat loss by convection and radiation. A power of 20 W was used throughout the experiment. The diameter of each ginger rhizome chopped out from the gingers supplied were 30 mm in diameter and 18 mm in thickness.

Procedure the Linear Heat Conduction experiment

- The room temperature was initially measured.
- The clip located in the middle of the insulated wooden rod (cylinder) was opened and the ginger rhizomes inserted.
- Thermal paste was applied between the adjacent faces of the wooden material to reduce temperature gradient across the joints. The ginger rhizome was then inserted into the central section at the middle of the cylinder.
- The main water supply was opened. The red valve was completely opened for water inlet to the cylinder.
- The power supply and the control board were switched on.
- The heater was switched on, and set to 20 W.
- The initial temperatures of the thermocouples $\boldsymbol{T}_1, \boldsymbol{T}_2, \boldsymbol{T}_3, \boldsymbol{T}_4, \boldsymbol{T}_5, \boldsymbol{T}_6,$ and $\boldsymbol{T}_7$ were measured.
- The temperature of the thermocouples at different time range were subsequently also measured.
- After the temperature measurements, the heater control was turned to zero; the heater, the control board and the main power supply were switched off. After waiting for approximately 5 min for the temperature of the water in the cylinder to cool down, the red valve and main water supply were turned off.

Procedure adopted during Drying

The Ginger Drying experiment was conducted according to ASAE Standard S352.2 (*Convection Oven*). Before the experiment started, the whole apparatus was operated for at least 15–30 min to stabilize the humidity, air temperature and velocity in the dryer. Drying was started at 08:00 am and continued until the specimen reached the final moisture content at time set for the experimental. The weight losses of the sample in the environmental chamber were recorded during the drying period of 2 and 24 h with electronic balance (EK-200 g, Max 200 ± 0.01 g). At the end of drying, the dried sample was collected for the measurement of its' thermal conductivity using the linear heat conduction equipment.

Determination of Moisture Contents

- The initial mass of the ginger sample was recorded using 0.00001 g "Analytical Plus Electronic Balance".
- The ginger was placed into an environmental chamber at constant temperature of 10–60 °C for a time period of 2 and 24 h.
- Then the mass of the dried ginger sample was recorded for the time periods 2 and 24 h.

Fig. 7.6 Powdered ginger rhizomes with thermal conductivity of 0.0503 W/m K

- Mass of the ginger sample was examined regularly till it reached an equilibrium value (Final mass).
- Moisture content of the ginger was computed.

Determination of Thermal Conductivity

Thermal conductivity of a material is defined as the capacity of a body or any material to transmit or conduct heat. Thermal conductivity is dependent on the following factors:

- Material structure
- Moisture content
- Density of material
- Pressure and temperature (operating conditions)

Figure 7.6 shows dried and ground (powdered) ginger rhizomes. The thermal conductivity is mathematically expressed as:

$$k = \frac{\mathrm{QD}}{\Delta \mathrm{TL}} \tag{7.7}$$

where

A Cross-sectional area
k thermal conductivity (K)
Q Amount of heat transfered through the material ($\frac{J}{}$) (W)
ΔT change in temperature (K)
L distance between $\mathrm{T}_4 - \mathrm{T}_3$ (m)

7.5 Results and Discussions

This study investigated two important features of convective drying of ginger rhizomes:

1. moisture content characteristics
2. thermal conductivity of each sample at varying drying times and temperatures using the linear heat conductions experimental unit.

The data on moisture contents of ginger rhizomes dried for 2 h and for 24 h respectively were plotted in Figs. 7.7 and 7.8 as a function of temperature. The best fit to the data was found to be a straight line. These figures represent the drying curves in terms of the moisture content. The reduction of moisture with increase in temperature is evidence of drying. The drying rate is given in moisture reduction per degree rise in temperature. The characteristics of these curves are given in Table 7.1.

Table 7.1 shows as expected that the ginger rhizomes dried for a longer time have higher average reduction in moisture given by the slopes of the graphs as 0.889/°C for 24 h drying and 0.4437/°C for 2 h drying, giving 50.1% in moisture reduction rate. The intercept which theoretically gives the initial moisture content at 0 °C is

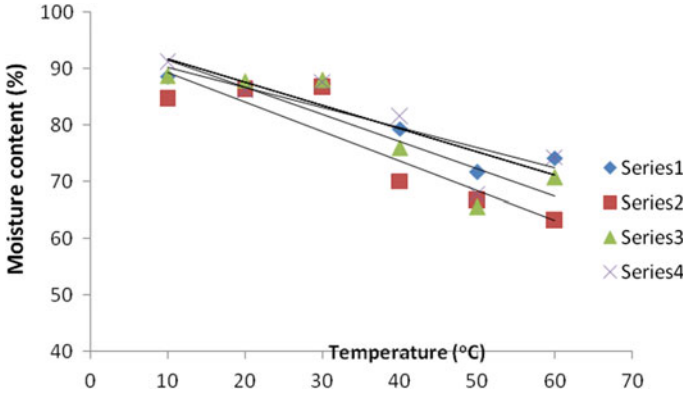


Fig. 7.7 Moisture content of rhizomes dried for 2 h, plotted as a function of temperature

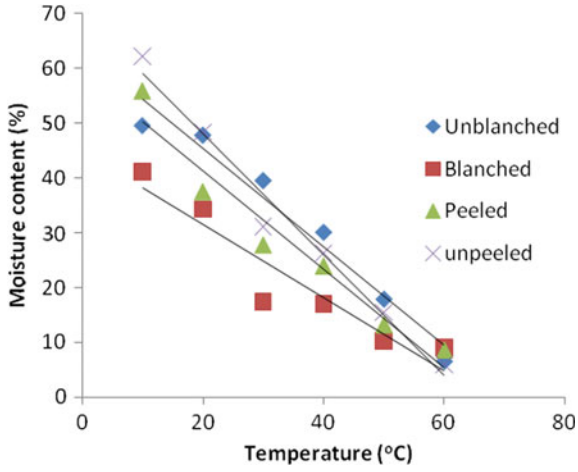


Fig. 7.8 Moisture content of rhizomes dried for 24 h, plotted as a function of temperature

Table 7.1 Data for moisture content

Drying time		Slope	Intercept	R-squared
2 h	Unbleached	0.3558	93.70	0.8395
	Blanched	0.5224	94.51	0.8219
	Peeled	0.4829	96.36	0.7974
	Unpeeled	0.4137	95.92	0.7693
Average		−0.4437	95.12	0.807
24 h	Unbleached	−0.896	63.30	0.9627
	Blanched	−0.6656	44.82	0.8949
	Peeled	−0.8955	59.15	0.9469
	Unpeeled	−1.099	70.05	0.9774
Average		−0.8890	59.33	0.9455

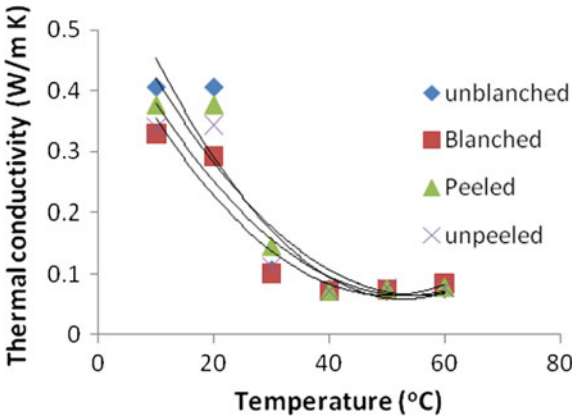


Fig. 7.9 Effects of temperature on the thermal conductivities of ginger rhizomes dried for 2 h

lower at 24 h drying (59.33%) compared to 95.12% on dry basis at 2 h of drying, as expected. The goodness-of-fit, on the average is higher for ginger dried for 24 h than for that dried for 2 h. The average moisture content for 2 h drying at 60 °C is 70.6% while for 24 h drying, it is an average of 7.55%, which is close to the target of 4–7% desired for this research. This is better than the result of 22.54% obtained at 50 °C under blanched condition drying for 32 h. Eze and Agbo reported that the principal processing of ginger rhizomes involves sorting, washing, soaking, splitting or peeling and drying to moisture content 7–12%. Our results are consistent with this range. These show the superiority of higher temperature drying and the use of the convective drying method. The results also show that the lowest moisture content (5.98%) was obtained for unpeeled ginger while the highest was the blanched (9.04%) all for 24 h-drying and at 60 °C.

The results for the thermal conductivities are presented in Figs. 7.9 and 7.10. The curves were fitted to polynomial functions of order two and the resulting equations are given on Table 7.2.

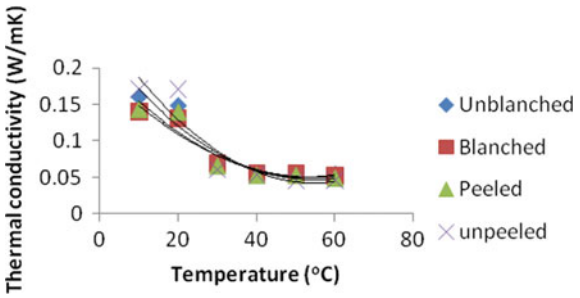


Fig. 7.10 Effects of temperature on the thermal conductivities of ginger rhizomes dried for 24 h

Table 7.2 (a) For products dried for 2 h (Fig. 7.9). (b) For products dried for 24 h (Fig. 7.10)

(a)		
Unblanched	$K = 0.0002T^2 - 0.0227T + 0.6588$	$R^2 = 0.6588$
Blanched	$K = 0.0002T^2 - 0.0181T + 0.5184$	$R^2 = 0.9106$
Peeled	$K = 0.0002T^2 - 0.019T + 0.5937$	$R^2 = 0.8804$
Unpeeled	$K = 0.0002T^2 - 0.0178T + 0.5409$	$R^2 = 0.8551$
(b)		
Unblanched	$K = 7 \times 10^{-5}T^2 - 0.0071T + 0.2367$	$R^2 = 0.9067$
Blanched	$K = 5 \times 10^{-5}T^2 - 0.0054T + 0.1974$	$R^2 = 0.9139$
Peeled	$K = 5 \times 10^{-5}T^2 - 0.0058T + 0.2074$	$R^2 = 0.8875$
Unpeeled	$K = 7 \times 10^{-5}T^2 - 0.008T + 0.2602$	$R^2 = 0.8614$

Figures 7.9 and 7.10 have similar shapes and can be seen as drying curves. The thermal conductivities were high at low drying times as was the case with moisture contents and decreased to almost asymptotic values at higher drying times at 60 °C. The intercepts which gave the expected conductivities at zero degree centigrade were higher for 2 h-dried ginger averaging to 0.578 W/mK than at 24 h of drying, which averaged to 0.225 W/mK, by a factor of 61.1%. The thermal conductivity for 24 h-dried ginger at 60 °C approximates to the thermal conductivity of dried ginger and it is 0.050 W/mK on the average, with unpeeled ginger giving the lowest value of 0.046 W/mK and unblanched ginger giving the highest value of 0.055 W/mK. For 2 h of drying the average value was 0.079 W/mK while the unblanched ginger gave

the lowest and blanched the highest. Previous studies concluded that peeled and blanched ginger allow a decrease in the resistance of this product to water transportation within the internal and external part because the outer skin of the rhizomes as observed from the unblanched and unpeeled provides slight resistance due to its' non-permeability which causes rigidity during the drying process therefore disallowing water easy transportation through it.

7.6 Conclusions

The following conclusions were drawn from this study:

- The results obtained for moisture content of ginger rhizomes clearly indicate that drying at significantly short time (say 2 h) will not reduce the moisture sufficiently to reduce the effects of pest and bacterial infections.
- The drying rate at higher drying times (24 h) was 0.889/°C and 0.4437/°C for 2 h drying, giving 50% in moisture reduction rate. The intercept which theoretically gives the initial moisture content at 0 °C is lower at 24 h drying (59.33%) compared to 95.12% on dry basis at 2 h of drying, as expected.
- The result of this study shows that the lowest moisture content (5.98%) is obtained for unpeeled ginger while the highest is the blanched (9.04%) all for 24 h-drying and at 60 °C.
- The average moisture content for 2 drying at 60 °C was 70.6% while for 24 h drying, it was an average of 7.55%. which is close to the target of 4–7% desired for this research. This is better than the result of 22.54% obtained at 50 °C under blanched condition drying for 32 h. Eze and Agbo reported that the principal processing of ginger rhizomes involved sorting, washing, soaking, splitting or peeling and drying to moisture content 7–12%.
- The significance of drying ginger for a longer time at even lower temperatures around 60 °C has been shown in this work. At higher temperatures ginger shrinkages and surface decolouration may occur. As can be seen, good results are achievable at temperature of 60 °C to sustain the quality of the products.
- The thermal conductivity for 24 h-dried ginger at 60 °C approximates to the thermal conductivity of dried ginger and it is 0.050 W/mK on the average, with unpeeled ginger giving the lowest value of 0.046 W/mK and unblanched ginger giving the highest value of 0.055 W/mK.
- The results show the superiority of the use of convective drying method over all other methods.

References

1. S. Adeyemi, L. Onu, *Development of Ginger Processing Machines-Ginger Peeler*. Annual Research Report (1997)
2. M.O. Afolayan, K. Adama, A. Oberafo, M. Omojola, S. Thomas, Isolation and characterization studies of ginger (Zingiber officinale) root starch as a potential industrial biomaterial. Am. J. Mater. Sci. **4**(2), 97–102 (2014), http://article.sapub.org/10.5923.j.materials.20140402.06.html. Accessed 12 March 2015
3. G. Akomas, E. Oti, Developing a technology for the processing of Nigerian ginger (Zingiber officinale Roscoe), in *Proceedings of the First National Ginger Workshop* (1988)
4. A. Bernard, Diseases, pest and other factors limiting ginger (Zingiber officinale Rose) production in River State (2008), http://uptonvilleoginstu.org/ginger.litm. Accessed 12 March 2015
5. C. Egbuchua, E. Enujeke, August). Growth and yield responses of ginger (Zingiber officinale) to three sources of organic manures in a typical rainforest zone, Nigeria. J. Horticult. For. **5**(7), 109–114 (2013)
6. O. Ekundayo, I. Laakso, R. Hiltumen, Composition of ginger (Zingiber officinale Roscoe) volatile oils from Nigeria. Flavour Fragr. J. **3**(2), 85–90 (2006). Accessed 19 Feb 2015
7. Z. Erbay, F. Icier, A review of thin layer drying of foods: theory, modeling, and experimental results. Crit. Rev. Food Sci. Nutr. **50**(5), 441–464 (2010). https://doi.org/10.1080/10408390802437-063
8. A.I. Gbasouzor, S.N. Omenyi, Studies of convective drying of ginger rhizomes, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco USA. Lecture Nate in Engineering and Computer Science (2017), pp. 957–964
9. KADP, Annual Report. Ministry of Agricultural Development, Kaduna State Agricultural Development. Kaduna State Agricultural Development Project, Kaduna, Nigeria (2004). Accessed 12 March 2015
10. KADP, *Production of Ginger: An Extension Guide*. Kaduna State Agriculture Development Project (2000). Accessed 12 March 2015
11. J. NdaNmadu, Efficiency of ginger production in selected local government areas of Kaduna State, Nigeria. Int. J. Food Agric. Econ. **1**(2), 39–52 (2014)
12. G. Nwandikom, B. Njoku, Design-related physical properties of Nigerian ginger, in *Proceedings of First National Ginger Workshop* (1988), pp. 101–107
13. G. Okafor, J. Okafor, Effects of pricking, sun-drying and sieving on ginger (Zingiber officinale Roscoe) colour and powder. Niger. Food J. **25**(1), 155–160 (2007)
14. B. Omeni, About Us: Agronigeria Ltd. (Agronigeria, Producer, & Agronigeria Ltd. (2015), From An Agronigeria Ltd. Website: http://agronigeria.com.ng/2014/01/09/steps-to-take-in-ginger-plantation/. Accessed 7 March 2015
15. L. Onu, Design and fabrication of manual ginger slicing machine. Annu. Res. Rep. **34**(1) (1997)
16. L.I. Onu, G.I. Okafor, Effect of physical and chemical factor variations on the efficiency of mechanical slicing of Nigerian ginger (Zingiber officinale Roscoe). J. Food Eng. **56**(1), 43–47 (2003). https://doi.org/10.1016/S0260-8774(02)00146-2
17. E. Oti, P. Okwuowulu, V. Ohiori, G. Chijioke, Biochemical changes in ginger (Zinginger officinale Roscoe) rhizomes stored under river sand and under dry grasses in pits in the humid tropics. Trop. Sci. **28**, 87–94 (1988)

Chapter 8
A Multi-agent Simulation Study for Congestion Minigation in a University Campus Restaurant

Takeshi Koide, Takeru Kobayashi and Maki Kikuda

Abstract Campus restaurants in many universities in Japan are crowded during lunch time. Managers of the restaurants usually consider to introduce some methods for the mitigation of the congestion and it is difficult to execute the introductions instantly due to the cost of the adoption. Numerical analysis on the effect of the methods will help the managers to determine if the methods should be adopted. In this study, a multi-agent simulation model is proposed to estimate the congestion at lunch break in a campus restaurant in our university. The model represents a series of customer's behaviors in the order and pick-up area in the restaurant and interactions among customers. Simulation study estimates the effect of introductions of new layouts on the congestion and the influence of the customers who make an order slowly at an order window.

Keywords Congestion mitigation · Multi-agent simulation · On-site survey
Practical application · Statistical analysis · University campus restaurant

8.1 Introduction

Our university established a new recreation facility building last year. A main cafeteria-style campus restaurant, operated by the cooperative at our university, is popular among university members and usually very crowded during lunch break. Students have to wait for their dishes in a queue to a pick-up window for around 5 min on average. Some members give up having a lunch in the restaurant in disgust of the

T. Koide (✉) · T. Kobayashi · M. Kikuda
Department of Intellegence and Informatics, Konan University, Okamoto 8-9-1, Higashinada-ku, Kobe 658-8501, Japan
e-mail: koide@konan-u.ac.jp

T. Kobayashi
e-mail: s1471045@s.konan-u.ac.jp

M. Kikuda
e-mail: s1371034@s.konan-u.ac.jp

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_8

waiting and the university cooperative loses sales as well as the trust of cooperative members. Many campus restaurants in Japan are in much the same circumstances. The congestion problem is to be tackled for university cooperatives not only from the managerial point of view but also for member's benefit.

Multi-agent simulation method is a powerful approach to model complex phenomena adopted in various study fields such as sociology, physics, and chemistry [1–3]. The simulation approach especially provides critical insights to understand emergent features produced by interactions among numerous agents. The multi-agent simulation models have been also applied in service industries, such as retailing [4] transportations [5], and restaurant industries [6] in order to estimate the performance of operations to improve quality of services provided.

A university campus restaurant was targeted in a simulation study in [7] to reduce customer wait time. The study is based on the simulation models for fast food restaurants [8–10]. Kitakami et al. [11] constructed a multi-agent simulation model for behaviors of customers in their university campus restaurant. Using the model, they observed the transition of usage situation of restaurant seats in every area of the restaurant. They analyzed the current situation of the target restaurant but did not propose any methods to mitigate the congestion. Koide and Kikuda [12] estimated the efficiency of a priority seat system, named speed seats, for the mitigation of congestion in a university campus restaurant by using a multi-agent simulation model.

This study extends the simulation model in [12] to apply the new cafeteria-style campus restaurant in our university in order to explore some solutions for congestion mitigation. The proposed model focuses on the area to order and pick-up inside the cafeteria restaurant and customer's behaviors are modeled more detailed than those in [12] while [12] targeted the whole area of another restaurant. In the proposed model, customer agents come to the order and pick-up area, select their menus, wait in a queue to an order window, order and pick a dish up, check out at a cash register, and exit from the area. Simulation experiments measure customer's dwell time in the area as an indicator of congestion in some different settings in order to estimate the effectiveness of layout changes and the influence of the order delays by some customers.

The rest of this paper is divided into four sections. In Sect. 8.2, the proposed multi-agent simulation model for the target restaurant area is shown. Section 8.3 explains the details of customer's behaviors in the simulation model. Section 8.4 shows the simulation experiments for the congestion and their results. Finally, Sect. 8.5 concludes the paper and discusses future works.

8.2 Multi-agent Simulation Model

This study targets a cafeteria-style campus restaurant in our university. The restaurant locates on the ground floor of a four-story building constructed in 2017 and its seating capacity is 1400 seats. The floor plan of the order and pick-up area of the restaurant is illustrated in Fig. 8.1. Customers come into the area from the entrances denoted

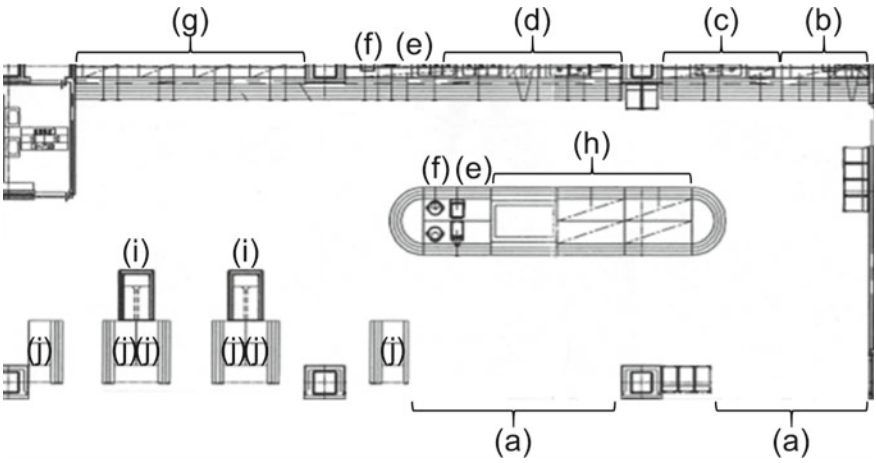


Fig. 8.1 Floor plan of the order and pick-up area of the cafeteria restaurant

by (a). The order counter has three main windows (b), (c), and (d) for noodles, rice bowls and curries (in short RB&Cs), and main dishes, respectively. Rice and miso soup are respectively served at (e) and (f) points. More than ten kinds of side dishes are lined up in the shelves denoted by (g). The island (h), called self-deli, is a sort of heating cabinet and supplies pre-cooked dishes and boxed lunches. Some dessert items are sold at the (i) shelves. There are totally six cash registers located at the positions denoted by (j). This study aims to assess the congestion in the order and pick-up area excluding the seating area.

The constructed multi-agent simulation model in this study adopts a continuous space model whose simulation space is shown in Fig. 8.2. The simulation space is constructed based on the floor plan given in Fig. 8.1. The components (b) through (j) are allocated at the same positions as in Fig. 8.1. The simulation space is composed of grid square cells with a side ten centimeter and the size of objects in the space, such as walls, posts, self-deli, cash registers, and dessert shelves, are adjusted into the unit cell size. Some way points are adopted in the space so that customer agents steer around obstacles. The customer agents are represented by circles filled with corresponding colors to their current destinations.

The target time period in this study is set to 12:10–12:30, the most crowded time period. The arrival rate of customer agents is computed for three main windows as well as self-deli based on the practical past results of their supply quantities. This model assumes that all customers arrive individually for the sake of shorthand nevertheless most customers actually come in a group. The elapsed time for restaurant staffs to serve a menu since a customer orders the menu was set to the average elapsed time measured by on-site surveys.

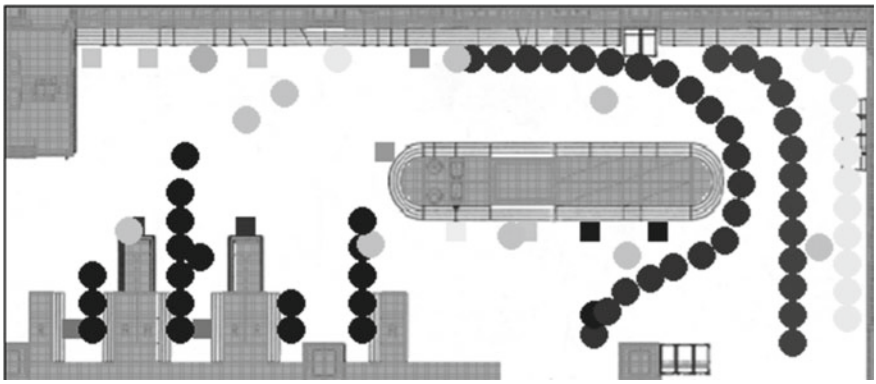


Fig. 8.2 Simulation space of the constructed simulation model

8.3 Behaviors of Customer Agents

The proposed model represents the customers in the target order and pick-up area as customer agents which have the following four-stage behaviors: (1) arrival (2) move (3) queuing and order (4) payment. Some customer agents repeat steps (2) and (3) to select more than one menus. This section mentions the behaviors in detail.

8.3.1 Arrival

A customer agent is randomly created in every 1.6 s. Customer agents determine a destination for their menus among noodles window, RB&Cs window, main dishes window, and self-deli island with a probability of 22%, 23%, 43%, 12%, respectively. The selection probability is determined based on the practical supplied amount of menus served at each destination. The three windows have each queue and the customers who select one of the three windows set his/her destination to the position of the end of the queue. The customers who select self-deli island proceed there directly since self-deli does not have a queue. The customer agents arrive at the simulation space from the entrance closer to their destination.

8.3.2 Move

After arriving at the entrance, a customer agent walks toward its destination. While a customer agent walks to the end of a queue, another customer agent is sometimes added to the end of the queue and the position of the end of the queue varies. In order

to deal with the situation, customer agents periodically confirm the position of their destination and update it.

The walking scheme is designed to simplify that in [13]. A customer agent basically travels at 40 cm/s and slows down if there are some other customers in the circular area 40 cm ahead of the customer with a radius of 35 cm. When the number of customers in the area is from one to three, from four to six, from seven to eight, and more than eight, the travelling speed is declined by 60%, 70%, 80%, and 90%, respectively. When there is an obstacle in the direction of forward movement, the traveling direction is changed to either left or right direction without obstacles. The simple avoidance scheme sometimes occurs unrealistic quirky movements. In order to overcome the problem, some way points are installed in the simulation space and customer agents travel through the way points to avoid obstacles.

8.3.3 Queuing and Order

When a customer agent arrives at the end of its target queue, the agent starts to wait in the queue. The outlines of the queues to each window are designated in the simulation model and the customer agents in a queue make a line along the corresponding outline.

The head consumer agent in a queue orders a menu at a window and waits for the menu to be served. The elapsed time to serve a menu is determined based on the on-site surveys, shown in Table 8.1. The side dishes and desserts are actually not served and the elapsed time in Table 8.1 means the time to choose and pick up a dish from shelves. The head consumer is relieved of the queue after receiving the ordered menu. The next consumer in the queue moves in front of the window and the other consumers fill up the front.

Some customers consider to pick up other menus, such as rice, miso soup, side dishes, and desserts, after their first menu. The probability to consider the additional menus is set as shown Table 8.2. The customer who selected a main dish, for instance, consider to pick up rice, miso soup, side dishes, and desserts with a probability of 90%, 80%, 50%, and 10%, respectively. When a customer who considers to pick up side dishes, the customer actually picks up from zero to four dishes. The number of picked side dishes is assumed to have the binomial distribution with occurrence probability 30%. When a customer who considers the other types of additional menus, the customer always picks up one item of the type.

Table 8.1 Elapsed time for a menu at windows and spots (s)

Noodles	RB&Cs	Main dishes	Self-deli
18	37	6	5
Side dishes	Rice	Miso soup	Desserts
5	3	5	3

Table 8.2 Probability to select additional menu

First/additional	Rice (%)	Miso soup (%)	Side dishes (%)	Desserts (%)
Noodles	1	5	50	10
RB&Cs	0	20	30	10
Main dishes	90	80	50	10
Self-deli	0	10	20	10

8.3.4 Payment

Customer agents having their all menus travel toward a cash register to make a payment. The customers select the register with the shortest queue among six registers. The closer register is selected if there are some registers with the same smallest length of queue. The head customer in a queue waits for the payment processing whose elapsed time is set to 15 s determined based on the on-site surveys. The customer agent leaves from the order and pick-up area and is eliminated from the simulation space.

8.4 Numerical Experiments

The proposed multi-agent simulation model was developed on artisoc [14], the multi-agent simulation platform made by Kozo Keikaku Engineering Inc. Two kinds of numerical experiments using the model have been conducted. The first is to estimate the effect of new layouts for the order and pick-up area on the lunch break congestion. The second is to investigate the effect of time delay by customers who make an order slowly. Through the execution of simulation, dwell time of customers is measured as the time period from arriving the order and pick-up area to leaving from it. This study adopts the dwell time as the indicator of the target congestion since customer's dwell time becomes longer due to heavier congestion.

8.4.1 Effect by Layout Changes

Preliminary experiments confirmed that the right side of the order and pick-up area, where there are three queues to the main order windows, has more customers than the left side as shown in Fig. 8.2. The customers in the two queues from the right, for a noodle or RB&C, have to cross other queues in walking to a cash register. The crossing extends the walking time and it might occur an accident to collide other customers. Increasing the space between queues sounds efficient to relax the congestion in the right side of the area.

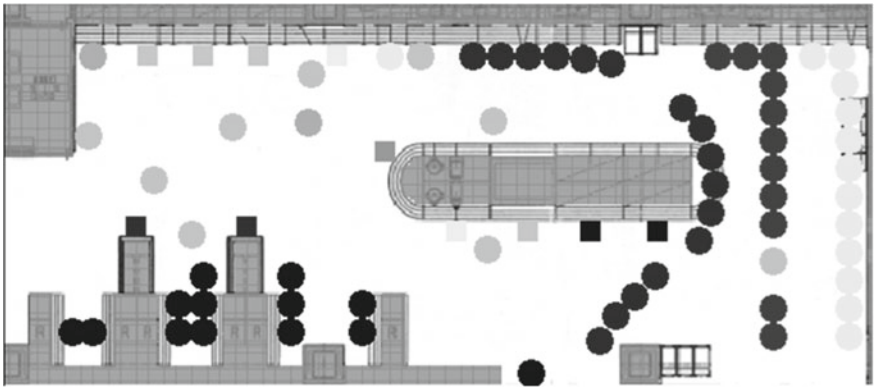


Fig. 8.3 Simulation space of a new layout, named Layout (1)

On the contrary, the left side of the area has enough space for customers to travel in Fig. 8.2. The preliminary experiments also confirmed that there are not so many customers in the queues for cash registers. It is preferable for the restaurant manager to close one cash register because the staff for the closed register is free and able to be assigned to another task. This numerical experiment also estimates the effect the register closure on the customer's congestion.

A new layout is shown in Fig. 8.3 where a right part of the self-deli island is detached so as to extend the spaces among the three queues for noodles, RB&Cs, and main dishes. The queues for RB&Cs and main dishes are partially separated to make space for customers who cross the queue. The separation reduces two and six customers in the respective queues for RB&Cs and main dishes inside the area. Since the elapsed time to order a menu at the head of the queues is given as shown Table 8.1, 74 (37 times 2) s and 36 (6 times 6) s should be reduced in the dwell time of customers in the queues for RB&Cs and main dishes.

The layout is proposed to makes customers to be easy to travel after obtaining their menus and called Layout (1) in the following. The original layout is named Layout (0) as a target for comparison of Layout (1). This experiment introduces two more layouts where the rightmost cash register is closed in Layouts (0) and (1). The layouts are named Layouts (0-) and (1-), respectively.

The simulation has been executed fifty times for the four layouts to measure the dwell time of customers. The results of the simulation are summarized in Fig. 8.4.

Figure 8.4 indicates that the customers for RB&Cs require longest dwell time followed by noodles, main dishes, and self-deli. The dwell time for main dishes is shorter than that for noodles because of the difference between the corresponding elapsed times shown in Table 8.1 while the length of queue for main dishes tends to be longer than that for noodles.

In Layout (1), customer's dwell time for all types of menu except for self-deli is shortened in comparison with in Layout (0). The t-test to estimate the difference between the population means of dwell time in Layouts (0) and (1) has been employed

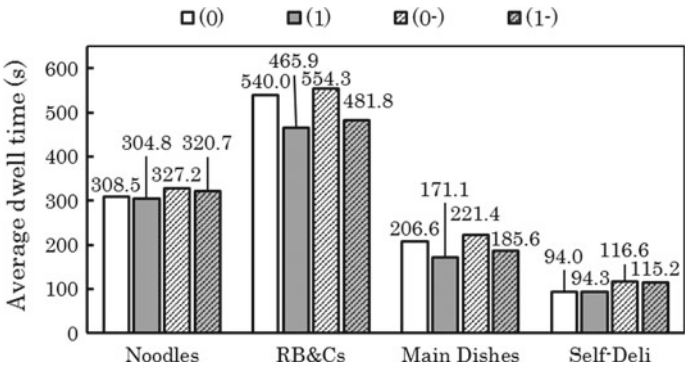


Fig. 8.4 Average dwell time of customers for each first choice in the four layouts

with considering the inevitable decrease by 74 s and 36 s for RB&Cs and main dishes, respectively. The test using a level of significance of 0.05 concludes that Layout (1) reduces the average dwell time for noodles by 2 s compared with Layout (0), and causes no significant change for the other first choices. In comparison with Layout (0-) to Layout (1-), only noodles have significant difference on the average dwell time of customers by 5 s at a level of significance of 0.05.

The change from Layout (0) to Layout (0-), which means closing one cash register, increases the overall average dwell time by 17.0 s, which is 107% of that for Layout (0). The degrees of the increase for noodles and self-deli are respectively 18.8 s and 22.6 s, greater than those for RB&Cs and main dishes. This is because the closed cash register locates on the rightmost and the register is convenient for customers who select noodles and self-deli. The t-test with the same significance condition concludes that the closing of the register increases customer's dwell time by 17, 12, 14, and 21 s for noodles, RB&Cs, main dishes, and self-deli, respectively.

The above discussion concludes that the new layout shown in Fig. 8.3 itself has little contribution to the reduction of customer's dwell time except for the customers for noodles, while the generated space between the queues might decrease collusions and relax customer's mind. The reduction of one cash register makes a staff free while customers wait 17 s longer to obtain their meals on average.

8.4.2 Effect by Customer's Order Delay

At the windows in the cafeteria counter, a menu is served in a dish after a customer make an order. Some customers, who are usually guests outside the campus or new students, take long time to decide what to order at the head of a queue. The order delay directly increases waiting time for other customers in the queue. This subsection estimates the influence of the time delay.

The experiments in this subsection have two kinds of parameters α and β where α is the rate of the slow customers and β is the time delay on the second time scale. The values of α and β have been set from 10% to 50% with 10% increments in between and integers from 0 to 5, respectively. For all combinations of the values of α and β, the simulation has been executed fifty times to measure customer's dwell time.

Figure 8.5 represents the simulation results of customer's average dwell time for noodles, RB&Cs, and main dishes where the result for self-deli is omitted since self-deli does not have its queue. In Fig. 8.5, the average dwell time for all types of menus increases approximately linearly with respect to time delay β for any values of α. The average dwell time for main dishes is more sensitive than the others to the value of β since the length of the queue for main dishes is longer than those of the others. Figure 8.5 also shows that the increment with respect to β is greater for greater value of α for any types of menus.

The regression lines for each data series in Fig. 8.5 have been derived and the slopes of the regression lines are displayed in Fig. 8.6. Figure 8.6 clarifies that the customers for main dishes are influenced more strongly than the other types of menus by the existence of slow customers to order. The data series in Fig. 8.6 approximately vary linearly with the value of α and the sloped for the data series also have been derived. As a result, the average dwell times T is approximately expressed by Eq. (8.1).

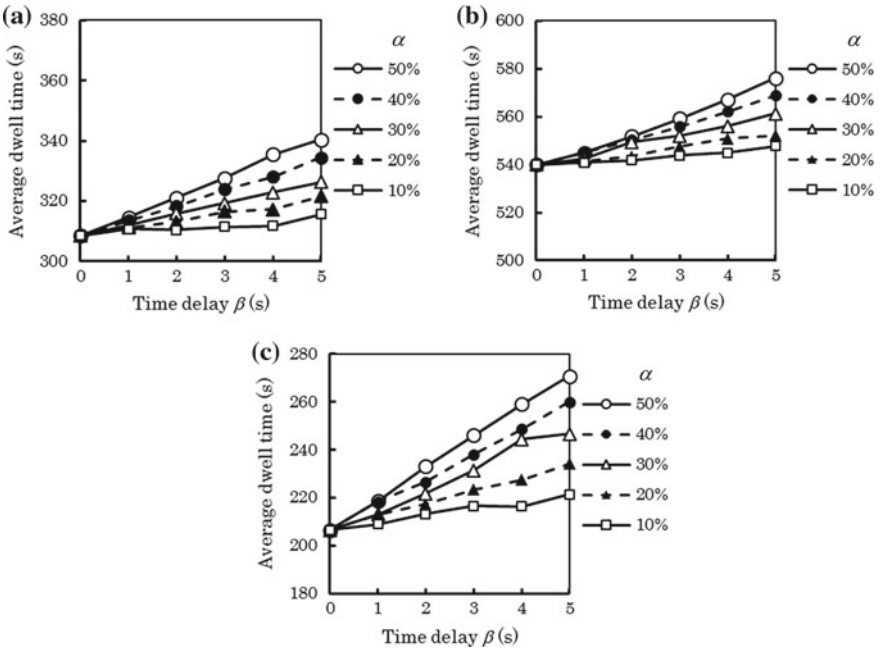


Fig. 8.5 Average dwell time of customers with delays **a** noodles, **b** RB&Cs, **c** main dishes

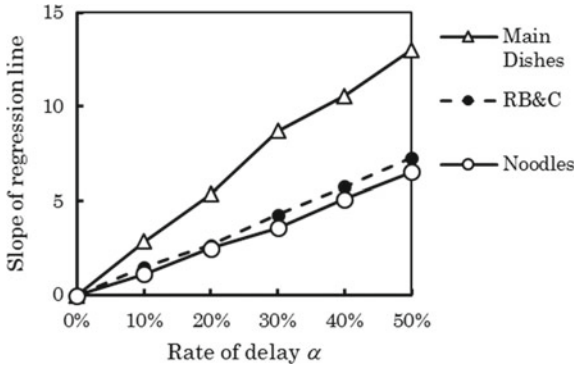


Fig. 8.6 Slope of regression lines for data series in Fig. 8.5

$$T = \begin{cases} 13.0\alpha\beta + 308.5 & \text{(noodles)} \\ 14.5\alpha\beta + 539.3 & \text{(RB\&Cs)} \\ 26.1\alpha\beta + 206.7 & \text{(main dishes)} \end{cases} \tag{8.1}$$

In Eq. (8.1), the first terms in the right-hand side mean the influence by the order delay and the constants mean the dwell time of customers with no order delay. A decrease of 1% in α or β reduces the influenced time by 1% in any type of menus. The decrease of the value of α or β for main dishes is approximately twice as effective as that for noodles and RB&Cs.

From the managerial point of view, it is important to display a menu list clearly so that even indecisive customers instantly decide their menu and their order delay becomes shorter. Bigger signs of menus might be effective to help customers in a queue to decide their orders before reaching the head of the queue.

8.5 Conclusion and Future Works

This paper constructed a multi-agent simulation model to estimate the congestion in a campus restaurant at our university. Experimental simulation study using the model has been conducted to investigate the effect on the congestion by changing the present area layout and the influence by the existence of customers who make an order slowly. The simulation study has revealed that the proposed new layout has little effect on the congestion and the closure of a cash register surely increases the dwell time of customers in the order and pick-up area. The influence on customer's dwell time in queues to three main windows by order delays has been expressed as an equation of the rate of the slow customers and the time delay.

The proposed model could be modified to deal with the queues outside the simulation space in the model. In a real situation, some queues extend up to the outside of the target area and some customers change their preferences on menus to select

another shorter queue so that their dwell time makes shorter. Furthermore, other evaluation scales could be considered to estimate the congestion in the restaurant, such as waiting time in each queue and times of near collisions, which will be discussed in a forthcoming paper.

Acknowledgements The authors are grateful to the manager and the staffs in Konan University Cooperative for cooperation of on-site surveys and useful discussion. We also thank Tetsuya Miyoshi for his helpful comments on the pedestrian modeling and for introducing his related papers. This work was partially supported by MEXT, Japan.

References

1. J.M. Epstein, R. Axtel, *Growing Artificial Societies—Social Science from the Bottom Up* (The Brookings Institution, 1996)
2. M. Resnick, *Turtles, Termites, and Traffic Jams—Explorations in Massively Parallel Microworlds* (The MIT Press, 1994)
3. A.M. Uhrmacher, D. Weyns, *Multi-agent Systems: Simulation & Applications* (CRC Press, 2009)
4. K. Miwa, S. Takakuwa, Simulation modeling and analysis for in-store merchandizing of retail stores with enhanced information technology, in *Proceedings of the 40th Conference on Winter Simulation Conference, WSC'08*, Miami, USA, 7–10 December 2008, pp. 1702–1710
5. B. Chen, H.H. Cheng, A review of the application of agent technology in traffic and transportation systems. IEEE Trans. Intell. Transp. Syst. **11**(2), 485–497 (2010)
6. T. Tanizaki, T. Shimmura, Modeling and analysis method of restaurant service process. Procedia CIRP **62**, 84–89 (2017)
7. S.A. Curin, J.S. Vosko, E.W. Chan, O. Tsimhoni, Reducing service time at a busy fast food restaurant on campus, in *Proceedings of the 2005 Winter Simulation Conference, WSC'05*, Orlando, USA, 4–7 December 2005, pp. 2628–2635
8. S.L. Jaynes, J.O. Hoffman, Discrete event simulation for quick service restaurant traffic analysis, in *Proceedings of the 1994 Winter Simulation Conference, WSC94*, Orlando, USA, 11–14 December 1994, pp. 1061–1066
9. K. Farahmand, A.F.G. Martinez, Simulation and animation of the operation of a fast food restaurant, in *Proceedings of the 1996 Winter Simulation Conference, WSC96*, Coronado, USA, 8–11 December 1996, pp. 1264–1271
10. D.M. Brann, B.C. Kulick, Simulation of restaurant operations using the restaurant modeling studio, in *Proceedings of the 2002 Winter Simulation Conference, WSC'02*, San Diego, USA, 8–11 December 2002, pp. 1448–1453
11. Y. Kitakami, A. Kato, Y. Taniwaki, F. Khasawneh, A study of ambient behavior in campus restaurant using multi-agent simulation, in *AIJ Tokai Chapter Architectural Research Meeting*, vol. 45 (2007), pp. 529–532 (in Japanese 2007)
12. T. Koide, M. Kikuda, An analysis on the effect of speed seats for congestion mitigation in a campus restaurant using multi-agent simulation, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 853–857
13. T. Miyoshi, H. Nakayasu, Y. Ueno, P. Patterson, An emergency aircraft evacuation simulation considering passenger emotions. Comput. Ind. Eng. **62**, 746–754 (2012)
14. artisoc, Multi-Agent Simulation Platform [Online]. http://www.kke.co.jp/en/solution/theme/artisoc.html

Chapter 9
Dynamic Virtual Bats Algorithm with Probabilistic Selection Restart Technique

Ali Osman Topal, Yunus Emre Yildiz and Mukremin Ozkul

Abstract Nature inspired algorithms have gained increasing attention as a powerful technique for solving optimization problems. Dynamic virtual bats algorithm (DVBA) is a relatively new nature inspired optimization algorithm. DVBA, like Bat Algorithm (BA), is fundamentally inspired by bat's hunting strategies, but it is conceptually very different from BA. In DVBA, a role based search is developed to avoid deficiencies of BA. Although the new technique outperforms BA significantly, there is still an insufficiency in DVBA regarding its exploration, when it comes to high dimensional complex optimization problems. To increase the performance of DVBA, this paper presents a novel, improved dynamic virtual bats algorithm (IDVBA) based on probabilistic selection. The performance of the proposed IDVBA is compared with seven meta-heuristic algorithms on a suite of 30 bound-constrained optimization problems from CEC 2014. The experimental results demonstrated that the proposed IDVBA outperform, or is comparable to, its competitors in terms of the quality of final solution and its convergence rates for high dimensional problems.

Keywords Bat algorithms · Bio-inspired algorithms · Dynamic virtual bats algorithm · Global numerical optimization · Meta heuristic algorithm · Nature inspired algorithms

A. O. Topal (✉) · M. Ozkul
Computer Engineering, Epoka University, Tirana, Albania
e-mail: aotopal@epoka.edu.al

M. Ozkul
e-mail: mozkul@epoka.edu.al

Y. E. Yildiz
Computer Engineering, Metrapolitan University, Tirana, Albania
e-mail: theyeyildiz@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_9

9.1 Introduction

Bat Algorithm was introduced by Yang as a new metaheuristic method that was based on the echolocation behavior of bats. It is more like a combination of Particle Swarm Optimization and Harmony Search [22]. Although BA does not imitate the real bats successfully, due to its simplicity and effectiveness a large number of BA variations have been developed and applied to a wide range of real problems [9, 12, 14, 23, 25]. However, as in most of the stochastic algorithms, the standard BA suffers from the premature convergence problem and it needs improvements in exploration. The random walk size and the pulse rate parameters play very important role on exploration ability of BA. So most of the researchers focused on these parameters. Recently, Local Memory Search Bat Algorithm (LMSBA) [25], Novel Adaptive Bat Algorithm (NABA) [8], Adaptive Bat Algorithm (ABA) [21], and Chaotic Local Search-based Bat Algorithm (CLSBA) are developed towards improving BA's exploration ability.

DVBA, by Topal and Altun [16, 17], is another meta-heuristic algorithm inspired by bats ability to manipulate frequency and wavelength of sound waves emitted during their hunt. DVBA is not a BA variation, it is a new simulation of the bat's hunting strategies. In DVBA, a role-based search is developed to improve the diversification and intensification capability of the Bat Algorithm. There are only two bats: explorer and exploiter bat. While the explorer bat explores the search space, the exploiter bat makes an intensive search of the local with the highest probability of locating the desired target. During the search bats exchange the roles according to their positions.

Recently, DVBA has been tested on well-known test functions [16, 17] and supply chain cost problem [18]. Experimental results show that, DVBA is suitable for solving most of the low dimensional problems. However, DVBA, similar to other evolutionary algorithms, has some challenging problems. For example, the convergence speed of DVBA is slower than other population-based algorithms like PSO [10], GA [5], BF [13], and BA. The convergence issue of these algorithms was resolved through the latest variants [4, 24]. Additionally, in high dimensional multimodal problems, escaping from the local optima traps becomes a difficult task for DVBA. In fact, most stochastic optimization algorithms, including particle swarm optimizer (PSO) [10] and genetic algorithm (GA) [5], suffer from the curse of dimensionality [20]. Simply put, this implies that their performance deteriorates as the dimensionality of the search space increases. Therefore, accelerating convergence speed and avoiding the local optima have become two of the most important issues in DVBA.

To minimize the impact of this weakness, an improved version of DVBA is proposed which accelerates convergence speed and avoids the local optima trap. To achieve both goals, we introduce a new search mechanism for the explorer bat. This new search mechanism improves the search performance and gives DVBA more powerful exploitation capabilities. Simulations and comparisons based on several well-studied benchmarks demonstrate the effectiveness, efficiency and robustness of the proposed IDVBA [19].

The rest of this paper is organized as follows. Section 9.2 summarizes BA, variants of BA, and DVBA. The improved DVBA algorithm is presented in Sect. 9.3.

Section 9.4 presents numerical experiments and results on the use of IDVBA for solving CEC 2014 test functions. Finally, conclusions are drawn in Sect. 9.5.

9.2 Bat Inspired Algorithms

In this section we will present a brief theory of the bat algorithm, 4 state of art modified versions of BA algorithms, and dynamic virtual bats algorithm.

9.2.1 Bat Algorithm (BA)

Bat algorithm was introduced by Yang in 2010 [22]. It is a population based algorithm which uses bat's echolocation ability to get optimum solution for tough optimization problems. Echolocation is a typical sonar system which bats use to detect prey, avoid obstacles, and locate bats' roosting crevices in the dark. Bat emits sound pulses and listens to the returning echoes by using the delay time, loudness of the response and the time difference between its ears, it can tell the shape, size, and the velocity of the prey. Bat has also the ability to change the way it emits the sound pulses. If it emits the sound pulses with high frequency, they will not travel longer but give detailed information about its close environment which helps bat to detect the prey position precisely. When bat emits sound pulses with low frequency, they will travel farther, and give rough information about its surroundings [1, 2].

The algorithm starts by placing n bats randomly in the search space. Velocity V_i, frequency f_i, pulse rate r_i, and loudness A_i are initialized for each bat at the beginning. In each iteration of the main loop, by using Eqs. (9.1), (9.2), and (9.3) bat's position and velocity are updated.

$$f_i = f_{min} + (f_{max} - f_{min})\beta, \tag{9.1}$$

$$V_i^t = V_i^{t-1} + (x_i^t - x_*)\, f_i, \tag{9.2}$$

$$x_i^t = x_i^{t-1} + V_i^t, \tag{9.3}$$

where $\beta \in [0, 1]$.

Then the algorithm evaluates the finesses (solutions) and chooses the current best position x_*. After these updates, if the bat's pulse rate r_i is low (it means that the bat is far away from the prey), with a high probability $(rand() > r_i)$ it will move near to the current best bat and make a random short fly there. If its pulse rate is high then it should be near the prey and it will just make a random fly around its current position. New position x_{new} is obtained by Eq. (9.4).

$$x_{new} = x_{old} + \varepsilon A^t, \tag{9.4}$$

where $\varepsilon \in [-1, 1]$ a random number, $A^t = \frac{1}{n}\sum_{i=1}^{n} A_i^t$ is the average loudness of all the bats at this time step, and n is the number of bats. After the random fly, if the bats position is better than the current global best and its loudness is loud enough to be greater than a random number ($rand() < A_i$), the bat will fly to this position and current global best will be updated with the new one. The bats pulse rate r_i will be increased and loudness A_i will be decreased as shown in Eq. (9.5). Then again bats will be evaluated and the current best x_* will be updated.

$$A_i^{t+1} = \alpha A_i^t, \quad r_i^{t+1} = r_i^0[1 - exp(-\gamma t)], \tag{9.5}$$

9.2.2 Local Memory Search Bat Algorithm (LMSBA)

In LMSBA [25], Yuanbin introduced a new random walk equation in which the bat's personal best position is added in the random walk equation to update the bat's position. However, the traditional random walk equation is not removed, it is used for comparison with the new equation and the best one is chosen to update the bat's position.

By using this method, LMSBA intends to create local extreme search in BA local search. However, this method alone might lead the bats to a local optima and they can be trapped there easily. Because, once they fly to the vicinity of the best solution, their personal best will be same as the best solution and the alternative solution won't help them to escape from local optima trap.

9.2.3 Novel Adaptive Bat Algorithm (NABA)

NABA [8] incorporates two techniques within BA, which include the Rechenberg's 1/5 mutation rule and the Gaussian/Normal probability distribution to produce mutation step size.

NABA, like LMSBA, offers new equation to generate new solution rather than doing just random walk. It modifies the random walk equation (Eq. 9.4) in line 11 of original Bat algorithm. NABA controls the random walk step size by the variance of Gaussian/Normal distribution. The modified equation is as follows:

$$x_{new} = x_{old} + \varepsilon A^t N(0, \sigma) \tag{9.6}$$

where σ is the standard deviation.

Kabir used the Rechenberg's 1/5 mutation rule [15] to adaptively change the random walk step size and pulse rate to control the exploration and exploitation. if the success-rate is less than 1/5, NABA is exploring too much and it intends to move the bats near to the best solution and decreases the step size. If the success-rate is more than 1/5, NABA is exploiting local optima too much and it increases the

step size. That will help bats to reach the local optima faster, but the accuracy will decrease because of the long step size.

9.2.4 Adaptive Bat Algorithm (ABA)

Wang presented an improved bat algorithm [21] to solve BA's premature convergence problem. Same as NABA and LMSBA, he improved the random walk equation, but he also modified the frequency and the velocity equations in BA. In BA, each bat uses the same frequency increment for the velocity which makes bat's flight behavior lack of flexibility. In ABA, Wang proposed a new method to let each bat dynamic and adaptively adjust its flight speed and its flight direction. ABA is targeting to increase the speed of the bat which is farther from the prey. The farther the distance between the bat and its prey (global best solution), the faster the speed flying to its prey.

Secondly, Wang improved the random walk by combining it with shrinking search. Bats which are far away from the current best, make random fly near to the current best. The step size of random fly within the range $[-1, 1]$ in BA while it shrinks in ABA as the iteration proceeds. ABA targets to increase the intensification of the search by shrinking search method.

9.2.5 Chaotic Local Search-Based Bat Algorithm (CLSBA)

CLSBA [2] is a combination of the standard BA with chaotic sequences generated by the logistic map. The use of chaos makes the frequency adaptive and more random in nature to balance the trade-off between exploration and exploitation. The frequency of pulse emission is modified as follows in CLSBA:

$$f_i(t+1) = \mu \times f_i(t) \times (1 - f_i(t)) \tag{9.7}$$

While the step size of random walk parameters ε (Eq. 9.3) varies within the range $[-1, 1]$ in BA, it varies within the range $[-SF(t), SF(t)]$ in CLSBA. SF is the scaling factor which changes dynamically according to the Rechenberg's mutation rule. Having a lower value of SF(t) increases the exploitation, a larger value of SF(t) accelerates the exploration.

9.2.6 Dynamic Virtual Bats Algorithm (DVBA)

Dynamic virtual bats algorithm (DVBA) [16], proposed in 2014 for global numerical optimization, is a recently introduced optimization algorithm which imitates the bats echolocation behavior in nature. When bats search out prey, they burst sound pulses

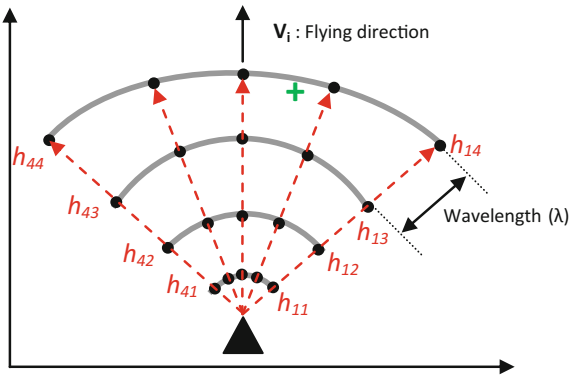


Fig. 9.1 Exploration: explorer bat is searching for prey with a wide search scope

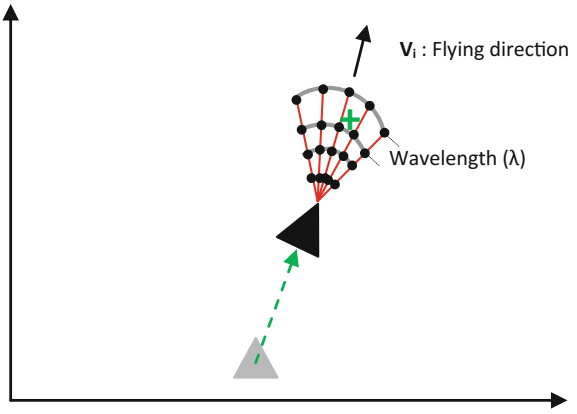


Fig. 9.2 Exploitation: exploiter bat is chasing prey with a narrow search space

with lower frequency and longer wavelengths so the sound pulses can travel farther. In this long range mode it becomes hard to detect the exact position of the prey; however, it becomes easy to search a large area. When bats detect prey, the pulses will be emitted with higher frequency and shorter wavelengths so that bats are able to update the prey location more often [1, 7]. In DVBA, two bats are used to imitate this hunting behavior. Each bat has its own role in the algorithm and during the search they exchange these roles according to their positions. These bats are referred as explorer bat and exploiter bat. The bat that is in a better position becomes the exploiter meanwhile the other becomes the explorer. While the exploiter bat increases the intensification of the search around the best solution, the explorer bat will continue to explore other solutions.

In Figs. 9.1 and 9.2, the hunting strategy of a bat is simulated. The black triangle is the current solution (bat location), the black circles are the visited solutions on the search waves in this iteration. During the search, the explorer bat's search scope gets in its widest shape; the distance between the search waves and the angle between the wave vectors (red dashed arrows) get larger (see Fig. 9.1). On the contrary, if the bat becomes the exploiter bat, its search scope gets its narrowest shape; the distance

Algorithm 1 DVBA pseudo code. f_{gbest} is the global best solution and d is the number of dimensions. [16]

Objective function $f(x)$, $x = (x_1, ..., x_d)^T$
Initialize the bat population $x_i (i = 1, 2)$ and v_i
Initialize wavelength λ_i and frequency f_i
Initialize the number of the waves
while ($t <$ Max number of iterations) **do**
  for each bat **do**
    Create a sound waves scope
    Evaluate the solutions on the waves
    Choose the best solution on the waves, h_*
  if ($f(h_*) < f(x_i)$) **then**
    Move to new solution
    Decrease λ_i and increase f_i
  else if ($f(x_i) > f_{gbest}$) **then**
    Change the direction randomly
    Increase λ_i and decrease f_i
  else if ($f(x_i) = f_{gbest}$) **then**
    Minimize λ_i and maximize f_i
    Change the direction randomly
  end if
  Rank the bats and find the current best x_{gbest}
end while

between the search waves and the angle between the wave vectors get smaller (see Fig. 9.2). The number of the visited solutions is same for both bats, just the distance between the solutions changes dynamically by using wavelength and frequency. The wavelength and the distance between the solutions are proportional. The frequency and the angle between the wave vectors are inversely proportional.

The algorithm determines the best solution in the bat's search scope. If it is better than the current location (solution), the bat will fly to the better solution, decrease the wavelength and increase the frequency for the next iteration. These changes are targeting to increase the intensification of the search. Unless there is no better solution than the current solution in the search scope, the bat will stay on it, turn around randomly and keep scanning its nearby surrounding space. It will keep spinning in this position and expanding its search scope until it finds a better solution. Bats also exchange the roles according to their position.

The basic steps of the algorithm for minimizing an objective function $f(x)$ are shown in Algorithm 1.

9.3 Proposed Novel Improved Dynamic Virtual Bats Algorithm (IDVBA)

9.3.1 The Weakness of DVBA

In DVBA, the explorer bat's search scope size is limited by the wavelength. This search scope might not be large enough to detect better solutions near its surrounding space. Thus, it is very likely that the explorer bat will be trapped in a local optimum. In addition, the exploiter bat's search scope size becomes very small during the exploitation and it moves very slowly. Therefore it might need too much time to reach the global optima. These problems in DVBA have been eradicated by introducing probabilistic selection restart techniques in IDVBA.

The concept of IDVBA is analogous to the idea behind the Micro-Particle Swarm Optimizer [6] and the Micro-Genetic Algorithm [11], where a set of restart operations are executed after the population has converged. However, IDVBA uses restart operations according to the bats' stagnancy during the search and it does not blacklist the inferior solutions as in micro-genetic algorithm.

9.3.2 Improved Dynamic Virtual Bats Algorithm (IDVBA)

To improve the search performance and give DVBA more powerful search capabilities, we introduce two probabilistic selections: *R*-random flying probability and *C*-convergence probability. If the explorer bat is stuck in large local minima, it chooses to fly away from the trap randomly with a probability *R* related to number of unsuccessful attempts. In addition, it chooses to fly near to the exploiter bat with a probability *C* related to the number of escapes attempted from the traps. *R* is calculated as follows:

$$R_i^{t+1} = R_i^t[1 - exp(-trial_i\gamma)], \tag{9.8}$$

where γ constant and $trial_i$ denotes the number of unsuccessful attempts. Obviously, the higher the $trial_i$ is, the greater the probability that the explorer bat might fly away from the trap to a random solution in the search space. As the unsuccessful attempts increase, the random flying probability R_i decreases and the possibility of $rand() < R_i$ being true (line 17 in Algorithm 2) increases. This can help the explorer bat to escape from the local optima trap rapidly. However, the explorer bat should not leave the trap without exploring the nearby surrounding space. Thus, γ should be chosen carefully.

The convergence probability *C* gives possibility to the explorer bat to converge with the exploiter bat, instead of exploring a random position. Thus, the exploitation speed will be increased rapidly around the best position. Time after time the explorer bat visits the exploiter bat to speed up the exploitation process then flies

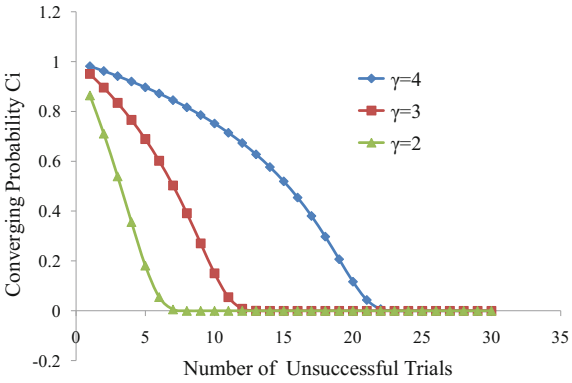


Fig. 9.3 Characteristics of Eq. 9.9. C-convergence probability changes as the unsuccessful trials increases for $\gamma = 2, 3$ and 4

away randomly to keep up the exploration process. This also increases the exploitation capability of IDVBA. C is calculated as follows:

$$C_i^{t+1} = C_i^t[1 - exp(-rfly_i\gamma)], \tag{9.9}$$

where $rfly_i$ denotes the number of random restarts. As shown in Eq. 9.9, the convergence probability C is inversely proportional with the number of random restarts $rfly_i$ done by the explorer bat. Thus, increasing random restarts will increase the probability of $rand() > C$ that allows the explorer bat to visit the exploiter bat often (line 25–26 in Algorithm 2).

Figure 9.3 shows the characteristics of Eq. 9.9 for different values of γ as the unsuccessful attempts of the explorer bat increases. As it can be seen from Fig. 9.3 and Algorithm 2, the higher γ value will keep the explorer bat longer in local minima. Our preliminary studies have suggested that the best value for γ is in the range of [2, 4] for the test functions we use in this work. R is calculated with the same equation used for C Eq. 9.9 but with different values of γ. Therefore Fig. 9.3 shows the characteristics of Eq. 9.8 as well. By properly choosing the value of γ, we can prevent bats from local minima traps while allowing them to explore and exploit the nearby surrounding space without experiencing too many unnecessary random jumps. In our experiments, we set $\gamma = 2$ for R and $\gamma = 3$ for C. According to all these approximations and improvements IDVBA can be given as in Algorithm 2. The differences from DVBA are shown in boldface font.

9.4 Numerical Experiments and Results

This section presents an extensive comparison among the performances of eight algorithms in two groups. In the first group, PSO, BA, DVBA, and IDVBA were compared. In the second group, IDVBA was compared with four state-of-the-art BA variants: LMSBA, NABA, ABA, and CLSBA.

Algorithm 2 IDVBA pseudo code. f_{gbest} is the global best solution and d is the number of dimensions. The code we discuss in the text is in boldface.

```
Objective function f(x), x = (x_1, ..., x_d)^T
Initialize the bat population x_i (i = 1, 2) and v_i
Initialize wavelength λ_i and frequency f_i
Initialize the number of the waves
C_i = R_i = 1.0
while (t < Max number of iterations) do
    for each bat do
        Create a sound waves scope
        Evaluate the solutions on the waves
        Choose the best solution on the waves, h_*
    if f(h_*) < f(x_i) then
        Move to new solution
        Decrease λ_i and increase f_i
        trial_i = 0
    else if f(x_i) > f_gbest then
        Calculate the random flying probability R_i by Eq. 9.8
        if rand() < R_i then
            Change the direction randomly
            trial_i = trial_i + 1
        else
            Restart the search from a random position
            Reset R_i and trial_i
            rfly_i = rfly_i + 1
            Calculate the C_i by Eq. 9.9
        if rand() > C_i then
            Produce a new solution around the exploiter bat.
            Reset C_i and rfly_i
        Increase λ_i and decrease f_i
    else if f(x_i) = f_gbest then
        Minimize λ_i and maximize f_i
        Change the direction randomly
        Reset rfly_i and trial_i
    end if
    Rank the bats and find the current best x_gbest
end while
```

In our experimental studies, parameter settings of the algorithms are the same as in their original papers. The population size of the algorithms has been kept 50 to regardless to the dimension size of the problem.

9.4.1 Comparison Experiments

This paper aims to test the quality of the final solution and the convergence speed at the end of a fixed number of function evaluations (FEs). According to the instructions

in CEC 2014 special session we set the maximum number of FEs 3×10^5 and the dimensions of the problems 30-D. In the tables, the test results are shown in terms of the mean error (MeanErr) and the standard deviation (STDEV) of the results. The mean error values are found according to $(F(x) - F(x^*))$ for evaluating the success of algorithms, where x is the best found value in a run and x^* is the global best of the test function. The best mean error values are typed in bold.

Furthermore, we used t-tests [3] to compare the means of the results produced by the IDVBA and the other algorithms at the 0.05 level of significance. The statistical significance level of the results are shown in the Tables 9.1 and 9.2. There '+' indicates that IDVBA is significantly more successful than selected one at a 0.05 level of significance by two-tailed test, '$\sim$' means the difference of means is not statistically significant and '$-$' indicates that IDVBA could not win the corresponding algorithm.

9.4.2 Experimental Results and Discussion

The functions f_1, f_2, and f_3 are unimodal and non-separable plate shape problems. From Tables 9.1 and 9.2, it can be seen that these functions are very hard to optimize and the algorithms did not show a significant success. By analyzing their t-test values, we can see that the solutions with IDVBA are significantly better than that with other algorithms, followed by NABA; PSO has the worst solutions. For the function f_3, NABA has the best solutions in terms of Mean Error, followed by CLSBA.

The test functions in second group ($f_4 - f_{16}$) are simple multi-modal. The t-test values show that the performance of IDVBA is significantly better than that of other seven algorithms. In this group, DVBA remained the toughest competitors of IDVBA in most of the functions.

The third group ($f_{17} - f_{22}$) has six hybrid functions which are almost same as real-world optimization problems. On these functions, IDVBA exhibits better performance than seven other algorithms. For function f_{20}, BA, LMSBA, NABA, and ABA have less mean error than that of IDVBA. LMSBA, CLSBA, and NABA performed comparable to IDVBA on f_{19}. In this group, LMSBA and NABA also find competitive solutions compared with IDVBA, as its t-test values reflect.

In group four, there are eight composition functions. The composition function merges the properties of the sub-functions better and maintains continuity around the global/local optima. We can observe that, the performance of IDVBA is superior overall to that of seven competitors except on the function f_{30} which LMSBA have better solutions.

It is explicitly observed that IDVBA found effective results in most cases, which means that IDVBA using improved search mechanism is more competent in tackling complex problems.

Additionally, we can observe that among the BA's variants, LMSBA has the best performance in terms of the mean, followed by ABA and NABA.

Table 9.1 Comparison of PSO, BA, DVBA, and IDVBA over 30 test functions of CEC 2014

Func.	PSO		BA		DVBA		IDVBA
	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)
f_1	5.41E+07 (2.24E+07)	+	6.48E+06 (6.82E+06)	+	1.05E+05 (4.28E+05)	≈	**6.65E+04** (1.20E+04)
f_2	2.42E+08 (4.23E+08)	+	8.74E+07 (3.76E+06)	+	2.93E+04 (1.32E+04)	+	**2.31E+04** (1.11E+04)
f_3	5.97E+04 (1.07E+04)	+	2.61E+04 (3.08E+03)	+	1.59E+04 (6.25E+03)	+	**9.19E+03** (6.25E+03)
f_4	1.66E+02 (8.29E+01)	+	9.89E+01 (2.15E+01)	≈	1.08E+02 (3.31E+01)	+	**9.51E+01** (2.14E+01)
f_5	2.08E+01 (9.01E+01)	+	2.09E+01 (1.25E−01)	+	2.01E+01 (6.11E−02)	+	**2.00E+01** (5.41E−02)
f_6	2.58E+01 (6.40E+01)	+	3.01E+01 (2.02E+00)	+	2.34E+01 (3.16E+00)	+	**5.48E+00** (9.96E−01)
f_7	9.28E+01 (1.23E+01)	+	1.61E+00 (6.43E−02)	+	1.07E+00 (1.12E−02)	+	**1.06E+00** (8.12E−03)
f_8	5.09E−02 (1.11E+00)	≈	5.10E−01 (3.29E−01)	+	3.00E−02 (1.76E−01)	≈	**2.37E−03** (8.42E−04)
f_9	3.18E+01 (1.28E+02)	≈	5.68E+01 (2.22E+01)	+	2.14E+01 (9.52E+00)	≈	**1.86E+01** (3.93E+00)
f_{10}	7.61E+02 (7.61E+02)	+	1.24E+03 (6.88E+02)	+	7.90E+02 (1.63E+02)	+	**6.16E+02** (1.28E+02)
f_{11}	6.83E+03 (1.93E+03)	+	5.39E+03 (7.14E+02)	+	3.93E+03 (5.23E+02)	+	**6.67E+02** (3.47E+02)
f_{12}	2.86E+00 (2.77E+01)	+	2.10E+00 (1.33E−01)	+	9.80E−01 (3.07E−01)	≈	**9.43E−01** (1.36E−01)
f_{13}	6.40E−01 (4.44E−01)	+	4.80E−01 (7.50E−02)	≈	5.10E−01 (1.03E−01)	+	**4.34E−01** (8.54E−02)
f_{14}	4.21E−01 (1.57E−01)	+	2.49E−01 (3.28E−02)	+	**2.00E−01** (2.05E−02)	≈	**2.00E−01** (1.85E−02)
f_{15}	4.47E+00 (1.92E+00)	+	2.62E+00 (1.24E+00)	≈	2.49E+00 (6.33E−01)	≈	**2.22E+00** (1.75E−01)
f_{16}	1.29E+01 (4.82E−01)	+	1.28E+01 (2.08E−01)	+	1.27E+01 (4.27E−01)	+	**1.22E+01** (3.28E−01)
f_{17}	3.26E+05 (2.57E+05)	+	1.33E+05 (4.89E+04)	+	8.09E+04 (3.42E+04)	+	**6.39E+03** (4.67E+03)
f_{18}	1.02E+06 (2.95E+06)	+	7.09E+05 (1.71E+05)	+	7.90E+03 (9.35E+03)	+	**8.25E+02** (7.28E+02)
f_{19}	2.82E+02 (1.98E+01)	+	1.68E+01 (2.03E+00)	+	1.81E+01 (2.33E+00)	+	**1.50E+01** (1.08E+00)
f_{20}	1.08E+04 (2.18E+03)	+	**5.01E+02** (8.42E+01)	−	8.29E+02 (3.22E+02)	+	5.51E+02 (3.22E+02)

(continued)

Table 9.1 (continued)

Func.	PSO		BA		DVBA		IDVBA
	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)
f_{21}	7.54E+05 (4.21E+05)	+	7.20E+05 (3.07E+04)	+	8.45E+04 (4.28E+04)	≈	**6.40E+04** (2.42E+04)
f_{22}	8.24E+02 (5.24E+02)	+	9.40E+02 (5.43E+02)	+	4.70E+02 (1.76E+02)	+	**3.70E+02** (1.02E+02)
f_{23}	3.42E+02 (6.24E+00)	+	3.17E+02 (4.06E−01)	+	**3.16E+02** (3.61E−01)	≈	**3.16E+02** (5.71E−01)
f_{24}	2.05E+02 (2.41E−01)	+	2.83E+02 (3.67E+01)	+	2.49E+02 (1.04E+01)	+	**1.40E+02** (7.43E+00)
f_{25}	2.20E+02 (4.51E+00)	+	**2.09E+02** (2.99E+00)	≈	2.19E+02 (9.11E+00)	+	2.11E+02 (5.29E+00)
f_{26}	**1.00E+02** (2.42E−01)	≈	**1.00E+02** (4.33E−01)	≈	**1.00E+02** (7.53E−02)	≈	**1.00E+02** (3.19E−02)
f_{27}	2.51E+03 (4.82E+02)	+	7.91E+02 (4.05E+02)	≈	5.70E+02 (2.61E+02)	≈	**4.50E+02** (7.21E+02)
f_{28}	1.81E+03 (4.91E+02)	+	2.57E+03 (3.15E+02)	+	1.56E+03 (3.08E+02)	+	**9.70E+02** (3.71E+02)
f_{29}	8.77E+07 (3.24E+07)	+	5.44E+05 (2.27E+05)	+	1.84E+04 (2.20E+04)	≈	**1.48E+04** (3.61E+04)
f_{30}	4.11E+05 (1.87E+04)	+	5.25E+04 (4.12E+04)	+	1.19E+04 (5.09E+03)	+	**4.20E+03** (4.26E+03)
+	27		24		19		
−	0		0		0		
≈	3		6		11		

"+", "−", and "≈" denote that the performance of the corresponding algorithm is significantly worse than, better than, and significantly not different to that of IDVBA, respectively

9.5 Conclusion

In order to apply DVBA to solve complex high dimensional problems efficiently, this paper proposes a novel improved dynamic virtual bats algorithm, namely IDVBA. The proposed algorithm employs two probabilistic selections—random flying probability R and convergence probability C. They are used to prevent the bats from falling into the local minimum traps, accelerate the convergence speed, and increase the accuracy. Results show that we achieved our goal efficiently with the new search mechanism.

To prove the effectiveness and robustness of the proposed algorithms, the proposed IDVBA was compared with the PSO, BA, DVBA, ABA, CLSBA, LMSBA, and NABA on all 30 bound-constrained numerical 30-D optimization problems from CEC-2014. The results demonstrate that the IDVBA achieves a good balance between exploration and exploitation and has the best universality on different type

Table 9.2 Comparison of LMSBA, NABA, CLSBA, ABA, DVBA, and IDVBA on all 30 test functions of CEC 2014

Func.	LMSBA		NABA		CLSBA		ABA		IDVBA
	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)
f_1	6.25E+6 (2.25E+6)	+	6.18E+6 (1.19E+6)	+	6.38E+6 (1.22E+6)	+	7.33E+6 (2.21E+6)	+	**6.65E+4** (1.20E+4)
f_2	7.53E+7 (2.61E+6)	+	5.46E+7 (2.92E+6)	+	6.47E+7 (6.56E+6)	+	7.18E+7 (4.93E+6)	+	**2.31E+5** (2.11E+4)
f_3	1.36E+3 (5.39E+2)	−	**2.48E+2** (2.73E+1)	−	5.20E+2 (1.02E+2)	−	1.54E+3 (8.97E+2)	−	9.19E+3 (3.21E+3)
f_4	1.02E+2 (4.45E+1)	≈	1.26E+2 (3.98E+1)	+	1.70E+2 (7.83E+1)	+	1.29E+2 (3.40E+1)	+	**9.51E+1** (1.23E+1)
f_5	2.08E+1 (3.92E−2)	+	2.11E+1 (3.08E−2)	+	2.12E+1 (3.13E−2)	+	2.09E+1 (3.07E−2)	+	**2.03E+1** (2.16E−2)
f_6	2.79E+1 (4.52E+0)	+	2.93E+1 (3.11E+0)	+	2.93E+1 (4.01E+0)	+	2.90E+1 (1.14E+0)	+	**5.48E+0** (9.96E−1)
f_7	1.58E+0 (9.03E−2)	+	1.48E+0 (4.48E−2)	+	1.56E+0 (9.90E−2)	+	1.59E+0 (6.72E−2)	+	**1.06E+0** (7.53E−3)
f_8	2.32E−1 (2.82E−1)	+	2.32E−1 (2.43E−1)	+	2.49E−1 (5.52E−1)	+	3.18E−1 (4.87E+1)	−	**2.37E−3** (8.42E−4)
f_9	5.67E+1 (4.97E+1)	+	5.83E+1 (5.87E+1)	+	4.60E+01 (3.03E+1)	≈	6.00E+1 (5.32E+1)	+	**3.86E+1** (3.93E+0)
f_{10}	1.51E+3 (6.57E+2)	+	1.05E+3 (6.02E+2)	+	1.69E+3 (6.34E+2)	+	1.73E+3 (5.72E+2)	−	**6.16E+2** (1.28E+2)
f_{11}	5.04E+3 (3.78E+2)	+	5.26E+3 (3.57E+2)	+	5.14E+3 (6.10E+2)	+	5.20E+3 (6.37E+2)	+	**6.67E+2** (3.47E+2)
f_{12}	2.11E+0 (1.99E−1)	+	1.93E+0 (1.54E−1)	+	2.03E+0 (1.94E−1)	+	1.87E+0 (3.16E−1)	−	**9.43E−1** (1.36E−1)
f_{13}	9.87E−1 (3.05E−2)	−	4.90E−1 (1.27E−1)	≈	4.90E−1 (7.33E−2)	+	**4.30E−1** (6.11E−2)	≈	4.34E−1 (8.54E−2)
f_{14}	3.10E−1 (3.48E−2)	≈	2.60E−1 (3.74E−2)	≈	2.70E−1 (5.46E−2)	≈	2.90E−1 (4.71E−2)	≈	**2.30E−1** (2.14E−1)
f_{15}	2.14E+1 (1.58E+0)	+	5.87E+0 (1.26E+0)	+	1.86E+1 (1.16E+0)	+	2.11E+01 (6.11E−1)	+	**2.22E+0** (1.75E−1)
f_{16}	1.41E+1 (1.76E−1)	+	1.33E+1 (4.86E−1)	+	1.34E+1 (3.58E−1)	+	1.30E+1 (7.35E−1)	+	**1.22E+1** (5.74E−1)
f_{17}	1.02E+5 (3.27E+5)	+	1.26E+5 (6.67E+4)	+	1.29E+5 (1.84E+5)	+	1.70E+5 (3.09E+5)	+	**8.39E+3** (4.67E+3)
f_{18}	9.91E+4 (2.77E+5)	+	4.15E+5 (2.9E+5)	+	1.04E+5 (1.17E+4)	+	2.17E+5 (1.36E+5)	+	**8.25E+2** (7.28E+2)
f_{19}	1.47E+1 (1.78E+0)	≈	**1.41E+01** (1.97E+0)	−	1.49E+1 (7.02E−1)	≈	2.80E+1 (2.71E+1)	+	1.50E+1 (1.28E+0)

(continued)

Table 9.2 (continued)

Func.	LMSBA		NABA		CLSBA		ABA		IDVBA
	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)	t-test	MeanErr (StdDev)
f_{20}	**3.52E+2** (5.14E+1)	−	3.87E+2 (1.16E+2)	−	5.30E+2 (5.34E+1)	≈	4.06E+2 (1.07E+2)	−	5.51E+2 (1.44E+2)
f_{21}	8.41E+4 (6.62E+4)	≈	**7.57E+4** (3.51E+4)	≈	1.13E+5 (6.87E+4)	+	1.36E+5 (8.29E+4)	+	9.40E+4 (4.70E+4)
f_{22}	8.97E+2 (1.67E+2)	+	1.00E+3 (2.59E+2)	+	1.03E+3 (8.22E+1)	+	1.01E+3 (2.06E+2)	+	**3.70E+2** (1.63E+2)
f_{23}	3.21E+2 (4.98E−1)	+	3.17E+2 (3.63E−1)	+	3.17E+2 (3.74E−1)	+	3.19E+2 (2.49E+0)	+	**3.16E+2** (6.31E−2)
f_{24}	2.33E+2 (6.72E+0)	+	2.43E+2 (1.31E+1)	+	2.55E+2 (3.28E+1)	+	2.39E+2 (7.93E+0)	+	**1.40E+2** (7.43E+0)
f_{25}	2.33E+2 (1.10E+1)	+	2.17E+2 (1.96E+1)	+	2.19E+2 (1.13E+1)	+	2.29E+2 (3.73E+1)	+	**2.11E+2** (2.46E+0)
f_{26}	1.03E+2 (5.83E−2)	+	1.20E+2 (4.01E+1)	+	1.41E+2 (4.90E+1)	+	**1.00E+2** (9.34E−2)	≈	**1.00E+2** (4.73E−2)
f_{27}	**5.47E+2** (2.77E+2)	≈	1.35E+3 (9.83E+1)	+	1.25E+3 (9.83E+1)	+	1.06E+3 (3.40E+2)	+	6.50E+2 (3.09E+2)
f_{28}	2.13E+3 (7.81E+2)	+	2.52E+3 (6.41E+2)	+	3.78E+3 (3.81E+2)	+	2.50E+3 (1.12E+2)	+	**1.27E+3** (1.49E+2)
f_{29}	1.79E+6 (5.14E+6)	+	6.71E+6 (9.27E+6)	+	1.87E+6 (3.63E+6)	+	4.06E+6 (5.06E+6)	+	**1.78E+4** (8.01E+3)
f_{30}	**4.47E+3** (6.03E+2)	−	1.11E+4 (4.67E+3)	≈	1.19E+4 (8.17E+3)	+	1.48E+4 (1.73E+3)	+	9.20E+3 (3.78E+3)
+	21		23		25		22		
−	4		3		1		5		
≈	5		4		4		3		

"+", "−", and "≈" denote that the performance of the corresponding algorithm is significantly worse than, better than, and significantly not different to that of IDVBA, respectively

of problems. And we can say that IDVBA in general performs better or comparable to other algorithms.

References

1. M. Airas. Echolocation in bats, in *Proceedings of Spatial Sound Perception and Reproduction. The Postgrad Seminar Course of HUT Acoustics Laboratory* (2003), pp. 1–25
2. C. Chandrasekar et al. An optimized approach of modified bat algorithm to record deduplication. Int. J. Comput. Appl. **62**(1) (2013)
3. S. Das, A. Abraham, U.K. Chakraborty, A. Konar, Differential evolution using a neighborhood-based mutation operator. IEEE Trans. Evol. Comput. **13**(3), 526–553 (2009)
4. D. Handayani, N. Nuraini, O. Tse, R. Saragih, J. Naiborhu, Convergence analysis of particle swarm optimization (pso) method on the with-in host dengue infection treatment model, in

AIP Conference Proceedings, vol. 1723 (AIP Publishing, 2016), p. 030013
5. J.H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (U Michigan Press, 1975)
6. T. Huang, A.S. Mohan, Micro-particle swarm optimizer for solving high dimensional optimization problems (μpso for high dimensional optimization problems). Appl. Math. Comput. **181**(2), 1148–1154 (2006)
7. L. Jakobsen, A. Surlykke, Vespertilionid bats control the width of their biosonar sound beam dynamically during prey pursuit. Proc. Nat. Acad. Sci. **107**(31), 13930–13935 (2010)
8. M.W.U. Kabir, N. Sakib, S.M.R. Chowdhury, M.S. Alam. A novel adaptive bat algorithm to control explorations and exploitations for continuous optimization problems. Int. J. Comput. Appl. **94**(13) (2014)
9. A. Kaveh, P. Zakian, Enhanced bat algorithm for optimal design of skeletal structures. Asian J. Civial Eng. **15**(2), 179–212 (2014)
10. J. Kennedy, C. Eberhart, iparticle swarm optimization, in *Proceedings of IEEE International Conference on Neural Network* (1995), 1942Y1948–1995
11. M. Koppen, K. Franke, R. Vicente-Garcia, Tiny gas for image processing applications. Comput. Intell. Mag. IEEE **1**(2), 17–26 (2006)
12. J.-H. Lin, C.-W. Chou, C.-H. Yang, H.-L. Tsai, A chaotic levy flight bat algorithm for parameter estimation in nonlinear dynamic biological systems. Comput. Inf. Technol. **2**(2), 56–63 (2012)
13. K.M. Passino, Biomimicry of bacterial foraging for distributed optimization and control. Control Syst. IEEE **22**(3), 52–67 (2002)
14. B. Ramesh, V.C.J. Mohan, V.V. Reddy, Application of bat algorithm for combined economic load and emission dispatch. Int. J. Electr. Eng. Telecomm. **2**(1), 1–9 (2013)
15. I. Rechenberg, Evolutionsstrategie 94, in *Werkstatt Bionik und Evolutionstechnik*, vol. 1 (Frommann-Holzboog, Stuttgart, 1994)
16. A.O. Topal, O. Altun, Dynamic virtual bats algorithm (dvba) for global numerical optimization, in *2014 International Conference on Intelligent Networking and Collaborative Systems (INCoS)* (IEEE, 2014), pp. 320–327
17. A.O. Topal, O. Altun, A novel meta-heuristic algorithm: dynamic virtual bats algorithm. Inf. Sci. **354**, 222–235 (2016)
18. A.O. Topal, O. Altun, E. Terolli, Dynamic virtual bats algorithm (dvba) for minimization of supply chain cost with embedded risk, in *Proceedings of the 2014 European Modelling Symposium* (IEEE Computer Society, 2014), pp. 58–64
19. A.O. Topal, Y.E. Yildiz, M. Ozkul, Improved dynamic virtual bats algorithm for global numerical optimization, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, USA, 2017), pp. 462–467
20. F. Van den Bergh, A.P. Engelbrecht, A cooperative approach to particle swarm optimization. IEEE Trans. Evolut. Comput. **8**(3), 225–239 (2004)
21. X. Wang, W. Wang, Y. Wang, An adaptive bat algorithm, in *Intelligent Computing Theories and Technology* (Springer, 2013), pp. 216–223
22. X.-S. Yang, A new metaheuristic bat-inspired algorithm, in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)* (Springer, 2010), pp. 65–74
23. X.-S. Yang, A.H. Gandomi, Bat algorithm: a novel approach for global engineering optimization. Eng. Comput. **29**(5), 464–483 (2012)
24. Y.E. Yldz, O. Altun, Hybrid achievement oriented computational chemotaxis in bacterial foraging optimization: a comparative study on numerical benchmark. Soft Comput. 1–17 (2015)
25. M. Yuanbin, Z. Xinquan, X. Shujian, Local memory search bat algorithm for grey economic dynamic system. TELKOMNIKA Ind. J. Electr. Eng. **11**(9), 4925–4934 (2013)

Chapter 10
Clustering for Binary Featured Datasets

Peter Taraba

Abstract Clustering is one of the most important concepts for unsupervised learning in machine learning. While there are numerous clustering algorithms already, many, including the popular one—k-means algorithm, require the number of clusters to be specified in advance, a huge drawback. Some studies use the silhouette coefficient to determine the optimal number of clusters. In this study, we introduce a novel algorithm called Powered Outer Probabilistic Clustering, show how it works through back-propagation (starting with many clusters and ending with an optimal number of clusters), and show that the algorithm converges to the expected (optimal) number of clusters on theoretical examples.

Keywords Binary valued features · Clustering · Emails · k-Means · Optimal number of clusters · Probabilities

10.1 Introduction

For over half a century, clustering algorithms have generated massive research interest due to the number of problems they can cover and solve. As an example [1], the authors use clustering to group planets, animals, etc. The main reason for the popularity of clustering algorithms is that they belong to unsupervised learning and hence do not require manual data labeling, in contrast to supervised learning, which can require the cooperation of many people who often disagree on labels. As an example one can even mention Pluto, the celestial body no longer considered a planet as of 2006. The advantage of using unsupervised over supervised learning is that rather than relying on human understanding and labeling, clustering algorithms rely purely on the objective properties of the entities in question. A good clustering algorithm survey can be found in [2].

The largest obstacle for clustering algorithms is finding the optimal number of clusters. Some results on this topic can be found in [3], where the authors com-

P. Taraba (✉)
Berkeley, CA 94710, USA
e-mail: taraba.peter@mail.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_10

pare several algorithms on two to five distinct, non-overlapping clusters. As another example, one can mention [4], where the authors use a silhouette coefficient (c.f. [5]) to determine the optimal number of clusters.

In this study, we construct three theoretical datasets: one with clusters that are clearly defined, a second with clusters that have randomly generated mistakes and last but not least clusters which have majority of features noisy. We then show that the algorithm introduced in this chapter converges to the expected number of clusters for all three examples.

This chapter is organized as follows. In Sect. 10.2, we describe the Powered Outer Probabilistic Clustering (POPC) algorithm. In Sect. 10.3, we confirm on three theoretical examples that the algorithm converges to the expected number of clusters. In Sect. 10.4, we present an additional example of real email dataset clustering. In Sect. 10.5, we dive deeper into the algorithm properties. And finally in Sect. 10.6, we draw conclusions.

10.2 Algorithm Description

The POPC algorithm, originally introduced in [6], relies on computing discounted probabilities of different features belonging to different clusters. In this section, we use only features that have binary values 0 and 1, which means the feature is either active or not. One can write the probability of feature f_i belonging to cluster k as

$$p(cl(f_i) = k) = \frac{c(s_j(f_i) = 1, cl(s_j) = k)C_m + 1}{c(f_i = 1)C_m + N},$$

where c is the count function, cl is the clustering classification, N is the the number of clusters, $k \in \{1, \ldots, N\}$ is the cluster number, s_j are samples in the dataset, $s_j(f_i) \in \{0, 1\}$ is the value of feature f_i for sample s_j, and C_m is a multiplying constant (we use $C_m = 1000$ in this study). If we sum over all clusters for one feature, we get:

$$\sum_{k=1}^{N} p(cl(f_i) = k) = 1,$$

and subsequently further over all features f_i we get:

$$\sum_{i=1}^{F} \sum_{k=1}^{N} p(cl(f_i) = k) = F. \tag{10.1}$$

If we used (10.1) as the evaluation function, we would not be able to optimize anything, because the function's value is constant no matter how we cluster our samples s_j. Hence, instead of summing over probabilities, our evaluation function J uses higher powers of feature probabilities as follows:

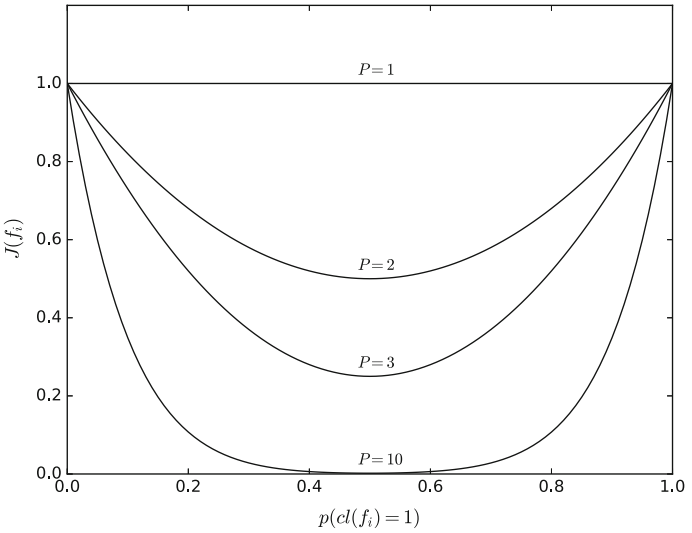


Fig. 10.1 Values of the evaluation function for a hypothetical case where there is one feature and two clusters. If the feature is distributed between two different clusters, the evaluation function J has a lower value (near 0.5 on the x-axis). In case the feature belongs to only one cluster, or mostly to one cluster, J has a higher value for evaluation function (left near 0.0 or right near 1.0 on the x-axis—values close to optimum $J(f_i) = 1$)

$$J = \sum_{i=1}^{F} J(f_i) = \sum_{i=1}^{F} \sum_{k=1}^{N} p^P(cl(f_i) = k) \leq F, \tag{10.2}$$

where P is the chosen power and we choose $P > 1$ in order to have a non-constant evaluation function. The main reason why this is desired can be explained with reference to a hypothetical case where there is only one feature and two clusters. We want samples with $s_j(f_1) = 1$ to belong only to one of two clusters. In the case that samples are perfectly separated into the two clusters on the basis of the one feature, we maximize $\sum_{k=1}^{N} p^P(cl(f_i) = k)$ (value very close to 1 as we use discounting). In the case that some samples belong to one cluster and some to the other, so the separation is imperfect, we get a lower evaluation score (value significantly lower than 1) for the feature. The higher the power P, the lower the score we obtain when one feature is active in multiple clusters. This is displayed in Fig. 10.1 for different powers $P \in \{1, 2, 3, 10\}$ and two clusters. For results reported in this section we use $P = 10$.

The algorithm starts by using the k-means algorithm to assign each sample s_j a cluster number. The number of clusters is set to half the number of samples in the dataset. Then the algorithm proceeds to reshuffle samples s_j into different clusters $\{1, \ldots, N\}$. If the evaluation function J increases as a result of reshuffling, the new cluster assignment is preserved. The algorithm ends when reshuffling no longer increases the evaluation function J.

The algorithm can be summarized in the following steps:

1. Using k-means clustering, assign each sample s_j cluster $cl(s_j) = k$, where $k \in \{1, \ldots, N\}$ and N—the number of clusters—is chosen to be half the number of data samples.
2. Compute $J_{r=0}$ for the initial clustering, where r denotes the iteration of the algorithm.
3. Increase r to $r := r + 1$, start with $J_r = J_{r-1}$.
4. For each sample s_j, try to assign it into all the clusters to which it does not belong to $(cl(s_j) \neq k)$
   a. If the temporary evaluation score J_T with the temporarily assigned cluster improves over the previous value, i.e. $J_T > J_r$, then assign $J_r := J_T$ and move sample s_j to the new cluster.
5. If J_r is equal to J_{r-1}, then stop the algorithm. Otherwise go back to step 3.

Remark 1 (Implementation detail) The algorithm can be made faster if one saves the counts of different features for different clusters and updates these counts only if and when the evaluation score improves. This is just an implementation detail, which speeds up computations significantly and does not influence the result.

10.3 Examples

In this section, we create three theoretical examples, one perfect, one slightly imperfect and one with a majority of noisy features.

For the perfect example, we create a dataset containing 200 samples with 100 features and assign each sample to a random cluster ($N = 7$) with a uniform discrete distribution, so that every cluster has approximately the same number of samples. In the same way, we assign every feature to a cluster. A feature is active (its value is 1) only if the feature belongs to the same cluster as the sample and only 80% of the time.

In Fig. 10.2 top, we show the k-means evaluation function depending on the number of clusters. There is a break-point at exactly $N = 7$, the expected number of clusters. Figure 10.2 bottom shows the number of clusters created by the POPC algorithm. We can see that if we start with a number of clusters larger or equal to 7, we always end with the expected number of clusters 7.

The first example clustered with POPC algorithm is displayed in Fig. 10.3. As shown, clusters are found as expected.

For theoretical example 2, we create also noisy features, which do not belong to any cluster. These features can be active for samples belonging to any cluster. They are active 20% of time. The results for example 2 are displayed in Fig. 10.4. The POPC algorithm introduced in this chapter once again yields the expected number of clusters.

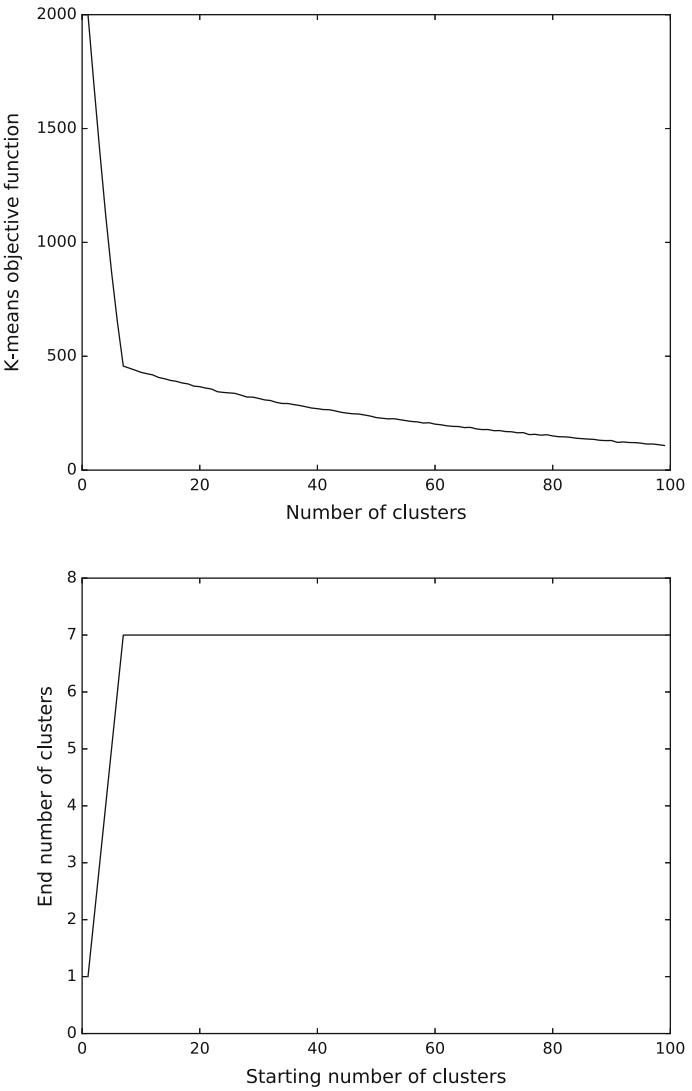


Fig. 10.2 (Top) K-means evaluation function depending on the number of clusters for example 1. (Bottom) Number of clusters created by the POPC algorithm depending on the number of starting clusters

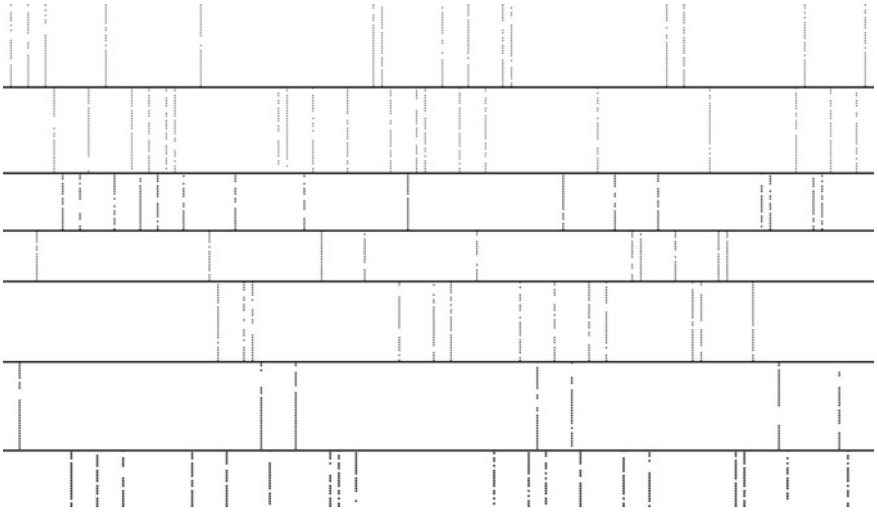


Fig. 10.3 Theoretical example 1—seven perfect clusters clustered by POPC. On vertical axis we have different samples belonging to different clusters (separated by lines) and on horizontal axis we have different features

The second example clustered with POPC algorithm is displayed in Fig. 10.5. As shown, clusters are found as expected despite having a small amount of imperfect features.

For theoretical example 3, we create 7 clusters with every cluster having 30 samples with 20 features. Every cluster has exactly one feature, which is active only for the cluster it belongs to. The remaining 13 features are random features, which do not belong to any cluster and are activated 50% of time. This is the main example, which shows why the POPC algorithm is so useful. K-means algorithm, due to a majority of features being randomly active, cannot find features which are significant features for clusters and finds clusters which are not the way we would expect them. Evaluation function $J_{k-means}$ is very close to 0 even if we knew we are looking for 7 clusters. This is displayed in Fig. 10.6.

On the other hand, when we use POPC algorithm we find exactly 7 clusters we expected and more importantly J_{POPC} is close to 7, which is the amount of features that are significant to respective clusters. This is displayed in Fig. 10.7. This last theoretical example explains why in real life situations, POPC is not only able to find the correct number of clusters, but also to find significantly better clusters than k-means for binary feature datasets.

All the theoretical examples introduced in this section can be generated with C# code which can be downloaded from [7]. Same code contains implementation of popc algorithm.

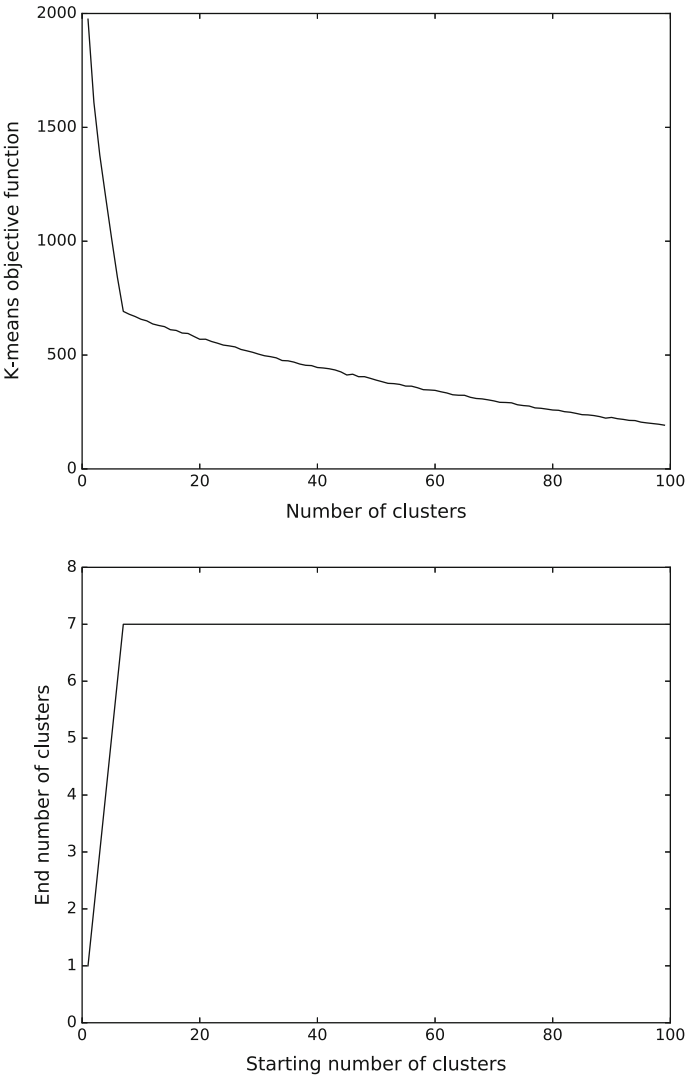


Fig. 10.4 (Top) K-means evaluation function depending on the number of clusters for example 2. (Bottom) Number of clusters created by the POPC algorithm depending on the number of starting clusters

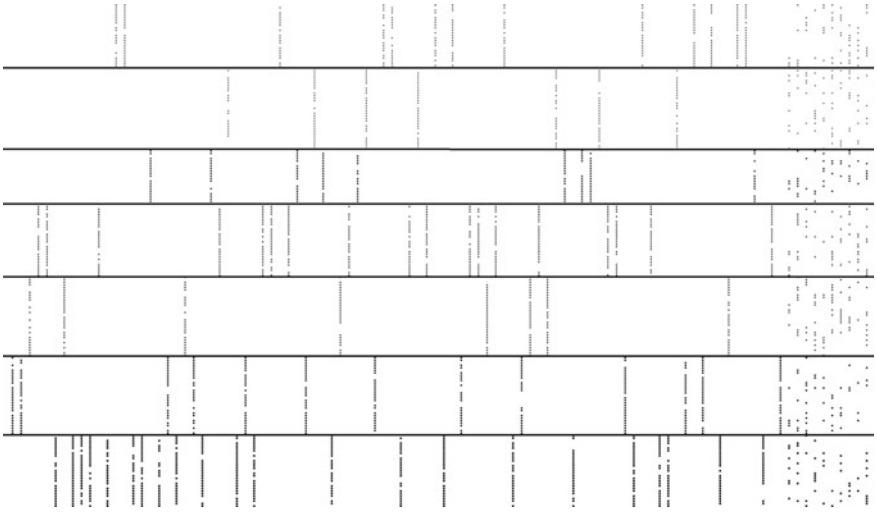


Fig. 10.5 Theoretical example 2—seven imperfect clusters clustered by POPC. On vertical axis we have different samples belonging to different clusters (separated by lines) and on horizontal axis we have different features. Last 10% features (on right) are the features, which do not belong to any cluster and are active 20% of time

10.4 Real Life Example—Email Dataset

As a last example, we will use an example from real life: email data. Each email is one sample in the dataset. The features are the people included on the email. Results are displayed in Fig. 10.8. As shown in the graph on the top, there is no clear break-point as in the previous two examples, showing why it is so hard in real life situations to find the optimal number of clusters. Despite this fact, when we start with the number of clusters larger than or equal to 93, the algorithm introduced in this chapter settles on a final clustering with 26 clusters.

The email dataset contains 270 emails (samples) and 66 people (features). The final score for the evaluation function with 26 clusters is $J = 52.19$ out of the maximum possible value 66. Even if we knew the correct number of clusters (which we do not) and used it with k-means, the evaluation function introduced in this chapter would be $J_{k-means} = 32.91$, which is significantly lower than 52.19. This represents an improvement in cluster quality of 29.21%. The score reflects approximately the number of people who belong only to a single cluster. Email clusters for the email dataset are displayed in Fig. 10.9. Improvement over k-means was confirmed on emails of two other email datasets belonging to other people. Due to privacy issues, the email dataset is not shared as it belongs to a private company.

Remark 2 (Interesting detail) While the maximum evaluation function value can be achieved by assigning all samples to the same single cluster, when starting with a large number of clusters, the algorithm does not converge even for real life datasets

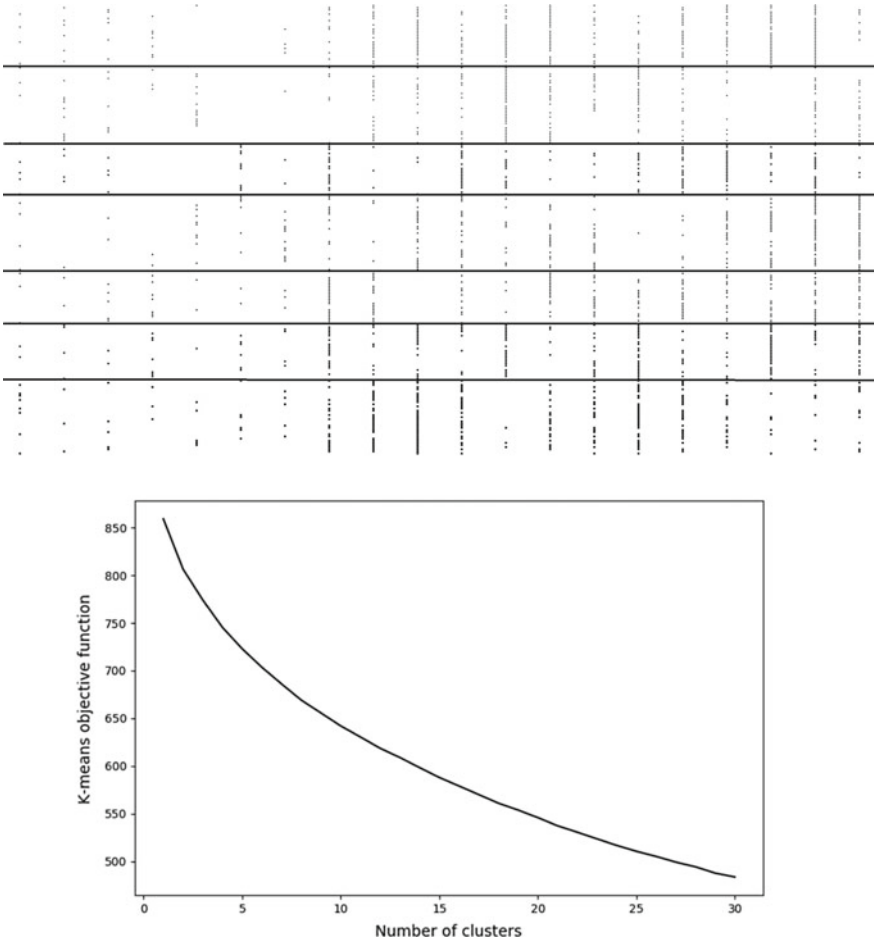


Fig. 10.6 Theoretical example 3—seven imperfect clusters clustered by k-means. (Top) On vertical axis we have different samples belonging to different clusters (separated by lines) and on horizontal axis we have different features. (Bottom) K-means evaluation function based on number of clusters—no clear break of evaluation function as for two previous theoretical examples

to a single cluster due to the use of back-propagation and the presence of local maximums along the way which result from reshuffling only a single sample at a time. This provides unexpected, but desired functionality. For the same reason, we are not able to start the algorithm with one cluster and subsequently subdivide the clusters, because this would only lower the evaluation function's value. Hence, the algorithm works only backwards.

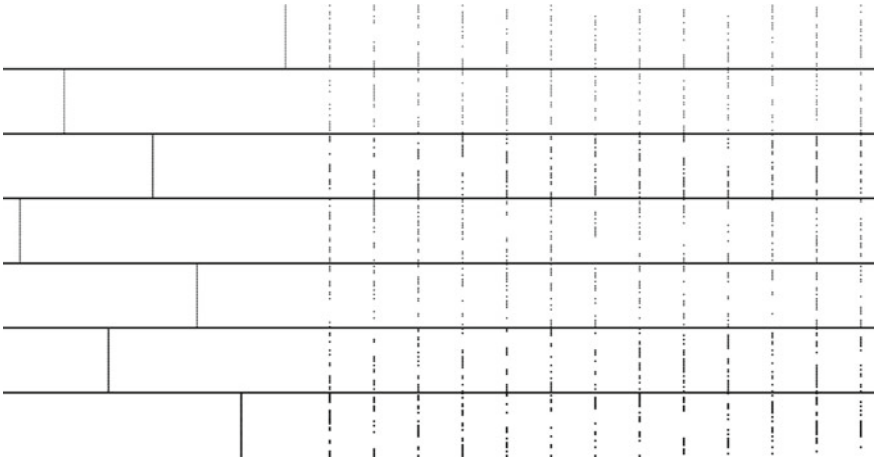


Fig. 10.7 Theoretical example 3—seven imperfect clusters clustered by POPC. On vertical axis we have different samples belonging to different clusters (separated by lines) and on horizontal axis we have different features. First 7 features (from left) are the significant features of respective clusters. Remaining 13 features are non-significant features, which do not belong to any cluster and are active 50% of time

10.5 Deeper Dive into the Algorithm

In this section we dive deeper into properties of the algorithm introduced in this paper (but for simplicity, we omit multiplying constant C_m and $\frac{\cdots+1}{\cdots+N}$ technique which is used to achieve non-zero probabilities).

As a first example, we consider the simple example of two clearly separated clusters, which features do not intersect as it is displayed in Table 10.1. Evaluation of the function is the same whether we have one or two clusters:

$$J_{2clusters} = n\left(\left(\frac{k}{k}\right)^P + \left(\frac{0}{k}\right)^P\right) + m\left(\left(\frac{l}{l}\right)^P + \left(\frac{0}{l}\right)^P\right) = n + m = J_{1cluster}$$

But if we start with two clusters, the evaluation function will not allow joining these two clusters, as we move only one sample at a time and the evaluation function would not improve for different rounds r of algorithm ($J_r > J_{r+1}$). Consider moving sample from cluster 2 to cluster 1:

$$J_r = n + m > n + m\left(\left(\frac{l-1}{l}\right)^P + \left(\frac{1}{l}\right)^P\right) = J_{r+1}$$

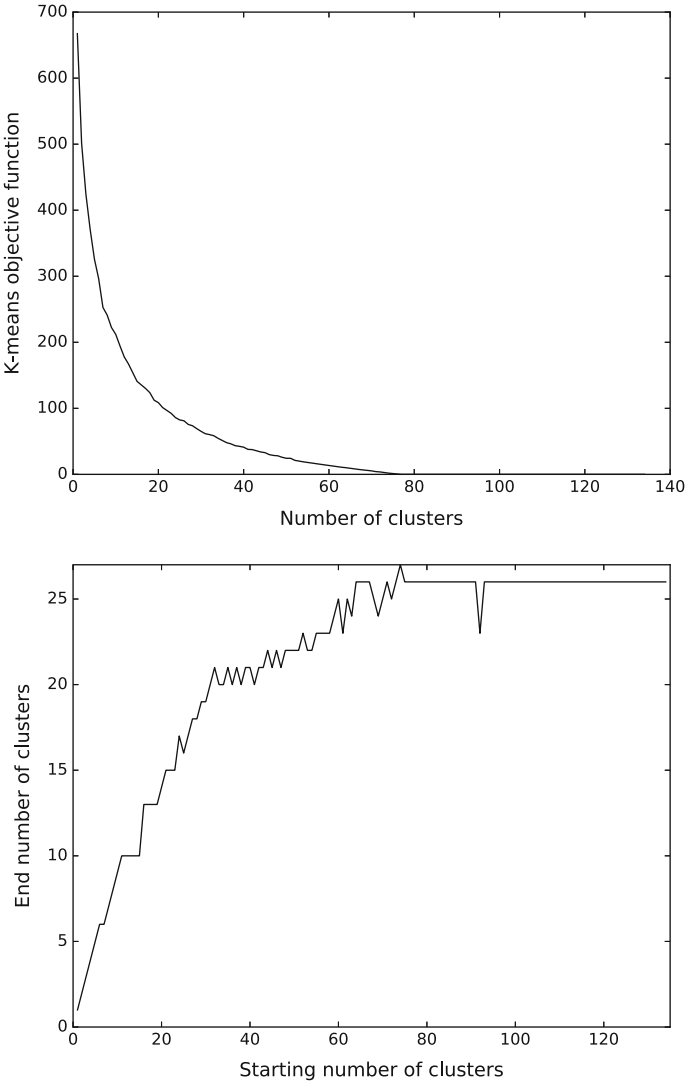


Fig. 10.8 (Top) K-means evaluation function depending on the number of clusters for real life example (emails). (Bottom) Number of clusters created by the POPC algorithm depending on the number of starting clusters

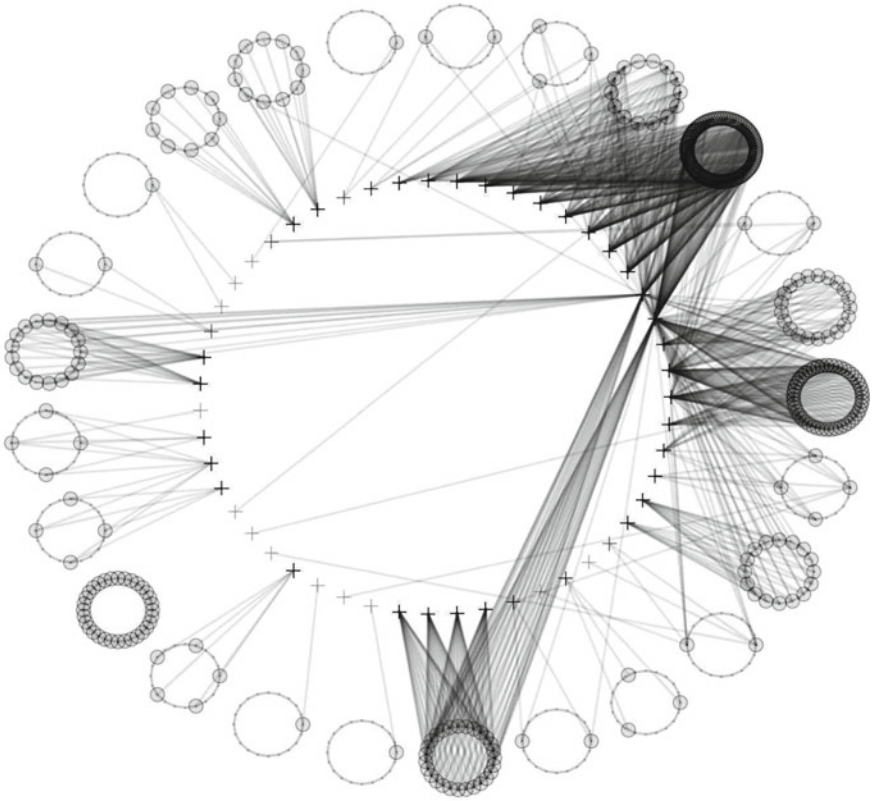


Fig. 10.9 Email clustering example. Emails are displayed as 'o' in different cluster circles. People are displayed as '+' and connections between emails and people included on these emails are displayed as lines between them

for $P > 1$, hence sample from cluster 2 will not move to cluster 1 and it could be shown the same way for sample from cluster 1 to cluster 2. This is the reason these two clusters will not be joined for $k > 1$ and $l > 1$.

Now, we explore a slightly more complicated situation when we have two clusters (Email Group 1 and 2). In one cluster, we have communication with all people (features 1 to $n + m$), and in the other email group we communicate only with a subgroup of people in the first email group (features $n + 1$ to $n + m$). This is shown in Table 10.2.

Now if we ask if these two clusters should be together or not, without optimizing our evaluation function by moving only one sample at a time, the answer is they should be together as

$$J_{2clusters} = n + m \left(\left(\frac{k}{k+l} \right)^P + \left(\frac{l}{k+l} \right)^P \right) < n + m = J_{1cluster}$$

Table 10.1 Two groups of emails with $k + l$ samples and $n + m$ people (features)—completely separated

Sample number	Email group	F_1	...	F_n	F_{n+1}	...	F_{n+m}
1	1	1	...	1	0	...	0
...	1	1	...	1	0	...	0
k	1	1	...	1	0	...	0
k + 1	2	0	...	0	1	...	1
...	2	0	...	0	1	...	1
k + l	2	0	...	0	1	...	1

Table 10.2 Two groups of emails with $k + l$ samples and $n + m$ people (features)—intersected features

Sample number	Email group	F_1	...	F_n	F_{n+1}	...	F_{n+m}
1	1	1	...	1	1	...	1
...	1	1	...	1	1	...	1
k	1	1	...	1	1	...	1
k + 1	2	0	...	0	1	...	1
...	2	0	...	0	1	...	1
k + l	2	0	...	0	1	...	1

for $P > 1$, which we consider.

So the question to answer is why, when using back-propagation, even though $J_{1cluster} > J_{2clusters}$ for the situation in Table 10.2, we do not end up with one cluster. Consider we have two clusters, and we move one sample from group two to group one only if the evaluation function is increased:

$$\begin{aligned} J_r &= n + m\left(\left(\frac{k}{k+l}\right)^P + \left(\frac{l}{k+l}\right)^P\right) \\ &< n + m\left(\left(\frac{k+1}{k+l}\right)^P + \left(\frac{l-1}{k+l}\right)^P\right) = J_{r+1} \end{aligned}$$

which holds if

$$k^P + l^P < (k+1)^P + (l-1)^P$$

and for case when $P = 2$ it is if

$$l < k + 1.$$

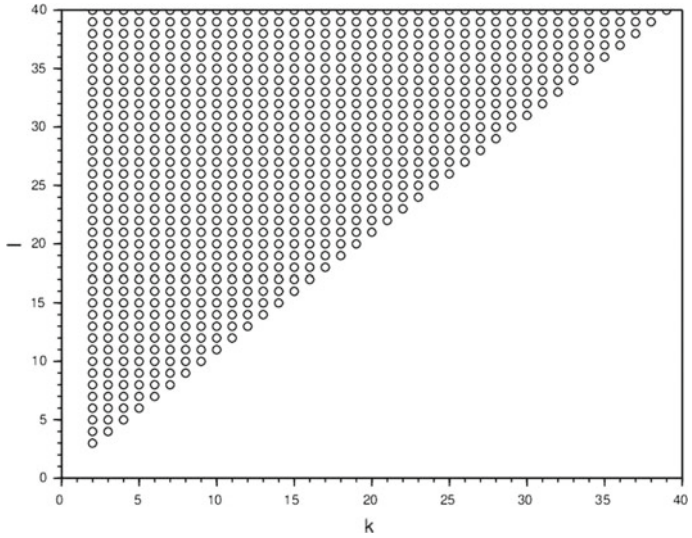


Fig. 10.10 The condition $k^P + l^P < (k+1)^P + (l-1)^P$ is the same for $P = 2$ and $P = 10$ for $k \in \{2, ..., 40\}$ and $l \in \{2, ..., 40\}$. Circle indicates condition is not fulfilled, while empty space opposite.

This means if there are enough samples l of the second group, at least $k + 1$, the member of group two will not move to group one and the clustering algorithm ends up with two clusters. Even for $P = 10$, the condition seems to be unchanged and is displayed in Fig. 10.10.

Another question for this example is why does the sample from group one not move to group 2. This can be explained by

$$\begin{aligned} J_r &= n + m\left(\left(\frac{k}{k+l}\right)^P + \left(\frac{l}{k+l}\right)^P\right) \\ &> n\left(\left(\frac{k-1}{k}\right)^P + \left(\frac{1}{k}\right)^P\right) + m\left(\left(\frac{k-1}{k+l}\right)^P + \left(\frac{l+1}{k+l}\right)^P\right) = J_{r+1} \end{aligned}$$

Even this simple situation can get complicated, but for $n \gg m$, we can simplify the equation to

$$\frac{J_r}{n} \approx 1 > \left(\left(\frac{k-1}{k}\right)^P + \left(\frac{1}{k}\right)^P\right) \approx \frac{J_{r+1}}{n}$$

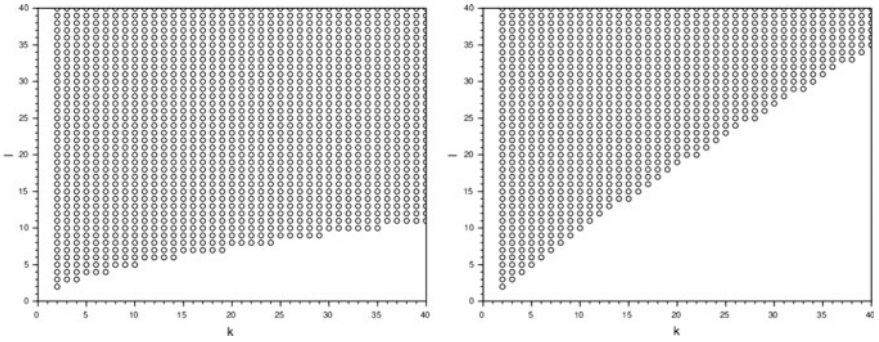


Fig. 10.11 The condition $J_r = n + m\left(\left(\frac{k}{k+l}\right)^P + \left(\frac{l}{k+l}\right)^P\right) > n\left(\left(\frac{k-1}{k}\right)^P + \left(\frac{1}{k}\right)^P\right) + m\left(\left(\frac{k-1}{k+l}\right)^P + \left(\frac{l+1}{k+l}\right)^P\right) = J_{r+1}$ is not the same for $P = 2$ (left) and $P = 10$ (right) for $k \in \{2, ..., 40\}$ and $l \in \{2, ..., 40\}$. For this example, we chose $n = m = 10$. Circle indicates condition is not fulfilled, while empty space opposite

which is true for $P > 1$ and $k > 1$. So if n is significantly larger than m, we have no reason to move the sample from group 1 to group 2. When this condition is fulfilled versus not is displayed in Fig. 10.11 also for case $n = m = 10$.

This is when parameter P starts to play its role unlike before (Fig. 10.10) and whether two clusters are joined or not depends on P.

10.6 Conclusions

We introduced a novel clustering algorithm POPC, which uses powered outer probabilities and works backwards from a large number of clusters to the optimal number of clusters. On three theoretical examples, we show that POPC converges to the expected number of clusters. In a real life example with email data, we show that it would be difficult to determine the optimal number of clusters based on the k-means evaluation score, but when the algorithm introduced in this chapter is used, it settles on the same number of clusters as if we had started with a large enough initial number of clusters. Importantly, the clusters are of higher quality in comparison with those produced by k-means even if we happened to know the correct number of clusters ex ante. Software *Small Bang*, which clusters emails and uses POPC algorithm, can be downloaded from [8].

Acknowledgements The authors would like to thank David James Brunner for many fruitful discussions on knowledge workers information overload as well as proofreading of the first draft. The authors would also like to thank anonymous reviewers for providing feedback which led to significant improvement of this chapter.

References

1. J. Hartigan, *Clustering Algorithms* (Wiley, 1975)
2. R. Xu, D. Wunsch, Survey of clustering algorithms. IEEE Trans. Neural Netw. **16** (2005)
3. G. Milligan, M. Cooper, An examination of procedures for determining the number of clusters in a data set. Psychometrika **50** (1985)
4. G. Frahling, C. Sohler, A fast k-means implementation using coresets. Int. J. Comput. Geom. Appl. **18** (2008)
5. P. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Comput. Appl. Math. **20** (1987)
6. P. Taraba, Powered outer probabilistic clustering, in *Proceedings of the World Congress on Engineering and Computer Science 2017, 25–27 October, 2017, San Francisco, USA*. Lecture Notes in Engineering and Computer Science (2017), pp. 394–398
7. P. Taraba, Popc examples [Online] (2017), https://github.com/pepe78/POPC-examples
8. P. Taraba, Small bang [Online] (2017), http://www.frisky.world/p/small-bang.html

Chapter 11
Addressing the Challenges of Igbo Computational Morphological Studies Using Frequent Pattern-Based Induction

Olamma U. Iheanetu and Obododimma Oha

Abstract Computational studies of Igbo language are constrained by non-availability of large electronic corpora of Igbo text, a prerequisite for data-driven morphological induction. Existing unsupervised models, which are frequent-segment based, do not sufficiently address non-concatenative morphology and cascaded affixation prevalent in Igbo morphology, as well achieving affix labelling. This study devised a data-driven model that could induce non-concatenative aspects of Igbo morphology, cascaded affixation and affix labelling using frequent pattern-based induction. Ten-fold Cross Validation (TCV) test was used to validate the propositions using percentages. An average accuracy measure of 88% was returned for the developed model. Ten purposively selected Igbo first speakers also evaluated samples of 100 model-analysed words each and the mean accuracy score of 82% was recorded. We conclude that morphology induction can be realized with a modestly sized corpus, demonstrating that electronic corpora scarcity does not constrain computational morphology studies as it would other higher levels of linguistic analysis.

Keywords Computational morphology · Frequent pattern-based morphology
Igbo computational morphology · Igbo morphology · Rule-based learning
Morphology induction

11.1 Introduction

Statistical morphological analysis of natural languages is a more viable option than traditional rule-based approach, especially for resource-scarce languages [2] due to the fact that these set of languages are less studied or under-studied. Statistical approaches to learning are largely data-driven. A very favoured statistical learning

O. U. Iheanetu (✉)
Department of Computer and Information Sciences, Covenant University, Ota, Nigeria
e-mail: olamma.iheanetu@covenantuniversity.edu.ng

O. Oha
Faculty of Arts, Department of English, University of Ibadan, Ibadan, Nigeria

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_11

method is the Unsupervised Learning Method (ULM). ULMs require large volumes of data to train the learner in other to make accurate classifications of unseen objects. Unfortunately, large volumes of electronic data are a scarce commodity for resource-scarce languages such as Igbo, and thus a major set-back in adopting data-driven approaches for computational studies [15].

The underlying fact is that most of the existing models of morphology induction are designed based on inherent behaviour of language which are arbitrary occurrence of word segments and frequent occurrence of word segments. This results in somewhat universal models although some uncommon peculiarities of some languages can render such models inappropriate/ insufficient for those languages. However, [5] argued that existing ULMs are not sufficient for Bantu languages. Igbo may not be a Bantu language but [3] believes that it belongs to the same language phylum as Bantu languages, causing it to share some similarities with them. Creutz [4] are of the opinion that ULMs are unsuitable for languages with sparse linguistics data in computer readable form.

Although some of the existing ULMs correctly analyse inflected Igbo words, results show that these models cannot sufficiently breakdown words produced from reduplication, circumfixation, compounding and interfixation morphological processes in Igbo. A further complication results from the highly agglutinative nature of Igbo, according the language the capability of having as high as four cascades of affixes, thus posing a challenge for the analyses of all affixes [15]. These situations and circumstances frustrate the morphological induction of Igbo and other resource-scare languages, which are mostly African languages. Proffered methods, as suggested by [16], should take into cognizance the challenge of scarcity as well as other peculiar characteristics of the Igbo morphology.

The underlying language behaviour for inflection is frequent word segments prevalent in a given corpus, hence most existing ULMs are frequent segment-based. As established in [16], frequent pattern-based ULMs could work better for most derived words, especially if they share some similarities with Igbo.

11.2 Previous Literature

Most unsupervised morphology models carter for languages having orthographic inconsistencies which necessitate adjustments in the heuristics used. Such adjustments may not be necessary for Igbo because of the high level of regularity in its orthography [15]. A consensus in ULM is that affix classification is still elusive [13]. Arbitrary Character Assumption (ACA) and Frequent Flyer Assumption (FFA) was postulated by [12] for the extraction of morpheme boundaries, on the premise that words are normally arbitrarily occurring segments. The frequently occurring segments which are equi-length with the segments that occur arbitrarily have a high probability of being affixes. Goldsmith [11] identified morpheme boundaries using the principle of Minimum Description Length (MDL) which describes the morphology that offers the least minimum description length. Harris [14] is a foundational

study in unsupervised morphology induction that bootstraps the problem of identifying morpheme boundaries using a heuristic approach; *successor frequency*. These ULMs are only representative of the general approaches to the unsupervised learning of morphology which are mostly frequent-segment based.

A brief background on Igbo language shows that Igbo is a tonal language spoken in South-east part of Nigeria [16] having an approximate thirty dialects [24, 26] but Standard Igbo is widely spoken and understood, and as well generally accepted among Igbo speakers. Igbo language is a member of the West Benue-Congo languages [3], formerly classified under the Niger-Congo Kwa language family; a language family characterized by two tones (high and low) and a down step that apply different meanings to the same set of phones [9]. Igbo exhibits a rich agglutinative morphology [23, 26] and Igbo verbs also inflect for aspect [26]. Igbo features a wide variety of highly productive concatenative and non-concatenative morphological processes. Owing to the agglutinative nature of Igbo morphology, cascaded affixation, a highly productive morphological process, becomes a common occurrence.

Most Igbo morphological processes can be generalized as affixation. An affix can come before the root word (prefix), or in between (infix or interfix, depending on where it splits the root word) or after the root word (suffix). Ndimele [22] noted that the position and function of an affix when it is attached to a root is definitive of the category of that affix. Hence there exist prefixes, suffixes, infixes, interfixes, circumfixes, superfixes and suprafixes in the positional classification of affixes. For Igbo language, prefixation, suffixation, interfixation [7, 20], superfixation and circumfixation apply [20]. Non-concatenative morphological operations involve root modification [25] truncation, subtractions, conversion and transfixation, which according [18] cannot be resolved by recurrent partials. Although most Igbo morphological processes are concatenative, some manifest majorly as non-concatenative morphology. These include interfixation, circumfixation and compounding.

11.3 The Problem

The cost of rule-based morphology models is very high in terms of data annotation cost, labour and the time involved. In addition, RBMs are subject to human error. The demands of UML approach are unattainable for resource-scarce languages like Igbo. ULMs require many examples to learn from. These examples are offered by the availability of very large amounts of electronic data or corpora. Igbo at present has very sparse linguistic data in computer readable form available for computational studies.

Simple affixations in Igbo may be discovered by known approaches to unsupervised learning; however, these approaches may not work well for multiple affixations, which is prevalent in some Igbo morphological processes. While experimenting with *Linguistica*, it was observed that *Linguistica* could not accurately analyse derived words like compounds and words having multiple extensional

suffixes. This was quite explainable judging that *Linguistica* was tested on English and French corpora; languages that rarely exhibit multiple suffixation.

Non-concatenative processes like reduplication and compounding pose a challenge for existing unsupervised learning models. This is because the affixes of words realized from these morphological processes may not occur as frequently in the corpus as the relevant threshold may demand. It is possible therefore that some relevant segments may record lone occurrence in the corpus and therefore may not be properly identified as valid morphemes. It is pertinent therefore to devise new methods that can cater for Igbo words that feature compounding and reduplication.

Igbo, Dghweɗe [6] and Yoruba languages [1] exhibit interfixation; a rare morphological behaviour. Interfixes may not feature as frequently occurring segments and therefore may not be captured by frequency-based morpheme boundary identification methods. Typical Igbo interfixes are merely Igbo phonemes represented in writing as *l, m*, *r*, and so on, stand between two copies of a stem. Some morphological processes of interest in this study include the following

Extensional suffixation	Where *wa, zie, nụ and kwa* are extensional suffixes that extend the meaning of the verb. See examples below
kpachapụkwaranụ	kpa-cha-pụ-kwa-ra-nụ
laghachikwaara	la-ghachi-kwa-a-ra

Reduplication—full reduplication is achieved by duplicating the stem word, although some scholars have argued that because the words do not undergo any morphological processing, it is mere repetition. Another way of realizing gerunds in Igbo is by reduplicating verb stems having the—CV structure and prefixing a harmonizing ọ or ò vowel and appropriate reduplicating vowel—i, i, u, or ụ. This is known as partial reduplication

ọ + lụ (marry) + lụ → ọlụlụ (marriage)
ò + zù (steal) + zù → òzùzù (act of stealing)

Partially reduplicated word is described as a word having a prefix and two identical roots, or a word having a prefix and two roots that differ only in the vowel of one of the roots, only if such differing vowel is a member of the list of valid reduplicating vowels—*i*, ị, *u*, and *ụ*. $v \in \{i, ị, u, ụ\}$.

Given an Igbo root verb $R = r_1, r_2, \ldots, r_n$, a partially reduplicated word $w = pr_1vR$, such that $p \in \{o, ọ\}$, $\{i, ị, u, ụ\} \in v$ and $R \in VocR$ where $VocR$ is a set of Igbo verb roots.

ọ-gb-ụ-gbọ	reduplicating vowel = ụ
o-gb-u-gbu	reduplicating vowel = u
o-l-i-lo	reduplicating vowel = i
ọ-g-ị-ga	reduplicating vowel = ị

Interfixation—This is a word that has an affix (interfix) which splits the word into two identical segments.

Given a vocabulary Voc and words $w \in Voc$, if w can be broken into segments such that $w = w_1, s_1, w_1$ there is a probability $P_{infx}(w)$ that s_1 is an interfix between

two segments of w_1 if $w_1 = w_1$ and $s_1 \in \{da, gh, l, m, r\}$. Where $P_{infx}(w)$ = Probability that w is an interfixed word [15]. In a way, interfixation can be viewed as full reduplication having an interfix between the words, if the interfix belongs to the list of true interfixes comprising *l*, *m*, *r*, *da* or *gh*.

ede-m-ede interfix = m
ogo-l-ogo interfix = l
egwu-r-egwu interfix = r

Compounding—Compounds are gotten from the coexistence of two stems, base forms or morphemes as a single word. In general terms, a compound comprises two or more stems that together to form a word.

obi + ọma → obiọma (good heart)
chi + ma → chima (God knows).

Given a vocabulary *Voc* and words $w \in Voc$, if w can be broken into segments such that $w = w_1, w_2, \ldots, w_n$, there is a probability $P_{(cpd)}(w)$ if $w_1, w_2, \ldots, w_n$ and $w_1 \neq w_2 \neq w_3, \ldots \neq w_n \in Voc$. Where $P_{cpd}(w)$ = Probability that w is a compound word. [15].

Circumfixation—The phenomenon of Igbo circumfixation morphological process is still very controversial and hence not well studied. It involves two disparate affixes one at each end of stem [21]. The vowel-syllabic nasal circumfix and the vowel-incorporated preposition circumfix exist [21]. However, no rule(s) that governs the choice of the initial vowel or syllabic nasal was presented by [21]. Given a vocabulary *Voc* and words $w \in Voc$, if w can be broken into segments such that $w = w_1, s_1, w_2$ there is a probability $P_{cfx}(w)$ if $w_1, w_2 \in Voc$ and $s_1 \in \{m, n\}$. Where $P_{cfx}(w)$ = Probability that w is a circumfixed word [15].

n-dụ-m-ọdụ circumfix = n-m → (advice)
e-zu-m-ike circumfix = e-m → (rest/holidays)

The notion of resource scarcity in the scientific study of a language is yet to acquire a clear definition [15]. The term is used interchangeably with other terms such as resource-starved, under-resourced, resource scarce, less studied, least developed, under developed, under resourced, and so on. In computational linguistics, these terminologies have been used to describe languages that have insufficient or no electronic texts in written or spoken form which are readily available for computational studies in that language [15]. The concept of resource scarcity needs to be formally addressed and properly re-addressed given that linguistic resources have become incredibly valuable and data availability plays a fundamental role in NLP applications and technologies [17]. Any language that cannot boast of large amounts of appropriately diacritized electronic spoken or written texts, which is easily available for use, is a resource-scarce language. For such languages, computational studies become cumbersome because computer readable texts have to be first generated and developed. The general impression is that resource-scarce languages, like Igbo, cannot be subjected to data-driven morphological analysis, as earlier echoed by [4].

Lack of diacritized texts—Most of the available computer-readable texts in Igbo are not diacritized, and when they are, the tone marks are either incomplete, incorrect, or irregular with the standard tone-marking convention. Some claim it is easier to publish text without the necessary diacritics, while others submit that the technology to include those diacritics do not exist or are not available. This challenge presents ambiguity issues. For morphological studies, the impact of this may not be as drastic, but for syntactic, semantic and higher levels of linguistic analysis, it may be catastrophic. Homonyms are words that share the same spellings but have different meanings based on the respective tone-marks on the words. Hence, without the appropriate tone-marks, it would be almost impossible to analyse homonyms appropriately. The following nouns are known homonyms in Igbo.

akwa → (cry) [High – High]; àkwà → (bed) [Low – Low]
àkwa → (egg) [High – Low]; akwà → (cloth) [Low – High]

11.4 Proferred Solution

Data Source—Data for this study was got from *Baibụl Nsọ* (*Nhazi Katọlik*); a current Igbo version of the holy Bible imprimatured by Bishop A. K. Obiefuna in 2000, an electronic version of Baibụl Nsọ by Bible Society of Nigeria, a story book; *Jụọchi*, one edition of the now defunct *Ogene* newspaper, and two years of Odenigbo lecture transcripts. The hardcopy story book and the newspaper were typed in order to have an electronic copy while *Baibụl Nsọ* (*Nhazi Katọlik*) was scanned with Fine reader, an optical character recogniser so as to make available to the study, an electronic copy of the text. Although not all of these Igbo texts were written in standard Igbo, all texts were consistent with the Ọnwụ orthography; the standard orthography. The resulting wordlist after data cleaning and processing was a 29,191 wordlist. The words in the wordlist were converted to strings of Cs and Vs and then these *word label* strings were manually segmented according to their respective inherent morphological structures.

Frequent Pattern Theory—[16] have established that non-concatenative and cascaded affixation in Igbo language can be analysed by ULM that are frequent pattern-based as against frequent segment-based. They developed the Frequent Pattern Theory (FPT), on the premise that the segments of some derived Igbo words do not manifest as frequent segments in a wordlist. This study applied frequent pattern mining for morphological analysis of especially non-concatenative aspects of Igbo morphology. The introflected segments of words emanating from morphological processes like partial reduplication, circumfixation, and interfixation cannot be identified by any morphological analysis model that is frequent segment-based [16]. A linguistic phenomenon was adopted, where words in a wordlist are represented as a combination of Cs and Vs, having number subscripts for unique identification of the letters. C represents consonants and V represents vowels. A Python program was written to automate the transcription of our wordlist into strings of Cs and Vs. Hence, C_0, C_1, C_2, is a substitute for the first, second and third consonants of a given word.

Table 11.1 Patterns of identified Igbo morphological processes

S/no	Morphological process	Pattern
1	Unbound morpheme	Root
2	Prefixation	Prefix-Root
3	Suffixation	Root-Suffix
4	Interfixation	Root-Interfix-Root
5	Circumfixation	Prefix-Root1-Cmfx-Root2
6	Partial reduplication	Prefix-MRoot-Root
7	Full reduplication	Root-Root
8	Compounding	Root1-Root2

MRoot—Modified Root, Cmfx—Circumfix

Likewise, V_0, V_1, V_2 stands for the first, second and third vowels that occur in a given word [16].

The identified morphological processes in this study and the associated patterns embedded in the words that they produce are shown in Table 11.1.

The first test carried out tried to determine if there are embedded features in Igbo words which can be used to induce the non-concatenative aspects of Igbo morphology. A PROLOG program was used to extract words from the wordlist according to the individual patterns of the identified Igbo morphological process.

Results showed that identified Igbo morphological processes clearly manifest distinguishable patterns which are stamped on the words produced by these morphological processes. It was observed that when the pattern of a given morphological process was used to extract words from the wordlist, most of the words extracted conformed to the pattern of the appropriate morphological process being tested. However, this observation was not uniform. Compound words was the most challenging word pattern to describe because the pattern description was rather fluid, extracting verbs like *abanye*, made up of two different words.

This test gives credence to the fact that Igbo words have embedded patterns, owing to the morphological process that produced such a word, which can be deployed as a feature for inducing that word. Some of the confused words from this test are shown in Table 11.2.

In another test, a comparison of the *word label* clusters and the word patterns which words belonging to the same cluster manifest was undertaken. A maximum of twenty words each were randomly selected from each class of word labels to verify if word labels as textual proxy of patterns were productive for capturing word patterns. Standard accuracy measure based on true positives and false positives was calculated.

From Table 11.3, results showed that some word label clusters like $V_0C_0C_1V_1$, $C_0C_1V_0C_0V_1C_2V_1$, and $V_0C_0V_1C_0V_1V_2$, for PRt1, Cmfx1 and PRedp2 morphological processes respectively, did not have as many as 20 words. A high level of consistency between word labels and their structures was observed. Some word labels did

Table 11.2 Morphological processes versus corresponding extracted words

S/no	Morphological processes being tested				
	Full reduplication	Circumfixation	Interruption	Compounding	Partial reduplication
1	*ksam-ksam	**i-ds-m-miri	**kwa-do-kwa	**aba-cha	*n-gi-ge
2	*ngwa-ngwa	**n-gwa-mma	**ara-ch-ara	*aba-nti	**i-iu-tu
3	*kpom-kpom	*n-si-m-ike	*ede-m-ede	*nwa-nza	**n-ga-gide
4	*wara-wara	*a-wa-m-anya	*iri-gh-iri	*nwa-mmefu	*o-di-di
5	*bara-bara	*e-nye-m-aka	*oko-m-oko	*odo-mmiri	*o-wu-wu

*True positives, ** False positives

not capture their morphological processes well. For example, $V_0C_0V_0C_1V_0C_0V_0$ clusters interfixed words better than $C_0V_0C_1V_0C_0V_0$, judging from the accuracy scores of 75% and 2% respectively. Prefixation, compounding and suffixation morphological processes each had an accuracy of 100%, 95% and 85% respectively, demonstrating that some word labels cluster words of a particular morphological process better than other word labels.

The third test was undertaken to understand how accurately word labels reflect morphological structure when used as textual proxies of word patterns. This study proposes that the use of word labels as a textual representation of the structural patterns of Igbo words constitutes a productive feature for the induction of the non-concatenative aspects of Igbo morphology. A model of Igbo morphology was developed using a determined textual proxy for the representation of the structural patterns of Igbo words, that is the morphological structure of a word. The model was evaluated in order to verify the accuracy of the model output [15].

According to [27], the ⌊ random i.i.d for this study is 29191. Word labels automatically clustered words according to their morphological structures based on a manually classified and segmented word label table. Using this table, the model learned the morpheme boundaries based on word patterns as captured by the segmented word labels presented in the manually segmented word label table. After the training, the developed model was able to predict segmentation of unseen words based solely on the word labels, thereby demonstrating the ability to generalize. The model also suggests the morphological structure of the segmented words, and this information doubles as the affix label of that word.

The Ten-fold Cross Validation (TCV) test, implemented in Visual Basic was used to evaluate the model's accuracy at predicting morpheme boundaries of unseen words. The TCV test involves four main steps namely:

- Partitioning the study wordlist into 10 data subsets
- Training the model using nine of the ten data subsets and testing it with the remaining one data subset to determine the accuracy of the model

Table 11.3 Accuracy of the use of word labels to cluster words produced by the same morphological process

Morphological processes	Intfx1	Intfx2	PRt1	PRt2	Cmfx1	Cmfx2
Word labels	$C_0V_0C_1V_0C_0V_0$	$V_0C_0V_0C_1V_0C_0V_0$	$V_0C_0C_1V_1$	$C_0C_1V_0C_2V_1$	$C_0C_1V_0C_0V_1C_2V_1$	$C_0C_1V_0C_0V_1C_2V_2$
True positive	4	15	4	20	3	11
False positive	16	5	0	0	4	7
Accuracy	0.2	0.75	1	1	0.43	0.61

Table 11.4 Resuit of the TVC test

Tests	True positives (1)	False positives (2)	Empty cells (3)	Uncertain (4)	Total = {(1) + (2) + (3) + (4)}	Accuracy score = (1)/Total
Test 1	282	14	4	0	300	0.94
Test 2	267	22	11	0	300	0.89
Test 3	270	22	7	1	300	0.90
Test 4	267	24	9	0	300	0.89
Test 5	261	28	10	1	300	0.87
Test 6	273	19	7	1	300	0.91
Test 7	267	24	9	0	300	0.89
Test S	258	30	12	0	300	0.86
Test 9	267	24	9	0	300	0.89
Test 10	243	41	13	3	300	0.81
Mean	265.5	24.8				**0.88**

- Repeating step 2 above nine more times, using different combinations of the training and testing data subsets each time
- Computing the mean accuracy of the ten tests.

The wordlist was divided into ten subsets, giving an approximate of 2919 words in each subset. The morphological analyser model was then trained ten times with nine different data subsets and tested ten times with one data-subset, which was different for each test. The training required the model to read a word, convert it to a corresponding word label, and segment the word label based on the manual classification table. The output of the training is a series of segmented word labels. For the testing, the model reads in a word, determines an appropriate segmentation of the word based on the segmented word label which is the output of the training, and automatically segments the word based on a determined word label segmentation from the training.

The accuracy of each of the ten tests was calculated based on true positives and false positives. The average accuracy score of the ten tests was then determined and used as the overall accuracy of the analyser model. Table 11.4 shows the test output from the developed analyser model, including the morphological structure (affix label) and correctness of the segmented words.

The mean score of the TVC test gave 88% accuracy for the developed model, an indication that the developed model could segment Igbo words based on their embedded patterns.

To strengthen this claim, 10 Igbo first language speakers consisting of 3 Igbo primary school teachers, 3 Igbo linguists and 4 Igbo postgraduate students, who were purposively selected, evaluated the output of the model. The assessors manually cross-checked 100 samples each of the model output which were randomly selected. Based on the number of "Yes" from each assessor, an accuracy score was calculated. Table 11.5 captures this evaluation. More assessors would have been sought but for

Table 11.5 Model accuracy score as assessed by 10 Igbo native speakers

Assessors	Yes	No	Accuracy
1	95	5	0.95
2	76	24	0.76
3	74	26	0.74
4	74	26	0.74
5	74	26	0.74
6	94	6	0.94
7	92	8	0.92
8	89	11	0.89
9	63	37	0.63
10	89	11	0.89
Average			**0.82**

the challenge of finding native speakers who understand morphological segmentation in a non-Igbo speaking location.

Thus, the mean accuracy of the model as evaluated by 10 native Igbo speakers was 82%, further buttressing the fact that pattern segments can be used to induce non-concatenative aspects of Igbo morphology.

According to [10], *Linguistica* achieved an accuracy score of 72% on the first 200,000 and 300,000 words of the Brown corpus; Syromorph developed by [19] had an accuracy score of 90% while [8] model had 98% accuracy for Arabic words and 85.3% for English words. Based on [27], a measure of the analyser model's accuracy is implied in the loss or discrepancy measure between the response y of the supervisor and that of the learning machine using the same input as shown below.

$$l(y, (f(x, \alpha))) = \left\{ \begin{array}{l} 0 \text{ if } y = f(x, \alpha) \\ 1 \text{ if } y = f(x, \alpha) \end{array} \right\}$$

where y = response of the supervisor, x = a certain input, and $f(x, \alpha)$ = response of the learning machine.

If $f(x, \alpha) = y$, then, the output of the model is same as the linguist (supervisor) and invariably, the accuracy score of the model is quite high. If $f(x, \alpha) \neq y$, then the guess or prediction of the learner model is not very close to that of the supervisor, in this case the linguist, and may imply a low accuracy score. This is shown in Table 11.5. Therefore, every instance ("Yes") an assessor agrees with the segmentation offered by the analyser model $f(x, \alpha)$, the output of the developed analyser model equals to zero because y, which is the word segmentation as suggested by Igbo linguists, is the same as the word segmentation that the analyser model predicted. That is, $f(x, \alpha) = y$, implying that the discrepancy between $f(x, \alpha)$ and y is zero. However, for every 'no' the implication is that the response of the analyser model, $f(x, \alpha)$ equals one because y is different from the word segmentation predicted by the analyser model.

That is $f(x, \alpha) \neq y$, implying that there is some discrepancy between $f(x, \alpha)$; the guess of the analyser model and y; the prediction of Igbo native speakers.

By representing words with their patterns, pattern repetitions within a word are evident, regardless of the length of the word. These repetitions suggest the presence of certain morphological activity, and as such, could possibly contain some morphological information. Such patterns that are driven by a morphological activity(ies) occur more frequently in the corpus than some others. Hence, patterns within a cluster share some commonalities like similar segment patterns. This could be as short as a single character or as long as three or four characters.

11.5 Limitations

A major setback encountered in this study is the unavailability of computer readable Igbo texts. Unfortunately, few electronics Igbo texts that were available lacked the necessary diacritics that would qualify them as valid Igbo words. Baibụl Nsọ (Nhazi Katọlik) was scanned using Fine Reader, an Optical Character Recognizer (OCR). Fine Reader either inconsistently replaced the UTF-8 characters with invalid Igbo characters, some illegible characters and symbols, as well as numerals or it completely deleted the sub dots from their characters. As a result, some of the scanned words were totally incomprehensible. This challenge prolonged the data cleaning stage because the process could not be automated due to inconsistencies and irregularities in the data.

11.6 Conclusion

The results documented in this study give credence to the fact that word labels depict patterns embedded in words and also present as a valuable feature for clustering both concatenative and non-concatenative Igbo words according to their morphological structures without supervision. The FPT can be viewed as a counterpart to the frequent segment postulates and assumptions on which the various approaches to the induction of concatenative morphology found in the literature have so far been based on. Using FPT, the study established a nexus between Igbo words and Igbo morphological processes, via word patterns, word labels, morphological structures and ultimately morphological processes.

This study formally showed that resource scarcity does not affect morphological computational studies much as it would other levels of linguistic analysis like syntax or semantics. Formally strengthening [11] claim that a 5,000-word corpus may be adequate for *Linguistica*. Igbo morphology was induced with a wordlist of 29,191 based on frequent pattern-based induction. This refutes [4] assertion that unsupervised learning was unsuitable for languages with sparse data. Theoretically, resource scarcity should not have as drastic an effect on morphology as it should have on syntax and other levels of linguistic analysis.

Hammarström [12] and Hammarström and Borin [13] assert that unsupervised induction of non-concatenative morphology may not be achieved in the near future because such an algorithm runs in exponential time, most unsupervised induction model of concatenative morphology run in quadratic time. In this study, FPT was applied to the induction of non-concatenative aspects of Igbo morphology. The proposed FPT is novel and clusters non-concatenative Igbo words in linear time.

From observations, length and repetition are two key features that interplay in FPT. One affects the other inversely. The longer a label the higher the choice of the symbol to be added which means that such labels have a higher probability of being a unique word label. Conversely, for word labels whose length result from cascaded suffixes, the choice of symbols for the word labels are constrained because there is repetition. Therefore, longer word labels tend to have more frequent flyers. Hammarström [12] ACA fails here. This is because the segments that form a cascaded suffixation word are frequent in the wordlist.

The idea that the distribution of certain characters in a word establishes patterns, stands to reason that the distribution of characters may have morphological significance. However, if a character is repeated in a word, such a repetition may not only be due to morphology, it may also be due to chance. Fortunately, due to the equi-probability of occurrence of events that are due to chance as expressed in the ACA proposed by [12], using the FPT, we expect that the frequency of repetition of characters due to chance is expected to be relatively lower than the frequency of the repetition of characters due to morphology.

Acknowledgements I acknowledge the entire management and staff of Covenant University, Ota, Nigeria for financing the publication of this material.
The tests and results presented here are those contained in an unpublished thesis of Iheanetu (2015). I acknowledge all my supervisors for their immense contributions to this study.
The Catholic Arch Bishop of Owerri, His Grace, Dr. Amarachi Obinna is highly appreciated for the release and permission to use electronic prints of Odenigbo lecture series. Finally, I acknowledge the management and staff of Africana-Fep publishers for the permission to use Baibụl Nsọ Nhazi Katọlik for the purposes of this study.

References

1. O. Awobuliyi, *Eko Iseda-Oro Yoruba* (Montem Paper Backs, Akure, Ondo state, 2008)
2. K.R. Beesley, L. Karttunen, *Finite State Morphology* (CSLI Publications, Stanford, United States of America, 2003)
3. R. Blench, *Atlas of Nigerian languages*, 3rd edn. (2012) Retrieved from 9 June 2015, www.rogerblench.info/Language/Africa/Nigeria/Atlas%20of%20Nigerian%20Languages-%20ed%20III.pdf
4. M. Creutz, Induction of the morphology of natural language: unsupervised morpheme segmentation with application to automatic speech recognition. Ph.D. Thesis, Computer and Information Science Department, Helsinki, University of Technology, Espoo (2007), xi+110 pp.
5. G. De Pauw, G. De Schryver, Improving the computational morphological analysis of a Swahili corpus for lexicographic purposes. Lexikos Afr. Assoc. Lexicogr. (AFRILEX)-reeks Series **18**, 303–318 (2008)

6. N. Emenanjo, The interfix: an aspect of universal morphology. J. West Afr. Lang. XII **1**(1982), 77–88 (1982)
7. N. Emenanjo, *Elements of Modern Igbo Grammar* (University Press Limited (UPL), Ibadan, 1987)
8. M.A. Fullwood, T.J. O'Donnell, Learning non-concatenative morphology, in *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, 8 August 2013, Sofia, Bulgaria, pp. 21–27
9. Gale Group Inc., Igbo. Junior Worldmark Encyclopedia of World Cultures (1999). Retrieved from 10 August 2010, Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-3435900354.html
10. J. Goldsmith, Unsupervised learning of the morphology of a natural language. Mass. Inst. Technol. (MIT) Press J. **27**(2), 153–198 (2001). https://doi.org/10.1162/089120101750300490
11. J. Goldsmith, An algorithm for the unsupervised learning of morphology. Nat. Lang. Eng. **1**(1) (2005). Cambridge University Press. Retrieved from, http://hum.uchicago.edu/~jagoldsm/Papers/algorithm.pdf
12. H. Hammarström, Unsupervised learning of morphology and the languages of the world, Doctoral thesis, Department of Computer Science and Engineering, Chalmers University of Technology and University of Gothenburg, Sweden (2009), 284 pp.
13. H. Hammarström, L. Borin, Unsupervised learning of morphology. MIT Press J. **37**(2), 309–350 (2010)
14. Z. Harris, *Morpheme Boundaries within Words: Report on a Computer Test*. Transformations and Discourse Analysis Papers (1967), p. 73
15. O.U. Iheanetu, Data-driven model of Igbo morphology. Doctoral thesis, Africa Regional Centre for Information Science, (ARCIS), University of Ibadan, Nigeria (2015) 284 pp.
16. O.U. Iheanetu, O. Oha, Some salient issues in the unsupervised learning of Igbo morphology, in *World Congress on Engineering and Computer Science 2017*. Lecture Notes in Engineering and Computer Science, 25–27 October 2017, San Francisco, USA, pp. 389–393
17. P. Lambert, M. Costa-jussa, R.E. Banchas, Introduction, in *Workshop on Creating Cross-Language Resources for Disconnected Languages and Styles*, 27th May 2012. Istanbul, Turkey (2012)
18. J.J. McCarthy, A prosodic theory of nonconcatenative morphology. Linguist. Inq. **12**(3), 373–418 (1981)
19. P. MacClanahan, G. Busby, R. Haertel, K. Heal, D. Lonsdale, K. Seppi, E. Ringer, A probabilistic morphological analyser for Syriac, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (2010), pp. 810–820
20. B.M. Mba, *A Minimalist Theory and Application to Igbo* (Catholic Institute for Development Justice and Peace (CIDJAP) Press, Enugu, 2011)
21. B.M. Mba, Circumfixation: interface of morphology and syntax in Igbo derivational morphology. IOSR J. Humanit. Soc. Sci. (JHSS) **5**(6), 1–8 (2012)
22. O.M. Ndimele, *A First Course on Morphology and Syntax* (Emhai Printing and Publishing Company, Port Harcourt, 1999)
23. B.I.N. Osuagwu, G.I. Nwaozuzu, G.A. Dike, V.N. Nwaogu, L.C. Okoro, *Fundamentals of linguistics* (Colon Concept Ltd, Owerri, 1997)
24. L.M. Paul, G.F. Simons, C.D. Fennig (eds.), *Ethnologue: Languages of the World, Eighteenth Edition* (SIL International, Dallas, Texas, 2015). Retrieved from 20 June 2015, http://www.ethnologue.com/language/ibo
25. A.K. Simpson, The origin and development of nonconcatenative morphology. Ph.D. Thesis. Graduate Division of the Department of California (2009), 194 pp.
26. University of California, Los Angeles (UCLA) Language Materials Project 2009. Igbo. UCLA Language Materials Project. Retrieved from 20 October 2010, http://www.lmp.ucla.edu/Profile.aspx?LangID=13&menu=004
27. V.N. Vapnik, An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**(5), 988–999 (1999)

Chapter 12
Application of Box-Jenkins Model in Predicting Road Traffic Crashes in Nigeria

Benjamin Ufuoma Oreko and Stanley Okiy

Abstract The current road traffic accident (RTA) trends in Nigeria had not been modelled. Modelling the pattern of RTAs would assist in estimating future occurrences. The aim of this work is to develop a mathematical model capable of forecasting road traffic accident cases in Nigeria. The method adopted was a time series analysis of Box-Jenkins Autoregressive Integrated Moving Average (ARIMA). RTA data from the Federal Road Safety Corps FRSC, Abuja, were employed in analysing a 57-year accidents time series from 1960–2016. Dual statistical software namely: Microsoft Excel 2010 and SPSS 16.0 versions were employed in the computation and to ascertain the adequacy of the predictive model developed. From the results obtained, ARIMA (1, 1, 1)1 model fit the pattern of road traffic accident cases in Nigeria. The result obtained was used to assess the performance of FRSC on road traffic casualty reduction in Nigeria. A 5-year forecast of road traffic accident occurrences from 2017 to 2021 was estimated. It was observed that, if the current situation remains, road traffic accident cases and road traffic casualties annually in Nigeria would approximate 8000–12,000 and 34,000–40,000, respectively. The outcome of this research would be very useful to Road Safety Agencies and Nigeria Government in policy formulations and implementation, aimed at road traffic accident prevention and control in Nigeria.

Keywords ARIMA models · Box-Jenkins · FRSC · Intervention · Road traffic accidents · Times series

B. U. Oreko (✉)
Department of Mechanical Engineering, Federal University of Petroleum Resources, Effurun, Nigeria
e-mail: oreko.benjamin@fupre.edu.ng

S. Okiy
Department of Welding Engineering and Offshore Technology, Petroleum Training Institute, Effurun, Nigeria
e-mail: stanleyokiy2002@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_12

12.1 Introduction

The Performance Assessment of Federal Road Safety Corps in Road Traffic Casualty Reduction in Nigeria has been studied [1]. As a result of its precarious situation in Nigeria, concerted efforts had been made by road traffic safety organizations and researchers to identify and find the best approach in reducing road traffic accidents in Nigeria. Equally too, statistical techniques had been employed in analyzing road traffic accident data with a view in assessing the situation in Nigeria [2–5]. However, it appears that no substantial work had been done currently in modelling and forecasting the pattern of road traffic accident cases in Nigeria using accident data from year 1960–2016. Available studies that tend to modelled RTA data in Nigeria were applied to town/city, State or country with limited number of years [6–12]. Hence, this work seeks to develop a mathematical model capable of estimating and predicting total road traffic accident cases in Nigeria using a 57-year data.

12.2 Research Design and Method Employed

Attempt had been made to examine the population under consideration (RTA data). The sampling unit is geographical, involving the 36-State structure and the Federal capital, Abuja, Nigeria. The technique adopted in this research is the non-experimental research design concept. This concept deals mainly on data collections. The main data used were primarily from the Federal Road Safety Corps (FRSC). This agency, apart from the obligation to prevent or reduce road traffic accidents situation, it is also saddled with the responsibility of recording and keeping data observed from the scene of road traffic accidents [1, 13]. The main focus of this research was to develop a model for evaluating and forecasting RTA pattern using Box-Jenkins Autoregressive Integrated Moving Average models (ARIMA) and to assess the performance of FRSC on road traffic casualty reduction in Nigeria [1]. Some theoretical background of ARIMA time series was highlighted. A Fifty-five (57) year data (1960–2016) of total road accident cases consisting of Fatal, Serious and Minor cases were analyzed [14, 15]. Fifty four (54) years data were used for the computations and three years out-of -sample data were used to assess the performance of the model developed. Excel 2010 and SPSS 16.0 version software were employed for the computational analysis. ARIMA model developed was fitted to total road traffic accident cases. Statistical test was done to confirm the model adequacy.

12.2.1 ARIMA Modelling Procedures

According to [16], the aim of ARIMA analysis is to find a model that could accurately represent the past and future patterns of a time series. It is an empirically driven methodology of systematically identifying, estimating, diagnosing and forecasting time series data [17]. Hence, in ARIMA term, a time series is a linear function of past values and random shocks (the error terms), expressed mathematically as;

$$Y_t = \text{Pattern} + e_t \tag{12.1}$$

where, Y_t = Actual values; e_t = residual (error)

In time series analysis, an autoregressive model is represented in the following form [16, 17];

$$Y_t = \varnothing_1 y_{t-1} + \varnothing_2 y_{t-2} + \cdots + \varnothing_p y_{t-p} + e_t \tag{12.2}$$

Similarly, a moving average model is represented in the following form

$$Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \tag{12.3}$$

Integrating the autoregressive and moving average models with appropriate differencing, stationarity could be achieved. In order words, ARIMA extends the combination of Autoregression (AR) and moving average (MA) process to non-stationary processes.

From Eqs. (12.2) and (12.3), Eq. (12.4) is obtained as;

$$Y_t = \varnothing_1 y_{t-1} + \cdots + \varnothing_p y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \tag{12.4}$$

where, θ_0, $\varnothing_1$ are constant and coefficients, respectively chosen to minimize the sum of squared errors;$|\varnothing_1| < 1$ = bound of stationarity; Y_{t-1} = Previous values

In order to realize the ARIMA model based on Eq. (12.4), a plot of the 54-year total road traffic accident data was done using SPSS software. After the plot, the data was investigated for stationarity, using the plots of the autocorrelation functions (ACF) and Partial autocorrelation functions (PACF). The road traffic accident cases series derived from the plots were found to be non-stationary, hence differencing was used to achieve stationarity. Thereafter, a univariate model was fitted to Y_t. According to [16], ARIMA components could be expressed as follows;

i. ARIMA (1, 0, 0) or AR (1) model;

$$Y_{(1,0,0)} = \theta_0 + \varnothing_1 Y_{t-1} + e_t \tag{12.5}$$

Using backshift notation, this becomes

$$(1 - \varnothing_1 B) Y_t = e_t \tag{12.6}$$

ii. ARIMA (0, 1, 0) could be expressed as;

$$Y_{(0,1,0)} = Y_{t-1} + e_t \tag{12.7}$$

iii. ARIMA (0, 0, 1) is expressed

$$Y_{(0,0,1)} = \mu + \theta_1 e_{t-1} + e_t \tag{12.8}$$

where, μ = mean; θ_1 = Estimated error coefficient; $|\theta_1| < 1$ = bound of invertibility; e_{t-1} = Previous forecast error; e_t = Actual forecast error. A combination of the Eqs. (12.5)–(12.8) could give additional model components of ARIMA.

12.2.2 Model Test Adequacy

i. Root Mean Squared Error (RMSE)—Also known as Standard Deviation is the square root of the mean of the squared deviations. It helps to measure the potential error distribution when using mean for forecasting [16], Sample standard deviation S, in term of X, could be expressed as;

$$S_x = \sqrt{\frac{\sum\left(\mathrm{X_t} - \overline{\mathrm{X}}\right)^2}{n-1}} \tag{12.9}$$

Sample standard deviation S, in term of Y, could be expressed as;

$$\mathrm{S}_y = \sqrt{\frac{\sum\left(\mathrm{Y_t} - \overline{\mathrm{Y}}\right)^2}{n-1}} \tag{12.10}$$

ii. Covariance—it measures the level of association between two variables and is express as;

$$\mathrm{COV(X, Y)} = \frac{\sum\left(X_t - \overline{\mathrm{X}}\right)\left(Y_t - \overline{\mathrm{Y}}\right)}{n-1} \tag{12.11}$$

iii. Pearson correlation coefficient—it measures the proportion of the covariation of X and Y to the product of their standard deviation. This could be expressed as;

$$\mathrm{r}_{xy} = \frac{\mathrm{cov\,(X, Y)}}{\mathrm{S_x S_y}} \tag{12.12}$$

iv. Coefficient of determination, R^2—It is used to measure model adequacy. It is the proportion of the variation that could be explained by the regression equation. It is used to assess how well a model explain or predict future outcomes. This could be stated as;

$$R^2 == 1 - \frac{\sum\left(Y - \hat{Y}\right)^2}{\sum\left(Y - \overline{Y}\right)^2} \tag{12.13}$$

v. Adjusted coefficient of determination, $\overline{R^2}$—It equals the proportion of the variance in the dependent variables Y that is explained or eliminated through the relationship with the independent variables X, thus:

$$\overline{R^2} = 1 - \frac{\frac{\sum\left(Y-\hat{Y}\right)^2}{n-1}}{\frac{\sum\left(Y-\overline{Y}\right)^2}{n-1}} \tag{12.14}$$

vi. Mean absolute percentage error—it is used to compare the accuracy of prediction methods. It is expressed as:

$$\text{MAPE} = \frac{\sum_{t=1}^{n} |PE_t|}{n} \tag{12.15}$$

where,

$$\text{PE} = \text{percentage error} = \frac{\left(\text{Y} - \hat{\text{Y}}\right)}{(\text{Y})} \times 100 \tag{12.16}$$

vii. Bayesian information criterion, BIC—Bayesian information criterion could help to provide best model guideline when there are competing models and this models had confusing differences in the fits, model complexities and trade-off [16].

$$\text{BIC} = \text{Bayesian information criterion} = \text{nLog(SSE)} + \text{kLog(n)} \tag{12.17}$$

where, k = Number of parameters that are fitted in the model;
Log = logarithm; n = Number of observation; SSE = Sum of the squared errors

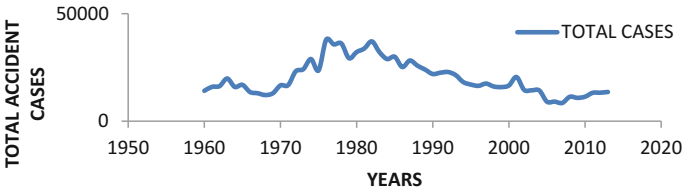


Fig. 12.1 Plot of total accident cases in Nigeria versus years from 1960–2013

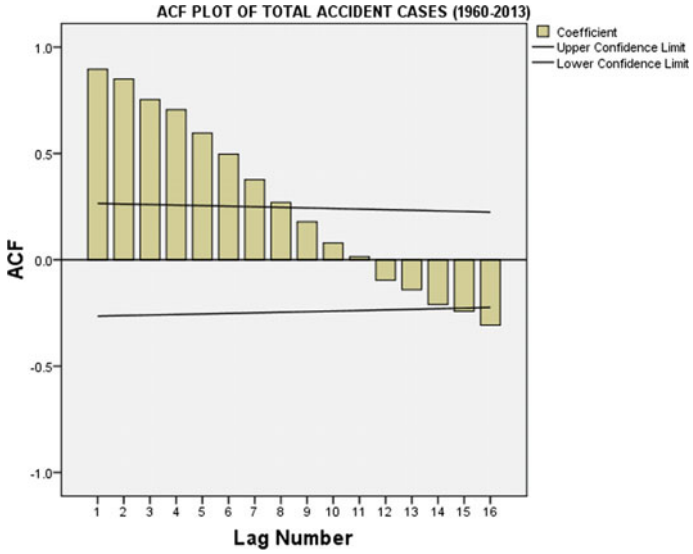


Fig. 12.2 Plot of ACF versus lag of RTA

12.2.3 Basis for Model Selection

Attempt will be made to consider Bayesian Information Criterion BIC, as the first criterion (performance measure) in selecting the best ARIMA component model that would be employed in forecasting road traffic accidents cases. Whenever ties exist in BIC, other model adequacy; R^2, RMSE and MAPE would be considered, accordingly, in the model fit selection.

12.3 Results and Discussions

A time series plot of total road traffic accident cases from 1960 to 2013 is shown in Fig. 12.1. The plot shows an irregular pattern of in road accident cases from 1960–1987. However, discernible decline was observed around 1988, which shows an irregular decreasing pattern from 1989 to 1913, occasioned with sharp increase in some years. In order to determine whether the series is stationary and non-seasonal, its Autoregression functions (ACF) and Partial Autoregression functions (PACF) were evaluated and plotted as shown in Figs. 12.2 and 12.3 respectively. The plots indicate an existence of ARIMA components model. To determine the best ARIMA models that would fit the pattern of total road traffic accidents cases in Nigeria, various components of ARIMA model were analysed and evaluated for model adequacy.

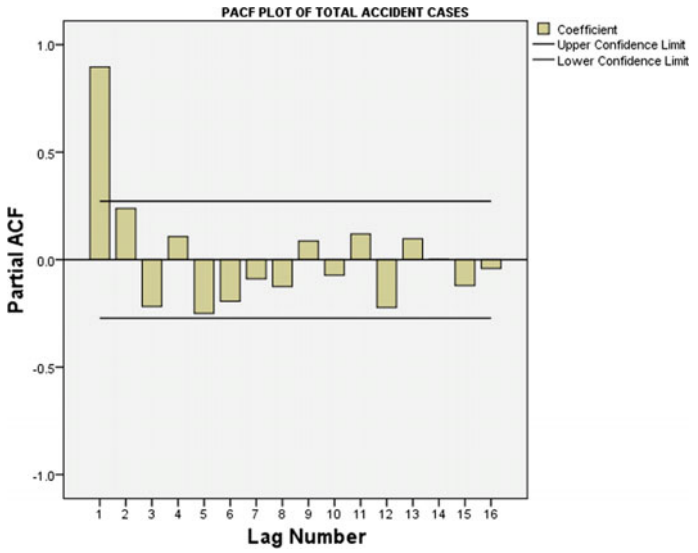


Fig. 12.3 Plot of PACF versus lag of RTA

The ACFs and PACFs of each components of ARIMA model and their respective residual values and plots were analysed and closely examined.

12.3.1 Selection of ARIMA Model for Forecasting Total Accident Cases in Nigeria

A summary of coefficient estimation for ARIMA components models for forecasting total road accident cases are tabulated in Table 12.1.

Descriptive and test statistics for model's goodness of fit were employed in order to confirm the model adequacy of each components in Tables 12.2 and 12.3.

From Table 12.3, Bayesian information criterion BIC was the first model adequacy employ to determine the best performance measure. It was observed that ARIMA (1, 1, 1)1 has the least value of 16.360. Also, the Coefficient of determination R^2, which examined the proportion of variation explained by the regression equation in each of the competing models shows that ARIMA (1, 1, 1)1 has the highest value of 84.1%. The root mean squared error (RMSE), has least value of 3253. Hence, ARIMA (1, 1, 1)1 perform better when compared to other ARIMA component models. Furthermore, to determine the effectiveness of ARIMA (1, 1, 1)1, its ACFs were assessed using Ljung-Box diagnostics test, as shown in Table 12.4.

Determining ARIMA (1, 1, 1)1 ACFs using Ljung Box diagnostics test, we have:

$$Q = 16.219 \leq \text{chi} - \text{square table}_{\alpha=0.05}^{\text{df}=16} = 26.296 \tag{12.18}$$

Table 12.1 Coefficient estimation for ARIMA components for total accident cases in Nigeria (1960–2013)

ARIMA (1, 0, 0)1		Estimate	SE
Constant		1.859×10^4	4.070×10^3
AR	Lag 1	0.902	0.057
ARIMA (1, 1, 0)1		Estimate	SE
Constant		−19.759	353.947
AR	Lag 1	−0.295	0.134
Difference		1	
ARIMA (1, 0, 1)1		Estimate	SE
Constant		18,050	4656
AR	Lag 1	0.936	0.050
MA	Lag 1	0.193	0.152
ARIMA (1, 1, 0)1		Estimate	SE
	Constant	−0.797	0.194
AR	Lag 1	−0.295	0.134
	Difference	1	
MA	Lag 1	−0.557	0.270
ARIMA(1, 1, 1)1		Estimate	SE
Constant		−20.628	365.670
AR	Lag 1	−0.797	0.194
	Difference	1	
MA	Lag 1	−0.557	0.270
ARIMA (0, 1, 1)1		Estimate	SE
Constant		−20.628	365.670
	Difference	1	
MA	Lag 1	0.214	0.137
ARIMA (0, 0, 1)1		Estimate	SE
Constant		20,090	1256
MA	Lag 1	−0.614	0.111

Table 12.2 Descriptive statistics for total accident cases

No. of years	Range	Min. value	Max. value	Sum	Mean		Std. deviation
					value	Std. error	
54	29,400	8477	37,881	1,090,751	20,199	1.085×10^3	7972

Table 12.3 Model adequacy of ARIMA models for total road accident cases in Nigeria

Total road accident cases	BIC	R^2	RMSE	MAPE
ARIMA (0, 0, 0)1	18.041	5.5×10^{-16}	7972	37.701
ARIMA (1, 0, 0)1	16.429	0.818	3432	11.939
ARIMA (0, 1, 0)1	16.368	0.814	3451	11.556
ARIMA (1, 1, 0)1	16.368	0.831	3325	12.023
ARIMA (1, 1, 1)1	16.360	0.841	3253	11.646
ARIMA (0, 1, 1)1	16.396	0.826	3372	11.923
ARIMA (0, 0, 1)1	17.471	0.485	5777	24.499
ARIMA (1, 0, 1)1	16.470	0.828	3374	12.017

Table 12.4 Ljung box diagnostics test of ARIMA models for total road accident cases

Model	Ljung-Box Q(18)		
	Statistic	DF	Sig.
Total cases-ARIMA(0, 0, 0)1	243.029	18	0.000
Total cases-ARIMA(1, 0, 0)1	39.455	17	0.002
Total cases ARIMA (0, 1, 0)1	36.933	18	0.005
Total cases-ARIMA (1, 1, 0)1	16.219	17	0.508
Total cases-ARIMA(1, 1, 1)1	12.477	16	0.711
Total cases-ARIMA(0, 1, 1)1	21.485	17	0.205
Total cases-ARIMA(0, 0, 1)1	0.0	0	0.000
Total cases-ARIMA(1, 0, 1)1	26.439	16	0.048

Since Ljung Box Q of 16.219 < chi-square table of 26.296, we could infer that ACF patterns are not statistically different than those of white noise. Therefore ARIMA (1, 1, 1)1 could be employ in forecasting the total road accidents cases in Nigeria.

12.3.2 Fitting Total Accident Cases Model Using ARIMA (1, 1, 1)1

ARIMA (1, 1, 1)1 model was fitted into road traffic accident cases in Nigeria from 1960 to 2013. This model appears stationary and non-seasonal with one AR term and one MA term being differenced with a constant mean. Figure 12.4 was the residual (error) plot of ACF and PACF versus lag of total accident cases for ARIMA (1, 1, 1)1 from 1960–2013. Figure 12.5 was the residual (error) plot of ARIMA (1, 1, 1)1 for total accident cases versus years (1960–2013). And Fig. 12.6 was the actual value versus the fitted value of ARIMA (1, 1, 1)1 for total accident cases (1960–2013). Hence, the ARIMA (1, 1, 1)1, from Eq. 12.4 could be expressed as;

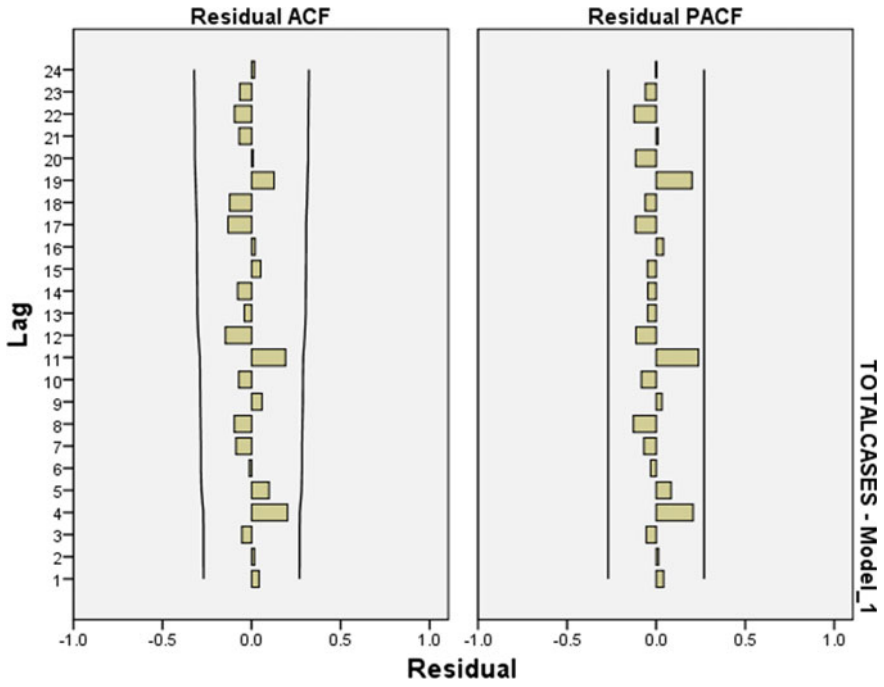


Fig. 12.4 Residual plot of ACF and PACF versus lag of total accident cases (ARIMA 1, 1, 1)1 (1960–2013)

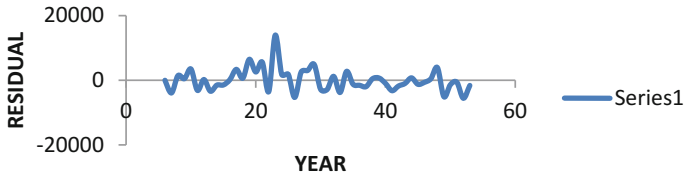


Fig. 12.5 Residual (error) plot of ARIMA (1, 1, 1)1 for total accident cases versus years (1960–2013)

$$Y_{(t)} = \theta_0 + (1 + \varnothing_1)Y_{t-1} - \varnothing_1 Y_{t-2} - \theta_1 e_{t-1} + e_t \tag{12.19}$$

From Table 12.1, ARIMA (1, 1, 1)1 constant coefficients were computed as $\theta_0 = -20.870$ and $\varnothing_1 = -0.797$ and $\theta_1 = -0.557$, the final model for ARIMA (1, 1, 1)1 becomes

$$Y_{(t)} = -20.870 + (0.203)Y_{t-1} + 0.797Y_{t-2} + 0.557e_{t-1} + e_t \tag{12.20}$$

In forecasting form Eq. (12.20) becomes;

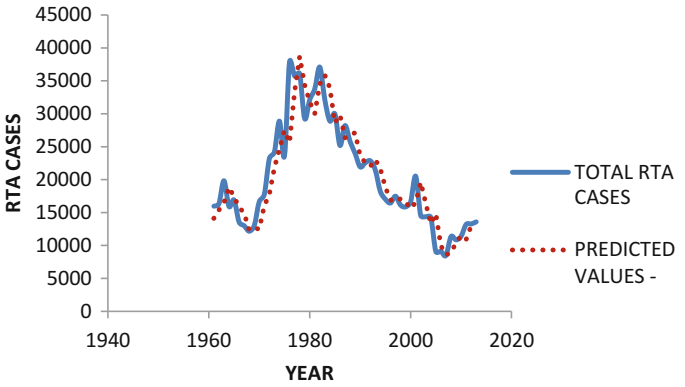


Fig. 12.6 Actual versus fitted value of ARIMA (1, 1, 1)1 for total accident cases (1960–2013)

$$\widehat{\mathrm{Y}}_{(t)} = -20.870 + (0.203)\mathbf{Y_{t-1}} + 0.797\mathbf{Y_{t-2}} + 0.557\mathbf{e_{t-1}} \tag{12.21}$$

12.3.3 Model Validation

The year 2014, 2015 and 2016 for total road accident case data were withheld as out-of-sample data, in order to assess the performance of the model developed. The actual value for total road accident cases in 2014, 2015 and 2016 as 10,383, 9734, 9694 cases, respectively [15]. Using the model developed in Eq. (12.21), we have;

$$\widehat{\mathrm{Y}}_{(2014)} = -20.870 + 0.203(13583) + 0.797(13262) + 0.557(265) \tag{12.22}$$

$$\textbf{For 2014, } \widehat{\mathbf{Y}}_{(2014)} = \textbf{13475 Cases} \tag{12.23}$$

where, $Y_{(t)}$ = 2014 actual value for total road accident cases = 10,380 cases; Y_{t-1} = actual value for 2013 total road accident cases = 13,583; Y_{t-2} = actual value 2012 total road accident cases = 13,262; Error e_{t-1} = 265. Error value e_t for 2014 = −3095

$$\textbf{For 2015} = \widehat{\mathbf{Y}}_{\mathbf{(2015)}} = \textbf{10932 cases} \tag{12.24}$$

$$e_t = \text{error value for } 2015 = 9734 - 10932 = -1198 \text{ cases} \tag{12.25}$$

$$\textbf{For 2016} = \widehat{\mathbf{Y}}_{\mathbf{2016}} = \textbf{9035 cases} \tag{12.26}$$

$$e_t = \text{error value for } 2016 = 9694 - \mathbf{9035} = -1198 \text{ cases} \tag{12.27}$$

Table 12.5 Predicted Result for total road traffic accident cases in Nigeria from 2017–2021

S/no.	Year	Model	Predicted RTA cases
1	2017	$-20.870 + (0.203)Y_{2016} + (0.797)Y_{2015} + 0.557e_{2016}$	= 9706
2	2018	$-20.870 + (0.203)\hat{Y}_{2017} + (0.797)Y_{2016} + 0.557e_{2016}$	= 9675
3	2019	$-20.870 + (0.203)\hat{Y}_{2018} + (0.797)\hat{Y}_{2017} + 0.557e_{2018}$	= 9679
4	2020	$-20.870 + (0.203)\hat{Y}_{2019} + (0.797)\hat{Y}_{2018} + 0.557e_{2019}$	= 9655
5	2021	$-20.870 + (0.203)\hat{Y}_{2020} + (0.797)\hat{Y}_{2019} + 0.557e_{2020}$	= 9653

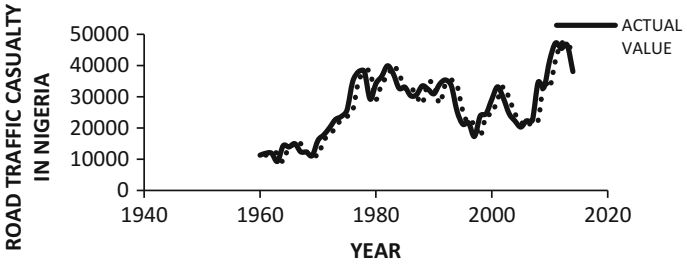


Fig. 12.7 Plot of actual value of road traffic casualty with intervention measures versus fitted model response. Extract from [1]

Since the forecast values of 13,475, 10,932, 9035 cases for year 2014, 2015, 2016 respectively, fall between the 95% prediction intervals of 5662–21,288, the model developed could be used to estimate and predict road traffic accident occurrences in Nigeria. Forecasting the next 5 years (2017–2022), we have the following result as shown in Table 12.5.

From Table 12.5, we observe that road traffic accident cases from 2017 to 2021 would be on the declining trend and in Fig. 12.7, [1] observed that road traffic casualty occurrences in Nigeria were on the decreasing profile. If the current situation remains, road traffic accident cases and road traffic casualties annually in Nigeria would approximate 8000–12,000 and 34,000–40,000, respectively. Likewise, the predicted estimate for road traffic accidents on average appeared less than the average RTA cases of 19,676, from year 1960 to 2016 [15].

12.4 Conclusion

Attempt had been made in this work to establish a model that could predict road traffic accidents cases in Nigeria. This model, if fully explore could assist road safety agencies to re-strategize in their policy implementation in order to reduce road traffic accident occurrences in Nigeria. It is expected that Box-Jenkins model applied in modelling road traffic accident cases in Nigeria would be of benefit to Federal Government of Nigeria and its established safety agencies such as the Federal Road Safety Corps, the Nigeria Police Force, other stakeholders and the general public to make projections of road traffic accident occurrences in Nigeria and be able to proffer stringent policies that would help make significant impact on road traffic accident reduction in Nigeria.

References

1. B.U. Oreko, S. Okiy, An Intervention Theoretic Modelling Approach on the Performance Assessment of Federal Road Safety Corps in Road Traffic Casualty Reduction in Nigeria, in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress On Engineering and Computer Science 2017*, 25–27 October, 2017, San Francisco, USA, pp 771–776
2. B.U. Oreko, C.C. Nwobi-Okoye, S, Okiy and A.C. Igboanugo, *Modelling the Impact of Intervention Measures on Total Accident Cases in Nigeria Using Box-Jenkins Methodology: A Case Study of Federal Road Safety Commission,* vol. 4:1345043 (Cogent Engineering, part of Taylor & Francis Group, 2017). ISSN 2331-1916
3. B.U. Oreko, C.C. Nwobi-Okoye, S, Okiy and A.C. Igboanugo, A Multivariate Linear Regression Theoretic Approach to Modelling Road Traffic Accidents in Nigeria. J. Niger. Assoc. Math. Phys. **35**, 193–200 (2017) ISSN 1116-4336
4. A.C. Igboanugo, E.F. Ekheumelo, Intervention analysis of road traffic accident, advance material research, **18**–**19**, 375-382 (2007); C.C. Nwobi-Okoye, Game theoretic analysis of road traffic problems in Nigeria. Afr. Res. Rev. **5**(3), 184-199 (2011)
5. J. Ohakwe, I.S. Iwueze, D.C. Chikezie, Analysis of road traffic accidents in Nigeria: a case study of Obinze/Nekede/Iheagwa road in Imo State, Southeastern Nigeria. Asian J. Appl. Sci. **4**, 166–175 (2011)
6. N. Eke, E.N. Etubu, S.O. Nwosu, Road accident mortalities in Port Harcourt, Nigeria. J. Forensic Med. Toxicol. **1**(2) (2000)
7. F.J. Jegede, Spatio-temporal analysis of road traffic accidents in Oyo State, Nigeria. Accid. Anal. Prev. **24**(2), 149–155 (1988)
8. O.E, Charles-Owaba, K.A. Adebiyi, On the performance of the FRSC: the case with the Oyo State sector command. Niger. J. Eng. Manag. **2**(3), 50–56 (2001)
9. S.O. Ismaila, O.G. Akanbi, N.O. Adekunle, O.E. Charles-Owaba, Current trends of vehicular accidents in Nigeria. Pac. J. Sci. Technol. **10**(2), 262–268 (2009)
10. O.E. Afamdi, Trends and characteristics of road traffic accidents in Nigeria. J. Accid. Anal. Prev. **106**(1), 27–29 (1986)
11. S.E. Asogwa, C.N. Obionu, Urban pedestrian accidents and the efficiency of counter measure. Central Afr. J. Med. **31**, 62–65. PubMed (1986)
12. K.U. Nnadi, C.C. Ibe, An econometric study of the road safety impact of Nigeria's federal road safety commission. Int. J. Soft Comput. **2**(2), 279–284 (2007)
13. Federal Road Safety Corps, Annual Report, Abuja (2014): FRSC. www.frsc.gov.ng. Assessed on 4 Dec 2015

14. Federal Road Safety Corps Annual Report, Abuja (2015): FRSC. www.frsc.gov.ng. Assessed on 19 Nov 2016
15. Federal Road Safety Corps Annual Report 2016, Abuja: FRSC. www.frsc.gov.ng. Assessed on 5 Sept 2017
16. S.A. De-Lurgio, *Forecasting Principle and Applications* (McGraw-Hill Publishers, ISBN, 1998), p. 0071159983
17. G. Box, G. Jenkins, C. Reinsel, *Time Series Analysis: Forecasting and Control*. 4th edn. (Prentice Hall, New Jersey, 2008). ISBN-13: 978-0130607744

Chapter 13
Gumbel Distribution: Ordinary Differential Equations

Hilary I. Okagbue, Olasunmbo O. Agboola, Abiodun A. Opanuga, Jimevwo G. Oghonyon and Pelumi E. Oguntunde

Abstract In this chapter, homogenous ordinary differential equations (ODES) of different orders were obtained for the probability density function, quantile function, survival function inverse survival function, hazard function and reversed hazard functions of Gumbel distribution. This is possible since the aforementioned probability functions are differentiable. Differentiation and modified product rule were used to obtain the required ordinary differential equations, whose solutions are the respective probability functions. The different conditions necessary for the existence of the ODEs were obtained and it is almost in consistent with the support that defined the various probability functions considered. The parameters that defined each distribution greatly affect the nature of the ODEs obtained. This method provides new ways of classifying and approximating other probability distributions apart from Gumbel distribution considered in this chapter. In addition, the result of the quantile function can be compared with quantile approximation using the quantile mechanics.

Keywords Differential calculus · Gumbel distribution · Hazard function
Inverse survival function · Quantile function · Quantile mechanics
Reversed hazard function · Survival function

H. I. Okagbue (✉) · O. O. Agboola · A. A. Opanuga · J. G. Oghonyon · P. E. Oguntunde
Department of Mathematics, Covenant University, Ota, Nigeria
e-mail: hilary.okagbue@covenantuniversity.edu.ng

O. O. Agboola
e-mail: olasunmbo.agboola@covenantuniversity.edu.ng

A. A. Opanuga
e-mail: abiodun.opanuga@covenantuniversity.edu.ng

J. G. Oghonyon
e-mail: jimevwo.oghonyon@covenantuniversity.edu.ng

P. E. Oguntunde
e-mail: pelumi.oguntunde@covenantuniversity.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_13

13.1 Introduction

Calculus is a very key tool in the determination of mode of a given probability distribution and in estimation of parameters of probability distributions, amongst other uses. The method of maximum likelihood is an example.

Differential equations often arise from the understanding and modeling of real life problems or some observed physical phenomena. Approximations of probability functions are one of the major areas of application of calculus and ordinary differential equations in mathematical statistics. The approximations are helpful in the recovery of the probability functions of complex distributions [1–4] especially in quantile approximations.

Apart from mode estimation, parameter estimation and approximation, probability density function (PDF) of probability distributions can be expressed as ODE whose solution is the PDF. Some of which are available. They include: beta distribution [5], Lomax distribution [6], beta prime distribution [7], Laplace distribution [8] and raised cosine distribution [9].

The aim of this research is to develop homogenous ordinary differential equations for the probability density function (PDF), Quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of Gumbel distribution. This will also help to provide the answers as to whether there are discrepancies between the support of the distribution and the necessary conditions for the existence of the ODEs. Similar results for other distributions have been proposed, see [10–22] for details.

Gumbel distribution is often used in modeling the distribution of the minimum and maximum of different distributions. The distribution was proposed by Gumbel [23, 24] and had undergone modifications such as its generalization [25, 26], beta Gumbel distribution [27], exponentiated Gumbel distribution [28], Kumaraswamy Gumbel distribution [29], exponentiated generalized Gumbel distribution [30], McDonald Gumbel distribution [31] and transmuted exponentiated Gumbel distribution [32]. Some aspects of the distribution studied by several authors such as in Bayesian analysis [33] and interval estimation [34]. The distribution has been applied in different fields and areas such as: modeling annual distribution of flood [35, 36], fitting extreme wind speeds [37–39], modeling and predicting storm [40], modeling the frequency of earthquakes [41], extreme rainfall data analysis by [42–44], estimate the probability of pipe wall perforation [45], extreme tsunami heights [46], irrigation analysis [47], estimation of the mean weight of fish in aquaculture cages [48], modeling and estimating risk of disease transmission [49] and modeling corrosion [50].

Differential calculus was used to obtain the results.

13.2 Probability Density Function

The probability density function of the Gumbel distribution is given as;

$$f(x) = \frac{1}{\sigma} \exp\left\{-\left[\frac{x-\mu}{\sigma} + \exp\left(-\frac{x-\mu}{\sigma}\right)\right]\right\} \tag{13.1}$$

To obtain the first order ordinary differential equation for the probability density function of the Gumbel distribution, differentiate Eq. (13.1), to obtain;

$$f'(x) = -\frac{1}{\sigma}\left(1 - \exp\left(-\frac{x-\mu}{\sigma}\right)\right) f(x) \tag{13.2}$$

The condition necessary for the existence of equation is $\sigma > 0$, $\mu, x \in \mathbb{R}$. Equation (13.2) is differentiated in an attempt to obtain a simplified ODE.

$$f''(x) = \frac{f'^2(x)}{f(x)} - \frac{f'(x)}{\sigma} - \frac{f(x)}{\sigma^2} \tag{13.3}$$

Simplify Eq. (13.3) using Eqs. (13.1) and (13.2). The outcome is the second order ordinary differential equation for the probability density function of the Gumbel distribution given by;

$$\sigma^2 f(x) f''(x) - \sigma^2 f'^2(x) + \sigma f(x) f'(x) + f^2(x) = 0 \tag{13.4}$$

with initial value conditions;

$$f(0) = \frac{1}{\sigma} \exp\left\{-\left[-\frac{\mu}{\sigma} + \exp\left(\frac{\mu}{\sigma}\right)\right]\right\} \tag{13.5}$$

$$f'(0) = -\frac{1}{\sigma^2}\left(1 - \exp\left(\frac{\mu}{\sigma}\right)\right)\left\{\exp\left\{-\left[-\frac{\mu}{\sigma} + \exp\left(\frac{\mu}{\sigma}\right)\right]\right\}\right\} \tag{13.6}$$

13.3 Quantile Function

The Quantile function of the Gumbel distribution is given as;

$$Q(p) = \mu - \sigma \ \ln(-\ln \ p) \tag{13.7}$$

To obtain the first order ordinary differential equation for the Quantile function of the Gumbel distribution, differentiate Eq. (13.7), to obtain;

$$Q'(p) = -\frac{\sigma}{p \ \ln \ p} \tag{13.8}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$.

Differentiate Eq. (13.8), to obtain;

$$Q''(p) = \left[\frac{\sigma}{p^2(\ln\ p)^2} + \frac{\sigma}{p^2 \ln\ p}\right] \tag{13.9}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$. Square both sides of Eq. (13.8);

$$Q'^2(p) = \frac{\sigma^2}{p^2(\ln\ p)^2} \tag{13.10}$$

$$\frac{Q'^2(p)}{\sigma} = \frac{\sigma}{p^2(\ln\ p)^2} \tag{13.11}$$

Also dividing both sides of Eq. (13.8) by p;

$$\frac{Q'(p)}{p} = -\frac{\sigma}{p^2 \ln\ p} \tag{13.12}$$

Substitute Eqs. (13.11) and (13.12) into Eq. (13.9);

$$Q''(p) = \left[\frac{Q'^2(p)}{\sigma} - \frac{Q'(p)}{p}\right] \tag{13.13}$$

The second order ordinary differential equation for the Quantile function of the Gumbel distribution is given by;

$$\sigma p Q''(p) - p Q'^2(p) + \sigma Q'(p) = 0 \tag{13.14}$$

with the initial value conditions; $Q(0.1) = \mu - 0.834\sigma,\ Q'(0.1) = 4.343\sigma$.

13.4 Survival Function

The survival function of the Gumbel distribution is given as;

$$S(t) = 1 - \exp\left\{-\left[\exp\left(-\frac{t-\mu}{\sigma}\right)\right]\right\} \tag{13.15}$$

To obtain the first order ordinary differential equation for the survival function of the Gumbel distribution, differentiate Eq. (13.15), to obtain;

$$S'(t) = -\frac{1}{\sigma}\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right)\exp\left\{-\left[\exp\left(-\frac{t-\mu}{\sigma}\right)\right]\right\} \tag{13.16}$$

The condition necessary for the existence of equation is $\sigma > 0, \mu, t \in \mathbb{R}$. Equation (13.15) can be written as;

$$\exp\left\{-\left[\exp\left(-\frac{t-\mu}{\sigma}\right)\right]\right\} = 1 - S(t) \tag{13.17}$$

Substitute Eq. (13.17) into Eq. (13.16);

$$S'(t) = -\frac{1}{\sigma}\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right)(1 - S(t)) \tag{13.18}$$

Differentiate Eq. (13.18);

$$S''(t) = -\frac{1}{\sigma}\left\{-\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right)S'(t) - \frac{1}{\sigma}\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right)(1 - S(t))\right\} \tag{13.19}$$

The condition necessary for the existence of equation is $\sigma > 0, \mu, t \in \mathbb{R}$

$$S''(t) = \frac{1}{\sigma}\exp\left(-\frac{t-\mu}{\sigma}\right)\left\{S'(t) + \frac{1}{\sigma}(1 - S(t))\right\} \tag{13.20}$$

Equation (13.18) can be simplify as;

$$\frac{1}{\sigma}\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right) = -\frac{S'(t)}{1 - S(t)} \tag{13.21}$$

Substitute Eq. (13.21) into Eq. (13.20);

$$S''(t) = -\frac{S'(t)}{1 - S(t)}\left(S'(t) + \frac{1}{\sigma}(1 - S(t))\right) \tag{13.22}$$

The second order ordinary differential equation for the survival function of the Gumbel distribution is given by;

$$\sigma(1 - S(t))S''(t) + \sigma S'^2(t) + (1 - S(t))S'(t) = 0 \tag{13.23}$$

with the initial value conditions;

$$S(0) = 1 - \exp\left\{-\left[\exp\left(\frac{\mu}{\sigma}\right)\right]\right\} \tag{13.24}$$

$$S'(0) = -\frac{1}{\sigma}\left(\exp\left(\frac{\mu}{\sigma}\right)\right)\exp\left\{-\left[\exp\left(\frac{\mu}{\sigma}\right)\right]\right\} \tag{13.25}$$

13.5 Inverse Survival Function

The inverse survival function of the Gumbel distribution is given as;

$$Q(p) = \mu - \sigma \ \ln(-\ln(1-p)) \tag{13.26}$$

To obtain the first order ordinary differential equation for the inverse survival function of the Gumbel distribution, differentiate Eq. (13.35),

$$Q'(p) = \frac{\sigma}{(1-p)\ln(1-p)} \tag{13.27}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$. Differentiate Eq. (13.27), to obtain;

$$Q''(p) = \left[\frac{\sigma}{(1-p)^2(\ln(1-p))^2} + \frac{\sigma}{(1-p)^2\ln(1-p)}\right] \tag{13.28}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$. Square both sides of Eq. (13.27);

$$Q'^2(p) = \frac{\sigma^2}{(1-p)^2(\ln(1-p))^2} \tag{13.29}$$

$$\frac{Q'^2(p)}{\sigma} = \frac{\sigma}{(1-p)^2(\ln(1-p))^2} \tag{13.30}$$

Also dividing both sides of Eq. (13.27) by 1 − p;

$$\frac{Q'(p)}{p} = \frac{\sigma}{(1-p)^2\ln(1-p)} \tag{13.31}$$

Substitute Eqs. (13.31) and (13.30) into Eq. (13.28);

$$Q''(p) = \left[\frac{Q'^2(p)}{\sigma} + \frac{Q'(p)}{1-p}\right] \tag{13.32}$$

The second order ordinary differential equation for the inverse survival function of the Gumbel distribution is given by;

$$\sigma(1-p)Q''(p) - (1-p)Q'^2(p) - \sigma\, Q'(p) = 0 \tag{13.33}$$

with initial value conditions; $Q(0.1) = \mu + 2.25\sigma$, $Q'(0.1) = -10.5458\sigma$.

13.6 Hazard Function

The hazard function of the Gumbel distribution is given as;

$$h(t) = \frac{\frac{1}{\sigma}\exp\left\{-\left[\frac{t-\mu}{\sigma} + \exp\left(-\frac{t-\mu}{\sigma}\right)\right]\right\}}{1 - \exp\left\{-\left[\exp\left(-\frac{t-\mu}{\sigma}\right)\right]\right\}} \tag{13.34}$$

To obtain the first order ordinary differential equation for the hazard function of the Gumbel distribution, differentiate Eq. (13.34);

$$h'(t) = \left\{-\frac{1}{\sigma}\left(1 - \exp\left(-\frac{t-\mu}{\sigma}\right)\right) + h(t)\right\}h(t) \tag{13.35}$$

The condition necessary for the existence of equation is $\sigma > 0, \mu, t \in \mathbb{R}$. Differentiate Eq. (13.35);

$$\begin{aligned} h''(t) = &\left\{-\frac{1}{\sigma}\left(1 - \exp\left(-\frac{t-\mu}{\sigma}\right)\right) + h(t)\right\}h'(t) \\ &+ \left\{-\frac{1}{\sigma^2}\left(\exp\left(-\frac{t-\mu}{\sigma}\right)\right) + h'(t)\right\}h(t) \end{aligned} \tag{13.36}$$

The condition necessary for the existence of equation is $\sigma > 0, \mu, t \in \mathbb{R}$. Simplify Eq. (13.36) to obtain;

$$h''(t) = \frac{h'^2(t)}{h(t)} - \frac{h(t)}{\sigma^2} - \frac{h'(t)}{\sigma} + \frac{h^2(t)}{\sigma} + h(t)h'(t) \tag{13.37}$$

The second order ordinary differential equation for the hazard function of the Gumbel distribution is given by;

$$\sigma^2 h(t)h''(t) - \sigma^2 h'^2(t) + (\sigma h(t) - \sigma^2 h^2(t))h'(t) + h^2(t) - \sigma h^3(t) = 0 \tag{13.38}$$

with initial value conditions;

$$h(0) = \frac{\frac{1}{\sigma}\exp\left\{-\left[-\frac{\mu}{\sigma} + \exp\left(\frac{\mu}{\sigma}\right)\right]\right\}}{1 - \exp\left\{-\left[\exp\left(\frac{\mu}{\sigma}\right)\right]\right\}} \tag{13.39}$$

$$h'(0) = \left\{-\frac{1}{\sigma}\left(1 - \exp\left(\frac{\mu}{\sigma}\right)\right) + h(0)\right\}h(0) \tag{13.40}$$

13.7 Reversed Hazard Function

The reversed hazard function of the Gumbel distribution is given as;

$$j(t) = \frac{2}{\sigma}\exp\left(-\frac{t-\mu}{\sigma}\right) \tag{13.41}$$

To obtain the first order ordinary differential equation for the reversed hazard function of the Gumbel distribution, differentiate Eq. (13.41);

$$j'(t) = \frac{1}{\sigma^2}\exp\left(-\frac{t-\mu}{\sigma}\right) \tag{13.42}$$

The condition necessary for the existence of equation is $\sigma > 0$, $\mu, t \in \mathbb{R}$.

$$\sigma j'(t) = \frac{1}{\sigma}\exp\left(-\frac{t-\mu}{\sigma}\right) \tag{13.43}$$

$$\sigma j'(t) = -j(t) \tag{13.44}$$

The first order ordinary differential equation for the reversed hazard function of the Gumbel distribution is given by;

$$\sigma j'(t) + j(t) = 0 \tag{13.45}$$

with initial value condition; $j(0) = \frac{1}{\sigma}\exp\left(\frac{\mu}{\sigma}\right)$.

13.8 Conclusion

Ordinary differential equations (ODEs) has been obtained for the probability density function (PDF), quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of the Gumbel distribution.

This differential calculus, modified product rule and efficient algebraic simplifications were used to derive the various classes of the ODEs. The parameter and the supports that characterize the Gumbel distribution determine the nature, existence, orientation and uniqueness of the ODEs. The results are in agreement with those available in scientific literature. Furthermore several methods can be used to obtain desirable solutions to the ODEs. This method of characterizing distributions cannot be applied to distributions whose PDF or CDF are either not differentiable or the domain of the support of the distribution contains singular points.

Acknowledgements This work was supported by Covenant University, Nigeria.

References

1. G. Steinbrecher, W.T. Shaw, Quantile mechanics. Euro. J. Appl. Math. **19**(2), 87–112 (2008)
2. H.I. Okagbue, M.O. Adamu, T.A. Anake, Quantile approximation of the chi-square distribution using the quantile mechanics, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science*, *25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 477–483
3. H.I. Okagbue, M.O. Adamu, T.A. Anake, Solutions of chi-square quantile differential equation, in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 813–818
4. Y. Kabalci, On the Nakagami-m inverse cumulative distribution function: closed-form expression and its optimization by backtracking search optimization algorithm. Wireless Pers. Comm. **91**(1), 1–8 (2016)
5. W.P. Elderton, *Frequency Curves and Correlation* (Charles and Edwin Layton, London, 1906)
6. N. Balakrishnan, C.D. Lai, *Continuous Bivariate Distributions*, 2nd edn. (Springer, New York, London, 2009)
7. N.L. Johnson, S. Kotz, N. Balakrishnan*, Continuous Univariate Distributions*, vol. 2, 2nd edn. (Wiley, 1995)
8. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994). ISBN: 0-471-58495-9
9. H. Rinne, Location scale distributions, linear estimation and probability plotting using MATLAB (2010)
10. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of Fréchet distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 186–191
11. H.I. Okagbue, P.E. Oguntunde, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated generalized exponential distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 192–197
12. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of cauchy, standard cauchy and log-cauchy distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 198–204
13. H.I. Okagbue, S.A. Bishop, A.A. Opanuga, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Burr XII and Pareto distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 399–404
14. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of Gompertz and Gamma Gompertz distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 405–411
15. H.I. Okagbue, M.O. Adamu, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of 3-parameter Weibull distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 539–545
16. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of exponentiated Fréchet distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 546–551

17. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, S.A. Bishop, Classes of ordinary differential equations obtained for the probability functions of Half-Cauchy and power Cauchy distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 552–558
18. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of exponential and truncated exponential distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 858–864
19. H.I. Okagbue, O.O. Agboola, P.O. Ugwoke, A.A. Opanuga, Classes of Ordinary differential equations obtained for the probability functions of exponentiated Pareto distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 865–870
20. H.I. Okagbue, O.O. Agboola, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of Gumbel distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 871–875
21. H.I. Okagbue, O.A. Odetunmibi, A.A. Opanuga, P.E. Oguntunde, Classes of ordinary differential equations obtained for the probability functions of half-normal distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 876–882
22. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, E.A. Suleiman, Classes of ordinary differential equations obtained for the probability functions of Harris extended exponential distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 883–888
23. E.J. Gumbel, Les valeurs extremes des distributions statistiques. Annal. I'Institut Henri Poincaré **5**(2), 115–158 (1935)
24. E.J. Gumbel, The return period of flood flows. Ann. Math. Stat. **12**, 163–190 (1941)
25. S. Adeyemi, M.O. Ojo, A generalization of the Gumbel distribution. Kragujevac J. Math. **25**, 19–29 (2003)
26. K. Cooray, Generalized Gumbel distribution. J. Appl. Stat. **37**(1), 171–179 (2010)
27. S. Nadarajah, S. Kotz, The beta Gumbel distribution. Math. Probl. Eng. **2004**(4), 323–332 (2004)
28. S. Nadarajah, The exponentiated Gumbel distribution with climate application. Environmetrics **17**(1), 13–23 (2006)
29. G.M. Cordeiro, S. Nadarajah, E.M.M. Ortega, The Kumaraswamy Gumbel distribution. Stat. Meth. Appl. **21**(2), 139–168 (2012)
30. T. Andrade, H. Rodrigues, M. Bourguignon, G.M. Cordeiro, The exponentiated generalized Gumbel distribution. Rev. Colomb. Estad. **38**(1), 123–143 (2015)
31. E. de Brito, G.O. Silva, G.M. Cordeiro, C.G.B. Demétrio, The McDonald Gumbel model. Comm. Stat. Theo. Meth. **45**(11), 3367–3382 (2016)
32. D. Deka, B. Das, B.K. Baruah, Transmuted exponentiated gumbel distribution (TEGD) and its application to water quality data. Pak. J. Stat. Oper. Res. **13**(1), 115–126 (2017)
33. R.A. Chechile, Bayesian analysis of Gumbel distributed data. Comm. Stat. Theo. Methods **30**(3), 485–496 (2001)
34. A. Asgharzadeh, M. Abdi, S. Nadarajah, Interval estimation for Gumbel distribution using climate records. Bull. Malay. Math. Sci. Soc. **39**(1), 257–270 (2016)
35. M.N. Leese, Use of censored data in the estimation of Gumbel distribution parameters for annual maximum flood series. Water Res. Res. **9**(6), 1534–1542 (1973)
36. N. Mujere, Flood frequency analysis using the Gumbel distribution. Int. J. Comp. Sci. Eng. **3**(7), 2774–2778 (2011)

37. E. Simiu, N.A. Heckert, J.J. Filliben, S.K. Johnson, Extreme wind load estimates based on the Gumbel distribution of dynamic pressures: an assessment. Struct. Safety **23**(3), 221–229 (2001)
38. E.C. Pinheiro, S.L.P. Ferrari, A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data. J. Stat. Comput. Simul. **86**(11), 2241–2261 (2016)
39. M.P. Pes, E.B. Pereira, J.A. Marengo, F.R. Martins, D. Heinemann, M. Schmidt, Climate trends on the extreme winds in Brazil. Renew. Energy **109**, 110–120 (2017)
40. S. Yue, The Gumbel logistic model for representing a multivariate storm event. Adv. Water Res. **24**(2), 179–185 (2000)
41. A. Kijko, Modified form of the first Gumbel distribution: model for the occurrence of large earthquakes, II: Estimation of parameters. Acta Geophys. Polon. **31**(2), 147–159 (1983)
42. I. Vidal, A Bayesian analysis of the Gumbel distribution: an application to extreme rainfall data. Stoc. Env. Res. Risk Assess. **28**(3), 571–582 (2014)
43. J.L. Ng, S. Abd Aziz, Y.F. Huang, A. Wayayok, M.K. Rowshon, Analysis of annual maximum rainfall in Kelantan. Malaysia, Acta Hortic. **1152**, 11–17 (2017)
44. N. Boudrissa, H. Cheraitia, L. Halimi, Modelling maximum daily yearly rainfall in northern Algeria using generalized extreme value distributions from 1936 to 2009. Meteo. Appl. **24**(1), 114–119 (2017)
45. Z.S. Asadi, R.E. Melchers, Extreme value statistics for pitting corrosion of old underground cast iron pipes. Relia. Eng. Syst. Safety **162**, 64–71 (2017)
46. S. Dong, J. Zhai, S. Tao, Long-term statistics of extreme tsunami height at Crescent City. J. Ocean Uni. China **16**(3), 437–446 (2017)
47. N. Gaj, C.A. Madramootoo, Long-term simulations of the hydrology for sugarcane fields in the humid tropics: case study on Guyana's coastland. J. Irrig. Drain. Eng. **143**(8), art. 05017002 (2017)
48. E. Soliveres, P. Poveda, V.D. Estruch, I. Pérez-Arjona, V. Puig, P. Ordóñez, J. Ramis, V. Espinosa, Monitoring fish weight using pulse-echo waveform metrics. Aquacult. Eng. **77**, 125–131 (2017)
49. J. Weusten, H. van Drimmelen, M. Vermeulen, N. Lelie, A mathematical model for estimating residual transmission risk of occult hepatitis B virus infection with different blood safety scenarios. Transfusion **57**(3), 841–849 (2017)
50. J.O. Okeniyi, C.C. Nwadialo, F.E. Olu-Steven, S.S. Ebinne, T.E. Coker, E.T. Okeniyi, A.S. Ogbiye, T.O. Durotoye, E.O.O. Badmus, C3H7NO2S effect on concrete steel-rebar corrosion in 0.5M H_2SO_4 simulating industrial/microbial environment, in *AIP Conference Proceedings, 1814*, Art. no. 020035 (2017)

Chapter 14
Half-Normal Distribution: Ordinary Differential Equations

Hilary I. Okagbue, Oluwole A. Odetunmibi, Sheila A. Bishop, Pelumi E. Oguntunde and Abiodun A. Opanuga

Abstract In this chapter, homogenous ordinary differential equations (ODES) of different orders were obtained for the probability density function, quantile function, survival function inverse survival function, hazard function and reversed hazard functions of half-normal distribution. This is possible since the aforementioned probability functions are differentiable. Differentiation and modified product rule were used to obtain the required ordinary differential equations, whose solutions are the respective probability functions. The different conditions necessary for the existence of the ODEs were obtained and it is almost in consistent with the support that defined the various probability functions considered. The parameters that defined each distribution greatly affect the nature of the ODEs obtained. This method provides new ways of classifying and approximating other probability distributions apart from half-normal distribution considered in this chapter. In addition, the result of the quantile function can be compared with quantile approximation using the quantile mechanics.

Keywords Differential calculus · Half-normal distribution · Hazard function
Inverse survival function · Quantile function · Quantile mechanics
Reversed hazard function · Survival function

H. I. Okagbue (✉) · O. A. Odetunmibi · S. A. Bishop · P. E. Oguntunde · A. A. Opanuga
Department of Mathematics, Covenant University, Ota, Nigeria
e-mail: hilary.okagbue@covenantuniversity.edu.ng

O. A. Odetunmibi
e-mail: oluwole.odetunmibi@covenantuniversity.edu.ng

S. A. Bishop
e-mail: sheila.bishop@covenantuniversity.edu.ng

P. E. Oguntunde
e-mail: pelumi.oguntunde@covenantuniversity.edu.ng

A. A. Opanuga
e-mail: abiodun.opanuga@covenantuniversity.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_14

14.1 Introduction

Calculus in general and differential calculus in particular is a very key tool in the determination of mode of a given probability distribution and in estimation of parameters of probability distributions, amongst other uses. The method of maximum likelihood is an example.

Differential equations often arise from the understanding and modeling of real life problems or some observed physical phenomena. Approximations of probability functions are one of the major areas of application of calculus and ordinary differential equations in mathematical statistics. The approximations are helpful in the recovery of the probability functions of complex distributions [1–4] especially in quantile approximations.

Apart from mode estimation, parameter estimation and approximation, probability density function (PDF) of probability distributions can be expressed as ODE whose solution is the PDF. Some of which are available. They include: beta distribution [5], Lomax distribution [6], beta prime distribution [7], Laplace distribution [8] and raised cosine distribution [9].

The aim of this research is to develop homogenous ordinary differential equations for the probability density function (PDF), Quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of half-normal distribution. This will also help to provide the answers as to whether there are discrepancies between the support of the distribution and the necessary conditions for the existence of the ODEs. Similar results for other distributions have been proposed, see [10–23] for details.

Half-normal distribution is a normal distribution with a mean set at zero and parameterized to the domain of positive real numbers and zero being the lower bound. Pewsey [24, 25] worked on the improved statistical inference for the distribution while [26] proposed unbiased estimators for the parameters of the distribution, which according to them, performs better than the traditional maximum likelihood. Some generalizations are available for the distribution such as: the extended generalized half-normal distribution [27], beta generalized half-normal distribution [28], generalized half-normal distribution [29, 30], discrete half-normal distribution [31], an extension of the half-normal distribution called the slashed half-normal distribution [32], Kumaraswamy generalized half-normal distribution [33], beta generalized half-normal geometric distribution [34], gamma half-normal distribution [35] and alpha half-Normal Slash distribution [36]. Also available are epsilon half-normal distribution [37]. The distribution is a sub-model of exponentiated generalized gamma distribution proposed by [38] and generalized half-t distribution by [39]. Also available is the odd log-logistic generalized half-normal lifetime distribution [40].

In addition, the distribution was generalized with the Airy model to obtain the M-Wright distribution [41]. Details of the new method of generating the distribution is given in [42] and its application to quality control were highlighted by [43]. Lang [44] used the distribution to model wage gap between immigrants and natives in Germany.

The distribution was among those used by Schoenberg et al. [45] in modeling of the distribution of the sizes of wildfire.

Differential calculus was used to obtain the results.

14.2 Probability Density Function

The probability density function of the half-normal distribution is given by;

$$f(x) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{14.1}$$

To obtain the first order ordinary differential equation for the probability density function of the half-normal distribution, differentiate Eq. (14.1);

$$f'(x) = -\frac{x\sqrt{2}}{\sigma^3\sqrt{\pi}} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{14.2}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

Simplify Eq. (14.2) using Eq. (14.1);

$$f'(x) = -\left(\frac{x}{\sigma^2}\right)\frac{\sqrt{2}}{\sigma\sqrt{\pi}} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} = -\frac{x}{\sigma^2} f(x) \tag{14.3}$$

The first order ordinary differential for the probability density function of the half-normal distribution is given as;

$$\sigma^2 f'(x) + x f(x) = 0 \tag{14.4}$$

with initial value condition; $f(1) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \mathrm{e}^{-\frac{1}{2\sigma^2}}$.

To obtain the second order ordinary differential equation for the probability density function of the half-normal distribution, differentiate Eq. (14.2);

$$f''(x) = \frac{\sqrt{2}}{\sigma^3\sqrt{\pi}} \left\{ \frac{x^2}{\sigma^2} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} - \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \right\} \tag{14.5}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

Two differential equations can be obtained from the simplification of Eq. (14.5). They are presented as Eqs. (14.6) and (14.7).

$$\sigma^4 f''(x) + (\sigma^2 - x^2) f(x) = 0 \tag{14.6}$$

$$\sigma^2 f''(x) + x f'(x) + f(x) = 0 \tag{14.7}$$

To obtain the third order ordinary differential equation for the probability density function of the half-normal distribution, differentiate Eq. (14.5);

$$f'''(x) = \frac{\sqrt{2}}{\sigma^3\sqrt{\pi}}\left\{-\frac{x^3}{\sigma^4}e^{-\frac{x^2}{2\sigma^2}} + \frac{3x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right\} \tag{14.8}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

Six differential equations can be obtained from the simplification of equations. They are listed as Eqs. (14.9–14.14).

$$\sigma^6 f'''(x) - (3\sigma^2 x - x^3)f(x) = 0 \tag{14.9}$$

$$\sigma^4 f'''(x) + (3\sigma^2 - x^2)f'(x) = 0 \tag{14.10}$$

$$\sigma^4 f'''(x) - x^2 f'(x) - 3xf(x) = 0 \tag{14.11}$$

$$\sigma^4 f'''(x) + \sigma^2 x f''(x) - 2xf(x) = 0 \tag{14.12}$$

$$\sigma^2 f'''(x) + xf''(x) + 2f'(x) = 0 \tag{14.13}$$

$$\sigma^4 f'''(x) + \sigma^2 x f''(x) + \sigma^2 f'(x) - xf(x) = 0 \tag{14.14}$$

$$f'(1) = -\frac{\sqrt{2}}{\sigma^3\sqrt{\pi}}e^{-\frac{1}{2\sigma^2}} \tag{14.15}$$

To obtain the fourth order ordinary differential equation for the probability density function of the half-normal distribution, differentiate Eq. (14.8);

$$f^{'v}(x) = \frac{\sqrt{2}}{\sigma^3\sqrt{\pi}}\left\{\frac{x^4}{\sigma^6}e^{-\frac{x^2}{2\sigma^2}} - \frac{6x^2}{\sigma^4}e^{-\frac{x^2}{2\sigma^2}} + \frac{3}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\right\} \tag{14.16}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

Twelve differential equations can be obtained from the simplification of Eq. (14.16);

$$\sigma^8 f^{'v}(x) - (x^4 - 6\sigma^2 x^2 + 3\sigma^4)f(x) = 0 \tag{14.17}$$

$$x\sigma^6 f^{'v}(x) + (x^4 - 6\sigma^2 x^2 + 3\sigma^4)f'(x) = 0 \tag{14.18}$$

$$\sigma^6 f^{'v}(x) + (x^3 - 6\sigma^2 x)f'(x) - 3\sigma^4 f(x) = 0 \tag{14.19}$$

$$\sigma^6 f^{'v}(x) - (\sigma^2 x^2 - 3\sigma^4)f''(x) + 2x^2 f(x) = 0 \tag{14.20}$$

$$\sigma^4 f^{'v}(x) - (x^2 - 3\sigma^2)f''(x) - 2xf'(x) = 0 \tag{14.21}$$

$$\sigma^6 f^{'v}(x) + x\sigma^4 f'''(x) - 3(\sigma^2 - x^2)f(x) = 0 \tag{14.22}$$

$$x\sigma^4 f^{'v}(x) + x^2\sigma^2 f'''(x) + 3(\sigma^2 - x^2)f'(x) = 0 \tag{14.23}$$

$$\sigma^2 f^{'v}(x) + xf'''(x) + 3f''(x) = 0 \tag{14.24}$$

$$\sigma^4 f^{'v}(x) - (x^2 - 2\sigma^2)f''(x) - 3xf'(x) - f(x) = 0 \tag{14.25}$$

$$\sigma^4 f^{'v}(x) + \sigma^2 x f'''(x) - 3xf'(x) - 3f(x) = 0 \tag{14.26}$$

$$\sigma^6 f'^{v}(x) + \sigma^4 x f'''(x) + 2\sigma^4 f''(x) + (x^2 - \sigma^2) f(x) = 0 \tag{14.27}$$

$$x\sigma^4 f'^{v}(x) + \sigma^2 x^2 f'''(x) + 2\sigma^2 x f''(x) - (x^2 - \sigma^2) f'(x) = 0 \tag{14.28}$$

14.3 Quantile Function

The Quantile function of the half-normal distribution is given by;

$$Q(p) = \sigma\sqrt{2} erf^{-1}(p) \tag{14.29}$$

To obtain the first order ordinary differential equation for the Quantile function of the half-normal distribution, differentiate Eq. (14.29);

$$Q'(p) = \frac{\sigma\sqrt{2\pi}}{2} \mathrm{e}^{[erf^{-1}(p)]^2} \tag{14.30}$$

The condition necessary for the existence of the equation is

$$\sigma > 0, 0 \leq p < 1.$$

Simplify Eq. (14.30) using Eq. (14.29), however Eq. (14.29) becomes;

$$\frac{Q(p)}{\sigma\sqrt{2}} = erf^{-1}(p) \tag{14.31}$$

$$Q'(p) = \frac{\sigma\sqrt{2\pi}}{2} \mathrm{e}^{\frac{Q^2(p)}{2\sigma^2}} \tag{14.32}$$

$$\ln Q'(p) = \ln\left(\frac{\sigma\sqrt{2\pi}}{2}\right) + \frac{Q^2(p)}{2\sigma^2} \tag{14.33}$$

$$2\sigma^2 \ln Q'(p) - Q^2(p) - 2\sigma^2 g = 0 \tag{14.34}$$

where $g = \ln\left(\frac{\sigma\sqrt{2\pi}}{2}\right)$ with initial value condition; $Q(0) = 0$.

To obtain the second order ordinary differential equation for the Quantile function of the half-normal distribution, differentiate Eq. (14.30);

$$Q''(p) = \frac{\sigma\pi\sqrt{2}}{2} erf^{-1}(p)\left(\mathrm{e}^{[erf^{-1}(p)]^2}\right)^2 \tag{14.35}$$

The condition necessary for the existence of the equation is

$$\sigma > 0, 0 \leq p < 1.$$

Substitute Eq. (14.29) into Eq. (14.35);

$$Q''(p) = \frac{\pi}{2} Q(p) \left(e^{[erf^{-1}(p)]^2} \right)^2 \tag{14.36}$$

The following equations obtained from the simplification of Eq. (14.30) are needed to simplify Eq. (14.36);

$$Q'^2(p) = \frac{\pi \sigma^2}{2} \left(e^{[erf^{-1}(p)]^2} \right)^2 \tag{14.37}$$

$$\frac{Q'^2(p)}{\sigma^2} = \frac{\pi}{2} \left(e^{[erf^{-1}(p)]^2} \right)^2 \tag{14.38}$$

Substitute Eq. (14.38) into Eq. (14.36);

$$Q''(p) = Q(p) \frac{Q'^2(p)}{\sigma^2} \tag{14.39}$$

$$\sigma^2 Q''(p) - Q(p) Q'^2(p) = 0 \tag{14.40}$$

14.4 Survival Function

The survival function of the half-normal distribution is given by;

$$S(t) = 1 - erf\left(\frac{t}{\sigma \sqrt{2}} \right) \tag{14.41}$$

To obtain the first order ordinary differential equation for the Survival function of the half-normal distribution, differentiate Eq. (14.41);

$$S'(t) = -\frac{\sqrt{2}}{\sigma \sqrt{\pi}} e^{-\frac{t^2}{2\sigma^2}} = -f(t) \tag{14.42}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

The second and third order ordinary differential equations for the Survival function of the half-normal distribution can also be obtained using Eq. (14.42);

$$\sigma^2 S''(t) + t S'(t) = 0 \tag{14.43}$$

$$S'(0) = -\frac{\sqrt{2}}{\sigma \sqrt{\pi}} \tag{14.44}$$

$$\sigma^4 S'''(t) + (\sigma^2 - t^2) S'(t) = 0 \tag{14.45}$$

$$\sigma^2 S'''(t) + \sigma t S''(t) + S(t) = 0 \tag{14.46}$$

14.5 Inverse Survival Function

The inverse survival function of the half-normal distribution is given by;

$$Q(p) = \sigma\sqrt{2}erf^{-1}(1-p) \tag{14.47}$$

To obtain the first order ordinary differential equation for the Quantile function of the half-normal distribution, differentiate Eq. (14.47);

$$Q'(p) = -\frac{\sigma\sqrt{2\pi}}{2}\mathrm{e}^{[erf^{-1}(1-p)]^2} \tag{14.48}$$

The condition necessary for the existence of the equation is

$$\sigma > 0, 0 \leq p < 1.$$

Simplify Eq. (14.48) using Eq. (14.47), however Eq. (14.47) becomes;

$$\frac{Q(p)}{\sigma\sqrt{2}} = erf^{-1}(1-p) \tag{14.49}$$

$$Q'(p) = -\frac{\sigma\sqrt{2\pi}}{2}\mathrm{e}^{\frac{Q^2(p)}{2\sigma^2}} \tag{14.50}$$

$$\ln\ Q'(p) = -\ln\left(\frac{\sigma\sqrt{2\pi}}{2}\right) - \frac{Q^2(p)}{2\sigma^2} \tag{14.51}$$

$$2\sigma^2 \ln\ Q'(p) + Q^2(p) + 2\sigma^2 g = 0 \tag{14.52}$$

where $g = \ln\left(\frac{\sigma\sqrt{2\pi}}{2}\right)$ and with initial value condition; $Q(0) = 0$.

14.6 Hazard Function

The hazard function of the half-normal distribution is given by;

$$h(t) = \left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}}\mathrm{e}^{-\frac{t^2}{2\sigma^2}}\right)\left(1 - erf\left(\frac{t}{\sigma\sqrt{2}}\right)\right)^{-1} \tag{14.53}$$

To obtain the first order ordinary differential equation for the hazard function of the half-normal distribution, differentiate Eq. (14.53);

$$h'(t) = -\left\{ \frac{t}{\sigma^2} + \frac{\left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}} \mathrm{e}^{-\frac{t^2}{2\sigma^2}} \right)}{\left(1 - erf\left(\frac{t}{\sigma\sqrt{2}} \right) \right)} \right\} h(t) \tag{14.54}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

$$h'(t) = -\left\{ \frac{t}{\sigma^2} + h(t) \right\} h(t) \tag{14.55}$$

$$\sigma^2 h'(t) = -(t + \sigma^2 h(t))h(t) \tag{14.56}$$

$$\sigma^2 h'(t) + \sigma^2 h^2(t) + t h(t) = 0 \tag{14.57}$$

with initial value condition: $h(0) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}}$.

Higher order ordinary differential equations can also be obtained such as;

$$\sigma^2 h''(t) + (t + 2\sigma^2 h(t))h'(t) + h(t) = 0 \tag{14.58}$$

$$\sigma^2 h'''(t) + (t + 2\sigma^2 h(t))h''(t) + 2\sigma^2 h'^2(t) + h'(t) + h(t) = 0 \tag{14.59}$$

14.7 Reversed Hazard Function

he reversed hazard function of the half-normal distribution is given by;

$$j(t) = \frac{\frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{\frac{t^2}{2\sigma^2}}}{erf\left(\frac{t}{\sigma\sqrt{2}} \right)} = \left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{t^2}{2\sigma^2}} \right) \left(erf\left(\frac{t}{\sigma\sqrt{2}} \right) \right)^{-1} \tag{14.60}$$

To obtain the first order ordinary differential equation for the reversed hazard function of the half-normal distribution, differentiate Eq. (14.60);

$$j'(t) = -\left\{ \frac{t}{\sigma^2} + \frac{\left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{t^2}{2\sigma^2}} \right)}{\left(erf\left(\frac{t}{\sigma\sqrt{2}} \right) \right)} \right\} j(t) \tag{61}$$

The condition necessary for the existence of the equation is $\sigma > 0$.

$$j'(t) = -\left\{ \frac{t}{\sigma^2} + j(t) \right\} j(t) \tag{62}$$

$$\sigma^2 j'(t) + \sigma^2 j^2(t) + tj(t) = 0 \tag{63}$$

with initial value condition; $j(1) = \frac{\frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{1}{2\sigma^2}}}{erf\left(\frac{1}{\sigma\sqrt{2}}\right)}$

14.8 Conclusion

Ordinary differential equations (ODEs) has been obtained for the probability density function (PDF), quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of the half-normal distribution.

This differential calculus, modified product rule and efficient algebraic simplifications were used to derive the various classes of the ODEs. The parameter and the supports that characterize the half-normal distribution determine the nature, existence, orientation and uniqueness of the ODEs. The results are in agreement with those available in scientific literature. Furthermore several methods can be used to obtain desirable solutions to the ODEs. This method of characterizing distributions cannot be applied to distributions whose PDF or CDF are either not differentiable or the domain of the support of the distribution contains singular points. Furthermore, the differential equations can be used to model phenomena described by the distribution.

Acknowledgements This work was supported by Covenant University, Ota, Nigeria.

References

1. G. Steinbrecher, W.T. Shaw, Quantile mechanics. Euro. J. Appl. Math. **19**(2), 87–112 (2008)
2. H.I. Okagbue, M.O. Adamu, T.A. Anake, Quantile approximation of the chi-square distribution using the quantile mechanics, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 477–483
3. H.I. Okagbue, M.O. Adamu, T.A. Anake, Solutions of chi-square quantile differential equation, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 813–818
4. Y. Kabalci, On the Nakagami-m inverse cumulative distribution function: closed-form expression and its optimization by backtracking search optimization algorithm. Wireless Pers. Comm. **91**(1), 1–8 (2016)
5. W.P. Elderton, *Frequency Curves and Correlation* (Charles and Edwin Layton, London, 1906)
6. N. Balakrishnan, C.D. Lai, *Continuous Bivariate Distributions*, 2nd edn. (Springer, New York, London, 2009)

7. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2. 2nd edn (Wiley, 1995)
8. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994). ISBN: 0-471-58495-9
9. H. Rinne, Location scale distributions, linear estimation and probability plotting using MATLAB (2010)
10. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of Fréchet distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 186–191
11. H.I. Okagbue, P.E. Oguntunde, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated generalized exponential distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 192–197
12. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of cauchy, standard cauchy and log-cauchy distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 198–204
13. H.I. Okagbue, S.A. Bishop, A.A. Opanuga, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Burr XII and Pareto distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 399–404
14. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of Gompertz and Gamma Gompertz distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 October, 2017* (San Francisco, U.S.A., 2017), pp. 405–411
15. H.I. Okagbue, M.O. Adamu, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of 3-parameter Weibull distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 539–545
16. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of exponentiated Fréchet distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 546–551
17. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, S.A. Bishop, Classes of ordinary differential equations obtained for the probability functions of Half-Cauchy and power Cauchy distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 552–558
18. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of exponential and truncated exponential distributions, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 858–864
19. H.I. Okagbue, O.O. Agboola, P.O. Ugwoke, A.A. Opanuga, Classes of Ordinary differential equations obtained for the probability functions of exponentiated Pareto distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 865–870
20. H.I. Okagbue, O.O. Agboola, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of Gumbel distribution, in *Lecture Notes in*

Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017 (San Francisco, U.S.A., 2017), pp. 871–875
21. H.I. Okagbue, O.A. Odetunmibi, A.A. Opanuga, P.E. Oguntunde, Classes of ordinary differential equations obtained for the probability functions of half-normal distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 876–882
22. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, E.A. Suleiman, Classes of ordinary differential equations obtained for the probability functions of Harris extended exponential distribution, in *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science, 25–27 Oct 2017* (San Francisco, U.S.A., 2017), pp. 883–888
23. H.I. Okagbue, M.O. Adamu, T.A. Anake, Ordinary differential equations of the probability functions of weibull distribution and their application in ecology. Int. J. Engine. Future Tech. **15**(4), 57–78 (2018)
24. A. Pewsey, Large-sample inference for the general half-normal distribution. Comm. Stat. Theo. Meth. **31**(7), 1045–1054 (2002)
25. A. Pewsey, Improved likelihood based inference for the general half-normal distribution. Comm. Stat. Theo. Meth. **33**(2), 197–204 (2004)
26. A.G. Nogales, P. Perez, Unbiased estimation for the general half-normal distribution. Comm. Stat. Theo. Meth. **44**(7), 3658–3667 (2015)
27. J.J. Duarte Sanchez, W.W. da Luz Freitas, G.M. Cordeiro, The extended generalized half-normal distribution. Braz. J. Prob. Stat. **30**(3), 366–384 (2016)
28. R.R. Pescim, C.G.B. Demétrio, G.M. Cordeiro, E.M.M. Ortega, M.R. Urbano, The beta generalized half-normal distribution. Comput. Stat. Data Anal. **54**(4), 945–957 (2010)
29. K. Cooray, M.M.A. Ananda, A generalization of the half-normal distribution with applications to lifetime data. Comm. Stat. Theo. Methods **37**(9), 1323–1337 (2008)
30. W.J. Huang, N.C. Su, A study of generalized normal distributions. Comm. Stat. Theo. Methods **46**(11), 5612–5632 (2017)
31. A.W. Kemp, The discrete half-normal distribution, in *Advances in Mathematical and Statistical Modeling* (Birkhäuser, Boston), pp. 353–360
32. N.M. Olmos, H. Varela, H.W. Gómez, H. Bolfarine, An extension of the half-normal distribution. Stat. Papers **53**(4), 875–886 (2012)
33. G.M. Cordeiro, R.R. Pescim, E.M.M. Ortega, The Kumaraswamy generalized half-normal distribution for skewed positive data. J. Data Sci. **10**(2), 195–224 (2012)
34. T. Ramires, E.M. Ortega, G.M. Cordeiro, G. Hamedani, The beta generalized half-normal geometric distribution. Studia Scient. Math. Hunga. **50**(4), 523–554 (2013)
35. A. Alzaatreh, K. Knight, On the gamma half-normal distribution and its applications. J. Modern Appl. Stat. Methods **12**(1), 103–119 (2013)
36. W. Gui, An alpha half-normal slash distribution for analyzing non negative data. Comm. Stat. Theo. Methods **44**(22), 4783–4795 (2015)
37. L.M. Castro, H.W. Gómez, M. Valenzuela, Epsilon half-normal model: properties and inference. Comput. Stat. Data Anal. **56**(12), 4338–4347 (2012)
38. G.M. Cordeiro, E.M.M. Ortega, G.O. Silva, The exponentiated generalized gamma distribution with application to lifetime data. J. Stat. Comput. Simul. **81**(7), 827–842 (2011)
39. B.Y. Murat, J. Ahad, F.Z. Dogru, O. Arslan, The generalized half-t distribution. Stat. Interf. **10**(4), 727–734 (2017)
40. G.M. Cordeiro, M. Alizadeh, R.R. Pescim, E.M. Ortega, The odd log-logistic generalized half-normal lifetime distribution: properties and applications. Comm. Stat. Theo. Methods **46**(9), 4195–4214 (2017)
41. D.O. Cahoy, Minkabo, Inference for three-parameter M-Wright distributions with applications. Model. Assist. Stat. Appl. **12**(2), 115–125 (2017)
42. A.G. Glen, L.M. Leemis, D.J. Luckett, Survival distributions based on the incomplete gamma function ratio, in *Proceedings Winter Simulation Conference*, Article. Number 7822105 (2017)

43. C.Y. Chou, H.R. Liu, Properties of the half-normal distribution and its application to quality control. J. Industr. Technol. **14**(3), 4–7 (1998)
44. G. Lang, The difference between wages and wage potentials: earnings disadvantages of immigrants in Germany. J. Econ. Inequal. **3**(1), 21–42 (2005)
45. F.P. Schoenberg, R. Peng, J. Woods, On the distribution of wildfire sizes. Environmetrics **14**(6), 583–592 (2003)

Chapter 15
Sources of Stressors Among Physics Education Undergraduates of Chukwuemeka Odumegwu Ojukwu University, Nigeria

Theresa U. Okafor

Abstract The study was aimed at examining the role of inadequate physics facilities/accommodation, socio-economic status and poor time management as stressors among Chukwuemeka Odumegwu University (COOU), Anambra State Physics Education undergraduates. The study was guided by five research questions. The survey design was adopted. All the physics Education undergraduates of COOU, Anambra State totalling forty students was used for the study. The researcher used questionnaire instrument targeted to physics Education students of COOU. The research questions were answered using percentages. It was found that low socio-economic status was the highest contributor of stress among COOU Physics Education undergraduates Anambra State with 80% of the respondents indicating that it was a major contributor of stress. It was recommended that counselling services should be rendered to Physics Education students with a view of helping them improve their socio-economic status by helping them getting vacation jobs in their institutions and other places and also helping them manage the time available to them effectively. Moreover conducive atmosphere for learning, adequate Physics facilities and well equipped medical units should be provided by the government and university authorities in order to reduce stress among physics Education undergraduates, so as to meet their academic challenges.

Keywords Economic status · Emotional problems · Inadequate physics facilities Laboratories · Stress · Stressors

T. U. Okafor (✉)
Department of Science Education, Chukwuemeka Odumegwu University, Uli, Nigeria
e-mail: teresaokafor@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_15

15.1 Introduction

Physics Education aims at promoting intellectual and calculative skills of the individual and his ability to use such skills, to create and manipulate material resources to create wealth, promote health and societal development. Okafor [11] observes that physics Education aims at the following:

- Facilitate a transition for scientific concepts and techniques acquired in integrated science with physics;
- Provide students with basic knowledge in physics concept and principles through efficient selection of content and sequencing;
- Show progress in its link with industry, everyday life benefits and hazards; and show; physics in its inter-relationship with other subjects.
- Provide a course, which is complete for pupils not proceeding to higher education, while it is at the same time a reasonable adequate foundation for a post secondary physics course.

However, when the student is burdened by stressors, these wonderful aims may be an illusion. Stress means pressure, tension or worry resulting from problems in one's life while stressor means an environmental condition or influence of stress (i.e. causes stress for) an organism. Stress may be unavoidable in human existence. Stress is a part of everyday life, including that of students. However, it could be managed if certain steps are taken. Stress has to do with day to day events and how an individual reacts to them. Any alteration in an individual's life, whether pleasant, or unpleasant, usually, requires some kind of human readjustment. Alike [1] opined that when this readjustment negatively affects the normal psychological or physiological well being of an individual, he or she experiences stress. Stress could arise as a result of physical exhaustion, anxiety, ill-health and financial problems.

Stress is a regular feature associated with students' life on campuses of Nigerian Universities.

This problem tends to undermine the educational achievement and by implication the economic growth and development of the country. The degrees of stress physics students undergo as a result of change in environment, hostel, access to computers, science equipments, laboratories, classroom accommodation, missing result scores, health problems, family problems and inability to manage time available to them would be overwhelming [1]. Consequently, Maisamari [8] described stress as a state of discomfort, tension or emotional pain which arises when an individual is faced with situation which present a demand that is important for the individual to meet but for which his capacities and resources are inadequate, he contended further that cognitively, stress interferes with thinking, concentration and memory [9]. He also observed that poor time management is a sort of stress. When students fail to manage the time available to them effectively, they may undergo some sort of stress. When it is time for examination, by staying awake all night, reading in order to cover the course outline. This may lead to break down their health condition prior to or during examination, thereby leading to academic underachievement.

Moderate stress motivates the individual to action necessary to adapt in rapidly changing world. When it comes to stress and its bad effect, it is seen as disabling, disorganising and disorienting. Woolfolk (2005) identified that behaviourally, stress is characterised by anxiety, anger, depression, intrinsic, thoughts/images, obsessions, altered motivation, impaired intellectual activities, aggression, substance abuse, avoidance, lack of concentration and loss of interest in physics education activities that were eagerly sought for. Also UNICEF [16] reported that countries like Japan and Malaysia made tremendous leap into economic and technological advancement because they focus on the development of their human resources, especially in physics, mathematics and technology, if Nigeria must borrow a leaf from there nations. Physics and technology education should occupy a position of pre eminence in national developmental projects.

Socio-economic status has been observed to be a form of stressor in the study of physics education. Sociologists define social class, or socio-economic status (SES) in terms of an individual's income, occupation, education and prestige in society [15]. Students from working class or lower-class background are less likely than middle class students to enter school knowing how to count, to name letters, to cut with scissors or name colours. They are less likely to perform better than children from middle-class homes [10, 14]. Equally Alika and Egbochukwu [2] found out that the socio-economic status of an individual exerts a lot of influence on the academic attainment of the individual. This implies that an undergraduate of Physics Education whose parental socio-economic status is low will find it difficult to manage stressors that he may likely encounter in the university due to the handicap posed by his socio-economic background. This is more apparent, as we have also many students on campus who are training themselves in the universities. In ablution, Franklin [5] found that individual who suffer extreme hunger for a long period of time due to poverty, experience a wide spread effect such as depression, poor concentration, hostile and irritable disposition. Leandro [7] in Awake publication found that when faced with financial distress it affects one's mood.

Physics students can be stressed as a result of lack of financial backup especially as regards purchase of basic necessities. This may raise tension and stress and as a result affect their academic achievement.

i. **Sources of Stress in Physics Education**

Alika [1] categorised major sources of stress in physics Education into physical, psychological and environmental stresses. Ikeotuonye [6] also categorised stressors to include the following: physical and health problems, financial and social problems, sexual, psychological, moral, family, environmental and vocational problems. Onyemerekeyo [13] also contended that modern day issues resulting from our lifestyle, tasks and challenges amongst others are all sources of stress. Above all, the sources of stress in physics education are the gate way by which the emotional, cognitive and physiological activity of the individual is disrupted. Understanding these stress sources, then possible interference to physics education students' well-being is important, since it is informative and educative.

ii. **Management of stress in physics education**

Neidharelt, Malcolm and Robert in Alike [1, p. 19], advanced for coping strategies in the management of stress, they include:

- Build up general health: through proper nutrition, rest, exercise and other positive health practices.
- Change the situation: this has to do with avoiding the source of stress.
- Change your mind: that is your perceptions of or thought about stressors.
- Change your body: this has to do with the ability to substitute relaxation responses for stress responses.

Edstrom [3] opined that stress manifestations include the physical cognitive, emotional and behavioural dimensions. Excessive stress may cause physical and mental health problems, reduce students self esteem and may affect students academic achievement.

One of the objectives of the national policy on Education [4] is inculcation of the right types of values and attitudes for the survival of the individual and the society at large. Irrespective of this noble objective, it appears that the stress which physics education undergraduates' face is on the increase. There is the need for universities to make adequate preparation at ensuring the provision of adequate laboratories. Computers physics equipment, conducive atmosphere for learners, development of a positive self image, self-direction and the acquisition skill in coping with stress, failure to ensure these issues are addressed may jeopardise physics education students' ambition and goals. The pressure of anxiety amongst physics

Education undergraduates may lead to despair and fear towards challenges that they may encounter. In universities and colleges stressors may take the form of unaccustomed activities. Sharing a nom with a stranger makes demands on the student, a new form of academic activity is demanding, rushing to secure the use of computers and laboratories ahead of time, financial resources and health issues can be stressors. Physics Education Undergraduates of the universities of Chukwuemeka Odumegwu Ojukwu University (COOU) are faced with a lot of stress, which range from academic, physical, emotional, poor time management and socio-economic problems amongst others.

15.2 Statement of the Problem

Psychologists assert that stress is an inevitable aspect of human existence, therefore, there is the need to exercise some sort of control over it, in order to minimize its damaging effects especially as regards its impact o Physics Education undergraduates. The advancement of any nation has to do with her human resources, UNICEF [16]. Therefore, it becomes imperative that efforts should be made to reduce stress among Physics Education students who may contribute to the future technological growth of the country.

The Physics Education undergraduates in the Chukwuemeka Odumegwu Ojukwu University experience a lot of discomfort ranging from a change of environment, academic pressures, inadequate laboratories, computers, accommodation, poor time management, health issues, poor socio-economic status and transportation. There is need to ensure that these issues are addressed by the various university authorities because failure to ensure that this is done may jeopardize Physics Education Undergraduates' ambition and goals. Without doubt, the problem of inadequate laboratories and accommodation in our campuses is a source of stress not only to students but to their parents. Usually in the Chukwuemeka Odumegwu Ojukwu University, priority is given to first year and final years students in term accommodation. Chronic stressors on campuses expose Physics Education Undergraduates to unwanted behavioural tendencies. When stress becomes unbearable to an individual it may lead to distress, which may cause some damage to the entire well-being of the individual. Therefore, a study of this nature is imperative in Nigerian universities in order to determine the sources of stress and strategies that should be adopted to eliminate stressors among Physics Education Undergraduates, in order to build a strong egalitarian and virile nation [12].

15.3 Purpose of Study

The main purpose of this study was to find out the sources of stress among Chukwuemeka Odumegwu Ojukwu University Physics Education undergraduates and proffer strategies that could be adopted in order to minimize or eliminate stress on university campuses.

C Research questions

The following research questions were posed to guide the study.

1. Does inadequate science facilities/accommodation constitute a source of stress to Physics education Undergraduates in Chukwuemeka Odumegwu Ojukwu University, Anambra State?
2. To what extent will poor health condition be a source of stress to Physics Education undergraduates in Chukwuemeka Odumegwu Ojukwu University, Anambra State?
3. Is emotional problem a source of stress to Chukwuemeka Odumegwu Ojukwu University, Anambra State Physics Education Undergraduates?
4. Is low socio-economic status a source of stress to Chukwuemeka Odumegwu Ojukwu University, Anambra State Physics Education Undergraduates?
5. To what extent will poor time management constitute a source of stress to Chukwuemeka Odumegwu Ojukwu University, Anambra State.

15.4 Methodology

The survey method was adopted for this study. All the Physics Education Undergraduates were used for the study, from one to final year totaling forty students. All the Physics Education students were used because they are not so large, they understand the system and also are in better position to express their experiences.

i. **Instrument**

The instrument used for this study was questionnaires. The questionnaire was designed based on the research questions for the purpose of this study. The instrument was made up of twenty items. All items represent statement to which they responded to on a 4 point modified Likert scale, ranging from strongly agree, agree, disagree and strongly disagree.

ii. **Validations of the instrument**

The questionnaire was face validated by two specialists in Chukwuemeka Odumegwu Ojukwu University (COOU) and two experts in industrial physics in the Chukwuemeka Odumegwu Ojukwu University, Anambra State. The specialist approved the face validity of the instrument and endorsed the questionnaire as having content validity after removing items that were found to be less importance to the study.

iii. **Reliability of the Instrument**

The reliability of the instrument was determined by using the test retest method. The questionnaire was administered to fifteen students in selected departments. The 2 weeks interval was to ensure that respondents do not remember exactly their previous responses and that the items being measured are relatively stable among respondents in order to show consistency in scores in both tests. The scores obtained were correlated and the reliability co-efficient of 0.89 was obtained, thus indicating that the instrument was adequate for the study.

vi. **Method of Data Collection**

The researcher with the aid of two research assistants administered instruments so to ensure orderliness and avoidance of attrition instruction on how to fill the questionnaire followed their distributors. The forty respondents filled the questionnaire immediately and returned them on the spot. The questionnaires contained strongly agree, agree, disagree and strongly disagree. The data generated from the respondents were analyzed by testing the five research questions formulated for the study. The statistical method employed was simple percentage.

15.5 Results

15.5.1 Research Question 1

Does inadequate science facilities/accommodation constitute a source of a stress to Physics Education Undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State.

Table 15.1 shows that out of 40 respondents, 10 and 14 students responded to strongly agree and agree respectively while 9 and 7 responded disagree and strongly disagree. When the mean percentage was computed, it was 60%. Thus implies that inadequate science facilities/accommodation is a source of stressor to Physics Education Undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State.

15.5.2 Research Question 2

To what extent will poor health condition be a source of stress Physics Education Undergraduates of University of Chukwuemeka Odumegwu Ojukwu University, Anambra State?

Table 15.2 showed that 8 and 22 respondents strongly agreed and agreed respectively that poor health is a source of stressor among Physics Education Undergraduate and this number of respondents constitute 75.0% of the respondents. This implies that poor health to a great extent is a source of stressor among Physics education students of Chukwuemeka Odumegwu Ojukwu University, Anambra State.

Table 15.1 Inadequate science condition

Variable	No. of respondent	Strongly agree	Agree	Disagree	Strongly disagree	%
Inadequate facilities/accommodation	40	10	14	9	7	60

Table 15.2 Poor health condition

Variable	No. of respondents	Strongly agree	Agree	Disagree	Strongly disagree	%
Poor health	40	18	22	7	3	75.0

Table 15.3 Emotional problem

Variable	No. of respondents	Strongly agree	Agree	Disagree	Strongly disagree	%
Poor health	40	18	22	7	3	75.0

Table 15.4 Socio-economic status

Variable	No. of respondents	Strongly agree	Agree	Disagree	Strongly disagree	%
Poor health	40	7	25	5	3	80

15.5.3 Research Question 3

Is emotional problem a source of stress to physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University?

Table 15.3 showed that 8 and 22 respondents strongly agreed and agree respectively that emotional problem is a source of stressor among Chukwuemeka Odumegwu Ojukwu University Anambra State, physics education undergraduates. The mean percentage of the figure above was computed to be 75.0%. This implies that emotional stressor is a problem among physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University Anambra State.

15.5.4 Research Question 4

Is low socio-economic status a source of stress to Chukwuemeka Odumegwu Ojukwu University, Anambra State, physics education undergraduates?

Table 15.4 showed that 7 and 25 respondents strongly agreed and agreed respectively that low socio-economic status of the respondents is a strong determinant of stress among physics education undergraduates. When the mean percentage was computed it was found to be 80%. This implies that low socio-economic status of a family to a large extent is a source of stressor among physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State.

15.5.5 Research Question 5

To what extent will poor time management constitute a source of stress to physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State?

Table 15.5 showed that 10 and 20 respondents strongly agreed and agreed respectively that poor time management constitute a source of stress to physics education

Table 15.5 Poor time management

Variable	No. of respondents	Strongly agree	Agree	Disagree	Strongly disagree	%
Poor time management	40	10	20	7	3	75

undergraduates. When the mean percentage was computed, it was found to be 75%, thus implies that poor time management is a source of stressor among physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State.

15.6 Discussion

From the analysis and interpretation of data, it was found out that inadequate science facilities/accommodation is a source of stress to physics education undergraduates. This finding is in line with the Alika [1] who found that modern day challenges, inadequate laboratories, computers and accommodation inclusive constitutes stressor in our campuses. When physics education undergraduates are faced with a challenge such as inadequate facilities like laboratories, computers, and accommodation, their academic performance may be affected. When this happens, the observation of Maisamari [10] comes into play, he contended that cognitively, stress interfaces with thinking, concentration and memory; this may lead to poor academic achievement.

The study also showed that poor health condition could be a source of physics education undergraduates. This finding is in agreement with the observation of Ikeotunye [6] who noted that physical and health problems constitute sources of stress among individuals. This study also showed that socio-economic status could be a source of stressor among physics education undergraduates, thus finding is in line with the observation of Alika [1] who described family problems as a source of stressor. The finding is also in line with the observation of Maisamari [8] when he described stress as a state of discomfort, tension or emotional pain, which arises when an individual if faced with situations which present a demand that's important for the individual to meet but for which his capacities and resources are inadequate. It is also in line with the finding s of Franklin [5] who found that individuals who suffer extreme hunger as a result of poverty for a long period of time experience a wide spread affect such depression, poor concentration, hostility and irritable disposition, thus on the long run may-effect academic achievement. This study also showed that emotional problems could be a source of stress among physics education undergraduates. This finding is in line with the assertion of Woolfolk [17] who noted that behaviourally, stress is characterized by anxiety, anger, depression, obsession, altered motivation, impaired intellectual activities, and lack of concentration and loss of interest amongst others.

Moreover, the study showed that poor time management is the highest contributor of stress among physics education undergraduates. This finding is in line with that of Maisamari [9] who identified poor time management as a source of stress. This is so because students who fail to manage their time effectively may express some form of stress especially when it is time for test and examination, they find themselves reading all night, in order to cover the course outline, this more often than not gets them stressed, some fall sick and are rushed to the hospital or the university medical centre. This is invariably may lead to poor academic achievement.

15.7 Conclusion

The conclusion drawn from this study is that inadequate laboratories, computers, physics equipments, students' hostel and lecture hall accommodation, poor health condition, emotional problem, low socio-economic status and poor time management constitute stressors to physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University, Anambra State. However, the study showed that low socio-economic status is the highest contributor to stress among physics education undergraduates because 80% of the respondents indicated that poor time management is a major source of stress among physics education undergraduates.

Recommendations

The stake holders of Education, the Federal Government and State Government agencies such as Education trust fund and the University authorities should ensure that adequate science facilities/accommodation are provided for students in university campuses. Health workers should visit the university campuses from time to time so as to ensure that adequate healthy conditions prevail on university campuses. University Health/Medical Centers should be well funded and equipped. Guidance and Counselling services should be provided in order to assist in counseling students with emotional problems and also in providing guidance and counselling services on effective time management to our young undergraduates, especially those in physics education.

Ministry of Education agencies, non-governmental agencies etc. should collaborate among themselves in providing incoming-generating ventures, factories, industries that will serves as a boost to availability of vacation jobs in order to improve the socio-economic status of students who need money in order to meet their social and educational challenges while in the university.

University guidance counsellors should identify indigent physics education undergraduate and help them negotiate some form of financial assistance through scholarship for them or work study programs on campuses. Counselling services should be emphatically rendered to physics education undergraduates and how to effectively manage their time in order to improve on their academic achievement, resolve emo-

tional problems and adjust effectively thereby develop their potential to improve on their academic achievement and contribute towards national development. They should adequately render their counselling services to physics education undergraduates in order to help them manage their stress so as to help them develop functional education.

References

1. I.H. Alike, Sources of stressors among Science Education Undergraduates of University of Benin: The counseling for enhanced academic achievement in science. Afr. J. Sci. Technol. Math. Educ. (AJSTME) **1**(1) (2011)
2. H.I. Alika, E.O. Egbochukwu, Benin J. Gend. Stud. **1**(1), 31–36 (2008)
3. M.S. Edstrom, *Conquering Stress* (Baron's Educational Services Incorporated, New York, 1993), pp. 146–166
4. Federal Republic of Nigeria, *National Policy on Education Yaba*, Rev. ed. (NERD Press, 2004), 8 pp
5. D. Franklin, *Roles of Schools in Mental Health* (Basic Book Inc., New York, 1998)
6. A.I. Ikeotuonye, *Stress Management Kit* (Gwagwalada; Corporate Concept, Abuja, Nigeria, 1993)
7. R. Leandor, *Managing Financial Stress*, April ed. (Awake Publication, 2005), pp. 26–27
8. J.Y. Miasamari, Some stress management technique: implication for counseling. Couns. J. Counsel. Assoc. Niger. (JCASSOXj Mil)
9. J.Y. Miasamari, *Stress and Stress Management Strategies* (Joyce Printers and Publishers, Abuja, Nigeria, 2002), pp. 7–19
10. G. Natriello, At-risk student, in *Education and Sociology: An Encyclopedia*, ed. by D.L Levinson, P.W. Cookson Jr., A.R. Sadovnik (Routledge Falmer, New York, 2002), pp. 49–54
11. T.U. Okafor, *Readability of Recommended Physics Textbook in Anambra State Secondary School*, An unpublished Ph.D. thesis, University of Nigeria, Nsukka, 2009
12. T.U. Okafor, Sources of stressors among physics education undergraduates of Chukwuemeka Odumegwu Ojukwu University, in *Proceeding of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Fracisco U.S.A. Lecture Note in Engineering and Computer Science (2017), pp. 232–237
13. N.P. Oyemerekeya, A review of some common methods of stress management. The counselor. J. Couns. Assoc. Niger. **14**(2) (1996)
14. S. Sirin, Socio-economic status and academic achievement. A meta-analytic research. Rev. Educ. Res. **75**(3), 417–453 (2005)
15. W. Thompson, J. Hickey, *Society in Focus* (Pearson, Boston, 2008)
16. UNICEF, The state of world children. Girls' Educ. Dev. **14**, 34–35 (2004)
17. E. Woolfolk, *Education Psychology* (Allyn and Bacon, Boston, 1995)

Chapter 16
Reactive Power Loss Minimization on an Interconnected Electric Power Network

Uche Chinweoke Ogbuefi, Boniface Onyemaechi Anyaka and Muncho Josephine Mbunwe

Abstract The inability of power system to maintain a proper balance of reactive power is the major cause of voltage collapse. A system can be saved from voltage collapse by reducing the reactive power load or by adding additional reactive power into the system. The electric power system is afflicted with continuous load shedding due to inadequate generation and transmission capacities. To maximize the amount of real power that can be transferred over a network, reactive power flow must be minimized. Thus, sufficient reactive power should be provided locally in the system to keep bus voltages within stipulated ranges to satisfy customers' equipment ratings. This paper presents an overview in reactive power compensation skills which remains as research challenges in this area. Newton-Raphson's solution method was used to carry out the analysis because of its fast convergence, sparsity, and simplicity attributes when compared to other solution methods, with relevant data obtained from Power Holding Company of Nigeria (PHCN). MATLAB/SIMULINK was used to carry out the simulation analysis. It is observed that the application of compensation on the unified system jointly has effect on the other buses. This is confirmed by a step-by-step application of compensation at 5% intervals. The effects were noticed in Bus (20) where voltage decreased from 0.9568 to 0.9329 p.u. about 2.39%, bus (19) from 0.998 to 1.1035 p.u. and others. These results indicate undershoot and overshoot that will cause damage to the system, and may lead to system collapse if no contingency control is installed. It is also observed that compensation should be done on weak buses only for better results. The results indicate the enhancement in voltage profile in addition to reduction in the network losses and more balanced system. Active and reactive power control greatly influence the electricity grid, thus, need adequate attention with the recent advent of integration of renewable energy into the grid.

U. C. Ogbuefi · B. O. Anyaka · M. J. Mbunwe (✉)
Department of Electrical Engineering, University of Nigeria, Nsukka, Enugu State, Nigeria
e-mail: muncho.mbunwe@unn.edu.ng

U. C. Ogbuefi
e-mail: uche.ogbuefi@unn.edu.ng

B. O. Anyaka
e-mail: boniface.anyaka@unn.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_16

Keywords Active power · Electric power · Interconnected network
Loss minimization · Power system · Reactive power · System collapse

16.1 Introduction

The demand of electricity is increasing day-by-day and this results in an increase in per unit production cost. Again, maintaining uninterrupted power flow and providing quality power during disturbance becomes challenging. Voltage ampere reactive (VAR) compensation is the management of reactive power to improve the performance of ac power systems. The concept of VAR compensation embraces a wide and diverse field of both system and customer problems, especially related with power quality issues since most power quality problems are attenuated or solved with an adequate control of reactive power [1].

In general, the problem of reactive power compensation is viewed from two aspects: load compensation and voltage support. In load compensation the objectives are to increase the value of the system voltage (power factor improvement), balance the real power drawn from the ac supply, compensate voltage regulation and eliminate current harmonics. In voltage support, the idea is for sustenance and to maintain stable voltage flow in the network. For power flow studies the frequency should remain nearly constant, because significant drop in frequency could result in high magnetizing currents in induction motors and transformers [2, 3]. The flows of active and reactive powers in a transmission network are fairly independent of each other and are influenced by different control actions. Active power control is closely related to frequency control, and reactive power control is closely related to voltage control [2, 3]. Since constancy of frequency and voltage are important factors in determining the quality of power supply, then the control of active and reactive power is vital to the satisfactory performance of a power system [3, 4].

Since electrical energy is normally generated at the power stations far away from the urban areas where consumers are located and are delivered to the ultimate consumers through a network of transmission and distribution, the terminal voltage vary substantially. Wider variation in voltage may cause erratic operation or even malfunctioning of consumers' appliances. The main cause for voltage variation is the variation in load on the supply system. With the increase in load on the supply system the voltage at the consumer premises falls due to increase in voltage drop in: (I) Alternator synchronous impedance, (ii) Transmission lines, (iii) Feeders and (iv) Distributors [5–7].

A well designed power system is one that gives good quality and reliable supply. By good quality it means the voltage levels are within reasonable limits. Naturally all the equipment on the power system are designed to operate satisfactorily only when the voltage levels in the system correspond to the rated voltage or at the most the variation are within $\pm 5\%$ of rated value [7]. Thus, compensation could be beneficial in in this aspect. The benefits of compensations are enormous and include the following: reactive power compensation in a transmission system improves the stability of the ac

system by increasing the maximum active power that can be transmitted. It also helps to maintain a substantially flat voltage profile at all levels of power transmission if properly harnessed. It also increases transmission efficiency. It controls steady-state and temporary over-voltages and can avoid disastrous blackouts [7–10]. Objectively, reactive power loss minimization and the study of the effect of joint compensation on an interconnected network are the main issue of this work and the result obtained showed the disparity.

In an interconnected power system of n-buses the power injected into the nodes is given by a set of 2n nonlinear simultaneous equations represented in Eqs. (16.1) and (16.2).

$$P_i = \sum_{n=1}^{N} |V_i V_k Y_{ik}| \cos(\theta_{ik} + \partial_k - \partial_i) \tag{16.1}$$

$$Q_i = -\sum_{n=1}^{N} |V_i V_k Y_{ik}| \sin(\theta_{ik} + \partial_k - \partial_i) \tag{16.2}$$

Both active and reactive powers of the loads vary as a function of voltage magnitude. Compensating devices are normally added to supply or absorb reactive power and thereby control the reactive power balance in a desired form [2].

16.1.1 Reactive Power Flow on Line Voltage Drop

Voltage variation is due to imbalance in generation and consumption of reactive power in the system. If the generated reactive power is more than the consumed power, the voltage levels go up and vice versa. However, if the two are equal, then the voltage profile becomes flat and it happens only when the load is equal to the natural load. Unfortunately the reactive power in the system keeps varying because of different types of load and if the reactive power generation is simultaneously controlled, a more or less flat voltage profile could be maintained. Too wide variation of voltage causes excessive heating of distribution transformers thereby reducing the transformer capacity [2, 9].

16.2 Compensation Plans

The device commonly used for voltage and reactive power control includes: shunt reactors, shunt capacitors, synchronous generators, synchronous condensers, FACTS DEVICES etc.

Shunt Reactors are inductive current element connected between lines and neutral to compensate for inductive load/(current) from transmission lines. They are used

to compensate for the effects of line capacitance, particularly to limit voltage rise on open circuit or light load [2, 6]. Shunt Capacitors supply reactive power and boost local voltages. The major advantages of shunt capacitors are their low cost and their flexibility of installation and operation [2]. They are readily applied to various points in the system, thereby contributing to efficiency of the power transmission and distribution. The disadvantage is that their reactive power output is proportional to the square of the voltage. The reactive power output is reduced at low voltages when it is likely to be needed most [2, 11, 12]. Shunt capacitors alleviate the excess loading of system and enable further active loads to be drawn from the same system. The reduction in line currents results in reduction of system losses. It is economical to carry out reactive power compensation from capacitors ideally located in the vicinity of reactive load e.g. motors and low voltage side of transformers.

16.2.1 Real and Reactive Power Control

A synchronous machine that is connected to an infinite bus has fixed speed and terminal voltage. The control variables are the field current and the mechanical torque on the shaft. The variation of the field current (I_f), referred to as excitation system control is applied to either a generator or a motor to supply or absorb a variable amount of reactive power. Since the synchronous machine runs at a constant speed, the only means of varying the real power is through control of the torque imposed on the shaft by either the prime mover in the case of a generator or the mechanical load in the case of a motor. The complex power delivered to the system by the generator is given in per unit as in Eqs. (16.3) and (16.4)

$$S = P + jQ = V_i I_a^* = |V_i||I_a|(\cos\,\theta + j\,\sin\,\theta) \tag{16.3}$$

And for real and imaginary quantities we obtain

$$P = |V_i||I_a|\cos\,\theta; \quad Q = |V_i||I_a|\sin\,\theta \tag{16.4}$$

16.2.2 Shunt Compensation—Static-Var Compensation

Shunt compensation is the use of shunt capacitor or and shunt reactors in the line to avoid or reduce voltage instability [1, 2, 8] Shunt compensators are connected in shunt either directly to a bus bar or to the tertiary winding or to the main transformer and sometimes at mid-point of the lines (in some countries) to minimize the voltage drop and the losses. Shunt compensators are installed near the local terminals in factory substations, in the receiving substations, at switching substations etc to provide leading volt ampere-reactive (MVar) and thus to reduce the line current and total kVA loading of substation transformer [4, 6, 7, 13].

16.2.3 Control of Voltage and Reactive Power

For a transmission line where X ≫ R and R is negligibly small, therefore

$$|\Delta V| = \frac{XQ_r}{X}, \quad Q_r = \frac{V_r}{X}|\Delta V| \tag{16.5}$$

This relationship shows that the reactive power Q_r is proportional to the magnitude of the voltage drop in the line. Thus voltage control and reactive power control are interrelated. The reactive power generated should be exactly equal to the reactive power consumed. Any mismatch in the reactive power balance affects the bus voltage magnitudes [1, 6, 7].

16.2.4 Reactive Power Compensation in the Nigeria 330 kV Network

VAr compensation is the management of reactive power to improve the performance of ac power systems. The concept of VAr compensation encircles a wide and diverse field of both system and customer problems, especially related with power quality issues. Most of the power quality problems can be attenuated or solved with an adequate control of reactive power.

System voltage is highly dependent on the flow of reactive power. The long transmission lines in the National Grid generate considerable reactive MVars which constitute serious problems in maintaining system voltages within statutory limits especially during light loads, system disturbances and or major switching. The Nigerian PHCN has many reactors in various locations in the country, some of which are shown in Table 16.1. Some of these reactors were incorporated in the system to carry out the compensation to control the effect of reactive Mar. The major cause of voltage variation or drop in the line is the flow of reactive power. More of over reactive currents causes I^2R losses in the system but produces no revenue.

16.2.5 Reactive Power Management in Electric Power System

An important factor in the control of voltage in a power system depends on the reactive power production or absorption. Reactive power is required to excite consumer's equipment and transmission network which consists of capacitive and inductive elements. It is important that a balance of reactive power be maintained in the operation of a system because control of voltage can be lost if this is not achieved [2, 3, 14]. The reactive power flow is minimized so as to reduce I^2R and (I^2X) losses to a practical minimum. This ensures that the transmission system operates efficiently. The rating of capacitor can be calculated with the simplified equation as;

Table 16.1 Status of reactors in Nigeria power network PHCN 330 kV system

Station	Reactor nomenclature	Rating		Remarks
		kV	MVAR	
Kaduna	3R3	330	75	Good
Jebba	2R1	„	75	Good
Kano	R1	„	75	„
Gombe	R1	„	50	„
	R2	„	50	„
Oshogbo	4R1	„	75	„
Benin	6R2	„	75	„
Ikeja-west	R1	330	75	Good
Makurdi	R1	330	75	Good
Maiduguri	2R1	330	75	Good

$$C = \frac{Q_C}{\omega V^2} \tag{16.6}$$

Equation 16.5 shows that the capacitance required to improve the system efficiency is inversely proportional to $\boldsymbol{V}^2$. Note that at high voltages power capacitors or capacitor bank values are rated in kilo volt-ampere reactive (kvar or mvar). For three phase system, the equation for the capacitor in delta connection, where ($V_P = V_L$)

Is given by Eq. 16.7.

$$C_\Delta = \frac{Q_C}{\omega V_P^2} = \frac{Q_C}{\omega V_L^2} \tag{16.7}$$

Compensation added to the network is given by Eq. 16.8, [7, 15]

$$Q_C = \frac{P}{Pf_1} \times \sin(\cos^{-1}(pf_1)) - \frac{P}{Pf_2} \sin(\cos^{-1}(pf_2) \tag{16.8}$$

where P = real power specified at the buses, Pf_1 = 0.85 power factor, Pf_2 = 0.95 power factor, Q_C = value of shunt capacitance to be added to the network to boost the system voltage. Hence the capacitor required per three phase in star connection is equal to three times the capacitance required per phase when the capacitors are connected in delta. Also, the capacitors for the star-connected bank have a working voltage equal to $\frac{1}{\sqrt{3}}$ times that for the delta-connected bank. For this reason, the capacitors are connected in delta in three-phase systems for improvement of the system stability. The installation of a capacitor bank can be used to avoid the need to change a transformer in the event of a load increase. System behavior is affected by the characteristics of every major element of the system. The representation of these elements by means of appropriate mathematical models is critical to the successful analysis of the system behavior [16, 17].

In an ac circuit, current 'I' leads voltage in a capacitive circuit and I lags voltage in an inductive circuit. A capacitor is used to generate Var and an inductor is used to reduce reactive power. In power system analysis there is a strong inter relationship between the real power and angle and also between the reactive power and the voltage magnitude which aids us in compensation analysis. These relationships below exist for real and reactive powers flow.

$$S = P + jQ,\ P = VI\text{Cos}\,\theta \text{ or } P = I^2R = \frac{V^2}{R},$$
$$\text{and } Q = VI\text{Sin}\,\theta \text{ or } Q = I^2X = \frac{V^2}{X}\,X_C = \frac{1}{j\omega C}$$

and with low Var in a system, the voltage becomes low and vice versa. These relationships are used for the result analysis [8, 9, 18].

16.3 Network Description

The Nigerian power network like many practical systems in developing countries consists of a few generating stations mostly sited in remote locations near the raw fuel sources which are usually connected to the load centers by long transmission lines.

The National Electric Power Authority (NEPA) now known as Power Holding Company of Nigeria (PHCN) has the sole statutory functions of generation, transmission, distribution and marketing of electricity, before the partial unbundling of the power sector. Nigeria national electricity grid at present consists of nine generating stations comprising of three (3) hydro and six (6) thermal plants with a total installed generating capacity of 6500 MW. The thermal stations are mainly in the southern part of the country located at Afam, Okpai, Delta (Ughelli), Egbin and Sapele. The hydroelectric power stations are in the country's middle belt and are located at Kainji, Jebba and Shiroro. The transmission network is made up of 5000 km of 330 kV lines, 6000 km of 132 kV lines, 23 of 330/132 kV sub-stations and 91 of 132/33 kV sub-stations [7, 9, 19].

Although, the installed capacity of the existing power stations is 6500 MW the maximum load ever recorded was 4000 MW. Presently, most of the generating units have broken down due to limited available resources to carry out the needed level of maintenance. The transmission lines are radial and overloaded. The switchgears are obsolete while power transformers have not been maintained for a long time. The present installed generating capacity in Nigeria is shown in Table 16.2. The PHCN has only once been able to generate a maximum of 4700 MW, for a country of more than 160 million people [20–22].

Table 16.2 Existing power stations

S/no.	Power station name	Location/state	Status	Capacity (MW)
1	Egbin Thermal Power Station	Lagos	Operating	1320
2	Afam Thermal PS	Rivers	Operating	969.6
3	Sapele Thermal PS	Delta	Operating	1020
4	Ijora Thermal PS	Lagos	Operating	40
5	Delta Thermal PS	Delta	Operating	912
6	Kainji Hydro PS	Niger	Operating	760
7	Jebba Hydro PS	Niger	Operating	578
8	Shiroro Hydro PS	Niger	Operating	600
9	AES Thermal PS	Lagos	Operating	300
Total capacity =				6500

16.3.1 Methodology

In this work, the method used is based on the work of [1]. The existing 330 kV, 30 bus system of Nigeria transmission network with Egbin power station as the slack bus is used, and an in-depth examination of the Nigeria Integrated Power Plant Network was carried out. The parameters of all the generators and other system components were obtained. Equations for the power flow analysis are then formulated incorporating these parameters. The algorithm for the Newton-Raphson's method was developed. The Newton-Raphson's solution method represented with Eqs. (16.9) and (16.10) was used to carry out the analysis because of its sparsity, fast convergence and simplicity attributes as compared to other solution methods using the relevant data as obtained from Power Holding Company of Nigeria (PHCN). MATLAB m-file program and SIMULINK model were developed and used for the simulation analysis.

$$\Delta \mathrm{P} = J_{11} \Delta \delta \tag{16.9}$$

$$\Delta Q = J_{22} \Delta |V| \tag{16.10}$$

16.3.1.1 Line Flows and Line Losses Model

At the end of the iterative solution of the bus voltages, the calculation of the line flows and losses for both compensated and uncompensated conditions were carried out. The following formulae are used considering a two bus system labeled 1 and 2 (or i and j, and then to nth buses). First the Line current is given as:

$$I_{12} = y(V_1 - V_2), \quad I_{21} = -I_{12} \tag{16.11}$$

$$I_{13} = y_{13}(V_1 - V_3), \quad I_{31} = -I_{13} \tag{16.12}$$

Line flows are calculated using

$$S_{12} = V_1 I_{12}^*; S_{21} = V_2 I_{21}^* \tag{16.13}$$

$$S_{13} = V_1 I_{13}^*; \quad S_{31} = V_3 I_{31}^* \tag{16.14}$$

Line losses

$$S_{L12} = S_{12} + S_{21}; \quad S_{L13} = S_{13} + S_{31} \tag{16.15}$$

16.3.2 Draft of Nigeria 330 kV Transmission Network Used as Case Study

The single-line diagram of the existing 330 kV Nigeria transmission network used as the case study is as shown in Fig. 16.1. It has 30 buses with nine generating station. The Egbin power station was chosen as the slack bus because of its capacity and location in the network.

16.3.3 Data Assembly—Line Data

The input data for the power flow analysis include the bus data, transmission line data (impedance of lines), voltages and transformer/load data obtained from Power Holding Company of Nigeria (PHCN) are as presented in Tables 16.3, 16.4 and 16.5.

The load and generation data expressed in per unit values are given as $\frac{MW+jMVAr}{base.value}$. where the Slack Bus is Egbin Generating Station. As in Table 16.4

Base value = 100 MVA

Base voltage = 330 kV, Per Unit Value = $\frac{MVA}{Base\ Voltage}$ as presented in Table 16.4.

16.3.4 Shunt Capacitor Compensation Algorithm

The flow chart in Fig. 16.2 is the procedural method applied to achieve the desired compensation results.

First, the base solution is obtained using Newton-Raphson's method. Check bus voltages range. Identify the problem buses by checking the bus voltages outside ±5% of the normal values (that is, 0.95–1.05) per unit. Calculate the capacitor values

Table 16.3 Transmission line data (of Bison, two conductors per Phase and 2 × 350 mm^2 X-section Conductor) for 330 kV lines

B/no.	Bus name		Length (km)	R_1 (p.u.)	X_1 (p.u.)	Shunt
	From	To				
1	Akamgbe	Ik-West	17	0.0006	0.0051	0.065
2	Ayede	Oshogbo	115	0.0041	0.0349	0.437
3	Ik-West	Egbin	62	0.0022	0.0172	0.257
4	Ik-West	Benin	280	0.0101	0.0799	1.162
5	Oshogbo	Jebba	249	0.0056	0.477	0.597
6	Oshogbo	Benin	251	0.0089	0.0763	0.954
7	JebbaTs	JebbaGs	8	0.003	0.0022	0.033
8	Jebba TS	Shiroro	244	0.0087	0.0742	0.927
9	Jebba TS	Kainji	81	0.0022	0.0246	0.308
10	Kainji	B.Kebbi	310	0.0111	0.942	1.178
11	Shiroro	Kaduna	96	0.0034	0.0292	0.364
12	Kaduna	Kano	320	0.0082	0.0899	0.874
13	Jos	Gombe	265	0.0095	0.081	1.01
14	Benin	Ajaokuta	195	0.007	0.056	0.745
15	Benin	Sapele	50	0.0018	0.0139	0.208
16	Benin	Onitsha	137	0.0049	0.0416	0.521
17	Onitsa	N.Heaven	96	0.0034	0.0292	0.0355
18	Onitsha	Alaoji	138	0.0049	0.0419	0.524
19	Alaoji	Afam	25	0.009	0.007	0.104
20	Sapele	Aladja	63	0.0023	0.019	0.239
21	Delta	Aladja	30	0.0011	0.0088	0.171
22	Kainji GS	Jebba TS	81	0.0022	0.0246	0.308
23	Ayede	lk West	137	0.0049	0.0416	0.521
24	Egbin TS	Aja	27.5	0.0022	0.0172	0.257
25	Egbin TS	Aja	27.5	0.0022	0.0172	0.257
26	Kaduna	Jos	197	0.007	0.0599	0.748
27	Jos	Makurdi	275	0.0029	0.0246	0
28	Oshogbo	lk West	252	0.0049	0.0341	0.521
29	Benin	Delta	107	0.0022	0.019	0.239
30	Onitsha	Okpai	80	0.009	0.007	0.104
31	Geregu	Ajokuta	5	0.0022	0.0172	0.257
32	Shiroro	Kaduna	96	0.0034	0.0292	0.364

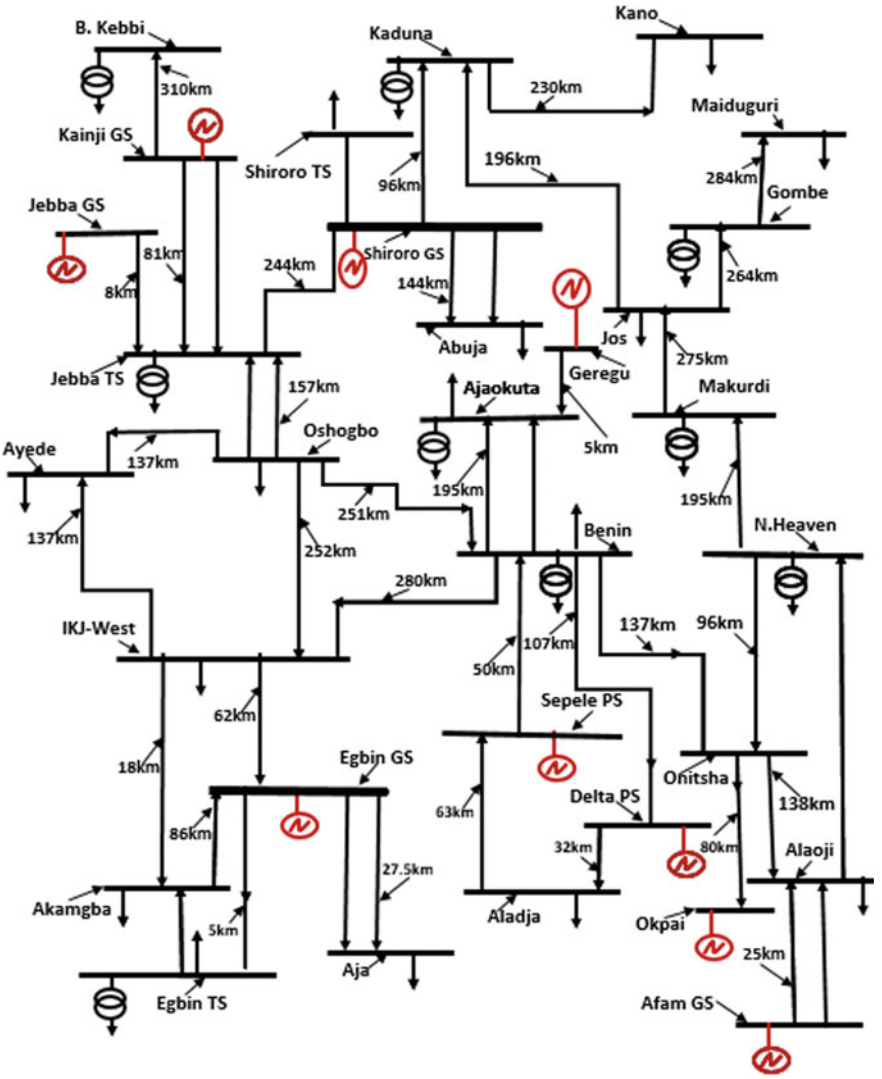


Fig. 16.1 One line diagram of the PHCN 330 kV 30 bus interconnected network

using this equation $\left(C = \frac{Q_C}{\omega V^2}\right)$ and apply compensation using this $Q_C = \frac{P}{Pf_1} \times \sin(\cos^{-1}(pf_1)) - \frac{P}{Pf_2} \sin(\cos^{-1}(pf_2))$. Where P is real power specified at the buses, Pf_1 and Pf_2 are power factors, while Q_C is value of shunt capacitance to be added to the network to boost the system voltage. Finally output result and stop. These procedures were simulated using MATLAB/SIMULINK. The results from the Newton-Raphson iterative method give the bus voltages, line flows, and power losses under normal (uncompensated) condition as shown in Table 16.7. The voltages at buses 14, 17, 18,

Table 16.4 Bus data values

B/no.	Bus name	Generation		Load		V (volts)	Angle (degree)	Remarks
		P (MW)	Q (MVar)	P (MW)	Q (MVar)			
1	Egbin	–	–	0.0000	0.0000	1.02	0.0000	Slack
2	Delta Ps	55.000	28.160	–	–	1.0000	0.0000	PV bus
3	Okpai	220.000	112.700	–	–	1.0000	0.0000	PV bus
4	Sapele	75.000	38.420	–	–	1.0000	0.0000	PV bus
5	Afam	479.000	245.390	–	–	1.0000	0.0000	PV bus
6	Jebba	322.000	164.960	–	–	1.0000	0.0000	PV bus
7	Kainji	323.000	165.490	–	–	1.0000	0.0000	PV bus
8	Shiroro	280.000	143.440	–	–	1.0000	0.0000	PV bus
9	Geregu	200.000	102.440	–	–	1.0000	0.0000	PV bus
10	Oshogbo	–	–	120.370	61.650	1.0000	0.0000	Load bus
11	Benin	–	–	160.560	82.240	1.0000	0.0000	Load bus
12	Ikj-West	–	–	334.000	171.110	1.0000	0.0000	Load bus
13	Ayede	–	–	176.650	90.490	1.0000	0.0000	Load bus
14	Jos	–	–	82.230	42.129	1.0000	0.0000	Load bus
15	Onitsha	–	–	130.510	66.860	1.0000	0.0000	Load bus
16	Akamgbe	–	–	233.379	119.560	1.0000	0.0000	Load bus
17	Gomgbe	–	–	74.480	38.140	1.0000	0.0000	Load bus
18	Abuja (katamkpe)	–	–	200.000	102.440	1.0000	0.0000	Load bus
19	Maiduguri	–	–	10.000	5.110	1.0000	0.0000	Load bus
20	Egbin TS	–	–	0.000	0.000	1.0000	0.0000	Load bus
21	Aladja	–	–	47.997	24.589	1.0000	0.0000	Load bus
22	Kano	–	–	252.450	129.330	1.0000	0.0000	Load bus
23	Aja	–	–	119.990	61.477	1.0000	0.0000	Load bus
24	Ajaokuta	–	–	63.220	32.380	1.0000	0.0000	Load bus
25	N.Heaven	–	–	113.050	57.910	1.0000	0.0000	Load bus
26	Alaoji	–	–	163.950	83.980	1.0000	0.0000	Load bus
27	Jebba TS	–	–	7.440	3.790	1.0000	0.0000	Load bus
28	B.Kebbi	–	–	69.990	35.850	1.0000	0.0000	Load bus
29	Kaduna	–	–	149.77	76.720	1.0000	0.0000	Load bus
30	ShiroroTS	–	–	73.070	37.430	1.0000	0.0000	Load bus

19, 22, 29 and 30 are outside the limit, and in order to ensure that they are within acceptable limits shunt capacitive compensation were introduced into the buses. Based on Power Holding Company of Nigeria (PHCN) power factor of 0.85 and 0.95 for transmission lines are used. The MVAr capacities of the various capacitors required to carry out compensation of the network at the buses were determined using

$$Q_C = \frac{P}{Pf_1} \times \sin(\cos^{-1}(pf_1)) - \frac{P}{Pf_2} \sin(\cos^{-1}(pf_2))$$

Table 16.5 Line data

S/no.	Circuit (Buses)		Length (km)	Impedance Z (p.u.)	Admittance Y (p.u.)	Shunt $\frac{y}{2}$ (p.u.)
	From	To				
1	Akamgbe	Ik-West	17	0.0006 + j0.0051	22.75 − j19.32	0.065
2	Ayede	Oshogbo	115	0.0041 + jo.o349	3.333 − j38.37	0.437
3	Ik-West	Egbin	62	0.0022 + j0.0172	7.308 − j57.14	0.257
4	Ik-West	Benin	280	0.0101 + j0.0799	1.637 − j12.626	1.162
5	Oshogbo	Jebba	249	0.0056 + j0.477	0.0246 − j3.092	0.597
6	Oshogbo	Benin	251	0.0089 + j0.0763	1.508 − j12.932	0.954
7	Jebba TS	Jebba GS	8	0.003 + j0.0022	3.174 − j1.594	0.033
8	Jebba TS	Shiroro	244	0.0087 + j0.0742	1.559 − j13.297	0.927
9	Jebba TS	Kainji	81	0.0022 + j0.0246	3.607 − j40.328	0.308
10	Kainji	B.Kebbi	310	0.0111 + j0.942	1.235 − j10.478	1.178
11	Shiroro	Kaduna	96	0.0034 + j0.0292	3.935 − j3.379	0.364
12	Kaduna	Kano	320	0.0082 + j0.0899	1.657 − j14.17	0.874
13	Jos	Gombe	265	0.0095 + j0.081	1.923 − j15.456	1.01
14	Benin	Ajaokuta	195	0.007 + j0.056	1.429 − j12.180	0.745
15	Benin	Sapele	50	0.0018 + j0.0139	3.194 − j17.555	0.208
16	Benin	Onitsha	137	0.0049 + j0.0416	2.8 − j33.771	0.521
17	Onitsa	N.Heaven	96	0.0034 + j0.0292	3.935 − j3.379	0.0355
18	Onitsha	Alaoji	138	0.0049 + j0.0419	2.754 − j33.553	0.524
19	Alaoji	Afam	25	0.009 + j0.007	59.230 − j53.846	0.104
20	Sapele	Aladja	63	0.0023 + j0.019	5.284 − j51.913	0.239
21	Delta	Aladja	30	0.0011 + j0.0088	13.995 − j1.119	0.171
22	Kainji GS	Jebba TS	81	0.0022 + j0.0246	3.607 − j40.328	0.308
23	Ayede	lk West	137	0.0049 + j0.0416	2.8 − j33.771	0.521
24	Egbin TS	Aja	27.5	0.0022 + j0.0172	7.316 − j57.2036	0.257
25	Egbin TS	Aja	27.5	0.0022 + j0.0172	7.316 − j57.2036	0.257
26	Kaduna	Jos	197	0.007 + j0.0599	1.924 − j16.469	0.748
27	Jos	Makurdi	275	0.0029 + j0.0246	4.726 − j40.093	0
28	Oshogbo	lk West	252	0.0049 + j0.0341	4.128 − j28.732	0.521
29	Benin	Delta	107	0.0022 + j0.019	6.013 − j51.935	0.239
30	Onitsha	Okpai	80	0.009 + j0.007	59.230 − j53.846	0.104
31	Geregu	Ajokuta	5	0.0022 + j0.0172	7.316 − j57.203	0.257
32	Shiroro TS	Kaduna	96	0.0034 + j0.0292	3.935 − j3.379	0.364

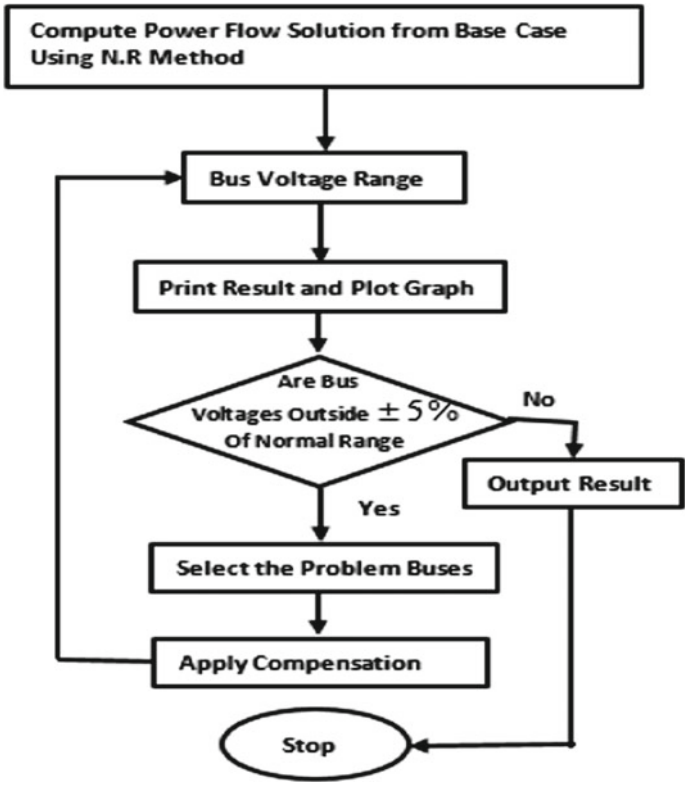


Fig. 16.2 Flow chart for the shunt capacitor compensation analysis algorithm

The following capacitor sizes were selected for the various lines. Jos bus (30 MVAr), Gombe bus (30 MVAr), Abuja bus (60 MVAr), Kano bus (40 MVAr), Kaduna bus (40 MVAr), and Makurdi bus (30 MVAr). These were introduced into the network to examine their effect on the system. The weak buses were identified as presented in Table 16.6, and the plots of the results are as shown in Figs. 16.3 and 16.4.

16.4 Results

It is also recorded during the compensation of the entire system jointly, that some buses that were normal are affected. Some buses values decreased from tolerable values while some over increased. Some of the pictorial graphs were as presented in Fig. 16.5a–f.

Table 16.6 Bus voltages for compensated and uncompensated

B/no.	Bus name	Bus vtgs. with compensation Volts (p.u.)	Without compensation Volts (p.u.)
1	Egbin-GS (Slack)	1.0000	1.0000
2	Delta-PS	1.0000	1.0000
3	Okpai-PS	1.0000	1.0000
4	SAP/PS	1.0000	1.0000
5	AFAM-GS	1.0000	1.0000
6	Jebba-GS	1.0000	1.0000
7	KAINJI-GS	1.0000	1.0000
8	Shiroro-PS	1.0000	1.0000
9	Geregu (PS)	1.0000	1.0000
10	Oshogbo	1.0035	0.9919
11	Benin	0.9998	0.9957
12	Ikeja-West	0.9969	0.993
13	Ayede	0.9967	0.9792
14	Jos	0.9823	0.8171
15	Onitsha	0.9793	0.9748
16	Akangba	0.9931	0.9859
17	Gombe	1.0242	0.8144
18	Abuja (Katampe)	0.9667	0.9402
19	Maiduguri	1.0455	0.8268
20	Egbin TS	0.9469	0.9816
21	Aladja	1.0006	0.9994
22	Kano	0.9338	0.7609
23	Aja	0.9692	0.9838
24	Ajaokuta	0.9999	0.9997
25	N.Heaven	0.9721	0.9582
26	Alaoji	0.9598	0.9564
27	Jebba-TS	0.9993	0.9988
28	B.Kebbi	1.0075	0.9873
29	Kaduna	0.9654	0.8738
30	Makurdi	0.9943	0.8247

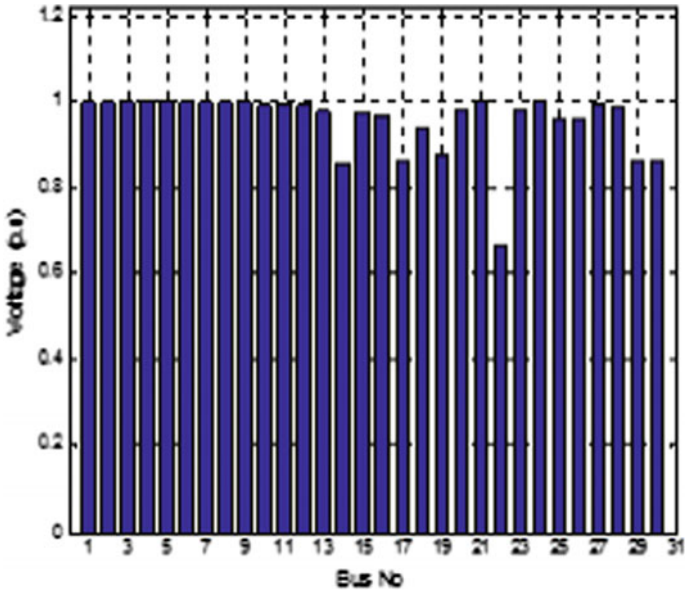


Fig. 16.3 Plot of bus voltages under normal (un compensated) condition

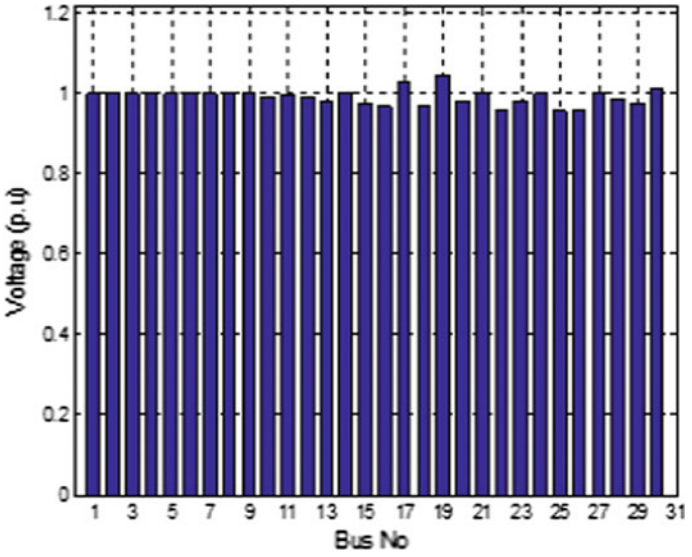


Fig. 16.4 Bar plot of bus voltages with compensation

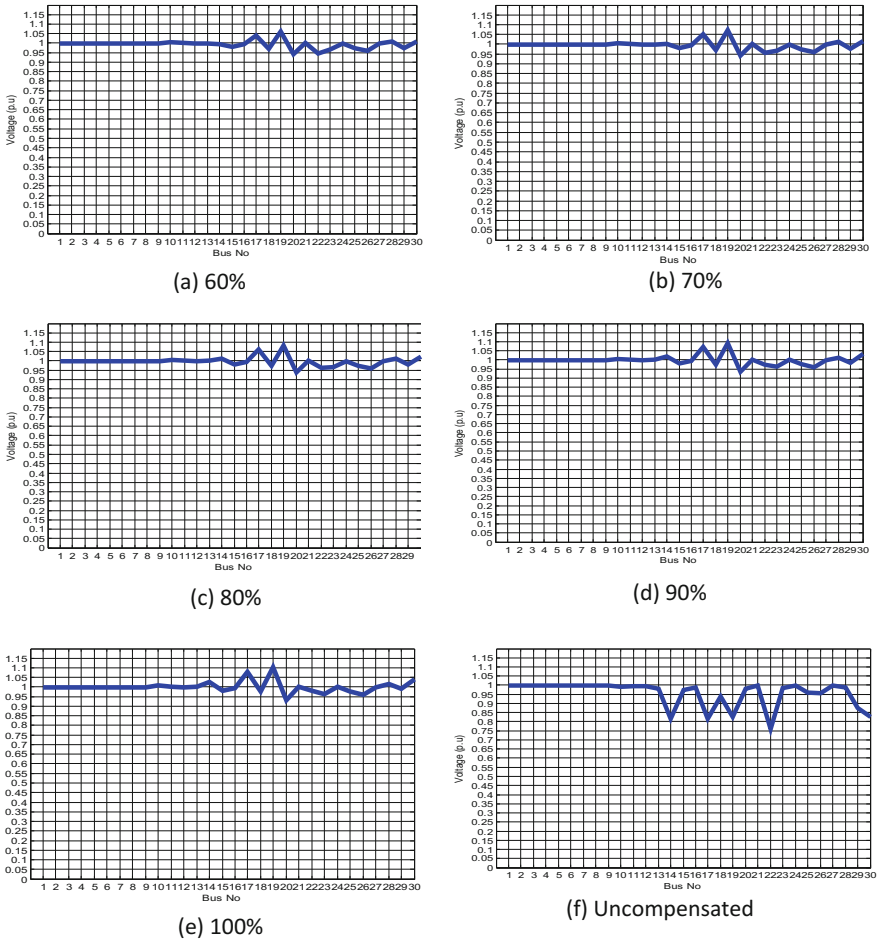


Fig. 16.5 **a–f** Graph of voltage versus bus nos. at different levels of percentage compensation

16.4.1 Discussions

The analysis of Nigeria 330 kV 30 bus network using Newton-Raphson's power flow solution algorithm with MATLAB/SIMULINK software was effectively completed. The results obtained revealed the weak buses with values outside the standard limit of 0.95 p.u. (313.5 kV) and 1.05 p.u. (346.kV). The recorded results are; Bus 14 (Jos) with value 0.8171 p.u., bus 17 (Gombe) 0.8144 p.u bus 18 (Abuja) 0.9402 p.u., bus 19 (Maiduguri) 0.8268 p.u., bus 22 (Kano) 0.7609 p.u., bus 29 (Kaduna) 0.8738 p.u., and bus 30 (Makurdi) 0.8247 p.u. under normal uncompensated condition as presented in Fig. 16.4.

The compensation was carried out on the weak buses. At 45% capacitive shunt compensation on those buses indicated improved performance, Kano and Jos were

still at the weak positions due to their distances in the national grid. With sixty (60) per cent compensation a better result was recorded as buses 14 (Jos) improved to 0.9823 and 22 (Kano) 0.9338. The compensated results are as displayed in Fig. 16.5a–f. It was observed that the application of compensation on the entire system jointly has negative effect on the other buses—which is part of the main aim of this work. This was verified by a step-by-step application of compensation at 5% intervals. **It was observed that compensating the whole network jointly affects some of the other buses that were within the tolerable range**. For instance, at Bus (20) the value decreased from 0.9568–0.9329 p.u. about 2.39% decrease. This can cause damage to the "**system's equipment and consumers' appliance**" if no proper security for contingency analysis control was installed. Also, bus (17) increased from 0.9786 to 1.0799 p.u. and bus (19) from 0.998 to **1.1035** p.u. and so on, which show undershoot and overshoot respectively which may lead to system collapse if not monitored. System efficiency was improved from 65% (uncompensated) to 85% after compensation.

16.5 Conclusion

In this work compensation of interconnected electric power network was studied using Shunt capacitive compensation. Shunt capacitive compensation is widely recognized as one of the prevailing methods to combat the problems of voltage drops, power losses, and voltage flicker in power system networks. Though each compensating technique has its area of proficiency and limit of application, but shunt capacitor compensation method was used because of its outstanding performance especially in long transmission lines and its control of reactive power flow. Though they associated with high cost implication but control voltage *directly* and also control temporary over voltage fast. **It was observed also that application of compensation on the interconnected system jointly has side effect on the other buses**. The results indicated that control of active and reactive power has high influence on the Nigeria network, hence adequate attention must be placed on it. Also with innovation/advent of renewable energy integration into the grid, if adequate control measure of reac-

Table 16.7 Bus voltages at various compensation levels for the compensated buses

B/no.	50%	65%	75%	90%
14	0.964547	0.976966	0.984707	0.995664
17	0.995396	1.009312	1.017975	1.030226
18	0.954925	0.954925	0.954925	0.954925
19	1.014868	1.029198	1.038118	1.05073
22	0.810757	0.834761	0.849712	0.945864
29	0.928746	0.93879	0.945069	0.953981
30	0.975601	0.98823	0.9961	1.007241

Table 16.8 Line current, line flow, and line losses (uncompensated)

B/sequence		Line current (p.u.)		Line flows (p.u.)		Line losses (p.u.)	
From	To	Real	Imaginary	Real	Imaginary	Real	Imaginary
16	12	0.1842	−1.4405	0.1783	−1.3941	−0.0046	0.0363
12	1	0.0514	−0.4023	0.05109	−0.3994	0.0511	−0.3995
12	11	0.0047	−0.0364	0.0047	−0.0361	−0.00001	0.0001
12	13	−0.0386	0.3280	−0.0384	0.3257	−0.0005	0.0045
13	10	0.0423	−0.3605	0.0415	−0.3530	−0.0005	0.0046
10	11	0.0059	−0.0512	0.0059	−0.0507	−0.00002	0.0002
10	27	0.0166	−0.1417	0.0165	−0.1406	−0.0001	0.0009
12	6	1.5244	−1.1179	1.5137	−1.1101	−0.011	0.0078
27	8	0.0016	−0.0177	0.0017	−0.0177	−0.0000	0.00002
27	7	0.0059	−0.0504	0.0059	−0.0503	−0.0000	0.00006
7	28	−0.0156	0.1328	−0.0156	0.1328	−0.0002	0.0017
8	29	−0.5455	4.6854	−0.5456	4.6854	−0.0756	0.6497
29	22	−0.3214	2.7405	−0.2769	2.3605	−0.0624	0.5322
14	17	0.0054	−0.0462	0.0046	−0.0396	−0.00002	0.0002
11	24	0.0085	−0.0678	0.0084	−0.0676	−0.00003	0.0003
11	4	0.0380	−0.2938	0.0378	−0.2926	−0.0002	0.0012
11	15	−0.0587	0.4984	−0.0584	0.4963	−0.0012	0.0105
15	25	−0.0579	0.5641	−0.0565	0.5498	−0.0009	0.0094
15	26	−0.0508	0.4346	−0.0496	0.4237	−0.0009	0.0080
26	5	3.0208	−2.3495	2.8890	−2.2470	−0.1318	0.1025
4	21	−0.0039	0.0327	−0.00396	0.0327	−0.0000	0.00002
2	21	−0.0039	0.0327	−0.0039	0.0327	−0.0000	0.00002
1	23	−0.1185	0.9261	−0.1184	0.9261	−0.0019	0.0149
29	14	−0.0094	0.0807	−0.0081	0.0695	−0.0000	0.0004
14	30	0.0137	−0.1170	0.0117	−0.1002	−0.0001	0.0009
10	12	0.0030	−0.0256	0.0029	−0.0254	−0.0000	0.0000
11	2	0.0261	−0.2173	0.0259	−0.2145	−0.0001	0.0009
15	3	1.7427	−1.356	1.6988	−1.3213	−0.0439	0.0341
8	18	−0.1486	0.6769	−0.1486	1.3112	−0.00890	0.0784
9	24	−0.0567	0.2838	−0.0566	0.2833	−0.00002	0.0001
19	17	−0.0408	0.2265	−0.0357	0.1432	−0.0006	0.0023
20	23	0.0158	−0.1236	0.0155	−0.1214	−0.0000	0.0003
27	26	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

tive power is not put in place there will be no much success. **Thus, it is advised that concentrating the compensation on the problem buses gives best result like buses** 14, 17, 18, 19, 22, 29, 30, at 65% recorded 0.976966, 1.009312, 0.954925. 1.029198, 0.834761, 0.93879 and 0.98823 and others as shown in Table 16.7. This reduces cost as well. Losses in the system were minimized as can be seen through Tables 16.8, 16.9, 16.10 and 16.11.

Table 16.9 Summary of total loss (uncompensated)

Line current (p.u.)		Line flows (p.u.)		Line losses (p.u.)	
Real	Imaginary	Real	Imaginary	Real	Imaginary
5.2816	3.4557	5.1362	3.2760b	−0.2427	0.6670

Table 16.10 Line current, line flows, and line losses (compensated)

B/sequence		Line current (p.u.)		Line flows (p.u.)		Line losses (p.u.)	
From	To	Real (I)	Imagi (I)	Real (P)	Imag (P)	Real (L)	Imag (L)
16	12	0.0859	−0.7304	0.0853	−0.7254	−0.0003	0.0028
12	1	0.0226	−0.1768	0.0225	−0.1762	0.0225	−0.1762
12	11	0.0047	−0.036	0.0047	−0.0359	0.000	1E−04
12	13	−0.0006	0.0048	−0.0006	0.0048	0.0000	0.0000
13	10	0.0224	−0.1908	0.0223	−0.1901	−0.0002	0.0013
10	11	−0.0056	0.0477	−0.0056	0.0479	0.0000	0.0002
10	27	−0.01	0.085	−0.010	0.0853	0.0000	0.0003
12	6	0.6698	−0.4912	0.6677	−0.4897	−0.0021	0.0015
27	8	0.0009	−0.0093	0.0009	−0.0093	0.0000	0.0000
27	7	0.0031	−0.0263	0.0031	−0.0263	0.0000	0.0000
7	28	0.0093	−0.079	0.0093	−0.079	−1E−04	0.0006
8	29	−0.1362	1.1695	−0.1362	1.1695	−0.0047	0.0405
29	22	−0.0522	0.4453	−0.0504	0.4299	−0.0016	0.014
14	17	0.0599	−0.5103	0.0588	−0.5013	−0.0025	0.0214
11	24	0.0004	−0.0031	0.0004	−0.0031	0.0000	0.0000
11	4	0.0022	−0.0167	0.0022	−0.0167	0.0000	0.0000
11	15	−0.0571	0.4845	−0.0571	0.4844	−0.0012	0.0099
15	25	−0.0252	0.2455	−0.0247	0.2405	−0.0002	0.0018
15	26	−0.0539	0.4605	−0.0527	0.451	−0.0011	0.009
26	5	2.785	−2.1661	2.673	−2.079	-0.112	0.0871
4	21	0.0037	−0.0305	0.0037	−0.0305	0.0000	0.0000
2	21	0.0037	−0.0305	0.0037	−0.0305	0.0000	0.0000
1	23	−0.2254	1.7619	−0.2254	1.7619	−0.0069	0.0543
29	14	0.0326	−0.2792	0.0315	−0.2695	−0.0006	0.0047
14	30	0.0198	−0.1691	0.0195	−0.1661	−0.0002	0.0020
10	12	−0.0183	0.1552	−0.0183	0.1557	−1E−04	0.0010
11	2	0.0015	−0.0123	0.0015	−0.0123	0.0000	0.0000
15	3	1.431	−1.113	1.4014	−1.09	−0.0296	0.023
8	18	−0.0827	0.7296	−0.0827	0.7296	−0.0027	0.0243
9	24	−0.0117	0.0586	−0.0117	0.0586	0.0000	0.0000
19	17	−0.0623	0.2498	-0.0651	0.2612	−0.0013	0.0053
20	23	0.1632	−1.2761	0.1546	−1.2083	−0.0036	0.0285
27	26	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 16.11 Summary of total loss (compensated)

Line current (p.u.)		Line flows (p.u.)		Line losses (p.u.)	
Real	Imaginary	Real	Imaginary	Real	Imaginary
4.5805	−1.4488	4.4256	−1.2589	−0.1485	0.1574

References

1. U.C. Ogbuefi, B.O. Anyaka, M.J. Mbunwe, T.C. Madueme; Compensation effect on the interconnected Nigeria electric power grid, in *Proceedings of the World Congress on Engineering and Computer Science 2017, Vol I WCECS 2017*, San Francisco, USA, 25–27 October, 2017, pp. 249–255
2. F. Iiiceto, E. Cinieri, Comparative analysis of series and shunt compensation schemes for AC transmission systems. IEEE Trans. **PAS-96**, 1819–1821 (1991)
3. NIPP In-House Grid Studies 330 kV and 132 kV Transmission Line Data (2006)
4. P. Shankar, Kundur, *Power System Stability and Control II*, 2nd edn. (McGraw-Hill Books, New York, 1994), pp. 581–595
5. J. Arrilaga, C.P. Arnoid, *Computer Analysis of Power System* (Wiley, UK, 1994), pp. 135–145
6. J.B. Gupta, A Course in Electrical Power (Sanjeev Kumai Kataria, Sarak Delhi, 2007–2008)
7. J.D. Glover, M.S. Sarma, *Power System Analysis and Design*, 3rd edn. (Wadsworth Group Books Cole, a division of Thomson learning Inc, 2002)
8. U.C. Ogbuefi, Power flow analysis of Nigerian power system with compensation on some buses. Ph.D Thesis, Department of Electrical Engineering, University of Nigeria, Nsukka, pp. 37–40, 75, 129, Aug 2013
9. P.O. Ewesor, Practical *Electrical Systems Installation Work & Practice*, 2nd edn. (Electrical Inspectorate Services Department, Federal Ministry of Power Abuja, Petvirgin Partners Publishing Co. Benin City, 2010), pp. 269–281
10. A. Husain, *Electric Power Systems*, 5th edn. (CBS Publisher & Distributors New Delhi Aligarh (India), 2011), pp. 323–352
11. B.R. Gupta, *Power System Analysis and Design* (S. Chand & Company Ltd, 2011), p. 473
12. S.S. Venkata, W.C. Guyker, W.K. Booth; Six-phase (mult-phase) power transmission systems, fault analysis. IEEE Trans. Power App Syst. **PAS-96**(3) (1977)
13. W.F. Tinney, C.E. Hart, Power flow solution by Netwons method. IEEE Trans. Power App. Syst. **PAS-86**, 1439–1458 (1967)
14. S.A. Marshall, *Introduction to Control Theory* (MacMillian Press, 1998)
15. R.S. Er. Dahiya, Sub-station engineering, design, concepts & computer applications (S.K. Kateria & Sons® Publishers of Engineering and Computer Books 4760–61/23, Ansari Road, Daryaganj, New Deili, 2010), p. 202
16. O.A. Komolafe, A. Momoh, Omoigui, Reliability investigation of the Nigerian electric power authority transmission network in a deregulated environment, in *Conference Record of the IEEE Industry Applications Conference*, vol. 2, 12–16 Oct 2003, pp 1328–1335
17. F. Dawalibi, W.G. Finney, Transmission line tower grounding performance in non-uniform soil. IEEE Trans. **PAS-99**(2), 471–989 (1980)
18. Nigerian National Daily, NEPA Blames Outages on Erosion, Others,—The Punch Newspaper, 23 July 1999
19. PHCN National Control Centre Oshogbo, Generation and transmission grid operations, Annual Technical Report for 2005 (PHCN publisher, 2006)

20. A. James Momoh, *Electrical Power Systems Application of Optimization* (Howard University, Washington D.C., Marcel Dekker Inc., 2001)
21. S. Onahaebi, S.T. Apeh, "Voltage instability in electrical network", a case study of the Nigeria 330 kV transmission network. Res. J. Appl. Sci. (RJAS) **2**(8), 855–874 (2007)
22. D.P. Kothari, J. Nagrath, *Modern Power System Analysis*, 3rd edn. (Tata McGraw-Hill Publishing Company Ltd., New Delhi, 2006)

Chapter 17
Application of Remote Telemetry for Improving Formula SAE Car Performance

Masoud Fathizadeh and Anan Ayyad

Abstract Formula Society of Automotive Engineers (FSAE) competition provides opportunity for students to enhance their engineering design and project management skills by applying learned classroom theories in a challenging competition. The engineering design goal for teams is to develop and construct a single-seat race car for the non-professional weekend autocross racer with the best overall package of design, safety, construction, performance and cost. One such improvement has been the implementation of a telemetry and data acquisition system. Data acquisition can be further used to improve the performance of the car on the fly. A telemetry and data acquisition system allows the collection and interpretation of data from sensors, actuators and other relevant critical sub-systems with the car. The collected information enables the team to diagnose and solve issues with the car as well as programming tools, simulation tools, and other procedures used to create a working telemetry and data acquisition system. The system is based on an Arduino Mega microcontroller and its shields, which gather and transmit data from multiple sensors that measure parameters such as suspension travel, throttle/brake position, steering angle, fuel pressure, lateral/longitude/vertical acceleration, engine coolant temperature and many more sensor-values from the Engine Control Module (ECU) via CAN-BUS. Though, many of these devices were not meant to work directly with one another, the use of communication protocols allowed the system to successfully relay data back to the pit via graphical display for assessment by the participating FSAE team. An overview of the construction of the Formula SAE car, telemetry and data acquisition system, equipment and utilized software will be given. The goal of this article is to explain the concepts and design of the system; components, sensors, actuators, relevant software, calculations, challenges, and the methodologies used to overcome the problems with the system.

M. Fathizadeh (✉) · A. Ayyad
College of Engineering Technology, Purdue University Northwest, Hammond, USA
e-mail: fathizad@pnw.edu

A. Ayyad
e-mail: aayyad@pnw.edu

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_17

Keywords Hardware design · Interfacing · Remote sensing · SAE · Race car Simulation · Telemetry

17.1 Introduction

The Formula SAE student organization, members are primarily mechanical engineering preparing for a national competition to demonstrate their knowledge and capability in design and improvement of a race car. The car is powered by a lightweight 610 cc motorcycle engine in accordance with Formula SAE Michigan regulations, and it is among the competition's 120 registered vehicles. But the competition is more than auto racing. Crafted by college students from across the globe, the competing vehicles will be tested for endurance, speed, handling and acceleration. Before any car's rubber hits the track, each team will compete in static categories based on design, cost and fuel efficiency. The goal of the team is to make significant improvements on its systems and designs. One such improvement has been the implementation of a telemetry and data acquisition system.

A telemetry and data acquisition system allows for the collection and interpretation of data from sensors on the car, which enables the team to not only diagnose and solve issues with the other systems of the car, but to fine-tune and optimize the geometry of the mechanical systems as well as making suggestions to the driver based on data. The team is required to measure various parameters including, but not limited to: suspension travel of all 4 wheels, throttle and brake position, steering angle, fuel pressure, lateral/longitude/vertical acceleration, engine/coolant temperature, and other relevant data from engine control module (ECU). To accomplish this task, various devices, computer programming tools and software were used. The telemetry system in the car is based largely on an Arduino Mega microcontroller and compatible shields. Some devices used in this project were not necessarily made to be used together, but through the use of communication protocols such as CAN-BUS and SPI, the system was able to produce a reliable flow of data. This data is then transmitted via wireless communications and displayed graphically using LabView. The end result was a reliable, affordable data acquisition system. The goal of this article is to explain the concepts, design, components, sensors, software, calculations, challenges and remedies to overcome the problems as well as safety needed to complete the project.

17.2 System Description and Layout

17.2.1 Controller

The Arduino Mega 2560, a microcontroller board based on ATmega2560, controls the collection of data for the telemetry system through measurements taken by the

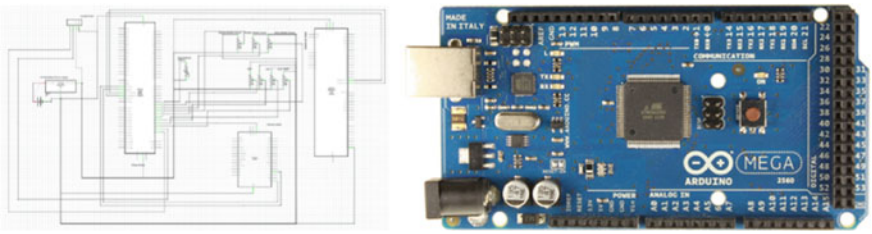


Fig. 17.1 Conroller layout and actual board

potentiometers and data logged and transmitted through the shields. The Arduino Mega 2560 is 101.52 mm long, 53.3 mm wide, and weighs in at 37 g. The microcontroller has 54 digital input/output pins (15 provide PWM output), 16 analog input pins, 4 UARTs (hardware serial ports), a 16 MHz crystal oscillator, a USB connection, a power jack, an ISCP header, and a reset button. It operates at 5VDC, and its recommended input voltage is 7–12 V, but it has a limit of 6–20 V. The DC current per input/output pin is 20 mA, while the DC current for the 3.3 V pin is 50 mA. It has a flash memory of 256 KB of which 8 KB is used by the bootloader. It has a static random access memory (SRAM) of 8 KB, while its read only memory (EEPROM) size is 4 KB. Its clock operates at a speed of 16 MHz [1]. Figure 17.1 shows the controller board layout as well as the actual board.

17.2.2 Buck Converter Voltage Regulator

A DROK 090029 regulated Buck Converter power supply, shown in Fig. 17.2, is used to provide power to the critical data collection devices. Steady and constant voltage is very important since variation of supply voltage can cause error in data collection and transmission. This power supply can have an input of 5–32 VDC and maintain a regulated output of 0–30 VDC. It is rated at 30 W with the maximum output current of 1.5 A. The output voltage preset resolution is 0.1 VDC. The voltage testing accuracy is $(\pm 1\%, +2)$ digits. The output is set to a constant 5 VDC which should not change as the system load changes drastically [2].

17.2.3 CAN-BUS

The CAN-BUS shield for Arduino allows for the logging of data from the ECU of the vehicle. The data can then be stored and displayed on a screen. The CAN-BUS shield is 101.6 mm long, 6.35 mm tall, and 101.6 mm wide as shown in Fig. 17.3. The shield features CAN v2.0B up to 1 Mb/s. It uses the Microchip MCP2515 CAN controller and the MCP 2551 transceiver. A standard 9-way sub-D, used with an

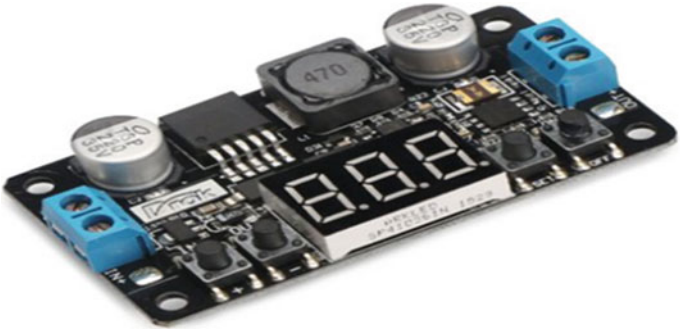
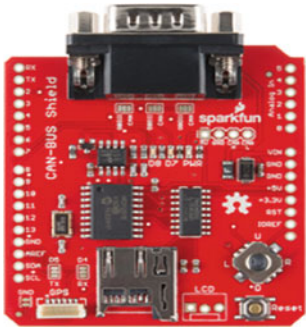


Fig. 17.2 DROK buck converter voltage regulator

Fig. 17.3 Sparkfun CAN-BUS



OBD-II cable, enables the CAN connection. The shield has a micro-SD card holder, a serial LCD connector, and a connector for an EM506 GPS module. It has a high speed SPI interface of 10 MHz on the shield, there is a reset button, joystick menu navigation control, and two LED indicators. Power can be supplied to Arduino by the sub-D via a resettable fuse and reverse polarity protection [3].

17.2.4 Triple-Axis Accelerometer Breakout

The triple-axis accelerometer as shown in Fig. 17.4 is used to determine acceleration forces in longitude, latitude, and vertical directions within the FSAE car. The accelerometer has a supply voltage of 1.95–3.6 V. Its interface voltage is 1.6–3.6 V. Its current consumption is 6–165 μA. The accelerometer possesses an I^2C digital output interface that operates to 2.25 MHz with a 4.7 kΩ pullup. It is manufactured by Adafruit Part identification number: MMA8451QT [4].

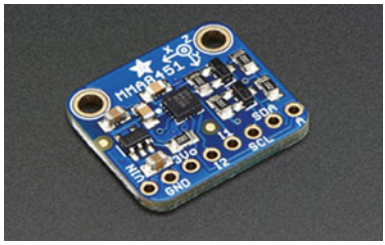


Fig. 17.4 Triple-axis accelerometer

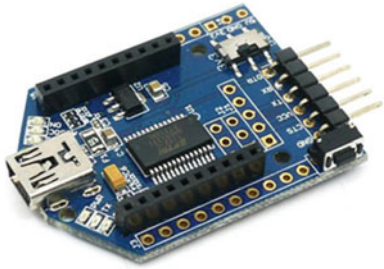


Fig. 17.5 UartSBee v4.0 shield

17.2.5 UartSBee v4.0 Shield

The UartSBee v4.0 connects the XBee-Pro 900 HP RF module, to the universal serial interface of the computer (USB). The UartSBee v4.0 shield (Fig. 17.5), where connected to the XBee-Pro 900 HP RF module (Fig. 17.6), is used to transmit and receive data collected by the Arduino Mega 2560. The data, which is stored on an SD card of the data logging shield, is also transmitted a computer running LabView software which simulates a virtual instrument panel for displaying types of data. The printed circuit board is 3.1 cm by 4.1 cm. Its microprocessor is FT232RL. The shield's interface is a mini-B USB and a 2.54 mm pitch pin header. The communication protocols for this shield are UART, eight Bit-bang inputs/outputs, and SPI. Its adapter socket is XBee compatible with a 2.0 mm pitch female pin header. The shield is FTDI compatible and has a USB 2.0 compatible Serial Interface. It has 3.3 and 5 V inputs/outputs and 3.3 and 5 V dual power outputs. It's typical and maximum input voltage is 5Vdc, and it has a current consumption of 500 mA. Its minimum output voltage is 3.3Vdc, while its maximum output voltage is 5Vdc it is manufacture red by Seeed Studio Part identification number: INT110B2P [5].

17.2.6 XBEE-Pro-900 Module

The XBee-Pro 900 HP RF module allows wireless connectivity to devices. The RF module receives and transmits data from the Arduino Mega 2560. The data is stored on the SD card of the data logging shield and simultaneously transmitted to a

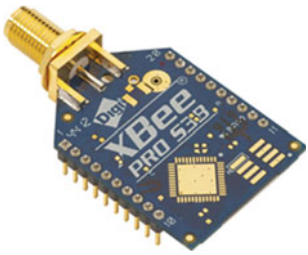


Fig. 17.6 XBEE Pro 900 HP

computer running the LabView software which simulates a virtual instrumentation panel for displaying relevant data. In order to set up the XBee-Pro 900 HP RF module, XCTU software is used. The specified antenna used for this project is RP-MSA. It has fifteen digital inputs/outputs, four 10-bit ADC inputs, and two PWM outputs. The RF module transmits data at a rates of 10 Kbps or 200 Kbps. At the lower the transmit speed, a longer the range and the slower the transmission and collection of data. The higher the transmit speed can be achieved. Also, with a shorter range a faster transmission and collection of data can be attained. The data rate operates between those two values for the purposes of the Formula SAE vehicle. The frequency range needed is 902–928 MHz. The supply voltage is 3.6Vdc. The transmit power is 250 mW. Its transmit current is 215 mA. Its reception current is 29 mA. Its outdoor/line-of-sight range is 10 kbps or up to 9 miles (15.5 km). This device is manufactured by Digi International Part identification number: 602-1299-ND [6].

17.2.7 Transmitting and Receiving Antennas

The transmitting antenna is mounted to the vehicle and handles the transmission of the custom bit string out of the initial processor (Arduino Mega). This 7-in. antenna is a straight whip type that requires panel mount. It has a frequency group of Ultra High Frequency (300 MHz–1 GHz) and a frequency of 900 MHz. The gain is 2.1 dBi.

The receiving antenna receives the custom bit string and is mounted to the trailer at the pit. This 6-in. antenna is a tilt whip type that requires a connector mount. It has a frequency group of Ultra High Frequency (300 MHz–1 GHz) and a frequency of 900 MHz. Its frequency range is 902–928 MHz. The gain is 2.15 dBi. These antennas are shown in Figs. 17.7 and 17.8 respectively [6].

Fig. 17.7 Transmitter antenna

Fig. 17.8 Receiver antenna

17.3 Sensors and Transducers

17.3.1 Linear and Rotary Potentiometer Sensors

The 75 mm linear travel potentiometer senses the linear displacement between two points when connected directly to the points of measure. These linear potentiometers can be used for measuring damper compression or extension and steering rotation, which is measured by rack displacement. Each device has its own specific software management. The linear potentiometers for the Formula car are utilized to measure damper compression. The linear potentiometer is housed in aluminum and has a spherical bearing of 5 mm. Its cable is a 450 mm Raychem 55A 24 AWG. Its environmental sealing is rated IP65, meaning the component is totally protected against dust ingress and limited ingress of low pressure water jets from any direction. The component best operates at a temperature range of −40 to 125 °C degrees. It has a mechanical life of greater than 25 million cycles and operates at a maximum speed of 10 m/s. Its maximum supply voltage is 40Vdc. The component's resolution is essentially infinite. The repeatability, or retest reliability, of the component is less than 0.01 mm and its independent linearity is less than ±0.5% [7].

The 10G rotary potentiometer is used for Formula SAE. The rotary potentiometer, once installed and calibrated, will measure angular displacement. For the purposes of the Formula SAE competition, the rotary potentiometer will measure and collect data on the angular displacement of the steering wheel. The rotary potentiometer has a nominal resistance of 10 kΩ, with a tolerance of ±5%. It has a high precision

Fig. 17.9 Linear potentiometer

Fig. 17.10 Rotary potentiometer-10G

of 0.030%. Mechanically, its displacement is 1080°/5 or 10 laps. The component operates at a temperature range of −55 to 125 °C. At 40 °C, the 10G potentiometer dissipates 2.4 W [7]. These potentiometers are shown in Figs. 17.9 and 17.10.

17.3.2 Fuel Transducer

A pressure transducer is used to measure fuel pressure that is given in psi. The part weighs 119.07 g with dimensions of 14.48 cm × 11.68 cm × 2.54 cm. The recommended supply voltage is 5 V, and its output voltage is approximately 0.5–4.5 V. The working pressure range is from 0 to 100 psi. Its performance data claims it was tested at pressures ranging from 0 to 435.11 psi. The sensor will burst at a pressure of 725.19 psi and above. Its overall accuracy is 1% full scale, and its long term stability is 0.2% full scale per year. The sensor operates at a temperature range of −40 to 125 °C [8].

17.3.3 Throttle Potentiometer

The throttle potentiometer will measure the displacement of the throttle. Its nominal resistance is 5 kΩ linear. The potentiometer has a tolerance of ±20% and a linearity of ±2%. It allows an electrical displacement of up to 106°. Mechanically, the displacement is up to 130°. The sensor fatigues after 10^6 complete cycles. Its cable length is 240 mm, as shown in Fig. 17.11 [7].

Fig. 17.11 Throttle potentiometer

17.4 Calibration

17.4.1 Fuel Pressure Transducer

The process of calibration for the fuel pressure sensor began by connecting a pressure sensor gauge, this gauge served as the regulator for the pressure of the air supply. The gauge is wired to an Arduino Uno, which allowed for the reading and display of analog input and pressure via the serial monitor of the Arduino code. To display the correct values for pressure, an equation relating analog input from the Arduino microcontroller to the pressure read on the pressure sensor had to be found and added to the Arduino code. The original equation for the pressure sensor readings, psiCalc, had to be adjusted to return proper values for the pressure reading.

Originally with this equation, all pressure readings read by the microcontroller code were incorrect. To solve this issue, a linear equation relating analog input and pressure was established. There were several steps involved in finding the proper equation. First, a print function was added to the Arduino code to display the analog input on the serial monitor. This function was used to determine what analog input is read when the pressure sensor reads a certain pressure. At zero psi, the analog input was 96. When the gauge read 60 psi, the serial monitor displayed an analog input of 591. Using a graphing calculator program Geogebra, the points (97, 0) and (591, 60) were plotted on an analog input versus pressure gauge reading graph. A best fit line was calculated for these two points using the features provided by the Geogebra program. This line was found to be $y = 0.12x - 11.64$. When applied to the Arduino code, the program displayed much more accurate pressure values on the serial monitor. For better accuracy, four points were taken at zero psi, 20 psi, 40 psi, and 60 psi. The analog input readings for these points were 96, 246, 420, and 591, respectively. Using the best fit line feature, the equation returned was $y = 0.12x - 10.73$. Though the equation derived using the regulator gauge was more accurate than the first two equations, another method of obtaining pressure readings was applied for more accuracy. Another gauge was attached to the first and left fully open. The leftmost gauge regulated the amount of air passing through the sensor, but the pressure reading was taken from the rightmost gauge. Using this constructed jig, nine points of data were taken and plotted in Geogebra. The gauge reading and analog input used for these points are shown in Table 17.1. After inputting this

Table 17.1 Pressure gauge calculated and actual reading with relevant % error

Gauge reading (psi)	Analog input	Serial monitor reading (psi)	Error %
0	97	0.00	0.00
10	182	9.89	1.11
20	265	19.97	0.15
30	348	29.93	0.23
40	429	39.29	1.80
50	515	49.73	0.54
60	598	59.81	0.31
70	683	69.53	0.67
80	765	79.25	0.94
90	844	89.93	0.07
Average % Error			0.65

equation in the Arduino code, the serial monitor reading in psi was compared to the actual reading on the pressure gauge in psi. To calculate the percent error of the serial monitor reading, the following equation was used:

$$error\,\% = \frac{(|Serial\ Monitoring\ Reading - Gauge\ Reading|)}{Gauge\ Reading} \times 100\% \qquad (17.1)$$

An average of all the percent errors for each point was calculated and is provided in Table 17.1. It was 0.6516%. Given that the average error for the readings were so low, it can be inferred that the program and formula found is accurate.

17.4.2 Suspension Displacement Potentiometer

The first step in the calibration process for the linear suspension potentiometer was converting its length to imperial units. The length of 75 mm corresponds to 2.95 in. Next, the length of displacement at full extension and full compression of the potentiometer had to be measured with a set of calipers. Fully extended, the length of displacement of the potentiometer was 4.1185 in. Fully compressed, the length of extension of the potentiometer was 1.1315 in. From these two values, a midpoint was found by subtracting 1.1315 in. from 4.1185 in. to find 2.9870 in., the total length between those two measurements. That length of 2.9870 in. was divided by 2 to find the midpoint of that length, which was 1.4935 in. This value was then added to 1.1315 in. to find the midpoint of the entire length of the potentiometer once fully extended. The point halfway between the midpoint and starting length of the potentiometer was found to be 1.8783 in. after subtracting 1.1315 in. from 2.6250 in., dividing the resulting value by two, and adding 1.1315 in. to the value after

the division by two. The point halfway between the midpoint of the potentiometer and the point of the potentiometer when fully extended was found to be 3.3718 in. after subtracting 2.6250 in. from 4.1185 in., dividing the resulting length by two, and adding the resulting length to 2.6250 in. The points were corrected by subtracting the starting position, 1.1315 in., from all points. Thus, the true midpoint of the suspension potentiometer was found to be 1.4635 in. The midpoint was set as the zero point for the purposes of measuring damper compression and extension. When the potentiometer extends outwards from the midpoint, or zero point, this is determined to be negative displacement. When the potentiometer compresses inwards from the midpoint, it is positive displacement. For these parameters, maximum extension would be −1.4635 in., and maximum compression would be 1.4635 in. At each point of measurement, an analog input value was displayed on the serial monitor and recorded in each point.

By dividing the midpoint 1.4635 in by the analog value of 513 measured at that point, a multiplier of 0.0028528265 was calculated. This multiplier was applied to the function

$$\Delta X = \left[(\text{suspension A1x } 0.0028528265) - 1.4635) * 100)\right]$$

This equation enabled the proper reading of the displacement of the sensor in both positive and negative directions relative to the midpoint.

Once an equation was established and implemented into the code, the error of each point was considered. Five measurements in both the positive and negative directions were taken. The output value from the serial monitor was compared to the actual caliper reading. Percent error between the caliper reading and the displacement output from the sensor was estimated using the formula:

$$Percent\ Error = \left(\frac{|Serial\ Reading - Caliper\ Reading|}{Caliper\ Reading}\right) * 100\%. \qquad (17.2)$$

The average percent errors for the negative and positive directions on the sensor were 0.49% and 0.04%, respectively. These percent errors indicate that the calibration was highly accurate.

17.4.3 Steering Angle

Displacement Potentiometer

The calibration process for the rotary potentiometer began with the calculation of a multiplier and a function to use in the Arduino code during testing. The rotary potentiometer has the capability of 3600° rotation, meaning that the center position would be at 1800°. Since this potentiometer is linear and the maximum analog input from the Arduino is known to be 1023, the analog input at the maximum 3600° rotation would also be 1023. At the center position of 1800°, the analog input would

read 511.5. To find the ratio of degrees to bits, the center position of 1800° is divided by the analog input of 511.5 to return a ratio of 3.5° to one bit. However, the diameters of the gears are known to be different, so this ratio must be adjusted accordingly. The diameter of the first rotary gear (d_1), which would be located on the steering shaft, was measured with calipers and found to be 1.227 mm. The diameter for the second gear (d_2) was found to be 1.100 mm. Using these numbers, the gear ratio was calculated to be 1.115 mm. This gear ratio was multiplied by 3.5° to find the actual degree-to-bit ratio, which is 3.904° to one bit. So, at an angle of 90°, using the degree-to-bit ratio, we know that the analog input should read 23.053. However, the steering wheel will be limited to rotating only to the right (positive direction) 90° and to the left (negative direction) 90°, meaning that the actual analog reading for the sensor at the positive 90° displacement should be 511.5 + 23.053 or approximately 535. At zero degree displacement, our center position would have an analog input reading of 511.5, and at negative 90° displacement, the analog input would be 511.5 − 23.053 or approximately 488. From these values, we know that the formula should be $\Delta X = [(\text{SteeringAngleA3} * 3.5) - 1800]$ to return the proper displacement value. This function was used for the Arduino code and was tested. The percent error of these values was found to be 6.5%. Using the best fit line function, the following formula was calculated to be used in the Arduino code:

$$\Delta X = \left[(\text{Steering\_Angle} * 3.05) - 1564.37\right] \tag{17.3}$$

17.4.4 Brake Position Sensor

The brake position sensor was first mounted to the calibration. The analog reading was established on the serial monitor and values were read from 0 to 1023. The values of the analog reading were used to determine the physical extremes of the sensor. The positions of these extremes were marked on the calibration bench. The pedal maximum was found and the new extreme value relevant to the pedal position was noted. Within the Arduino code, the new minimum and maximum analog reading values were scaled from 0 to 100 to represent 0–100% deflection.

17.5 Signal Processing

LabView is an electrical simulation and data display software produced by National Instruments. The software utilizes graphical code to ease the process of programming its functions. The following LabView Front Panel is functioning as the graphical user interface for the system. It is displaying the gauges and readings of each data channel being obtained from the vehicle. LabView receives a custom string of bits over a serial COM port and imports the string into this program. There are two main

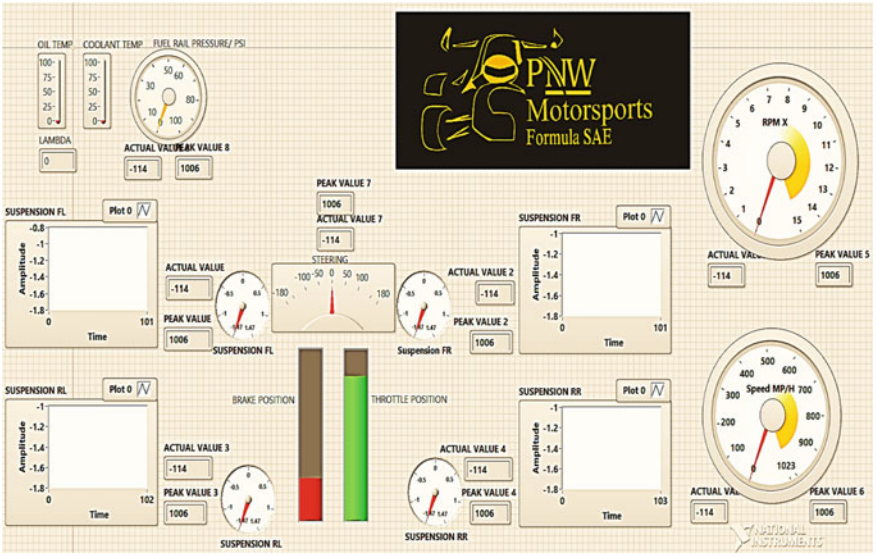


Fig. 17.12 LabView front panel

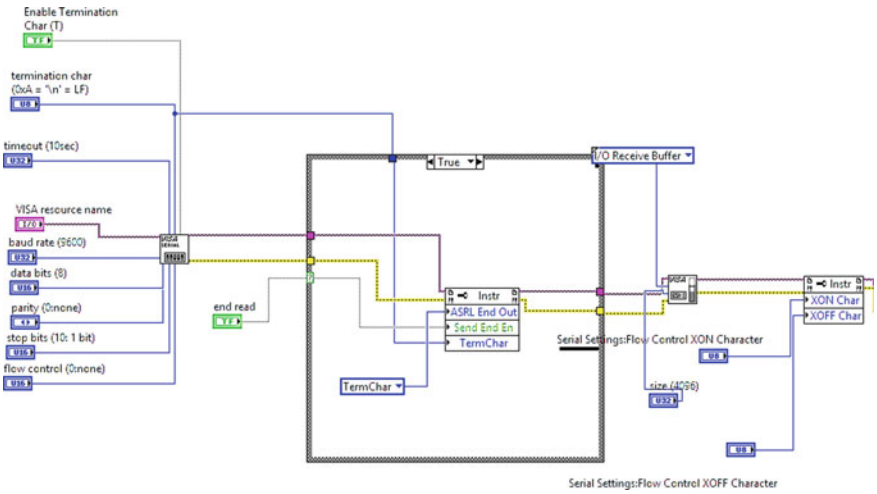


Fig. 17.13 The block diagram (VISA function)

panels for LabView: the Front Panel and the Block Diagram Panel which are shown in Figs. 17.12 and 17.13 respectively.

17.6 System Impact

17.6.1 Suspension

The design relies heavily on dynamics simulations performed on specific software. Data acquired from these simulations are then saved to be used as reference for future designs. However, with the introduction of a data acquisition system incorporated into the FSAE car, real time data can be analyzed and collected to improve the handling of the car. Sensors such as the shock LVDTs and accelerometer display what roll gradient the car experiences when driving around sharp corners and can be adjusted for quicker or slower damper response. Based on data collected, different suspension setups can be prepared ahead of time for each dynamic event at the FSAE events. The addition of a data acquisition system will aid in verifying computer simulation designs and provide data to improve immediate handling as well as future suspension designs.

17.6.2 Drive Train

The data acquisition system not only helps fine-tuning and adjusting the mechanical systems of the vehicle, it will also allow the FSAE team to diagnose any faults that may occur especially the ones that may lead to catastrophic failures. For instance, detecting abnormally high engine temperature readings or low oil pressure readings at the right time can help avert a major engine failure.

17.7 Conclusion

In summary, the time, labor, and funds that contributed to the construction of the telemetry and data acquisition system resulted in a functional system that returns reliable data. This system allowed for the monitoring of multiple parameters of the vehicle. From the data gathered, Purdue Northwest Motorsports could make necessary adjustments to the suspension and the drive train systems as well as detection and prediction of mechanical failures before they occur. The data acquisition provides sufficient information for future suggestions to the drivers to improve their performance. The overall cost of this system was very reasonable. This project showed that a functional telemetry system can be made in-house at an affordable price with outstanding functionality. A Brief description of this project is available at Ref. [8].

References

1. Arduino, *Arduino MEGA 2560 & Genuino MEGA 2560* (2016). From Arduino: https://www.arduino.cc/en/Main/ArduinoBoardMega2560. Accessed 20 Dec 2016
2. DROK, *DROK LM2596 Numerical Control Voltage Switching Regulator DC Buck Converter* (2013), From Amazon: https://www.amazon.com/DROK-Numerical-Switching-Regulator-Adjustable/dp/B00BYTEHQO
3. Sparkfun Electronics, *CAN-BUS Shield* (n.d.), From Sparkfun: https://www.sparkfun.com/products/13262. Accessed 20 Dec 2016
4. https://www.adafruit.com/?gclid=EAIaIQobChMIh_DktqrJ2QIVkrrACh0eJg88EAAYASAAEgKMAvD_BwE
5. Seeed Development Limited, *UartSBee V4* (2016), From SeeedWiki: https://seeeddoc.github.io/UartSBee_V4/. Accessed 20 Dec 2016
6. Digi International Inc., *XBee-Pro 900HP* (2015), From Digi: https://www.digi.com/pdf/ds_xbeepro900hp.pdf. Accessed 20 Dec 2016
7. AiM Infotech, *Car/Bike Linear Potentiometer Diameter 13* (n.d.), From AiM Sportline: http://www.aim-sportline.com/download/doc/eng/sensors/Dtasheet_CarBikeLinearPotentiometer13_100_eng.pdf. Accessed 20 Dec 2016
8. A. Ayyad, M. Fathizadeh, Telemetry and data collection to improve Formula SAE car, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October, 2017, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 317–321

Chapter 18
Reliable Energy Generation Using Hybrid System Technology to Improve Standard of Living in the Rural Area

Muncho Josephine Mbunwe and Uche Chinweoke Ogbuefi

Abstract The provision of reliable electricity supply in both developed and developing nation is one of the primary needs for socio economic development of that nation. This paper turn's to solve by way of introducing indigenous technology hybrid solar/Wind/Diesel Power system that will harness the natural renewable energies from the Sun and Wind to generate electricity. Renewable energy resources are a favourable alternative for rural energy supply. In order to handle the fluctuating nature, hybrid systems can be applied. The 2007–2016, 9 years solar radiation data, and the 2007–2016 data for wind speed for Potiskum in Yobe State of Nigeria, were collected. The two parameter Wielbull distribution was used to simulate power in watts per meter square densities for the 9 years period. Load estimates of a typical rural community and for rural ICT infrastructures were estimated. The step by step design of a 15 kW solar power supply system and a 10 kW wind power was done as a sample case. The results showed the average exploitable wind power density of 54.5 W/m^2 average mean speed of 8.04 m/s. It was found that with the exception of monsoon months, August and September, solar energy can be utilized throughout the year for North-Eastern region of Nigeria. Therefore, the development of hybrid wind-solar system for off grid communities will go a long way to improve socio economy lives of people in that community.

Keywords Density · Diesel · Hybrid power system · Photovoltaic Renewable energy · Wind

M. J. Mbunwe · U. C. Ogbuefi (✉)
Department of Electrical Engineering, University of Nigeria, Nsukka, Nsukka, Enugu State, Nigeria
e-mail: uche.ogbuefi@unn.edu.ng; ucchsamo@gmail.com

M. J. Mbunwe
e-mail: mamajoesix@gmail.com; muncho.mbunwe@unn.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_18

18.1 Introduction

Energy is essential to our society to ensure quality life and to underpin all other elements of our economy. It is one of the basic requirements of human society and vital for human life and technological advancement. The availability and consumption of energy is an index of prosperity in any nation. It is a known fact that the high rate of industrial growth of any country is a function of the amount of energy available in that country and the extent to which it is utilized [1]. Many societies across the world have developed a large interest for electrical energy [2]. This interest has been stimulated by the relative ease with which electricity can be generated, distributed, and utilized, and variety of its applications. It is arguable whether the consumption of electricity should be allowed to grow unchecked, but the fact is that there is an ever-increasing demand for this energy form. Clearly, if this demand is to be met, then the world's electricity generating capacity will have to increase [1, 2].

The objective of the energy system is to provide energy services. Energy services are the desired and useful products, processes or indeed services that results from the use of energy, such as for lighting, provision of air-conditioned indoor climate, refrigerated storage, transportation, appropriate temperatures for cooking, etc. The energy chain to deliver these services begins with the collection or extraction of primary energy, which is then converted into energy carriers suitable for various end-uses [1, 3]. The African continent is home to large untapped energy resources; however per capita energy use in Africa is less than a third of the global average, and an estimated 600 million people lack access to electricity [4]. The case of Nigeria is not so different from other African countries. Nigeria's energy is supplied from different hydro-power and thermal power stations. The country, located between longitude 8°E and latitude 10°N, and has two major seasons, wet and dry. The seasonality makes the extent of water availability at the different hydro power stations variable, leading to intermittent supply at times of low water levels during the dry season [1, 5].

Also, the thermal power stations have been bedeviled by lack of adequate supplies of natural gas from the various Niger Delta gas wells, thereby making continuous energy production from these installations difficult. This has left Nigerians at the mercy of private alternative power generation through the use of diesel and petrol generators. The emissions from these generating sets have also been subjects of critical global discussions because they release a lot of greenhouse gases to the atmosphere. As at 2007, about 25% of the 774 local government areas of Nigeria were still not connected to the national grid and today, more than 85% of these areas are still not connected [1, 6]. A national projection based on 13% Gross Domestic Product growth rate revealed that energy demand will increase from 5746 MW in 2005 to 297,900 MW in 2030 while supply should increase from 6440 MW to above 300,000 MW within the same period of years [7]. To accomplish this, requires an additional 11,686 MW every year to meet demand and costing for the period about $US484.62 billion. However, current energy production within the country is

less than 4000 MW due to fluctuations in the availability and poor maintenance of generating equipments. Thus, Nigeria still has a long way to go in achieving energy sufficiency [6].

18.2 Power Generation in Nigeria at a Glance

Under colonial rule, Nigeria started electricity generation and supply in 1896. In 1929, it set up the Nigeria Electricity Supply Company (NESCO) as an electricity utility company operating a hydroelectric power station near Jos, Plateau state. In 1951, the Electricity Corporation of Nigeria (ECN) was established. The first 132 kV line was built in 1962, to link Ijora power station in Lagos to Ibadan power station [5, 6]. Since then, there has been increase in electricity infrastructure and changes both in the nomenclature and operations of the regulating agencies. The Niger Dams Authority (NDA) was established in 1962 with a mandate to develop the hydro-power sub-sector. It was merged with the ECN in 1972. It was followed by the coming of the National Electric Power Authority (NEPA) then National Electricity Regulatory Commission (NERC) and Power Holding Company of Nigeria (PHCN) as the search for stable power supply in the country continues [1, 7].

The Federal Government in the year 2000 adopted an approach of restructuring the power sector and privatizing business units unbundled from NEPA to PHCN. This development NEPA has, since January 2004, been unbundled into seven generation companies, one transmission company and eleven distribution companies with a holding company that will unwind the unbundled companies. This arrangement is expected to encourage private sector investment particularly in generation and distribution [6]. Electricity production in Nigeria over the last 45 years has varied from gas-fired, oil fired, hydroelectric power stations to coal fired stations with hydroelectric power systems and gas fired systems taking precedence. Before now there are a total of 16 power plants in Nigeria 10 owned by Power Holding Company of Nigeria and another six plants belonging to independent power producers. This is expected to generate 6,426 MW of electricity in order to achieve the presidential target of 6,000 MW of electricity by December 2009. The Nigerian Power Generation sector can be detailed into the following subsectors [8] as shown in Tables 18.1, 18.2 and 18.3 respectively [1]:

(a) Existing Federal Government of Nigeria (FGN) Power Generation facilities.
(b) Independent Power Projects.
(c) National Integrated Power Projects (NIPP).

Table 18.1 Existing FGN hydro power station

S/N	Name of generation company	Year of construction	Location	Installed capacity (MW)	Available capacity (MW)
1	Kainji/Jebba Hydroelectric PLC-Kainji Power Station	1968	Kainji, Niger state	760	480
2	Kainji/Jebba Hydroelectric PLC Jebba Power Station	1985	Jebba, Niger state	540	450
3	Shiroro Hydroelectric PLC	1989	Shiroro, Niger state	600	450
Total				1,900	1,380

Table 18.2 Existing FGN thermal power station

S/N	Name of generation company	Year of construction	Location	Installed capacity (MW)	Available capacity (MW)
1	Egbin Power PLC	1986	Egbin, Lagos State	1,320	1,100
2	Geregu Power PLC	2007	Geregu, Kogi State	414	276
3	Olorunsogo Power PLC	2008	Olorunsogo, Ogun State	304	76
4	Delta Power PLC	1966	Ughelli, Delta State	900	300
5	Sapele Power PLC	1978	Sapele, Delta State	1,020	90
6	Afam Power PLC	1963	Afam, Rivers State	726	60
7	Calabar Thermal Power Station	1934	Calabar, Cross River State	6.6	NIL
8	Oji River Power Station	1956	Oji River, Achi, Enugu State	10	NIL
9	Omotosho Power PLC	2007	Omotosho, Ondo State	304	76
Total				5,004.6	1,978

Table 18.3 Independent power project

S/N	Name of power plant	Location	Installed capacity (MW)	Available capacity (MW)
1	AES Power Station	Egbim, Logos State	224	224
2	SHELL-Afam VI Power Station	Afam, Rivers State	650	650
3	AGIP-Okpai Power Station	Okpai, Delta State	480	480
4	ASG-Ibom Power Station	Akwa-Ibom State	155	76
5	RSG-TRANS AMADI Power Station	Port Harcourt, Rivers State	100	24
6	RSG-Omoku Power Station	Omoku, Rivers State	150	30
Total			1,759	1,484

18.2.1 Technologies for Rural Energy Supply

Generally, power supply in developing countries for rural areas takes place in three different ways:

1. Locally, by supplying single consumers and load groups;
2. Decentralize, by erecting or extending stand-alone regional mini-grids;
3. Centrally, by expansion of interconnected grids.

The approaches of local and decentralized electrification are obviously closely related and can be met by similar technologies [1, 9].

18.2.2 Decentralized Electrification

In highly fragmented areas or at certain distances from the grid, the decentralized approaches of regional mini-grid systems or local supply of single consumers can become competitive due to lower investments and maintenance costs compared to large scale electrification by expanding interconnected grids. Different technological options are in practice, most commonly diesel generating sets and renewable energies [9].

18.2.2.1 Diesel Generating Sets

In this cases security of supply is not of major importance, single diesel gensets are can be applied for electrification, accepting that no electricity can be supplied when the genset is out of commission, that is, due to repair or maintenance [10]. This problem can be met by using a group of diesel gensets, with the other gensets providing backup. The voltage of the generator is often adjusted to be higher than the required 220 V for the household because of high losses within the local distribution lines. Diesel gensets have problems with short durability, which is due to the fact that they work very inefficiently when running just at fractions of their rated capacity. Moreover, frequent start-up and shut-down procedures decrease their lifetime as well [11]. Diesel gensets are typically just operated for around 4 h in the evenings, and very often old motors from cars are used for the purpose of electrification [12]. Many, especially, rural areas are far away or isolated (that is islands) from higher developed regions so that the regular supply with diesel fuel becomes a logistical problem and an important financial burden even in countries, where fuel is heavily subsidized. Moreover, the transportation of diesel fuel can result in severe environmental damage [1, 9, 13].

18.2.2.2 Renewable Energy Technologies

The use of renewable energy technologies is a very promising approach towards meeting environmental, social as well as economic goals associated with small town electrification [6, 8, 14]. Renewable energy resources are of various forms with varying potentials, some of which are:

Biomass energy, solar energy, Wind energy, geothermal energy, Hydrogen and fuel cells, Ocean energy, and Hydro-power.

18.2.3 Prospect for Solar Energy in Nigeria

It is estimated that there is solar radiation of 3.5–7.0 kW h/m^2/Day across Nigeria daily, 3.5 kW h/m^2/day at coastal areas like Rivers and Bayelsa state and about 7.0 kW h/m^2/day at the extreme northern part of Nigeria. This amount to about, 5.53 kW h/m^2/Day, for the average global and solar radiations unutilized. This average amount represents a huge prospect for Solar energy generation if a total capacity can be developed for solar PV generation. The northern areas have an average daily sunrise time of 06:15 h and sunset time of 18:38 h [15]. In Nigeria less than 20% of 160 million populations are connected to the national electricity grid with 10% of urban population and 5% of rural population have access to the National Electricity Grid. With this, it is clear that solar energy installation and utilization in Nigeria is at very low basis. This implies that prospect for solar energy demand and utilization should be great in the near future if developed for domestic and industrial application

[16]. The mirage of problems encountered by Power Holding Company of Nigeria (PHCN) resulting in her inability to supply the required electricity to the Nigerian nation, due to lower generation capacity, out dated equipment, and other similar factors inclusive, makes the future of alternative (solar) energy very bright [1, 17].

18.2.4 Concept of Solar Energy

Solar energy is energy from the Sun. It is renewable, inexhaustible and environmental pollution free. Nigeria, like most other countries is blessed with large amount of sunshine with an average sun power of 490 W/m^2/day [1, 8]. Solar charged battery systems provide power supply for complete 24 h a day irrespective of bad weather. Moreover, power failures or power fluctuations due to service part of repair as the case may be, is non-existent.

There are two basic categories of technologies that convert sunlight into useful forms of energy. Firstly, solar photovoltaic (PV) modules convert sunlight directly into electricity. Secondly, solar thermal power systems use focused solar radiation to produce steam, which is then used to turn a turbine producing electricity [10, 11]. The advantage of using solar energy is that beyond initial installation and maintenance, solar energy is free solar energy does not require expensive and ongoing raw materials on like oil or coal, and requires significantly lower operational labour than conventional power production. Life expectancy ranges between manufacturers, but many panels produced today carry a 25–30 year warranty—with a life expectancy of up to 40 years [11].

18.2.4.1 Decentralization of Power

Solar energy offers decentralization in most (sunny) locations, meaning self-reliant societies [12].

18.2.4.2 Solar Avoidance of Politics and Price Volatility

Solar energy has the ability to avoid the politics and price volatility that is increasingly characterizing fossil fuel markets. The sun is an unlimited commodity that can be sourced from many locations, meaning solar is less vulnerable to the price manipulations and politics that have more than doubled the price of many fossil fuels in the past decade up till now. Solar energy production is set to become even cheaper in the near future as better technology and economies of scale take effect [11].

Table 18.4 Application of hybrid system technology

Energy services	Application	Benefits
Lighting	Household, public and community lighting	Improvement of health care
Communication	Internet/telephone TV broadcast/cinema	Facilitation of security/safety and emergency assistance
Refrigeration	Refrigeration of medicine Cold storage, ice making	Reduction of marginalization
Water supply	Drinking water, irrigation, purification	Increase of employment
Motive power	Drying and food preservation, textile and dyeing	Enhancement of social life
Health services	Sterilization of medical equipment, electric diagnosis	Generation of added value product and increase of productivity

18.2.5 Hybrid System Technology

Hybrid systems are another approach towards decentralized electrification, basically by combining the technologies presented above. They can be designed as stand-alone mini-grids or in smaller scale as household systems. Hybrid power system combines two or more modes of electricity generation together, usually, renewable energy technologies, such as Solar Photovoltaic (PV) and Wind turbines [1, 12–15]. Hybrid systems provide a high level of energy security through the mix of generation methods and often will incorporate a storage system (battery and fuel cell) or small fossil fueled generator to ensure maximum supply reliability and security. Hybrid systems allow for all productive uses of electricity (domestic, public and income-generating uses), permitting the scaling up of productive applications and rural development [11]. A case study is as shown in Table 18.4.

18.2.6 Solar Resource Assessment of the Site

The state lies mainly in the dry Savannah agro-ecological region and it is dry and hot for most part of the year, except in the southern part which has a milder climate. The state is predominantly rural with few towns like Nguru, Damaturu, Fika, and Potiskum. The land is generally a flat plain and nucleated settlement patterns have been developed mainly on stabilized sand dunes and around water points or where oases exist. Rural communities are well spaced out on the plain to allow each community enough land to support agriculture as the economic activity. Groundnut, beans, and cotton being the predominant crops cultivated. The energy needs by household in the state are mainly for cooking, where about 97.7% is gotten from fuel wood, for

Table 18.5 Solar irradiation level for potiskum in (kW h/M^2/Day) for 2007–2016

Month	Sunshine hours, S	Day length, S_{MAX}	% sunshine hours S/S_{MAX}	Hg (MJ/m^2/day)	H_0 (MJ/m^2/day)	KT
January	6.9	11.51	0.59	16.3	32.32	0.50
February	8.8	11.76	0.75	22.4	34.71	0.64
March	8.0	12.09	0.66	28.1	36.99	0.76
April	6.8	12.28	0.55	26.1	37.91	0.69
May	8.3	12.61	0.66	27.9	37.37	0.75
June	8.8	12.55	0.70	26.0	36.74	0.71
July	6.6	12.51	0.53	24.8	36.89	0.65
August	6.5	12.21	0.53	21.9	37.46	0.58
September	7.1	11.89	0.51	22.8	37.15	0.61
October	7.8	11.48	0.68	27.9	35.35	0.79
November	9.4	11.41	0.82	21.4	32.79	0.65
December	10.1	11.37	0.89	17.4	31.42	0.56

Hg the monthly average daily global solar radiation falling on a horizontal surface at a particular location
H_0 the monthly mean daily radiation on a horizontal surface in the absence of atmosphere
S the monthly mean daily number of observed sunshine hours
S_{max} the monthly mean value of day length and
KT the clearness index at a particular location

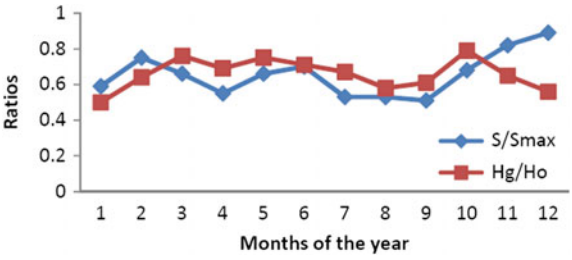


Fig. 18.1 A graph showing the clearness index for the different months of the year and the percentage sunshine hours

lighting, where kerosene lamp is the major source; and agricultural activities which is usually done manually. Machines and equipments used in micro-scale enterprises like tailoring, weaving etc. are powered manually [1, 12]. The average monthly solar irradiation level for the site, Potiskum, is as shown in Table 18.5 and Fig. 18.1.

From this it can be deduce that the average sunshine hours is approximately 8 h for an average day length of 12 h. From the observation of clearness index, the presence of clouds is very rare even in monsoon months. This is the favorable condition for solar energy utilization. Atmospheric clearness is indicated by fraction of extraterrestrial radiation that reaches the earth surface as global solar radiation, KT. The clearness index is the measures of the degree of clearness of the sky. From the estimated value of the global radiation, Hg, the clearness index, KT, is calculated. With the above

Table 18.6 Wind data for Potiskum for 16 years (1997–2012)

Month	V (m/s)	Σ (m/s)	C (m/s)	K	Power density (W/m^2)	Energy density ($kW\ h/m^2$)
January	8.57	2.28	9.57	4.23	55.52	43.11
February	9.07	2.46	10.14	4.20	60.01	42.58
March	9.07	2.25	10.09	4.57	73.34	72.59
April	8.94	1.74	9.81	5.95	69.73	68.85
May	8.32	1.57	9.11	6.16	68.35	55.39
June	8.30	2.16	9.26	4.34	65.27	50.89
July	7.71	2.02	8.60	4.31	63.59	48.72
August	6.58	2.19	7.43	3.32	49.85	37.92
September	6.29	1.70	7.03	4.16	25.09	15.87
October	6.41	1.54	7.12	4.73	39.92	30.25
November	8.54	2.25	9.53	4.28	43.89	28.79
December	8.69	2.52	9.76	3.80	39.32	23.32

V the average wind speed
Σ the standard deviation
C the average scale factor
K the shape factor

results and considerations, the maximum values of global solar radiation appear in March, April and May respectively during dry season while minimum values were observed in August and September, respectively during wet season.

18.2.7 Wind Resource Assessment

The wind speed data were obtained from Nigeria meteorological Agency Oshodi. The wind speed data obtained are for 20 years (1997–2016) at a height of 10 m from the ground surface. The data were averaged monthly wind speed over this period. The probability distribution functions of the wind speed and the duration function were calculated. Mean monthly values of wind speed were used for the calculation of Weibull distribution parameter C, D and K. Mean monthly values of wind speed were used to calculate wind energy and power densities [1, 13] as shown in Table 18.6 and Fig. 18.2.

From the Table 18.6, it is deduce that the average mean speed is 8.04 m/s with an average power density of 54.5 W/m. This site corresponds to the wind power class of 1 and therefore, this particular site is ideal for non grid connected applications.

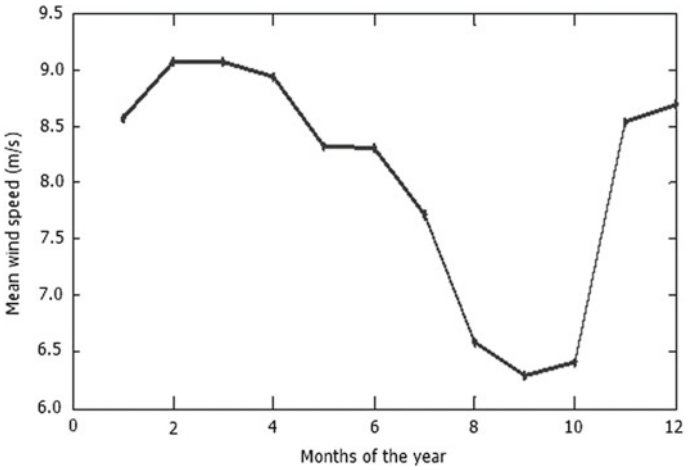


Fig. 18.2 A graph of wind speed for the months of the year

18.3 Methodology of the Design Technology

With most forms of electricity production, the primary fuel is "dispatchables", which can be converted to electrical energy at a rate which is controlled by the operator. This allows the electric utilities industry to adjust power output to meet demand as it fluctuates throughout the day. The combination of solar and wind with diesel generator(s) and a bank of batteries, is included for backup purposes. Power conditioning units, such as converters, are also a part of the supply system. The solar/wind hybrid is reliable. In order to ensure the continuity of supply, maximize the lifetime of components by reducing the stress on the system (especially the battery), and since solar and wind are dependent on seasons, diesel/gasoline/LPG generators are commonly used as complementary sources/backup. Finally, an Energy Management System (EMS) combined with inverter and coordinating battery, generator, and load management is essential for the optimum operation of a mini grid power system, especially hybrid systems. Basic operation of the hybrid system components are as follows.

18.3.1 PV Module

To obtain high power, numerous cells are connected in series and parallel circuits on a panel (module), to generate the required current and voltage. PV Modules can be wired together to form a PV array. By wiring modules in series the available voltage is increased and by wiring in parallel, the available current is increased. Either way, the power produced is the same sun. It is important to note that losses of voltage

occur due to the temperature rise of the cells in the heat of the sun and also that a 12 V battery typically needs about 14 V in order to be charged. To determine the size of PV modules, the required energy consumption must be estimated. Therefore, the PV module size (Ms) in Watt Power is calculated as:

$$Ms = \frac{Ec}{Is \times \eta} \tag{18.1}$$

where,

Ms module size in watt power
Ec Daily energy consumption in watts or kilowatts
Is Insulation in kW h/m^2/day
η efficiency

The PV modules mounting can be a ground mount that works either on rooftops or the ground, or pole mount for getting them up in the air. Both are angle adjustable so that PV array will face the sun as near to perpendicular as possible. Many owners will adjust their mounting racks two to four times a year to get maximum exposure as the sun changes its angle during seasons. If the rooftop has a good angle to the sun, the modules could be mounted solidly to the roof without an adjustable rack. Trackers are another PV mounting option, which are pole mounts that automatically adjust themselves so that the PV faces the sun throughout the day.

18.3.2 Small Wind Turbine

Wind turbine works the opposite of a fan as it is used to create electricity. Most turbines have either two or three blades. These three-bladed wind turbines are operated "upwind", with the blades facing the wind. The other common wind turbine type is the two bladed, downwind turbines. The wind turns the blades, which spin a shaft, which is connected to a generator and produces electricity. Utility scale turbines range in size from 50 to 750 kW. Single small turbines, below 50 kW, are used for homes, telecommunications dishes, or water pumping. The turbine should be subjected to non-turbulent wind and mounted higher than trees and other obstacles. A tall wind turbine tower is needed (9 m above anything within 120 m). The design principles of smaller wind turbines are somewhat different to the larger ones in that the small wind turbines produces power frequently over short periods, like for battery charging. It is important that small turbines generate in weak winds and respond quickly when harnessable winds occur. An important factor in how much power wind turbine will produce is the height of its hub.

Various mathematical models have been developed to assist in the predictions of the output power production of wind turbine generators (WTG). A statistical function known as Wiebull distribution function has been found to be more appropriate for this purpose [11]. The function is used to determine the wind distribution in the

selected site of the case study and the annual/monthly mean wind speed of the site. The Wiebull distribution function has been proposed as a more generally accepted model for this purpose. The two parameter Wiebull distribution function is expressed mathematically in Eq. (18.2) as:

$$F(v) = \frac{k}{c}\left[\frac{v}{c}\right]^{k-1} \exp\left[\left[\frac{v}{c}\right]k\right] \tag{18.2}$$

It has a cumulative distribution function as expressed in the Eq. (18.3):

$$M(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)\right] \tag{18.3}$$

where v is the wind speed, K (dimensionless) is the shape parameter and C (m/s), the scale parameter of the distribution.

To determine K and C, the approximations widely accepted are given in equations

$$k = \left[\frac{\sigma}{v'}\right]^{-1.09} \tag{18.4}$$

$$C = v' \times \frac{k^{2.664}}{0.184 + 0.186\,k^{2.735}} \tag{18.5}$$

where

C standard deviation of the wind speed for the site (m/s)
v′ mean speed (m/s)

18.3.2.1 Power Content of the Wind

The amount of power transferred to a wind turbine is directly proportional to the area swept out by the rotor, to the density of the air, and the cube of the wind speed.

The power, P, in the wind is given by:

$$P = 1/2Cp \cdot \rho \cdot A \cdot v'^{(3)} \tag{18.6}$$

where:

Cp turbine power coefficient.
A rotor swept area (m^2), proposed theoretical maximum value of 0.593 for Cp
ρ air density (kg/m^3)

$$A = \frac{\pi D^2}{4} \tag{18.7}$$

where

D rotor blade diameter (m) and
V′ mean wind speed (m/s)

18.3.2.2 Wind Speed Variation with Height

Wind speed near the ground changes with height. This requires an equation that predicts wind speed at a height in terms of the measured speed at another. The most common expression for the variation of wind speed with height is the power law expressed as follows:

$$\frac{v_2}{v_1} = \left[\frac{h_2}{h_1}\right]^{\alpha} \tag{18.8}$$

where v_1 and v_2 are the mean wind speed at height h_2 and h_1 respectively. The exponent factor depends on factors such as surface roughness and atmospheric stability. The power available in the wind is proportional to the cube of its speed. This means that if wind speed doubles, the power available to the wind turbine increases by a factor of eight, (2^3). Since wind speed increases with height, increase to the tower height mean enormous increase in the amount of electricity generated by a wind turbine. To make wind energy feasible in a given area, it requires minimum wind speeds of 3 m/s for small wind turbine and 6 m/s for large turbines.

18.3.3 Diesel Gensets

A diesel generator is simply a normal electric generator driven by a diesel engine (prime mover). An electrical generator is an electromechanical system that converts mechanical energy into electrical energy through the interaction of electromagnetic and electrostatic fields within the system. According to the reliability level desired in this paper, one can decide to install one or more generators in order to be able to provide full service, even during maintenance periods. Regarding the diesel generator, major maintenance operations should be considered with respect to the cost and unavailability of the equipment. Gensets in the range of 30–200 kVA would typically need a major maintenance operation after 15,000–25,000 running hours. Moreover, diesel generating sets are rather sensitive to climatic and geographic conditions. The decrease in efficiency is 1% for every 100 m above sea level, and another 1% for every 5.5 °C above a temperature of 20 °C.

18.3.4 Batteries of the Hybrid System

The case study uses lead acid batteries with tubular plates, either vented lead acid (VLA: flooded batteries with liquid electrolyte to be regularly refilled) or valve-regulated lead acid (VRLA: maintenance free batteries). The chosen battery should be specifically designed for solar applications. The battery cycling should be designed in order that the battery is able to store excess PV and wind power production and that it cycles within a discharge depth that allows enough cycles for the battery to last for at least 6 years (typically 2200 cycles at 50–60% depth of discharge, considering one cycle per day) and ideally for 8–10 years. Because of the strong impact of temperature on battery service (life a temperature increase of 5 °C decreases the service life by 15–20%), design of the battery room should ensure that batteries are kept at the lowest temperature possible. The battery to be used must:

(a) be able to withstand several charge and discharge cycle
(b) be at low self-discharge rate
(c) be able to operate with the specified limits.

The battery lifespan (measured in number of cycles) depends on the depth of discharge reached at every cycle: the deeper the battery is discharged at each cycle, the shorter its lifespan. The smaller the battery capacity, the cheaper the initial battery costs; however a smaller battery would be more deeply discharged and its lifetime reduced with its replacement cost increased. Because the battery recharge process ends with a very low power load (constant voltage, diminishing current) it is preferable to have this end-of-cycle charged by the genset rather than by the PV array, and at a time when the genset operates at a good load factor. This operating mode allows for maximizing the usage of solar energy.

A battery's total capacity (the total amount of current it can deliver multiplied by the duration of that supply) is reduced when discharged with a high current. This means that if the power load that the battery has to supply increases every year, after a few years the battery will not only be depleted more quickly and in addition, its total capacity gets reduced by higher discharge current. This has to be taken into account when considering locations with significant demand growth. The system designer should ensure that a safety margin is included in the sizing of the battery and that its operation mode optimizes its lifespan.

18.3.5 Inverter

A multifunctional inverter system which comprises of a device that controls the operating point of the PV array and optimizes its output will be introduced. It is a device that inverts DC current into AC and rectifies the AC current into DC to charge the battery. It also controls the charging of the battery to extend its lifespan. These functionalities can be split between several distinct units or combined in a central

piece of equipment. Seasonal variations of the load and its yearly growth should be taken into account when specifying its rated capacity. The multifunctional inverter devices are controlling the operations of the different energy sources of the system. A failure in one of its components would significantly hamper the functioning of the entire system. Improper settings for the various thresholds that control the shift between sources may affect the lifespan of the battery or the efficiency of solar energy use.

For PV Solar Systems, the inverters are incorporated with some inbuilt protective devices such as:

i. Overload protections
ii. Miniature Circuit Breaker Trip Indicator (MCB)
iii. Low-battery protection
iv. Constant and trickle charging system
v. Load status indicator.

The charging system (charge controller) should be able to manage the various charge steps, including regular equalization and float charge to maximize battery lifespan. It should have these features:

(a) Prevent feedback from the batteries to PV modules
(b) Have a connector for DC loads
(c) Have a work mode indicator.

For hybrid systems, there are two distinguish scenarios: firstly, if the hybrid systems relies on renewable energy technologies for power supply alone (like PV/Wind hybrid systems), the control of charge and discharge will work as it does in systems with just one renewable energy resource. The main objective of applying charge control is to maximize the battery's lifetime. Secondly, the situation is different for hybrid systems using diesel gensets as a backup. Since the genset is switched on the renewable energy resource cannot meet the demand, the objective of system control, in addition to maximize the battery's lifetime, is also to minimize costs for diesel fuel and maintenance.

18.3.6 Configuration for Hybrid Systems Technology

(a) Electricity generation coupled at DC bus line: All electricity generating components are connected to a DC bus line from which the battery is charged. AC generating components need a DC/AC inverter. The battery, controlled and protected from overcharge and discharge by a charge controller, then supplies power to the DC loads in response to the demand. AC loads can be optionally supplied by an inverter (Fig. 18.3).
(b) Electricity generation coupled at AC bus line: All electricity generating components are connected to an AC bus line. AC generating components may be directly connected to the AC bus line or may need an AC/AC converter to

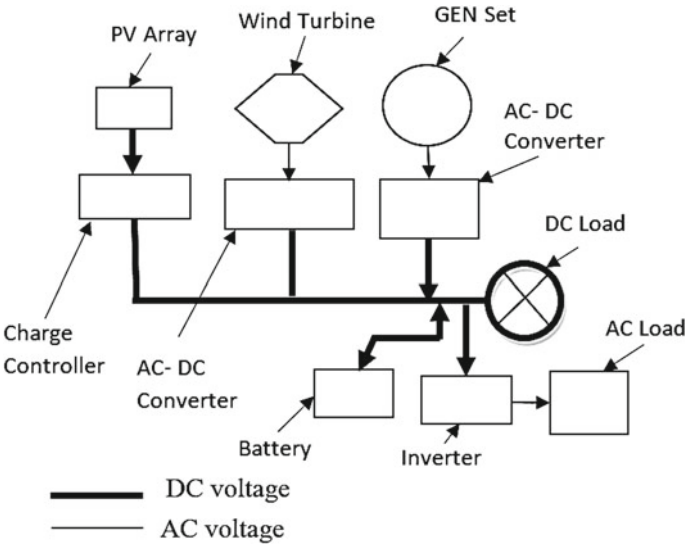


Fig. 18.3 Electricity generation coupled at DC bus line for charging the battery

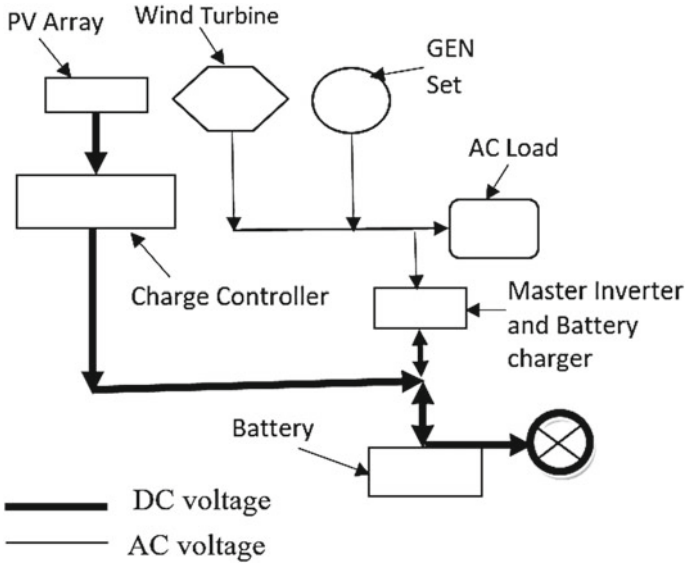


Fig. 18.4 Electricity generation coupled at AC bus line to enable a bidirectional master inverter control of energy supply to the AC loads and the battery charging

enable stable coupling of the components. In both options, a bidirectional master inverter controls the energy supply for the AC loads and the battery charging. DC loads can be optionally supplied by the battery (Fig. 18.4).

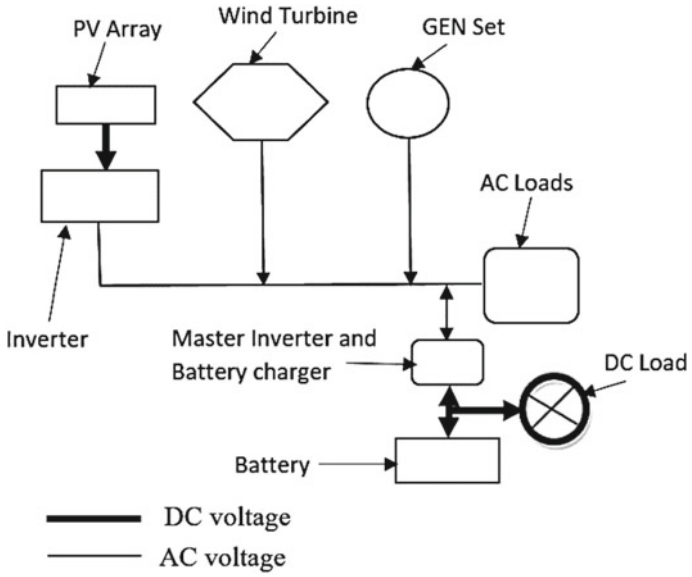


Fig. 18.5 Electricity generation coupled at the AC/DC bus line to enable component stability

(c) Electricity generation coupled at the AC/DC bus line.

DC and AC electricity generating components are connected at both sides of a master inverter, which controls the energy supply of the AC loads. DC loads can be optionally supplied by the battery. On the AC bus line, AC generating components may be directly connected to the AC bus line .or may need an AC/AC converter to enable stable coupling of the components (Fig. 18.5).

18.3.7 Load Estimate of a Typical Community Including Information Communication Technology (ICT) Services

The typical load curve for a rural village is generally composed of a prominent peak in the evening corresponding to lighting use, a morning/midday peak, and a base load. The base load is generally present in the morning, and in some cases extends to night hours. The energy demand in rural areas during night hours is quite limited (or non-existent in small villages) and hence the load level during the night is generally very low compared to the evening and morning peaks. Small hybrid systems are suitable for supplying the power needs of a small rural village where the energy consumption is quite limited; for instance a Potiskum village with none or very few, productive or commercial activities.

Assuming there are twenty families resident here, two health centers, two schools and an ICT centre. With the generation of electricity to this village, it is assumed that the 20 household's resident there will own electrical appliances. The load specification for the next 3 years is used in this analysis.

i. **Lighting Circuit Assessment**
Assume that each household will use 6, 40 W bulbs, therefore power demand for lighting is for 8 h in a day for a week, P_w, will be:

$$P_w = 6 \times 40 \times 20 \times 8 = 38.4 \text{ kW h} \tag{18.9}$$

ii. **Power Circuit Assessment**
Assume also that each household has 1 television set, Ts; 1 radio set, Rs; and 1 refrigerator, Ref, at power ratings of 120 W, 20 W and 250 W respectively. Therefore, Power for the Television sets:

$$P_{Ts} = 1 \times 120 \times 20 \times 8 = 19.2 \text{ kW h} \tag{18.10}$$

Power for radio set:

$$P_{Rs} = 1 \times 20 \times 20 \times 8 = 3.2 \text{ kW h} \tag{18.11}$$

Power for refrigerator:

$$P_{Ref} = 1 \times 250 \times 20 \times 18 = 90 \text{ kW h} \tag{18.12}$$

iii. Power demand for water pumping
Assume that the entire village will use two 1.5 hp pumping machine, pm.
Power demand:

$$P_{pm} = 2 \times 1.5 \times 0.746 \times 2 = 4.476 \text{ kW h} \tag{18.13}$$

Total power demand, P_{Td}:

$$\begin{aligned} P_{Td} &= P_w + P_{Ts} + P_{Rs} + P_{pm} + P_{Ref} \\ P_{Td} &= 38.4 + 19.2 + 3.2 + 4.476 + 90 \\ P_{Td} &= 155.276 \text{ kW} \end{aligned} \tag{18.14}$$

iv. The energy needed for a typical ICT center, Hospital center and School in rural/remote environment are as shown in Tables 18.7, 18.8 and 18.9.

All items use AC current. The total daily electricity demand is 185.496 kW h; the annual demand is 67.487 MW h; the maximum load for the system is 12.5 kW; and the study assumes a peak of 15 kW.

Table 18.7 Energy needed for a typical ICT centre in rural environment

Qty	Description of item	Load (W per unit)	Total load (W)	Daily/Hour of actual utilization (h)	Weekly watts/hours (Wh)
1	Router	25	25	8	1,200
1	Port fast Switch	15	15	8	720
2	Wireless Access Point	12	24	8	1,152
1	Server (plus accessories)	150	150	8	7,200
1	RF (Radio Communication)	40	150	8	1,920
10	Laptops (with security cables)	40	400	8	18,200
2	VOIP Phones	20	40	6	640
2	HP desk jet 5943	44	88	8	704
1	Laser Printer	100	100	7	700
4	Lighting	15	60	8	2,880
4	Ceiling fans	60	240	8	11,520
Total				10,156	46,836

Table 18.8 Energy needed for a typical hospital service in rural environment

Description of item	Qty	Load (W per unit)	Total load (W)	Daily hour of actual utilization (h)	Weekly watts hours (Wh)
Cold chain storage (fridge)	1	45	45	8	1,800
Lighting for the operating theatre	3	7.5	22.5	8	900
Lighting for ward	10	3	30	8	1680
Premises lighting/street light	4	15	60	12	5040
Television color	1	100	100	6	4200
Fans	6	15	90	12	1920
Total				4340	12,660

Table 18.9 Energy needed for a typical school in rural environment

Description of item	Qty	Load (W per unit)	Total load (W)	Daily hour of actual utilization	Weekly hour of actual utilization (W)
Fans	8	60	480	8	19,200
Lightning for street	6	15	30	9	1890
Lightning for classroom	12	15	300	8	12,000
Laser printer	2	100	200	5	5000
HP desk jet 5943	2	44	88	4	704
Total				7862	
2 school				15,724	

18.3.8 Design Results

The capacity ratio of solar to wind power units is 60:40. The plant is required to produce 200 kW h/day, which means that solar power plant should generate up to 120 kW h/day and the wind plant should generate 80 kW h/day. Thus the annual generation should be up to 73 MW h. The choice of components for the hybrid plant is as follows:

a. The solar plant is expected to operate for 8 h,
b. The wind plant is expected to operate for 24 h
c. The battery system is used during night hours.

The choice of 15 kW is a sample case and this can be extended to any required capacity. To achieve a solar power capacity of 15 kW the capacities of Solar panel, Charging Controller, bank of battery and Inverter are determined. The values cannot be picked abstractly and hence, their ratings and specification have to be determined through calculations in other for the system to perform to required specifications. For this design 8 h was assumed for the duration of the operation and the calculations is done as indicated below:

(1) ***Solar Panel***

Total load = 15 kW.

Period of operation or duration = 8 h.

Then, Total Watt-Hour = $15 \times 8 = 120$ kW h.

The period of the solar panel exposed to the sun = 8 h (Averagely between 8 am and 5 pm).

Therefore solar panel wattage = 120 kW h/8 h = 15 kW.

Hence solar panel of 15 kW will be needed for this design.

Since the highest power produced is 242.11 kW h/m^2, so for 8 h of average sunshine, we have 40.34 kW/m^2. This is much for this design, considering a 15 kW system is needed.

For this design, the Sun module SW 275 mono, Maximum power P_{max} 275 Wp, Open circuit voltage V_{oc} 39.4 V, Maximum power point voltage V_{mpp} 31.0 V, Short circuit current I_{sc} 9.58 A, Maximum power point current I_{mpp} 8.94 A, Cells per module 60, Cell type Mono crystalline, Cell dimensions (65.94 cm × 39.41 cm), weight 46.7 lbs (21.2 kg) with an efficiency of 16.4%. The study assumes that a maximum point tracker control method will be used and uses the minimum number of PV modules that will deliver 100% electricity to the load.

Therefore, No of panel = 15000 W/275 W = 56.

This shows 56 of 275 W solar panel will be required for this design.

(2) ***Charging Controllers***

For this design of 15 kW solar power supply,

$$P = IV \tag{18.15}$$

where:

I the expected charging current and
V the voltage of the battery (12 V)
P the power supply rating (15 kW)

Hence,

$$I = \frac{P}{V} = \frac{15{,}000}{12} = 1250 \text{ A} \tag{18.16}$$

A Charging controller of 625 A will be used for this study.

(3) ***Battery capacity***

The battery will supply 30% of the total load for 10 h during the night hours, that is, 3.75 kW.

Operational period = 10 h.

Watt-hour capacity = 37.5 kW/h.

To make the chosen battery to last long it is assumed that only a quarter (¼) of the battery capacity will be made used of, so that it will not be over discharged therefore the required battery capacity will be 37.5 × 4 = 150 kW/h.

Now the choice of battery hour depends on A-H rating of the storage battery. For example, for 1000 AH, 12 V battery the number of batteries that will be needed is 150,000/1000 = 150 batteries. Hence, for this design and to avoid too much weight and occupying unnecessary space, a 2490 AH 12 V, with size as 40 by 13.5 by 33 and weight of 1690 lb, a product of Solar-One technology battery is used. Therefore the total number of storage battery required for 150 kW solar power supply system = 60. The batteries are connected in parallel to increase the A-H rating, and the voltage is

increased using an inverter with a step up transformer to increase it to utility voltage of 240 V.

Since the total load is 12.5 kW it is advisable to size the required inverter to be 15 kW as designed for solar panel ratings. Hence 15 kW pure sign wave inverter is recommended with 90% efficiency, in other to prolong the lifespan of the inverter.

(4) ***Wind turbine details***

The wind turbine is estimated to produce 10 kW, in order to estimate the power extracted from the wind by the turbine, using the formula:

$$W = 1/2\rho A V_1^3 \times 0.59 \tag{18.17}$$

where,

$$A = \frac{\pi D^2}{4} \tag{18.18}$$

Using Eqs. (18.17) and (18.18), an average power output of 10 kW at a blade diameter of 8.2 m can be obtained from the study site, where mean wind speed, V = 8.04 m/s.

The cut in and cut out speed rating of a turbine also plays an important role in selecting a wind turbine, since the turbine can only produce its rated power at the rated speed.

For this study the 2 A air AWT wind turbine rated at 6 kW, with a cut in wind speed of 3 m/s and a rated wind speed of 11 m/s and a rotor diameter of 5.5 m is recommended as it can deliver optimal power generation in low and moderate wind speed. This turbine was chosen because the total power generated was approximately 10 kW.

(5) ***Diesel genset***

With the diesel generating set being used intermittently to supply the peak load for at least 10 h, the recommended rating of the gen set should be 15 kW.

(6) ***Control system for power management and Optimization of operational strategy***

Operating strategy for the hybrid PV/Wind/Diesel system with battery implies when the renewable sources produce less energy than what is demanded (that is, the wind speed, the solar radiation are low), the deficit power should be supplied by the battery bank. When the state of charge of the battery bank reaches its minimal level, (that is, 1/4 of the battery capacity), the diesel generator functions normal. A good operation of a hybrid system can be achieved only by a suitable control of the interaction in the operation of the different devices. An exhaustive knowledge of the management strategies to be chosen in the preliminary stage is therefore fundamental to optimize the use of the renewable sources, minimize the wear of batteries, and consume the smaller possible quantity of fossil fuel. Hybrid controller switches the

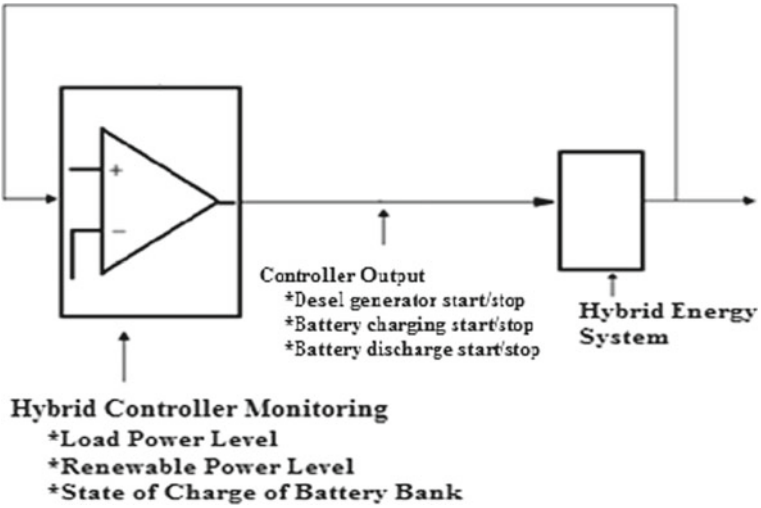


Fig. 18.6 Hybrid system controller block diagram for switching the batteries into charging mode or discharging mode

batteries into charging mode whenever excess power is available from the renewable sources, and switch to discharging mode whenever there was a short over from sources. The block diagram is as shown in Fig. 18.6.

That is the battery power indicates operating strategy for the hybrid system will show charging (power positive) or discharging (power negative), indicating that the hybrid controller utilizes the battery bank effectively.

18.4 Data Analysis on Study Location

The proposed site for the project is Potiskum, Yobe State, located in the North-East geopolitical zone of Nigeria. The area is located at 11.42°N, 11.02°E with an altitude of 414.8 m, its hottest month are March, April and May with temperature ranging from 30° to 42° and has an average annual solar irradiation of 6.12 kW h/m^2/day as shown in Table 18.1. The state lies mainly in the dry Savannah agro-ecological region and it is dry and hot for most part of the year, except in the southern part with milder climate. The state is predominantly rural with few towns like Nguru, Damaturu, Fika, and Potiskum. The energy needs by household in the state are mainly for cooking, where about 97.7% is gotten from fuel wood, for lighting, where kerosene lamp is the major source; and agricultural activities which is usually done manually. Machines and equipments used in micro-scale enterprises like tailoring, weaving etc. are powered manually.

18.5 Conclusion and Recommendation

With decreasing PV prices, PV/wind/diesel hybrid mini grids attract significant attention from institutions in charge of rural electrification and donor agencies to mitigate fuel price increases, deliver operating cost reductions, and offer higher service quality than traditional single source generation systems. The combining of technologies provides interesting opportunities to overcome certain technical limitations. The future deployment of hybrid technology in developing countries will be driven by different factors according to the type of application addressed. The micro hybrid system range for use as a reliable and cost effective power source for telecom base stations continues to develop and expand. Capacity building and access to concessional financing will be the key enablers for the development of this segment. Medium size distributed hybrid systems need political momentum to foster the involvement of the private sector. Larger isolated mini grids require substantial investments and then appropriate profitability. The result obtained from the analysis indicates that the solar energy utilization has bright prospects in Potiskum. The estimated values of global radiation can efficiently be used to compensate for energy deficits. From the studies it has been found that with the exception of monsoon months, August and September, solar energy can be utilized throughout the year for North-Eastern region of Nigeria. The site corresponding to the wind power class of 1 is also ideal for non grid connected applications.

References

1. M.J. Mbunwe, U.C. Ogbuefi, C. Nwankwo, Solar hybrid for power generation in a rural area: its technology and application, in *Proceedings of the World Congress on Engineering and Computer Science 2017, WCECS 2017*, 25–27 October 2017, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 309–316
2. C.O. Nebo, *Revitalizing an Effective Power System for Sustainable Development in Nigeria*, Delivered at the Herbert Macauley Lecture, University of Nigeria, Nsukka, June 2014, pp. 2–4
3. S. Kumar, V.K. Grag, A hybrid model for solar-wind power generation system. Int. J. Adv. Res. Electr. Electron. Instrumen. Eng. 1–2 (2013)
4. U.S. Akpan, Electricity Access in Nigeria: is off grid electrification using solar photovoltaic panels economically viable, in *A Sustainability, Policy and Innovative Development Research Solutions, Nigeria Project*, March 2013, pp. 4–10
5. Alliance for Rural Electrification, *Hybrid Power System Based on Renewable Energy: A Suitable and Cost Competitive Solution for Rural Electrification* (2013), pp. 8–9
6. Presidential Task Force on Power, Power Generation (Status and Outlook), in *Electric Power Investors Forum by Bureau of Public Enterprise (BPE)* (2012), pp. 4–8
7. A.B.J. Haney, *PV System Operations and Maintenance Fundamentals* (Next Phase Solar Inc., Lagos, 2013), pp. 12–17
8. M. Abdullah, S.K. Singh, Global Solar radiation evaluation for some selected stations in North Eastern Nigeria **2**(3), 1–8 (2014)
9. G.A. Duvuna, S.A Akanji, B.T. Abur, *Statistical Analysis of Wind Energy Potential Based on Weibull and Rayleigh Model in North Eastern*
10. Bureau of Public Enterprise (BPE), *Presidential Task Force on Power, Power Generation (Status and Outlook)* (2012), pp. 4–8

11. Alliance for Rural Electrification, *Hybrid Power System Based on Renewable Energies System* (2013)
12. S.L. Trama Tecno Ambienta, A brochure on hybrid power system based on renewable energies, in *Alliance for Rural Electrification* (2012), pp. 5–9
13. S.G. Oyagbirin, F.G. Akinboro, I.A. Adejumobi, *Hybrid Solar and Wind Power: An Essential for Information and Communication Technology, Infrastructure and People in Rural Communities*, vol. 9 (2011), pp. 12–17
14. G.San Miguel, D. Lopez, J. Servert, Hybrid solar-biomass plant for power generation: technical and economic assessment. Glob. NEST J. **13**(2), 266–276 (2011)
15. Oluseyi O. Ajayi, The potential of wind energy in Nigeria. Wind Eng. **34**(2), 304–305 (2010)
16. F.I. Ibitoye, A. Adenikinju, Future Demand for electricity in Nigeria. Appl. Energy **84**, 492–504 (2007)
17. T. Guul, *Integrated Analysis of Hybrid for Rural Electrification in Developing Countries*, TRITA-LWR Master Thesis 1651-064x, Royal Institute of Technology, Stockholm, 2004

Chapter 19
Android Mobile Malware Classification Using a Tokenization Approach

Intan Nurfarahin Ahmad, Farida Ridzuan, Madihah Mohd Saudi, Sakinah Ali Pitchay, Nurlida Basir and N. F. Nabila

Abstract Android is one of the most commonly targeted platforms in terms of mobile malware attacks on the part of many users worldwide. Different types of attacks and exploitations have been developed to masquerade as genuine mobile applications in order to obtain confidential information from the victim's smartphone. Therefore, to overcome these challenges, a new mobile malware classification based on system calls and permissions using a tokenization approach is developed in this paper. The experiment was conducted in a controlled lab environment by using static and dynamic analyses to extract permissions and system calls from call logs. A total of 5560 samples from Drebin were used as training dataset, and 500 samples from Google Playstore were used as testing dataset. The new classification involving the use of a tokenization approach produced a 99.86% accuracy rate and has outperformed existing methods. This new classification can be used as guidance, and reference for other researchers with the same interests. In the future it can be used as input for the formation of a mobile malware detection model.

Keywords Call logs · Hybrid analysis · Mobile malware classification Permissions · System call · Tokenization

19.1 Introduction

Nowadays smartphones are no longer limited to making calls and sending messages [1]. They also integrate multiple wireless networking technologies to support other functionalities and services such as social networking, web browsing, online banking

I. N. Ahmad · F. Ridzuan · M. M. Saudi (✉) · S. A. Pitchay · N. Basir · N. F. Nabila
Faculty of Science and Technology (FST),
Universiti Sains Islam Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Malaysia
e-mail: madihah@usim.edu.my

F. Ridzuan · M. M. Saudi · S. A. Pitchay · N. Basir
CyberSecurity and Systems Unit, Islamic Science Institute (ISI),
Universiti Sains Islam Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Malaysia

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_19

and Global Positioning Systems (GPS) [2]. A piece of research reported by an e-marketer shows that the number of mobile phone users continues to increase every year worldwide, and by 2019 more than half of the world's phone users will access the Internet regularly, and more than 4 billion people will be online in 2021 [3]. This will expose users to different kinds of cyber threat, especially in the form of mobile malware attacks. In this context, Android has become the most popular operating system for users and the one which is most commonly targeted by attackers [4]. Furthermore, Android users have the privilege of being able to conduct any customization and extensions in terms of downloaded applications, and the system allows users to install applications from various sources such as third-party markets, torrents or direct downloads [4, 5]. This privilege has created a loophole in terms of the user's security environment, and provides an opportunity for attackers to embed malware into an application by camouflaging it as a genuine application. The first Android malware in the form of code injection, entitled *Dvmap*, was found in April 2017. It has a rooting capability and has been downloaded from Google Playstore more than 50,000 times to date [6]. It is also capable of injecting malicious code into the system's runtime libraries [6]. In addition, *Skygofree* has surveillance and spy capabilities, can gain privileges, and steal call logs, GPS and text messages stored in the victim phone's memory. The application was spread using a fake website and has infected several Italian Android users [7]. These two are examples of malware that usually targets the Android platform. Most of the attacks are motivated by profit [2].

Many researchers have proposed methods for the classification and detection of Android mobile malware, such as the identification of expert features, system call behaviour, permission and API usage. Mobile malware classification and detection can be categorized into two techniques—static and dynamic-based analyses. Static-based analysis involves extracting information from the application's manifest file without executing it [8]. Meanwhile, dynamic-based analysis extracts the malware behaviour during its execution in an emulator environment, and allows researchers to obtain additional detailed information from the suspected applications [8]. Previous work which have used dynamic-based analysis, have managed to generate behaviour patterns from malicious applications [9–11]. However, there is still room for improvement in terms of accuracy and efficiency perspectives to optimize the performance.

In this paper, a hybrid approach which combines static-based (permissions features) and dynamic-based (system call sequence features) analysis is implemented. Hybrid-based analysis is used to extract permissions and system calls features related to Android call log exploitation. A tokenization approach is implemented in this research with the aim of producing a consistent pattern's string size. A combination of hybrid-based analysis and tokenization is proposed in this paper to increase the accuracy rate.

This paper is organized as follows. Section 19.2 presents the previous work related to Android mobile malware classification and detection techniques. Section 19.3 presents the research methodology and the proposed mechanism. Section 19.4 presents the overall experiment results, and Sect. 19.5 presents the summary and the potential for future research.

19.2 Related Work

Malware analysis is a very time-consuming activity, and one should have an in-depth knowledge in order to handle it. A good malware analysis approach is a great weapon to fight the dark side of the information society [12]. Malware analysis usually involves static and dynamic-based analyses or a combination of both entitled hybrid analysis [5].

19.2.1 Static, Dynamic and Hybrid-Based Approaches

Static-based analysis performs observations based on source codes or binaries without actually running the suspected program. The results usually show suspicious patterns and behaviours that exist in the program [8]. In the case of Android malware detection, the researcher usually extracts features such as Requested Permissions, API calls, Operation Code and system calls [13]. These features are used to detect a malicious payload and profile malware threats [13]. MAMA (Manifest Analysis for Malware Detection in Android) is designed to extract several features from the Android manifest file and has achieved a high true positive rate of 94.83% [14]. On the other hand, PUMA makes use of user permissions as the selected feature, and has managed to produce an 86% accuracy rate [15]. Meanwhile, Drebin uses several features such as hardware components, requested permissions, application components, filtered intents, API calls and network addresses, and produces a 94% of accuracy rate based on different malware families [16].

Static-based analysis, however, is not suitable for the detection of malware that employs cover-up techniques such as code polymorphism and obfuscation. It is slightly unstructured, and relies heavily on experience and personal skills [8]. Therefore, dynamic-based analysis or, as it is also known, behavioural-based analysis, is the alternative approach to counter the weaknesses of static-based analysis. Dynamic-based analysis observes suspicious behaviour in a running application. CopperDroid uses system call and binder information to detect bad behaviour on the part of suspected applications. It created four artificial malwares and obtained 100% classification accuracy based on its system call features application [10]. However, this research focuses only on author-created malware and the results show a high false positive for real world malware [9]. In the case of MALINE, it was used to detect malicious system call patterns on 4289 android samples, and produced a 93% accuracy rate. This has been evaluated using histograms and a Markov chain approach [17].

Nowadays, many researchers use hybrid-based analysis to enhance malware detection. Tools such as AASandbox and Droid-Sec are examples that integrate hybrid analysis to extract and analyse Android features such as permissions and Java code from the APK file [18, 19]. ProfileDroid has successfully discovered new unknown behaviours of mobile malware characteristics based on hybrid-based anal-

ysis [20]. Although these pieces of research have produced high positive rates of classification and detection, most of the behaviour-based patterns produced by these researchers are inconsistent in terms of their string size. In addition, they also suffer from a lack of ability when it comes to handling huge amounts of features collected from the various applications. These weaknesses can lead to low classification and detection accuracy rates. Therefore, a suitable approach is needed to increase the mobile malware classification and detection rate.

19.2.2 Tokenization Concept in Mobile Malware Classification

A major problem in terms of mobile malware classification and detection is the huge size of the dataset to be analysed. This places a huge burden on current malware detectors when it comes to obtaining an accurate result [21]. Thus, the indexing rule is implemented to increase the accuracy rate of malware classification and detection. A signature matching algorithm was proposed in order to give a more accurate result in the case of obfuscation programs [22]. Other researchers have applied Locality Sensitive Hashing to the behaviour profile, and successfully achieved efficient and scalable malware clustering [23]. Several researchers have also implemented the Function Call Graph approach in order to support efficient indexing techniques in malware analysis [24, 25]. Yet, these approaches are easily evaded by more advanced obfuscation techniques. Tokenization is a popular classification technique and one that is used widely in the information retrieval research area and lexical analysis [26]. Generally, tokenization is a process of classifying or breaking up a sequence of strings into pieces involving input characters, sub-characters, or subgroups [27]. Others have used a tokenization approach to minimize data length and complexity, thus reducing the cost of data handling [28].

In terms of information retrieval, tokenization is used to tokenize all words, numbers and characters into a subgroup from a sequence of words or texts. This will shorten the text and only the important words will remain, thus increasing the accuracy of the information retrieval process [26]. An N-gram algorithm has been implemented by several researchers in the mobile malware classification and detection field [8, 29]. They tokenized the set of malware's feature behaviour found in an application into n groups to produce a new classification and detection model which can produce more accurate results.

Therefore, this paper proposed an experiment that implements a tokenization approach that produces unique behaviour patterns based on permission and system call sequences with the combination of hybrid-based analysis with the aim of achieving a higher accuracy rate and better performance.

19.3 Research Methodology

The tokenization approach is used to produce a unique system call sequence pattern. This research implements a hybrid-based analysis approach where the permissions feature is extracted during static analysis and the system call feature is extracted during dynamic analysis. An initial experiment was conducted using 5560 Android malware samples from the Drebin dataset [16]. The experiment was conducted in a controlled laboratory environment as illustrated in Fig. 19.1. In this experiment, an emulator from Genymotion [30] was used with Android version 4.1.1 and API level 16. This emulator runs on Windows 8 with 8 GB of RAM. This experiment proposed an Android mobile malware classification based on system calls and permissions that are expected to exploit call logs. Four main phases were involved in this experiment. First, in Sect. 19.3.1 we discuss the permissions-based analysis phase. Then, in Sect. 19.3.2 we discuss the system call-based analysis phase where the process of extracting system calls from suspected samples were conducted. Meanwhile, in Sect. 19.3.3, the tokenization approach is used to build a new classification model. This produces unique behaviour patterns based on system call sequences. Finally, a new classification model of Android mobile malware based on system calls and permissions was produced, and its classification accuracy was tested using the WEKA tool and is explained in Sect. 19.3.4.

19.3.1 Permissions-Based Analysis Phase

We conducted static analysis to extract permissions from the mobile applications. In this phase, all permissions related to call log exploitation were extracted. Firstly,

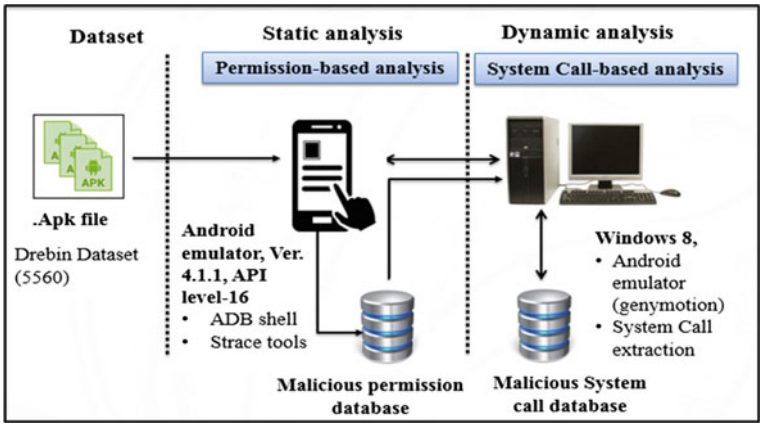


Fig. 19.1 Lab architecture

Table 19.1 Permissions that are expected to exploit call logs

Permissions	Function
CALL_PHONE	Malicious application may place unnecessary and illegal calls
CALL_PRIVILEGED	Malicious application may place unnecessary and illegal calls to emergency services
PROCESS_OUTGOING_CALLS	Malicious application can automatically make a phone call
READ_CALL_LOG	Malicious application is able to read, save and share call log data without user notice
READ_CONTACTS	Malicious application is able to read user's contacts
READ_PHONE_STATE	Malicious application is able to access features such as phone numbers and serial numbers of the phone

5560 malicious samples from Drebin [16] were extracted using static analysis. Then, applications with a minimum of one specific permission related to user call logs exploitation were further analysed. Table 19.1 shows the list of permissions that were expected to trigger user call logs exploitation. These permissions are chosen based on their function and their ability to perform suspicious activities in user call logs. The dynamic-based analysis is carried out once permissions-based analysis is complete.

19.3.2 System Call-Based Analysis Phase

The system call-based analysis phase consists of two main processes—system call recorder and system call analyzer. In the first process, the system call sequences of all suspected applications are extracted. This process is conducted in a Genymotion VM environment [30]. The system call can only be triggered through user interaction with the running application. Therefore, a monkey tool is used to generate pseudo-random gestures such as keystrokes, touches, and gestures on a device or on an emulator [17]. With this tool, the researcher was able to obtain consistent results as it allows the user to manipulate the command based on requirements. Next, each application went through 1000 random events or gestures generated at a time. The process of the system call recorder involved the following steps:

- Suspected application is installed in the emulator
- ADB shell monkey is used to trigger a system call event
- Strace tools is used to record the system call
- The extracted system call is stored as an Strace log for further processing

Finally, all recorded system calls for each suspected malicious application are stored as an Strace log file for further analysis. The system call analyzer consists of two main sub-processes. Firstly, the Strace log files are transferred to a system call sequenced patterns database. Each of the recorded system calls is noted as 1 to indicate the presence of the system call, while 0 indicates the absence of the system call. This step produces strings of binary numbers to represent the initial system call patterns for each application.

Next, each of the binary patterns is compared individually to avoid redundancy. Only unique patterns of system call sequences related to call logs exploitation are produced.

19.3.3 Tokenization of the System Call Sequence Phase

In this phase, the extracted malicious system call sequence is converted into new unique patterns using the tokenization approach. If we compare this with the existing work that used covering algorithms, it manages to produce 60 patterns from malicious system call sequences, but shows inconsistencies in terms of the pattern's string size [2]. The system call trigger was different for each application, hence various system call string lengths were produced. Therefore, the tokenization approach is implemented in this research with the aim of producing a consistent pattern string size with high data processing flexibility. Figure 19.2 shows the work flow of the system call sequence classification using the tokenization approach.

The tokenization approach consists of three main processes. First, the raw system call sequence extracted during the dynamic analysis is converted into binary values and a binary pattern. Next, the binary pattern is converted into hexadecimal values. This process reduces the string size and produces a consistent pattern string. Finally, the tokenization approach is implemented with regard to the new hex-value patterns, to classify or break them up into smaller pieces of input string [31]. Figure 19.3 shows the processes of the tokenization approach implemented in this experiment. The hex-value patterns are divided into a five-n different dataset, which includes

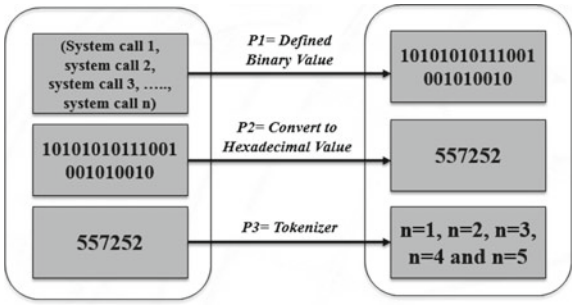


Fig. 19.2 System call sequence classification using the tokenization approach

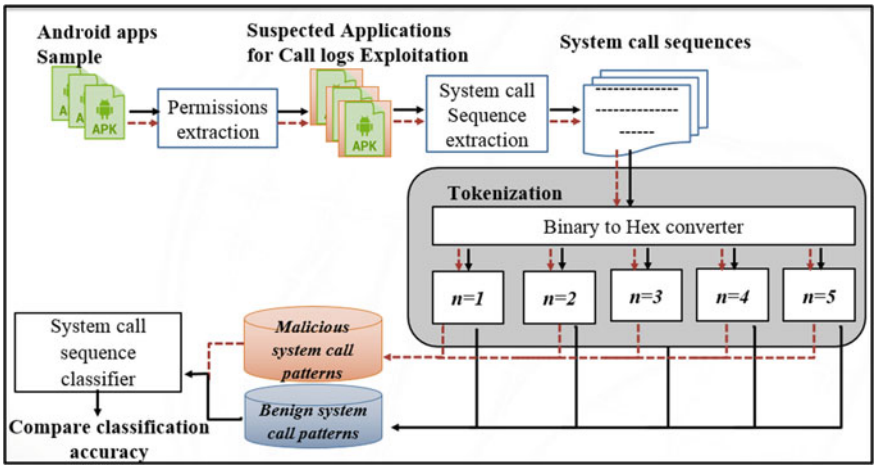


Fig. 19.3 Tokenization process for system call sequence classification

n=1, n= 2, n= 3, n=4 and n =5, where n indicates the number of tokenization groups. Later the output from this phase is used as the input for the next phase.

19.3.4 Behaviour-Based Classification the Evaluation Phase

Based on the tokenization approach, the output from the previous phase is used as the input for this phase. This phase evaluates the new malicious system call patterns that are expected to exploit call logs. The unique malicious system call patterns are further classified as being either malicious or benign, prior to finding the optimum n-value, classifier and accuracy of the classification performance. Four classifiers—Support Vector Machine (SVM), Random Forest, Naïve Bayes, and J48—were run using WEKA 3.8.10 [32]. These classifiers have been widely used by previous researchers [11, 16, 33] as they are able to deal with large instances, and features such as those found in text classification, pattern analysis and bioinformatics [29].

The new system call patterns are evaluated based on their classification accuracy. The classification accuracy is determined based on the number of patterns that are correctly classified as malicious patterns. The best n value and classifier are chosen based on the number of system call patterns that generate the highest classification accuracy.

19.4 Results and Discussion

In this section, the results related to sample extraction analysis and system call classification are presented and discussed.

19.4.1 New Unique Patterns

For this experiment, static and dynamic-based analyses are performed on 5560 samples of malicious application collected from Drebin [16]. Specific features were extracted in different phases. The first phase is the permissions-based analysis where any application with permission features that are related to call log exploitation as shown in Table 19.1, are categorized as malicious applications. Next, dynamic analysis is carried out on the suspected malicious applications. In this phase, the system calls from running applications were extracted. This phase identifies the behaviour of each application based on user interaction. Then, the extracted raw system calls are transferred to the system call sequence database. Next, each system call sequence log is compared to one another to eliminate redundancy, thus only unique system calls patterns remain. From 5560 malicious application samples used in this experiment, 464 patterns of malicious system call sequences were generated, related and suspected to involve call log exploitation.

A tokenization approach is implemented in this experiment to reduce the pattern's string size, and to produce a consistent system call string length for each pattern. Furthermore, this approach increases data flexibility by implementing a unique hex-value for each pattern, hence optimizing the performance of the mobile malware classification and detection approach. Figure 19.4 shows examples of system call patterns in terms of binary values which have been converted to hex-values.

19.4.2 Behaviour-Based Classification Accuracy

In this experiment, four popular classifiers in the form of SVM, Random Forest, Naïve Bayes, and J48 were used to classify $n = 5$ hex-values to different datasets produced during the tokenization phase, including the patterns in binary values. Each classifier is validated using cross-validation. The value is set as 10, where the cross validation is divided into 10 subsets and the holdout method is repeated 10 times, where 9 subsets are used for training and the last piece is used for testing. A total of 464 malicious patterns and the classification performance are shown in Fig. 19.5.

The classification accuracy rate A_i depends on the number of samples that are correctly classified as true positive and true negative over the sets of data. The evaluation is based on formula (19.1):

Samples	Binary-patterns	Hexadecimal-patterns (N=1)
1	1110111110011000000000100000000011111001100000100010100000000000	03BE600803E608A000
2	1010100001110000000100000000000011111010010000000000000011100001	02A1C04003E90000E1
3	1101101110011100011000010000000000000001100000000010100000000000	036E7184000600A000
4	1000011100001000000000000000000000000101100100100010100000000000	021C2000001648A000
5	1111111110011000000000100001000011111001100000100010100000000000	03FE600843E608A000
6	1110111110011000000000100000000000000000000000000000000000000000	03BE60080000000000
7	1111111111011101111000000000000000000010110010010001010000000000	03FF7780001648A000

Fig. 19.4 Example of 464 malicious system call patterns in binary, converted into hexadecimal

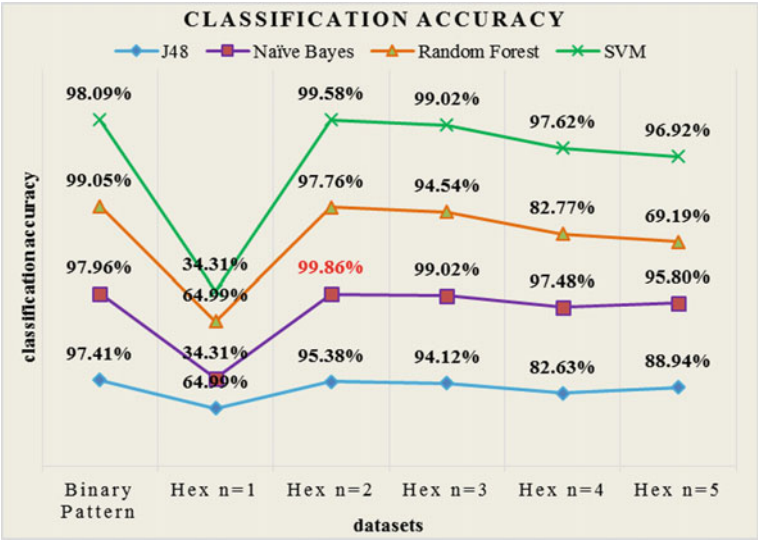


Fig. 19.5 Results in terms of classification accuracy

$$A_i = \frac{t}{n}(100) \tag{19.1}$$

where t is the number of sample cases correctly classified, and n is the total number of sample cases.

Therefore, in this research, the implementation of a suitable approach is meant to increase the number of samples that can be correctly classified by producing a consistent pattern string size. This consistent string size is defined as each pattern

having the same size and type of data. For example, in this research, the data are in the form of binary or hexadecimal values with array lengths of 66. This reduces the matching complexity between data.

This experiment aims to classify system call patterns into malicious or benign applications. Based on Fig. 19.5, the hex-value of the system call patterns using a Naïve Bayes classifier where $n=2$ resulted in the highest classification accuracy of 99.86%. Interestingly, when patterns are classified higher than $n=2$, the accuracy shown with regard to all classifiers is decreased. This is due to the fact that the increasing number of groups for each pattern produced in the tokenization phase generates a sparse vector element [29]. This element mostly holds zero values, resulting in a lower classification accuracy for each pattern. Other than that, the classification accuracy drops drastically once the patterns are converted from binary patterns to $n=1$ hex value patterns. The conversion of binary values to hexadecimal values has shortened the pattern's string length. Java programming was set up with a default one-dimensional array value based on the number of system call features extracted, to make the pattern size consistent for each application. The dataset $n=1$ hex value patterns was generated based on its default one-dimensional array value. As a result, a high spare vector element has been generated in the dataset, thus producing low classification accuracy for dataset $n=1$ hex value patterns [29].

In terms of the classifier used, Naïve Bayes successfully produced the highest classification accuracy compared to other classifiers. This is due to the nature of Naïve Bayes which is able to handle huge numbers of datasets, and is not sensitive to irrelevant features [34]. In this experiment, the binary patterns were able to produce a high classification accuracy, but by implementing the tokenization approach, the result was improved and increased up to a 99.86% accuracy rate, with $n=2$ hex values being chosen as the best dataset.

19.4.3 Comparison with Existing Work

In order to highlight the significance of this research result, a comparison is made with existing work. Table 19.2 shows how the results of this research measure against the best results in terms of classification accuracy rates found in the work of Lin et al. [8], Masud et al. [29], and Canfora et al. [35]. These comparative results show that this research approach has achieved some improvement in terms of classification accuracy rates. This is highlighted in the last column of Table 19.2.

In Table 19.2, the comparison is made based on the features used and the analysis approach of each work. Our work used hybrid-based analysis to extract the permission and system call features. Furthermore, the tokenization approach was implemented in this experiment to increase the mobile malware classification accuracy rate. The conversion of binary to hexadecimal values produced a consistent string size with the same data type. Moreover, the tokenization approach helps to increase classification performance by shortening the pattern string size to an n-group. Patterns with consistent string sizes and data types can reduce the matching complexity between data.

Table 19.2 Comparison with previous work

Author/work	Dataset	Features used	Analysis approach	Accuracy rate (%)
Masud et al. [29]	100 benign 102 malicious	System calls	Dynamic	97.06
Canfora et al. [35]	200 benign 200 malware	System calls, permissions	Hybrid	80
Lin et al. [8]	933 benign 265 malware	Permissions system call	Hybrid	98
Current work	5560 malware 500 benign	Permissions, system Calls	Hybrid	99.86

Less complexity increases the number of samples that can be correctly classified, thus increasing the classification accuracy rate. Therefore, with the implementation of the tokenization approach, this experiment successfully achieved a 99.86% accuracy rate. In conclusion, the current work outperformed existing work by Lin et al. [8], Masud et al. [29], and Canfora et al. [35].

19.4.4 Additional Discussion

During the system call-based analysis phase, the dynamic analysis was conducted in a controlled lab environment. To sustain the consistencies of the results from each suspected samples, the emulator is set up with 4.1.1 Android version and 16 API level. However, in terms of this experiment, not all samples can be executed. Some of them can be installed and run, while some of them can only be installed, but are unable to be executed (run in the background). On the other hand, some of them cannot be installed or executed in the emulator. The number of samples that have these conditions can be seen in Table 19.3. This condition might occurs due to the period of sample collection from Drebin which was from August 2010 to October 2012 [16]. During that period of time, most Android devices were using a lower than 4.1.1 Android version and API level 16. Therefore, some of the samples collected might only be compatible with the older and lower version of Android and API level, thus making it unsuitable for this experimental environment. Therefore, only compatible samples were used for dynamic analysis.

19.5 Conclusion and Future Work

This paper presents an Android mobile malware classification system based on permission and system call sequence patterns that are suspected of attempting to exploit user call logs using a tokenization approach. The utilization of such an approach

Table 19.3 List of samples' condition

Condition	No. of samples	Explanation
Installed and executed	5122	Compatible with proposed experiment environment
Installed and run in the background	357	Lower API level or Android Version
Unable to be installed	81	Lower API level or Android Version

increased the mobile malware classification and detection performance, and successfully produced a unique and consistent string length for each pattern. Thus, the size of patterns stored for data processing is also reduced, leading to a more efficient performance in terms of classification with a 99.86% accuracy rate. For future work, our work could be used as the input or basis for the creation of an Android malware detection model, and could provide guidance for other researchers in terms of implementing it with different Android application features.

Acknowledgements The authors would like to express their gratitude to Ministry of Higher Education (MOHE) Malaysia and Universiti Sains Islam Malaysia (USIM) for the support and facilities provided. The research paper is supported by MOHE: [FRGS/1/ICT04/USIM/02/1] and (PPP/USG-0116/FST/30/11616).

References

1. I.N. Ahmad, F. Ridzuan, M.M. Saudi, S.A. Pitchay, N. Basir, N.F. Nabila, Android mobile malware classification using tokenization approach based on system call sequence, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 85–90
2. M.M. Saudi, F. Ridzuan, N. Basir, N.F. Nabila, S.A. Pitchay, I.N. Ahmad, Android mobile malware surveillance exploitation via call logs: proof of concept, in *UKSIM '15 Proceedings of the 2015 17th UKSIM-AMSS International Conference on Modeling Simulation* (2015), pp. 176–181
3. e-Marketer Report, *Worldwide Internet and Mobile Users: eMarketer â€™ s Updated Estimates and Forecast for 2017–2021 Executive Summary* (2017)
4. M. Dimjasevic, S. Atzeni, I. Ugrina, Z. Rakamaric, Android malware detection based on system calls, in *Uucs* (2015)
5. X. Wang, Y. Yang, Y. Zeng, Accurate mobile malware detection and classification in the cloud. Springerplus **4**(1), 583 (2015)
6. R. Unuchek, Dvmap: the first Android malware with code injection, in *Securelist* (2017) [Online], https://securelist.com/dvmap-the-first-android-malware-with-code-injection/78648/
7. C. Smith, Android malware has 'never-before-seen' spying capabilities, in *BGR.com* (2018), pp. 1–5
8. Y.D. Lin, C.Y. Huang, Y.N. Chang, Y.C. Lai, Three-phase detection and classification for android malware based on common behaviors. J. Commun. Softw. Syst. **12**(3), 157–165 (2016)

9. I. Burguera, U. Zurutuza, S. Nadjm-Tehrani, Crowdroid: behavior-based malware detection system for android, in *Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices—SPSM '11* (2011), p. 15
10. A. Reina, A. Fattori, L. Cavallaro, A system call-centric analysis and stimulation technique to automatically reconstruct android malware behaviors, in *ACM European Workshop on Systems Security (EuroSec)* (2013), pp. 1–6
11. W. Yu, H. Zhang, L. Ge, R. Hardy, On behavior-based detection of malware on Android platform, in *GLOBECOM—IEEE Global Telecommunication Conference* (2013), pp. 814–819
12. H. Dornhackl, K. Kadletz, R. Luh, P. Tavolato, *Automatic Intelligent Analysis of Malware Behaviour*, vol. 8, no. 4 (2015), pp. 1225–1229
13. S.Y. Yerima, S. Sezer, G. McWilliams, I. Muttik, A new android malware detection approach using Bayesian classification, in *2013 IEEE 27th International Conference on Advanced Information Networking and Application* (2013), pp. 121–128
14. B. Sanz, I. Santos, X. Ugarte-pedrero, C. Laorden, J. Nieves, P. Garc, in *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, vol. 239 (2014), pp. 469–478
15. B. Sanz, I. Santos, C. Laorden, X. Ugarte-Pedrero, P.G. Bringas, G. Álvarez, PUMA: permission usage to detect malware in android, in *Advances in Intelligent Systems and Computing, AISC*, vol. 189 (2013), pp. 289–298
16. D. Arp et al., Drebin: effective and explainable detection of android malware in your pocket, in *Proceedings 2014 Network and Distributed System Security Symposium* (2014), pp. 23–26
17. M. Dimja, S. Atzeni, I. Ugrina, Z. Rakamari, M. Dimja, Android malware detection based on system calls android malware detection based on system calls, in *Iwspa@Codaspy* (2015), pp. 1–8
18. T. Bläsing, L. Batyuk, A.D. Schmidt, S.A. Camtepe, S. Albayrak, S. Bläsing, T. Batyuk, L. Schmidt, A.D. Camtepe, S.A. Albayrak, An android application sandbox system for suspicious software detection, in *2010 5th International Conference on Malicious unwanted Software (MALWARE)* (2010), pp. 55–62
19. Z. Yuan, Y. Lu, Z. Wang, Y. Xue, Droid-Sec, in *Proceedings of the 2014 ACM Conference on SIGCOMM—SIGCOMM '14* (2014), pp. 371–372
20. X. Wei, L. Gomez, I. Neamtiu, M. Faloutsos, ProfileDroid: multi-layer profiling of android applications, in *Proceedings of the 18th Annual International Conference on Mobile Computer Networks* (2012), pp. 137–148
21. W. Jin, *Practical, Large-Scale Detection of Obfuscated Malware Code via Flow Dependency Indexing* (2014)
22. A. Aiken, Y. Feng, S. Anand, I. Dillig, A. Aiken, Apposcopy: semantics-based detection of android malware through static analysis, in *FSE 2014* (2014), pp. 16–22
23. U. Bayer, P.M. Comparetti, C. Hlauschek, C. Kruegel, E. Kirda, Scalable, behavior-based malware clustering. Sophia **272**(3), 51–88 (2009)
24. J. Kinable, O. Kostakis, Malware classification based on call graph clustering. J. Comput. Virol. **7**(4), 233–245 (2011)
25. H. Gascon, F. Yamaguchi, K. Rieck, D. Arp, *Structural Detection of Android Malware using Embedded Call Graphs Categories and Subject Descriptors* (2013)
26. V. Singh, B. Saini, *An Effective Tokenization Algorithm for Information Retrieval Systems* (2014), pp. 109–119
27. Technopedia, Tokenization, in *Technopedia Technology Dictionary* (2017) [Online], https://www.techopedia.com/definition/13698/tokenization
28. M. Rouse, Tokenization definition, in *TechTarget* (2017) [Online], http://searchsecurity.techtarget.com/definition/tokenization
29. M.Z. Masud, S. Sahib, M.F. Abdollah, S.R. Selamat, R. Yusof, An evaluation of n-gram system call sequence in mobile malware detection. ARPN J. Eng. Appl. Sci. **11**(5), 3122–3126 (2016)
30. *Genymotion* (2014) [Online], https://www.genymotion.com/
31. C.D. Manning, P. Raghavan, H. Schütze, *An Introduction to Information Retrieval,* vol. 21 (2009)

32. Machine Learning Group at University of Waikato, *Weka 3: Data Mining Software in Java* (2014) [Online], http://www.cs.waikato.ac.nz/ml/weka/index.html
33. B. Wolfe, K.O. Elish, D. Yao, *Comprehensive Behavior Profiling for Proactive Android Malware Detection* (2014), pp. 328–344
34. E. Keogh, *Naïve Bayes Classifier* (2006)
35. G. Canfora, F. Mercaldo, C.A. Visaggio, A classifier of malicious android applications, in *Proceedings of the 2013 International Conference on Availability, Reliability and Security, ARES 2013* (2013), pp. 607–614

Chapter 20
Non-taxonomic Relation Extraction Using Probability Theory

N. F. Nabila, Nulida Basir and Mustafa Mat Deris

Abstract Ontology is a representation of knowledge with a pair of concepts and relationships within a particular domain. Most of extracting techniques for non-taxonomic relation only identifies concepts and relations in a complete sentence. However, this does not represent the domain completely since there are some sentences in a domain text that have a missing or an unsure term of concepts. To overcome this issue, we propose a new algorithm based on probability theory for ontology extraction. The probability theory will be used to handle the incomplete information system, where some of the attribute values in information system are unknown or missing. The new proposed method will calculate and suggest the relevant terms, such as subject or object, that are more likely to replace the unsure value. The proposed method has been tested and evaluated with a collection of domain texts that describe tourism. Precision and recall metrics have been used to evaluate the results of the experiments. The output of this proposed method will be significantly used in the conceptualization process of the ontology engineering process to assist ontology engineers and beneficial to obtain valuable information from a variety of sources of natural language text description such as journal, structured databases of any domain, and also enable to facilitate big data analysis.

Keywords Concept · Incomplete sentence · Non-taxonomic relation
Ontology extraction · Probability theory · Relation extraction

N. F. Nabila (✉) · N. Basir
Universiti Sains Islam Malaysia (USIM), 71800 Nilai, Negeri Sembilan, Malaysia
e-mail: fatin@usim.edu.my

N. Basir
e-mail: nurlida@usim.edu.my

M. M. Deris
Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia
e-mail: mmustafa@uthm.edu.my

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_20

20.1 Introduction

Various works have focused on ontologies as they have potential in many application areas, such as text mining, information retrieval, knowledge management and the Semantic Web. Ontology provides a description of a certain domain of concern that consists of several components, such as axioms, instance, concept and relation. However, most of the ontologies are constructed manually, which is a difficult task, costly and time-consuming [15].

Considerable works have been completed to construct an ontology from domain texts. Nevertheless, most of these works have only focused on extracting concepts and taxonomic relationships, such as is-a relation only, with very little work on extracting non-taxonomic relationship. The extraction of non-taxonomic relationships are considered as one of the most challenging and important tasks [6, 8, 21]. Some works have been proposed for identifying non-taxonomic relations, focusing on identifying a given specific relationship, such as part-whole [19, 20] and cause-effect [3]. However, these works are not able to identify other relationships that are crucial for the domain.

Existing research [2, 5, 11, 17] on extracting non-taxonomic relations between two concepts (terms) focus on terms that appear as subject and object in a single sentence. For example, in the sentence, "Each student need to read 3 books every week". The terms student and book are identified as subject and object of a sentence and read is a relationship that relates subject student and object book. A problem arises if the two concepts (student and book) do not exist in the same sentence, then these two concepts will not be considered in the construction of ontologies. However, if the sentence is irregular: either an object or a subject is missing or not clear, then the relationship between concepts is not extracted. For example, the sentence, "The students read it every day" is considered as an irregular sentence as it does not have a clear object. It is not clear what the object "it" is referring to in the sentence. In this case, it is assumed that the missing value is "unsure" or as an "unknown" condition. The object "it" may have been described in the previous sentence, but in the existing techniques, relations from sentences that have an uncertain value are not extracted. As a result, the domain texts can be considered as not properly represented, as some relations cannot be identified. Therefore, this paper is to investigate the feasibility of developing an alternative technique to overcome the issue of missing potential relations, which are not handled by the previous methods. This work is a continuation of our previous works [10].

The paper is organized as follows: Sect. 20.2 describes literature review. In Sect. 20.3, the proposed method is explained. Section 20.4 presents the experiment of the method. Finally, Sect. 20.5 presents the conclusions and future work.

20.2 Literature Review

Ontology has been used in many areas such as information retrieval and semantic web. According to [16], an ontology can be generated from different sources such as set of documents, existing ontologies, or from a combination of both sources. It can also be generated from scratch. Figure 20.1 shows the general process of ontology construction from text. In Fig. 20.1, the general process of ontology construction from texts consists of extracting ontological components (such as concept, relation) from text, which then creates an ontology.

The relations between concepts are known as taxonomic and non-taxonomic relations. A relationship between concepts is needed to explain more about the domain. A taxonomic relation represents a hierarchy of concepts, i.e. is-a relation. For example, mother is-a person. A non-taxonomic relation represents relation other than an "is-a" relation that exists in texts. The extraction of non-taxonomic relationships are considered as one of the most challenging and important tasks [11, 12, 21]. Many existing techniques [4, 11–13] focused on extracting relationships between two concepts focus on terms that appear as subject and object in a single sentence. Villaverde et al. [17] proposed a technique that identifies nouns as concepts and verbs that hold a place between two nouns that occur in a single sentence, as a relationship. All nouns and verbs are identified by using parts-of-speech (POS) taggers that were applied to each sentence from the documents collection to fulfill the pattern: <term><verb><term>, where the terms are identified nouns that also exist in ontology concepts. In both works, if the concepts do not exist in ontology concepts, then the relation is not identified. Punuru and Chen [11 and Serra and Giradi 13] also used the predicate as a relation between two concepts. In contrast to [17], this work does not refer to ontology concepts to find the relevant concepts. Punuru and Chen [11] proposed the Subject-Verb-Object (SVO) Triples method to identify non-taxonomic relations between two concepts, where the concepts must appear as the subject and the object of a sentence. They used MINIPAR dependency parser to determine the appearance of concepts. Then, the verb that occurred together with the concept pair was identified. Serra and Girardi [13] proposed a technique that used an NLP approach and data mining technique to identify potential non-taxonomic relationships from textual sources. Serra et al. [14] proposed a semi-automatic method called PARNT to extract non-taxonomic relations from texts. Ribeiro [12] proposed a framework that used A Nearly New Information Extraction System (ANNIE), Stanford dependency parser

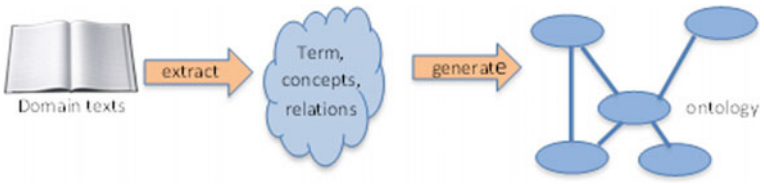


Fig. 20.1 Ontology construction processes from text

and association technique, to enrich ontology relations. The framework extracted relations between two concepts, where the concepts must appear as the subject and the object of a sentence, as similar to [11]. This work extracted non-taxonomic relationships of the domain of tennis sport collected from various sources. Muzaffar et al. [9] proposed a hybrid framework that used four features such as the bag of word model, NLP approach, Lexical and semantic based UMLS, to extract relation from biomedical datasets collected from MEDLINE database. This work extracted all verb phrases that occur between treatments and disease entities in the sentence. All these techniques only able to extract non-taxonomic relations between two concepts i.e. subject and object that appear in the same sentence. However, the existing methods do not cover if the subject or the object of a sentence is "unclear" or "missing".

Therefore the domain texts may improperly presented, as some concepts and relations cannot be identified. Hence, this paper proposed a technique based on probability thoery for suggesting the relevant concept for missing concept in incomplete sentence.

20.3 The Method

This section describes the proposed method for extracting concepts and relationship for constructing the ontology from domain texts. Figure 20.2 shows the flow of the approach. This approach involves the following steps: (1) The Extraction of Concept and Predicate, (2) Generating Subject-Object Pair and (3) Relation Labeling.

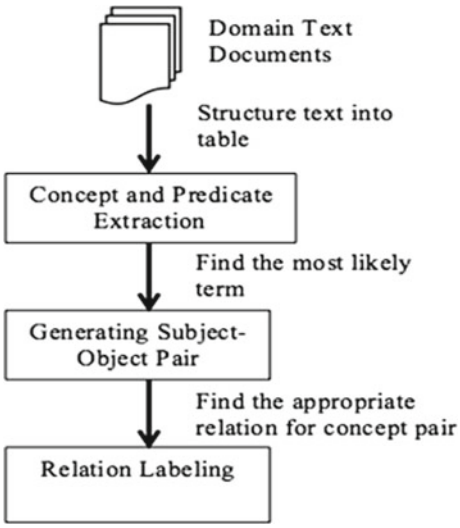


Fig. 20.2 Flow of the approach design

20.3.1 The Extraction of Concept and Predicate

In this step, three main text-preprocessing steps, such as, part-of-speech (POS) tagging, stop word removal, and morphological analysis, are used to extract terms from texts. Statistical analysis is used to determine the relevant term as a concept in domain text. Each term is produced during a preprocessing step and calculated using a statistical measurement such as term frequency, to determine its relevancy to the domain. Then, the dependency pairs between these terms (i.e., grammatical relation between subject and/or object with the predicate) are identified using the Minipar shallow parser [7]. All identified terms and their relations are presented in the information table. In this table, all incomplete sentences with missing object or subject are highlighted by '*'. This "*" is known as a "unclear" value of the sentences.

20.3.2 Generating Subject-Object Pair

This phase determines the most likely terms to replace the "unsure" value of the concept (i.e. Subject or object) in order to complete the ontology component table. Two steps were involved to identify most-likely term: (1) Synonym Predicate Match, (2) probability of most likely term. The first step is used when the predicates in the ontology component have the synonym of predicates. Then, both predicate can be used as relations between the concepts. If the ontology component table still has unsure value after Step 1, then the second step is used to replace the uncertain subject or object with the most likely term suggested by probability theory. As an example, the voting machine text was used as a case study. All sentences were extracted and presented in the ontology component table as shown in Table 20.1. In Table 20.1, u1, u2, …, u10 represent sentences. C is an attribute of sentences that consist of subject, object and predicate of sentence.

Table 20.1 Part of sentences in voting machine dataset

Sentence ID	C		
	Subject	Object	Predicate
u1	Voter	Machine	Trust
u2	Company	*	Supply
u3	Machine	Paper	Produce
u4	Voter	*	Check
u5	Machine	Record	Produce
u6	*	Record	Produce
u7	Government	Machine	Provide
u8	Voter	*	Trust
u9	*	Name	Verify
u10	Machine	Record	Evaluate

Step 1: Synonym Predicate Match

This phase identified the most-likely terms by using the synonymous predicate as defined in Definition 20.1. Here, synonym predicate consists of different predicates, but has a similar meaning. The synonymous of predicate is referred to WordNet.

Definition 20.1 Let Ci = {si, oi, pi} be a complete regular sentence and Cj = {sj, *, pj} be a irregular sentence in text. If (pi is equivalent to pj) then * = oi.

Based on Definition 20.1, if pi is equivalent to pj, then this work handles two scenarios as follows:

(1) Scenario 1

For example, in Table 20.1, the predicate supply in u2 is synonymous with the predicate provide in u7. Since the predicates are synonym, then in this work, both predicate can be considered as relations between the subject and object even when the subject and object appear in different sentences. Thus, the object in u7 (i.e. machine) is selected as most-likely term to replace the uncertain value of object in u2. Since the uncertain value of object in u2 is replaced with machine, then u2 will has one set of triples.

$$u2 = \{company, *, supply\}, u7 = \{government, machine, provide\},$$
$$If(supply \text{ is equivalent to } provide), then * = machine.$$
$$Thus, u2 = \{company, machine, supply\}.$$

(2) Scenario 2

For example, in Table 20.1, both u4 and u9 are irregular sentences where u4 has uncertain value of object, while u9 has uncertain value of subject. Since the predicate check in u4 is synonymous with the predicate verify in u9, the object name in u9 is selected as most-likely term to replace the uncertain value of object in u4. While, the subject voter in u4 is selected as most-likely term to replace the uncertain value of subject in u9.

$$u4 = \{voter, *, check\}, u9 = \{*, name, verify\}$$
$$If(check \text{ is equivalent to } verify), then * \text{ in } u9 = voter, *\text{in } u4 = name$$
$$Thus, u4 = \{voter, name, check\}, u9 = \{voter, name, verify\}$$

Table 20.2 shows the updated incomplete ontology component table after step one. Based on Table 20.2, the table is incomplete ontology component table because the uncertain values still exist. Therefore, the second step, i.e. probability of most likely term, will be used in the following section.

Table 20.2 Incomplete ontology component table after phase 1

Sentence ID	Subject	Object	Predicate
u1	Voter	Machine	Trust
u2	Company	Machine	Supply
u3	Machine	Paper	Produce
u4	Voter	Name	Check
u5	Machine	Record	Produce
u6	*	Record	Produce
u7	Government	Machine	Provide
u8	Voter	*	Trust
u9	Voter	Name	Verify
u10	Machine	Record	Evaluate

Step 2: Probability of Most Likely Term

The second step of the technique is to find the most likely term to replace the unsure value. In probability, the two events X and Y are called independent if the occurrence of Y had no impact of the occurrence of X. Otherwise, X and Y are called dependent. Theorem 20.1 is used to calculate the probability if there are more than two independent events in an experiment [18].

Theorem 20.1 *If in an experiment, the events, X_1, X_2, X_3, …, X_k are independent, then*

$$P(X_1 \cap X_2 \cap X_3 \cap \ldots \cap X_k) = P(X_1)P(X_2)P(X_3)\ldots P(X_k).$$

In this work, the probability theory is used to handle the irregular pattern of a sentence. In our work, event X_1 refers to the term that appears as the subject in a sentence, event X_2 is referring to the term that appears as the object of a sentence and X_3 is referring to the term that appears as the predicate of a sentence. We assumed X_1, X_2 and X_3 are independent. In this work, Theorem 20.1 is used to calculate the probability of subject, object and predicate independent events will occur in domain texts. The highest probability value is selected as most-likely term to replace the uncertain value.

For example, consider the probability of {computer, program, execute} from 10 sentences where there are 7 computer as subject, 5 program as object and 5 execute as predicate in the sentences. Based Theorem 20.1, X_1 is referring to computer, X_2 is referring to program and X_3 is referring to execute. Thus, the probability that independent events, X_1, X_2 and X_3 will occur in sentences is

$$\mathrm{P}(\mathrm{X}_1 \cap \mathrm{X}_2 \cap \mathrm{X}_3) = \mathrm{P}(\mathrm{X}_1)\mathrm{P}(\mathrm{X}_2)\mathrm{P}(\mathrm{X}_3) = \frac{7}{10} * \frac{5}{10} * \frac{5}{10} = 0.175.$$

Thus, in this work, the unsure value can be calculated using the probability theory defined as follows:

Algorithm: Determination of most-likely term
Input: incomplete ontology component table
Output: most-likely term
Begin
Step 1. Calculate the probability for each attribute (i.e. subject, object, predicate)
Step 2. Calculate the probability of set C where C = {subject, object, predicate}
Step 3. Determine the probability of same subject and object pair
Step 4. Determine the most-likely term
End

Fig. 20.3 The determination of most-likely term algorithm

Definition 20.2 The probability of most-likely term is denoted as P(Dm). This can be defined as

$$P(Dm) = D/|U|$$

where,

- D is probability that the same subject and object with predicate occur
- |U| is total number of sentences in ontology component table

The highest probability value was selected as most-likely term to replace the uncertain value. The algorithm for selecting most-likely term using probability is given in Fig. 20.3. In Fig. 20.3, the algorithm used incomplete ontology component table as an input. The determination of most-likely term algorithm consists of several main steps. Each step details in the algorithm are described in Fig. 20.3.

(a) Calculating the probability of each attribute of set C.

The probability of each attribute of set C (i.e. subject, object and predicate) that occurs in a sentence will be calculated. Definition 20.3 is to calculate the probability of attributes that exist in a sentence. Definition 20.4 is used to calculate the attribute if the attribute has uncertain value.

Definition 20.3 Let C = {s, o, p}. If an attributes x has value other than '*', then P(x) = 1, i.e., the probability of x is 1.

Example 20.1 From Table 20.2 the probability of subject voter, P(voter) in u1 is P(voter) = 1.

Similarly for probability of object machine, P(machine) in u1 is 1.

Definition 20.4 Let C = {s, o, p}. If there exists '*' value (i.e. uncertain value) for attributes, then the probability of not '*' attributes (i.e. subject/object) in the table is calculated based on the formula below.

$$P(x) = (|(x)|/U,\ x \in C)$$

where

- x is not '*' value that exist in the correspondence attribute (subject/object),
- (||) represents the cardinality of the sets,
- U represents number of sentences in ontology component table.

Example 20.2 From Table 20.2, the object (i.e. '*') of u8 is a missing value, thus the probability of other terms that could appear as object (i.e. machine, paper, name and record) are calculated as follows:

For the set of attribute object = {*, machine, paper, name, record},
P(*) = 1/10,
P(machine) = 3/10,
P(paper) = 1/10,
P(name) = 2/10
P(record) = 3/10

In this example, attribute object has uncertain value and other terms that appear as object in Table 20.2 are machine, paper, name and record. Thus, the probability of other terms that could appear as object was calculated.

(b) Calculate the probability of set C

The probability that three independent events (i.e. subject, object and predicate) will occurs in a sentence, P(Ci), is calculated by using a formula defined below.

Definition 20.5 The events s, o and p are independent, then P(Ci) = P(si) * P(oi) * P(pi)

where

- c is the three independent attribute (i.e. subject, object, predicate)
- i is referring to number of sentence the triplet occurs
- Event s is referring to term that appear as subject in a sentence
- Event o is referring to term that appear as object of a sentence
- Event p is referring to term that appear as predicate of the subject

Example 20.3 Based on Table 20.2, the probability of C for u1 and u8 are calculated as follows:

- For u1,
  $C = \{\text{voter},\ \text{machine},\ \text{trust}\}, P(Cu1) = 1 * 1 * 1 = 1$

SentenceID	Subject-object	Predicate	D
u6	(Voter, record)	produce	4/10
	(machine, record)		13/10
	(company, record)		1/10
	(government, record)		1/10
u8	(voter, machine)	trust	13/10
	(voter, paper)		1/10
	(voter, name)		2/10
	(voter, record)		3/10

Fig. 20.4 The result of step 3

- For u8,
  $C = \{voter,\ machine,\ trust\}, P(Cu8) = 1 * 3/10 * 1 = 3/10$
  $C = \{voter,\ paper,\ trust\}, P(Cu8) = 1 * 1/10 * 1 = 1/10$
  $C = \{voter,\ name,\ trust\}, P(Cu8) = 1 * 2/10 * 1 = 2/10$
  $C = \{voter,\ record,\ trust\}, P(Cu8) = 1 * 3/10 * 1 = 3/10$

In this example, the decision value for u1 is 1 and u8 has four values since there exists four probabilities of terms (i.e. machine, paper, name and record) that appear as object in Table 20.2 to replace the object is '*'.

(c) Determine the probability of same subject and object pair

The probability that has the same subject and object with predicate in step b was identified using the formula below:

$$D = \sum_{\text{same s,o}} P(C_i)$$

Figure 20.4 show the result of step c for Table 20.2.

(d) Determine the most-likely term

The most likely term for uncertain value was calculated by using Definition 20.3.

Example 20.4 Based on Fig. 20.4, the results of probability of most likely term for u6 are

(voter, record) $\Rightarrow$ oduce, P(Dm) = 0.4/10 = 0.04
(machine, record) $\Rightarrow$ oduce, P(Dm) = 1.3/10 = 0.13
(company, record) $\Rightarrow$ oduce, P(Dm) = 0.1/10 = 0.01
(government, record) $\Rightarrow$ oduce, P(Dm) = 0.1/10 = 0.01

Here, the probability of machine and record with predicate produce is the highest probability value. Therefore, machine is considered as most-likely term for uncertain value of object in u6. Table 20.3 shows the complete ontology component table for Table 20.1. In this table, all the uncertain values are replaced with the most likely terms for the uncertain value.

Table 20.3 Complete ontology component table

Sentence ID	Subject	Object	Predicate
u1	Voter	Machine	Trust
u2	Company	Machine	Offer
u3	Machine	Paper	Produce
u4	Voter	Machine	Check
u5	Machine	Record	Produce
u6	Machine	Record	Produce
u7	Government	Machine	Provide
u8	Voter	Machine	Trust
u9	Voter	Name	Verify
u10	Machine	Record	Evaluate

20.3.3 Relation Labeling

In this phase, the support and confidence metric in association rule [1] have been used to identify the most appropriate relations among concepts (i.e. subject-object pair). In this paper, the predicate that has the highest value of confidence for the subject-object pair is selected as the most appropriate relation for the pair. The confidence is high if the subject-object pair co-occurred frequently with the predicate in domain texts.

20.4 Experiments

For conducting the experimental evaluation, a tourism datasets was used. Tourism corpus was collected from Wikipedia websites and the Los Angeles Time website and consisted of 65 texts and over 29,000 words describing tourism. To evaluate our method, the prototype based on the proposed method was developed using Java and Javascript. The existing work, namely methods by [14], was developed and tested by using the same texts. In this paper, all domain texts were given to the experts for them to identify all relevant relations of the domain texts manually and the results were used as benchmarks for the system. The results produced by the proposed solution, and [14], were compared with the results produced by domain experts. The experiments' results were then analyzed using precision and recall to measure the relevancy and quality of the extracted relations.

The evaluation goal is to analyze how much the extracted relations produced by the proposed method improved the coverage of knowledge of domain texts. In this evaluation, all texts in the same domain are uploaded to the prototype and then the relations (i.e. subject, object, and predicate) are extracted. The relations from the tourism dataset that are extracted are shown in Fig. 20.5. In Fig. 20.5, the sen-

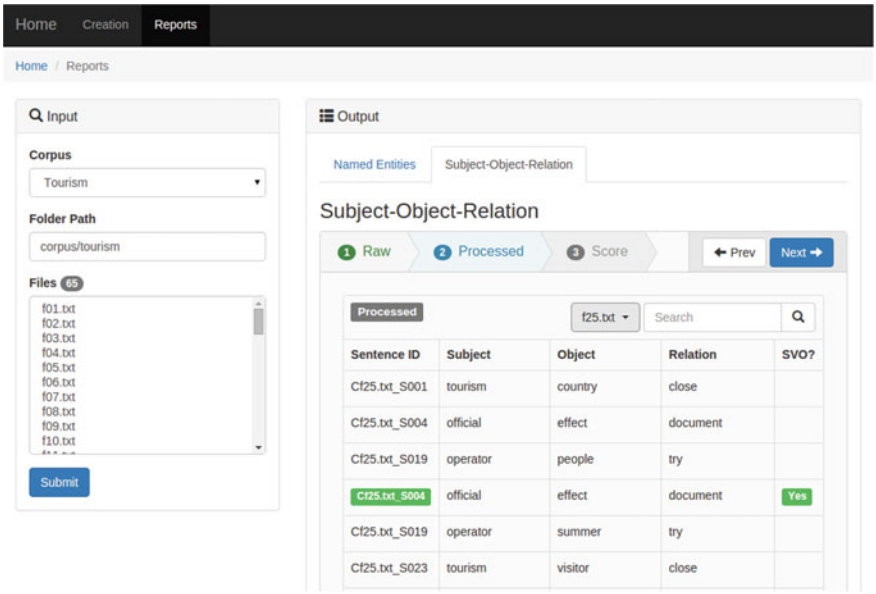


Fig. 20.5 The extracted relations for tourism corpus

tence ID highlighted in green indicates the complete sentences that have the subject, object and predicate co-occur in the same sentence. While the remaining lists are irregular sentences in raw texts, which have an uncertain term (i.e. subject/object). Based on the proposed solution, it extracted the relations and determined the possible subject/object to replace the uncertain term by using synonyms of predicate and probability.

In this paper, the support and confidence metric in association rule have been used to identify the most appropriate relations among concepts (i.e. subject-object pair). A part of the confidence value of relations between subject-object pair with the predicate in tourism corpus is illustrated in Fig. 20.6. In this work, the predicate that has the highest value of confidence for the subject-object pair is selected as the most appropriate relation for the pair. The confidence is high if the subject-object pair co-occurred frequently with the predicate in domain texts. For example, in Fig. 20.6, confidence values between (tourism, economy) with predicate affect is 0.4, which is higher than the other predicate such as diversity and suffer. Thus, in this work, predicate affect has been selected as the most appropriate relation to label the relationship between tourism and economy.

The evaluation results for both methods is presented in Table 20.4. From Table 20.4, for a tourisim corpus, the Serra method extracted 232 relations of which 127 are correct relations and 105 are incorrect relations. In contrast the proposed method extracted 456 relations of which 356 are correct relations, which is higher than the Serra method. Based on the same domain texts, the domain expert has

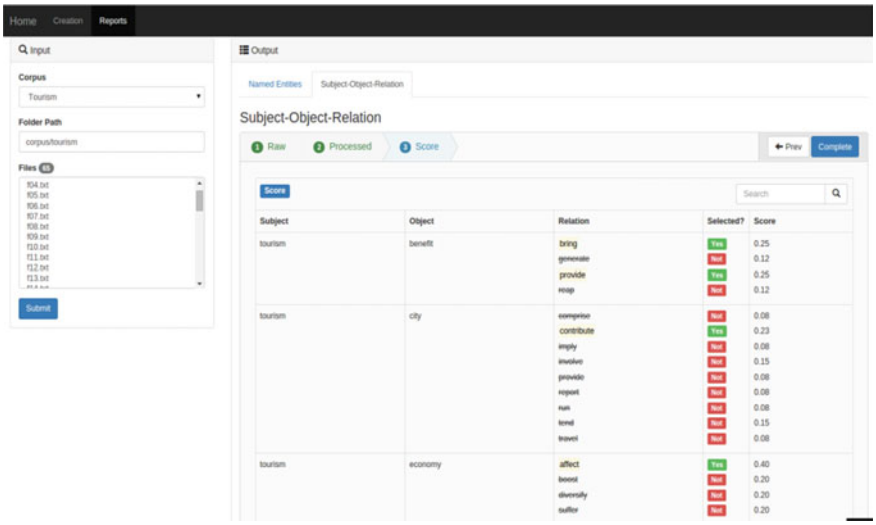


Fig. 20.6 Part of the confidence value of relations between subject-object pairs with the predicate in the tourism corpus

Table 20.4 Performance of proposed method

Method	# of extracted relations	# of correct relations	# of incorrect relations	Precision	Recall (%)
Serra method	232	127	105	54.74%	34.51
Proposed method	456	356	100	78.07	96.74
Domain expert	368	1634	–	100%	100

identified 368 valid relations. Based on domain expert, the proposed solution obtained 96.74%, which is higher than the Serra method [14]. Meanwhile, for precision value, the proposed method has achieved 78.07% (based on Expert 1), which is slightly higher than the Serra method.

Table 20.5 shows that the correct relations extracted from the proposed method based on synonym predicate and probability theory are able to extract non-taxonomic relationships between subjects and objects that occur not only in the same sentence, but also in different sentences, and for unsure values in irregular sentences. Therefore, based on the evaluation results, it shows that the proposed method helps in the enrichment of the domain ontology to represent the corpus.

Table 20.5 Results for non-taxonomic relations

Expert	# of correct relations	# of relation where the two concepts in	
		Same sentence	Different sentences
Proposed method	356	127	229
Serra	232	127	–

20.5 Conclusion and Future Work

In conclusion, this paper has presented an approach to extract ontology component, which focus on non-taxonomic relationships. The proposed method calculates the accuracy of the terms to suggest the most-likely term to replace the unsure subject or object. The evaluation of experimental results shows that the use of synonymous predicate and probability are able to increase the number of relations or knowledge to represent the domain.

Acknowledgements Universiti Sains Islam Malaysia (USIM). Grant: PPP/USG-0116/FST/30/11616.

References

1. R. Agrawal, T. Imieliński, A. Swami, Mining association rules between sets of items in large databases. ACM SIGMOD Rec. **22**(2), 207–216 (1993)
2. A. Akbik, J. Brob, Wanderlust: extracting semantic relations from natural language text using dependency grammar patterns, in *WWW Workshop* (2009)
3. N. Asgha, *Automatic Extraction of Causal Relations from Natural Language Texts: A Comprehensive Survey*. arXiv:1605.07895
4. T.V. Geetha, Unsupervised domain ontology learning from text, in *4th International Conference, MIKE 2016 on Mining Intelligence and Knowledge Exploration*, Mexico City, Mexico, November 13–19, 2016, Revised Selected Papers, vol. 10089 (Springer, 2017), p. 132
5. A. Imsombut, A statistical approach for semantic relation extraction, in *Eighth International Symposium on Natural Language Processing, 2009. SNLP'09* (IEEE, 2009), pp. 54–58
6. N. Kaushik, N. Chatterjee, A practical approach for term and relationship extraction for automatic ontology creation from agricultural text, in *2016 International Conference on Information Technology (ICIT),* (IEEE, 2016) pp. 241–247
7. D. Lin, *Dependency-Based Evaluation of MINIPAR* (In Treebanks, Springer, Netherlands, 2003), pp. 317–329
8. L. Liu, X. Ren, Q. Zhu, S. Zhi, H. Gui, H. Ji, J. Han, *Heterogeneous Supervision for Relation Extraction: A Representation Learning Approach*. arXiv:1707.00166 (2017)
9. A.W. Muzaffar, F. Azam, U. Qamar, A relation extraction framework for biomedical text using hybrid feature set. Computat. Math. Methods Med. (2015)
10. N. F. Nabila, B. Basir, A. Mamat, M.M. Deris, Improving knowledge extraction from texts by generating possible relations, in *Proceedings of the World Congress on Engineering and Computer Science*, *Lecture Notes in Engineering and Computer Science* 2017, 25–27, San Francisco, USA, pp. 55–60

11. J. Punuru, J. Chen, Learning non-taxonomical semantic relations from domain texts, J. Intell. Informat. Syst. 191–207 (2012)
12. Ribeiro, *Extraction of Non-taxonomic Relations from Texts to Enrich a Basic Ontology* (2014)
13. I. Serra, R. Girardi, A process for extracting non-taxonomic relationships of ontologies from text. Intell. Informat. Manag. **3**(4) (2011)
14. I. Serra, R. Girardi, P. Novais. PARNT: a statistic based approach to extract non-taxonomic relationships of ontologies from text, in *2013 Tenth International Conference on Information Technology: New Generations (ITNG)*, (IEEE, 2013) pp. 561–566
15. M. Shamsfard, A.A. Barforoush, Learning ontologies from natural language texts. Int. J. Hum Comput Stud. **60**(1), 17–63 (2004)
16. M. Uschold, Creating, integrating and maintaining local and global ontologies, in *Proceedings of the First Workshop on Ontology Learning (OL-2000) in Conjunction with the 14th European Conference on Artificial Intelligence (ECAI-2000)* (2000)
17. J. Villaverde, A. Persson, D. Godoy, A. Amandi, Supporting the discovery and labeling of non-taxonomic relationships in ontology learning. Expert Syst. Appl. **36**(7), 10288–10294 (2009)
18. D.E. Walpole, R.H. Myers, S.L. Myers, *Essentials of Probability and Statistics for Engineers and Scientists* 1st ed. (Prentice Hall, 2012)
19. F. Xia, C. Cao, Extracting part-whole relations from online encyclopedia, in *International Conference on Intelligent Information Processing* (Springer, Berlin, Heidelberg, 2014), pp. 57–66
20. T. Yıldız, S. Yıldırım, B. Diri, Extraction of part-whole relations from Turkish corpora, in *International Conference on Intelligent Text Processing and Computational Linguistics*, (Springer, Berlin, Heidelberg, 2013) pp. 126–138
21. R. Zarrad, N. Doggaz, Zagrouba, E., Using heading hierarchy for non-taxonomic relation extraction, in *2016 IEEE/ACS 13th International Conference of IEEE Computer Systems and Applications (AICCSA)*, pp. 1–7 (2016)

Chapter 21
Insider Threat Veracity Issues

William R. Simpson and Kevin E. Foltz

Abstract Enterprise Level security (ELS) has no accounts or passwords, and consequently identity is an important issue. All person and non-person entities in ELS are registered and known. PKI credentials are issued, and when necessary, multi-factor authentication is used to improve the assurance of the identity. Because the next step in ELS is claims-based access and privilege, many data owners are worried about the trustworthiness (sometimes called reputation) of the identified requesters (this applies to person and non-person entities within the enterprise). Individuals are vetted periodically, and a baseline is established by those instances; however, activities that occur between those vetting events may provide clues about the trustworthiness of the individuals. Similarly, pedigrees in software and hardware entities are established periodically. Because the terms trust and integrity are overloaded, we refer to these data as veracity. Further, when requested, the veracity that applies to certain categories will be provided as counter-claims along with the claims. These counter-claims may be used by the applications and services for increased levels of surveillance and logging and perhaps even limitation of privilege. The computation of veracity brings about security concerns and requires special handling. This paper reviews the data categories, data requirements, security issues, and data resources that apply to entity veracity, as well as the counter-claim structures and issues associated with their tracking and usage. The paper then presents findings and recommendations, along with the future work necessary to complete this evolution.

Keywords Behavior · Claims · Counter-Claims · Insider threat · Integrity Reputation · Motivation · Veracity

W. R. Simpson (✉) · K. E. Foltz
Institute for Defense Analyses, 4850 Mark Center Drive, 22311 Alexandria, VA, USA
e-mail: rsimpson@ida.org

K. E. Foltz
e-mail: kfoltz@ida.org

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_21

21.1 Introduction

This work is based partly upon a paper presented at WCECS17 [1]. The insider threat must be monitored and assessed, especially for those subject to executive orders. Since Edward Snowden [2], Bradley Manning [3], and others [4], we simply have no choice but to assess our own insider threat situation. An insider threat is:

> … a malicious threat to an organization that comes from people within the organization, such as employees, former employees, contractors or business associates, who have inside information concerning the organization's security practices, data and computer systems. [5]

The manifestation of the threat may come from any entity in the environment, person or non-person. The spate of insider activity has led to a U.S. executive order [6] that requires, in part, federal agencies and enterprises to:

> …perform self-assessments of compliance with policies and standards issued pursuant to sections 3.3, 5.2, and 6.3 of this order, as well as other applicable policies and standards, the results of which shall be reported annually to the Senior Information Sharing and Safeguarding Steering Committee established in section 3 of this order….

For Enterprise Level Security (ELS) [7] federal applications, we must include these self-assessments. The requirement has led to the development of new products and an overwhelming volume of white papers and other research telling us how some vendors would do this assessment, and a number of patents pending [8–11]. All of this leads to a number of product offerings to perform the analysis of entity veracity within the enterprise. A summary of these techniques (through 2011) is provided in [12]. The basic idea is to gather information concerning the trustworthiness of an entity in our system, as shown in Fig. 21.1. This is an ELS adaption of the figure presented in patent application [9].

Figure 21.1 shows the security issues associated with just the computation of an entity's trustworthiness. Access to public records and sources that are not vetted opens vulnerabilities not tolerated in a high assurance environment such as ELS. The initial implementation will be done in isolation from the enterprise and data will be ported to the enterprise. The figure shows the desired ultimate architecture where the computation is isolated and the enterprise may be provided a read only interface of the results (which may initially be a mirror of the actual veracity store). Two additional concerns in the figure include a read only interface from the computation environment to the enterprise attribute store (which may initially be a mirror of the actual enterprise attribute store), and a read only interface from the computation environment to the enterprise support desk behavioral data (which may initially be a mirror of the actual enterprise support desk behavioral data). Paranoia is warranted when dealing with unclean data and the entire insider threat analysis system will be heavily monitored, and sanitized often with complete software re-installation at periodic intervals.

This paper presents a form of self-assessment that evaluates veracity from the ELS perspective rather than from the perspective of the product's baseline. This paper also

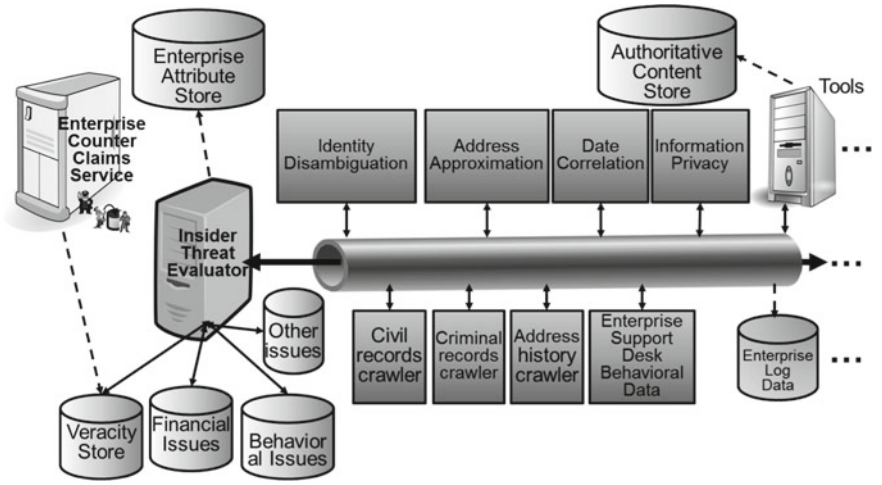


Fig. 21.1 Data gathering for insider threat analyses

addresses the issues associated with the self-assessment, and it provides a framework and a process for using veracity information within ELS. To do this, we examine integrity, reputation, and veracity as they apply to the problem of the insider threat.

21.2 Integrity, Reputation, and Veracity

Generally, the determination of trustworthiness of an individual is based upon an assessment of the integrity of that individual. One definition of integrity is given below:

> Integrity is the quality of being honest and having strong moral principles; moral uprightness. It is generally a personal choice to hold oneself to consistent moral and ethical standards. In ethics, integrity is regarded by many people as the honesty and truthfulness or accuracy of one's actions. [13]

Social media would define this as reputation, which is good because integrity is already over-used in the information technology (IT) literature. However, the literature defines reputation as a soft issue.

> Reputation is the estimation in which a person or thing is held, especially by the community or the public generally. [14]

Microsoft has refined reputation by adding trust:

> Reputation Trust represents a party's expectation that another party will behave as assumed, based upon past experience. Reputation Trust is bidirectional and can be split into Consumer Reputation Trust and Provider Reputation Trust. [15]

But trust is an overloaded term in information technology and requires a great deal of context. The dictionary description of veracity comes closer to the target, and it is not used in any of the IT contexts associated with ELS:

> Veracity is the quality of being truthful or honest. [16]

From the IT standpoint, we have adopted the concept of veracity and tailored its definition to be more amenable to self-assessment in ELS environments:

> *Entity Veracity* is the degree to which an entity is worthy of trust as demonstrated by resistance to or avoidance of factors that denigrate trust or compromise reliability. Positive factors may enhance veracity, and negative ones may reduce veracity. Veracity is based upon recognized accomplishments and failures, along with the associated stress factors or other trust debilitating factors present. A history of actions in difficult circumstances provides strong evidence for or against veracity.

The next step is to determine which of the factors need to be measured. But first we need to understand how identity and access control are handled within ELS.

21.3 Enterprise Level Security

The ELS design is a distributed security approach (see Fig. 21.2) that addresses five security principles derived from the basic design concepts. We address only two here, and the interested reader is directed to [7] for a more complete treatment:

- Know the Players—this is done by enforcing bi-lateral end-to-end authentication;
- Separate Access and Privilege from Identity—this is done by an authorization credential.

21.3.1 Know the Players

In ELS, the identity certificate is an X.509 Public Key Infrastructure (PKI) certificate [17]. This identity is required for all requesters and providers of services (active entities), both person and non-person, e.g., services, as shown in Fig. 21.3. PKI certificates are verified and validated. Ownership is verified by a holder-of-key check. Supplemental (in combination with PKI) authentication factors, such as identity-confirming information or biometric data, may be required from certain entities.

21.3.2 Separate Access and Privilege from Identity

ELS can accommodate changes in location, assignment, and other attributes by separating the use of associated attributes from the identity. Whenever changes to

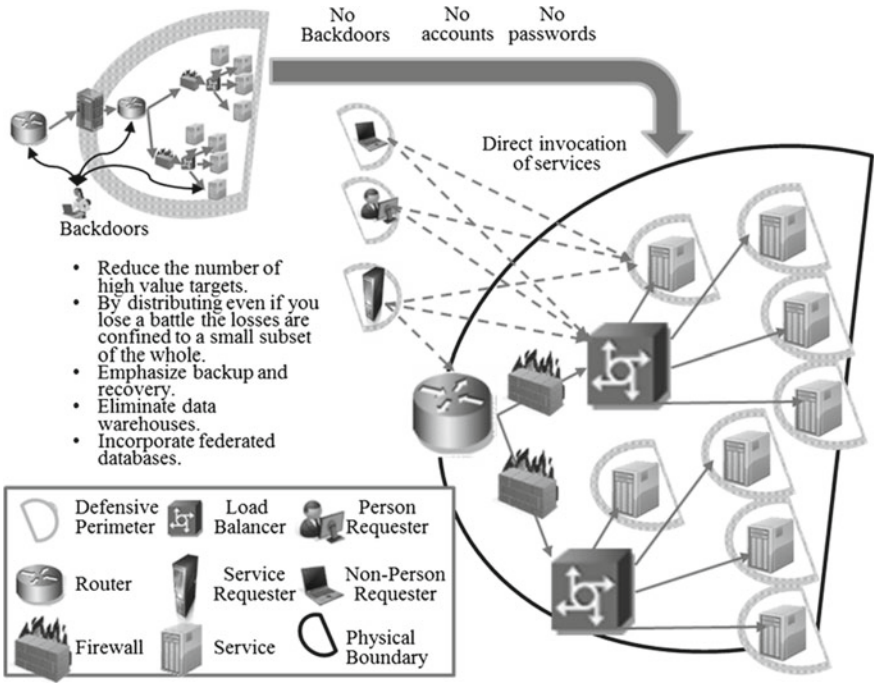


Fig. 21.2 Distributed security architecture

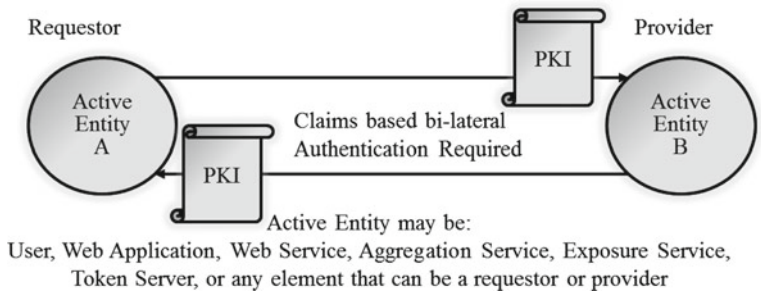


Fig. 21.3 Bi-lateral authentication

attributes occur, claims are recomputed based on the new associated attributes (see Sect. 21.3), allowing immediate access to required mission information. As shown in Fig. 21.4, access control credentials utilizing the Security Assertion Markup Language (SAML) (SAML authorization tokens differ from the more commonly used single-sign-on (SSO) tokens, and in ELS, they are not used for authentication.) [18]. SAML tokens are signed, and the signatures are verified and validated before acceptance. The credentials of the signers also are verified and validated. The credential for

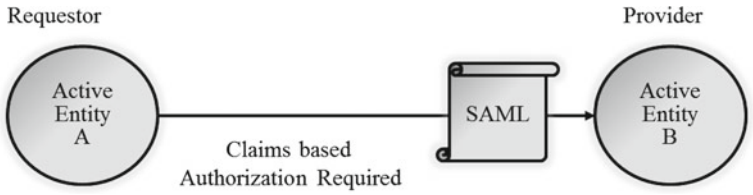


Fig. 21.4 Claims-based authorization

access and privilege is bound to the requester by ensuring a match of the distinguished name used in both the authentication and the authorization credentials.

21.4 Elements of Veracity

A list of indicative events may be formulated by category and data sources [19, 20]. We start categorization with person entities because this is required in the self-assessment, but veracity extends to all entities within the enterprise because non-person entities may actually be under insider threat control. For all entities, we assume a default value of 1.0 for veracity before detailed veracity computations are made. This is the minimum value needed to pass periodic re-evaluations, so it is assumed that all entities in the enterprise possess this value unless veracity factors indicate otherwise.

21.4.1 Person Entities

Person entity factors cover a variety of data about the person and his behaviors and these may come from a variety of sources. These data cannot be considered unless they derive from designated (by the enterprise) authoritative sources. Entity veracity factors are assigned, initially, unit values and may be combined from a number of sources. Unit values may be positive or negative (either increasing or decreasing veracity), and they are applied to veracity measures in a later section. Any previously resolved issues (through vetting or supervisor administrative judgement) may be discarded. Five categories (each with a number of subcategories and each instance is a factor) are delineated below:

Category 1 Community information—characteristics or events that add to the veracity of a person. Each adds a fixed value to overall veracity.

a. Ties within the community (positive or negative),
b. Recent job title change (positive or negative),
c. Recent relevant awards or job punishments (positive or negative),

d. Direct support or doubt from notable entities (Trust transitivity) (positive or negative).

Category 2 Financial information. Degree of debt or other financial burdens since last vetting. These may be age-and source-sensitive, and they may be attribution-sensitive, as discussed in the next section.

a. Issues with credit cards (negative),
b. Large number of credit reports (negative),
c. Recent suspicious loan activity (negative),
d. Sudden explained or unexplained wealth (negative),
e. Debt exceeds ability to pay (negative).

Category 3 Legal issues or other stress factors. These may be age-and source-sensitive, and they may be attribution-sensitive, as discussed in the next section.

a. Recent death in family (negative),
b. Poor job performance rating (negative),
c. Divorce (negative),
d. DUI (negative),
e. Felony or misdemeanor charges (negative).

Category 4 Discovered secrets. These may be age- and source-sensitive, and they may be attribution-sensitive, as discussed in the next section.

a. Attempts to hide sexual issues (negative),
b. Uncovered alternate identities (negative),
c. Residential ambiguity or multiple residences in a locale (negative).

Category 5 Unusual behavior. These will generally be from the Enterprise Support Desk Records and may be considered authoritative.

a. Non-cleared travel (negative);
b. Unusual and unexplained IT usage (negative),
    i. Unusual downloads (negative),
    ii. Unusual hours of usage (negative),
    iii. Many open applications at same time (negative);
c. Sharing of credentials (negative);
d. Frequent use of backup methods (negative);
e. Unusual delegations (negative);
f. Extended on-line absence followed by high activity (negative);
g. Unusual hours or time on-line (negative).

21.4.2 Non-person Entities

These factors will generally be from the Enterprise Support Desk Records and may be considered as authoritative. All are negative.

Category 6. Non-Person Veracity

a. Recent attacks. These are considered unless complete teardown and rebuild has happened since the attacks.
b. Recognized misuse of privilege. This may be documented through the enterprise support desk.
c. The host server is physically moved outside (or into) a protected area. All enterprise assets are registered, and the registration must be updated when any changes occur.
d. Call-out to unknown URLs. This is a known sign of exploitation, and unless the device is being used in counter-cybersecurity, it should be considered for a complete teardown and rebuild.
e. Missing log records.
f. Lenient access and privilege requirements. Privileges granted to the device may be greater than the device uses for its own access.
g. Available software interfaces that are not authorized. One clear step with ELS is to close all interfaces not being used and remove the software behind those interfaces where possible.
h. Non-uniform identity requirements on interfaces. All interfaces in use should have the same identity assurance requirements.
i. Missing current patches that are authorized. One example is Industrial Control Systems (ICS) not being patched until they have to be taken off-line.

21.5 Issues Based on Elements of Veracity

1. It is not easy to discern where an entity is facing issues that may lead to an insider problem. At the same time, acquisition of the data may have ethical and legal implications. We will briefly discuss some of the issues associated with computing veracity. Data Sources:

    a. Public and private mix. Confirming sources is problematical, and public web sites readily trade information, which obscures the original source.
    b. Attribution. (Source www.whitepages.com (3 November 2016): 190 possible matches for Frank Jones in Virginia, 50 in Richmond area, 12 are 50–65 years old, 5 of these are Frank E. Jones.) Many public sources are subject to error and may or may not have enough confirming information to provide unimpeachable attribution.
    c. Few vetted sources or authoritative content stores.
    d. PII issues. Privacy affects not only the acquired data, but the confirming data. Use of social security or other private information may assist with attribution but cross privacy lines.

2. Veracity of the veracity data:

a. In ELS, entity attributes are meticulously screened. The attribute data is required to come from authoritative sources (meaning that an organization is tasked with maintaining the accuracy and currency of the data). Even then, the data is not completely trusted and must pass sanitization and mediation before it is accepted into the attribute store. Public sources may have little more than a data entry clerk and no checks for accuracy or completeness.
b. In the public domain, sources feed one another and veracity checks may or may not be made. Old events may acquire new dates, and some of the details may get further confused.
c. Errors in the public private data. Even when data correction paths are available,
d. The data may get regenerated as in (b) above.
e. Can reputations be rebuilt? This question is real, and we do not currently have the answer. ELS users are periodically revetted, and we can assume issues that happened before that vetting will be resolved during the vetting process. One finding is to limit (by configuration) how far back insider data is considered.

3. Veracity adjustments to access or privilege may affect getting the job done. This can lead quickly to unrecoverable situations.
4. Delegation is an issue. ELS claims that are discretionary in nature may be delegated. Delegation itself may be considered a veracity issue.

21.6 Creating a Veracity Model and Counter-Claims

A simplified model is developed as a start. While weightings may be applied to the various values of veracity factors, it is best to await some actual experience with the representation before beginning that modification. In Sect. 21.4, we delineated five basic categories of veracity for person users and a single category for non-person users for evaluation, subject to data sources and correlation. Accordingly veracity is described as an n-tuple shown below:

For Persons:

$$\begin{aligned}\text{Veracity} = (&\text{Community} = \text{V1},\ \text{Financial} = \text{V2},\ \text{Legal} = \text{V3},\\ &\text{Discovered Secrets} = \text{V4},\ \text{and Behavior} = \text{V5})\end{aligned} \tag{1}$$

For Non-Persons:

$$\text{Veracity} = \text{V6} \tag{2}$$

Further, each value, Vi, has a default value of 1.0 which is appreciated by ΔV for each of the unique factors in each category. For example, using category 1:

$$(\Delta \mathrm{V}1)_k = (\pm 0.1) * \text{source factor1} * \text{source factor 2} \tag{3}$$

for every unique occurrence, k, of a factor in paragraph 4.1 item #1.

The default value of 1.0 is reduced by ΔV for each of the unique factors in categories 2–6 where applicable.

$$(\Delta \mathrm{V}i)_k = (\pm 0.1) * \text{source factor1} * \text{source factor 2} \tag{4}$$

where $i = 2 - 6$ for every unique occurrence, k, of each subcategory in paragraph 4.1 or 4.2.

Source factor 1 is 0.5 for publicly derived data, and 0.25 for publicly derived data without source citation or date of item. Source factor 1 is 1.0 for authoritative source data. Source factor 2 is 0.5 where attribution is approximate and 1.0 where attribution is certain.

$$\mathrm{V}i = 1.0 + \sum\nolimits_k (\Delta \mathrm{V}i)_k \tag{5}$$

Counter claims will be provided when requested by the data owner in the registration of his/her service. The counter claims will be given as a vector of values:

$$\text{Counter Claim for a person} = (\mathrm{V1},\ \mathrm{V2},\ \mathrm{V3},\ \mathrm{V4},\ \mathrm{V5},\ \text{none}) \tag{6}$$

$$\text{Counter Claim for a non-person} = (\text{none},\ \text{none},\ \text{none},\ \text{none},\ \text{none},\ \mathrm{V6}) \tag{7}$$

Supervisors and data owners will have claims for access to component data from the insider threat server for subordinates (in the case of supervisors) and for application and service users (in the case of data owners). Issues may be marked as resolved at the supervisor's discretion (subject to attribution and logging). An example would be at periodic vetting, the supervisor may mark some issues resolved.

Actions possible:

1. Threshold for denial of access to resources. Not recommended.
2. Threshold for notification to supervisors and data owners (Recommended).
3. Reduce privilege. Not recommended. This may affect performance reviews and cause the value of veracity to further decline in a self-generated spiral.
4. Upon notification, set up a counseling session with the individual or the owner of the asset to review the issues and seek corrections (Recommended).
5. After review, the data may be manually reset, if desirable, by providing rationale and obtaining appropriate authority.

In all cases, when requested by the data owner, the counter claim will be passed in the SAML.

21.7 Conclusions and Future Work

The formulation of entity veracity provides a method to monitor insider threats, which is required by presidential directive for some but desirable by all organizations. Certain findings are appropriate at this point:

1. For persons, the data associated with information generated prior to the last formal vetting of the person may be marked as resolved at the supervisor's discretion.
2. For persons, it is not felt that automated responses are warranted at this time.
3. For persons, manual resolutions of unfavorable veracities should be implemented at this time.
4. For non-persons, automated responses may be appropriate.
5. Thresholds and responses should be worked out over time with experience.
6. Self-assessment—data as required by executive order 13587 should be summarized and reported.

The next step is a trial instantiation and the working of the unique security issues discussed in the introductory section of this paper as well as the ethical and legal issues discussed in Sect. 21.5. The veracity measures can provide a management view into the insider threat and can be used to satisfy the requirement for self-assessment. This work is part of a body of work for high-assurance enterprise computing using web services. Elements of this work are described in [7, 21–29].

Acknowledgements This work was supported in part by the U.S. Secretary of the Air Force and The Institute for Defense Analyses (IDA). The publication of this paper does not indicate endorsement by any organization in the Department of Defense or IDA, nor should the contents be construed as reflecting the official position of these organizations.

References

1. W.R. Simpson, K.E. Foltz, Enterprise level security: insider threat counter-claims, in *Proceedings of the World Congress on Engineering and Computer Science 2017, 25–27 October, 2017.* Lecture Notes in Engineering and Computer Science (San Francisco, USA, 2017), pp. 112–117
2. B. Gertz, *The Cyber Threat: Snowden—Ultimate Insider Threat Missed by NSA Security* (2016), http://freebeacon.com/national-security/cyber-threat-snowden-insider-threat-at-nsa/. Accessed 17 Apr 2017
3. S. Fishman, New York Magazine, *Bradley Manning's Army of One* (2011), http://nymag.com/news/features/bradley-manning-2011–7/. Accessed 17 Apr 2017
4. R. Francis, *CSO Online, 9 Employee Insiders Who Breached Security* (2014), http://www.csoonline.com/article/2692072/data-protection/data-protection-165097-disgruntled-employees-lash-out.html. Accessed 17 Apr 2017
5. Wikipedia, Insider Threat (2016), https://en.wikipedia.org/wiki/Insider_threat
6. B. Obama, *Executive Order 13587—Structural Reforms to Improve the Security of Classified Networks and the Responsible Sharing and Safeguarding of Classified Information* (2011)
7. W.R. Simpson, *Enterprise Level Security—Securing Information Systems in an Uncertain World*. Auerbach Publications, CRC Press. ISBN 9781498764452, p. 397 (2016)

8. A.M. Strosaker, M.T. Strosaker, *Determining Veracity of Data in a Repository Using a Semantic Network*, US Patent 8108410 B2, International Business Machines Corporation (2012), https://www.google.com/patents/US8108410. Accessed 17 Apr 2017
9. L. Geoffrey, *Candidate-Initiated Background Check and Verification*, US Patent 20050055231 A1, published 10 Mar 2005, http://www.google.com/patents/US20050055231
10. E. Shapiro, S. Mintz, System and Method for Providing Access To Verified Personal Background Data, US Patent 20040168080 A1, 26 Aug 2004, http://www.google.com/patents/US20040168080
11. Y. Zhao, L. Jianqiang, Hierarchy extraction from the websites, US Patent 20090327338 A1, Nec (China) Co., Limited. 31 Dec 2009, https://www.google.com/patents/US20090327338
12. J. Hunker, C.W. Probst, Insiders and insider threats—an overview of definitions and mitigation techniques. J. Wireless Mobile Netw. Ubiquitous Comput. Dependable Appl. (JoWUA) **2**, 4–27 (2011)
13. Wikipedia, Integrity, https://en.wikipedia.org/wiki/Integrity, Accessed November 2016
14. Dictionary.com, http://www.dictionary.com/browse/reputation, reputation. Accessed November 2016
15. G.J. van der Geest, Carmen de Ruijter Korver, Microsoft, The Architect. J. *Managing Identity Trust for Access Control*, July 2008, https://blogs.msdn.microsoft.com/nickmac/2009/05/21/the-architecture-journal/
16. Merriam-Webster, Veracity, http://www.merriam-webster.com/dictionary/veracity, accessed November 2016
17. X.509 Standards: a. DoDI 8520.2, Public Key Infrastructure (PKI) and Public Key (PK) Enabling, 24 May 2011; b. JTF-GNO CTO 06-02, Tasks for Phase I of PKI Implementation, 17 January 2006; c. X.509 Certificate Policy for the United States Department of Defense, Version 9.0, 9 February 2005; d. FPKI-Prof Federal PKI X.509 Certificate and CRL Extensions Profile, Version 6, 12 October 2005; e. RFC Internet X.509 Public Key Infrastructure: Certification Path Building, 2005; f. Public Key Cryptography Standard, PKCS #1 v2.2: RSA Cryptography Standard, RSA Laboratories, Oct 27, 2012; g. PKCS#12 format PKCS #12 v1.0: Personal Information Exchange Syntax Standard, RSA Laboratories, June 1999, http://www.rsa.com/rsalabs/node.asp?id=2138 PKCS 12 Technical Corrigendum 1, RSA laboratories, Feb 2000
18. Organization for the Advancement of Structured Information Standards (OASIS) open set of Standards: h. Ragouzis et al., Security Assertion Markup Language (SAML) V2.0 Technical Overview, OASIS Committee Draft, March 2008; i. P. Mishra et al. Conformance Requirements for the OASIS Security Assertion Markup Language (SAML) V2.0. OASIS Standard, March 2005; j. S. Cantor et al. Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0, OASIS Standard, March 2005
19. J.W. Butts, R.F. Mills, R.O. Baldwin, Developing an insider threat model using functional decomposition, in Computer Network Security, ser. Lecture Notes in Computer Science, V. Gorodetsky, I. Kotenko, V. Skormin, Eds. Springer Berlin/Heidelberg, 2005, vol. 3685, pp. 412–417, http://dx.doi.org/10.1007/11560326 32
20. R. Chinchani, A. Iyer, H.Q. Ngo, S. Upadhyaya, Towards a theory of insider threat assessment, in *Proceedings of the 2005 International Conference on Dependable Systems and Networks (DSN'05)*, Yokohama, Japan. IEEE, June–July 2005, pp. 108–117
21. W.R. Simpson, K.E. Foltz, Assured Identity for Enterprise Level Security, Lecture Notes in Engineering and Computer Science, in *Proceedings of the World Congress on Engineering*, July 2017, Imperial College, London, pp. 440–445
22. W.R. Simpson, Enterprise Level Security—Securing Information Systems in an Uncertain World, by Auerbach Publications, CRC Press. ISBN 9781498764452, May 2016, p. 397
23. W.R. Simpson, K.E. Foltz, Enterprise level security: insider threat counter-claims, in *Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science: 2017, 25–27 October, 2017, San Francisco, USA, pp. 112–117
24. W.R. Simpson, K.E. Foltz, Escalation of access and privilege with enterprise level security, in *Proceedings of the 22nd International Command and Control Research and Technology Symposium* (ICCRTS), Los Angeles, CA. September 2017, pp. TBD

25. W.R. Simpson, K.E. Foltz, Enterprise level security with homomorphic encryption, in *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017)*, Volume 1, pp. 177–184, Porto, Portugal, 25–30 April, 2017, SCITEPRESS – Science and Technology Publications
26. K. Foltz, W.R. Simpson, Enterprise considerations for ports and protocols, in *Proceedings of the World Congress on Engineering and Computer Science 2016*. Lecture Notes in Engineering and Computer Science: 19–21 October, 2016, San Francisco, USA, pp. 124–129
27. K.E. Foltz, W.R. Simpson, Simplified key management for digital access control of information objects, in *Proceedings of the World Congress on Engineering 2016*. Lecture Notes in Engineering and Computer Science: 29 June–1 July, 2016, London, U.K., pp. 413–418
28. K.E. Foltz, W.R. Simpson, Enterprise level security—basic security model, in *Proceedings of the 20th World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI,* vol. I, WMSCI 2016, Orlando, Florida, 8–11 March 2016, pp. 56–61
29. K.E. Foltz, W.R. Simpson, Access and Privilege in Secure Big Data Analysis, in *Proceedings of the International Conference on Big Data, BIG DATA 2016*, Wessex Institute. 3–5 May 2016, Alicante, Spain, pp. 193–205

Chapter 22
Why Should A Senior Citizen Be A Facebook User?

Reasons for A Senior Citizen's Positive Attitude Towards Facebook

Ramiro Augusto Rios Paredes

Abstract This analytical study proposes a Decalogue for the elderly, considering that different investigations around the world, have given beneficial credit to social interaction for the Senior Citizens' quality of life. This Decalogue seeks to shape their attitude towards Information and Communication Technologies, so that they tend to opt for the use of Facebook for social interaction with their loved ones and friends. Each one of the proposed statements of the Decalogue is based on conceptual bases, which derived from documentary research initially carried out with a heuristic perspective, and later with hermeneutical analysis. The process that was carried out in this study matches with that of an inductive approach. It was developed through qualitative research techniques, which coordinately applied in order to approach the research object in a progressive way.

Keywords Decalogue · Facebook · Psychosocial theories of aging
Quality of life · Qualitative research · Senior citizen · Social interaction

R. A. Rios Paredes (✉)
Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas ESPE, Av. General Rumiñahui S/N, 171-5-231B Sangolquí, Ecuador
e-mail: rarios@espe.edu.ec

R. A. Rios Paredes
Universidad Nacional de La Plata, La Plata, Argentina

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_22

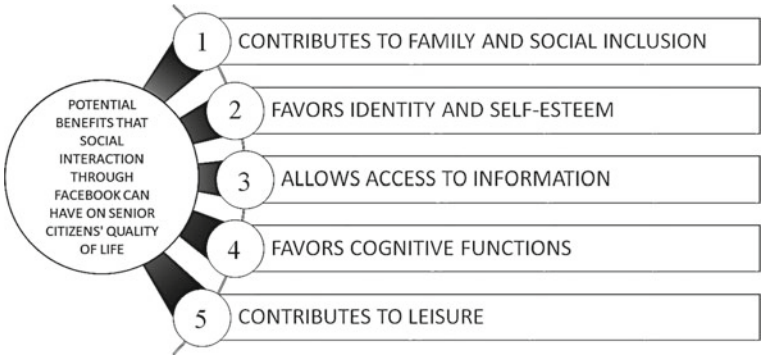


Fig. 22.1 Potential benefits that social interaction through Facebook can have on senior citizens' quality of life

22.1 Introduction

This chapter focuses on the study presented in [1], which belongs to the Social Computing Research Area ("an area of computer science that is the intersection of social behavior and computational systems, that implies two components: a social behavior component and a computational system or technical component. The technical component provides the environment in which people interact.") [2], resulted in delivering a proposed Decalogue. This Decalogue intends to shape senior citizens' attitude to lead them into social interaction through Facebook so that quality of life's enhancements get to be experienced. The research, that was carried out to obtain the Decalogue, was motivated by the following transcendental events: (1) The newest opportunities and services that Information and Communication Technologies (ICT) offer to people. (2) The current importance of enhancing the quality of life (QoL) of older adults, which might be analyzed through several approaches such as social interaction through Facebook, as it is focused in this study. (3) All the benefits regarding seniors' quality of life that have been credited to Facebook by several research studies around the globe, which may be found on Fig. 22.1, and are well-founded by publications displayed in Table 22.1.

22.2 Research Problem

22.2.1 Objective

Establishing ten key messages, based on conceptual bases, for a Senior Citizen (SC) to feel motivated to take advantage of Social Interaction through Facebook, so that he/she gets to achieve satisfactory interpersonal relationships.

Table 22.1 Analyzed studies that allowed the author to identify the benefits of Facebook social interaction for senior citizens

References	Potential benefits that have been sustained				
	1	2	3	4	5
[3]	Yes	Yes	Yes		Yes
[4]	Yes	Yes	Yes		
[5]		Yes		Yes	
[6]	Yes				
[7]	Yes	Yes	Yes		
[8]	Yes	Yes	Yes		
[9]	Yes	Yes			Yes
[10]				Yes	
[11]	Yes				Yes
[12]	Yes	Yes	Yes	Yes	Yes
[13]	Yes				
[14]	Yes		Yes	Yes	Yes
[15]	Yes	Yes	Yes	Yes	Yes

22.2.2 Research Questions

What are the messages, which should be established within the Decalogue that can encourage a senior citizen to socially interact through Facebook to achieve satisfactory interpersonal relations?

What are the conceptual bases of the messages that should be established within the aforementioned Decalogue?

22.2.3 Defining the Research Context

The research context was initially conformed by the author's closest SCs. In order to shape and sustain the messages of this Decalogue, through conceptual bases, the research context was retargeted towards diverse publications published online by reliable sources. These sources include ITU, CEPAL, and several indexed journals, focused on: Theories of Aging, senior citizens, their quality of life and well-being, and key components that might define the quality of a senior's interpersonal relationships.

22.2.4 Research Justification

This analytical study presents the conceptual bases that support each one of the messages of the Decalogue. This Decalogue intends to shape the SC's attitude to encourage their interaction through Facebook, in order to achieve the acknowledged benefits that social media has on the quality of life. It also contributes in reducing the deficiency stated in [16]: "There has been relatively little data collection and analytical work done on the impacts of ICT access and use by households and individuals." Due to the author's knowledge, resources, and adequate expertise, this study was feasible, letting him comprehend the proposals and facts from the analyzed publications.

22.2.5 Methods Utilized in This Research

The process maintained in this study had an inductive focus, which was developed by following qualitative research methods: direct observation, and documentary research. These methods were applied complementarily and coordinately to approach the research object in a progressive manner.

Direct Observation: In the beginning of the study, this method was applied to the author's immediate surroundings, in order to record the attitudes and reactions of nearby senior citizens from a chronological 65–90-year-old age range, family members and friends, towards the use of Facebook.

Documentary Research: Initially performed with a heuristic perspective, and subsequently followed by a hermeneutical analysis of the aforementioned documentary context.

22.3 Results

22.3.1 From Direct Observation

After direct observation was performed within the researcher's surroundings, two specific queries appeared:

1. What kind of circumstances allow a SC to feel comfortable when socially interacting throughs Facebook?

Whilst several SCs, specially the younger ones, proved to be thrilled when socially interacting through electronic devices, there was another group that did not showed any interest in social interaction through Facebook, arguing they already had enough interpersonal relations and a high quality of life.

When analyzing this query, direct observation led to the following concept: The first factor that motivated a SC to handle ICT alternatives, such as Facebook, is the social connection that may be achieved with family members, particularly with those ones living far away.

Being motivated, a SC turns out to feel capable of surpassing any mental barrier related to their age or their gaps of technology knowledge. Consequently, a SC becomes confident to look for help from reliable people, and learn whatever it takes to satisfy that communication interest.

As a result, a second query appeared: Is a SC still capable of learning through digital literacy? The answer has been detailed below.

22.3.2 From Documentary Research

In addition to extracting the impacts that Social Interaction through Facebook might have on a SC's quality of life (displayed in Fig. 22.1), documentary research allowed to shape the messages of the Decalogue in order for a SC to maintain a positive attitude towards Facebook, and revealed the conceptual bases of those messages.

The messages of the Decalogue have been arranged, into an ascending way, according to the following criteria 'The first messages will intend to have an influence over the SC's conception of elderly in order for him/her to project a positive attitude towards their QoL. Subsequently, those messages, which will lead the SCs to value positively their interpersonal relationships with close relatives and friends, follow up with the Decalogue. The last messages will attempt to encourage the SC towards the use of Facebook as a means of obtaining family and social inclusion, as well as other satisfactions'.

Decalogue for a senior citizen's positive attitude towards Facebook:

1. I am capable of remaining as active as I was in previous periods of my life.
2. I am capable of being the architect of my own quality of life.
3. Social activity is beneficial and provides me with satisfaction.
4. Being in touch with my loved ones is an important part of my life.
5. With Facebook, I can avoid the social isolation and loneliness that would affect my general well-being (emotional, physical, and social).
6. Facebook is a valid alternative for me to remain as a socially active human being with my loved ones, especially with those living far away from me.
7. Facebook allows me to maintain interpersonal relationships with different meaning and intensity.
8. I am capable of learning how to use Facebook.
9. Facebook has functionalities that, when fully exploited, will provide me with very satisfying virtual social interactions.
10. I might find other ways to employ Facebook and the Internet, and they will bring me satisfaction.

Conceptual Bases of the Decalogue:

- For message No. 1: Psychosocial theories of aging related to elderly found in [17]: Modernity with a Dynamic Vision, Modernity with a Life Cycle Perspective, and Theory of Activity.
  The theory of Modernity—A dynamic Vision, has as a starting hypothesis, which establishes that "… as individuals age, they socially, psychologically, and biologically change, they change roles, and they add up knowledge, attitudes and experiences. As new cohorts begin to appear, they become older in different times and respond to unique historical experiences until they fade away". "… History shapes up individuals in a different way according to the year they were each born, exposing them to the influence of several events, and providing them with some determined combinations of resources so that they can develop their lives, and even giving them the possibility to have their own and unique way of interpreting reality".
  The theory of Modernity—The Life Cycle Perspective, introduces the idea that "elderly is another stage within the total life cycle process. In other words, the old age stage does not necessarily is meant to be seen as a disruption of time or entering a final stage, but it is rather a part of a process (and a process itself), where the individual continues to interact with society, just as he did during his prior life stages. Similar to these prior stages (childhood, youth, and adult life), the old age stage, in some way, has its own set of rules, roles, expectations and statuses, and society is there to establish a social agreement in regards of what age range corresponds to this cycle. The situations and the social position, which get to be experienced during the old age stage, are determined by the events, decisions, and behaviors of individuals during their prior life stages.
  Keeping into consideration that elderly—seen as another stage of the life cycle—which is conditioned by restrictions and privileges (like any other life stage); would not have, by definition, to be a stage of social exclusion." This is the most significant benefit here.
  The theory of Activity, whose main principle indicates that "normal aging corresponds to maintaining the individual's habitual activities and attitudes as long as possible and, therefore, satisfactory aging consists of remaining as the adult age. This approach acknowledges the loss of roles (due to retirement, becoming widowed, or the emancipation of their children, among other circumstances) to be the main source of a senior citizen's incapability of adapting to the system. This approach also began to discuss the loss, reassignment and meaning of the roles during the old age stage."
- For message No. 2: There is not a unique definition of QoL, but what has come to understanding is that there are several objective and subjective components, that may be classified into different types of life wellbeing, such as: emotional wellbeing, physical wellbeing, spiritual wellbeing, intellectual wellbeing, social wellbeing and material wellbeing. All these types of wellbeing are developed throughout diverse ambits of the social system, which a person has grown up into: his family, his school, his job, his community, the environment.

"The World Health Organisation defines Quality of Life (QL) as individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns" [18]. QoL is a 'multidimensional concept that encompasses objective and subjective components'. It also includes diverse ambits of life, reflects the objective cultural rules of wellbeing, and provides each ambit of life with a specific importance. All these aspects may be considered more significant to some individuals than to others [19].
"The QoL is not something that a person can have or not, it is needed to consider it as something in a scale; a person can have a low or high QoL. QoL can be evaluated and increased" [18].
"Any stimulus may modify the individual's construction of their quality of life" [20].

- For messages No. 3 and No. 6: "Social activity is beneficial in itself and results in greater satisfaction in life. Social interaction is important in the development of the concept of self in old age. All decreases in social interaction in old age are best explained by poor health or disability. If retirement or the limitations of age make participation impossible, people will find substitutes for the roles or activities that they have had to renounce" [21].
- For message No. 4: "The networks conformed by Family, friends and acquaintances, not only encourage senior citizens to keep their social identity, but they also provide with services, information and emotional/material support" [22].
  Within this matter, the concept of social/family inclusion fits perfectly, which is why it was conceived through an adaptation of what was stated by [23]: "Social acceptance of the elderly within their familiar environment and their community settings (subjective measurements), in order to get social interaction, relationships, and social networks (objective measurements)."
- For message No. 5: The definitions of social isolation, and loneliness found in [24].
  "Social isolation is the distancing of an individual, psychologically or physically, or both, from his or her network of desired or needed relationships with other persons. Isolation is the experience of diminished social connectedness is measured by the quality, type, frequency, and emotional satisfaction of social ties."
  "Loneliness: is absence of meaningful social connections that are an inherent need that all human beings have."
- For message No. 7: The starting point is the definition of interpersonal relationship that according to [25] is: "A reciprocal knowledge and commitment based on the interactions that give rise to specific forms of trust."
  The intensity of the reciprocal commitment existing in an interpersonal relationship in this study is referred to as the strength of the relation or tie-strength, for which there are important statements such as the following:
  "Strength of a tie is a quantifiable property that characterizes the link between two nodes [in a social network]" [26].

Supported by the approach exposed in [27], it is possible to affirm that the factors that have an effect onto the strength of a personal relation are classified as indicators and predictors.

"Indicators are actual components of tie-strength (closeness, duration and frequency, breadth of topics and mutual confiding)" [26].

The indicators will contribute with their weights to define a specific value for the tie-strength, according to the mathematical definition of weighted average for a non-empty data series: X = {X1, X2, X3, … Xn} to which matches with the weights W = {W1, W2, W3, … Wn}, following the Eq. (1):

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{x_i w_i}^{n} w_i} = \frac{x_1 w_1 + x_2 w_2 + \cdots + x_n w_n}{w_1 + w_2 + \cdots + w_n} \tag{1}$$

"Predictors are contextual contingencies (neighborhood, affiliation, similar socioeconomic status, workplace, and occupation prestige). Predictors are related to tie-strength but not components of it" [26].

- For message No. 8: "Learning capability lasts, during normal aging, up until 80 years old or even beyond. Learning how to employ and handle ICT systems is one of the skills that might be developed by stimulating cognitive processes… During digital literacy training, SCs tend to show an enthusiastic attitude while handling a computer and accessing the internet" [5].

  "It has not been proved that elders are incapable of carrying out intellectual activities. According to the American National Institute of Aging, a senior's brain activity is as high and efficient as a young person's brain. Old age keeps a high learning capability, which might be enhanced by wisdom and previous experiences" [12].
- For message No. 9: The full use of the features offered by Facebook for social interaction refers to the exploitation of all the characteristics that the social network platform puts at the service of the SCs, to help them establish satisfactory interpersonal relationships: Social Presence; Ubiquity; Transmission Methods: Unicast, Multicast and Broadcast; Multimedia communication Services; and Asynchronous/Synchronous Communication Modes. These terms are conceptualized as follows:

  Social Presence: [28] affirm that "The term social presence has no unique definition within the literature, but it is rather continuously re-defined. In one of the initial definitions, social presence is defined as "the degree to which people are perceived as 'real'. …we consider social presence as 'the perceived user experience of being together when communicating and interacting over distance, enabled through a system's capabilities that allow conveying a variety of non-verbal cues such as facial expressions, postures and gestures, thus offering a realistic animation of human behaviour". Dasgupta [2] indicates that: "Social presence is a critical attribute of a communication medium that can determine the way people interact and communicate. Further, people perceive some communication media as having a higher degree of social presence than other communication media."

Ubiquity: In [29], this term is interpreted for a Web Application as an "application that enforces the notion of a session, and suffers from the anytime/anywhere/anymedia syndrome (it must run "as is" on a variety of platforms)".
Usability: According to [30] "is a quality attribute that assesses how easy user interfaces are to use. It may be defined by five components: Learnability, Efficiency, Memorability, Satisfaction and Errors." A non-exhaustive analysis of the Facebook Interface, considering the five previous components and the definitions provided by [30], proves that the service is:

1. Simple for newbie users to operate and execute basic tasks of social interaction (Learnability);
2. A quick tool to execute desired tasks once the users get used to the social network interface (Efficiency);
3. Easy to reestablish control over the interface once the users get back to it after a period of not operating it (Memorability);
4. Pleasant to use (Satisfaction).
5. Open for users to make mistakes, due to lack of comprehension or attention to the privacy policies proposed by Facebook, and the type of adjustments provided for the users to control their privacy settings (Errors).
6. Along with Usability comes the Utility component, which involves functionality. It can be a certain fact that the Facebook interface offers services to enhance the QoL of the SC user, particularly those ones that lead to social and familiar inclusion. These terms, cited within the study, signify: Social acceptance of elders within their family circle and their community (subjective component) to accomplish a social interaction, relationships and social networking (objective components) [31].

Broadcasting Methods Unicast, Multicast and Broadcast: The unicast is a one-to-one method, which allows the data transmission to be performed between one only transmitter and one only receiver.
Multicast is a one-to-many method, which allows the data transmission to be performed among one transmitter and several receivers simultaneously within a specific group.
The broadcast transmission means data is transferred from a source to all the nodes of the network.
Multimedia communication Services: Refer to the services which simultaneously altogether offer diverse ways of expression, such as text, graphics, images—photos, animations, real-time or pre-filmed video and audio.
Asynchronous/Synchronous Communication Modes: Communication which does not/does need the interlocutors to be simultaneously logged into the system.

- For message No. 10: In addition to feeling socially active with loved ones and keeping current friendships (social/family inclusion), a SC will come to realize that Facebook interaction will also allow him to get the benefits displayed on Fig. 22.1.

22.4 Discussion

In order for a SC to interact successfully through Facebook, there are several requisites that should be met, such as:

1. A SC should feel motivated.
2. A SC should be in the adequate technological environment: budget to cover internet access expenses, technological devices to access social media services, reliable digital literacy or help, and a positive attitude (the proposed Decalogue is expected to have an influence on this particular attribute).
3. The SC should live within their own house. In contrast to people residing at a nursing house, seniors living at home may experience several advantages, such as: (1) availability of ICT devices set up within the SC's home will depend on the SC or their closest relatives. (2) Internet access at home allows a SC to have time flexibility to get online. (3) If the SC does not have enough computer skills, getting reliable help will be easier.
4. In addition, the location of the SC residence should guarantee the availability of broadband Internet access provided by ISPs, the existence of technology providers, and the presence of computer training centers.
5. Facebook, as a social networking service, should enhance certain usage features. According to senior citizens, they may experience several shortcomings alike [14, 32, 33] pointed out in addition to irrelevant information or excessive advertisement.

22.5 Conclusion

Facebook is a valid option for a healthy, independent SC to keep active social communication. Furthermore, it turns out to be an inclusive tool that might aid them in maintaining familiar and social inclusion.

The proposed Decalogue constitutes a compendium of previous research focused on psychosocial theories of aging, concepts and attributes related to SC, which were assembled through this study, in a logical manner in order to sensitize the SCs towards the change of attitude regarding the social interactions through Facebook.

References

1. R. Rios, Decalogue for a Senior Citizen's Positive Attitude towards Facebook, in *Proceedings of The World Congress on Engineering and Computer Science 2017, 25–27 October, 2017.* Lecture Notes in Engineering and Computer Science (San Francisco, USA, 2017), pp. 153–157
2. S. Dasgupta, Social Computing: ©, no. June (2010)

3. Age Uk, Technology and Older People Evidence Review, no. 1128267 (2011), https://www.ageuk.org.uk/Documents/EN-GB/For-professionals/Research/Evidence_Review_Technology.pdf?dtrk=true
4. K. Ala-mutka, D. Broster, R. Cachia, C. Centeno, C. Feijóo, A. Haché, S. Kluzer, S. Lindmark, W. Lusoli, G. Misuraca, C. Pascu, Y. Punie, J. a Valverde, A.K. Ala-mutka, D. Broster, R. Cachia, C. Centeno, C. Feijóo, A. Haché, S. Kluzer, S. Lindmark, W. Lusoli, G. Misuraca, C. Pascu, Y. Punie, J. a Valverde, E. Y. Punie, *The Impact of Social Computing on the EU Information Society and Economy* (2009), http://ftp.jrc.es/EURdoc/JRC54327.pdf
5. G.A. González, L.G. Gómez, A.J. Mata, Information and communication technologies (ICT) as an alternative for the stimulation of cognitive processes in old age—Las Tecnologías de la Información y Comunicación (TIC) como alternativa para la estimulación de los procesos cognitivos en la vejez, Scientif. J. Netw. Latin Am. Caribbean, Spain and Portugal, Scientific Information System (2012), http://www.redalyc.org/pdf/2831/283121840008.pdf
6. N. Haris, R.A. Majid, N. Abdullah, R. Osman, The role of social media in supporting elderly quality daily life, in *2014 3rd International Conference on User Science and Engineering (i-USEr)* (Shah Alam, 2014), pp. 253–257
7. D. Hartnett, Elizabeth, M., Shailey, P. Jane, P. Marian, E. Shirley, M. Christopher Paul, D. Kathryn, M. Brendan, H. Tania, Roberts, Older people and online social interactions : an empirical investigation, in *UKAIS International Conference on Information Systems*, pp. 18–20 (2013). http://oro.open.ac.uk/36591/1/paper12.pdf
8. A. Hendry, L. Rodriguez Mañas, A. Cano, R. Bernabei, M. Lllario, W. Molloy, Building European commitment to prevent and tackle frailty. Luxembourg: Publications Office of the European Union (2015), https://ec.europa.eu/research/innovation-union/pdf/active-healthy-ageing/a3_decalogue_2015.pdf
9. A. Hope, T. Schwaba, A.M. Piper, Understanding digital and material social communications for older adults, in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI'14*, pp. 3903–3912 (2014)
10. Janelle W. Myhre, Matthias R. Mehl, Elizabeth L. Elizabeth L, Cognitive benefits of online social networking for healthy older adults. J. Gerontol. Series B **72**(5), 752–760 (2017). https://doi.org/10.1093/geronb/gbw025
11. C. Moser, V. Fuchsberger, K. Neureiter, W. Sellner, M. Tscheligi, Elderly's social presence supported by ICTs: investigating user requirements for social presence, in *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pp. 738–741 (2011)
12. S.A. Prado, Impact of ICT On Elderly In Asturias: Improvement of Self-Concept and Satisfaction—Impacto De Las Tic En Las Personas Mayores En Asturias : Mejora Del Autoconcepto Y De La Satisfacción, EDUTEC: Electronic magazine of educational technology. ISSN 1135-9250, pp. 1–13 (2013), http://edutec.rediris.es/Revelec2/Revelec44/pdf/Edutec-e_n44-Agudo-Fombona.pdf
13. D. Quinn, L. Chen, M.D. Mulvenna, R. Bond, Exploring the relationship between online social network site usage and the impact on quality of life for older and younger users: An interaction analysis. J. Med. Internet Res. **18**(9), 1–16 (2016). https://doi.org/10.2196/jmir.5377
14. D. Vilte, V. Saldaño, A. Martín, G. Gaetán, Evaluation of the use of social networks in the third age—Evaluación del Uso de Redes Sociales en la Tercera Edad (2013). http://conaiisi.unsl.edu.ar/2013/142-446-1-DR.pdf
15. F. Zhang, D. Kaufman, Social and emotional impacts of internet use on older adults. Eur. Sci. J. ESJ **11**(17), 1–15 (2015). http://www.eujournal.org/index.php/esj/article/view/5789
16. I.T.U. ITU, *Manual for Measuring ICT Access and Use by Households and Individuals* (2014). https://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-ITCMEAS-2014-PDF-E.pdf
17. P. Aranibar, Conceptual approach to the situation of the elderly in Latin America. Population and development series—Acercamiento conceptual a la situación del adulto mayor en América Latina (2001)

18. R.C. Esteban, R.F. Builds, R.E. Curto, Evaluation toolkit on seniors education to improve their quality of life, no. 2, pp. 1–21 (2012), http://edusenior.uji.es/data/outcomes/wp2/QEduSen-WP2-2.1.a.RequirementAnalysis-Quality of Life-v.02.pdf
19. L.M. Rojas, F. Campos, Guide quality of life in old age—tools to live longer and better—Guía Calidad de Vida en la Vejez - Herramientas para vivir más y mejor (2011). http://adultomayor.uc.cl/docs/guia_calidad_de_vida.pdf
20. Z. Gabriel, A. Bowling, Quality of life from the perspectives of older people. Ageing Soc. **24**(5), 675–691 (2004). https://doi.org/10.1017/s0144686x03001582
21. M. Oddone, Theoretical Background of Active Aging—Antecedentes teóricos del Envejecimiento Activo, Inf. Envejec. en red, vol. 4, no. 2340–566X, pp. 1–9 (2013). http://envejecimiento.csic.es/documentos/documentos/oddone-antecedentes.pdf
22. A. Concha, L. Olivares, L. Sepúlveda, Social Networks in the Third Age—Redes sociales en la tercera edad (2000). http://www.redadultosmayores.com.ar/buscador/files/DESAR017.pdf
23. S.C. Simplican, G. Leader, J. Kosciulek, M. Leahy, Defining social inclusion of people with intellectual and developmental disabilities: an ecological model of social networks and community participation. Res. Developm. Disabil. **38**, 18–29 (2015). https://doi.org/10.1016/j.ridd.2014.10.008
24. K. Elder, J. Retrum, *Framework for Isolation in Adults Over 50*, p. 48 (2012). http://www.aarp.org/content/dam/aarp/aarp_foundation/2012_PDFs/AARP-Foundation-Isolation-Framework-Report.pdf
25. M. Grossetti, What is a social relationship? A set of dyadic mediations—¿Qué es una relacion social ? Un conjunto de mediaciones diádicas, REDES: Hispanic Magazine for the analysis of social networks Vol.6, # 2, June 2009. http://revista-redes.rediris.es/pdf-vol16/vol16_2e.pdf
26. A. Petroczi, F. Bazsó, T. Nepusz, Measuring tie-strength in virtual social networks. Connections **27**(2), 39–52 (2007). http://eprints.kingston.ac.uk/2396/
27. Peter V. Marsden, Karen E. Campbell, Measuring tie strength. Soc. Forces **63**(2), 482–501 (1984). https://doi.org/10.1093/sf/63.2.482
28. A.P. Achilleos, C. Mettouris, G. A. Papadopoulos, K. Neureiter, C. Rappold, C. Moser, M. Tscheligi, L. Vajda, A. Toth, P. Hanak, O. Jimenez, R. Smit, The connected vitality system: Enhancing social presence for older adults, in *Proceedings of 12th International Conference Telecommunications ConTEL 2013*, pp. 199–206 (2013)
29. C. Finkelstein, A., Savigni, A., Kappel, G., Retschitzegger, W., Kimmerstorfer, E., Schwinger, W., Hofer, Th., Pröll, B, Feichtner, Ubiquitous web application development—a framework for understanding, in *6th World Multiconference System Cybern. Informatics*, no. 1, pp. 431–438 (2002)
30. J. Nielsen, Usability 101 : Introduction to Usability Why Usability is Important How to Improve Usability, Focus (Madison), pp. 1–3 (2003)
31. R. Rios, Understanding the potential of the technological context of the Ecuadorian elderly person living in their home to take advantage of social interaction through the Internet—Comprensión de la potencialidad del contexto tecnológico del adulto mayor ecuatoriano que vive en su casa para el aprovechamiento de la interacción social a través de Internet," Actas CIAIQ2017 - Investig. Cual. en Ing. y Tecnol.—ISBN978-972-8914-78-3, vol. 4, pp. 65–74 (2017)
32. F.J. Díaz, I. Harari, Evaluando la usabilidad de las redes sociales por parte de adultos mayores [en línea]. En: Interaction South America (ISA 14): 6ta. Conferencia Lationamericana de Diseño de Interacción; 2014 nov 19-22; Buenos Aires: Interaction Design Association; Asociación de Profesionales en Experiencia de Usuario; Internet Society; Universidad Católica Argentina. Disponible en: http://bibliotecadigital.uca.edu.ar/repositorio/ponencias/evaluando-usabilidad-redes-sociales.pdf
33. G. Gomes, C. Duarte, J. Coelho, E. Matos, Designing a facebook interface for senior users. Scientif. World J. (2014). https://doi.org/10.1155/2014/741567

Ramiro Augusto Rios Paredes Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas ESPE, Av. General Rumiñahui s/n, Sangolquí, Ecuador, 171-5-231B. He is currently a doctoral candidate at the Doctoral Program in Ciencias Informáticas, Universidad Nacional de La Plata, Calle 50 y 120, La Plata, Argentina.

Chapter 23
Archive Browsing System for the Roads with Extremely Delayed Recovery After the 2011 Tohoku Earthquake

Noriaki Endo, Jieling Wu, Bingzhen He and Satoru Sugita

Abstract In our previous study, we identified the municipalities and each road within these municipalities for which road recovery was extremely delayed after the 2011 Tohoku Earthquake. In this study, we built an archive system to monitor these roads following the 2011 Tohoku Earthquake to identify and confirm the vulnerabilities of roads and raise their resilience. This system allows us to browse video images of vulnerable roads that are narrow, steep-walled, and located in mountainous regions. To date, we could only identify the type of problematic roads, but we had not developed methods for increasing their resilience, which can be done using the proposed system. The next step is to confirm the vulnerability of road components in detail.

Keywords 2011 Tohoku Earthquake · Disaster archive system · Disaster risk reduction · Probe-car data · Resilience to disasters · Telematics data Vehicle-tracking map

N. Endo (✉) · J. Wu · B. He
Iwate University, Morioka, Iwate, Japan
e-mail: n_endo@gw2k.hss.iwate-u.ac.jp; endo-100@chopin.jp

J. Wu
e-mail: j_wu@gw2k.hss.iwate-u.ac.jp

B. He
e-mail: b_he@gw2k.hss.iwate-u.ac.jp

S. Sugita
Chubu University, Kasugai, Aichi, Japan
e-mail: satoru@isc.chubu.ac.jp

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_23

23.1 Introduction

23.1.1 Disaster Archive System of the 2011 Tohoku Earthquake

Immediately after disasters such as the 2011 Tohoku Earthquake [1] (Figs. 23.1 and 23.2), the people present in the disaster area are forced to focus on not only their survival but also the survival of their neighbors. After the situation settles down to some extent, people begin to consider sharing their experiences from the disaster with the next generation and utilizing these experiences for future disaster prevention. These disaster experiences could be of great interest to people for disaster prevention, not only in afflicted areas but also throughout Japan and the world. Therefore, we created an experimental archival system to preserve digital images from the disaster as archival material for the next generation [2].

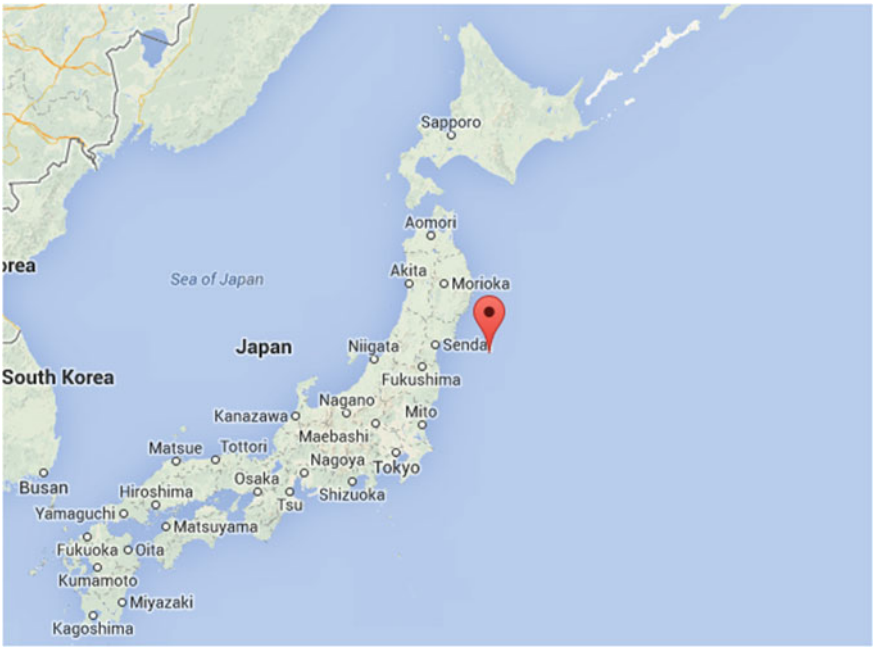


Fig. 23.1 The epicenter of the earthquake that occurred on March 11, 2011 in Tohoku (https://www.google.co.jp/maps/)

Fig. 23.2 The epicenter of the earthquake that occurred on March 11, 2011 in Tohoku (magnified, https://www.google.co.jp/maps/)

23.1.2 Road Usage Recovery Following the 2011 Tohoku Earthquake

In our previous work [3–5], we evaluated regional differences for road recovery in the Iwate Prefecture after the 2011 Tohoku Earthquake. For these studies, we used vehicle-tracking maps constructed from probe-car data that were made available on the Internet by Toyota Motor Corporation, Japan. In these studies [3–5], we also determined the municipalities and each road within these municipalities with extremely delayed road recovery after the earthquake.

23.1.3 Purpose of This Study

In this study, by combining the scopes of our studies [2–5], we built an experimental archival system for browsing roads with an extremely delayed recovery after the 2011 Tohoku Earthquake. The purpose of this system is to help identify the vulnerable road components to raise overall road resilience. This paper is the revised and extended version of the previous paper [6] in the WCECS 2017 proceedings.

23.2 Methodology

23.2.1 Research Area

All areas of the Iwate Prefecture afflicted by the 2011 Tohoku Earthquake (Fig. 23.3).

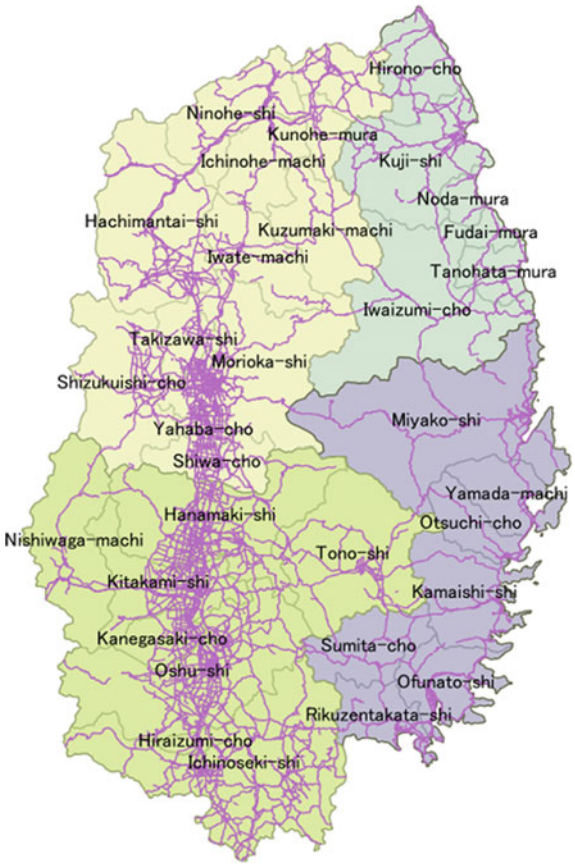


Fig. 23.3 The vehicle-tracking map of Iwate Prefecture is shown by the violet lines. The perimeter of a city is shown by a gray polygon. Iwate Prefecture has been divided into four areas: the Northern Costal, Southern Coastal, Northern Inland, and Southern Inland areas

23.2.2 Research Materials

In this study, we used video data for the roads in the Iwate Prefecture with extremely delayed recovery after the 2011 Tohoku Earthquake. The video images were captured in the afflicted area of the Iwate Prefecture, following post-disaster road recovery. We used video data available on the Internet through services such as YouTube with the consent of their creators.

23.2.3 System

23.2.3.1 The Server for Providing the Content

Hardware: a PC server assembled by the author.

Software: The operating system(OS) was Linux. Apache HTTP server was used as the HTTP server software.

23.2.3.2 The Client for Creating the Content

Hardware: a PC client.

Software: Notepad was used for writing and editing the code of the system. The system was deployed on a web browser such as Firefox or Internet Explorer.

23.2.3.3 The Role of the Local Server

We used the server for serving HTML, ASX files, and video data to the clients. The ASX file functions will be mentioned in the following section.

23.2.3.4 The Role of the Web Service

We used the Yahoo! JavaScriptMap API as the platform (the developing environment which was provided by Yahoo Japan) to draw background maps. Accordingly, the system was programmed in JavaScript. Using this web service, we could build a sophisticated system rather easily.

23.3 Construction of the Archive

The method used in this study was similar to the one used in our previous study [2].

By linking video data and high-resolution satellite images, captured in the target earthquake disaster area, with the digital map of the area, we built an archival system for browsing disaster images clearly and easily.

23.3.1 Specifying the Exact Route of the Video

Images uploaded on the Internet do not often have exact geospatial attributes and exact route attributes. Therefore, initially, we specified the start point, midpoints,

and end point of the video file. We watched the video and map of the location in the video simultaneously. The routes in the video were specified based on the signs of structures such as buildings and roads. We recorded the route information as a geospatial attribute.

23.3.2 Making Links from Each Node to the Video Scene

We measured the time from the initial node to each node and made ASX files to play the video between two nodes (beginning node to end node) by clicking on the beginning node. We used the Yahoo! JavaScript Map API to link the map images and ASX files. This has been illustrated in Figs. 23.4 and 23.5.

23.3.3 The Role of ASX Files

An ASX file is a type of Windows meta file that makes it possible link web pages and Windows Media Audio (WMA) servers or web server content. Using an ASX file, we were able to control the playback of Windows Media Video (WMV) files. We could specify the beginning of the video by using the StartTime tag. Additionally, we could specify the playback duration time of the video using the Duration tag. This made it possible to set multiple playback patterns for one video file. Therefore, in this study, we were able to consider a video without any geospatial attributes as one with geospatial attributes.

23.4 Results: Archive System Operation

Figure 23.4 depicts the example map containing the routes of the video data taken around Iwate Prefectural Route 7, 158, and 160. System users can play the video between two nodes (the beginning and end nodes) by clicking on the beginning node on the map (Fig. 23.5). When the cursor is pointed at each node, the number of the nodes appears at the bottom of the map.

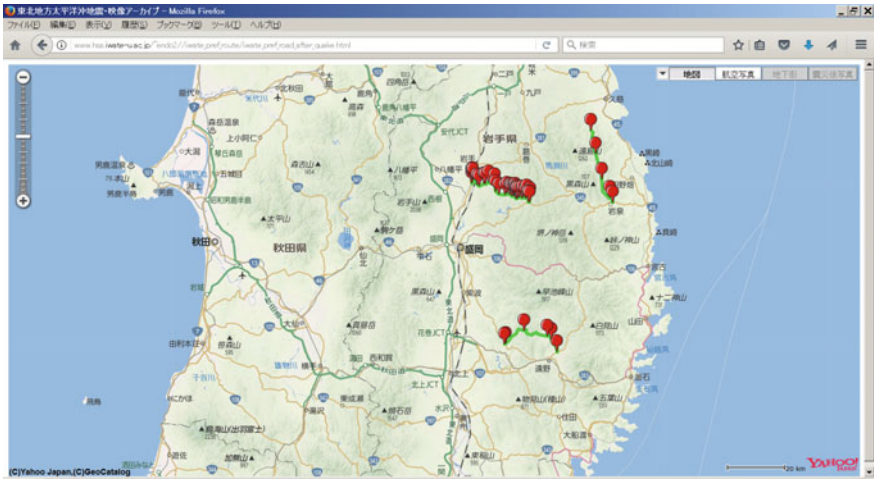


Fig. 23.4 An example map for browsing the roads with an extremely delayed recovery following the 2011 Tohoku Earthquake

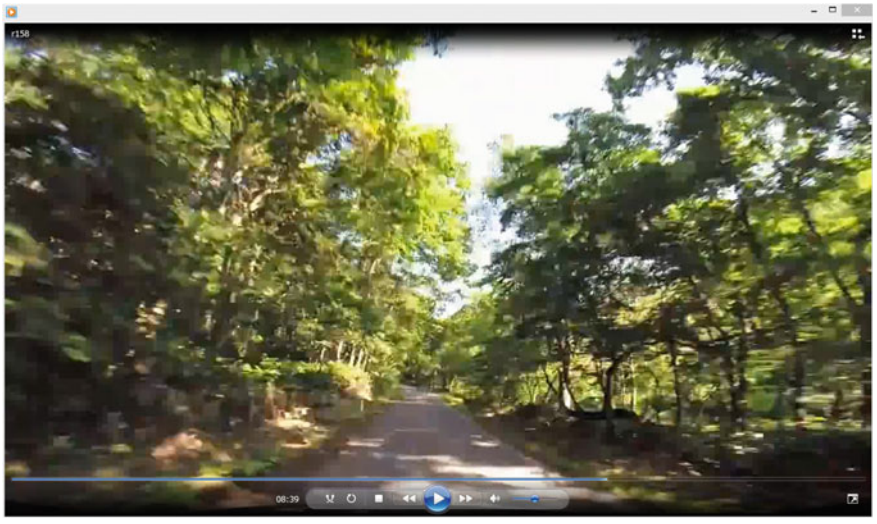


Fig. 23.5 A video image of the road with an extremely delayed recovery following the 2011 Tohoku Earthquake (Iwate Prefectural Road 158)

23.5 Discussion

23.5.1 Upgrading Public Internet Images to Geospatial Data

Although the images provided by the media, such as TV stations, were useful with regard to the 2011 Tohoku Earthquake, the images uploaded on the Internet proved to be more informative. However, images uploaded on the Internet usually have no exact geospatial attributes. In particular, the route information is completely absent. This means that they do not have much value as geospatial data.

Accordingly, in this study, we added routing data to images uploaded on the Internet to upgrade them for their use as geospatial data. Finally, we accumulated the data to build an earthquake disaster archive of the 2011 Tohoku Earthquake. We particularly looked for images from areas around main roads because we considered that most people would be interested in information about the roads. These images can be used for the disaster education of both children and adults and hopefully in preventing a future earthquake disaster.

23.5.2 About Research Materials

We received some responses for our previous study which was presented in the WCECS 2017 conference [6]. The responses were about the accuracy and the quality of video images which were uploaded to Internet (such as images on YouTube service).

As for the accuracy, we were usually able to confirm the road number by road signs. On the other hand, we quite agree to the comments of the participants for the video quality. It may be sometimes better for us to use video images which were taken for ourselves instead of the images uploaded to the Internet.

23.5.3 Problem of Copyright

In general, we can not edit and reuse video movies uploaded to the Internet without authors' permission. If an archive system has many data, it may be very complicated to get all authors' permission. So, if an author permit the edit and the reuse of his video works, the Creative Commons License may be recommended. This is one of the best solutions to avoid complicated processes when a person other than authors would like to build an archive.

23.5.4 Problem of Map Update

Using the Web service, we could build the Tsunami disaster archive rather easily and economically. Though the Web service system contributed very much for our system, the map update of the Web service will be contrarily rather unfavorable for the 3.11 disaster map. That's because we need the map layer when the 2011 Tohoku Earthquake occurred for our system. If the Web service would keep containing the map layer above described for a long time, we would appreciate it very much.

23.5.5 Roads Vulnerable to the Next Disaster

Using this system, we could browse the video images for vulnerable (problematic) roads. These roads were narrow, steep-walled, and located in mountainous regions. They were just what we had expected from the results of our previous study [5]. Until now, we were only able to identify the type of problematic roads but had not developed techniques to increase their resilience. The next step is to identify the vulnerable points on the roads in detail. To realize this, we have to collect the documents containing the recovery history of the roads and evaluate them in future studies.

Acknowledgements The authors are grateful to Ms.Junko Akita (Lecturer of Iwate University) for her contribution to our studies.

References

1. N. Mori, T. Takahashi, T. Yasuda, H. Yanagisawa, Survey of 2011 Tohoku Earthquake tsunami inundation and run-up. Geophys. Res. Lett. **38**(7) (2011). https://doi.org/10.1029/2011GL049210
2. N. Endo, A. Takehara, Creating a tsunami disaster archive of the Great Northeastern Japan earthquake using images uploaded to the Internet, in *Proceedings of the 8th International Symposium on Digital Earth, ISDE 8*, 26–29 August 2013, Kuching, Malaysia (2013). https://doi.org/10.1088/1755-1315/18/1/012079
3. H. Komori, N. Endo, Road distance traveled by vehicles following the 2011 Tohoku Earthquake, calculated by G-BOOK telematics data, in *Proceedings of the 9th International Conference on Signal-Image Technology and Internet-Based Systems, SITIS 9*, 2–5 December 2013, Kyoto, Japan (2013), pp. 870–874. https://doi.org/10.1109/SITIS.2013.142
4. N. Endo, H. Komori, Analysis of vehicle tracking maps in Iwate Prefecture following the 2011 Tohoku Earthquake, in *Proceedings of the International Multiconference of Engineers and Computer Scientists 2015, IMECS 2015*, 18–20 March 2015, Hong Kong. Lecture Notes in Engineering and Computer Science (2015), pp. 124–128, http://www.iaeng.org/publication/IMECS2015/IMECS2015_pp124-128.pdf
5. N. Endo, H. Komori, Regional difference on road recovery in Iwate Prefecture following the 2011 Tohoku Earthquake, in *Proceedings of the World Congress on Engineering and Computer Science 2015, WCECS 2015*, 21–23 October 2015, Berkeley, CA, USA. Lecture Notes in

Engineering and Computer Science (2015), pp. 162–167, http://www.iaeng.org/publication/WCECS2015/WCECS2015_pp162-167.pdf

6. N. Endo, J. Wu, B. He, S. Sugita, Archive system to browse the roads whose recovery was extremely delayed after the 2011 Tohoku Earthquake, in *Proceedings of the World Congress on Engineering and Computer Science 2017, WCECS 2017*, 25–27 October 2017, Berkeley, CA, USA. Lecture Notes in Engineering and Computer Science (2017), pp. 149–152, http://www.iaeng.org/publication/WCECS2017/WCECS2017_pp149-152.pdf

Chapter 24
Exponentiated Generalized Exponential Distribution: Ordinary Differential Equations

Hilary I. Okagbue, Pelumi E. Oguntunde, Paulinus O. Ugwoke, Abiodun A. Opanuga and Ezinne C. Erondu

Abstract In this chapter, homogenous ordinary differential equations (ODE) of different orders were obtained for the probability density function, quantile function, survival function inverse survival function, hazard function and reversed hazard functions of exponentiated generalized exponential distribution. This is possible since the aforementioned probability functions are differentiable. Differentiation and modified product rule were used to obtain the required ordinary differential equations, whose solutions are the respective probability functions. The different conditions necessary for the existence of the ODE were obtained and it is almost in consistent with the support that defined the various probability functions considered. The parameters that defined each distribution greatly affect the nature of the ODEs obtained. This method provides new ways of classifying and approximating other probability distributions apart from exponentiated generalized exponential distribution considered in this chapter. In addition, the result of the quantile function can be compared with quantile approximation using the quantile mechanics.

Keywords Differential calculus · Exponentiated distributions · Hazard function Inverse survival function · Quantile function · Quantile mechanics · Reversed hazard function · Survival function

H. I. Okagbue (✉) · P. E. Oguntunde · A. A. Opanuga · E. C. Erondu
Department of Mathematics, Covenant University, Ota, Nigeria
e-mail: hilary.okagbue@covenantuniversity.edu.ng

P. E. Oguntunde
e-mail: pelumi.oguntunde@covenantuniversity.edu.ng

A. A. Opanuga
e-mail: abiodun.opanuga@covenantuniversity.edu.ng

E. C. Erondu
e-mail: ezinneerondu@yahoo.com

P. O. Ugwoke
Department of Computer Science, University of Nigeria, Nsukka, Nsukka, Nigeria
e-mail: okeyugwoke@yahoo.com

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_24

24.1 Introduction

Calculus is a very key tool in the determination of mode of a given probability distribution and in estimation of parameters of probability distributions, amongst other uses. The method of maximum likelihood is an example.

Differential equations often arise from the understanding and modeling of real life problems or some observed physical phenomena. Approximations of probability functions are one of the major areas of application of calculus and ordinary differential equations in mathematical statistics. The approximations are helpful in the recovery of the probability functions of complex distributions [1–4] especially in quantile approximations.

Apart from mode estimation, parameter estimation and approximation, probability density function (PDF) of probability distributions can be expressed as ODE whose solution is the PDF. Some of which are available. They include: beta distribution [5], Lomax distribution [6], beta prime distribution [7], Laplace distribution [8] and raised cosine distribution [9].

The aim of this research is to develop homogenous ordinary differential equations for the probability density function (PDF), Quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of exponentiated generalized exponential distribution. This will also help to provide the answers as to whether there are discrepancies between the support of the distribution and the necessary conditions for the existence of the ODEs. Similar results for other distributions have been proposed, see [10–22] for details.

The distribution was proposed by Oguntunde et al. [23] as a three parameter model that can be used as one of the generalizations of the exponential distribution. The proposed model has also the generalized exponential and exponentiated exponential distribution as its sub-models. The distribution belongs to the exponentiated class of distributions which has seen a lot of research activities and modifications. Details on the general class of exponentiated distributions can be seen in [24–28].

In particular, some exponentiated distributions are available in scientific literature such as: exponentiated Gumbel type-2 distribution [29], exponentiated Weibull distribution [30–32], exponentiated generalized inverted exponential distribution [33], exponentiated generalized inverse Gaussian distribution [34], exponentiated generalized inverse Weibull distribution [35, 36], gamma-exponentiated exponential distribution [37], exponentiated Gompertz distribution [38, 39], beta Exponentiated Mukherjii-Islam Distribution [40], transmuted exponentiated Pareto-i distribution [41], gamma exponentiated exponential–Weibull distribution [42], exponentiated gamma distribution [43], exponentiated Gumbel distribution [44], exponentiated uniform distribution [45] and beta exponentiated Weibull distribution [46, 47].

Others are: exponentiated log-logistic distribution [48], McDonald exponentiated gamma distribution [49], exponentiated Generalized Weibull Distribution [50], beta exponentiated gamma distribution [51], exponentiated gamma distribution [52], exponentiated Pareto distribution [53], exponentiated Kumaraswamy distribution [54], exponentiated modified Weibull extension distribution [55], exponen-

tiated Weibull-Pareto distribution [56], exponentiated lognormal distribution [57], exponentiated Perks distribution [58] and Kumaraswamy-transmuted exponentiated modified Weibull distribution [59]. Also available are: exponentiated power Lindley–Poisson distribution [60], exponentiated Chen distribution [61], exponentiated reduced Kies distribution [62], exponentiated inverse Weibull geometric distribution [63], exponentiated geometric distribution [64, 65], exponentiated Weibull geometric distribution [66], exponentiated transmuted Weibull geometric distribution [67], exponentiated half logistic distribution [68], transmuted exponentiated Gumbel distribution [69], exponentiated Kumaraswamy-power function distribution [70], exponentiated-log-logistic geometric distribution [71], bivariate exponentaited generalized Weibull-Gompertz distribution [72] and so on.

24.2 Probability Density Function

The probability density function of the exponentiated generalized exponential distribution is given as;

$$f(x) = \alpha\beta e^{-\alpha\lambda x}[(1 - e^{-\lambda x})^{\alpha}]^{\beta-1} \tag{24.1}$$

To obtain the first order ordinary differential equation for the probability density function of the exponentiated generalized exponential distribution, differentiate equation (24.1), to obtain;

$$f'(x) = \left\{-\alpha\lambda + \frac{\alpha(\beta-1)\lambda e^{-\lambda x}}{(1 - e^{-\lambda x})}\right\} f(x) \tag{24.2}$$

The condition necessary for the existence of the equation is $x, \alpha, \beta, \lambda > 0$. Differentiate equation (24.2) to obtain;

$$\begin{aligned} f''(x) = &\left\{-\alpha\lambda + \frac{\alpha(\beta-1)\lambda e^{-\lambda x}}{(1 - e^{-\lambda x})}\right\} f'(x) \\ &- \left\{\frac{\alpha(\beta-1)\lambda^2(e^{-\lambda x})^2}{(1 - e^{-\lambda x})^2} + \frac{\alpha(\beta-1)\lambda^2 e^{-\lambda x}}{(1 - e^{-\lambda x})}\right\} f(x) \end{aligned} \tag{24.3}$$

The condition necessary for the existence of the equation is $x, \alpha, \beta, \lambda > 0$. Simplify Eq. (24.3) using Eq. (24.2) to obtain;

$$f''(x) = \frac{f'^2(x)}{f(x)} - \left\{\frac{1}{\alpha(\beta-1)}\left(\frac{f'(x)}{f(x)} + \alpha\lambda\right)^2 + \lambda\left(\frac{f'(x)}{f(x)} + \alpha\lambda\right)\right\} f(x) \tag{24.4}$$

The condition necessary for the existence of the equation is $x, \alpha, \lambda > 0, \beta > 1$.

The ordinary differential equations can be obtained for the particular values of the parameters.

24.3 Quantile Function

The Quantile function of the exponentiated generalized exponential distribution is given as;

$$Q(p) = \frac{1}{\alpha\lambda} \ln\left(\frac{1}{1 - p^{\frac{1}{\beta}}}\right) \tag{24.5}$$

To obtain the first order ordinary differential equation for the Quantile function of the exponentiated generalized exponential distribution, differentiate equation (24.5), to obtain;

$$Q'(p) = \left(\frac{p^{\frac{1}{\beta}-1}}{\alpha\beta\lambda(1 - p^{\frac{1}{\beta}})}\right) \tag{24.6}$$

The condition necessary for the existence of the equation is $\alpha, \beta, \lambda > 0, 0 < p < 1$.

$$\alpha\beta\lambda(1 - p^{\frac{1}{\beta}})Q'(p) - p^{\frac{1}{\beta}-1} = 0 \tag{24.7}$$

The ordinary differential equations can be obtained for given values of the parameters. Some cases are considered in shown in Table 24.1.

Table 24.1 Classes of differential equations obtained for the quantile function of exponentiated generalized exponential distribution for different parameters

α	β	λ	Ordinary differential equation
1	1	1	$(1-p)Q'(p) - 1 = 0$
1	1	2	$2(1-p)Q'(p) - 1 = 0$
1	2	1	$2(1-p)Q'(p) - 1 = 0$
1	2	2	$4(1-p)Q'(p) - 1 = 0$
2	1	1	$2\left(\sqrt{p}-p\right)Q'(p) - 1 = 0$
2	1	2	$4\left(\sqrt{p}-p\right)Q'(p) - 1 = 0$
2	2	1	$4\left(\sqrt{p}-p\right)Q'(p) - 1 = 0$
2	2	2	$8\left(\sqrt{p}-p\right)Q'(p) - 1 = 0$

24.4 Survival Function

The survival function of the exponentiated generalized exponential distribution is given as;

$$S(t) = 1 - [1 - e^{-\alpha\lambda x}]^{\beta} \tag{24.8}$$

To obtain the first order ordinary differential equation for the survival function of the exponentiated generalized exponential distribution, differentiate equation (24.8), to obtain;

$$S'(t) = -\alpha\beta\lambda e^{-\alpha\lambda x}[1 - e^{-\alpha\lambda x}]^{\beta-1} \tag{24.9}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \lambda > 0$.
Simplify Eq. (24.9) using Eq. (24.8) to obtain;

$$S'(t) + \alpha\beta\lambda(1 - (1 - S(t))^{\frac{1}{\beta}})(1 - S(t))^{1-\frac{1}{\beta}} = 0 \tag{24.10}$$

with initial condition;

$$S(1) = 1 - [1 - e^{-\alpha\lambda}]^{\beta} \tag{24.11}$$

The ordinary differential equations can be obtained for the different values of the parameters.
When $\beta = 1$, Eqs. (24.10) and (24.11) become;

$$S'(t) + \alpha\lambda S(t) = 0 \tag{24.12}$$

$$S(1) = e^{-\alpha\lambda} \tag{24.13}$$

When $\beta = 2$, Eqs. (24.10) and (24.12) become;

$$(\sqrt{1 - S(t)})S'(t) + 2\alpha\lambda(1 - \sqrt{1 - S(t)}) = 0 \tag{24.14}$$

$$S(1) = 1 - [1 - e^{-\alpha\lambda}]^2 \tag{24.15}$$

24.5 Inverse Survival Function

The inverse survival function of the exponentiated generalized exponential distribution is given as;

$$Q(p) = \frac{1}{\alpha\lambda} \ln\left(\frac{1}{1 - (1 - p)^{\frac{1}{\beta}}}\right) \tag{24.16}$$

Table 24.2 Classes of differential equations obtained for the inverse survival function of exponentiated generalized exponential distribution for different parameters

α	β	λ	Ordinary differential equation
1	1	1	$pQ'(p)+1=0$
1	1	2	$2pQ'(p)+1=0$
1	2	1	$2pQ'(p)+1=0$
1	2	2	$4pQ'(p)+1=0$
2	1	1	$2\sqrt{p}\left(1-\sqrt{1-p}\right)Q'(p)+1=0$
2	1	2	$4\sqrt{p}\left(1-\sqrt{1-p}\right)Q'(p)+1=0$
2	2	1	$4\sqrt{p}\left(1-\sqrt{1-p}\right)Q'(p)+1=0$
2	2	2	$8\sqrt{p}\left(1-\sqrt{1-p}\right)Q'(p)+1=0$

To obtain the first order ordinary differential equation for the inverse survival function of the exponentiated generalized exponential distribution, differentiate equation (24.16), to obtain;

$$Q'(p)=-\frac{p^{\frac{1}{\beta}-1}}{\alpha\beta\lambda(1-(1-p)^{\frac{1}{\beta}})} \tag{24.17}$$

The condition necessary for the existence of the equation is; $\alpha,\beta,\lambda>0, 0<p<1$.

Simplify Eq. (24.17) to obtain;

$$\alpha\beta\lambda(1-(1-p)^{\frac{1}{\beta}})Q'(p)+p^{\frac{1}{\beta}-1}=0 \tag{24.18}$$

The ordinary differential equations can be obtained for given values of the parameters. Some cases are considered and shown in Table 24.2.

24.6 Hazard Function

The hazard function of the exponentiated generalized exponential distribution is given as;

$$h(t)=\frac{\alpha\beta e^{-\alpha\lambda t}[(1-e^{-\lambda t})^{\alpha}]^{\beta-1}}{1-[1-e^{-\alpha\lambda t}]^{\beta}} \tag{24.19}$$

To obtain the first order ordinary differential equation for the hazard function of the exponentiated generalized exponential distribution, differentiate equation (24.19), to obtain;

$$h'(t) = -\frac{\alpha\lambda e^{-\alpha\lambda t}}{e^{-\alpha\lambda t}}h(t) + \frac{\alpha(\beta-1)\lambda e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}[(1-e^{-\lambda t})^{\alpha}]^{\beta-2}}{[(1-e^{-\lambda t})^{\alpha}]^{\beta-1}}h(t)$$
$$+\frac{\alpha\beta\lambda e^{-\alpha\lambda t}(1-e^{-\alpha\lambda t})^{\beta-1}[1-(1-e^{-\alpha\lambda t})^{\beta}]^{-2}}{[1-(1-e^{-\alpha\lambda t})^{\beta}]^{-1}}h(t) \tag{24.20}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \lambda > 0$.
Simplify to obtain;

$$h'(t) = \left\{-\alpha\lambda + \frac{\alpha(\beta-1)\lambda e^{-\lambda t}}{(1-e^{-\lambda t})} + h(t)\right\}h(t) \tag{24.21}$$

Differentiate equation (24.21) and simplify to obtain;

$$h''(t) = \frac{h'^2(t)}{h(t)} - \left\{\begin{array}{l}\frac{1}{\alpha(\beta-1)}\left(\frac{h'(t)}{h(t)} + \alpha\lambda - h(t)\right)^2 \\ +\lambda\left(\frac{h'(t)}{h(t)} + \alpha\lambda - h(t)\right) - h(t)\end{array}\right\}h(t) \tag{24.22}$$

The condition necessary for the existence of the equation is $t, \alpha, \lambda > 0, \beta > 1$.
The ordinary differential equations can be obtained for the particular values of the parameters.

24.7 Reversed Hazard Function

The reversed hazard function of the exponentiated generalized exponential distribution is given as;

$$j(t) = \frac{\alpha\beta\lambda e^{-\alpha\lambda t}}{\left[1-e^{-\lambda t}\right]^{\alpha}} \tag{24.23}$$

To obtain the first order ordinary differential equation for the reversed hazard function of the exponentiated generalized exponential distribution, differentiate equation (24.23), to obtain;

$$j'(t) = -\left\{\alpha\lambda + \frac{\alpha\lambda e^{-\lambda t}}{1-e^{-\lambda t}}\right\}j(t) \tag{24.24}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \lambda > 0$.
Differentiate equation (24.24) and simplify to obtain;

$$j''(t) = \frac{j'^2(t)}{j(t)} + \left\{\frac{1}{\alpha}\left(\frac{j'(t)}{j(t)} + \alpha\lambda\right)^2 - \lambda\left(\frac{j'(t)}{j(t)} + \alpha\lambda\right)\right\}j(t) \tag{24.25}$$

The ordinary differential equations can be obtained for the particular values of the parameters. Several analytic, semi-analytic and numerical methods can be applied to obtain the solutions of the respective differential equations. Also comparison with two or more solution methods is useful in understanding the link between ODEs and the probability distributions.

24.8 Conclusion

Ordinary differential equations (ODEs) has been obtained for the probability density function (PDF), quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of the exponentiated generalized exponential distribution.

This differential calculus, modified product rule and efficient algebraic simplifications were used to derive the various classes of the ODEs. The parameter and the supports that characterize the exponentiated generalized exponential distribution determine the nature, existence, orientation and uniqueness of the various ODEs. The results are in agreement with those available in scientific literature. Furthermore several methods can be used to obtain desirable solutions to the ODEs. This method of characterizing distributions cannot be applied to distributions whose PDF or CDF are either not differentiable or the domain of the support of the distribution contains singular points.

Acknowledgements This work was supported by Covenant University, Nigeria.

References

1. G. Steinbrecher, W.T. Shaw, Quantile mechanics. Euro. J. Appl. Math. **19**(2), 87–112 (2008)
2. H.I. Okagbue, M.O. Adamu, T.A. Anake, Quantile approximation of the Chi-square distribution using the quantile mechanics, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 477–483, 25–27 Oct 2017
3. H.I. Okagbue, M.O. Adamu, T.A. Anake, Solutions of Chi-square quantile differential equation, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 813–818, 25–27 Oct 2017
4. Y. Kabalci, On the Nakagami-m inverse cumulative distribution function: closed-form expression and its optimization by backtracking search optimization algorithm. Wireless Pers. Comm **91**(1), 1–8 (2016)
5. W.P. Elderton, *Frequency Curves and Correlation* (Charles and Edwin Layton, London, 1906)
6. N. Balakrishnan, C.D. Lai, *Continuous Bivariate Distributions*, 2nd edn. (Springer, New York, London, 2009)
7. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, Vol. 2, 2nd edn (Wiley, 1995)

8. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994), ISBN: 0–471-58495-9
9. H. Rinne, *Location Scale Distributions, Linear Estimation and Probability Plotting using MATLAB* (2010)
10. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of Fréchet distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 186–191, 25–27 Oct 2017
11. H.I. Okagbue, P.E. Oguntunde, P.O. Ugwoke, A.A, Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated generalized exponential distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 192–197, 25–27 Oct 2017
12. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of cauchy, standard cauchy and log-cauchy distributions, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 198–20425–27 Oct 2017
13. H.I. Okagbue, S.A. Bishop, A.A. Opanuga, M.O. Adamu. Classes of Ordinary differential equations obtained for the probability functions of Burr XII and Pareto distributions, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 399–404, 25–27 Oct 2017
14. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of Gompertz and Gamma Gompertz distributions, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 405–411, 25–27 Oct 2017
15. H.I. Okagbue, M.O. Adamu, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of 3-parameter Weibull distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 539–545, 25–27 Oct 2017
16. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of exponentiated Fréchet distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 546–551, 25–27 Oct 2017
17. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, S.A. Bishop, Classes of ordinary differential equations obtained for the probability functions of half-Cauchy and power Cauchy distributions, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp 552–558, 25–27 Oct 2017
18. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of exponential and truncated exponential distributions, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA), pp. 858–864, 25–27 Oct 2017
19. H.I. Okagbue, O.O. Agboola, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated Pareto distribution, *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science 25–27 Oct 2017 (San Francisco, USA pp. 865–870
20. H.I. Okagbue, O.O. Agboola, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of Gumbel distribution. in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 871–875, 25–27 Oct 2017

21. H.I. Okagbue, O.A. Odetunmibi, A.A. Opanuga, P.E. Oguntunde, Classes of ordinary differential equations obtained for the probability functions of half-normal distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp. 876–882, 25–27 Oct 2017
22. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, E.A. Suleiman, Classes of ordinary differential equations obtained for the probability functions of Harris extended exponential distribution, in *2017 Proceedings of the World Congress on Engineering and Computer Science*. Lecture Notes in Engineering and Computer Science (San Francisco, USA) pp 883–888, 25–27 Oct 2017
23. P.E. Oguntunde, A.O. Adejumo, K.A. Adepoju, Assessing the flexibility of the exponentiated generalized exponential distribution. Pacific J. Sci. Tech. **17**(1), 49–57 (2016)
24. G.M. Cordeiro, A.Z. Afify, H.M. Yousof, R.R. Pescim, G.R. Aryal, The exponentiated Weibull-H family of distributions: theory and applications. Medit. J. Math. **14**(4), 155 (2017)
25. S. Nadarajah, S. Kotz, The exponentiated type distributions. Acta Applic. Math. **92**(2), 97–111 (2006)
26. G.M. Cordeiro, E.M. Ortega, D.C.C. da Cuncha, The exponentiated generalized class of distributions. J. Data Sci. **11**, 1–27 (2013)
27. S. Rezaei, A.K. Marvasty, S. Nadarajah, M. Alizadeh, A new exponentiated class of distributions: Properties and applications. Comm. Stat. Theo. Methods **46**(12), 6054–6073 (2017)
28. R.D. Gupta, D. Kundu, Exponentiated exponential family: an alternative to gamma and Weibull distributions. Biometrical J. **43**(1), 117–130 (2001)
29. I.E. Okorie, A.C. Akpanta, J. Ohakwe, The exponentiated Gumbel Type-2 distribution: properties and application. Int. J. Math. Math. Sci. Art no. 5898356 (2016)
30. M. Pal, M.M. Ali, J. Woo, Exponentiated Weibull distribution. Statistica **66**(2), 139–147 (2006)
31. G.S. Mudholkar, D.K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Trans. Relia. **42**(2), 299–302 (1993)
32. M.N. Nassar, F.H. Eissa, On the exponentiated Weibull distribution. Comm. Stat. Theo. Methods **32**(7), 1317–1336 (2003)
33. P.E. Oguntunde, A.O. Adejumo, O.S. Balogun, Statistical properties of the exponentiated generalized inverted exponential distribution. Appl. Math. **4**(2), 47–55 (2014)
34. A.J. Lemonte, G.M. Cordeiro, The exponentiated generalized inverse Gaussian distribution. Stat. Prob. Lett. **81**(4), 506–517 (2011)
35. A. Flaih, H. Elsalloukh, E. Mendi, M. Milanova, The exponentiated inverted Weibull distribution. Appl. Math. Inf. Sci **6**(2), 167–171 (2012)
36. I. Elbatal, H.Z. Muhammed, Exponentiated generalized inverse Weibull distribution. Appl. Math. Sci. **8**(81), 3997–4012 (2014)
37. M.M. Ristić, N. Balakrishnan, The gamma-exponentiated exponential distribution. J. Stat. Comput. Simul. **82**(8), 1191–1206 (2012)
38. H.H. Abu-Zinadah, A.S. Aloufi, Some characterizations of the exponentiated Gompertz distribution. Int. Math. Forum **9**(30), 1427–1439 (2014)
39. H.H. Abu-Zinadah, Six method of estimations for the shape parameter of exponentiated Gompertz distribution. Appl. Math. Sci. **8**(85–88), 4349–4359 (2014)
40. S.A. Siddiqui, S. Dwivedi, P. Dwivedi, M. Alam, Beta exponentiated Mukherjii-Islam distribution: Mathematical study of different properties. Global J. Pure Appl. Math. **12**(1), 951–964 (2016)
41. A. Fatima, A. Roohi, Transmuted exponentiated Pareto-i distribution. Pak. J. Statist **32**(1), 63–80 (2015)
42. T.K. Pogány, A. Saboor, The Gamma exponentiated exponential–Weibull distribution. Filomat **30**(12), 3159–3170 (2016)
43. S. Nadarajah, A.K. Gupta, The exponentiated gamma distribution with application to drought data. Calcutta Stat. Assoc. Bul. **59**(1–2), 29–54 (2007)
44. S. Nadarajah, The exponentiated Gumbel distribution with climate application. Environmetrics **17**(1), 13–23 (2006)

45. C.S. Lee, H.Y. Won, Inference on reliability in an exponentiated uniform distribution. J. Korean Data Info. Sci. Soc. **17**(2), 507–513 (2006)
46. G.M. Cordeiro, A.E. Gomes, C.Q. da-Silva, E.M. Ortega, The beta exponentiated Weibull distribution. J. Stat. Comput. Simul. **83**(1), 114–138 (2013)
47. S. Hashmi, A.Z. Memon, Beta exponentiated Weibull distribution (its shape and other salient characteristics). Pak. J. Stat. **32**(4), 301–327 (2016)
48. K. Rosaiah, R.R.L. Kantam, S. Kumar, Reliability test plans for exponentiated log-logistic distribution. Econ. Qual. Control **21**(2), 279–289 (2006)
49. A.A. Al-Babtain, F. Merovci, I. Elbatal, The McDonald exponentiated gamma distribution and its statistical properties. SpringerPlus, **4**(1), Art no. 2 (2015)
50. P.E. Oguntunde, O.A. Odetunmibi, A.O. Adejumo. On the exponentiated generalized Weibull distribution: a generalization of the Weibull distribution, Indian J. Sci. Tech. **8**(35), 1–7, Art. no. 67611 (2015)
51. N. Feroze, I. Elbatal, Beta exponentiated gamma distribution: some properties and estimation. Pak. J. Stat. Oper. Res. **12**(1), 141–154 (2016)
52. A. I. Shawky, R. A. Bakoban. Exponentiated gamma distribution: Different methods of estimations, J. Appl. Math., vol. 2012, Art no. 284296, 2012
53. A.I. Shawky, H.H. Abu-Zinadah, Exponentiated Pareto distribution: Different method of estimations. Int. J. Contem. Math. Sci. **4**(11), 677–693 (2009)
54. A.J. Lemonte, W. Barreto-Souza, G.M. Cordeiro, The exponentiated Kumaraswamy distribution and its log-transform. Braz. J. Prob. Stat. **27**(1), 31–53 (2013)
55. A.M. Sarhan, J. Apaloo, Exponentiated modified Weibull extension distribution. Relia. Engine. Syst. Safety **112**, 137–144 (2013)
56. A.Z. Afify, H.M. Yousof, G.G. Hamedani, G. Aryal, The exponentiated Weibull-Pareto distribution with application. J. Stat. Theory Appl. **15**, 328–346 (2016)
57. C.S. Kakde, D.T. Shirke, On exponentiated lognormal distribution. Int. J. Agric. Stat. Sci. **2**, 319–326 (2006)
58. B. Singh, N. Choudhary, The exponentiated Perks distribution. Int. J. Syst. Assur. Engine. Magt. **8**(2), 468–478 (2017)
59. A. Al-Babtain, A.A. Fattah, A.H.N. Ahmed, F. Merovci, The Kumaraswamy-transmuted exponentiated modified Weibull distribution. Comm. Stat. Simul. Comput. **46**(5), 3812–3832 (2017)
60. M. Pararai, G. Warahena-Liyanage, B.O. Oluyede, Exponentiated power Lindley-Poisson distribution: Properties and applications. Comm. Stat. Theo. Methods **46**(10), 4726–4755 (2017)
61. S. Dey, D. Kumar, P.L. Ramos, F. Louzada, Exponentiated Chen distribution: properties and estimation. Comm. Stat. Simul. Comput. (2017)
62. C.S. Kumar, S.H. S. Dharmaja, The exponentiated reduced Kies distribution: properties and applications. Comm. Stat. Theo. Methods (2017)
63. Y. Chung, D.K. Dey, M. Jung. The exponentiated inverse Weibull geometric distribution, Pak. J. Stat. **33**(3), 161–178
64. S. Nadarajah, S.A.A. Bakar, An exponentiated geometric distribution. Appl. Math. Model. **40**(13), 6775–6784 (2016)
65. V. Nekoukhou, M.H. Alamatsaz, H. Bidram, A note on exponentiated geometric distribution: another generalization of geometric distribution. Comm. Stat. Theo. Methods **45**(5), 1575 (2016)
66. Y. Chung, Y. Kang, The exponentiated Weibull geometric distribution: properties and estimations. Comm. Stat. Appl. Methods **21**(2), 147–160 (2014)
67. A. Fattah, S. Nadarajah, A.H.N. Ahmed, The exponentiated transmuted Weibull geometric distribution with application in survival analysis. Comm. Stat. Simul. Comput. (2017)
68. W. Gui. Exponentiated half logistic distribution: different estimation methods and joint confidence regions, Comm. Stat. Simul. Comput. (2017)
69. D. Deka, B. Das, B.K. Baruah, Transmuted exponentiated Gumbel distribution (TEGD) and its application to water quality data. Pak. J. Stat. Oper. Res. **13**(1), 115–126 (2017)
70. N. Bursa, G. Ozel, The exponentiated Kumaraswamy-power function distribution. Hacettepe J. Math. Stat. **46**(2), 277–292 (2017)

71. N.V. Mendoza, E.M. Ortega, G.M. Cordeiro, The exponentiated-log-logistic geometric distribution: Dual activation. Comm. Stat. Theo. Methods **45**(13), 3838–3859 (2016)
72. A.H. El-Bassiouny, M.A. EL-Damcese, A. Mustafa, M.S. Eliwa, Bivariate exponentaited generalized Weibull-Gompertz distribution. J. Appl. Prob. **11**(1), 25–46 (2016)

Chapter 25
Quantifying the Mind: Worry, Tension, and Anxiety

Toru Yazawa

Abstract To quantify non-linear behavior of physiologic system such as the cardiovascular control system, we first used lobsters because we are invertebrate neurobiologists. After finding that the lobsters can display its emotion by changing the pattern of heart-beating, we extended the method to human: The heartbeat represents momently varying emotional tension. We show that this variation of the inner world is detectable and quantifiable using a long-time electrocardiogram (EKG). In each investigation, we captured 2000 heartbeats without missing a single beat. The heartbeat interval time series was analyzed by "modified detrended fluctuation analysis (mDFA)" technique, which we have recently developed by our group. The mDFA calculates the scaling exponent (SI, scaling index). A normal healthy heartbeat exhibits an SI of around 1.0. The heartbeat recorded from subjects who have stress and anxiety exhibited a lower SI. The values of SI changed one right after the other when circumstances and atmospheres surrounding subjects were changed. We report that the mDFA technique is a useful computation method for checking the mind and health.

Keywords Anxiety · EKG · Electrocardiogram · Fear · Heartbeat-interval time series · Modified detrended fluctuation analysis · mDFA · Quantitative measurement · Tension

25.1 Introduction

Nonlinear regulatory systems for controlling the organs, such as the heart, is operating in the state far from equilibrium: i.e., its functioning is not stable but dynamic. Maintaining constancy is not equivalent to maintaining equilibrium. This means that physiologists need another theory instead of a classical homeostasis theory.

T. Yazawa (✉)
Biological Science, Tokyo Metropolitan University, Minami-Ohsawa, Hachioji, Tokyo 192-0397, Japan
e-mail: yazawa-tohru@tmu.ac.jp; yazawatorujp@yahoo.co.jp

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_25

Researchers have tried to characterize the fluctuation of the heartbeat by a scaling exponent (SI, scaling index), and demonstrated that SIs were altered with disease and aging [1].

Goldberger et al. [1], Ivanov et al. [2] and Peng et al. [3] suggested that their method is applicable for the diagnosis of human heart disease, because they succeeded to show the difference between healthy and sick human subjects, although they did not study the transitory phenomena of subjects who were healthy in the past but are presently not healthy. Demonstrating the detection of long-range correlation [2] and multi-fractal analysis [3], they eventually proposed that the scaling nature of heartbeat fluctuation is potentially usable for the diagnosis. However, for a practical application, for example, predicting when in the near future the heart stops its beating, more experimental approaches are necessary [4].

Our approach is to make an index system that can describe the condition of the cardio-vascular system by a quantitative way. It is believed that beat-to-beat variations of the heart rate reflect modulation of cardiac control center [5]. If one has such a quantitative description method in hand, he/she can predict forthcoming illness. We consider that we hope to find the way for distinguishing sick conditions from normal conditions, based on electro-physiological heart data, electrocardiogram, (EKG, abbreviation after Willem Einthoven).

Not only in vertebrate animals like human, but also in invertebrates such as crustaceans, they have the heart. One may think that there are big differences between vertebrate and crustaceans in the anatomy of the cardio-vascular system. However, the heart is a pump that propels blood, and that is equipped with acceleration and breaking devises, the control of nerve fibers. Those are common features for both, in vertebrates and invertebrates. Therefore, fundamentally, all the cardio-vascular system operates in the same mode; ontogenetically designed as a pump equipped with positive and negative feedback controllers. Both vertebrates and invertebrates have a very complex cardio-vascular system in structure, function and regulation. Beneficial points for adopting invertebrates are; there are less-ethical problems and there exists a large amount of accumulated knowledge about the heart evolution, and all the hearts must have the same evolutional origin.

From above arguments we may conclude that there are common features between vertebrate (human) and invertebrate (lobster). Basic architecture of the heart and of its control mechanism could be fundamentally identical for all ‘hearted’ animals. Crustaceans supply, thus good model specimens, for the research on neurodynamical control of the heart.

We have observed that a healthy-looking crustacean heart stop unpredictably resulting in a failure of heart’s pumping-blood function [6]. This can happen to a human. It would be the worst-case scenario for the cardio-vascular system. Generally, cardiac failure has a principal underlying aetiology of ischemic damage arising from vascular insufficiency [7].

Meanwhile, we know that our (human) heart function directly reflects psychology. The autonomic nerves can change heart function according to stimuli such as stress that we receive. Thus, such stimuli from environment cause dynamic changes of our internal world and changes the heart function as well.

This study is a challenge to quantify the internal world using long-time EKG recordings. We have recently made a novel method for EKG analysis, which is the modified detrended fluctuation analysis (mDFA) [6]. In the present study, we show that mental changes which occur internally can be measured with the combination of EKG and mDFA. Although mDFA is not a method well known to the public, we would like to introduce mDFA, which is helpful for checking health and the mind as we exemplify in the present paper.

25.2 Heartbeat Recording and Ethics

We used a Power Lab system (ADInstruments, Australia) to record heartbeats. For recording human EKG, we used a set of three ready-made silver/silver chloride disposable electrodes (positive, negative, and ground) (Vitrode V, Nihon Koden, Japan), which is commercially available. Wires from the EKG electrodes were connected to our lab-made amplifier, which is activated by two button batteries at the voltage of 3.0 V. This equipment is only used in the "university laboratory condition" with a certificate of consent written by all subjects. The EKG signals were passed to the Power Lab system. Over 500 subjects have been tested so far, but no electrical accident has happened by the home-made amplifier recording.

The human heartbeats were recorded outside of a hospital, in for example university laboratories and convention halls (The Innovation Japan Exhibition). All subjects were treated as per the ethical control regulations of following universities (Tokyo Metropolitan University; Tokyo Women's Medical University; Universitas Advent Indonesia, Bandung; Universitas Airlangga, Surabaya, Indonesia).

25.3 Heartbeat Recording and Analysis

A computation method mDFA [6] is described in the paper appeared in the conference proceedings WCECS 2017 [8].

Briefly, we always use a routine mDFA [6]. Our routine works include: (1) obtaining a baseline-stable EKG at 1 kHz sampling rate (a lab-made amplifier with an input time-constant < 0.22 s), (2) detecting peaks of heartbeat, (3) measuring peak to peak interval (such as R-R peak interval of conventional EKG), (4) constructing inter-heartbeat interval time series, and (5) analyzing the time series using mDFA program.

mDFA uses a consecutively recorded ~2000 heartbeat data. The number 2000 could be ideal number of heartbeat if one wants to use mDFA. We tested a longer recording period, for example EKG for 2 h. Long data is not adequate to use, because subjects are NEVER stable. Nobody can keep a steady state, i.e., keeping sitting posture for 2 h. And more importantly, the heart seems NOT to keep staying at a stable condition for such a long period. Instead, the heart control by the brain is very

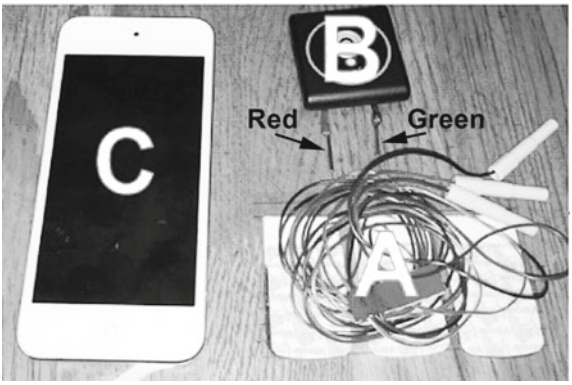


Fig. 25.1 Electrocardiogram recorder and mDFA calculator

variable and thus dynamic. We therefore fixed our data length for about 30 min or so, which is a period length for about 2000 beat. A ~2000 beat is key length of mDFA technology.

25.4 EKG-Recording-Computing Device

Figure 25.1 shows our data logging and mDFA computing devices. Figure 25.1a shows three individual electrodes for EKG recording, commercially available, in-hospital use, using for a prematurely-born baby in an incubator, Vitrode V, Nihon Koden, Tokyo, Japan. Figure 25.1b shows an EKG-amplifier, heartbeat interval calculator, and Bluetooth radio transmitter. Red and green end terminals (see arrows in Fig. 25.1) are lab-made input terminals from the EKG electrodes. Figure 25.1c shows an iPod (Apple, USA) which has a computation program, i.e., mDFA, that we made and was incorporated in it.

Figure 25.2 shows a practical view of iPod touch screen. To start recording, an operator can touch the rec button (Rec), and then after completing capture of 2000 beats, it automatically computes SI. As can be seen in the figure, it computes. SIs are computed from various box size ranges (see the reference [6, 8] in detail), [10; 30], [30; 70], [70; 140], [130; 270], [51; 100], [30; 140], and [30; 270]. For the final SI-result, we use the last one, here it is 0.531390, as explained in the reference [6, 8].

25.5 Results

Case 1: Driving safely

Volunteer (a man in his 60s) drove a car from his home to a town 150 km away to see his mother-in-low who is hospitalized. He has been driving the road so many

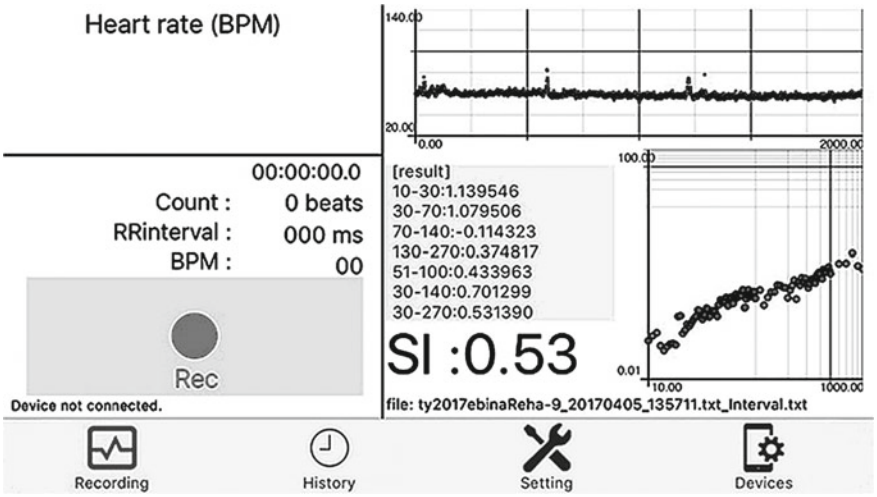


Fig. 25.2 An example screen view of iPod (lab-made, not for sale)

time, thus he is familiar with the road conditions every corner. Furthermore, he drove safely as possible as he can by obeying the speed limit. We recorded his EKGs while driving, computed the scaling exponents by the device (Figs. 25.1 and 25.2).

Driver's heart rate was monitored (Fig. 25.3a, b). Figure 25.3d shows an example result of mDFA computation. Figure 25.3b represents a 2000 beat recording. Figure 25.3a shows an expanded time series of heart rate recording (see arrows). The program instantaneously computed the scaling exponent (SI) (Fig. 25.3c). Figure 25.3d shows a summary of the characteristics of the data {i.e., the file-name, 37 min and 0.2 s recording in total-time for the 2000 beats, R-R interval value in the end of recording, 1046 ms, and heart rate (57 beat per min, BPM) of the last heartbeat.} Figure 25.3c indicates that it was proved that driving safely gives a perfect healthy scaling exponent near 1.0. Here, the SI is 0.99.

Table 25.1 shows an example of EKG-heart-rate monitor and the scaling exponent (SI) computation during the driving for many hours. We confirmed that we can indeed quantify the mind of driver, or internal world of any human-subjects, such as the vehicle-driver, because we consider that the heartbeat reflects the mind or that the heart is the window of the brain [6, 9].

In Table 25.1, when starting driving, SI showed a low value (SI = 0.84, see Data No. 1). This can be explained that the driver solved many worries about fuel Gas, driving route and so force. When taking express way, the driver kept its speed limit and enjoyed blue sky of a spring morning day (SI = 1.03, Data No. 2). Many vehicles over-took his car one right after the other although only some cars were followed his car. He continued driving safely (from No. 3 to No. 8). One can see that his safe-driving gave good SI, i.e., near 1.0 as can be seen the SI values from No. 2 to No. 8.

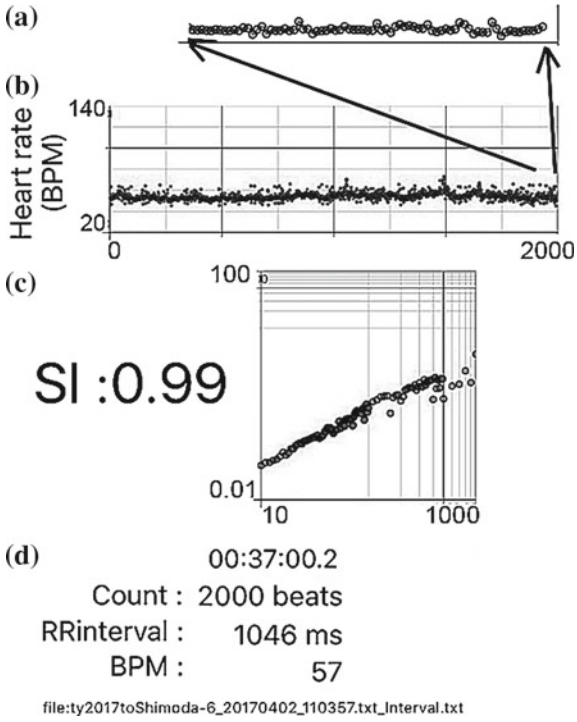


Fig. 25.3 EKG monitoring and mDFA results

We found that eating lunch decreased the SI value (SI = 0.61, Table 25.1, Data No. 9). We can explain this results as followings: the mind (his brain function, i.e., autonomic nerve function) concentrated to enjoy foods, digesting them in the stomach and even pay less attention to environment, the condition of which is a kind of state loosing dynamic response of the brain that momentarily controls the heartbeats every second.

We also confirmed that mDFA can capture anxiety/worry of a subject. One can see that nervous conditions put SI lower. Table 25.1 shows SIs of the subject who walked into the hospital and visited and stayed the room of his mother-in-law (see Data No. 10 and 11, SI = 0.64 and 0.53, respectively). After the hospital, the person's SI recovered during driving and shopping at the super market (Nos. 12 and 13).

And most interestingly, when meeting a new person (the drivers brother-in-low) to greet him, the SI showed a very low value (Data No. 14, SI = 0.77) again, which indicates that he is very nervous NOT to display an ungentlemanly attitude.

In conclusion, SI derived from mDFA can determine internal world (Fig. 25.4).

Technically before conducting mDFA, our three must do: (1) From heartbeat raw data, noise from power souse and body movement should be avoided. (2) But, skipping heartbeats and/or arrhythmic beats are not noise. Such irregularity is very important information for disentangling health care. We believe that this mDFA-technology can quantify acute psychological distress and/or jubilation although there

Table 25.1 EKG-heart-rate monitoring and mDFA results, SI

Data no.	Driving a car very safely to see relatives	SI
1	Talking about Gas and driving route	0.84
2	Driving legal speed, express way	1.03
3	Driving legal speed, local road	1.09
4	Driving legal speed, local road	1
5	Driving legal speed, local road	1.03
6	Driving legal speed, local road	0.99
7	Driving legal speed, local road	1.04
8	Driving legal speed, local road	1
9	Lunch, sweet-sour pork, special black vinegar	0.61
10	Visit hospital	0.64
11	Visit hospital	0.53
12	Driving, super market	0.84
13	Super market	0.99
14	Ggreeting brother-in-law	0.77

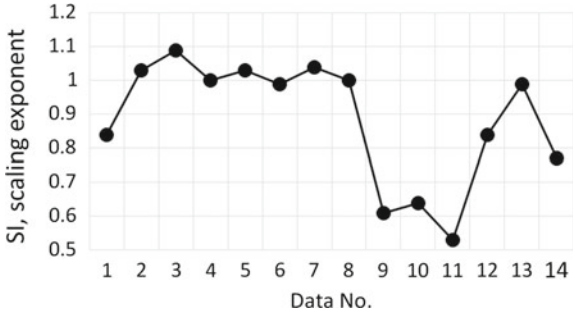


Fig. 25.4 SI derived form mDFA can determin internal world

are many more research questions. Irregular heartbeats should not be removed from raw data when constructing time series data. (3) We keep the size/length of data constant; a raw data must contain about 2000 beats, ranging from 1900 to 2100, or always exactly 2000.

Case 2: Overseas flight

Figure 25.5 shows 27 data while a volunteer (a man 66) traveled from the Narita-Tokyo Airport to the Washington Dulles International Airport to attend a conference held in the USA. Each analysis gives one SI from a 2000-beat-EKG as shown in Fig. 25.5.

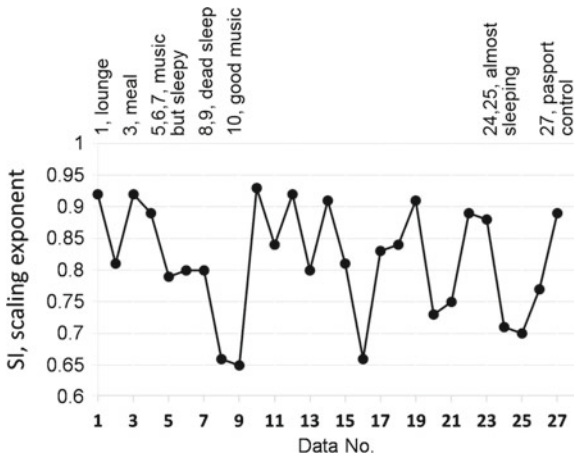


Fig. 25.5 Scaling exponent records while traveling

We confirmed that the SI values can represent the internal world of the subject. This subject did not get enough sleep before taking oversea flight. Therefore, SI in Fig. 25.5 always less than 1.0—please compare with the case shown in Fig. 25.4, where the driver took good sleep before driving.

When the subject was at an aroused state such as staying in the waiting lounge, SI shows a "healthy" SI, i.e., here it is 0.92. (Data No. 1). Meal service, SI is 0.92 (Data No. 3). When he enjoyed "good days music, 70 s hit chart," SI is 0.93 (Data No. 10). While he was sleeping, SI goes down lasting for about one hour (Data No. 8 and 9, SI = 0.66 and 0.65, respectively). Good music but not so exciting as "good days music" makes SI a lower value (Data Nos. 5, 6, and 7, SI = 0.79, 0.8, and 0.8, respectively). While he was watching a boring program, feeling sleepy, SI gets a low value (Data No. 24 and 25, SI = 0.71 and 0.70, respectively). At passport control after landing, SI = 0.89 (Data No. 27). In summary, it is surely observed that sleepy state lowers SI.

In conclusion, happy life could fundamentally guarantee a healthy exponent. Anxiety and stress lowered the scaling exponent. mDFA might reflect psychological and physical internal bodily state. mDFA might look at internal state through the heart. The heart is the window of the mind.

25.6 Discussion

This study suggests that the scaling exponents computed by mDFA can quantify the mind such as stress.

Although we need much more comprehensive examples, we propose that mDFA is helpful computation tool in the research on emotion, particularly fear and anxiety

disorders, understanding how emotion is encoded in the heartbeat time series, in animal models and humans.

If the body is tortured by stimuli from environment, and/or if some stimuli would harm us internally, which is invisible from outside, we would be upsetting for the nervous system, because we can realize that stimuli is distorting the autonomic nerve function, little of which has been understood by human being until today [4], although we spend everyday life under advanced science and technology. Using mDFA computation, we can numerically evaluate/quantify the state of our body, even it is invisible to us.

Although we (basic scientists, biologists) cannot make by ourselves, making a gadget is very rewarding. It is the right time to start making it. The gadget can work: (1) recording 2000 consecutive heartbeats without missing even a single pulse, (2) computing automatically the scaling exponent that can check the scaling exponent = 1.0, which is perfectly healthy state [6], and finally (3) the gadget would capture what is going on in front of, around, and inside our mind. It gives us health information, each time we use it, for example, on an everyday basis.

In the present paper, we would suggest that we have entered the world experiencing seeing inside without sight. Sometimes a new technology does not have to be supercomplicated. mDFA computation is a kind of high school level mathematics instead of sophisticated nonlinear measures and/or linear complex computation like the HRV, the heart rate variability. mDFA looks at how the brain communicate with the heart and also with the world. mDFA is a tool that enable us to explore previously uncharted territories.

Acknowledgements This work was supported by the JSPS Grant 17K01364.

References

1. A.L. Goldberger, L.A.N. Amaral, J.M. Hausdorff, P.C. Ivanov, C.-K. Peng, H.E. Stanley, Fractal dynamics in physiology: alterations with disease and aging. Proc. Natl. Acad. Sci. **99**(Suppl. 1), 2466–2472 (2002)
2. P.C. Ivanov, L.A.N. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z.R. Struzik, H.E. Stanley, Multifractality in human heartbeat dynamics. Nature **399**, 461–465 (1999)
3. C.-K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, Long-range correlation and non-Gaussian behavior of the heartbeat. Phys. Rev. Lett. **70**, 1343–1346 (1993)
4. K. Hu, P.C. Ivanov, M.F. Hilton, Z. Chen, R.T. Ayers, H.E. Stanley, S.A. Shea, Endogenous circadian rhythm in an index of cardiac vulnerability independent of changes in behavior. Proc. Natl. Acad. Sci. **101**(52), 18223–18227 (2004)
5. J.P. Saul, Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. News Physiol. Sci. **5**, 32–37 (1990)
6. T. Yazawa, *Modified Detrended Fluctuation Analysis, mDFA* (ASME monograph, 2015)
7. Nicola Smart, Catherine A. Risebro, Athalie A.D. Melville, Kelvin Moses, Robert J. Schwartz, Kenneth R. Chien, Paul R. Riley, Thymosin beta-4 induces adult epicardial progenitor mobilization and neovascularization. Nature **445**, 177–182 (2007)

8. T. Yazawa, Anxiety, worry, and fear: quantifying the mind using EKG-mDFA method, in *Proceedings of the World Congress on Engineering and Computer Science 2017*. Lecture Notes in Engineering and Computer Science, 25–27 Oct 2017, San Francisco, USA, pp. 519–524
9. T. Yazawa, Quantification of stress: a case study using modified detrended fluctuation analysis (mDFA) of the heartbeat, in *Proceedings of the 20th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2016),* July, Orlando, FL, USA, 5 pp.
10. C.-K. Peng et al., Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos **5**, 82–87 (1995)
11. T. Yazawa, Invisible emotion, anxiety and fear: quantifying the mind using EKG with mDFA, in *Proceedings of the 8th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2017)*, March, Orlando, FL, USA, pp. 201–205

Chapter 26
Half-Cauchy and Power Cauchy Distributions: Ordinary Differential Equations

Hilary I. Okagbue, Muminu O. Adamu, Patience I. Adamu, Sheila A. Bishop and Ezinne C. Erondu

Abstract In this chapter, homogenous ordinary differential equations (ODES) of different orders were obtained for the probability density function, quantile function, survival function inverse survival function, hazard function and reversed hazard functions of half-Cauchy and power Cauchy distributions. This is possible since the aforementioned probability functions are differentiable. Differentiation and modified product rule were used to obtain the required ordinary differential equations, whose solutions are the respective probability functions. The different conditions necessary for the existence of the ODEs were obtained and it is almost in consistent with the support that defined the various probability functions considered. The parameters that defined each distribution greatly affect the nature of the ODEs obtained. This method provides new ways of classifying and approximating other probability distributions apart from half-Cauchy and power Cauchy distributions considered in this chapter. In addition, the result of the quantile function can be compared with quantile approximation using the quantile mechanics.

Keywords Differential calculus · Half-Cauchy distribution · Hazard function Inverse survival function · Power Cauchy distribution · Quantile function Quantile mechanics · Reversed hazard function · Survival function

H. I. Okagbue (✉) · P. I. Adamu · S. A. Bishop · E. C. Erondu
Department of Mathematics, Covenant University, Ota, Nigeria
e-mail: hilary.okagbue@covenantuniversity.edu.ng

P. I. Adamu
e-mail: patience.adamu@covenantuniversity.edu.ng

S. A. Bishop
e-mail: sheila.bishop@covenantuniversity.edu.ng

E. C. Erondu
e-mail: ezinneerondu@yahoo.com

M. O. Adamu
Department of Mathematics, University of Lagos, Akoka, Nigeria
e-mail: madamu@unilag.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_26

26.1 Introduction

Calculus is a very key tool in the determination of mode of a given probability distribution and in estimation of parameters of probability distributions, amongst other uses. The method of maximum likelihood is an example.

Differential equations often arise from the understanding and modeling of real life problems or some observed physical phenomena. Approximations of probability functions are one of the major areas of application of calculus and ordinary differential equations in mathematical statistics. The approximations are helpful in the recovery of the probability functions of complex distributions [1–4] especially in quantile approximations.

Apart from mode estimation, parameter estimation and approximation, probability density function (PDF) of probability distributions can be expressed as ODE whose solution is the PDF. Some of which are available. They include: beta distribution [5], Lomax distribution [6], beta prime distribution [7], Laplace distribution [8] and raised cosine distribution [9].

The aim of this research is to develop homogenous ordinary differential equations for the probability density function (PDF), Quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of half-Cauchy and power Cauchy distributions. This will also help to provide the answers as to whether there are discrepancies between the support of the distributions and the necessary conditions for the existence of the ODE. Similar results for other distributions have been proposed, see [10–22] for details.

Half-Cauchy distribution is obtained by restricting the domain of the standard Cauchy distribution to only positive values or observations. Polson and Scott [23] argued for the replacement of inverse-gamma distribution with this distribution in Bayesian hierarchical models while Shaw [24] suggested that the distribution can be used in lieu of the exponential distribution in prediction and modeling. The distribution, according to [25], is one of few distributions related to the ratio of two folded normal distributions. The distribution is also related to the folded t distribution as proved by [26]. Half-Cauchy distribution is one of distributions that are self-decomposable [27] and infinitely divisible [28]. Details on the distribution can be found in [29]. Some modifications or proposed improved models of the distribution includes: Kumaraswamy half-Cauchy distribution [30], beta half-Cauchy distribution [31] and generalized odd half-Cauchy family of distributions [32]. Paradis et al. [33] applied the distribution in ecological modeling.

In an attempt to find a suitable distribution that effectively fits medical survival data, Power Cauchy Distribution was proposed by [34]. The distribution is a submodel of transformed beta family earlier proposed by [35]. Tahir et al. [36], proposed Poisson power Cauchy as an improved model over the power Cauchy distribution.

26.2 Half-Cauchy Distribution

26.2.1 Probability Density Function

The probability density function of the half-Cauchy distribution is given by;

$$f(x) = \frac{2}{\pi\sigma\left[1 + \left(\frac{x}{\sigma}\right)^2\right]} \tag{26.1}$$

To obtain the first order ordinary differential equation for the probability density function of the half-Cauchy distribution, differentiate equation (26.1);

$$f'(x) = -\frac{4x}{\pi\sigma^3\left[1 + \left(\frac{x}{\sigma}\right)^2\right]^2} \tag{26.2}$$

The condition necessary for the existence of equation is $\sigma, x > 0$.
Simplify;

$$\left(1 + \left(\frac{x}{\sigma}\right)^2\right) f'(x) = -\frac{2x}{\sigma^2} f(x) \tag{26.3}$$

The first order ordinary differential for the probability density function of the half-Cauchy distribution is given as;

$$(\sigma^2 + x^2) f'(x) + 2x f(x) = 0 \tag{26.4}$$

with initial value condition;

$$f(1) = \frac{2\sigma}{\pi(\sigma^2 + 1)} \tag{26.5}$$

26.2.2 Quantile Function

The Quantile function of the half-Cauchy distribution is given by;

$$Q(p) = \sigma \tan\left(\frac{\pi p}{2}\right) \tag{26.6}$$

To obtain the first order ordinary differential equation for the Quantile function of the half-Cauchy distribution, differentiate equation (26.6);

$$Q'(p) = \frac{\pi\sigma}{2} \sec^2\left(\frac{\pi p}{2}\right) \tag{26.7}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$. Applying the trigonometric identity which is;

$$\sec^2\left(\frac{\pi p}{2}\right) = \tan^2\left(\frac{\pi p}{2}\right) + 1 \tag{26.8}$$

$$Q'(p) = \frac{\pi\sigma}{2}\left(\tan^2\left(\frac{\pi p}{2}\right) + 1\right) \tag{26.9}$$

Equation (26.6) can be written as;

$$\left(\frac{Q(p)}{\sigma}\right)^2 = \tan^2\left(\frac{\pi p}{2}\right) \tag{26.10}$$

Substitute Eq. (26.10) into Eq. (26.9);

$$Q'(p) = \frac{\pi\sigma}{2}\left(\frac{Q^2(p)}{\sigma^2} + 1\right) \tag{26.11}$$

$$Q'(p) = \frac{\pi}{2\sigma}\left(Q^2(p) + \sigma^2\right) \tag{26.12}$$

The first order ordinary differential for the Quantile function of the half-Cauchy distribution is given as;

$$2\sigma\, Q'(p) - \pi(Q^2(p) + \sigma^2) = 0 \tag{26.13}$$

with initial value condition; $Q(0.1) = 0.1584\sigma$.

26.2.3 Survival Function

The survival function of the half-Cauchy distribution is given by;

$$S(t) = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right) \tag{26.14}$$

To obtain the first order ordinary differential equation for the survival function of the half-Cauchy distribution, differentiate equation (26.14);

$$S'(t) = -\frac{2}{\pi\sigma\left[1 + \left(\frac{t}{\sigma}\right)^2\right]} \tag{26.15}$$

The condition necessary for the existence of equation is $\sigma, x > 0$.

$$S'(t) = -\frac{2\sigma}{\pi(\sigma^2 + t^2)} \tag{26.16}$$

The first order ordinary differential for the survival function of the half-Cauchy distribution is given as;

$$\pi(\sigma^2 + t^2)S'(t) + 2\sigma = 0 \tag{26.17}$$

with initial condition; $S(1) = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{1}{\sigma}\right)$.

26.2.4 Inverse Survival Function

The inverse survival function of the half-Cauchy distribution is given by;

$$Q(p) = \sigma \tan\left(\frac{\pi(1-p)}{2}\right) \tag{26.18}$$

To obtain the first order ordinary differential equation for the inverse survival function of the half-Cauchy distribution, differentiate equation (26.18);

$$Q'(p) = -\frac{\pi\sigma}{2}\sec^2\left(\frac{\pi(1-p)}{2}\right) \tag{26.19}$$

The condition necessary for the existence of equation is $\sigma > 0, 0 < p < 1$.

Applying the same technique as obtained from the Quantile function, to obtain that;

$$Q'(p) = -\frac{\pi\sigma}{2}\left(\frac{Q^2(p)}{\sigma^2} + 1\right) \tag{26.20}$$

$$Q'(p) = -\frac{\pi}{2\sigma}\left(Q^2(p) + \sigma^2\right) \tag{26.21}$$

The first order ordinary differential for the inverse survival function of the half-Cauchy distribution is given as;

$$2\sigma Q'(p) + \pi(Q^2(p) + \sigma^2) = 0 \tag{26.22}$$

with initial condition; $Q(0.1) = 6.31375\sigma$.

26.2.5 Hazard Function

The hazard function of the half-Cauchy distribution is given by;

$$h(t) = \frac{2}{\pi\sigma\left[1+\left(\frac{t}{\sigma}\right)^2\right]\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)\right]} \tag{26.23}$$

To obtain the first order ordinary differential equation for the hazard function of the half-Cauchy distribution, differentiate equation (26.23);

$$h'(t) = \left(-\frac{2t}{\sigma^2+t^2} + h(t)\right)h(t) \tag{26.24}$$

The condition necessary for the existence of equation is $\sigma, t > 0$.

The first order ordinary differential for the hazard function of the half-Cauchy distribution is given as;

$$(\sigma^2+t^2)h'(t) - (\sigma^2+t^2)h^2(t) + 2th(t) = 0 \tag{26.25}$$

with initial condition; $h(1) = \frac{2\sigma}{\pi\left[\sigma^2+1\right]\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{1}{\sigma}\right)\right]}$.

26.2.6 Reversed Hazard Function

The reversed hazard function of the half-Cauchy distribution is given by;

$$j(t) = \frac{2}{\pi\sigma\left[1+\left(\frac{t}{\sigma}\right)^2\right]\left[\frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)\right]} \tag{26.26}$$

To obtain the first order ordinary differential equation for the probability density function of the half-Cauchy distribution, differentiate equation (26.26);

$$j'(t) = -\left\{\frac{\left(\frac{2t}{\sigma^2}\right)}{\left[1+\left(\frac{t}{\sigma}\right)^2\right]} + \frac{\frac{2}{\pi\sigma}\left[1+\left(\frac{t}{\sigma}\right)^2\right]^{-1}}{\left[\frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)\right]}\right\} j(t) \tag{26.27}$$

The condition necessary for the existence of equation is $\sigma, t > 0$.

$$j'(t) = -\left(\frac{2t}{\sigma^2+t^2} + j(t)\right)j(t) \tag{26.28}$$

The first order ordinary differential for the reversed hazard function of the half-Cauchy distribution is given as;

$$(\sigma^2 + t^2)j'(t) + (\sigma^2 + t^2)j^2(t) + 2tj(t) = 0 \tag{26.29}$$

with initial condition; $j(1) = \frac{2\sigma}{\pi[\sigma^2+1][\frac{2}{\pi}\tan^{-1}(\frac{1}{\sigma})]}$.

26.3 Power Cauchy Distribution

26.3.1 Probability Density Function

The probability density function of the power Cauchy distribution is given by;

$$f(x) = \frac{\frac{2\alpha}{\pi\sigma}\left(\frac{x}{\sigma}\right)^{\alpha-1}}{\left[1 + \left(\frac{x}{\sigma}\right)^{2\alpha}\right]} \tag{26.30}$$

To obtain the first order ordinary differential equation for the probability density function of the power Cauchy distribution, differentiate equation (26.30) and simplify;

$$f'(x) = \left(\frac{\alpha - 1}{x} - \pi f(x)\left(\frac{x}{\sigma}\right)^{\alpha}\right) f(x) \tag{26.31}$$

The condition necessary for the existence of equation is $\alpha, \sigma, x > 0$.

$$\sigma^{\alpha} x f'(x) - \sigma^{\alpha}(\alpha - 1) f(x) + \pi x^{\alpha+1} f^2(x) = 0 \tag{26.32}$$

Some cases are considered;
When $\alpha = 1$, Eq. (26.32) becomes;

$$\sigma f'(x) + \pi x f^2(x) = 0 \tag{26.33}$$

When $\alpha = 2$, Eq. (26.32) becomes;

$$\sigma^2 x f'(x) - \sigma^2 f(x) + \pi x^3 f^2(x) = 0 \tag{26.34}$$

When $\alpha = 3$, Eq. (26.32) becomes;

$$\sigma^3 x f'(x) - 2\sigma^3 f(x) + \pi x^4 f^2(x) = 0 \tag{26.35}$$

26.3.2 Quantile Function

The Quantile function of the power Cauchy distribution is given by;

$$Q(p) = \sigma\left(\tan\left(\frac{\pi p}{2}\right)\right)^{\frac{1}{\alpha}} \tag{26.36}$$

To obtain the first order ordinary differential equation for the Quantile function of the power Cauchy distribution, differentiate equation (26.36);

$$Q'(p) = \frac{\sigma}{\alpha}\left(\tan\left(\frac{\pi p}{2}\right)\right)^{\frac{1}{\alpha}-1}\frac{\pi}{2}\sec^2\left(\frac{\pi p}{2}\right) \tag{26.37}$$

The condition necessary for the existence of equation is $\alpha, \sigma > 0, 0 < p < 1$.

$$Q'(p) = \frac{\frac{\sigma}{\alpha}\left(\tan\left(\frac{\pi p}{2}\right)\right)^{\frac{1}{\alpha}}\frac{\pi}{2}\sec^2\left(\frac{\pi p}{2}\right)}{\tan\left(\frac{\pi p}{2}\right)} \tag{26.38}$$

Substitute Eq. (26.36) into Eq. (26.38);

$$Q'(p) = \frac{\frac{\pi}{2}\sec^2\left(\frac{\pi p}{2}\right)Q(p)}{\alpha\,\tan\left(\frac{\pi p}{2}\right)} \tag{26.39}$$

Equation (26.36) can be written as;

$$\left(\frac{Q(p)}{\sigma}\right)^{\alpha} = \tan\left(\frac{\pi p}{2}\right) \tag{26.40}$$

Substitute Eq. (26.40) into Eq. (26.39);

$$Q'(p) = \frac{\frac{\pi\sigma^{\alpha}}{2}\sec^2\left(\frac{\pi p}{2}\right)Q(p)}{\alpha Q^{\alpha}(p)} \tag{26.41}$$

$$Q'(p) = \frac{\pi\sigma^{\alpha}\sec^2\left(\frac{\pi p}{2}\right)Q^{1-\alpha}(p)}{2\alpha} \tag{26.42}$$

Applying the trigonometric identity which is;

$$\sec^2\left(\frac{\pi p}{2}\right) = \tan^2\left(\frac{\pi p}{2}\right) + 1 \tag{26.43}$$

Square the both sides of Eq. (26.40)

$$\left(\frac{Q(p)}{\sigma}\right)^{2\alpha} = \tan^2\left(\frac{\pi p}{2}\right) \tag{26.44}$$

Substitute Eq. (26.44) into Eq. (26.43), to obtain;

$$\sec^2\left(\frac{\pi p}{2}\right) = \left(\frac{Q(p)}{\sigma}\right)^{2\alpha} + 1 \tag{26.45}$$

Substitute Eq. (26.45) into Eq. (26.42);

$$Q'(p) = \frac{\pi\sigma^{\alpha}\left(\left(\frac{Q(p)}{\sigma}\right)^{2\alpha} + 1\right)Q^{1-\alpha}(p)}{2\alpha} \tag{26.46}$$

$$Q'(p) = \frac{\pi\left(Q^{2\alpha}(p) + \sigma^{2\alpha}\right)Q^{1-\alpha}(p)}{2\alpha\sigma^{\alpha}} \tag{26.47}$$

$$Q'(p) = \frac{\pi\left(Q^{1+\alpha}(p) + \sigma^{2\alpha}Q^{1-\alpha}(p)\right)}{2\alpha\sigma^{\alpha}} \tag{26.48}$$

$$2\alpha\sigma^{\alpha}Q'(p) - \pi\left(Q^{1+\alpha}(p) + \sigma^{2\alpha}Q^{1-\alpha}(p)\right) = 0 \tag{26.49}$$

26.3.3 Survival Function

The survival function of the power Cauchy distribution is given by;

$$S(t) = 1 - \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)^{\alpha} \tag{26.50}$$

To obtain the first order ordinary differential equation for the survival function of the power Cauchy distribution, differentiate equation (26.50);

$$S'(t) = -\frac{\frac{2\alpha}{\pi\sigma}\left(\frac{t}{\sigma}\right)^{\alpha-1}}{\left[1 + \left(\frac{t}{\sigma}\right)^{2\alpha}\right]} \tag{26.51}$$

The condition necessary for the existence of equation is $\alpha, \sigma, t > 0$.
Simplify to obtain;

$$\pi(\sigma^{2\alpha} + t^{2\alpha})S'(t) + 2\alpha\sigma^{\alpha}t^{\alpha-1} = 0 \tag{26.52}$$

The first order ordinary differential equations can be obtained for the particular values of the any given parameters. Some cases are considered and shown in Table 26.1.

Table 26.1 Classes of differential equations obtained for the survival function of the power Cauchy distribution for different parameters

α	σ	π	Ordinary differential equations
1	1	1	$(1+t^2)S'(t)+2=0$
1	1	2	$(1+t^2)S'(t)+1=0$
1	2	1	$(4+t^2)S'(t)+4=0$
1	2	2	$(4+t^2)S'(t)+2=0$
2	1	1	$(1+t^4)S'(t)+4t=0$
2	1	2	$(1+t^4)S'(t)+2t=0$
2	2	1	$(16+t^4)S'(t)+16t=0$
2	2	2	$(16+t^4)S'(t)+8t=0$

26.3.4 Inverse Survival Function

The inverse survival function of the power Cauchy distribution is given by;

$$Q(p)=\sigma\left(\tan\left(\frac{\pi(1-p)}{2}\right)\right)^{\frac{1}{\alpha}} \tag{26.53}$$

To obtain the first order ordinary differential equation for the inverse survival function of the power Cauchy distribution, differentiate equation (26.53);

$$Q'(p)=-\frac{\pi(Q^{2\alpha}(p)+\sigma^{2\alpha})Q^{1-\alpha}(p)}{2\alpha\sigma^{\alpha}} \tag{26.54}$$

$$2\alpha\sigma^{\alpha}Q'(p)+\pi(Q^{1+\alpha}(p)+\sigma^{2\alpha}Q^{1-\alpha}(p))=0 \tag{26.55}$$

The first order ordinary differential equations can be obtained for the particular values of the parameters. Some cases are considered and shown in Table 26.2.

26.3.5 Hazard Function

The hazard function of the power Cauchy distribution is given by;

$$h(t)=\frac{\frac{2\alpha}{\pi\sigma}\left(\frac{t}{\sigma}\right)^{\alpha-1}}{\left[1+\left(\frac{t}{\sigma}\right)^{2\alpha}\right]\left[1-\frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)^{\alpha}\right]} \tag{26.56}$$

To obtain the first order ordinary differential equation for the hazard function of the power Cauchy distribution, differentiate equation (26.56);

The condition necessary for the existence of equation is $\alpha,\sigma,t>0$.

Table 26.2 Classes of differential equations obtained for the inverse survival function of the power Cauchy distribution for different parameters

α	σ	π	Ordinary differential equations
1	1	1	$2Q'(p) + Q^2(p) + 1 = 0$
1	1	2	$Q'(p) + Q^2(p) + 1 = 0$
1	2	1	$4Q'(p) + Q^2(p) + 4 = 0$
1	2	2	$2Q'(p) + Q^2(p) + 4 = 0$
2	1	1	$4Q(p)Q'(p) + Q^4(p) + 1 = 0$
2	1	2	$2Q(p)Q'(p) + Q^4(p) + 1 = 0$
2	2	1	$16Q(p)Q'(p) + Q^4(p) + 16 = 0$
2	2	2	$8Q(p)Q'(p) + Q^4(p) + 16 = 0$

$$h'(t) = \left\{ \frac{\alpha - 1}{t} - \frac{2\alpha t^{2\alpha - 1}}{\sigma^{2\alpha} + t^{2\alpha}} + h(t) \right\} h(t) \tag{26.57}$$

The first order ordinary differential equations can be obtained for the particular values of the parameters.

26.3.6 Reversed Hazard Function

The reversed hazard function of the power Cauchy distribution is given by;

$$j(t) = \frac{\frac{2\alpha}{\pi\sigma}\left(\frac{t}{\sigma}\right)^{\alpha - 1}}{\left[1 + \left(\frac{t}{\sigma}\right)^{2\alpha}\right]\left[\frac{2}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right)^{\alpha}\right]} \tag{26.58}$$

To obtain the first order ordinary differential equation for the reversed hazard function of the power Cauchy distribution, differentiate equation (26.58);

$$j(t) = \left\{ \frac{\alpha - 1}{t} - \frac{2\alpha t^{2\alpha - 1}}{\sigma^{2\alpha} + t^{2\alpha}} - j(t) \right\} j(t) \tag{26.59}$$

The first order ordinary differential equations can be obtained for the particular values of the parameters.

The condition necessary for the existence of equation is $\alpha, \sigma, t > 0$

26.4 Conclusion

Ordinary differential equations (ODEs) has been obtained for the probability density function (PDF), quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of the half-Cauchy and power Cauchy distributions.

This differential calculus, modified product rule and efficient algebraic simplifications were used to derive the various classes of the ODEs. The parameter and the supports that characterize the half-Cauchy and power Cauchy distributions determine the nature, existence, orientation and uniqueness of the ODEs. The results are in agreement with those available in scientific literature. Furthermore several methods can be used to obtain desirable solutions to the ODEs. This method of characterizing distributions cannot be applied to distributions whose PDF or CDF are either not differentiable or the domain of the support of the distribution contains singular points.

Acknowledgements This work was supported by Covenant University, Nigeria.

References

1. G. Steinbrecher, W.T. Shaw, Quantile mechanics. Euro. J. Appl. Math. **19**(2), 87–112 (2008)
2. H.I. Okagbue, M.O. Adamu, T.A. Anake, Quantile approximation of the Chi-square distribution using the quantile mechanics, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 477–483
3. H.I. Okagbue, M.O. Adamu, T.A. Anake, Solutions of Chi-square quantile differential equation, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 813–818
4. Y. Kabalci, On the Nakagami-m inverse cumulative distribution function: closed-form expression and its optimization by backtracking search optimization algorithm. Wirel. Pers. Commun. **91**(1), 1–8 (2016)
5. W.P. Elderton, *Frequency Curves and Correlation* (Charles and Edwin Layton, London, 1906)
6. N. Balakrishnan, C.D. Lai, *Continuous Bivariate Distributions*, 2nd edn. (Springer, New York, London, 2009)
7. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2, 2nd edn. (Wiley, 1995)
8. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994). ISBN: 0-471-58495-9
9. H. Rinne, *Location Scale Distributions, Linear Estimation and Probability Plotting Using MATLAB* (2010)
10. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of Fréchet distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 186–191
11. H.I. Okagbue, P.E. Oguntunde, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated generalized exponential distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*,

25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 192–197
12. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Cauchy, standard Cauchy and log-Cauchy distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 198–204
13. H.I. Okagbue, S.A. Bishop, A.A. Opanuga, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Burr XII and Pareto distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 399–404
14. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of Gompertz and Gamma Gompertz distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 405–411
15. H.I. Okagbue, M.O. Adamu, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of 3-parameter Weibull distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 539–545
16. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of exponentiated Fréchet Distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 546–551
17. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, S.A. Bishop, Classes of ordinary differential equations obtained for the probability functions of half-Cauchy and power Cauchy distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 552–558
18. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of exponential and truncated exponential distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 858–864
19. H.I. Okagbue, O.O. Agboola, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated Pareto distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 865–870
20. H.I. Okagbue, O.O. Agboola, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of Gumbel distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 871–875
21. H.I. Okagbue, O.A. Odetunmibi, A.A. Opanuga, P.E. Oguntunde, Classes of ordinary differential equations obtained for the probability functions of half-normal distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 876–882
22. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, E.A. Suleiman, Classes of ordinary differential equations obtained for the probability functions of Harris extended exponential distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 883–888

23. N.G. Polson, J.G. Scott, On the half-Cauchy prior for a global scale parameter. Bayes. Anal. **7**(4), 887–902 (2012)
24. M.W. Shaw, Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proc. R. Soc. B **259**, 243–248 (1995)
25. H.J. Kim, On the ratio of two folded normal distributions. Commun. Stat. Theory Methods **35**(6), 965–977 (2006)
26. S. Psarakis, J. Panaretoes, The folded t distribution. Commun. Stat. Theory Methods **19**(7), 2717–2734 (1990)
27. A. Diédhiou, On the self-decomposability of the half-Cauchy distribution. J. Math. Anal. Appl. **220**(1), 42–64 (1998)
28. L. Bondesson, On the infinite divisibility of the half-Cauchy and other decreasing densities and probability functions on the nonnegative line. Scand. Actua. J. **1987**(3), 225–247 (1987)
29. E. Jacob, K. Jayakumar, On half-Cauchy distribution and process. Int. J. Statist. Math. **3**(2), 77–81 (2012)
30. I. Ghosh, The Kumaraswamy-half-Cauchy distribution: properties and applications. J. Stat. Theo. Appl. **13**(2), 122–134
31. G.M. Cordeiro, A.J. Lemonte, The beta-half-Cauchy distribution. J. Probab. Stat. Art. no. 904705 (2011)
32. G.M. Cordeiro, M. Alizadeh, T.G. Ramires, E.M. Ortega, The generalized odd half-Cauchy family of distributions: properties and applications. Commun. Stat. Theory Methods **46**(11), 5685–5705 (2017)
33. E. Paradis, S.R. Baillie, W.J. Sutherland, Modeling large-scale dispersal distances. Ecol. Model. **151**(2–3), 279–292 (2002)
34. B. Rooks, A. Schumacher, K. Cooray, The power Cauchy distribution: derivation, description, and composite models, in *NSF-REU Program Reports* (2010)
35. G. Venter, Transformed beta and gamma distributions and aggregate losses. Proc. Casualty Act. Soc. 156–193 (1983)
36. M.H. Tahir, M. Zubair, G.M. Cordeiro, A. Alzaatreh, M. Mansoor, The Poisson-X family of distributions. J. Stat. Comput. Simul. **86**(14), 2901–2921 (2016)

Chapter 27
3-Parameter Weibull Distribution: Ordinary Differential Equations

Hilary I. Okagbue, Muminu O. Adamu, Abiodun A. Opanuga, Jimevwo G. Oghonyon and Patience I. Adamu

Abstract In this chapter, homogenous ordinary differential equations (ODE) of different orders were obtained for the probability density function, quantile function, survival function inverse survival function, hazard function and reversed hazard functions of 3-parameter Weibull distribution. This is possible since the aforementioned probability functions are differentiable. Differentiation and modified product rule were used to obtain the required ordinary differential equations, whose solutions are the respective probability functions. 3-parameter Weibull distribution is an extension of the Weibull distribution with an extra parameter. The different conditions necessary for the existence of the ODEs were obtained and it is almost in consistent with the support that defined the various probability functions considered. The parameters that defined each distribution greatly affect the nature of the ODE obtained. This method provides new ways of classifying and approximating other probability distributions apart from 3-parameter Weibull distribution considered in this chapter.

Keywords Differential calculus · Hazard function · Inverse survival function Quantile function · Reversed hazard function · Survival function · Weibull distribution

H. I. Okagbue (✉) · A. A. Opanuga · J. G. Oghonyon · P. I. Adamu
Department of Mathematics, Covenant University, Ota, Nigeria
e-mail: hilary.okagbue@covenantuniversity.edu.ng

A. A. Opanuga
e-mail: abiodun.opanuga@covenantuniversity.edu.ng

J. G. Oghonyon
e-mail: jimevwo.oghonyon@covenantuniversity.edu.ng

P. I. Adamu
e-mail: patience.adamu@covenantuniversity.edu.ng

M. O. Adamu
Department of Mathematics, University of Lagos, Akoka, Nigeria
e-mail: madamu@unilag.edu.ng

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7_27

27.1 Introduction

Calculus is a very key tool in the determination of mode of a given probability distribution and in estimation of parameters of probability distributions, amongst other uses. The method of maximum likelihood is an example.

Differential equations often arise from the understanding and modeling of real life problems or some observed physical phenomena. Approximations of probability functions are one of the major areas of application of calculus and ordinary differential equations in mathematical statistics. The approximations are helpful in the recovery of the probability functions of complex distributions [1–4].

Apart from mode estimation, parameter estimation and approximation, probability density function (PDF) of probability distributions can be expressed as ODE whose solution is the PDF. Some of which are available. They include: beta distribution [5], Lomax distribution [6], beta prime distribution [7], Laplace distribution [8] and raised cosine distribution [9].

The aim of this research is to develop homogenous ordinary differential equations for the probability density function (PDF), Quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of 3-parameter Weibull distribution. This will also help to provide the answers as to whether there are discrepancies between the support of the distribution and the necessary conditions for the existence of the ODEs. Similar results for other distributions have been proposed, see [10–22] for details.

3-parameter Weibull distribution is a variant of the Weibull distribution and was obtained to improve the flexibility of modeling with Weibull distribution [23]. The distribution has been studied by [24], where they estimated the shape parameter of the distribution. Cran [25] studied extensively the properties of moment estimators of the distribution while [26] proposed the robust estimator for the 3-parameter Weibull distribution. Some other aspects that have been studied includes: conditional expectation [27], parameter estimation under defined censoring [28, 29], censoring sampling [30], posterior analysis and reliability [31, 32], minimum and robust minimum distance estimation [33, 34], three-parameter Weibull equations [35], confidence limits [36], quantile based point estimate of the parameters [37], percentile estimation [38], methods of estimation of parameters [39–43]. Strong computational techniques have now been used in the estimation of parameters of the distribution such as particle swarm optimization [44], differential evolution [45]. Li [46] applied the least square method in the estimation of the parameters of the distribution. Mahmoud [47] observed that the 3-parameter inverse Gaussian distribution can be used and apply as an alternative model for the 3-parameter Weibull distribution. The distribution has been used in the modeling of real life situations such as: fatigue crack growth [48], step-stress accelerated life test [49], ageing [50], helicopter blade reliability [51], cost estimation [52], time between failures of machine tools [53].

27.2 Probability Density Function

The probability density function of the 3-parameter Weibull distribution is given as;

$$f(x) = \frac{\beta}{\eta}\left(\frac{x-\alpha}{\eta}\right)^{\beta-1} \mathrm{e}^{-\left(\frac{x-\alpha}{\eta}\right)^{\beta}} \tag{27.1}$$

with the parameters $\alpha \in \mathbb{R}, \beta, \eta, > 0, x \geq 0$.

To obtain the first order ordinary differential equation for the probability density function of the 3-parameter Weibull distribution, differentiate equation (27.1), to obtain;

$$f'(x) = \left\{\frac{\beta-1}{x-\alpha} - \frac{\beta}{\eta}\left(\frac{x-\alpha}{\eta}\right)^{\beta-1}\right\} f(x) \tag{27.2}$$

The condition necessary for the existence of the equation is $x, \alpha, \beta, \eta > 0$. The differential equations can only be obtained for particular values of α, β and η.

When $\beta = 1$, Eq. (27.2) becomes;

$$f_a'(x) = \left(-\frac{1}{\eta}\right) f_a(x) \tag{27.3}$$

$$\eta f_a'(x) + f_a(x) = 0 \tag{27.4}$$

When $\beta = 2$, Eq. (27.2) becomes;

$$f_b'(x) = \left\{\frac{1}{x-\alpha} - \frac{2(x-\alpha)}{\eta^2}\right\} f_b(x) \tag{27.5}$$

$$\eta^2(x-\alpha) f_b'(x) - (\eta^2 - 2(x-\alpha)^2) f_b(x) = 0 \tag{27.6}$$

This can be obtain for any given parameter of the distribution.

To obtain a second order ODE, Eq. (27.2) is differentiated to obtain;

$$\begin{aligned} f''(x) = {} & \frac{f'^2(x)}{f(x)} - \frac{(\beta-1) f(x)}{(x-\alpha)^2} \\ & - \frac{(\beta-1)^2 f(x)}{(x-\alpha)^2} + \frac{(\beta-1) f'(x)}{x-\alpha} \end{aligned} \tag{27.7}$$

$$f''(x) = \frac{f'^2(x)}{f(x)} - \frac{\beta(\beta-1) f(x)}{(x-\alpha)^2} + \frac{(\beta-1) f'(x)}{x-\alpha} \tag{27.8}$$

The condition necessary for the existence of the equation is $x \geq 0, x - \alpha \neq 0, f(x) > 0, \beta, \eta > 0$.

The second order ordinary differential equation for the probability density function of the 3-parameter Weibull distribution is given by;

$$(x-\alpha)^2 f''(x) - (x-\alpha)^2 f'^2(x) + \beta(\beta-1)f^2(x) - (\beta-1)(x-\alpha)f(x)f'(x) = 0 \tag{27.9}$$

with the initial conditions given as;

$$f(0) = \frac{\beta}{\eta}\left(-\frac{\alpha}{\eta}\right)^{\beta-1} \mathrm{e}^{-\left(-\frac{\alpha}{\eta}\right)^{\beta}} \tag{27.10}$$

$$f'(0) = -\frac{\beta}{\eta}\left(-\frac{\alpha}{\eta}\right)^{\beta-1}\left\{\frac{\beta-1}{\alpha} + \frac{\beta}{\eta}\left(-\frac{\alpha}{\eta}\right)^{\beta-1}\right\}\mathrm{e}^{-\left(-\frac{\alpha}{\eta}\right)^{\beta}} \tag{27.11}$$

27.3 Quantile Function

The Quantile function of the 3-parameter Weibull distribution is given as;

$$Q(p) = \alpha - \eta(-\ln(1-p))^{\frac{1}{\beta}} \tag{27.12}$$

The parameters are: $\beta, \eta > 0, 0 < p < 1$.

To obtain the first order ordinary differential equation for the Quantile function of the 3-parameter Weibull distribution, differentiate equation (27.12), to obtain;

$$Q'(p) = -\frac{\eta}{\beta(1-p)}(-\ln(1-p))^{\frac{1}{\beta}-1} \tag{27.13}$$

The condition necessary for the existence of the equation is; $\beta, \eta > 0, 0 < p < 1$.

Some algebraic simplifications are done to obtain;

$$Q'(p) = -\frac{(\alpha - Q(p))^{1-\beta}\eta^{\beta}}{\beta(1-p)} \tag{27.14}$$

$$Q(0.1) = \alpha - \eta(-\ln(0.9))^{\frac{1}{\beta}} \tag{27.15}$$

The differential equations can only be obtained for particular values of α, β and η.

When $\beta = 1$, Eq. (27.14) becomes;

$$Q'_a(p) = -\frac{\eta}{(1-p)} \tag{27.16}$$

$$(1-p)Q'_a(p) + \eta = 0 \tag{27.17}$$

When $\beta = 2$, Eq. (27.14) becomes;

$$Q'_b(p) = -\frac{\eta^2}{2(1-p)(\alpha - Q_b(p))} \tag{27.18}$$

$$2(1-p)(\alpha - Q_b(p))Q_b'(p) + \eta^2 = 0 \tag{27.19}$$

When $\beta = 3$, Eq. (27.14) becomes;

$$Q_c'(p) = -\frac{\eta^3}{3(1-p)(\alpha - Q_c(p))^2} \tag{27.20}$$

$$3(1-p)(\alpha - Q_c(p))^2 Q_c'(p) + \eta^3 = 0 \tag{27.21}$$

27.4 Survival Function

The survival function of the 3-parameter Weibull distribution is given as;

$$S(t) = \mathrm{e}^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}} \tag{27.22}$$

To obtain the first order ordinary differential equation for the survival function of the 3-parameter Weibull distribution, differentiate equation (27.22), to obtain;

$$S'(t) = -\frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} \mathrm{e}^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}} \tag{27.23}$$

The condition necessary for the existence of the equation is $t \geq 0, \alpha \in \mathbb{R}, \beta, \eta > 0$.

$$S'(t) = -\frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} S(t) \tag{27.24}$$

The differential equations can only be obtained for particular values of α, β and η.

When $\beta = 1$, Eq. (27.24) becomes;

$$S_a'(t) = -\frac{1}{\eta} S_a(t) \tag{27.25}$$

$$\eta S_a'(t) + S_a(t) = 0 \tag{27.26}$$

When $\beta = 2$, Eq. (27.24) becomes;

$$S_b'(t) = -\frac{2}{\eta}\left(\frac{t-\alpha}{\eta}\right) S_b(t) \tag{27.27}$$

$$\eta^2 S_b'(t) + 2(t-\alpha) S_b(t) = 0 \tag{27.28}$$

When $\beta = 3$, Eq. (27.24) becomes;

$$S_c'(t) = -\frac{3}{\eta}\left(\frac{t-\alpha}{\eta}\right)^2 S_c(t) \tag{27.29}$$

$$\eta^3 S_c'(t) + 3(t-\alpha)^2 S_c(t) = 0 \tag{27.30}$$

27.5 Inverse Survival Function

The inverse survival function of the 3-parameter Weibull distribution is given as;

$$Q(p) = \alpha + \eta\left(\ln\frac{1}{p}\right)^{\frac{1}{\beta}} \tag{27.31}$$

To obtain the first order ordinary differential equation for the inverse survival function of the 3-parameter Weibull distribution, differentiate equation (27.31), to obtain;

$$Q'(p) = -\frac{\eta}{\beta p}\left(\ln\frac{1}{p}\right)^{\frac{1}{\beta}-1} \tag{27.32}$$

$$Q'(p) = -\frac{\eta\left(\ln\frac{1}{p}\right)^{\frac{1}{\beta}}}{\beta p\left(\ln\frac{1}{p}\right)} \tag{27.33}$$

The condition necessary for the existence of the equation is $\beta, \eta > 0, 0 < p < 1$. Simplify Eq. (27.33) using Eq. (27.31) to obtain;

$$\beta p Q'(p) + \eta^{\beta}(Q(p)-\alpha)^{1-\beta} = 0 \tag{27.34}$$

with initial value condition given as;

$$Q(0.1) = \alpha + \eta(\ln 10)^{\frac{1}{\beta}} \tag{27.35}$$

The differential equations can only be obtained for particular values of α, β and η.

27.6 Hazard Function

The hazard function of the 3-parameter Weibull distribution is given as;

$$h(t) = \frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} \tag{27.36}$$

To obtain the first order ordinary differential equation for the hazard function of the 3-parameter Weibull distribution, differentiate equation (27.36), to obtain;

$$h'(t) = \frac{\beta(\beta-1)}{\eta^2}\left(\frac{t-\alpha}{\eta}\right)^{\beta-2} \tag{27.37}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \eta > 0$.

$$h'(t) = \frac{(\beta-1)}{\eta}h(t) \tag{27.38}$$

The first order ordinary differential equation for the hazard function of the 3-parameter Weibull distribution is given by;

$$\eta h'(t) - (\beta-1)h(t) = 0 \tag{27.39}$$

with initial value condition given as;

$$h(0) = \frac{\beta}{\eta}\left(-\frac{\alpha}{\eta}\right)^{\beta-1} \tag{27.40}$$

To obtain the second order ordinary differential equation for the hazard function of the 3-parameter Weibull distribution, differentiate equation (27.37);

$$h''(t) = \frac{\beta(\beta-1)(\beta-2)}{\eta^3}\left(\frac{t-\alpha}{\eta}\right)^{\beta-3} \tag{27.41}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \eta > 0$

Two ordinary differential equations can be obtained from the simplification of Eq. (27.41). These are shown as Eqs. (27.42) and (27.43).

$$(t-\alpha)^2 h''(t) - (\beta-1)(\beta-2)h(t) = 0 \tag{27.42}$$

$$(t-\alpha)^2 h''(t) - (\beta-2)h'(t) = 0 \tag{27.43}$$

To obtain the third order ordinary differential equation for the hazard function of the 3-parameter Weibull distribution, differentiate equation (27.41);

$$h'''(t) = \frac{\beta(\beta-1)(\beta-2)(\beta-3)}{\eta^4}\left(\frac{t-\alpha}{\eta}\right)^{\beta-4} \tag{27.44}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \eta > 0$.

Three ordinary differential equations can be obtained from the simplification of Eq. (27.44). These are shown as Eqs. (27.45)–(27.47).

$$(t-\alpha)^3 h'''(t) - (\beta-1)(\beta-2)(\beta-3)h(t) = 0 \tag{27.45}$$

$$(t-\alpha)^2 h'''(t) - (\beta-2)(\beta-3)h'(t) = 0 \tag{27.46}$$

$$(t-\alpha)h'''(t) - (\beta-3)h''(t) = 0 \tag{27.47}$$

With additional initial values;

$$h'(0) = \frac{\beta(\beta-1)}{\eta^2}\left(-\frac{\alpha}{\eta}\right)^{\beta-2} \tag{27.48}$$

$$h''(0) = \frac{\beta(\beta-1)(\beta-2)}{\eta^3}\left(-\frac{\alpha}{\eta}\right)^{\beta-3} \tag{27.49}$$

27.7 Reversed Hazard Function

The reversed hazard function of the 3-parameter Weibull distribution is given as;

$$j(t) = \frac{\frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}}}{1 - e^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}}} \tag{27.50}$$

To obtain the first order ordinary differential equation for the reversed hazard function of the 3-parameter Weibull distribution, differentiate equation (27.50), to obtain;

$$j'(t) = \left\{\frac{\beta-1}{t-\alpha} - \frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right)^{\beta-1} - j(t)\right\} j(t) \tag{27.51}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \eta > 0$.

The differential equations can only be obtained for particular values of α, β and η.

When $\beta = 1$, Eq. (27.51) becomes;

$$j_a'(t) = \left(-\frac{1}{\eta} - j_a(t)\right) j_a(t) \tag{27.52}$$

$$\eta j_a'(t) + j_a(t) + \eta j_a^2(t) = 0 \tag{27.53}$$

When $\beta = 2$, Eq. (27.52) becomes;

$$j_b'(t) = \left\{\frac{1}{t-\alpha} - \frac{\beta}{\eta}\left(\frac{t-\alpha}{\eta}\right) - j_b(t)\right\} j_b(t) \tag{27.54}$$

$$\eta^2(t-\alpha)j_b'(t) + (\beta(t-\alpha)^2 - \eta^2)j_b(t) + \eta^2(t-\alpha)j_b^2(t) = 0 \tag{27.55}$$

Second order ODE can be obtained from the differentiation of Eq. (27.51) is differentiated to obtain;

$$j''(t) = \left\{ \frac{\beta - 1}{t - \alpha} - \frac{\beta}{\eta} \left(\frac{t - \alpha}{\eta} \right)^{\beta - 1} - j(t) \right\} j'(t) - \left\{ \frac{\beta - 1}{(t - \alpha)^2} - \frac{\beta(\beta - 1)}{\eta^2} \left(\frac{t - \alpha}{\eta} \right)^{\beta - 2} + j'(t) \right\} j(t) \tag{27.56}$$

The condition necessary for the existence of the equation is $t, \alpha, \beta, \eta > 0$.
Further simplifications are done to obtain;

$$j''(t) = \frac{j'^2(t)}{j(t)} + j(t)j'(t) - \frac{\beta(\beta - 1)j(t)}{(t - \alpha)^2} + \frac{(\beta - 1)j'(t)}{t - \alpha} + \frac{(\beta - 1)j^2(t)}{t - \alpha} \tag{27.57}$$

The ODEs can be obtained for the particular values of the distribution. Several analytic, semi-analytic and numerical methods can be applied to obtain the solutions of the respective differential equations. Also comparison with two or more solution methods is useful in understanding the link between ODEs and the probability distributions.

27.8 Conclusion

Ordinary differential equations (ODEs) has been obtained for the probability density function (PDF), quantile function (QF), survival function (SF), inverse survival function (ISF), hazard function (HF) and reversed hazard function (RHF) of the 3-parameter Weibull distribution.

This differential calculus, modified product rule and efficient algebraic simplifications were used to derive the various classes of the ODEs. The parameter and the supports that characterize the 3-parameter Weibull distribution determine the nature, existence, orientation and uniqueness of the various ODE. The results are in agreement with those available in scientific literature. Furthermore several methods can be used to obtain desirable solutions to the ODEs. This method of characterizing distributions cannot be applied to distributions whose PDF or CDF are either not differentiable or the domain of the support of the distribution contains singular points.

Acknowledgements This work was supported by Covenant University, Nigeria.

References

1. G. Steinbrecher, W.T. Shaw, Quantile mechanics. Euro. J. Appl. Math. **19**(2), 87–112 (2008)
2. H.I. Okagbue, M.O. Adamu, T.A. Anake, Quantile approximation of the Chi-square distribution using the quantile mechanics, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 477–483
3. H.I. Okagbue, M.O. Adamu, T.A. Anake, Solutions of Chi-square quantile differential equation, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 813–818
4. Y. Kabalci, On the Nakagami-m inverse cumulative distribution function: closed-form expression and its optimization by backtracking search optimization algorithm. Wirel. Pers. Commun. **91**(1), 1–8 (2016)
5. W.P. Elderton, *Frequency Curves and Correlation* (Charles and Edwin Layton, London, 1906)
6. N. Balakrishnan, C.D. Lai, *Continuous Bivariate Distributions*, 2nd edn. (Springer, New York, London, 2009)
7. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2, 2nd edn. (Wiley, 1995)
8. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994). ISBN: 0-471-58495-9
9. H. Rinne, *Location Scale Distributions, Linear Estimation and Probability Plotting Using MATLAB* (2010)
10. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of Fréchet distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 186–191
11. H.I. Okagbue, P.E. Oguntunde, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated generalized exponential distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 192–197
12. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Cauchy, standard Cauchy and log-Cauchy distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 198–204
13. H.I. Okagbue, S.A. Bishop, A.A. Opanuga, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of Burr XII and Pareto distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 399–404
14. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of Gompertz and Gamma Gompertz distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 405–411
15. H.I. Okagbue, M.O. Adamu, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of 3-parameter Weibull distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 539–545

16. H.I. Okagbue, A.A. Opanuga, E.A. Owoloko, M.O. Adamu, Classes of ordinary differential equations obtained for the probability functions of exponentiated Fréchet distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 546–551
17. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, S.A. Bishop, Classes of ordinary differential equations obtained for the probability functions of half-Cauchy and power Cauchy distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 552–558
18. H.I. Okagbue, P.E. Oguntunde, A.A. Opanuga, E.A. Owoloko, Classes of ordinary differential equations obtained for the probability functions of exponential and truncated exponential distributions, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 858–864
19. H.I. Okagbue, O.O. Agboola, P.O. Ugwoke, A.A. Opanuga, Classes of ordinary differential equations obtained for the probability functions of exponentiated Pareto distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 865–870
20. H.I. Okagbue, O.O. Agboola, A.A. Opanuga, J.G. Oghonyon, Classes of ordinary differential equations obtained for the probability functions of Gumbel distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 871–875
21. H.I. Okagbue, O.A. Odetunmibi, A.A. Opanuga, P.E. Oguntunde, Classes of ordinary differential equations obtained for the probability functions of half-normal distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 876–882
22. H.I. Okagbue, M.O. Adamu, E.A. Owoloko, E.A. Suleiman, Classes of ordinary differential equations obtained for the probability functions of Harris extended exponential distribution, in *Proceedings of the World Congress on Engineering and Computer Science 2017*, 25–27 October 2017, San Francisco, U.S.A. Lecture Notes in Engineering and Computer Science (2017), pp. 883–888
23. A.M. Razali, A.A. Salih, A.A. Mahdi, Estimation accuracy of Weibull distribution parameters. J. Appl. Sci. Res. **5**(7), 790–795 (2009)
24. M. Teimouri, A.K. Gupta, On the three-parameter Weibull distribution shape parameter estimation. J. Data Sci. **11**, 403–414 (2013)
25. G.W. Cran, Moment estimators for the 3-parameter Weibull distribution. IEEE Trans. Reliab. **37**, 360–363 (1988)
26. A. Adatia, L.K. Chan, Robust estimators of the 3-parameter Weibull distribution. IEEE Trans. Reliab. **34**(4), 347–351 (1985)
27. D. Kundu, M.Z. Raqab, Estimation of $R = P(Y < X)$ for three-parameter Weibull distribution. Stat. Prob. Lett. **79**, 1839–1846 (2009)
28. G.H. Lemon, Maximum likelihood estimation for the three parameter Weibull distribution based on censored samples. Technometrics **17**(2), 247–254 (1975)
29. M. Sirvanci, G. Yang, Estimation of the Weibull parameters under type I censoring. J. Am. Stat. Assoc. **79**, 183–187 (1984)
30. J. Wyckoff, L.J. Bain, M. Engelhardt, Some complete and censored sampling results for the three-parameter Weibull distribution. J. Stat. Comput. Simul. **11**(2), 139–151 (1980)
31. S.K. Sinha, J.A. Sloan, Bayes estimation of the parameters and reliability function of the 3-parameter Weibull distribution. IEEE Trans. Reliab. **37**(4), 364–369 (1988)
32. E.G. Tsionas, Posterior analysis, prediction and reliability in three-parameter Weibull distributions. Commun. Stat. Theory Methods **9**, 1435–1449 (2000)

33. J.R. Hobbs, A.H. Moore, R.M. Miller, Minimum-distance estimation of the parameters of the 3-parameter Weibull distribution. IEEE Trans. Reliab. **34**(5), 495–496 (1985)
34. M.A. Gallagher, A.H. Moore, Robust minimum-distance estimation using the 3-parameter Weibull distribution. IEEE Trans. Reliab. **39**(5), 575–580 (1990)
35. D.R. Wingo, Solution of the three-parameter Weibull equations by constrained modified quasi linearization (progressively censored samples). IEEE Trans. Reliab. **22**(2), 96–102 (1973)
36. H. Hirose, Maximum likelihood estimation in the 3-parameter Weibull distribution: a look through the generalized extreme-value distribution. IEEE Trans. Dielectr. Electr. Insul. **3**(1), 43–55 (1996)
37. D.M. Brkic, Point Estimation of the 3-Weibull parameters based on the appropriated values of the quantiles. Elektrotehnika Zagreb **28**(6), 335–342 (1985)
38. U. Schmid, Percentile estimators for the three-parameter Weibull distribution for use when all parameters are unknown. Commun. Stat. Theory Methods **26**(3), 765–785 (1997)
39. J.I. McCool, Inference on the Weibull location parameter. J. Qual. Tech. **30**(2), 119–126 (1998)
40. V.G. Panchang, R.C. Gupta, On the determination of three-parameter Weibull mle's. Commun. Stat. Simul. Comput. **18**(3), 1037–1057 (1989)
41. E.E. Afify, A method for estimating the 3-parameter of the Weibull distribution. Alex. Eng. J. **39**(6), 973–976 (2000)
42. D. Cousineau, Fitting the three-parameter weibull distribution: review and evaluation of existing and new methods. IEEE Trans Dielectr. Electr. Insul. **16**(1), 281–288 (2009)
43. G. Tzavelas, A study of the number of solutions of the system of the log-likelihood equations for the 3-parameter Weibull distribution. Appl. Math. **57**(5), 531–542 (2012)
44. H.H. Örkcü, V.S. Özsoy, E. Aksoy, Estimating the parameters of 3-p Weibull distribution using particle swarm optimization: a comprehensive experimental comparison. Appl. Math. Comput. **268**, 201–226 (2015)
45. H.H. Örkcü, E. Aksoy, M.I. Doıan, Estimating the parameters of 3-p Weibull distribution through differential evolution. Appl. Math. Comput. **251**, 211–224 (2015)
46. Y.M. Li, A General linear-regression analysis applied to the 3-parameter Weibull distribution. IEEE Trans. Reliab. **43**(2), 255–263 (1994)
47. M. Mahmoud, Bayesian estimation of the 3-parameter inverse Gaussian distribution. Trabajos de Estadistica **6**(1), 45–62 (1991)
48. H. Itagaki, T. Ishizuka, P.Y. Huang, Experimental estimation of the probability distribution of fatigue crack growth lives. Prob. Eng. Mech. **8**(1), 25–34 (1993)
49. L.C. Tang, Y.S. Sun, T.N. Goh, H.L. Ong, Analysis of step-stress accelerated-life-test data: a new approach. EEE Trans. Reliab. **45**(1), 69–74 (1996)
50. P. Praks, H.F. Bacarizo, P.E. Labeau, On the modeling of ageing using Weibull models: case studies. safety, reliability and risk analysis: theory, methods and applications, in *Proceedings of the Joint ESREL and SRA-Europe Conference*, vol. 1 (2009), pp. 559–565
51. A.R. Shahani, M. Babaei, Helicopter blade reliability: statistical data analysis and modeling. Aerosp. Sci. Technol. **55**(1), 43–48 (2016)
52. J. Zhao, G. Peng, H. Zhang, Schedule and cost integrated estimation for complex product modeling based on Weibull distribution, in *Proceedings of the IEEE 19th International Conference on Computer Supported Cooperative Work in Design, CSCWD*, Article number 7230971 (2015), pp. 276–280
53. X.Y. Xue, W. Xu, J.H. Li, Reliability modeling on time between failures of NC machine tools. Adv. Mater. Res. **145–150**, 2014 (1028)

Index

© Springer Nature Singapore Pte Ltd. 2019

S.-I. Ao et al. (eds.), *Transactions on Engineering Technologies*,
https://doi.org/10.1007/978-981-13-2191-7