Progress in IS

Volker Wohlgemuth
Stefan Naumann
Grit Behrens
Hans-Knud Arndt *Editors*

Advances and New Trends in Environmental Informatics

A Bogeyman or Saviour for the UN Sustainability Goals?

Springer

Progress in IS

"PROGRESS in IS" encompasses the various areas of Information Systems in theory and practice, presenting cutting-edge advances in the field. It is aimed especially at researchers, doctoral students, and advanced practitioners. The series features both research monographs that make substantial contributions to our state of knowledge and handbooks and other edited volumes, in which a team of experts is organized by one or more leading authorities to write individual chapters on various aspects of the topic. "PROGRESS in IS" is edited by a global team of leading IS experts. The editorial board expressly welcomes new members to this group. Individual volumes in this series are supported by a minimum of two members of the editorial board, and a code of conduct mandatory for all members of the board ensures the quality and cutting-edge nature of the titles published under this series.

More information about this series at https://link.springer.com/bookseries/10440

Volker Wohlgemuth · Stefan Naumann ·
Grit Behrens · Hans-Knud Arndt
Editors

Advances and New Trends in Environmental Informatics

A Bogeyman or Saviour for the UN Sustainability Goals?

Springer

Editors
Volker Wohlgemuth
HTW Berlin - University of Applied Sciences
Berlin, Germany

Stefan Naumann
Environmental Campus Birkenfeld
Trier University of Applied Sciences
Birkenfeld, Germany

Grit Behrens
FB Campus Minden
University of Applied Sciences Bielefeld
Minden, Germany

Hans-Knud Arndt
Faculty of Computer Science
Otto-von-Guericke University Magdeburg
Magdeburg, Germany

ISSN 2196-8705 ISSN 2196-8713 (electronic)
Progress in IS
ISBN 978-3-030-88062-0 ISBN 978-3-030-88063-7 (eBook)
https://doi.org/10.1007/978-3-030-88063-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book presents the main research results of the 35th edition of the long-standing and established international and interdisciplinary conference series on environmental information and communication technologies (EnviroInfo 2021).

The conference was held from 27 to 29 of September 2021. It was organized by the Technical Committee on Environmental Informatics of the Gesellschaft für Informatik e.V. (German Informatics Society—GI).

This book presents a selection of peer-reviewed research papers that describe innovative scientific approaches and ongoing research in environmental informatics and the emerging field of environmental sustainability. Combining and shaping national and international activities in the field of applied informatics and Environmental Informatics, the EnviroInfo conference series aims at presenting and discussing the latest state-of-the-art development on Information and Communication Technology (ICT) and environmental-related fields. A special focus of the conference was on the question whether Environmental Informatics is a bogeyman or a savior to achieve the UN Sustainable Development Goals.

The respective articles cover a broad range of scientific aspects including advances in core environmental informatics-related technologies, such as Sustainable Mobility Digital Sharing Economy and Sustainability, Sustainable Usability and User Experience, Earth System Observation and Computational Analysis for Sustainable Development, Modeling and Simulation in the Environmental and Earth Sciences Artificial Intelligence and Sustainability, Environmental Health Informatics, and other relevant topics in the field of Environmental Informatics.

We would like to thank all contributors for their submissions. Special thanks also go to the members of the program and organizing committees, for reviewing all submissions. In particular, we like to thank our cooperation partner German International University in Berlin and the team around Mr. Daniel Krupka from the German Informatics Society for the local organizing support.

Last but not least, a warm thank you to our sponsors that supported the conference.

Finally, we wish to thank Mrs. Barbara Bethke and Mr. Christian Rauscher from Springer and the entire Springer production team for their assistance and guidance in successfully producing this book.

Berlin, Germany Volker Wohlgemuth
Magdeburg, Germany Hans-Knud Arndt
Bielefeld, Germany Grit Behrens
Birkenfeld, Germany Stefan Naumann

Contents

Modelling and Simulation of Systems

Urban Environments and Information Systems

Sustainable Mobility

Sustainable Software Engineering

An Approach to Describe Design Pattern Structures for Sustainable Software Products

Benno Schmidt

Abstract Software affects society and environment in a variety of ways. In addition to the energy and resource requirements driven by software products, social factors are to be mentioned as sustainability-relevant aspects. In order to support software developers in creating products that can be positively assessed from a sustainability point of view, the question of practically usable assistance arises in this context. With this aim, in this paper the methodological approach of design patterns is examined. It is discussed how sustainability patterns can be described in a structured way and how they can be incorporated into practical software engineering processes. Particular attention is paid to aspects of completeness and consistency of the descriptions as well as the explicit inclusion of human action. It is recommended to provide pattern descriptions with a high formalization degree in order to enable verifiability for logical plausibility and completeness and to allow further machine-assisted processing. Prospectively appropriate pattern repositories could be set up to support practical software product development processes in the future.

Keywords Sustainable software products · Green software · Design patterns

1 Motivation

Our digital lifestyle is accompanied with significant positive and negative ecological and social effects [5, 9, 10]. The discussion about the use of information and communications technology (ICT) often focuses on hardware-related aspects such as energy and resource requirements with view to the manufacture and operation or the disposal of devices. In addition to ecological effects, the social dimension must also be considered, e.g. with respect to fair working conditions, questions of equal access to information, the right to informational self-determination, the creation of sustainability awareness, etc.

B. Schmidt (✉)
Bochum University of Applied Sciences, Am Hochschulcampus 1, 44801 Bochum, Germany
e-mail: benno.schmidt@hs-bochum.de

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_1

Software is required to run computer systems. And in the end it is software that significantly influences the energy requirements of the hardware, that establishes (possibly undesired) data connections to other hardware and software units, or that causes premature obsolescence of functional devices [11, 12]. Software therefore has a decisive influence on sustainability-related topics. (For a comprehensive concept of sustainability that goes beyond the simple Brundtland definition, see [4, 14, 23], for the concept of "sustainable software" see [12]). In fact, the products resulting from software development processes often can be assessed negatively from the sustainability perspective. From the point of view of those involved in the development process, this may sometimes happen unconsciously. (The aim here is to sensitize product managers, developers, clients, customers, and users to this issue!) However, often there is a lack of supportive approaches for product manufacturers and software developers.

One approach to deal this problem is to collect and provide appropriate practical options for action. First such recommendations and guidelines are given, e.g. by the industry association Bitkom [9] or by the German Federal Environment Agency with its specific criteria for awarding the environmental label "Blue Angel for resource- and energy-efficient software products" [16, 26]. The implementation into software engineering processes, for example in the context of system design or implementation activities and as concrete support for software architects and designers as well as programmers, is currently still difficult in practice.

However, another methodical approach, which is often used in the field of software engineering, is the use of so-called design patterns. The question arises as to whether special design patterns can be identified for the development of sustainable software products. Thus, the present paper discusses how sustainability patterns can be described in a structured way and how they can be incorporated into practical software engineering processes.

This paper is structured as follows: First, Chap. 2 briefly describes the concept of design patterns and pattern languages. Then a simple example is given in Chap. 3. Chapter 4 deals with the formal description of patterns. Chapter 5 focuses on the involvement of human actors and the proof of actual practical functioning into the descriptions. Finally, in Chap. 6, first experiences on the practical construction of a pattern collection are given.

2 Design Patterns for Sustainable Software

2.1 Software Design Process

The term "software design" is interpreted quite broadly in this paper. In a very general view, software design refers to those process activities that are carried out between the formulation of the product requirements and the subsequent implementation of the product. The focus of interest is the identification and description of basic software abstractions and their relationships [24] with the aim of implementing an executable software system that meets the product requirements and also the demand for sustainable development here. The activities in the software design process affect a variety of aspects, in particular the software-architectural design, the development of communication interfaces (e.g. for web services), the design of system components, the conception of algorithms and data structures, the design of user interfaces, user experience aspects (UX) and much more.

Within the scope of this paper, human actors who are involved in the production, operation and use of software as well as in the achievement of sustainability goals (or who are affected by the effects of software use) are also explicitly included in the consideration. In this respect, it is less a matter of developing a new notion or a new concept of the term sustainability for the field of software design here; cf. [19]. Rather, it is about the application of the already existing, general sustainability concepts to the socio-technical system *human—environment and society—software* (in accordance with [12]). In a further perspective, concepts and ideas from the areas of "product design", "industrial design" and "design research" can be methodologically included in the consideration, e.g. the theory of product language (artifacts as communication media), the study of interactions between technology and design or an understanding of design activities as forward-looking actions [2, 25].

2.2 Design Pattern Approach

In the field of software engineering, a well-known approach to describe established solutions and to support the selection of alternative implementation options is given by so-called "patterns" as proven, reusable solutions for recurring design problems [3]. Prominent examples are the object-oriented design patterns from Gamma et al. [8] as well as pattern collections for numerous other areas of software technology. Accordingly, the question arises whether such a pattern collection for the development of sustainable software products can be set up in order to support the respective actors in their actions and decisions and to further promote the discussion of alternative design and implementation options.

2.3 Characteristic Properties

It should be noted at this point that not all portions of design knowledge give patterns. Indeed, a distinction should be made between guidelines, standards, heuristics, requirements, algorithm efficiency, etc. [3]. In the scope of the present paper, the following aspects are considered relevant:

- A statement of the addressed problem and the corresponding design solution are given explicitly.

- A pattern contains an explanation of why the corresponding solution is proposed ("rationale").
- Here, a pattern is anchored in the software engineering domain as well as in the domain of sustainability science. I.e., the pattern does not only relate to software engineering aspects, but also explicitly addresses sustainability-related issues.
- The degree of abstraction of the proposed solution relates to decomposable systems. I.e., a pattern description includes at least two interacting software artifacts ("participants" and "collaborations" in [8]) as well as other elements that relate to the sustainability goals.
- A pattern should show how a positive contribution is actually made to sustainability issues (e.g. by naming verifiable qualitative measurands) and under what conditions. (Which actors actually have to become active so that practical effectiveness is ensured?).

2.4 Pattern Description Languages

Languages for describing such patterns include possibilities for describing usage contexts and relationships between patterns [1]. In addition to the designation of a pattern, essential descriptive elements include the following: An evaluation with regard to the degree of maturity of the pattern, the actual description of the problem, the specification of "forces" which indicate relevant design aspects in the sense of possibly competing design goals, the description of the solution as well as a summarizing diagram of the solution idea. (It should be noted that patterns often are described using mostly similar or additional elements.)

One of several ways to file such pattern descriptions and the relationships between the patterns is offered by the XML-based "Pattern Language Markup Language", or PLML for short (pronounced "pell-mell"; [3, 21]). The corresponding schema specification supports users in the documentation of patterns by enabling automated checks with regard to the correct use of the specified structural elements. The description of the functionality of a pattern is initially largely informal, so that, with a view to the problem under consideration in this paper, further description elements need to be added, or descriptions given in other description languages need to be incorporated subsequently (see Chap. 4 of this article).

3 A Simple Example

Let us consider an example problem that is deliberately kept very simple: An existing application *app* runs in an execution environment *e*. Now the execution environment *e* is replaced by a new environment *e'* (e.g. due to an update of the operating system, etc.). However, the application cannot be executed in *e'*, so that the user's hardware threatens to become obsolescent. (In view of the considerations below, object identifiers have already been introduced here.)

A solution could look like this: To operate *app,* another application *em* is provided, which can be run in *e'* and which emulates the old environment. Figure 1 illustrates this simple approach.

The corresponding PLML description has the following structure:

```
<plml>
  <pattern ...>
    <name>Emulator</name>
    ...
    <problem>
      An existing software application can be run inside
      a given execution environment. Now the execution
      environment is replaced by a new one ...
    </problem>
    <solution>
      Provide a further application which can be run
      inside the new execution environment and which ...
    </solution>
    <evidence>
      <rationale>
        The pattern is suggested because ...
      </rationale>
    </evidence>
    ...
  </pattern>
</plml>
```

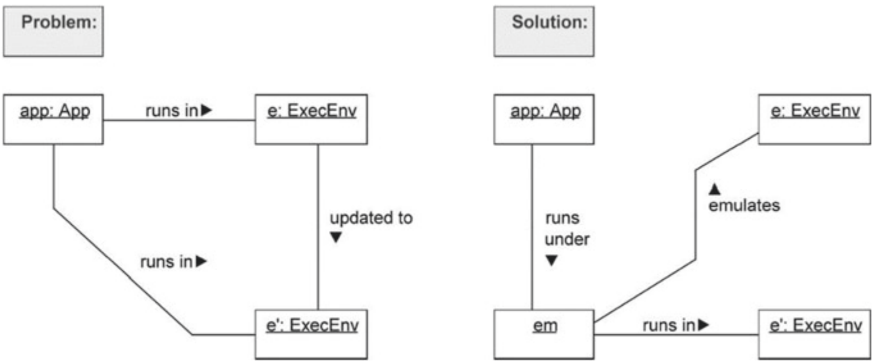


Fig. 1 Graphic illustration of the core of the emulator pattern

The structure will include numerous other elements; see [6, 21] for details.

4 Formal Pattern Descriptions

4.1 Advantages of Formal Descriptions

In the example under consideration, the description of the pattern initially has been given informally. Further structuring and a more formal description would be desirable, since this would help to achieve the following goals:

- Problem and solution are described less ambiguously (avoidance of space for interpretation); the solution is clarified.
- The interacting software-technical units become explicit, whereby the gap between the pattern description and the software development process is getting smaller.
- For descriptions with a high degree of formalization, machine processable “models” can be provided. So it becomes possible to check pattern descriptions for logical completeness and consistency.

Noteworthy, formal description techniques require appropriate skills on the part of the documenter. With the largely informal descriptions considered so far (here based on the PLML) and the possibility of optionally embedding formal descriptions in these, documenters with different prior knowledge can still have the opportunity to actively contribute solutions.

4.2 Example

The emulator example can be described axiomatically: After introducing the predicates “runs under” and “updated to”, we formally can write down in infix notation: For (*app* runs in *e*) $\wedge$ (*e* updated to *e’*), the following situation has to be realized: *app* runs in *e’*. And after the introduction of further predicates: (*em* emulates *e*) $\wedge$ (*em* runs in *e’*) $\wedge$ (*app* runs under *em*) $\rightarrow$ *app* runs $\text{inl}_{\text{em + e}}$ *e’*, whereby *app* runs $\text{inl}_{\text{em + e}}$ *e’* expresses that *app* only virtually runs directly in *e’*, but actually via *em* runs in *e’*.

With the help of suitable specification languages, this can be coded and thus becomes machine processable. For this purpose, transformational-predictive specification languages which allow to formulate corresponding pre- and post-conditions for model states and, if necessary, further “constraints”, can be used. Prominent examples of such languages are the “Object Constraint Language” OCL [27], Z-notation or VDM. Furthermore, algebraic languages such as CASL, OBJ and others or tools known from the field of ontology engineering can also be used. As an example of an algebraic description, an executable coding created using the OBJ successor CafeOBJ [15] is shown below. (The complete code is available via [7]).

```
module CompatibilityProblem {
  [ App ExecEnv ]
  pred _runs under_ : App ExecEnv
  pred _updated to_ : ExecEnv ExecEnv
}

module CompatibilitySolution {
  protected(CompatibilityProblem)
  pred _emulates_ : App ExecEnv
  pred _runs in|(_+_) _ : App App ExecEnv ExecEnv .
  pred _runs in_ : App App

  vars app em : App .
  vars e e' : ExecEnv .
  eq app runs in|(em + e) e' =
    (em emulates e) and (app runs in em) and
    (em runs under e') .
}
```

The simple code, which results directly from the mathematical terms specified above, can be checked immediately for its syntactic correctness, logical conclusiveness and completeness. After setting the appropriate preconditions, the interactive query.

reduce APP *runs in*|(EM + E) E' .

can be answered here with "true" by reducing the entered expression using a term rewriting mechanism [15]. (For a more complex specification example, see the description of a CO_2 emission compensation pattern given in [22]).

4.3 Embedding of Formal Descriptions into the Framework

Here, an element has been added to the XML schema definition of the PLML in order to be able to embed the formal pattern specification in the superordinate description ("formalization") [7]. The objects involved (participants) and required operations (collaborations) of the pattern can also be automatically extracted and documented from the formal description.

```
<plml>
  <pattern ...>
    ...
    <formalization language="CafeOBJ"
          url="../formal/Emulator.mod"/>
      <participant id="app" type="uml$App" .../>
      <participant id="em" type="uml$App" .../>
      <participant id="e" type="uml$ExecEnv" .../>
      ...
      <relation id="emulates" params="em, e" .../>
      <relation id="runs_in" param="app, em" .../>
      ...
    </formalization>
    ...
  </pattern>
</plml>
```

Within the product development process, the specified objects and the relations between them have to be mapped to the resulting implementation by the software developers.

5 Actors and Sustainability Goals

Certainly further additions are required in the description of the considered example. The fact that the user's hardware becomes useless if the application of the emulator runs unbearably slow in the new environment, i.e. the system becomes obsolescent due to the migration from *e* to *e'*, has not yet been described. This could be formulated as follows: perf(*app*, *e'*, *hw*) << perf(*app*, *e*, *hw*) → *hw* gets obsolescent by$|_{e,\, e'}$ *app*, where the function perf() determines the performance of an application when executed in an environment using a specific hardware and << denotes the relation "quantitatively significantly lower than".

Furthermore, it must be ensured that the actors involved in the process described actually act (as previously only implicitly assumed): The user must be willing to continue to use the existing hardware under *e'*. (This means that "psychological obsolescence" does not occur.) In addition, the emulator *em* must actually be made available by the software manufacturer *v*. Otherwise the described pattern will not work. Accordingly, the mathematical description can be supplemented by a further aspect. Overall, it must be verified that the following target condition is met: (*u* aware about *system*) ∧ (*v* supplies *em*) ∧ (*app* runs in$|_{em\,+\,e}$ *e'*) ∧ ¬ (perf(*app*, *e'*, *hw*) << perf(*app*, *e*, *hw*)).

In accordance with the considerations made, in Fig. 2 additional objects and object relationships have been added to the diagram from Fig. 1. For the graphic notation, the UML (Unified Modeling Language) is used. However, other forms of

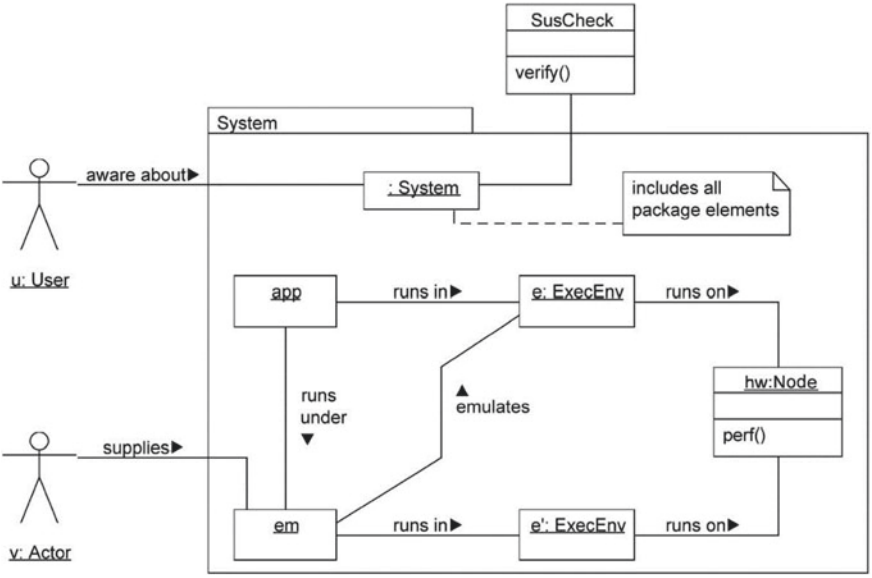


Fig. 2 Supplementation of the emulator pattern with actor objects and hardware relation

illustration also can be used. It should be noted at this point that the terms “execution environment” (ExecEnv) and “node” are part of the UML vocabulary [18].

The relations that are not explicitly included in the diagram can in turn be formally described in a simple manner, so that the now more complex description again can be checked automatically with regard to its correctness, conclusiveness, and completeness. (Ultimately, the logical argumentation is checked. However, the description is not validated against reality at this point.)

6 Practical Construction of Pattern Collections

6.1 Course of the Documentation

The course of the pattern documentation work could practically be structured as follows, whereby steps 2 to 4 initially can be viewed as optional.

1. General description of the pattern (largely informal specification of problem–solution pairs and other elements, e.g. according to the structure specification of the PLML).
2. Contextualization of the described patterns through formal information on pattern relations (here the patterns are related to each other, for example by

giving references to similar patterns, sub-patterns, contradicting patterns, etc.; see also [1, 21]).
3. Specification of the solutions by adding descriptions with a higher degree of formalization.
4. Derivation of further artifacts (e.g. reusable code frameworks, etc.).

If PLML-based descriptions were to be used more widely, support from a comfortable editor would be preferable to working on the XML document. If more extensive pattern collections were to be built within a collaborative working community, tools for discussing the pattern would still be helpful.

6.2 Sustainability Pattern Candidates

In the public GitHub repository [7], prototypically some candidates for sustainable software design patterns are listed. Most of them are given as PLML instances there. Partly, simple formal descriptions have been embedded.

Alongside the "emulator" pattern which has been taken as an example for this paper, the repository includes candidates such as "Compensate", "Backward compatibility", "Buffered dataflow", "Postpone into green phase", "Switch off unneeded processes" or "Component replaceability", just to mention a few. Regarding this, the reader is encouraged to join our "pattern mining" activities and to contribute and discuss further pattern candidates.

Many of the collected candidates ultimately are pure efficiency patterns and thus provide a "weak" sustainability quality only (referring to the terms "pseudo sustainability" and "weak sustainability" as defined in [14]). However, in case they directly relate to issues of energy and resource conservation, they have been included in this candidate collection. In this context, particularly reactive design patterns [9, 13] are worth being highlighted and need to be examined further.

Model candidates that lead to something that might be called "strong" or even "ultra" sustainability [14], should be mentioned explicitly. Here, candidates to be discussed might be "Communicate resource consumption", "Component leaflets", "Open source", "Incentive", or "Sustainability awareness increase" [7, 20]. In a special way, these candidates address the social dimension of sustainability. Maintaining natural resources for future generations, distributive effects between heterogeneous user groups and user cultures can potentially be achieved here. These approaches (which possibly are still too general to be called "patterns") can cause structural and sustainable changes to take place in software production processes.

6.3 Anti-patterns

The problems considered in this article might be viewed in a "complementary" way: Patterns indicate good and proven solutions. Often, however, recurring errors might also lead to unsustainable products. In this respect, it can be worthwhile to describe "anti-patterns" in addition to patterns in the future.

At this point just a few obvious anti-patterns and "bad habits" are mentioned as examples: Usage of unnecessarily complex libraries and frameworks by programmers and the prominent "bloatware" anti-pattern; missing algorithm information for offered API methods (this might result in inefficient and/or resource-consuming software products, e.g. "O(n^2)-solutions", see [7] for a discussion); unnecessary exploitation of performance gained by new computer hardware (which might lead to hardware obsolescence effects); intensive/careless logging (which can result in considerable energy demands, e.g. for intensively used microservices); etc.

6.4 Sustainability of Classic Design Patterns

The present paper primarily dealt with pattern candidates that specifically contribute to increasing the sustainability of software products. Conversely, however, it can also be examined how the use of conventional (initially not sustainability-relevant) design patterns might affect the sustainability of a product. Examples of this can be found in the literature. E.g., the energy footprint resulting from the application of the common patterns given by Gamma et al. [8] such as "Observer", "Decorator", "Factory", etc., is empirically investigated.

In this context, it is quite conceivable that automatic code transformations between patterns can lead to significant reductions in energy consumption [17]. Here it might be helpful to identify design patterns built into software products to identify further sustainability-relevant software construction elements with the help of suitable tools, cf. [17, 19].

6.5 Potential Difficulties in Using Patterns

The abstract description of a pattern does not guarantee that the problem-solving potential will be understood and recognized by the software developer or product designer. The question arises whether only experienced developers and designers can apply patterns effectively [24]. Therefore it is useful to give examples of the application of a pattern that show how the described approach can be incorporated into software designs in practice. (For the specification of examples, explicit description elements are provided in the PLML).

Moreover, it is important to provide a vocabulary for formulating the patterns that is understood by all those involved into the software design process (see also comments in Chap. 7 of this article).

Since not all design knowledge can be mapped onto patterns, it makes sense to provide additional design knowledge in different forms, e.g. in the form of recommendations, guidelines, checklists, tips etc.

Today, software development takes place in a highly complex technical environment, whereby often many further design decisions have to be made for the practical implementation of a pattern. The success of a pattern implementation might depend on whether these decisions are made correctly [19].

Regardless of this, design patterns often form a solid basis for discussing relevant design decisions and making proven design knowledge available to other software developers. Today it is impossible to imagine practical software development without them.

Finally, one more aspect should be mentioned: Pattern collections and other guidelines would possibly be a good instrument to sensitize actors in the field of software development to aspects of sustainable development and to provide inspiring suggestions for innovative product designs.

7 Summary and Outlook

With the help of structured templates (e.g. XML-based), patterns can be easily documented on the basis of mainly informal descriptions. However, there is still a need for descriptions that include, among other things, more precise documentation of the components and interactions involved in patterns. This can be achieved through formal descriptions, which also have the advantage of being machine processable.

The question arises to what extent the elements of the vocabulary used in the descriptions can be supplemented with more concrete “semantics”. For the elements used in the software engineering world, the Universal Modeling Language (UML) provides a largely clearly defined vocabulary [18]. However, it can be asked to what extent and how a comprehensive and internally consistent vocabulary can be set up for the largely interdisciplinary work area considered here, and which description elements are missing from a sustainability-scientific point of view. In the future, it would be desirable to involve sustainability experts in software development processes actively.

A rudimentary approach consists of adding relevant aspects to the pattern descriptions, e.g. in the form of actors through which required objects or operations are initiated, as well as in the formulation of post-conditions referring to verifiable target criteria. The basic technical feasibility has been demonstrated using a simple example here. However, there are further questions to be investigated. In addition to the aforementioned aspects, the methodological gap in the course of mapping patterns to specific implementations should be mentioned here explicitly.

A simple prototype collection of pattern candidates for the development of sustainable software products is available under [7]. It would be desirable to set up a more extensive, consolidated collection in the future. Active "pattern mining" is required here. Although some patterns are already named in the literature [9, 20, 22], these are not described in much detail. The task now is to identify and document further patterns and to provide other targeted innovative solutions to support the development of sustainable software products.

Acknowledgements Particularly the author thanks Petra Schweizer-Ries from Bochum University of Applied Sciences for her support during the past years.

References

1. Borchers, J.O.: A pattern approach to interaction design. In: Proceedings DIS 2000 International Conference on Designing Interactive Systems, Aug. 16–19, 2000. New York. ACM Press, New York (2000)
2. Bürdek, B.E.: Design: History, Theory and Practice of Product Design, 3rd edn. Birkhäuser, Basel (2015)
3. Dearden, A., Finlay, J.: Pattern languages in HCI: a critical review. Human-Comput. Interact. **21**, 49–102 (2006)
4. Dosch, K., Sachsen, K., et al.: Lexikon der Nachhaltigkeit. Aachener Stiftung Kathy Beys (2015). https://www.nachhaltigkeit.info. Accessed 10 May 2021
5. Duboc, L., Penzenstadler, B., Porras, J., Kocak, S.A., Betz, S., Chitchyan, R., Leifler, O., Seyff, N., Venters, C.C.: Requirements engineering for sustainability: an awareness framework for designing software systems for a better tomorrow. Requirements Eng. **25**, 469–492 (2020). https://doi.org/10.1007/s00766-020-00336-y
6. Fincher, S.: PLML: Pattern Language Markup Language. University of Kent, Canterbury, UK (2006). https://www.cs.kent.ac.uk/people/staff/saf/patterns/plml.html. Accessed 10 May 2021
7. Formal Descriptions of Sustainable Software Design Patterns. https://gitlab.com/bogeo/suspatterns, GitLab-Repository, Bochum University of Applied Sciences, Geovisualization Lab. Accessed 10 May 2021
8. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design patterns: elements of reusable object-oriented software. Addison-Wesley, Reading, MA (1994)
9. Geiger, L., Hopf, T., Loring, J., Renner, M., Rudolph, J., Scharf, A., Schmidt, M., Termer, F.: Ressourceneffiziente Programmierung. Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (Bitkom), Berlin (2020). https://www.bitkom.org/sites/default/files/2021-03/210329_lf_resouceneffiziente-programmierung.pdf. Accessed 10 May 2021
10. Gröger, J.: Digitaler ökologischer Fußabdruck. Wie nachhaltig ist unser digitaler Lebensstil? FIfF Kommunikation, **37**(3), Bremen, S. 22–24 (2020)
11. Hilty, L.: Software und Nachhaltigkeit. Wie Fremdbestimmung durch Software materielle Ressourcen entwertet. FIfF Kommunikation, **37**(3), Bremen, S. 31–35 (2020)
12. Kern, E., Guldner, A., Naumann, S.: Including software aspects in green IT: how to create awareness for green software issues. In: Kharchenko, V., Kondratenko, Y., Kacprzyk, J., (eds.) Green IT Engineering Social, Business and Industrial Applications, pp. 3–20. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-00253-4_1

13. Kuhn, R.: Reactive Design Patterns. Manning, Shelter Island (NY) (2017)
14. Martens J., Schilder, K.: Sustainable development. In: Krieger, J. (ed.) The Oxford Companion to Comparative Politics, 2nd edn. Oxford, pp. 813–815 (2013)
15. Nakagawa, A.T., Futatsugi, K., Preining, N.: CafeOBJ User's Manual. Ver. 1.5.6 (2016). http://www.cafeobj.org/documentation. Accessed 10 May 2021
16. Naumann, S., Guldner, A., Kern, E.: The eco-label blue angel for software-development and components. In: Kamilaris, A., Wohlgemuth, V., Karazas, K., Athanasiadis, I.N. (eds.) Advances and New Trends in Environmental Informatics, pp. 79–89. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-61969-5_6
17. Noureddine, A., Rajan, A.: Optimising energy consumption of design patterns. In: IEEE/ ACM 37th IEEE International Conference on Software Engineering (ICSE), pp. 623–626 (2015). https://doi.org/10.1109/ICSE.2015.208
18. Object Management Group. (eds.): OMG Unified Modeling Language, Ver. 2.5.1 (2017). https://www.omg./org/spec/UML. Accessed 10 May 2021
19. Robillard, M.O.: Sustainable software design. In: Proceedings of the 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE 2016), pp. 920–923 (2016). https://doi.org/10.1145/2950290.2983983
20. Roher, K., Richardson, D.: Sustainability requirement patterns. In: 3rd International Workshop on Requirements Patterns (RePa 2013), Rio de Janeiro, pp. 8–11 (2013). https://doi.org/10.1109/RePa.2013.6602665
21. Schmidt, B.: Konzeptuelle Elemente einer Beschreibungssprache für Usability-Muster in GDI-Anwendungen. Bochum University of Applied Sciences, Geovisualization Lab, report no. 18–01 (2018). https://www.researchgate.net/publication/327542721. Accessed 10 May 2021
22. Schmidt, B.: Towards a formal language to describe sustainable software design patterns. In: Schaldach, R., Simon, K.-H., Weismüller, J., Wohlgemuth, V. (eds.) Environmental Informatics, Adjunct Proceedings of the 33rd EnviroInfo Conference, pp. 128–135. Kassel, Germany (2019)
23. Schweizer-Ries, P.: Sustainability science and its contribution to IAPS: seeking for integrated sustainability. IAPS Bull. **40**(Autumn 2013), 9–12 (2013)
24. Sommerville, I.: Software Engineering, 8th edn. Addison-Wesley, Boston, MA (2011)
25. Steffen, D.: Praxisorientierende Designforschung und Theoriebildung. Dissertation, Bergische Universität, Wuppertal (2011). urn:nbn:de:hbz:468–20130118–124233–9
26. Umweltbundesamt, Hrsg.: Entwicklung und Anwendung von Bewertungsgrundlagen für ressourceneffiziente Software unter Berücksichtigung bestehender Methodik, Berlin (2018). https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2018-12-12_texte_105-2018-ressourceneffiziente-software_0.pdf. Accessed 10 May 2021
27. Warmer, J., Kleppe, A.: Object Constraint Language 2.0. mitp, Bonn (2004)

A Taxonomy About Information Systems Complexity and Sustainability

Yassin Bischoff, Robin van der Wiel, Bart van den Hooff, and Patricia Lago

Abstract With the increasing digitalization of all industry sectors, information systems are becoming more and more complex. At the same time, thanks to their crucial societal role, they have the potential to help (or hinder) organizations in their ambition to contribute to sustainability goals. In our work, we aim at helping them by identifying the important concerns that shape their information systems and that are found to influence complexity in certain sustainability aspects. To this aim, we perform a study in a mid-sized bank, and build a taxonomy of concerns that emerge from real projects and experience, and blend both complexity and sustainability.

Keywords Information systems · Complexity · Sustainability

1 Introduction

The continuously increasing pace of technological and environmental developments requires organizations to adapt and expand their Information Systems (IS) at an ever-higher rate. Information Technology (IT) departments cater to the demands of the business side by creating more and more powerful IS architectures[1] with increasing functionalities. However, over time this practice results in complex webs of com-

[1] An organizational IS architecture can be defined as *the set of an organization's IS components and their interdependencies* [1].

Y. Bischoff · R. van der Wiel (✉) · B. van den Hooff · P. Lago
Vrije Universiteit Amsterdam, Amsterdam, Netherlands
e-mail: r.r.vander.wiel@vu.nl

B. van den Hooff
e-mail: b.j.vanden.hooff@vu.nl

P. Lago
e-mail: p.lago@vu.nl

P. Lago
Chalmers University of Technology, Gothenburg, Sweden

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_2

ponents, (legacy) sub-systems and interdependencies that impose a barrier to future change. Put simply, the complexity introduced by meeting the business demands of today, can hinder an organization's ability to meet the demands of tomorrow.

The question of how to optimally manage IS complexity is a consistent priority of IT managers [2] and has resulted in an extensive and diverse literature stream [3]. Four prominent conceptualizations are: (i) *Structural Complexity*, which is measured by the number and heterogeneity of components and links in a system, (ii) *Subjective Complexity*, which is contingent on an individual's perception of how easily a system can be used, (iii) *Dynamic Complexity*, which emerges from the interaction between system components over time, and (iv) *Organizational complexity*, which encompasses the broader organizational context factors that influence IS. Most empirical studies adopt a single conceptualization of IS complexity as their theoretical lens, thereby narrowing their analytical focus to a limited set of the interdependent influencing factors. We argue that their joint consideration provides additional and often complementary foci for analysis that contribute to the understanding of complexity [4–6].

A second crosscutting concern that has been gaining increasing attention is sustainability. This trend has also reached predominantly technological areas such as software engineering [7]. The World Commission on Environment and Development (WCED) defines sustainability as *meeting current requirements without reducing future generations*' *ability to do likewise* [8]. In the light of this definition, the question emerges how the increasing complexity of an IS architecture influences an organization's capability to meet both current and future requirements, and thus to sustain itself over extended periods of time. Like complexity, sustainability is defined as a multi-dimensional concept that benefits from a comprehensive analysis: (i) *Economic sustainability*, i.e., the preservation of economic performance and value [7]; (ii) *Environmental sustainability*, i.e., the use of natural resources below their reproduction capacity; (iii) *Social sustainability*, i.e., the provision of equal or greater social capital to current and future generations [9], and (iv) *Technical sustainability*, i.e., the long-term use and appropriate evolution of technological components [7].

In essence, it is crucial for organizations to both manage the complexity of their IS architectures, and simultaneously be able to sustain their competitive capabilities in the long term. Nonetheless, empirical data on the relationship between IS complexity and sustainability is scarce, especially with regards to how projects that aim to reduce complexity can contribute to the achievement of sustainability goals. To fill this gap, this work aims to address the following research question: *How can we get a grip on the complexity of IS architectures so that we can understand how architectural decisions made over time influence IS-related sustainability?*

The remainder of this paper is organized as follows: Sect. 2 describes the theoretical background of this research and defines the different dimensions of IS complexity and sustainability. Section 3 presents the design and execution of this research, while Sect. 4 describes our findings. Next, we discuss them in Sect. 5 and then mention related works that investigate the combination of complexity and sustainability in Sect. 6. Finally, in Sect. 7 we set out our conclusions and discuss possible future work.

2 Theoretical Background

2.1 On IS Complexity

Organizations must continuously adapt their IS architectures to cater to fluctuating business demands and to keep up with the ever-rising pace of technological developments. Over time, myriads of interdependent components and links are accumulated; this results in increasingly complex architectures. As each system change mandates the careful consideration of all dependencies, successful management requires organizations to get a grip on their IS complexity [10]. From the literature, we derive four conceptualizations prevalent in complexity science, namely *structural*, *subjective*, *dynamic* and *organizational* complexity. These conceptualizations come with individual analytical strengths and limitations. Hence, it is argued that their joint consideration benefits the research question as this allows to analyse complexity from varying and complementing perspectives [6].

The *structural* conceptualization of IS complexity is the most widespread conceptualization in the literature [3, 11]. This states that the complexity of a system can be measured by the number and heterogeneity of its components and links [6, 11–14]. This conceptualization of complexity is static and objective [3, 15]. To reduce structural complexity, organizations may reduce the number of components and links, or implement modular IS architecture designs [6, 11]. Modularity means the creation of sub-systems which are only loosely coupled to one another through standardized interfaces. Such modules can be reused or replaced with relative ease, thus increasing an IS's adaptability [16, 17].

Dynamic complexity is based on the principle that systems are not static, but continuously evolve—adding a new dimension of complexity by introducing a temporal axis [14]. As the rate of change can vary, both over time and with regard to the number and variety of the components and links within an IS architecture, this take on complexity reflects the ambiguity, uncertainty and unpredictability that accompany IS complexity management [6, 18]. Based on Haki et al. [1] we define dynamic complexity in the following terms: "Dynamic complexity of an IS architecture arises from the interactions over time, between the different technological components of which this architecture consists, which dynamically and non-linearly influence one another and may lead to unpredictable outcomes.".

The *subjective* conceptualization posits that complexity is not a characteristic of the system, but of the relationship an individual human agent has with the system [19]. Thus, complexity is not a function of a system's components and links (i.e., of technological characteristics), but of the cognitive structures, heuristics and mental models that human agents use to evaluate a system (i.e., of social characteristics) [3, 6, 19].

In essence, subjective and structural complexity both regard complexity as being a mostly bounded and static occurrence. In practice, it is often useful to not analyze either system components or human agents, but rather how they interact and influence each other over time [11, 20]. IS and organizations together form a *complex* *adaptive system*, consisting of various subsystems that dynamically influence each other [21]. Therefore, it is important to also include *organizational* complexity in our analysis, which concerns the complexity of the organizational environment of which the IS architecture is a part. This concerns elements like the variety and number of, and interdependencies between, organizational goals and objectives, stakeholders, processes, decision makers, units and geographical locations as well as the degree of change in such elements [18, 22, 23]. In that sense, organizational complexity can be both structural and dynamic [18], while also inheriting a subjective element, as different actors may have different perceptions of how complex an organization is.

2.2 On Sustainability

While it is important that organizations adapt their IS architecture to respond to business demands and technological developments, it is also critical to consider what effects these changes have on the sustainability of the IS architecture.

Sustainability in general was defined by the WCED as meeting the current requirements without reducing future generations' ability to do likewise [8]. In particular, we consider the definition of sustainability by Lago et al. [24] in the context of software-intensive systems, where sustainability is expressed as the preservation of the function of a system over an extended period of time; this definition mainly points to the notion of technical sustainability. However, it is important to consider sustainability from a wider perspective than just the technical aspect, including economic, environmental, and social sustainability as well. Sustainability is all about addressing the bigger picture, not isolating one sustainability dimension. There seems to be a trade-off between the short-term interest and long-term benefits of software-intensive systems, resulting in poor sustainability of a system which can express itself in software aging and technical debt [25]. Lago et al. argue that a trade-off analysis is a suitable mechanism to simultaneously consider all sustainability concerns of relevance for a certain system [7].

In order to analyse all aspects of the sustainability of a software or information system, we distinguish between four different types of concerns, namely *economic*, *environmental*, *social*, and *technical* sustainability. Their definitions are adopted from Lago et al. [26] and adapted to the scope of this research.

First, *economic* sustainability regards the financial aspects of an organization, and more specifically aims at preserving capital and economic value over time [26]. For *environmental* sustainability, we are concerned with how software development and maintenance impact energy consumption, the usage of other natural resources, and ecologic awareness creation [26]. We consider and distinguish between being green *by* IT and green *in* IT: while green by IT focuses on software specifically designed for the domain of natural resource preservation, green in IT focuses on how to make software itself more sustainable, resulting in a more environmentally sustainable product [27]. *Social* sustainability is centered around the provision of equal or greater social capital to current and future generations [9]. As an example, social sustainability encompasses new IT solutions that directly support social communities in any domain [7]. Lastly, *technical* sustainability addresses the long-term use and appropriate evolution of technological components in an ever-changing environment [26]. This dimension entails e.g., the adaptability of an IS, hence allowing it to cater to ever-evolving business goals, resulting in said long-term use.

3 Study Design and Execution

3.1 Premise

The IS complexity domain builds upon an already extensive literature stream, and sustainability related research has gained considerable traction in the domain of IS in recent years. Nonetheless, how complexity may influence the achievement of sustainability objectives is still mostly unexplored. To gain a first understanding of this relationship, we investigated the relevant IS complexity and sustainability literature. Building upon this literature, we created a conceptual framework (see Fig. 1) that describes potential interrelationships between the two concepts.

The conceptual framework summarizes our basic assumptions in terms of how the different dimensions of complexity and sustainability may be interrelated. The "decisions" shown in the framework refer to the three cases we have studied, as explained below.

In order to build an analytical and explanatory theory [28], we adopted a case study design approach. This allows rich data collection through an in-depth investigation of interdependencies and contextual factors. Case studies are well recognized in IS research [29, 30] and are an especially appropriate research strategy for building theories in new topic areas [31].

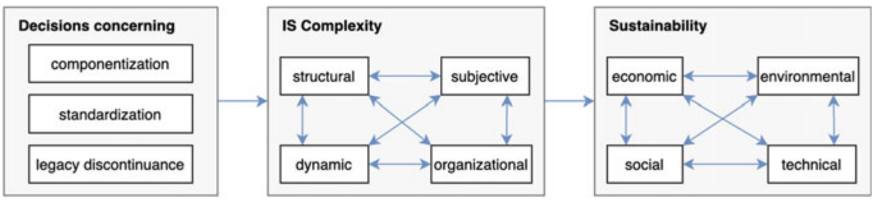


Fig. 1 Conceptual framework

3.2 Research Setting

We chose to conduct our study at Alfa Bank (anonymized), which is a mid-sized Western European bank with private, corporate and retail clients. Alfa Bank has a mixed software portfolio, part of which was built in-house and the other part licensed from larger companies (e.g., Microsoft). The share of software that was built in-house has become smaller over time, due to fast advancements in the software world and the lack of manpower to maintain this software. Alfa Bank has recently committed itself to the pursuit of Sustainable Development Goals (SDGs) as defined by the UN. Hereby, the existing IS architecture as well as the continuous digitalization of the company are considered to be key enablers of this pursuit. In addition, Alfa Bank started a long-term and company-wide IS complexity reduction program in 2013, as the complexity of their IS architecture was perceived as hindering business agility. This program encompassed decisions regarding three concrete interventions in the IS architecture: *componentization*, *standardization* and *legacy discontinuation*. We studied these three interventions as three embedded case studies within Alfa Bank.

The *componentization* decision aimed at resolving the complex web of customized point-to-point interfaces between system components by introducing a *Service Oriented Architecture* (SOA) with an Enterprise Service Bus as a standardized backbone for the IS architecture. The *standardization* initiative aimed at reducing the number of redundant or duplicate interfaces, systems and platforms by migrating them to a new centralized platform. Standardizing further corresponded with the decision to *discontinue legacy systems* to reduce maintenance effort and barriers to system adaptability. One of the authors of this study was involved in qualitative case studies conducted on each of the three decisions. These case studies were taken as a foundation for our conceptual framework and introduce a longitudinal perspective into our study, by enabling us to analyze how decisions made in the past influence the current level of IS complexity.

3.3 Data Collection and Analysis

The overall study design and execution is summarized in Fig. 2. It consists of four phases of data collection and analysis. The *first* research phase, a literature study, produced the conceptual framework introduced in Sect. 3.1 which served as a basis for framing the concepts of complexity and sustainability for the rest of the study. To reflect our findings, the conceptual framework was revised on multiple occasions throughout the study.

The three case studies described in Sect. 3.2 were extensively documented. Technical reports and Master theses based on these case studies were the primary source of data for the *second* phase, in which we analyzed the three case studies and mapped each case study's findings in so-called *Decision Maps* (detailed in Sect. 4.1),

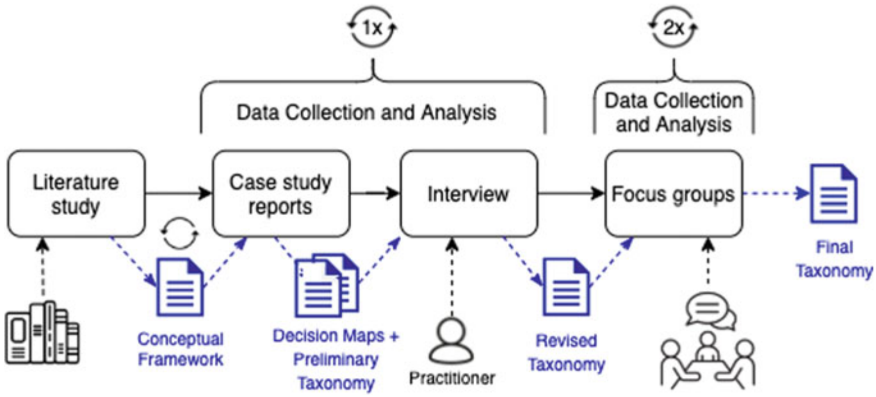


Fig. 2 Study design and execution

i.e., diagrams framing *concerns* and their interdependencies [7]. The decision maps helped identify a comprehensive list of the consequences (i.e., concerns) that followed the IS complexity related interventions. The case study analysis resulted in a preliminary taxonomy, where every decision related to IS complexity reduction was linked to a complexity and sustainability dimension, hence indicating what impact a decision has in what areas.

In the *third* phase of data collection we validated and expanded these first findings through an interview with a manager in the IT department of Alfa Bank who was involved in all three case studies. This interview resulted in minor updates to the preliminary taxonomy.

Finally, during the *fourth* phase, we organized two focus groups with employees from Alfa Bank with the objectives of (i) confirming the findings from phase two, and (ii) uncovering possible new complexity- or sustainability-related concerns that did not follow from the case studies directly. The first focus group was conducted with three Alfa Bank employees, and the second group with two. Four of the participants work in different sub-departments of the Platforms & Technologies office, while one of the participants is an IT business consultant for the bank. Each focus group lasted for one hour and started by presenting the sustainability and complexity dimensions. The participants were asked to think of examples where complexity and sustainability played a role in their day-to-day work and if there were links between the two domains. The focus groups were concluded by a presentation of the preliminary taxonomy which was then discussed and expanded based on the participants input.

4 Findings

4.1 Case Analyses

As explained in Sect. 3.2, we analyzed the cases by summarizing the network of concerns related to each of the three decisions (componentization, legacy discontinuance, and standardization) in a decision map. Figure 3 shows the decision map for the componentization case. In this section we discuss the main outcomes of these analyses, leading to our preliminary taxonomy. We discuss the analysis of the componentization decision in some more detail, to illustrate how we used the decision maps to link complexity dimensions and sustainability concerns. Because of space constraints, we discuss the other decisions less elaborately.

4.1.1 Componentization

There were three features of the componentization decision that affected complexity: the *level of granularity*, use of a *unified data model*, and use of an *enterprise service bus*.

Level of granularity refers to how much functionality is built into a service. A high degree of granularity means many services with very limited functionality, whereas a low degree of granularity means less services each with a broader range of built-in

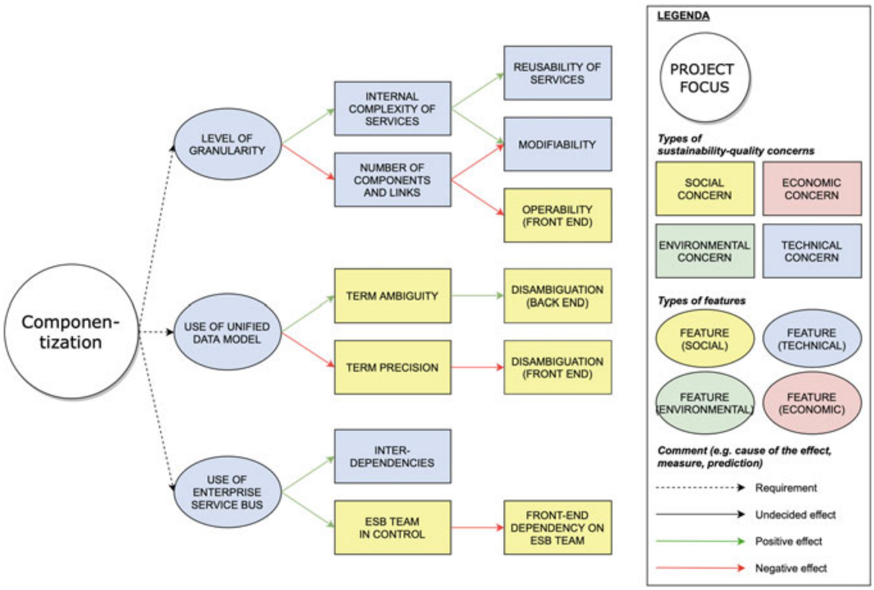


Fig. 3 Decision map for componentization

functionalities. The level of granularity influences the structural complexity of the IS architecture: first, by affecting the *Internal Complexity of Services*, where a broader range of functionality means a higher level of complexity; second, by affecting the *Number of Components and Links* in the architecture (where high granularity means more services and links between them). These structural complexity dimensions, in turn, affect the Reusability of services (e.g., it is easier to reuse services with less internal complexity), the *Modifiability* of the architecture (e.g., more components and links make it difficult to implement changes), and the *Operability* of the IS architecture from a user perspective (e.g., more components and links make it more difficult to conduct certain tasks).

Use of a Unified Data Model refers to the degree to which definitions of data are consistent across different actors. Unified data definitions impact the structural complexity of the architecture by affecting the technical concern *Term Ambiguity* (reducing the number of terms used for one and the same thing) and the *Term Precision* (as more complex definitions are typically the result). This, in turn, affects the social concern of *Disambiguation*, which has two connotations: reducing the ambiguity in the meaning of certain entities and their attributes, and resolving the uncertainty about the correct use of terms carrying complex definitions.

Use of an Enterprise Service Bus describes the degree to which applications are linked to one another through a single focal point: the Enterprise Service Bus (ESB). An ESB as central gateway for all applications reduces the *Number of links* (structural complexity), hence reducing worries about potentially harmful interdependencies between applications. Interestingly, we uncovered also a social concern (related to organizational complexity): as the *ESB team is in control*, there is an increased *dependency on the ESB team* from other actors.

4.1.2 Discontinuing Legacy Systems

A relatively straightforward way to reduce the structural complexity of an IS architecture is to discontinue *Legacy Systems*, incumbent systems that are approaching or are beyond their useful life in organisations. Our analysis of the decision to discontinue legacy at Alfa Bank reveals that such discontinuance has a positive influence on a number of concerns, such as *Maintenance Costs*. A prominent concern determining the success of legacy discontinuance is *System Interdependence*, the extent to which legacy systems are connected to other elements of the architecture. A number of more detailed concerns were found that influence this interdependence, such as *Centrality* (the number and depth of connections the legacy system has to other systems, indicating its central role in the architecture), *Stringency* (the degree to which the legacy system predetermines information flows), and *External Susceptibility*, defined as the extent to which a system can be affected through external developments. Susceptible systems will require more frequent adaptations and increase the risks of starting a chain reaction, whereby one change may create the need for further adaptations in other IS and applications. These concerns, in turn, are related to other concerns that directly influence the success of the discontinuance project. Examples of such concerns are the discontinuation *Budget*, the company's *Strategic focus*, the company's *Mergers & Acquisitions* (the need of integrating another company's systems), *Regulations & Legal Compliance* (which may force the company to continue its legacy systems), and *Technical Interdependencies* between systems (in terms of software, data, processes, etc.).

4.1.3 Standardization

The standardization decision in this case concerned the migration of applications from a variety of platforms to one and the same standardized cloud-based platform. Our analysis of this case shows an important role for the *Interdependence* of systems. This reflects systems' interconnectedness in terms of both *Horizontal Coupling* (links with other applications) and *Vertical Coupling* (links between an application and other architectural layers, such as data or technology). Vertical coupling was especially relevant here, because insufficient attention for the coupling between an application and its underlying layers during the migration to the cloud platform triggered all kinds of (unexpected) incidents (both in terms of *Priority of Incidents* and *Number of Incidents*). So while standardization on the one hand reduced structural complexity (reducing the variety in platforms), it increased dynamic complexity in terms of unpredictable interactions between architectural layers, leading to unpredictable outcomes. This, in turn, affected the technical concern *IS Application Instability* and the social concern *User satisfaction*—in the case of a poorly functioning IS, more incidents will be reported, resulting in a decreased user satisfaction.

4.2 The Taxonomy

In this section, we present the final taxonomy of concerns that followed as a result of the case studies, interview, and focus groups. The taxonomy depicted in Fig. 4 integrates the high level concepts described in the conceptual framework (i.e., complexity and sustainability dimensions). The model starts with *Complexity* in the center, then spreads out into four complexity conceptualizations and 10 sub-categories, which in turn contain leaves. These leaves represent the concerns identified during the data collection phases, where the colour-coding indicates the corresponding sustainability dimension. For the sake of traceability, the concerns identified during the focus groups were marked with an asterisk, whereas the leaves without asterisks emerged from the case studies.

All four conceptualizations of IS complexity derived from the literature are reflected in our taxonomy, each with particular relations to sustainability concerns. Hereby, the majority of the concerns affect technical sustainability (blue leaves in the taxonomy), followed by social sustainability (yellow leaves), with only a few economic (red) and environmental (green) concerns.

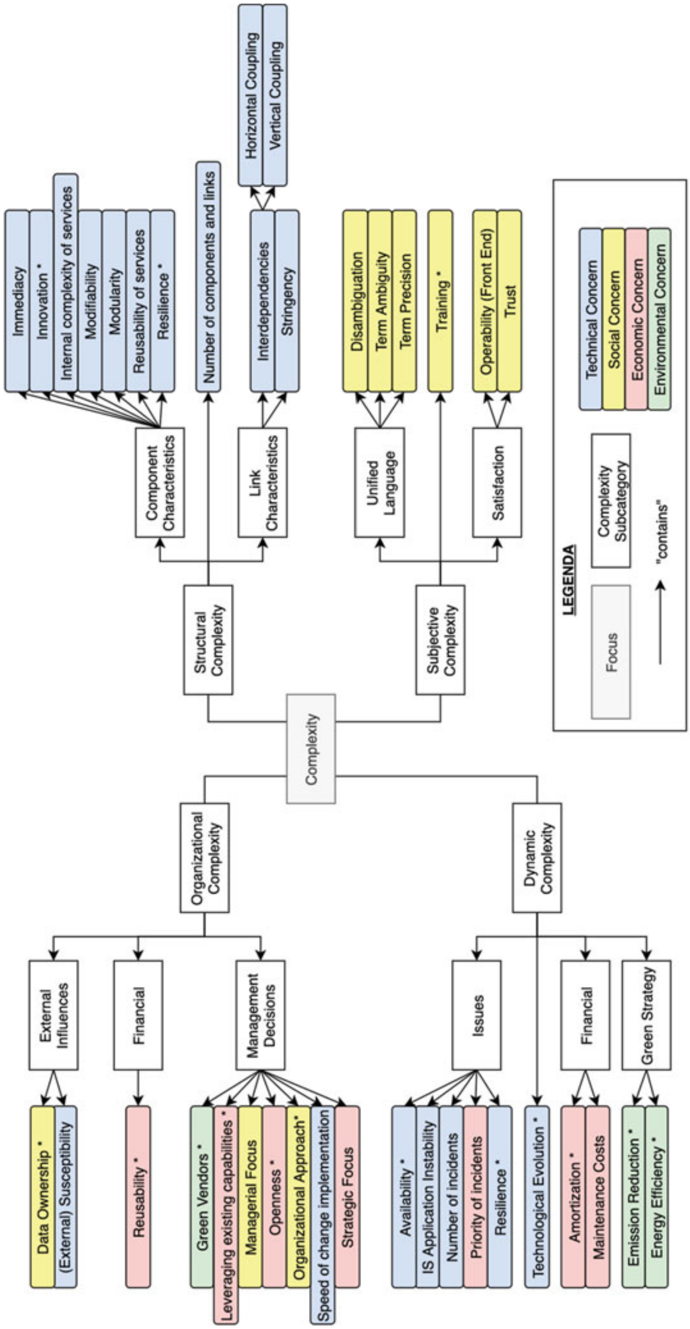


Fig. 4 Final taxonomy

4.2.1 Structural Complexity

This branch displays the highest number of technical sustainability concerns, due to the mainly technical perspective of the structural complexity dimension. The concerns found for *Structural Complexity* can be grouped into two sub-categories, namely *Component Characteristics* and *Link Characteristics*. The concern *Number of components and links* was found to be a universal indicator of structural complexity, not restricted to either components or links.

Component Characteristics pertain to the technical characteristics of the components in an architecture. We identified that these concerns mainly impact the technical sustainability of IS. For example, *Modifiability* of IS, refers to the degree to which IS can be effectively modified without introducing defects or degrading existing product quality. This inherently plays an important role in the long-term use of IS, thus having an impact on technical sustainability.

Link Characteristics relate to characteristics of links (i.e., connections) between and within IS. A higher number of *Interdependencies* of IS can increase structural complexity, which can in turn have a negative impact on the technical sustainability of the IS architecture. Interdependencies in this case can be a result of horizontal coupling (interfaces with many other systems) or vertical coupling (interfaces with other architectural layers of the same system) [4].

4.2.2 Subjective Complexity

Subjective Complexity is divided into *Unified Language Model* and *Satisfaction*. Both sub-categories lie in the eye of the beholder (i.e., they are subjective) and as such concern social sustainability.

The *Unified Language Model* describes how consistently terminology is used throughout an organization. This is a social concern, both because it affects how people communicate, but also because the rules established for this communication (e.g., in form of a Unified Data Model) impact subjective complexity. In this context, *Disambiguation* depicts the degree to which *Term Precision* and *Term Ambiguity* in the organizational language affect the clarity of terminology. For instance, precise terms may mitigate the risk of misinterpretations, but also increase the complexity involved with each term: instead of just referring to “an account”, it would be necessary to always specify what kind of account one is referring to. When among users in the same organizational context such details are assumed implicitly, their specification could hinder the natural communication flow and add subjective complexity.

Another re-occurring sub-category was *Satisfaction*. Satisfaction improved when easy system *Operability* was given for front-end users, e.g., because they could conduct all their tasks within a single application instead of having to switch continuously.

4.2.3 Dynamic Complexity

Dynamic Complexity describes the concerns that may arise through interactions between or within IS over time. This branch has three sub-categories, *Issues*, *Financial*, and *Green Strategy*.

The sub-category *Issues* describes concerns that are related to effectively preventing and handling problems. An example of preventing possible issues is to ensure a high *Availability* of IS, such that other systems that rely on these IS are able to function accordingly. Additionally, handling incidents according to their *Priority* is important from the perspective of economic sustainability. With limited resources at hand, it is crucial to prioritize fixing issues of the most important systems, in order to minimize the negative economic effects.

Financial concerns influence the economic sustainability of an IS over time. *Maintenance costs* of IS are especially interesting here. For example, keeping maintenance costs low might result in a system that is discontinued after a number of years due to poor maintenance. With more maintenance effort (and costs), the system might well have remained functional for a longer period of time, thus leading to a higher level of economic sustainability.

The final sub-category of dynamic complexity describes two concerns that have a positive effect on the environmental sustainability of IS. From the focus groups it became apparent that Alfa Bank is working on implementing a company-wide green strategy, which at this time is underrepresented in their decisions and actions in relation to IT. While Alfa Bank supports green initiatives in general, "green in IT" strategies are not yet incorporated in their daily practices. An example of a concern that was taken into account by Alfa Bank is the *Energy Efficiency* of IS. As a result, the trade-off between performance and energy efficiency has a positive impact on the environmental sustainability (focusing on the "best performance with respect to green").

4.2.4 Organizational Complexity

Organizational Complexity relates to the broader organizational context that IS are positioned in, such as *External Influences*, *Financial* aspects and *Management Decisions*. It is interesting to notice that this dimension is the only one including concerns in all four sustainability dimensions. This finding reflects the diversity in the types of concerns that can occur in the context of organizations.

Organizational complexity was found to most clearly represent the influences that occurred due to organizational systems not being bounded and isolated, but to also depend on external factors. Hereby, *External Influences* showed, for instance, that technology can be *Susceptible* to change, because governmental regulations and policies mandate a different procedure to ensure customer privacy. The number of external influences and the authority they have over an organization was found to increase organizational complexity.

Furthermore, it became apparent that identical sub-categories, e.g., *Financial*, can be subject to multiple complexity dimensions. To which dimension they belong depends on the context and motivation of the individual concern. For instance, the economic concern of *Reusability* regards how a certain part of the IS architecture can be used again to support other parts of the organization. Being able to reuse such parts helps organizations reduce the organizational complexity and resource investments.

Finally, due to the breadth of their scope, *Management Decisions* include concerns in all sustainability dimensions. Decisions were found to impact social concerns (e.g., the way users collaborate in projects due to a certain *Organizational Approach*), but also technical ones (e.g., the *Speed of Change Implementation*). Also, environmental concerns may influence how partnerships with external vendors were evaluated, e.g., *Green Vendors* were preferred over those that did not actively engage in green activities. However, the consideration of additional evaluation criteria, e.g., whether a potential partner qualifies as a "Green Vendor", increases the complexity involved in making a management decision.

5 Discussion

In answer to our main research question, i.e., "*How can we get a grip on the complexity of IS architectures so that we can understand how architectural decisions made over time influence IS-related sustainability?*", our findings show that different architectural decisions affect the structural, subjective, dynamic, and organizational complexity of IS architectures. Next, our taxonomy shows that changes in each of these four conceptualizations of IS complexity are related to the four dimensions of sustainability in different ways. The set of interrelationships between the dimensions of sustainability and complexity can be depicted as in Fig. 5. As such, using the taxonomy and the discovered interrelationships between these dimensions can help decision makers in identifying, measuring, and monitoring important concerns that are recurrent and impact the sustainability of IS projects, and analyze how different interventions in the complexity of IS architectures may affect the different dimensions of sustainability. In this context, a suitable tool for measurement would be the Software Sustainability Assessment Method (SoSA) by Condori & Lago as described in [32].

The concerns extracted from the cases were all confirmed during the focus groups (albeit with minor corrections and additions). However, it was especially interesting to observe that the overview provided by the taxonomy helped the participants reflect on the "big picture" and notice that (i) most complexity concerns were either technical or social; and (ii) no environmental concerns were present in spite of the significant ongoing investment of Alfa Bank in Green IT. This is a consequence of the fact that the bank's sustainability strategy is relatively recent. Still, the taxonomy worked as a trigger to reflect on what other concerns related to environmental sustainability were already in place throughout the organization, leading to uncovering the three

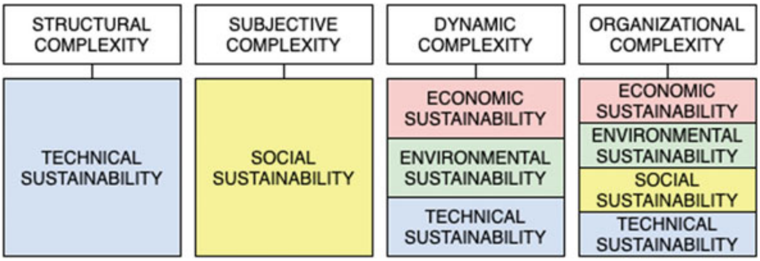


Fig. 5 Identified interrelations between complexity and sustainability dimensions

environmental concerns in Fig. 4. This shows that the taxonomy can be used to spark discussions, and to highlight areas of interest which were previously unnoticed (e.g., the opportunity of becoming more green *in* IT).

It is worth noting that the taxonomy neither claims to holistically represent all potential concerns that may follow IS related decisions, nor does the displayed representation of sustainability dimensions necessarily reflect the consequences such decisions can have in projects. The transferability of the model to other domains *may* not be guaranteed, as the practitioners with whom the taxonomy was created and validated all worked at Alfa Bank. However, the studied literature was not bounded to the domain of banking, and therefore the taxonomy consists of numerous concerns that are not related to the banking domain in particular, but pertain to general aspects of IS (e.g., reusability, or modularity).

6 Related Work

To the best of our knowledge, there are no previous works that contribute like we do, to the combination of complexity and sustainability. Some works, however, are related in that they focus on sustainability of information systems and are positioned in a context similar to ours. For example, Korte et al. [33] apply social, environmental and economic dimensions of sustainability to the management of IT and try to develop a "sustainable framework". Vessey and Ward [21], in turn, develop the idea of sustainable IS alignment, which they define as "the enduring fit between an organization's goals and its IS over time". They adapt the complexity theory worldview and conceptualize organizations as complex adaptive systems (CAS). Seidel et al. [29] go more in the direction of Green IS, i.e., how information systems can contribute to achieving environmental goals. In their work, however, complexity is just mentioned. Finally, from the perspective of sustainable development, Munasignhe [34] argues for the potential value of applying concepts from other disciplines, such as complex adaptive systems: "the concept of co-evolution of social, economic, and ecological systems within a larger, more complex adaptive system provides useful insights".

7 Conclusion and Future Work

This paper presents a taxonomy with the goal of aiding decision making with regard to complexity and sustainability management in IS architectures. The primary contribution resides in linking the two previously distinct domains of IS complexity and sustainability to one another, exploring potential interdependencies and influences.

As this research was intended as an exploratory study to open the floor for the investigation of interdependencies between IS complexity and sustainability, further refinement of the taxonomy is required. Three avenues for possible future research are: (i) an iterative expansion of the provided taxonomy to capture relevant concerns more comprehensively; (ii) the validation of the taxonomy in different organizational domains to assess the generalizability of the results; and (iii) an operationalization of the individual concerns and their correspondent categories and dimensions as metrics to allow quantified comparisons.

Acknowledgements This research was partially funded by the Network Institute - Vrije Universiteit Amsterdam under the Academy Assistants research program.

References

1. Haki, K., Beese, J., Aier, S., Winter, R.: The evolution of information systems architecture: An agent-based simulation model. MIS Q. **44**(1) (2020)
2. Kappelman, L., Torres, R., McLean, E., Maurer, C., Johnson, V., Kim, K.: The 2018 SIM IT issues and trends study. MIS Q. Executive **18**(1), 7 (2019)
3. Jochemsen, E., van den Hooff, B., Rezazade Mehrizi, M., Plomp, M.: Understanding information systems complexity: a systematic literature review. Amsterdam: School of Business and Economics, 28 p. (Research Memorandum; vol. 2020, no. 1). **2020**(1) (2020)
4. Jochemsen, E., van den Hooff, B., Plomp, M., Rezazade Mehrizi, M., Derks, L.: The effects of standardization on IS architecture complexity. In: Twenty-Ninth European Conference on Information Systems (ECIS), Marrakesh, Morocco (2021)
5. Mol, K.: The impact of componentization on IS complexity (2017)
6. Widjaja, T., Gregory, R.W.: Monitoring the complexity of IT architectures: design principles and an IT artifact. J. Assoc. Inf. Syst. **21**(3), 4 (2020)
7. Lago, P.: Architecture design decision maps for software sustainability. In: International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS). IEEE/ACM, pp. 61–64 (2019)
8. Brundtland, G., Khalid, M., Agnelli, S., Al-Athel, S., Chidzero, B., Fadika, L., Hauff, V., Lang, I., Shijun, M., Morino de Botero, M., Singh, M., Okita, S.: Our Common Future (Brundtland report). Tech. rep., World Commission on Environment and Development (1987)
9. Dyllick, T., Hockerts, K.: Beyond the business case for corporate sustainability. Bus. Strategy Environ. **11**(2), 130–141 (2002)
10. Ross, J.W., Weill, P., Robertson, D.: Enterprise Architecture as Strategy: Creating a Foundation for Business Execution. Harvard Business Press (2006)
11. Merali, Y.: Complexity and information systems: the emergent domain. J. Inf. Technol. **21**(4), 216–228 (2006)
12. Schütz, A., Widjaja, T., Gregory, R.: Escape from winchester mansion–toward a set of design principles to master complexity in IT architectures. In: International Conference on Information Systems (2013)

13. Schuetz, A., Widjaja, T., Kaiser, J.: Complexity in enterprise architectures-conceptualization and introduction of a measure from a system theoretic perspective. In: Proceedings of 21st European Conference on Information Systems (2013)
14. Schneberger, S.L., McLean, E.R.: The complexity cross: implications for practice. Commun. ACM **46**(9), 216–225 (2003)
15. Schneider, A., Zec, M., Matthes, F.: Adopting notions of complexity for enterprise architecture management. In: Twentieth Americas Conference on Information Systems, pp. 1–10 (2014)
16. Mikalef, P., Pateli, A., van de Wetering, R.: IT architecture flexibility and IT governance decentralisation as drivers of it-enabled dynamic capabilities and competitive performance: the moderating effect of the external environment. Eur. J. Inf. Syst. 1–29 (2020)
17. Yoo, Y., Henfridsson, O., Lyytinen, K.: Research commentary-the new organizing logic of digital innovation: an agenda for information systems research. Inf. Syst. Res. **21**(4), 724–735 (2010)
18. Xia, W., Lee, G.: Complexity of information systems development projects: conceptualization and measurement development. J. Manage. Inf. Syst. **22**(1), 45–83 (2005)
19. Fioretti, G.: A subjective measure of complexity. Adv. Complex Syst. **2**(04), 349–370 (1999)
20. Orlikowski, W.J.: Using technology and constituting structures: a practice lens for studying technology in organizations. Organ. Sci. **11**(4), 404–428 (2000)
21. Vessey, I., Ward, K.: The dynamics of sustainable IS alignment: the case for IS adaptivity. J. Assoc. Inf. Syst. **14**(6), 2 (2013)
22. Piccinini, E., Gregory, R., Muntermann, J.: Complexity in IS programs: a delphi study. In: Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv. Israel (2014)
23. Beetz, K.R., Kolbe, L.M.: Towards managing IT complexity: an IT governance framework to measure Business-IT responsibility sharing and structural IT organization. In: AMCIS (2011)
24. Lago, P., Penzenstadler, B.: Reality check for software engineering for sustainability—pragmatism required. J. Softw.: Evol. Process **29**(2) (2017)
25. Becker, C.: Sustainability and longevity: two sides of the same quality? In: Workshop on Requirements Engineering for Sustainable Systems (RE4SuSy), vol. 20, p. 21. ceur-ws.org (2014)
26. Lago, P., Koçak, S.A., Crnkovic, I., Penzenstadler, B.: Framing sustainability as a property of software quality. Commun. ACM **58**(10), 70–78 (2015)
27. Calero, C., García-Rodríguez De Guzmán, I., Moraga, M., García, F.: is software sustainability considered in the CSR of software industry? Int. J. Sustain. Dev. World Ecol. **2**6(5), 439–459 (2019)
28. Gregor, S.: The nature of theory in information systems. MIS Q. 611–642 (2006)
29. Seidel, S., Recker, J., Vom Brocke, J.: Sensemaking and sustainable practicing: functional affordances of information systems in green transformations. MIS Q. 1275–1299 (2013)
30. Majchrzak, A., Markus, M.L., Wareham, J.: Designing for digital transformation: lessons for information systems research from the study of ICT and societal challenges. MIS Q. **40**(2), 267–277 (2016)
31. Eisenhardt, K.M.: Building theories from case study research. Acad. Manage. Rev. **14**(4), 532–550 (1989)
32. Condori-Fernandez, N., Lago, P.: Characterizing the contribution of quality requirements to software sustainability. J. Syst. Softw. **137**, 289–305 (2018)
33. Korte, M., Lee, K., Fung, C.C.: Evolving IT management frameworks towards a sustainable future. In: Building Sustainable Information Systems. Springer US, pp. 271–284 (2013)
34. Munasinghe, M.: Sustainable development: basic concepts and application to energy. In: C.J. Cleveland (ed.) Encyclopedia of Energy. Elsevier, pp. 789–808 (2004)

Software Design for Sustainability

Microservice-Based Architecture for the Integration of Data Backends and Dashboard Applications in the Energy and Environment Domains

Jannik Sidler, Eric Braun, Christian Schmitt, Thorsten Schlachter, and Veit Hagenmeyer

Abstract This article presents a software architecture based on the onion architecture that uses the concept of application microservices in order to integrate data backends with dashboard applications. Its main goal is to reduce the complexity in the architecture's frontend and therefore to increase the performance of the application for the user. The concept of the added application layer as well as its interaction with the other parts of the architecture is described in detail. Then an evaluation of its advantages is presented which shows the benefits of the concept regarding performance and simplicity using a real-world use case in the energy and environmental domains.

Keywords Environmental information systems · Energy dashboards · Web application · Software architecture · Application microservice · Onion architecture

1 Introduction

Nowadays, climate change is one of the dominating and most difficult challenges mankind has to face. In this context, many measures are in progress of being accomplished in order to support the energy transition and the protection of the environment. During this process, the requirements for software engineers began to rise steadily, since one of the most important goals on the road to a successful deceleration of climate change is making the world's population understand how their behaviour can be relevant in the process of deceleration. To improve this understanding, there are various possibilities. As the internet has a great relevance in the process of gathering information, a great potential lies in the numerous websites that deal with topics such as environment protection. Consequently, an important aspect in the process of the population's understanding of the changes our world succumbs is the utility of

J. Sidler (✉) · E. Braun · C. Schmitt · T. Schlachter · V. Hagenmeyer
Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: jannik.sidler@kit.edu
URL: https://www.iai.kit.edu

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_3

these websites. Therefore, tools and applications are required which make the usage of renewable energy sources, the effects of the energy transition and the change of different sectors of the environment visible for everyone. Additionally, the influences of climate change on the environment have to be measured and visualised.

In the present article, a software architecture is presented which supports the formerly mentioned scenarios. The architecture consists of three parts. The first part is a set of web services which is called Generic Microservice Backend (GMB), using the microservice design pattern [5, 6]. GMB's competence lies in the provision of generic master data and timeseries data that is usable in different contexts and applications. GMB is based on previous work described in [1] and is used as a foundation in the present article. In general, these web services used in GMB are derived from a software architecture called *onion architecture* [2]. They fulfill the tasks of the domain services which are described in Sect. 2.

The second part is an application layer that also contains web services which are located on top of the previously described domain services. These web services are called application services in the onion architecture [2]. In contrast to the domain services, they perform tasks which are more specific to an existing application, where the domain services perform tasks that are of a more general nature which can be reused in multiple other scenarios. The application services also use the microservice design pattern [5, 6].

The third part is a flexible and reusable framework [7]. Its main goal is the visualisation of data by using lightweight web components which provide easy integration into existing websites and other web frameworks.

The main goals of this article can be summarised into three different aspects:

- Usage of application services to enhance the efficiency of an existing architecture
- Reusability of components
- Support for applications in the energy- and environment domain

The structure of this article is as follows: In Sect. 2, foundations and related work, which are principally connected to the presented work, are described and compared with the architecture at hand. In the third section, the mentioned architecture is described in detail and core concepts are explained. In Sect. 4, an evaluation regarding the usefulness of the achieved work is provided. Additionally, improvements of the current architecture are explained and discussed. The last section summarises the presented concepts and approaches and provides an outlook on further work that still has to be done in this context.

2 Foundations and Related Work

This section presents foundations of the given software architecture as well as similar work from different contexts. One important aspect in the scope of the present article is the term *microservice*. According to [5, 6], microservices are a design pattern in software engineering enhancing the encapsulation of functionality in distributed logical units and furthermore, the independence of these logical units of other functionality used in the same context. According to Fowler, microservices are "an approach to develop a single application as a suite of small services, each running in its own process and communicating with lightweight mechanisms" [5]. This approach has several advantages. First, the reusability of microservices is high. They provide an API that can be easily reworked, according to the context's requirements. Furthermore, they can be reused easily in different contexts due to their modular design. Second, the maintainability of microservices is very good. As they are independent of other microservices, they can be maintained without creating undesired effects on other components, as the only point of access to the microservice is its API. Finally, microservices fit very well in modern software deployment infrastructures. As they encapsulate only a certain, manageable amount of functionality, they can easily be deployed "by a fully automated deployment machinery" [5]. For these reasons, the architecture presented in this paper uses the microservice design pattern.

Moreover, the core concept used in the present article are the application services given by the onion architecture [2]. The onion architecture is a pattern which compares the structure of a software architecture with the one of an onion. In Fig. 1, this reference architecture is depicted in detail. The domain model is located in the middle. It is the core of the architecture, as every behaviour and state of the architecture depends on the way the domain is organised. The layer around the domain model is called domain services which are closely related to the domain model and implement the behaviour of the domain. Further outside, the application services connect the user interface and infrastructure (which are the layers on the outer edge) to the domain services. They contain business logic which defines the functionality of the application that is build using the domain layer. The outer layers are coupled with the next inner layer in an unidirectional manner, which means that the user interface is indirectly coupled to the domain services as well, but not vice versa.

In the present article, the onion architecture and the microservice design pattern are combined. This has already been done in [3] by using domain-driven design (DDD) [4] which strongly encourages using microservice architectures and can additionally be mapped to the onion architecture. Consequently, the application services will be adressed more specifically as application microservices in the following sections and in the context of the presented architecture.

3 Architecture for Applications Using Data Backends

In this section, the architecture for applications using generic data backends is explained. Figure 2 shows that the architecture is grouped into three different layers.

The domain layer mainly consists of three microservices [5]: The Master Data Service (MD) is responsible for managing master data, including functionality to read and write them from/to a persistent database. To ensure this functionality, the MD service has an appropriate REST API, providing all the functionality needed by

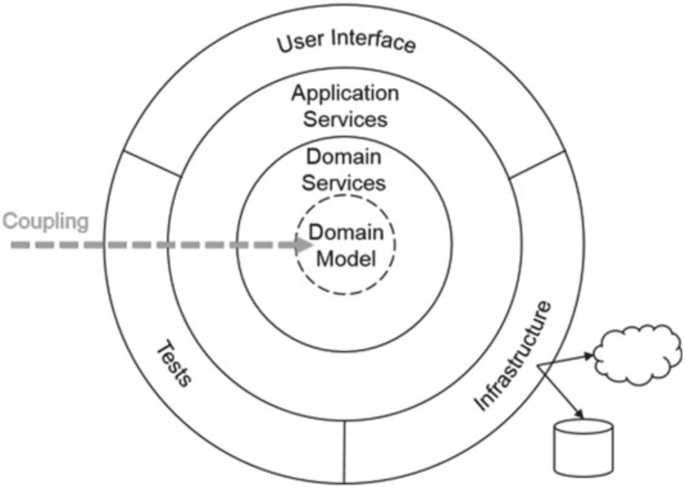


Fig. 1 The onion architecture [3]

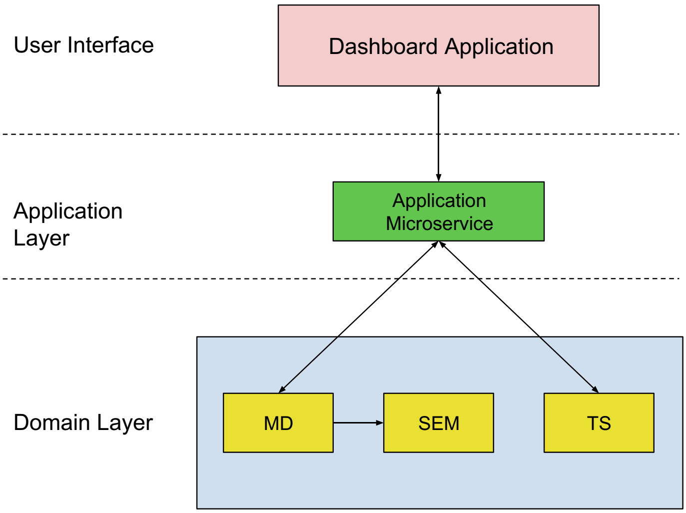


Fig. 2 Architecture of the given dashboard application

applications, e.g. providing requested data by using filters for more specific search queries and by providing the ability to sort and compute aggregations [1]. The Semantic Service's (SEM) goal is the validation of master data objects by using predefined schemas as references. These schemas define the structure of master data objects of a specific category. Therefore, the SEM service is required to verify the validity of master data objects [1]. Furthermore, the SEM service adds semantic information to master data objects, as for instance additional descriptions, data types and further meta information. The Timeseries Service (TS) manages timeseries data, e.g. measurements. Therefore, it uses optimised database tools for this kind of data. Additionally, its API provides all necessary functionality for aggregating and filtering time series data according to the application's requirements [1].

The user interface [7] represents the application that provides the top-level components which are responsible for user interactions. It is implemented by using a flexible framework which has been developed previously. More details can be obtained from [7].

The application layer is located between the previously described layers. The concept of the application microservices based on the onion architecture [2] has been described in Sect. 2. In the present paper, this concept is mainly used to increase the performance of the user interface by reducing its complexity and responsibilities. This extension is described in the following.

In order to improve performance, there are two main aspects which have to be considered. The first one is data processing. Components requesting data potentially need specific formats containing particular attributes and/or a fixed number of values. These special data formats usually cannot be provided by domain services, which means that a certain number of processing steps has to be added to the application layer. This data processing, if executed in the user interface, has to be performed by the end user's hardware resources, as without these processing steps, the data cannot be visualised by the user interface components. Consequently, if the user cannot afford these hardware requirements, the user interface will be slow and it takes longer until changes are visible, which leads to the user's dissatisfaction. Additionally, integrating business logic in the user interface is not consistent with best practices such as the model-view-controller pattern (MVC) [8]. The solution for this problem is to shift these data processing tasks from the user interface (and from the user's hardware) to the backend side, precisely to the application layer. The application microservice(s) can be used to accomplish processing tasks and consequently provide the specific data formats for the applications.

The second aspect that has to be considered in order to improve the user interface's performance is the network aspect. Without an application layer (or a similar gateway component), the user interface of the microservice-based architecture has to make many requests to get the required data from the backend since the data is distributed across different microservices and resources. Furthermore, in this case, the overhead from the data retrieving is much higher compared to retrieving the data with less requests. This leads to a higher network load, as constantly high numbers of data requests have to be made in the background. Combined with the previously described aspect of data processing, this leads to a significantly reduced performance. This problem can be tackled by using an additional application layer, too. The number of requests the user interface has to make can be reasonably reduced when it does not have to organise the required data by itself. Alternatively, the application microservice is used to orchestrate the different requests to the domain microservices. Afterwards, the data is filtered by accomplishing necessary processing tasks and by generating the required data format. At this point, an improvement regarding performance of the user interface and complexity of the single component layers is reached. Instead

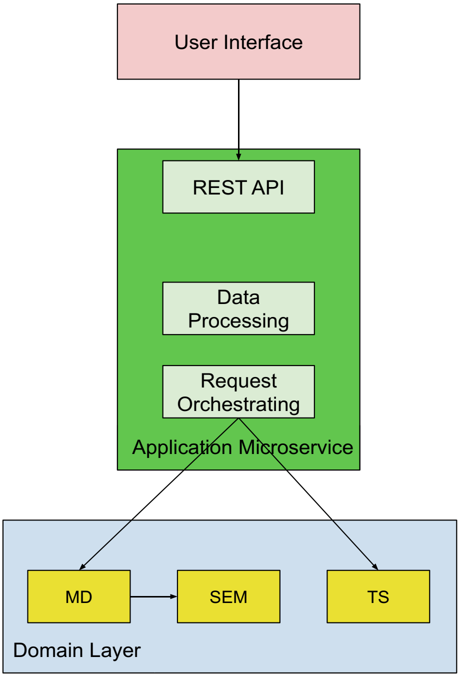


Fig. 3 Structure of the application microservice in collaboration with user interface and domain layer

of assigning the mentioned tasks to the user interface, an additional application layer is included being explicitly responsible for application-specific tasks (e.g. creating particular data formats for the user interface) that are independent of the components used in the user interface.

Figure 3 shows the structure of the application microservice. It consists mainly of three components. First, there is a REST API which provides endpoints encapsulating a certain set of functionality. In the context of this architecture, one endpoint is used to update the state of the whole application. Within this REST API which can be called by the user interface, certain other functions and classes are called containing the particular logic required for data processing. So on the one hand, there is a functional block which contains all necessary steps to perform the data processing and to create the appropriate data format. On the other hand, the data for this tasks is provided by the domain services. The requests to the domain services are orchestrated in a third part of the application microservice.

4 Evaluation

In the evaluation section, two different applications are presented and their purpose and functionality are described.

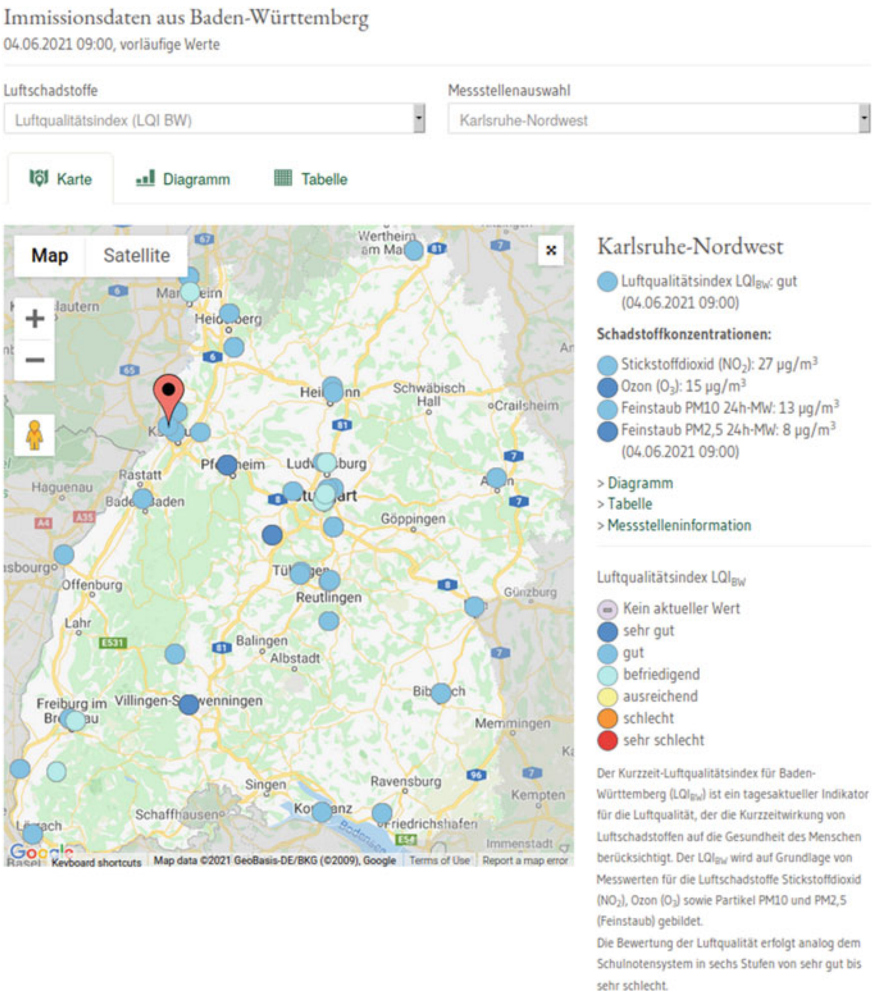


Fig. 4 Screenshot of the air measurement application [in German]

The first application that is shown in Fig. 4 is a measurement net application with the purpose to visualize data from measurement sensors which measure different air pollution parameters. The application allows the visualisation of different parameters, such as particular matter, nitrogen dioxide and ozone. Additionally, an air quality index can be selected which provides a calculation of the air quality based on the formerly mentioned air pollutants. These parameters can be selected by the corresponding dropdown menu ("Luftschadstoffe" → "air pollutants").[1] Similarly, the location of a measurement station can be selected by using another dropdown

[1] Pattern for die translations: "german term" → "english term".

menu (“Messstellenauswahl” → “choice of measurement station”). Alternatively, the location can also be selected by clicking a measurement station on the map component provided by the application. After choosing parameters for both menus, a legend appears right to the map, providing particular values that have been measured recently. Furthermore, the legend explains the colours of the measurement stations on the map, which provide an evaluation of the recently measured value, leading from very good (“sehr gut” → “very good”, colour deep blue) to very bad (“sehr schlecht” → “very bad”, colour red). Furthermore, there are two different views that can be selected. On the one hand, the diagram (“Diagramm” → “diagram”) view provides an overview over measured values of one measurement station from a certain time period (e.g., last seven days until today). On the other hand, the table (“Tabelle” → “table”) view displays the measured values of all stations in a table which makes comparing values from different stations more comfortable.

This application does not use application microservices (further referred to as AMS), which results in a high network load for clients who use the application. Additionally, after making all required requests, the user interface (further referred to as UI) still does data processing tasks which mainly means generating data formats for the different components. On the one hand, this process violates the MVC pattern [8], as the view components should not implement any application logic, nor should they do any processing tasks on the data they visualise. On the other hand, the high network load combined with the mentioned processing tasks slow the application and lead to an increased initial loading time.

The second application presented in this context is a dashboard application of the EnergyLab 2.0, a research project at Karlsruhe Institute of Technology (KIT), which is used for the evaluation of the concept presented in Sect. 3 and which shows the advantages of the concept compared with the formerly presented environmental application. Figure 5 shows the dashboard application with its different components. It has several components that visualise data of different energy storage systems and energy plants that make use of renewable energy technologies, e.g. a solar park, a redox flow battery and a wind park. They retrieve their data from sensors that are located in the field (in the area of KIT) which are connected to a backend described in the following. Furthermore, the user interface contains one or more data source components. Data source components are located in the background and therefore, they are invisible for the client. Their purpose is the fetching of data. This dashboard application is the user interface of the architecture presented in Sect. 3.

The domain layer (backend) consists of the microservices presented and described previously. Additionally, the application uses an AMS that orchestrates requests to the domain layer and creates a suitable data format for the UI based on the data collected from the domain microservices.

The main advantage of using an AMS in this architecture lies in the cost reduction of requests made by the UI. Requests can have different costs, based on the physical distance between the corresponding devices and on the data they contain. For example, a request from one device to another device in the same network is relatively cheap, while a request from one network to another one is more expensive. In the following, AMSs are always considered to be in the same network as the correspond-

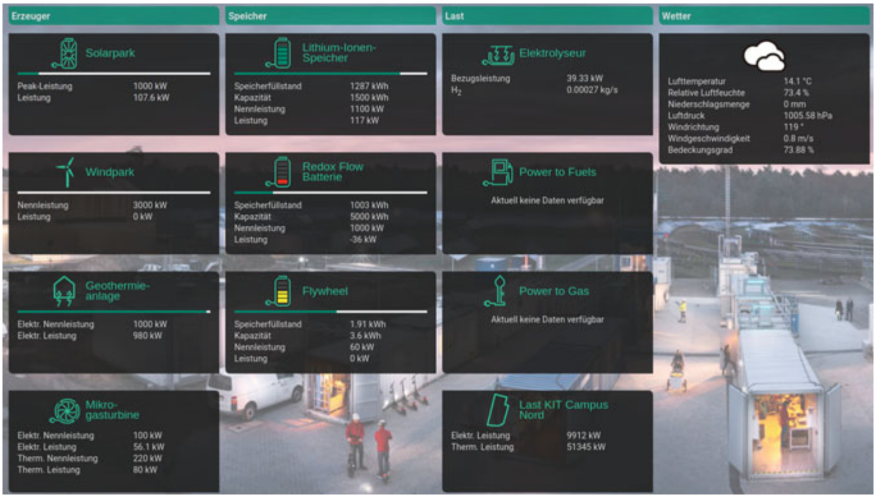


Fig. 5 Screenshot of the Energy Lab 2.0 dashboard

ing domain layer, which makes requests between UI and AMS more expensive than requests between AMS and the domain layer.

Moreover, there are different strategies how the UI is updated. As described previously, the UI consists of components that can potentially be dependent on different backend (domain layer) resources in the domain layer. Therefore, there are various approaches how the data in the components can be updated. To evaluate the previously presented concept of using AMSs, two different approaches are considered. The first one is a synchronous updating strategy. In this case, the UI components are updated together by using one data source component. The second strategy is an asynchronous update, where the components can be updated independently of each other. Each strategy results in a different total number of requests that is required to accomplish the update. Figure 6 shows the number of requests in the synchronous case without (a) and with using an AMS (b). The bold arrows show expensive requests, while the slim arrows show cheap requests. As an example scenario, an update of the four dashboard components solar park, lithium ion battery, wind park and weather is considered. While the solar park and the wind park components show the currently injected electrical power, the weather component shows local weather data and the lithium ion battery component shows the current level of power the battery still has, as well as the maximal capacity, the nominal capacity and general power.

The dashboard depicted in Fig. 5 implements the synchronous update strategy and uses an AMS. Compared to the case without a such one, the number of expensive requests is higher, as the data source has to request all necessary resources by itself, which results in a potentially higher number of expensive requests, depending on the number of resources that have to be requested. In the case with AMS, the UI's data source component has to make only one expensive request per update.

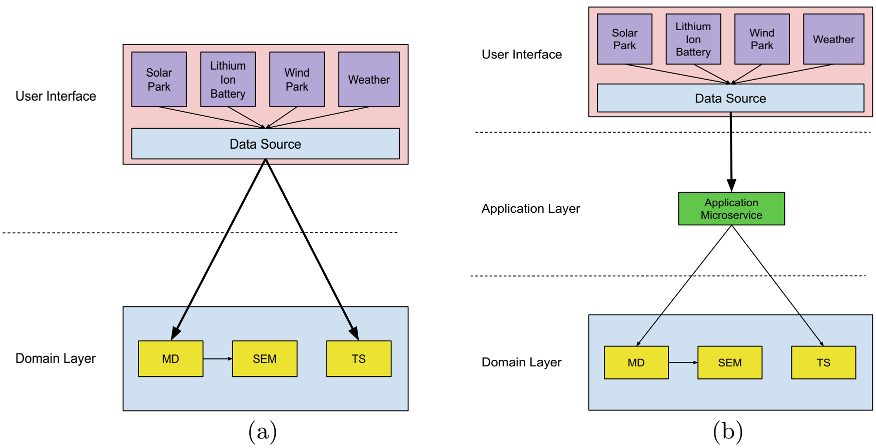


Fig. 6 Possible synchronous request orchestration without (**a**) and with AMS (**b**)

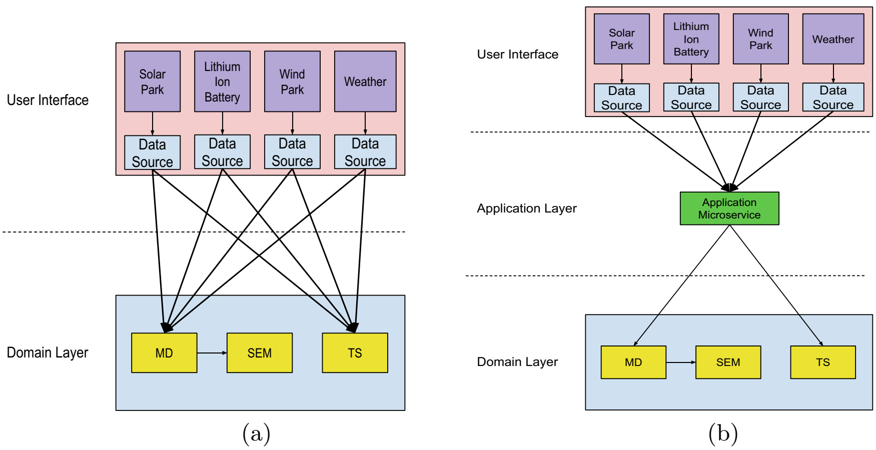


Fig. 7 Possible asynchronous request orchestration without (**a**) and with AMS (**b**)

Figure 7 shows the same scenario by implementing the asynchronous updating strategy. Although this updating concept is not used in the dashboard application shown in Fig. 5, it still can be reasonable, especially when there are components that need more frequent updates then other ones within the same application. In this case, the components are updated individually, which means that every component has its own data source, as depending on the update frequency, fewer data sources may become a bottleneck in the updating process. In this scenario, the number of expensive requests is higher as the data sources may have to request different resources from the domain layer, depending on the corresponding requirements of the requesting component.

Table 1 Number of expensive requests for the different updating scenarios

Type	Number of requests
Asynchronous update without AMS	8
Asynchronous update with AMS	4
Synchronous update without AMS	8
Synchronous update with AMS	1

As previously stated, the expensive requests are much more significant than cheap requests in terms of network connection. Therefore, the cheap requests are neglected and the focus of the following description lies in the expensive requests. To evaluate the presented concept from Sect. 3, it is assumed that each UI component accesses different resources from the domain layer which is the worst-case scenario as it requires the biggest number of expensive requests. Table 1 shows the number of expensive requests for the different scenarios. For the asynchronous cases, the usage of an AMS reduces the number of expensive requests by 50%, since it makes cheaper requests to the domain layer. In the synchronous case, without an AMS the UI has to make 8 requests, since it is assumed that each UI component requires different resources. By using an AMS, the number of expensive requests can be reduced by 87,5%. This is, when considering only the value, a great reduction of the network load. Still, it has to be mentioned that the increase of performance is most significant when the UI has to be updated frequently. When only sporadic updates are required, the saved costs and the relative increase of performance decrease. Regarding the two presented use cases from Figs. 4 and 5, it can be stated that the usage of the presented concept leads to an increase of performance, as regular updates are necessary in both cases. The result is a significant reduction of data transfer and therefore, a reduced network load for the client using the application. By achieving this reduction, and consequently by using the proposed architecture, the power consumption of client devices can be reduced.

5 Conclusion and Outlook

The present article describes the usage of application microservices, originally derived from the application services from the onion architecture [2] within a software architecture that formerly encapsulated application-specific tasks in its user interface. In this concept, the application microservices perform application-specific tasks as well as tasks usually executed by a gateway component, which are the orchestration of requests. By using this design pattern, the number of requests sent simultaneously from the user interface to the backend (domain layer) is reduced significantly. Additionally, data processing tasks are encapsulated in the application microservice to a high degree, which results in an improved performance for the the application. However, the processing efforts transferred from the user interface to the backend have to be made available on the server infrastructure. Moreover, two real applications (one from the energy domain, one from the environmental domain) are shown and discussed and the benefit of the provided concept is stated in their corresponding context.

Future tasks based on the presented concept are the design of a formal engineering process integrating the modified pattern into a microservice architecture.

Furthermore, the effort required to implement application microservices has to be examined. There are scenarios where many of them are required while the tasks they perform are relatively similar, so there can be a certain overhead that has to be compared with the benefit that is originally created by the usage of application microservices.

References

1. Braun, E., et al.: A generic microservice architecture for environmental data management. In: International Symposium on Environmental Software Systems (ISESS), pp. 383–394 (2017)
2. Palermo, J.: The Onion Architecture (2008). https://jeffreypalermo.com/2008/07/the-onion-architecture-part-1/. Accessed 16 July 2021
3. Hippchen, B., et al.: Designing microservice-based applications by using a domain-driven design approach. Int. J. Adv. Softw. 432–445 (2017)
4. Evans, E.: Domain-Driven Design: Tackling Complexity in the Heart of Software. Addison-Wesley (2004)
5. Fowler, M., Lewis, J.: Microservices—definition of this new architectural term. https://martinfowler.com/articles/microservices.html. Accessed 25 May 2021
6. Newman, S.: Building Microservices. O'Reilly Media Inc. (2015)
7. Braun, E., et al.: A Lightweight Web Components Framework for Accessing Generic Data Services in Environmental Information Systems. From Science to Society, pp. 191–201 (2017)
8. Fowler, M.: GUI Architectures (2006). https://martinfowler.com/eaaDev/uiArchs.html. Accessed 9 June 2021

A Literature Review to Analyze the State of the Art of Virtual Power Plants in Context of Information Security

Erfan Koza and Asiye Öztürk

Abstract Compared to conventional power plants, Virtual Power Plant is a new electrotechnical and information technology concept that is used for the interdisciplinary merging of electrical engineering and information technology for the central control and monitoring of decentralized renewable energy systems and components. A formal presentation of definitions and consensus building is seen as an instrument to specify the practice-oriented design of Virtual Power Plants. The present paper focuses on the Virtual Power Plants as key factor for sustainable energy—especially in its common understanding of terms. This compilation serves as a mutual knowledge base and function because of the "Shared Conceptualization" for the specialist user within the knowledge domain "Information technology" and "Electrical engineering." An Event-driven Process Chain is created for this purpose. In order to obtain such an Event-driven Process Chain, a qualitative analysis based on semi—systematic literature review led to a selection of suitable term definitions that are contained in the scientific databases. This investigation finally leads to the compilation of entities and their mutual relations, which are required to create the model. The conception of an Event-driven Process Chain is seen in this context as a contribution to the specification and structuring of the discussion about Virtual Power Plants. A potential contribution is, the ability to map a field of research, synthesize the state of knowledge, and create an agenda for further research in context of information systems and information security research.

Keywords Virtual power plants · Sustainable energy systems · Information security · Cyber security · Event-driven process chain · VPP · CVPP · TVPP

E. Koza (✉) · A. Öztürk
Hochschule Niederrhein, University of Applied Sciences, Mönchengladbach, Germany
e-mail: Erfan.Koza@hs-niederrhein.de

A. Öztürk
e-mail: Asiye.Oeztuerk@hs-niederrhein.de

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_4

1 Introduction

The efforts of the Paris Agreement [1] to reduce greenhouse emissions related to electricity generation have been leading to a fast increase in the deployment of renewable energy systems [2].

The increasing share of electricity generation from renewable energy sources creates several issues on power electricity system especially as regards demand and generation balance. In order to maintain power system stable new frameworks are available now. One of these new concepts is the use of Virtual Power Plants (VPP). In the relevant literary works, different perspectives and interpretations of the term VPP are given. Numerous contributions to VPP of the professional associations and institutes underpin this impression (Table 1).

This differentiated interpretation can essentially be associated with the respective assignment of the technical and logical functionalities that a VPP is able to perform. In this context there are partially isolated perspectives or perspectives mixed in different hybrid models, so that a different definition perspective will result from depending on the conceived understanding [3]. Thus, this paper attempts to identify the common intersection as well as the discrepancies between the individual scientific interpretations and to outline these in the sense of a standardized definition to arise an easily understandable explanation. Furthermore, the relevant properties that can be assigned to the respective definition descriptions should be determined. This differentiation is based on the functionalities assigned to the respective VPP. In analogy, the example of automobiles can be used here [4]. These are differentiated in their characteristics so that a visual and functional differentiation can be made between the individual entities [5]. Nonetheless, all entities fulfill the same purpose to guarantee private or public mobility. Can this point of view be transferred to the VPP or do VPP fulfill a different type of objective depending on the assigned functionality? Furthermore, a coherent denominator should be defined on which the definition of each VPP should be based on.

Table 1 Database and the number of selected works

Target	Language	Timeframe	Total number	Number of «elected for semantic analysis
ACM Digital Library	English	2015–2021	79	62
AIS Electronic Library	English	2015–2021	19	12
IEEE Xpktre Digital Library	English	2015–2021	15«	149
EBSCO	English	2015–2021	15	12
wtso	German	2015–2021	121	75
Springer Link	German	2015–2021	156	112
ECONBIZ	German	2015–2021		2
Total X				424

In addition, from the theoretical consideration (literature review) the possible gaps should be identified, which are also necessary in the area of security technology, explicitly in the sense of information security, as a key figure for the secure implementation and operation of VPP. A potential contribution is the ability to map a field of research, synthesize the state of knowledge, and create an agenda for further research in context the information system and information security research.

2 Research Methodology

It is an essential part of all scientific research activities to build the research by relating it to research results. The process of a literature review can be defined as a structured way to analyze and synthesize previous works [6]. A literature review also gives an overview of interdisciplinary areas. Here results are synthesized so that it can be shown in which areas in-depth research is required, which in turn is essential for the creation of theoretical frameworks and the conception of models [7].

In a first phase "design", for a scientific and reproducible process, this paper uses the methods of semi—systematic literature and in order to select the relevant and significant literary works from the various scientific databases. The semi-systematic approach describes a process in which different research groups from different disciplines consider the topics. In addition to the objective, studies are also listed here that show how a topic has developed over time or how it has progressed over various research traditions and bring them together using meta-narratives instead of effect size [8]. With the help of this analysis, certain topics, theoretical aspects, or common issues within a discipline or components of theoretical concepts can be pointed out [9]. In this context, the shown databases in (Table 1) were examined (phase: conduct).

The focus of this paper is placed on whether the existing literary works, in addition to the general consideration of the VPP, also consider IT aspects such as computer science to describe the information security architecture/data models of VPP or the topic of IT security in general in context of VPP. The results (phase: analysis) of the literature research are mapped to the empirical research (modeling and concepts of an information security architecture). Therefore there are the following artefacts in this work (phase: structuring and writing the review): Overview of the definitional context of VPP, the Event-driven Process Chain (EPC) model that represents the IT foundation of a VPP in order to be able to implement the individual main functions on this basis, a mapping table that shows possible functions of VPP and currently implemented industry-specific IT functions that are used by the operators of VPP and a main focus on the literature research that considers the relevance of IT security in context of VPP (Table 1) and gives an response to the following research question: "In addition to the electrotechnical research, are there empirical or theoretical research of information security also considered in context of VPP?".

3 Types of the Term "Virtual Power Plant"

3.1 Syntax of "Virtual Power Plant"

The focus of a VPP is on the conception and provision of novel and intelligent solutions for economically effective integration of decentralized energy generation systems (Distributed Energy Resources (DER)) in the planning, operation, and development of intelligent electricity networks [10]. The VPPs describe power plants that consist of several generation units, responsive loads or electricity storage systems and feed the generated electricity into the power grid in a controlled and bundled manner. In this way, they can make an important contribution to the energy transition. The switch from conventional energies (coal and gas) to renewables (diversity of technology and locations) is also associated with challenges. On the one hand, it is no longer just a few large power plants that produce electricity (Fig. 1), but many decentralized energy generation systems across power generation landscape. On the other hand, the electricity feed-in from renewable energies fluctuates strongly (weather conditions), [10].

A singular generating unit connected to a distribution network cannot offer efficient and profitable capacity, reliability, flexibility, and controllability in a liberalized and an open electricity market due to its unfavorable proportions in terms of

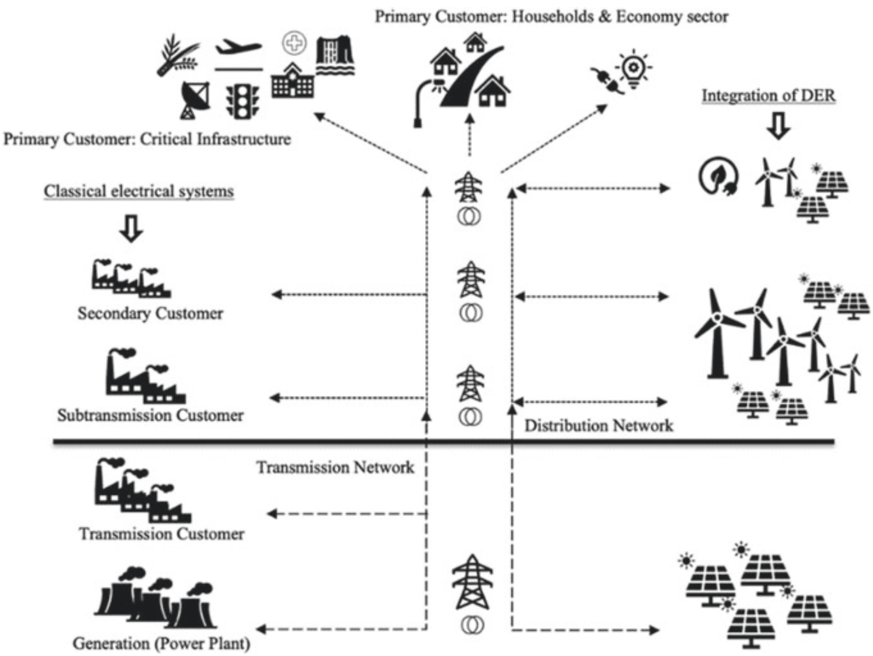


Fig. 1 The revolution in power grid with distributed energy resources

expenditure and income or meet the grid code requirements for power systems and networks' security (inter-temporal dependencies). So far, conventional power plants have stepped in when wind, sun or biomass cannot cover the electricity demand. Ultimately, their share should continue to increase during the energy transition. However, there can only be a secure power supply if the power grid is stable [11]. This is exactly where "VPP" come into play as a complex generating plant. Like a control center, they bundle the electricity from many small producers and feed it into the electricity power grid in exactly the amount that was previously agreed upon when the electricity was sold. This process requires complex planning and control incorporation. For example, in order to be able to predict as precisely as possible how much electricity a VPP can feed into the power grid, the amount of electricity available from the connected producers must be known.

This requires a monitoring system that monitors each individual generator in real time and knows exactly which systems are switched on and ready for operation and how much power they can currently provide. In the case of solar and wind power plants, e.g., the weather data must be taken into account in order to arrive at precise forecasts. The challenge is then to develop a strategy for optimal operation of the VPP, based on the use of the modern transmission techniques and IT. In order to ensure the power supply at all times system services are required, e.g. to keep frequency and voltage in the power grid stable. VPP can also contribute to this, e.g. by providing control power for frequency maintenance. This has already been successfully demonstrated in some scientific studies [12, 13]. At a higher level, a VPP can be defined as follows: A VPP is a flexible representation of a portfolio of DER that optimizes the economic value of the energy produced and offers reliability and capacity for ancillary services. A VPP would aggregate many diverse types of DER, distributed system operator (DSO), transmission system operator (TSO) and market simplify each unit's technical and economical parameters in order to be characterized in terms of a conventional centralized power plant set of parameters, such as active and reactive power capacity, scheduled generation, ramp rates and price. A commercial VPP (CVPP) represents a DER portfolio for the aggregation of individual DERs, which is used for effective integration and thus for participation in energy markets (Fig. 2).

This creates the possibility of integrating the DER in the analogy to the regular energy producers within the transmission network. This approach reduces the risk of imbalance, which is associated with the sole operation of a DER on the market and which is associated with the sole operation on the market and offers the advantages of the diversity of resources and the increase in capacity achieved through aggregation (bundling of the individual generation plants) [14].

In this context, the DER bundled in a CVPP are able to achieve economies of scale in market participation and benefit as a whole from information on market participation in order to maximize sales opportunities [15]. Here, the CVPP can be mapped in the system under certain contractual constellations from every geographic region and integrated into the markets. In cases where there is a contractual restriction, a CVPP can still represent DER units from different locations, with the clustering of resources

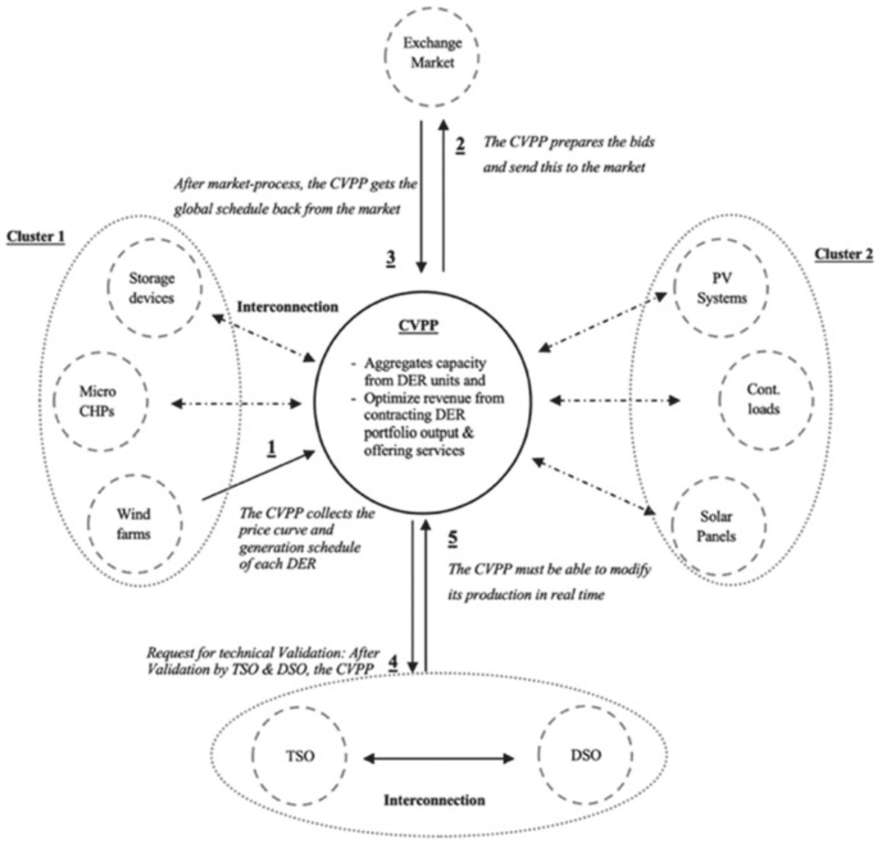


Fig. 2 A context of a CVPP

still taking place by location, which leads to the formation of a number of DER portfolios to be carried out the connected location are defined. For example, this scenario is foreseen in the markets with marginal location pricing and in markets where a zonal approach is mandatory for participation. Here, the CVPP can be mapped in the system under certain contractual constellations from every geographic region and integrated into the markets. [16]. In a commercial context, the VPP provides the following aspects:

- Visibility of DER in the energy markets,
- DER participation in the energy markets
- Maximization of value from DER participation in the markets and
- Frequency response and standing reserve services, to offer the opportunity for long term contract between TSO and aggregator for power reserve available in predefined time frames.

Figure 2 summarizes CVPP activity. After verifying the contractual and market law constellations, the CVPP transmits the confirmed DER contract plans and corresponding offers to the technical VPP (TVPP), [17]. Each DER integrated in the portfolio transmits the information (e.g., operating parameters) to the CVPP. This information is aggregated centrally in order to generate a singular VPP profile, which represents the total capacity of each individual DER individually and collectively. By using intelligent trading software, the CVPP is now able to optimize the portfolio's sales potential. This CVPP role can be assumed by several market players, including energy suppliers, independent third parties such as operators of VPP. DER is free to choose a CVPP to represent them in the wholesale market and in relation to the TVPP. The TVPP provides the following aspects [17]:

- Visibility of DER to the system operator(s)
- DER contribution to system management activities
- Optimal use of DER capacity to provide technically feasible system services incorporating local network constraints.

The TVPP aggregates and models the reaction properties of a control and monitoring system, which arises from the information technology conclusion of DER, in order to make the controllable loads and network states visualized and manageable in a single electrical-geographical (network), those of TSO and DSO would not be easy to implement. Figure 3 summarizes the TVPP activity. The hierarchical structure of the TVPP aggregations can therefore be systematically characterized in the low, medium and high voltage areas of a transmission and distribution network. The TVPP aggregates the operational items, parameters and cost data of each DER on the network along with detailed network information (topology, restriction information, etc.): Performing TVPP activities requires local network knowledge and network

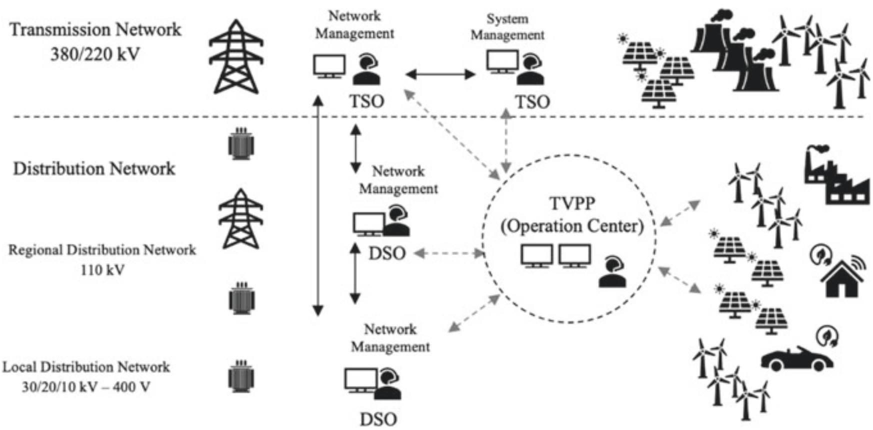


Fig. 3 A context of a TVPP

control functions. The DSO is best placed to perform this role. With this TVPP function, the DSO role can evolve so that the distribution network is actively managed in the same way as a transmission network operator [17].

Another alternative to this is the integration of the "Operation Center", which acts as an assistance system and thus as an instance between the DER and the DSO in order to be able to control and monitor the aggregated network conditions. The TVPP activities include the monitoring and control of the network aggregation of a defined distribution network area in order to be able to guarantee the specifications of the DSO and TSO for a stable and secure DER connection [18].

As a result, a VPP can be used for two different roles and responsibilities: CVPP, that elaborates the economical optimization of a VPP portfolio for the open electricity markets, and a TVPP that ensures optimized and safe operation of the power grid, taking into account physical constraints and potential services offered by VPP.

3.2 Reason for VPP Ambiguity and Consistency

Thus it can be stated that the reasons why VPPs are carried out in ambiguity are diverse: On the one hand, a VPP can be defined as a fundamental technical concept that can be traced back and projected on the meta-level to several differentiated areas of investigation with different focuses (Figs. 2 and 3). On the other hand, as already mentioned, this general concept can be deepened in terms of granularity in different technical directions without losing any fundamental consensus. The ambiguity of the term "VPP" relates to its parts of the word "Virtual" or "Power Plant", as this brings together various questions into one term. If, from an IT point of view, an object can be defined as a whole (higher-level system) or as several parts (several combined sub-systems) at a certain point in time, the following question arise: Is a VPP as an assistance system that can be integrated into the existing system landscape of the TSO or DSO or is a VPP a self-sufficient and independent system?

Nevertheless, the different forms of a VPP, shown in Sect. 3.1, have their justified technical and organizational right to exist. In order to reduce the internal technical and organizational dependencies and thus also the complexity, the principles of "Separation of Concerns" or "Separation of Duties" are used as a rule to reduce the individual areas and to generate a clearer organizational and technical task. In addition, with the use of Separation of Concern, the focus of the respective VPP expression is set on different results, which ultimately also guarantees the use of the specialist lists in the dedicated commercial and technical areas. Besides the aforementioned advantages, the separation of duties avoids the existing legal, contractual and regulatory conflicts of interest. Due to different actual, technical, and commercial (heterogeneous) data and time intervals, different ICT infrastructures with different software-technical solutions with different professional and personal know-how have to be used. Ultimately, this view can be implemented by separating organizational and technical tasks from a VPP into CVPP and TVPP. Since the term "VPP" follows different concepts, each of these concepts must be analyzed in its parts. These concepts (case-related) can be presented using a symmetrical presentation from the point of view of information and communication technology, whereby the following terms are opposed to each other: TVPP, CVPP, large-scale VPP, Central Control Computer (CCC), Operation Center (OC)), aggregation of DER. These terms show an association with the term VPP. For example, the term VPP is understood generally in relation to the higher-level centralization and IT virtualization of the DER in an ICT infrastructure from the IT point of view, while the association with CVPP describes the usefulness of a VPP by other actors and thus other activities are carried out to facilitate economic processes and efficiently counteracting challenges that have to do with the decentralization of the energy generation landscape [16, 24]. So, it turns out that without understanding the relationships between the descriptors and the system actors, VPP remains incomprehensible. The commonality of a CVPP and TVPP can ultimately be traced back to the information technology infrastructure that is used for centralizing, transforming, analyzing, and visualizing the decentralized data from DER, regardless of the task at hand.

4 Event-Driven Process Chain for a VPP

This paper draws on the knowledge that has been acquired within a design-oriented research project the basic properties and functions of a VPP and outlines this with the help of an EPC. In a processing sequence 00, the configuration and global setting values (limit ranges, limit values and default values) must first be initialized and determined. Overall, e.g., the basic settings for e.g., network planning, network operation, cost parameters, control settings, connection test settings, network monitoring settings and standard line types or as well as for other function must be defined (Fig. 4).

This data is checked for consistency and data errors, or data gaps are corrected through automated synchronization (data comparison). The sequence 01 offers various visualization and analysis functions for evaluating the data quality. The verification of the data can be divided into four steps. In the first step, the data from various source data systems are bundled and imported into the internal database. The import is carried out in a synchronized manner, only elements are checked that have been changed (added, changed, deleted) since the last synchronization run.

This data is checked for consistency and data errors, or data gaps are corrected through automated synchronization (data comparison), (Fig. 5). In addition, the results can be saved and downloaded using the "data export" function in the data analyzing (Table 2). Extended classic functions of a VPP are now listed in Table 2. These functions are assigned to relevant stakeholders who are essential for the process [19–23]. The list of functions and the assigned persons are not exhaustive. This means that a VPP can bring further functionalities and other actors such as state authorities, the Federal Network Agency or the Federal Office for Security in IT can be considered.

Fig. 4 Processing sequences 00—internal configuration

5 Information Security Architecture

From the information technology point of view a VPP can be defined as a centralized complex information and communication system to control over an aggregation of DER, controllable loads and storage devices. Its main function is to supervise, control, optimize and manage generation the electrical energy flow not only within the cluster, but also in exchange with a TSO, DSO and the market. For this purpose, a VPP must be able to record the resulting data in the DER as well as in other decentralized units in real time in order to analyze the aggregated data with algorithms, calculation models and data models and consequently to create suitable handling options and forecasts. In such a VPP, active control is achieved through an information and communication technology (ICT) infrastructure consisting of smart devices

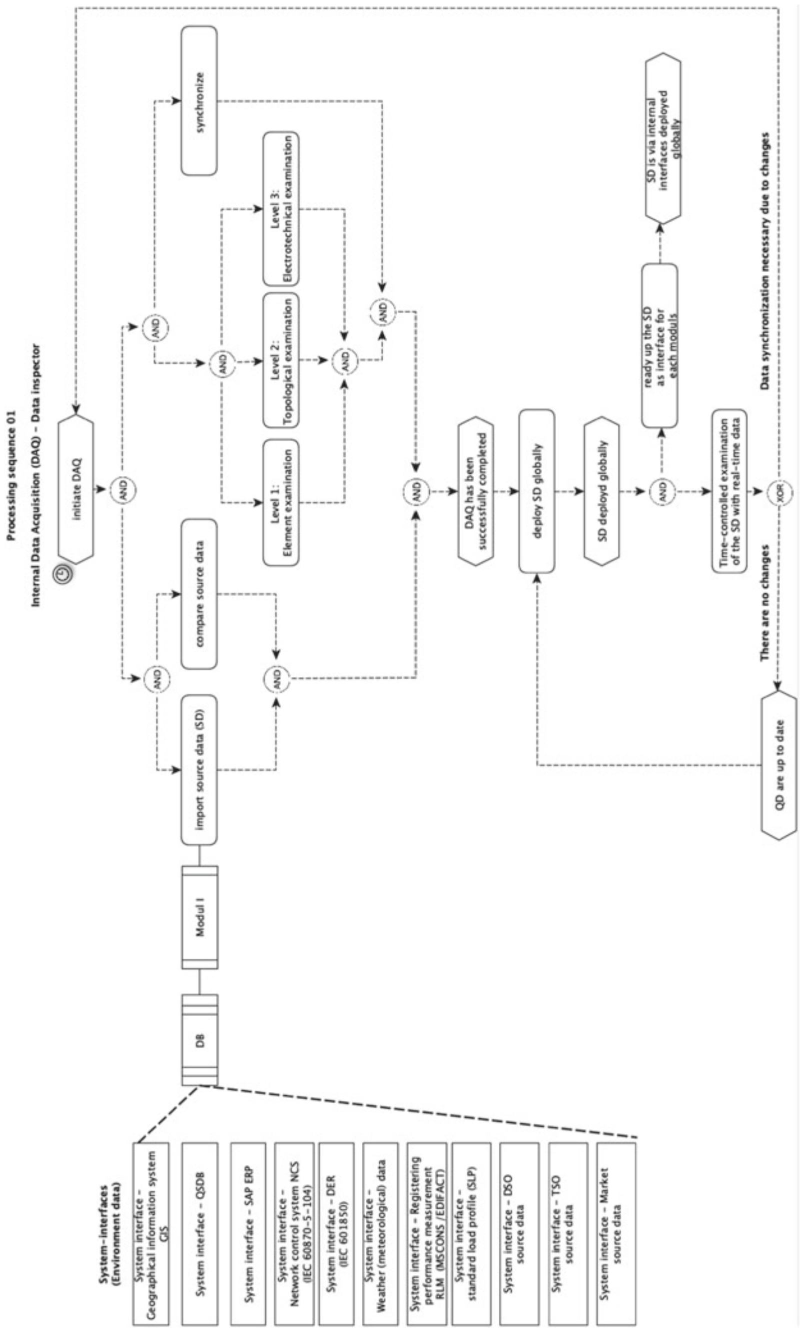


Fig. 5 Processing sequences 01—data inspector

Table 2 Function-asset-mapping

Function	Description	Assigned atlors					
		TVPP	CVPP	Maricet	TSO	DSO	DER-Owner
Trading	Market access, portfolio management		x	x			
Data analyze /visualization	Data from different systems are linked and checked for consistency and data error	x					
Network transparency	Power (low results and short-circuit Biwlyites for the network area arc permanently stored here	x			x		x
Network study	Scenarios for future supply tasks can be created and evaluated across the board for the entire own network. This enables potential bottlenecks in Lite network to be Identified al an early stage and long-term future strategies to br derived		x	x		x	x
Variant calculation	Future scenarios (e.g., additional loads, new lines) can be created manually for a single network bushes and to be evaluated in a non-registered basis	x					x

(continued)

Table 2 (continued)

Function	Description	Assigned atlors					
		TVPP	CVPP	Maricet	TSO	DSO	DER-Owner
Measurement daia analysts	Different types of measurement data can be displayed and analyzed transparently for the entire network area	x					x
Measuring devices placement	With the measuring device placement a selection of measuring device locations with minimal costs can be determined for recording Ibt network status	x					x
Network forecasl	Tbe network status is displayed based od forecast daia from loads or feeders	*					R
MoniloringConirol	Actual performance, storage suius and readiness of the systems is shown Live. Therefore, a precise load management ii possible	x	x	x			x
Flexibility management	Managing the flexibility by controlling tbe controllable local systems, DER. systems, loads and supporting the smart meter gateways	x	x	x	x	x	x

(continued)

Table 2 (continued)

Function	Description	Assigned atlors					
		TVPP	CVPP	Maricet	TSO	DSO	DER-Owner
Control energy management	Highly flexible parameterization, pooling of plants, Web interface for visualizing process information lor oool Dirtncrs	x	x	x	x	x	x
Optimization	Schedule optimization for consumers and generators. Ancillary services, demand response, extensive analysis, time series management, workflow control, modeling resources in detail		x	x			x
Scenario simulation	Simulation of future supply tasks. Scenarios are set up so that potential bottlenecks in the network can be identified. In addition, generating plants, new lines or loads can be determined		x	x		x	x

and smart meters, wireless and wired connections, a CCC, and software applications [24]. The network structure diagram (Fig. 6) of a VPP shows the primary data connections, which illustrate the interconnectivity of the VPP, TSO, DSO and DER. The recorded data communication paths only show the internal and external interfaces, which are essentially mapped in the Operation Technology networks (OT networks). The OT networks of the involved actors are segmented and structured in accordance with the applicable security guidelines, which means that access to other network areas or control systems is not possible. The exclusive interfaces for

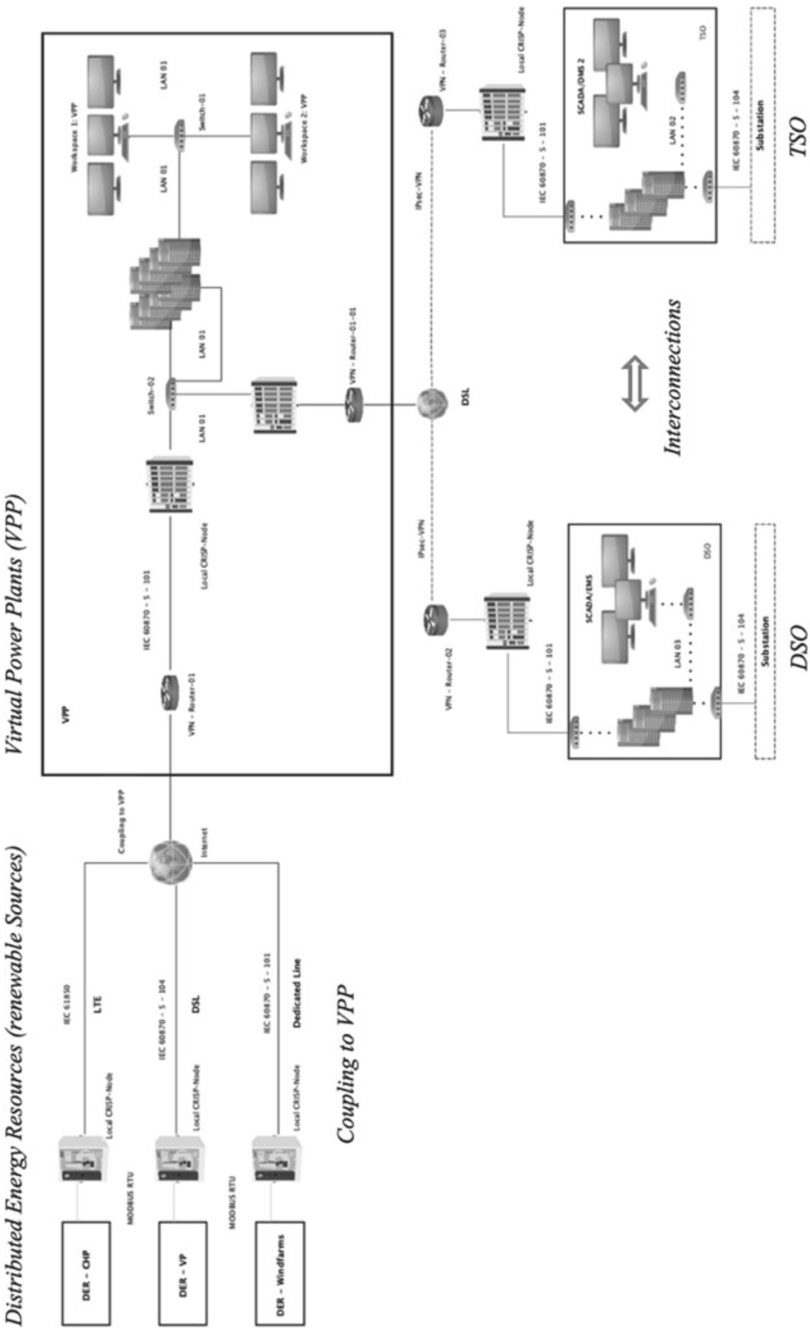


Fig. 6 A network structure diagram of a TVPP as a result from an empirical research project

the bilateral transfer of data in other dedicated networks (office networks and to other stakeholders) are not described here. This can be done in connection with the fact that the ETL processes are defined differently for certain predefined purposes. The extension of the interfaces therefore depends on the respective task that is to be fulfilled by a VPP. Figure 2 shows the base foundation of a VPP, on the basis the specifications (regarding commercial and technical tasks) can be set up. From the DER, TSO and DSO, measured values are generally extracted from the database archives, and sent to VPP using a database procedure via an IPsec VPN tunnel (e.g., X.509 Certificate (CA)) or via dedicated lines with different standardized data communication protocols. A self-signed root certificate authority is generated for the creation of the certificates, which exclusively uses IKEv2 (Internet Key Protocol Version 2) as a secure method for key management in IPsec-based virtual private networks. The telecontrol devices of the TSO and DSO, referred to as local CRISP-Node (Cross Industry Standard Process for Data Mining) in the diagram (Fig. 5), have two separate network interfaces, between which no routing takes place. One interface is used to connect the respective control center of TSO and DSO, the other for coupling to the VPN-Router. The telecontrol devices of DER have a firewall based on "*DENY-ALL*" strategy. There are two basic concepts for configuring a firewall, the "*ALLOW-ALL*" and the "*DENY-ALL*" strategy. In the case of an "*ALLOW-ALL*" strategy, any communication via the firewall is permitted in principle, specific services or stations can be assigned restrictions. With a "*DENY-ALL*" strategy, all communication is in principle blocked, individual stations can be excluded from this block. Both strategies can provide a sensible approach to the conception of the firewall in real environments. In the direction of an Internet connection of DER, TSO and DSO, it is advisable to use a deny-all concept in order to be able to control access to the insecure medium of the Internet completely at IP level. Using this concept, only services that are explicitly permitted by the administrator of the firewall can be used. The configuration thus offers the lowest risk of accidentally allowing unwanted communication. The data transfer between the network nodes is established via the integrated VPN Client of the local CRISP-Node of DER, TSO and DSO. The CRIPS-Node of DER, DSO and TSO independently set up encryption via IPSec to the VPN gateway in VPP.

The configuration database of the VPP is encrypted and password protected to ensure the integrity and confidentiality of the data and IT systems. A "role-based access control" (RBAC) is implemented in the configuration database of the VPP. At RBAC, access rights are assigned based on a defined role model. The defined user roles abstract the work processes in an organization and therefore vary from company to company. Furthermore, accesses from IEC 60,870–5-104, 101 and IEC 60,850 from unknown addresses are rejected. All services of CRISP-Node that are not required are switched off, such as the integrated web server or Dynamic Host Configuration Protocol. In this context, the data and information generated in the decentralized units (DER) can be centrally collected in VPP, adjusted, extracted, and released for further analysis.

6 Gaps of Information Security Aspects of VPP

In addition to the results of the information security architecture of a VPP with the associated possible functions, the security factor plays a central role. While all literary works identified and analyzed so far have dealt generally with the theoretical and empirical research of a VPP in terms of electrical engineering and functional safety (gained knowledge from literature review), the subject of information security and IT security does not receive significant attention. Some of the related works are in [25–28]. In the absence of IT security research in this area, it must be resigned that the use and integration of the new technologies can only be used efficiently in the long term if they have a certain acceptable level of security both in society and in politics. In particular, the VPPs show a high degree of ICT dependency and must therefore be protected with adequate technical, organizational, and individual security mechanisms in the sense of securing and maintaining the electrical energy supply. In terms of network technology, the VPPs are to be understood as central intelligent network communication nodes that are connected to the decentralized DER or other nodes via the defined interfaces, such as VPN technologies, in order to transfer certain data from the decentralized units in a bilateral data communication network and exchange information (e.g. metadata and switching commands). Due to the inevitable connection, which is made global and open by the medium of the Internet (the connection of the DER can be carried out independent of location in the sense of a CVPP, e.g.), a multitude of attack vectors emerge that are within the designed IT network in the sense of any time any places can be run. In particular, such cyber physical systems, which have an “always on status”, show a major weak point. In addition, the classic preventive security mechanisms from office IT can only be partially and effectively transferred to industrial production and control systems.

Many factors play a role here, such as the integration of the appropriate hardware technologies, network structure with the integration of security gateways, correct configuration of the network nodes, network access control, change management, patch management, identity access management, intrusion detection systems, Risk Establishment, Business Continuity Management, Incident Response Management up to cryptographic concepts and encryption principles play a key role, which are used for the sustainable and effective protection of the basic information security. If, e.g., the decentralized network nodes are set up in the sense of human–machine interfaces without physical perimeter protection (access control), logical access and access control (authentication and authorization and role authorization concepts) and can then sometimes lead to physical sabotage an unauthorized person gains access to a wind turbine. The attacker can now connect to the central control unit of a VPP via the open and logically unmonitored ports (early detection systems, use of device certificates to authorize network partners) in order to compromise the entire network. The intended remote access, which is used for the maintenance of the components, can also be used via exploit attacks, the Internet and the existing and not timely discovered vulnerabilities. Such scenarios lead directly to serious problems due to the electrotechnical and IT networking of the systems, networks, and actors in the energy industry. Due to the existing ICT dependencies of a VPP, such system impairments can lead to long-term system failures, under which the affected VPP and its connected DER clusters can suffer a blackout. However, this is not a local problem. This can also affect the upstream and downstream processes of the DSO.

The local failure effects can ultimately have negative effects on the regional power supply and sometimes lead to the impairment of the regional power supply, where millions of people can ultimately be affected by such failures. If such disruptions are a complex, planned and coordinated attack scenario, several networks can be affected by the so-called domino effect. The failure of the electrical power supply can also continue to spread to other critical infrastructures and sectors (health care, drinking water supply, finance, traffic engineering, etc.) due to the so-called cascade effect, which can lead to humanitarian and existence-threatening bottlenecks. Therefore, considering the aspects of information security and IT security in the sense of a holistic approach, taking into account the triangology of information security (technology, organization and human factor), plays a key role in establishing and accepting such technologies, especially with regard to trust in digitization to reinforce.

7 Conclusion

A VPP is not only an electrotechnical but also an IT concept that plays a central and significant role in the sense of digitizing and automating the integral control processes (cyber physical systems and IoT) and in the sense of Smart Grids. Installing the VPP plays a central task in the context of the energy transition. On the one hand, the VPP serve with their commercial characteristics for economic prosperity and efficient and relevant marketing of the within the new energy landscape, so that the energy, which is ultimately generated by renewable energy sources, is cost-effective and profitable at the same time. And on the other, the VPP will be used to control and monitor the electrical network and DER and ensure that despite increasing complexity (the decentration of the energy production plants, progressive rising number of network participants, etc.), the future electricity grid in the sense of the uninterruptible power supply remains stable. For this purpose, this electrotechnical and IT concept is carried out in two essential expressions with CVPP and TVPP [29]. However, two essential aspects must be further specified in this context. The first aspect includes a stapling of the technical procedure for the construction and operation of a VPP, which have already been carried out scientifically informed by means of numerous empirical studies (laboratory and field examinations, simulations, case studies and prototyping) [30]. The second aspect, which received less compliance, is the scientific discussion with the topic of information security and IT security. With regard to the research question the results of the literature analysis showed that the topic of information security receives little attention in the existing literature and is therefore neglected. Although the relevance of the topic is listed in some papers, there are few approaches and concepts on how the smart grid should be secured in the future. A safe and reliable power grid can only be achieved if security is also guaranteed.

Dependencies on ICT must also be considered so that the complex systems can be made manageable, because the fact that the system landscape is connected to a central connection with the DER means that the data connection is largely via the medium of the Internet. This means that there is no self-sufficient network, so that many dangers such as trojans, ransomware etc. are possible (colonial pipeline was blackmailed by a hacker group called DarkSide within a ransomware attack). A failure of these ICT with DER can lead to a cascade effect (domino effect) and a regional network area failing. In the long term, this can result in a blackout and the Critical Infrastructures can no longer provide their services. The practical and scientific responsibility of the designed VPPs must therefore be explored in the sense of ensuring and maintaining the ICT infrastructure. Currently, IT security topics for TSO and DSO can be dealt sustainably and efficiently by the implementation of an information security management system based on ISO/IEC 27,001 “Information security management system—Requirements” and ISO/IEC 27,019 “Information security controls for the energy utility industry.” However, such regulatory units and implementation models for VPP need to be examined to lead to these rules and models for the desired result, increase in the resilience of IT systems and reducing the vulnerability of IT systems. The personnel and technical capacities as well as the preventive, reactive, corrective and detective information security mechanisms also play an essential role that require a primary investigation. New safety implementation models, including a safety-related shared-concept, can be a substantial cornerstone. Not only the security implementation models (prevention), but also the corrective safety models, such as the attachment on a CERT requires a deeper examination. For this, too, explicit technical know-how must be used to effectively protect the ICT infrastructure and IT systems against the attacks from the Internet and against external vectors. In this context, the information security can be defined as a system indicator or key factor for the establishment and operation of such technologies.

References

1. United Nations. Adoption of the Paris agreement. In: Conference of the Parties (2015)
2. Farhangi, H.: The path of the smart grid. In: IEEE Power and Energy Magazine, **8** (2010)
3. Kuntschke, R., Specht, M., Van Amelsvoort, M.: Economic optimization in virtual power plants vs. stable grid operation—bridging the gap. In: 20th Conference on Emerging Technologies & Factory Automation (ETFA). Luxembourg, Luxembourg (2015)
4. Okpako, O., Rajamani, H., Pillai, P., et al.: A new performance index for evaluating community virtual power plant with domestic storage. In: IEEE Power & Energy Society General Meeting. Chicago, IL, USA (2018)
5. Fenix. A European Project Supported by the European Commission, FENIX deliverable D3.3 Report: Financial and Socio-Economic Impacts of Embracing the Fenix Concept Source: http://www.fenix-project.org. Fenix 09.20 (2009)
6. Baumeister, R.F., Leary, M.R.: Writing Narrative Literature Reviews. Ohio, 311–320 (1997)
7. Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br. J. Manage. **14**, 207–222 (2003)
8. Wong, G., Greenhalgh, T., Westhorp, G., et al.: RAMESES publication standards: meta-narrative reviews. In: MCD Medicine, vol. 11 (2013)

9. Ward, V., House, A., Hamer, S.: Developing a framework for transferring knowledge into action: a thematic analysis of the literature (2009)
10. Dulâu, L.I., Abrudean, M., Bicâ, D.: Applications of virtual power plants approaches. In: The International Conference and Exposition on Electrical and Power Engineering (EPE 2014). Iasi, Romania (2014)
11. Kieny, C., Berseneff, B., Hadjsaid, N., Besanger, Y., Maire, J.: On the concept and the interest of virtual power plant: some results from the European project Fenix. IEEE Power & Energy Society General Meeting (2009)
12. Fenix. Virtual Power Plant Concept in Electrical Networks. In: The 2nd Conference on Integration of Renewable and Distributed Energy Resources. Napa CA (2006)
13. Fraunhofer Institute for Energy Economics and Energy System Technology (IEE). Kombikraftwerk2, (2010–2013). Source: https://www.iee.fraunhofer.de/en/research_projects/search/2013/kombikraftwerk2.html
14. Kok, K., Akkermanns, H.: Transport network feasible solutions in market-based allocation of flow resources. In: Emerging Technologies in Robotics and Control Systems (2007)
15. Fenix. Flexible Electricity Networks to Integrate the expected Energy Evolution. Bilbao (2009)
16. Chalkiadakis, G., Robu, V., Kota, R., et al.: Cooperatives of distributed energy resources for efficient virtual power plants. In: The 10th Conference of Autonomous Agents and Multiagent Systems (AAAMAS). Taipei, Taiwan, pp. 787–794 (2011)
17. Pourghaderi, N., Fotuhi-Firuzabad, M., Kabirifar, M., et al.: Energy scheduling of a technical virtual power plant in presence of electric vehicles. In: The 25th Iranian Conference on Electrical Engineering (ICEE2017). Tehran, Iran (2017)
18. Kok, J.K., Warmer, C.J., Kamphuis, I.G.: PowerMatcher: multiagent Control in the Electricity Infrastructure. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems. Utrecht, Netherlands, pp. 75–82 (2005)
19. Essakiappan, S., Shoubaki, E., Koerner, M., et al.: Dispatchable virtual power plants with forecasting and decentralized control, for high levels of distributed energy resources grid penetration. In: The 8th International Symposium on Power Electronics for Distributed Generation Systems (PEDG). Florianopolis, Brazil (2017)
20. Lazaroiu, G.C., Dumbrava, V., Roscia, M., et al.: Energy trading optimization of a Virtual Power Plant on electricity market. In: The 9th International Symposium on Advanced Topics in Electrical Engineering. Bucharest Romania, pp. 911–916 (2015)
21. Powalko, M., Komarnicki, P., Rudion, K.: Enhancing virtual power plant observability with PMUs. In: The 5th International Conference on Critical Infrastructure. Beijing, China (2010)
22. Bagchi, A., Goel, L., Wang, P.: Generation adequacy evaluation incorporating an aggregated prohabilistic model of active distribution network components and features. IEEE Trans. Smart Grid, Singapore **8** (2018)
23. Lezama, F., Sucar, L.E., Munoz de Cote, E., et al.: Differential evolution strategies for large-scale energy resource management in smart grids. In: Proceedings of the Genetic and Evolutionary Computation Conference Companion. Berlin, Germany, pp. 1279–1286 (2017)
24. Heimgaertner, F., Menth, M.: Distributed controller communication in virtual power plants using smart meter gateways. In: IEEE International Conference on Engineering, Technology and Innovation (IECE/ITMC). Stuttgart, Germany (2018)
25. Stübs, M.: Student research abstract: IT-security in self-organizing decentralized virtual power plants. In: Proceedings of the 33rd Annual ACM Symposium of Applied Computing Hamburg. Germany, pp. 431–432 (2018)
26. Rosinger, C.: Informationssicherheit im Smart Grid, pp. 83–103. Springer, Berlin Heidelberg, In IT-Architekturentwicklung im Smart Grid (2012)
27. Lang, L., Dietmannsberger, M., Lücken, A., et al.: Modellierung und Bewertung einer autarken Notstromversorgung der Bevölkerung unterhalb der KRITIS-Schwelle unter Einsatz von Photovoltaik-Systemen. In: Nachhaltige Energieversorgung und Integration von Speichern. Hamburg, Germany, pp. 42–47 (2015)
28. Pavleska, T., Aranha, H., Masi, M., et al.: Drafting a cybersecurity framework profile for smart grids in EU: a goal-based methodology. In: Dependable Computing—EDCC, pp. 144–155 (2020)

29. Cheng, L., Zhou, X., Yun, Q., et al.: A review on virtual power plants interactive resource characteristics and scheduling optimization. In: The 3rd IEEE conference on Energy Internet and Energy System Integration. Changsha, China, pp. 514–519 (2019)
30. De Zotti, G., Pourmousavi, S.A., Morales, J.M., et al.: A control-based method to meet TSO and DSO ancillary services needs by flexible end-users. IEEE Trans. Power Syst. **35**, 1868–1880 (2020)

Artificial Intelligence and Sustainability

Rebound Effects in Methods of Artificial Intelligence

Martina Willenbacher, Torsten Hornauer, and Volker Wohlgemuth

Abstract Artificial intelligence (AI) is one of the pioneering driving forces of the digital revolution in terms of the areas of application that already exist and those that are emerging as potential. On the technical side, this paper deals with the energy requirements of artificial intelligence processes. It also identifies efficiency approaches in this sector. Increases in productivity often lead to an increased demand for energy, which is contrary to sustainability in terms of reducing CO_2 emissions. Therefore, it will be examined to what extent rebound effects can reduce the savings potential for energy in relation to methods of artificial intelligence and what the main factors of CO_2 emissions are.

Keywords Artificial intelligence · Rebound-effect · Resource and energy efficiency

1 Introduction

Methods of artificial intelligence have gained enormous popularity in recent years, also due to increasingly cheaper storage media and, in the medium term, graphic computing units. Artificial, multi-layer neural networks have received strong media attention and also achieved [1] a strong increase in popularity in the scientific sector within the last decade.

Particularly in the field of artificial neural networks, enormous amounts of data are processed based on non-linear functions, e.g., to allow the system to autonomously find correlations within the data set [2]. Convolutional neural networks (CNNs) for image recognition and multilayer natural language processing systems (NLPs) for speech recognition are particularly worthy of mention.

M. Willenbacher (✉)
Leuphana University Lüneburg, Institute of Environmental Communication, Universitätsallee 1, 21335 Lüneburg, Germany
e-mail: martina.willenbacher@htw-berlin.de

M. Willenbacher · T. Hornauer · V. Wohlgemuth
School of Engineering - Technology and Life, University of Applied Sciences HTW Berlin, Treskowallee 8, 10318 Berlin, Germany

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_5

Integer	
Add	
8 bit	0.03pJ
32 bit	0.1pJ
Mult	
8 bit	0.2pJ
32 bit	3.1pJ

FP	
FAdd	
16 bit	0.4pJ
32 bit	0.9pJ
FMult	
16 bit	1.1pJ
32 bit	3.7pJ

Memory	
Cache	(64bit)
8KB	10pJ
32KB	20pJ
1MB	100pJ
DRAM	1.3-2.6nJ

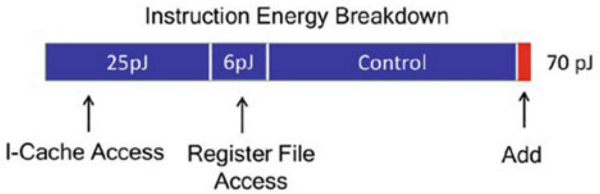


Fig. 1 Energy and time consumption for data access and movement [4]

The more extensive the training of the AI or the creation of the weightings, the more complex the calculations and the associated demands on the hardware. In general, this applies to Deep Learning, especially TPU, but also GPU. For the calculation of the algebraic basics, these processors are more suitable than a CPU [3].

Not surprisingly, the amount of material required increases with the same methodology if the time factor of the calculation is to be reduced. When using artificial intelligence methods, most of the power used is not required by the processing of the matrix multiplications themselves, but by accesses to the memory areas and the associated rewriting of the data to be processed. Using the example of an 8-core server CPU based on a 45 nanocenter infrastructure, it became apparent that both addition and multiplication processes require only a fraction of the energy of data accesses. Even the comparatively complex multiplication of floating-point numbers with a size of 32 bits requires hardly more than a third of the energy of the memory access to 8 KB in the processor's cache. If the data is in the main memory, the energy requirement increases by a factor of about 1000 [4] (Fig. 1).

As the size of the training data increases, the power consumption increases exorbitantly simply because of the necessary rewriting of the data.

2 Rebound Effects—Basics

As the amount of material and energy used increases, the question arises as to what extent the use of such a tool is purposeful. This question is even more pressing if artificial intelligence is to be used in a sector that is striving for greater resource efficiency with the help of these methods. If the increase in efficiency in the energy consumption of a sector leads to the savings achieved through the lower energy demand being reinvested elsewhere to consume more energy by other means, this is known as the rebound effect. Rebound effects are both a serious problem in consumer behavior and for production cycles when attempts are made to increase efficiency [5]. They can therefore occur in any form of endeavor to achieve energy efficiency. So too in the use of artificial intelligence methods. Due to the rebound effects existing at different levels, a wide range of efficiency-reducing concomitant effects arises. In general, a distinction can be made between direct, indirect, and macroeconomic rebound effects [5].

We talk about direct rebound effects when the energy savings of a product are reduced by the use or modification of the same product. So, if a workflow can be designed to be more energy efficient—but is subsequently carried out more frequently because it can now be afforded—the energy saving dwindles. One study assumes that the direct rebound effect reduces the energy efficiency of a product by approx. 10–30% [6].

If, on the other hand, the potential for energy savings is reduced through the use or modification of another product, this is indirect rebound. A more efficient work process is not used more often, but instead the money saved is used to enable and realize another production process. The same study concludes that indirect rebound effects reduce energy saving potentials by approx. 5–10%. The macroeconomic rebound effect broadly describes the credo of economic growth through energy efficiency: if less energy is needed for production within an industry, the effect can occur that a completely different industry benefits from the lower demand of the (initially seemingly non- correlated) industry through lower electricity prices and produces more itself. Similarly, a company could benefit from the supplier's more energy-efficient production and as a result increase its own production rate. The data on macroeconomic rebound effects currently seems to be rather sparse. Allan, G. et al. concludes in their study "UKERC Review of Evidence for the Rebound Effect" that a loss of more than 50% of the energy savings potential should not be considered exceptional and that backfire effects are also partially achieved [7]. Backfire effects mean that due to the side effects of energy savings, more energy is ultimately consumed than was saved by the energy efficiency achieved [7].

3 Energy Consumption and Greenhouse Potential of AI

In addition to energy efficiency, the emission of CO_2-equivalent greenhouse gases represents a serious environmental damage. How strong the greenhouse gas potential associated with AI processes is essentially depends on the question of how much the purchased electricity is composed of fossil primary energy sources and how efficiently the hardware works for the respective calculation [8]. Of course, there is also the question of raw material criticality regarding the hardware primary materials. In addition to transport and extraction, the manufacturing processes themselves have

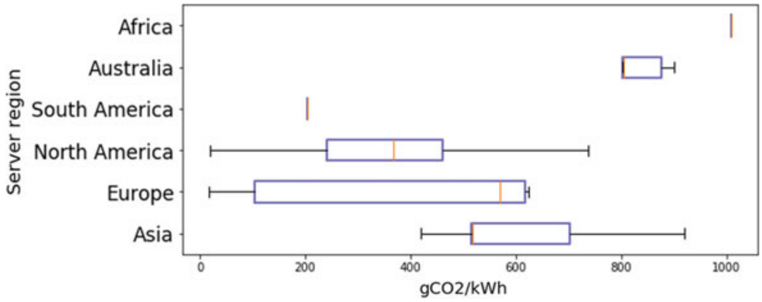


Fig. 2 Quantifying the carbon emissions of machine learning [10]

the potential to exacerbate regional conflicts or release [9] toxic chemicals into the environment.

Regarding the emission of CO_2-equivalent greenhouse gases, the sometimes regionally high differences should be mentioned, which are considered in the calculation of the Machine Learning Emissions Calculator: "Servers located in North America emit anywhere between 20 g CO_2 eq/kWh in Quebec, Canada to 736.6 g CO_2eq/kWh in Iowa, USA" [10] (FIg. 2).

4 More Differentiated Standard for Data Centers

Elementary for the energy demand of computations is the efficiency of the IT infrastructure. In the field of artificial intelligence, it is relatively common to have the calculations carried out by cloud services due to the acquisition costs. In this respect, data centers play an important role in the topic of energy requirements of artificial intelligence and the associated rebound effect. According to ISO/IEC 30,134–2:2016, the current standard for quantifying energy efficiency in data centers PUE (power usage effectiveness) is composed of the total energy of a data center in relation to the electricity used for computing and storing data [11]. A PUE of 1.0 would therefore mean that all incoming electricity can be converted into computing power. However, there is no obligation to provide information on how this data was determined by the operators of the data centers. At the European level, the "European Code of Conduct for energy efficiency in data centers" is a non-binding compendium for data center operators to make their own operations more energy efficient [12].

The Federal Environment Agency was commissioned by the Federal Ministry of Economics and Technology to research a standard that describes the energy and resource efficiency of data centers and cloud storage. The KPI4DCE (Key Performance Indicator for Data Center Efficiency), which is currently in the validation phase, includes significantly more comprehensive sustainability criteria for data centers than PUE. In addition to the energy required to manufacture the data storage, the longevity, transport route of the data to the user and the local climate conditions of the [13] data center are also taken into account. In some cases, operators also use opportunities to feed [14] server waste heat into heating networks (in accordance with the principle of the circular economy).

5 Energy Demand and Greenhouse Potential in NLP Models

Strubell et. al. from the University of Massachusetts estimated, among other things, the energy demand, the CO2 equivalent caused and the approximate computational costs for cloud providers for the training of several popular [15] natural language programming models. The values for a Transformer model were particularly remarkable. The basis of this model was the TensorFlow Library "Tensor2Tensor" [16, 17]. A model with 213 million parameters was used for the weighting. Neural architecture search (NAS) was built into the existing Transformer architecture as an additional feature. For the calculations, a TPUv2 (tensor processor) chip was used to train the AI. The TPU needed 10 h of time [18] for this (Fig. 3).

Strubell et al. calculated that, using 8 GPUs (Nvidia Tesla P100), this would correspond to a training time of just over 274 h. From this, in turn, the energy demand was anticipated, which in turn was estimated at a CO_2 equivalent of 0.954 pounds per kW/h. The result reads relatively devastating: implementing the AI in the large Transformer set including NAS is equivalent to 626,155 pounds of CO2. Which is roughly 5 times the CO2 equivalents of the product life of a mid-range car built in the mid-1990s and used for about 193,000 km [19].

In a study prepared in 2015 for the Energy Commission of New Zealand (EECA), a car with an electric drive and an estimated useful life of 210,000 km was calculated at approx. 51,940 pounds of CO2 equivalent within the scope of a life cycle assessment. In this calculation it is assumed that an electric car requires 2 batteries and ecoinvent v3.1 was used as a data basis for the material flows.

This somewhat more recent value is of course strongly dependent on the country's electricity mix. Nevertheless, it is clear that this is not a trivial environmental impact.

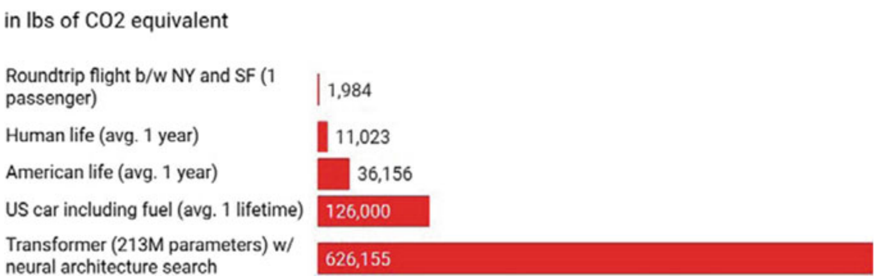


Fig. 3 CO_2 equivalents [18]

	Date of original paper	Energy consumption (kWh)	Carbon footprint (lbs of CO2e)	Cloud compute cost (USD)
Transformer (65M parameters)	Jun, 2017	27	26	$41-$140
Transformer (213M parameters)	Jun, 2017	201	192	$289-$981
ELMo	Feb, 2018	275	262	$433-$1,472
BERT (110M parameters)	Oct, 2018	1,507	1,438	$3,751-$12,571
Transformer (213M parameters) w/ neural architecture search	Jan, 2019	656,347	626,155	$942,973-$3,201,722
GPT-2	Feb, 2019	-	-	$12,902-$43,008

Note: Because of a lack of power draw data on GPT-2's training hardware, the researchers weren't able to calculate its carbon footprint.

Fig. 4 Energy and policy considerations for deep learning in NLP [18]

The described training of the AI corresponds roughly to dozens of times the CO2 equivalent of such an electric car over the complete product lifetime [20] (Fig. 4).

The NLP model GPT-2, published by OpenAI in February 2019, has 1.542 billion parameters in its largest variant. GPT-2 is supposed to be able to compose texts on its own—that is, to predict which word after a series of words is most likely to result in a text that is understandable to humans. For training, the developers of GPT-2 have specified 168 h on 32 TPUv3 chips. While TPU v2 performs 180 TeraFLOPS per chip, the computing power of TPU v3 is already at 420 TeraFLOPS [21]. In mid-2020, the successor model GPT-3 was released, which contains 175 billion parameters in its largest variant.

In the field of artificial intelligence, the computing effort of a training is also expressed in petaflop/s-day. One petaflop/s-day corresponds to a workload of [10, 15] neuronal operations per second (which would correspond to a total of 1020 computing operations for one day). The authors state the effort for pre-training as "several thousand petaflop/s- day", whereas for GPT-2 it would still be a dimensioning in the range of a few dozen petaflop/s-day seconds per day. However, the claim is also made that the computations for the actual language processing in the running operation of a trained GTP-3 would not have a significant energy requirement (approx. 0.4 kW/hour) [22] (FIg. 5).

This exponential trend in machine learning in terms of the demand for computing capacity by relatively popular sets can be observed roughly since 2012.

Before 2012, it was not common for GPU to be used for computation. While AlexNet has ensured that the approach of using the GPU has gained popularity, such approaches [24] existed in the field of image processing before that. The further back

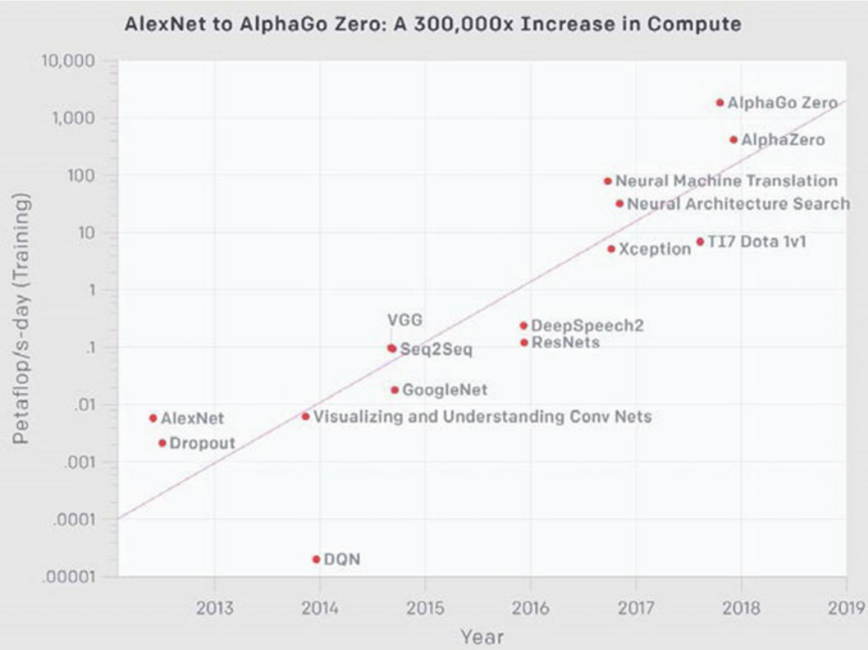


Fig. 5 Trend regarding the need for computing capacity [23]

in time one looks, the more relevant differences in the price of RAM and persistent memory are likely to have been. In addition, the development of machine learning may have been accelerated by more performant algorithms and by the popularity of economic applications (FIg. 6).

6 Energy Demand Forecasting via Frameworks and Models

AI techniques can also be used to predict the energy requirements of learning processes themselves. This is an attempt to be able to operate more energy-efficient machine learning in the actual computation process. This optimization can be done for both training and the use of the finished AI. This should increase the probability of selecting a model that makes sense in the context of the task. It can also make it easier to avoid producing unusable results if, for example, the result cannot function reliably on the target platform. It can also help to estimate the time and energy required for training processes more realistically in advance.

The “NeuralPower" framework makes it possible to predict how high the power and energy requirements of a CNN will be and how long calculation processes will take.

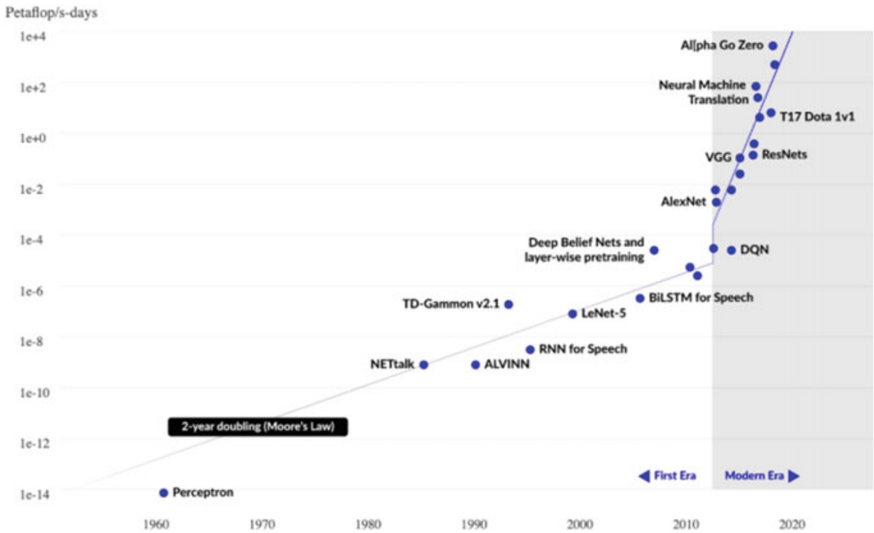


Fig. 6 Artificial neural networks with their amount of PetaFLOPS required for training [23]

Tests should already suffice for the prediction of the framework. The CNN itself does not have to be in operation, nor does it have to be implemented on the target platform. The prediction of power consumption was developed for GPUs. The software uses regression methods for this. It is possible to consider [25] the power requirements of individual layers in isolation.

The "SyNergy" framework serves to enable predictions of learning procedures on mobile hardware or architecture for embedded systems. Apart from the specialization on this form of hardware, the use cases seem to be similar to those of "NeuralPower". According to the authors, at the time SyNergy was programmed, this form of prediction by popular frameworks such as TensorFlow, Caffe or cuDNN would not have existed for this form of processor. SyNergy was also used for energy prediction of CNNs. Here, the "Streamline" performance tool developed by ARM was used to analyze the energy demand at layer level in a fine-grained manner. SyNergy also uses regression methods [26] for this purpose (Fig. 7).

Google uses its own AI "Deepmind" to optimize the cooling system of its own data centers. The company claims that this has resulted in an energy saving of 40% compared to the previous state [27]. Alphabet also boasts that it wants to [28] be CO_2 neutral in terms of energy consumption by around 2030.

The Group has thus set itself more ambitious goals. In its 2021 amendment to the Renewable Energy Sources Act, the German government set a target of 65% gross electricity consumption from renewable energies by 2030 [29].

The business model of Google's holding company Alphabet is nevertheless based on the fact that personalized advertising is massively distributed worldwide on its own platforms with the help of artificial intelligence, among other things. The purpose

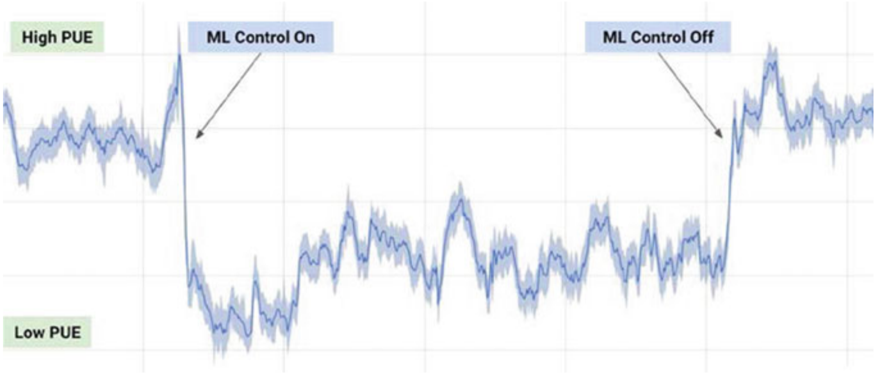


Fig. 7 DeepMind AI reduces Google cooling bill by 40% [27]

behind this is, of course, to create incentives for consumption. This, in turn, is not very helpful in advancing ecological goals in a society.

7 Green AI and Digitalization

In addition to the enormous environmental impact described above, the resource requirements of current AI models also represent another critical trend: At scales beyond the Evolved Transformer, the cost of electricity and hardware (or cloud service rental) can run into seven or eight figures. Outside the data science sector, models with a very large number of parameters (such as the aforementioned set from GTP-3) or experimental approaches with modified algorithms in this price range are hardly usable for SMEs or can represent a serious financial risk. This effect is exacerbated for public research institutions such as universities—especially in countries with lower education budgets. This can lead to a widening of financial class differences in research and education. The Green AI concept attempts to counteract these developments. The approach pursues the idea of bringing into the scientific discourse on AI not only the improvement of a performative best, but also the efficiency with which a certain result was achieved. The weighting could thus be composed of the number of examples used (E), the size of the training data set (D) and the number of AI trainings through hyperparameter experiments (H), thus scaling the result (R) or the accuracy of the AI [30].

$$Cost(R) \propto E \cdot D \cdot H$$

The resulting weighted costs place a strong focus on efficient AI, as the parameters E, D and H are directly reflected in the computational effort.

In addition, it is proposed to include the number of floating point operations (FPO) in the evaluation. FPO has the advantage that it makes no quantitative difference whether an AI has been trained with powerful hardware or not (e.g. in contrast to the time required). FPO allows relatively clear conclusions to be drawn about energy consumption, but not about the associated CO_2 emissions.

All in all, Green AI attempts to establish a more multi-layered approach. Not only the accuracy of an AI should come into focus, but also additional factors such as money, energy consumption and, derived from this, environmental impact. By reducing the amount of material used, this approach hopes to improve the reproducibility of institutions with smaller budgets. Green AI also positively highlights the trend of publishing models that have already been trained.

8 Discussion and Conclusion

The measures mentioned in the previous sections are important in their own way to contribute to energy saving. However, they remain only approaches to increasing efficiency. As necessary as this level is, pure efficiency is not sufficient as a solution with regard to rebound effects. Just as Green AI attempts to frame a more holistic picture of the AI community's objective, it is necessary from a sustainability perspective to expand efficiency to include levels such as consistency and sufficiency. In simple terms, consistency describes the implementation of a circular economy. The proportion of recyclable components should be as high as possible. Ideally, the same processed resources are recycled and reused again and again for new products [31]. A concrete example would be the "cradle to cradle" approach [32].

Since intangible goods are created in software development, this logic can only be transferred indirectly to the product through concepts such as frameworks. Above all, the same resources are not processed. Only in terms of knowledge and time savings can direct benefits be gained in this way. Instead, software without protective mechanisms is only bound to the hardware requirements in terms of its duplicability. The extent to which companies have an interest in making already trained AI available as a blueprint or simply reselling the resulting results depends on the respective business model.

Even though some efforts are being made in data centers to further utilize waste heat and more detailed KPIs are expected from the Federal Environment Agency in the near future, the circular economy in the area of IT still has a lot of room for improvement: This applies in particular to the recyclability of hardware. Even with the modularity of IT products, there are hardly any German companies that show a serious interest in the topic, apart from a few exceptions such as the company Shiftphones [33]. Probably little will change in this regard without political incentives. A high CO_2 price could help cloud service providers to be more motivated to use green electricity and efficient hardware and cooling systems. If providers consider a possible loss of image unimportant, however, this could only have the effect that such business models are outsourced to countries where the costs of CO_2-heavy business models are not internalized. In the worst case, a competitive situation could arise among countries that undercut each other with ecological guidelines—as is already the case in some EU countries in the area of corporate taxes [34, 35]. If there is no support through targeted subsidies, for example, more ecologically effective CO_2 pricing could also have a negative impact on the public education sector.

The use of artificial intelligence thus raises the question of whether techniques are used for their own sake or whether innovations are to be seriously advanced.

This aspect is also criticized by Strubell et al.: Very elaborately created AI systems with a homeopathic improvement in performance compared to the current state-of-the-art can be confronted with enormous, negative environmental influences. In the case of the Evolved Transformer, for example, an improvement in the BLEU score of only 0.1 points (compared to the state of the art at the time) was achieved for the machine translation from English into German [15, 18].

It can be assumed that donors will be more easily persuaded to make investments or funding available if a tool is proclaimed that is currently part of an industry or media trend. The sense of this remains doubtful if there is a frantic attempt to achieve tiny increases in efficiency with a method, only to require all the more energy and hardware at the same time.

Even highly resource-intensive AI approaches undoubtedly have their strengths. However, using it as a universal tool for all possible use cases can be tantamount to wastefulness, especially when technically skilled personnel make these misguided decisions with their eyes open. What is needed is a greater understanding of the conditions under which the material cost of individual methods of artificial intelligence is no longer in proportion to the yield. At the end of the day, there is conscious renunciation—sufficiency. This does not necessarily involve a ban or painful renunciations: "The basic understanding of sufficiency is that social problems cannot be solved by new technologies alone, but only in conjunction with behavioral changes"[5]. If these behavioral changes ultimately lead to people spending their time more sensibly or pleasantly, it is not necessarily a form of renunciation that must be unpleasant. In consumption, this can be unnecessary goods that restrict more than they really serve. In the case of a relatively superfluous use of AI compared to the effort involved, on the developer side this can correspond to a task that is only processed because it must be done. But not because it would be useful or interesting. If instead, from the perspective of the client and service provider, there is a more concrete idea of when an AI is so parameterized anyway that it is a waste for everyone involved, it is probably a better decision to work on something else instead. It could possibly be a sensible approach to precisely automate this assessment through AI. At least in basic terms, the AI frameworks described provide assistance for this. Both socially and as an individual, the question arises as to whether more or some form of growth is always better. In the reality of each individual's life, there is only limited space and a limited time window for consumption, just as the environment has planetary limited resources. Rebound effects can also be used as a synonym for growth ambition. This is of course also the case when economic growth is to be stimulated with the help of AI. As long as a society and its economic system is dependent on growth, there will also be rebound effects.

In combination with an efficient circular economy and conscious consumer behavior, the impact of rebound effects can possibly be reduced.

Innovative technology such as AI can certainly make useful contributions—but only if it is used in a targeted manner. More digitization does not necessarily lead to more efficiency.

References

1. Bianchini, S., Müller, M., Pelletier, P.: Deep Learning in Science. Université de Strasbourg, arXiv preprint arXiv:2009.01575,2020
2. Russell, S., Norvig, P.: Artificial Intelligence—A Modern Approach. Pearson Germany Ltd. (2012)
3. Wang, Y., Gu-Yeon, W., Brooks, D.: Benchmarking TPU, GPU, and CPU Platforms for Deep Learning. Harvard University, John A. Paulson school of engineering and applied sciences (2019)
4. Horowitz, M.: Computing's Energy Problem (and what we can do about it). In: International Solid-State Circuits Conference. San Francisco, USA (2014)
5. Santarius, T.: The Rebound Effect—On the Undesired Consequences of Desired Energy Efficiency. Wuppertal Institute for Climate, Environment and Energy GmbH (2012)
6. Sorrell, S.: The Rebound Effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency, UK Energy Research Centre (UK ERC) (2007)
7. Allan, G., et al.: UKERC Review of Evidence for the Rebound Effect: Technical Report 4: Computable general equilibrium modelling studies. UK Energy Research Centre (2007)
8. Fritsche, U.R.: Greenhouse gas emissions and abatement costs of nuclear, fossil and renewable electricity supply, Öko-Institut e.V. (2007)
9. Marschall, L., Holdinghausen, H.: Seltene Erden—Umkämpfte Rohstoffe des Hightech-Zeitalters. ekom verlag, München (2007)
10. Lacoste, A., et al.: Quantifying the Carbon Emissions of Machine Learning. Université de Montréal, arXiv:1910.09700 (2009)
11. ISO/IEC 30134: Information technology—Data centres—Key performance indicators—Part 2: Power usage effectiveness (PUE). 30134–2:2016, 2016
12. Acton, M., Bertoldi, P., Booth, J.: Best practice guidelines for the EU code of conduct on data centre energy efficiency. Joint Research Centre, European Commission, JRC **119571**, 2020 (2020)
13. Schödwell, B. Zarnekow, R.: Kennzahlen und Indikatoren für die Beurteilung der Ressourceneffizienz von Rechenzentren und Prüfung der praktischen Anwendbarkeit, Umweltbundesamt (2018)
14. Funke, T., et al.: Waste heat utilisation in the data centre. eco - Verband der Internetwirtschaft e. V. (2019)
15. Strubell, E., et al.: Energy and Policy Considerations for Deep Learning in NLP. University of Massachusetts Amherst, College of Information and Computer Sciences (2019)
16. Github repository. https://github.com/tensorflow/tensor2tensor. 2017–2021
17. Vaswani, A., et al.: Attention is all you need. Google Brain (2017)
18. So, D.R., Liang, C., Le, Q.V.: The Evolved Transformer. arXiv:1901.11117 (2019)
19. Keoleian, G.A., et al.: LCI Modeling challenges and solutions for a complex product system: a mid-sized automobile. In: The Engineering Society for Advanced Mobility Land Sea and Space International (SAE) (1998)
20. Arup Pty Ltd.: Energy Efficiency and Conservation Authority Life Cycle Assessment of Electric Vehicles - Final Report. Energy Efficiency and Conservation Authority, New Zeeland (EECA) (2015)

21. Norrie, T., et al.: Google's Training Chips Revealed: TPUv2 and TPUv3. Institute of Electrical and Electronics Engineers Hot Chips 32 Symposium (HCS) (2020)
22. Brown, T.B., et al.: Language Models are Few-Shot Learners. OpenAI (2020)
23. Amodei, D., Hernandez, D.: AI and Compute. OpenAI (2018)
24. Chellapilla, K.: High performance convolutional neural networks for document processing. Microsoft Research (2006)
25. Cai, E., et al.: NeuralPower: predict and deploy energy-efficient convolutional neural networks. In: Proceedings of Machine Learning Research (PMLR), Volume 77: Asian Conference on Machine Learning (2017)
26. Rodrigues, F.C., et al.: SyNERGY: an energy measurement and prediction framework for Convolutional Neural Networks on Jetson TX1. In: International Conference on Parallel and Distributed Processing Techniques & Applications (PDPTA) (2018)
27. Evans, R., Gao, J.: DeepMind AI reduces google data centre cooling bill by 40%. Alphabet Inc. (2016)
28. A message from our CEO—Our third decade of climate action: realizing a carbon-free future. Alphabet Inc. (2020)
29. Act on the Expansion of Renewable Energies—EEG 2021. §1, para. 2. Federal Ministry of Justice (2021)
30. Schwartz, R., et al.: Green AI. Allen Institute for AI (2019)
31. Geissdoerfer, M., et al.: The circular economy—a new sustainability paradigm?. J. Cleaner Prod. **143** (2017)
32. McDonough, W., Braungart, M.: Cradle to Cradle: Remaking the Way We Make Things. North Point Press, New York (2002)
33. https://www.shiftphones.com/pfand/. Accessed 23 April 2021
34. Gallego, C.S.: Offshore secrecy—but where is Bermuda, Luxembourg? New EU 'blacklist' omits major tax havens. In: International Consortium of Investigative Journalists (ICIJ) (2017)
35. Press Release European Commission: State aid: Ireland gave illegal tax benefits to Apple worth up to €13 billion. European Commission (2016)

Assisting PV Experts in On-Site Condition Evaluation of PV Modules Using Weather-Independent Dark IV String Curves, Artificial Intelligence and a Web-Database

Joachim Rüter, Felix Meyer, Grit Behrens, Konrad Mertens, and Matthias Diehl

Abstract Photovoltaic (PV) modules can make a huge contribution to achieve the Sustainable Development Goals of the United Nations. To be able to make that contribution, regular check-ups and evaluation of installed PV modules are necessary as they can develop faults and degenerate over time. In this project, we improve the *dark IV string curve method* used for on-site fault detection and module evaluation. We do so by training artificial intelligence (AI) models to predict the maximum power point and the bright IV curve of PV modules given the weather-independent dark IV string curve. We present some background on this topic, describe the data used for training and the developed models. The results are illustrated graphically. To make the models available for PV experts in practice and to support their decision-making process, we also developed the web-database-application *iPVModule* for storing historical PV Module data and integrated the AI-models.

Keywords Artificial intelligence · Deep learning · Photovoltaic · Dark IV string curve · IV curve · Maximum power point · Web · Database · Application · Predictive maintenance

J. Rüter (✉) · F. Meyer · G. Behrens
Solar Computing Lab, University of Applied Sciences Bielefeld, Artilleriestr. 9, 32425 Minden, Germany
e-mail: joachim.rueter@fh-bielefeld.de

K. Mertens
Testing Lab for Photovoltaics, Münster University of Applied Sciences, Electrical Engineering, Stegerwaldstraße 39, 48565 Steinfurt, Germany

M. Diehl
Photovoltaikbüro Ternius und Diehl GbR, Schünauerhofstrstraße 27, 65428 Rüsselsheim, Germany

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_6

1 Motivation

With the adoption of the "Agenda 2030 for Sustainable Development" and the 17 "Sustainable Development Goals" (SDGs), the United Nations ("UN") gave themselves goals for a sustainable development and for an increase of the prosperity of our world and its inhabitants. One of these SDGs is "affordable and clean energy" (SDG 7, [15]). But indirectly, this goal touches many of the others. As like the generation of clean water (SDG 6: "clean water and sanitation", [15]) or to cook food (SDG 2: "zero hunger", [15]), many everyday necessities and activities to improve ones life need energy. Because of this, affordable and clean energy is a central topic of the SDGs. This applies as well for developed countries that dump out too much CO_2 in their energy production as for developing countries in which many citizens do not have any access to energy.

For the generation of affordable and clean energy, many experts see the generation of energy from solar irradiation using photovoltaic (PV) systems as a promising approach. Theoretically, it is possible to produce all the energy the world currently needs with PV systems ([6], p. 61). Although in practice a combination of different technologies seems the most reasonable approach ([12], p. 159). When the current installation rate of PV systems holds, PV could produce up to 30% to 50% of the total worldwide generated energy by mid-century ([12], p. 159). This form of energy generation has benefits for developed countries as well for developing countries.

But besides its benefits, PV systems still have some open problems which have to be solved. As the production of PV modules has a high energy usage, they have to operate for a particular time to have a positive energy footprint in the big picture. However, during their lifetime, faults, damages and degeneration can occur that reduce the output of the modules. Because of that, regular check-ups are important to detect these in time or even better in a predictive manner to minimize power losses and to ensure an effective and high-yield operation.

But as current check-up methods are either expensive or very time-consuming and as the prices of PV modules have fallen over the past few years, modules are not checked but directly removed and destroyed instead of repaired as soon as power losses are detected. That reduces the environmental benefits of PV modules and its positive impact on achieving the SDGs. Because of that, PV experts need better tools and more assistance for an easier, faster and more efficient decision-making and fault-detection process.

In this work, we show that information and communication technologies (ICT) can help to make the check-up process of PV modules more efficient and effective by using artificial intelligence (AI), useful and interpretable data combined with communication technologies to make the information accessible for the PV experts on-site. We also developed a platform for them to connect and share their knowledge and findings on specific PV modules. We present some more background on fault detection of PV modules and describe the development of AI models that can help experts in their decision-making and evaluation process. To make the models and data accessible in practice we integrated them into our web-database-application *iPVModule*.

2 Background and Related Work

As described in Sect. 1, fault detection and condition prognosis of PV modules are very important aspects to ensure high-yield and long-term operation of PV plants. These consist of one or more PV strings which are formed by PV modules connected in series. A very accurate overview of the condition of PV modules can be achieved by measurements in laboratories using an artificial irradiation source. With these measurements, the *bright current-voltage curve* (*bright IV curve*) and some other characteristics of the module are measured. Examples of such curves are shown in Fig. 1. As the curve is very sensitive to its environmental conditions such as irradiation and temperature, the curves are always measured under the same conditions. The dependence on these conditions is exemplary shown in Fig. 1 (right). From these curves the *maximum-power-point* (MPP) and other characteristic properties such as the *short-circuit current* (I_{SC}) and the *open-circuit voltage* (U_{OC}) can be derived as shown in Fig. 1 (left). The MPP is the point at which the PV system generates the most power. It is the operating point at which the system should be operated for maximum yield and gives information about the condition of the system and its maximum energy conversion capacity. Deviations from the expected curve or from expected parameters can give information about faults and degeneration.

Because of a continuous reduction in the prices of PV modules over the last few years, the evaluation of PV modules in testing laboratories is becoming more and more unattractive once they are installed on-site as the process of packing, transporting, testing, transporting back and installing again is very time-consuming and expensive. For these reasons, some on-site measurement techniques have been developed to detect faults and power losses without transportation-time and -costs. Among these are mainly *peak power measurement*, *thermography* and *low-cost outdoor electroluminescence*. They offer the possibility to detect some faults and to determine the conditions and quality of a PV plant, however these methods have problems, too, like the need for good and stable weather conditions or being time consuming.

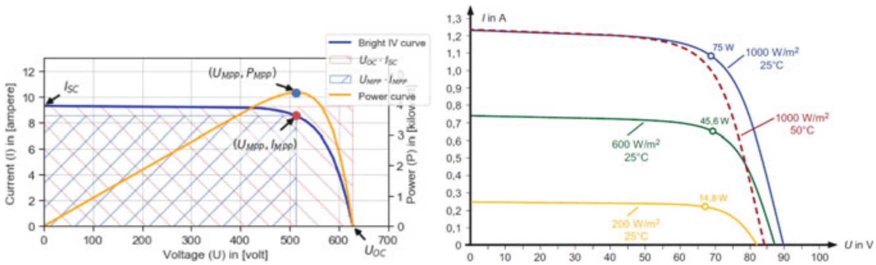


Fig. 1 Bright IV curves with annotated characteristics and dependencies. Left: Bright IV curve (blue line) with some annotated characteristics. It corresponds to the left axis. Yellow curve symbolizes the resulting power curve. It corresponds to the right axis. Adapted from: [3]. Right: Bright IV curves measured under different environmental conditions. *Source* [6], p. 175

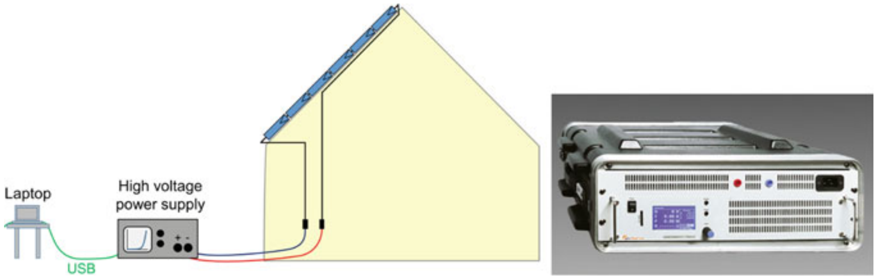


Fig. 2 Left: Set-up for dark IV string curve measurements. Adapted from: [7]. Right: A commonly used high voltage power supply for dark IV string curve measurements. *Source* [11]

A commonly used on-site evaluation method is peak power measurement. The goal of these measurements is to trace the above described bright IV curve of a PV string on-site using the sun as irradiation source. Passing clouds can make this procedure very time-consuming and varying module temperatures in a PV string can make the interpretation and comparison of curves difficult as described above and shown in Fig. 1 (right) and the reproducibility of the measurements can not be ensured. Additionally, the measurement devices need a minimal irradiation which is not always given (see [6], p. 279).

To solve these problems of weather-dependence and of time-consumption, [7] respectively [8] introduced the *dark IV string curve* measurement technique. In these technique, the *dark IV curve* of a whole string is measured at night without irradiation. To do so, a programmable high voltage power supply (0 – 1000 V, 0 – 5 A) is connected to the string which applies a rising positive voltage and measures the resulting current at the same time. The whole process takes less than a second (see [7]) and, as all modules have the same irradiation (0 W/m^2) and the same temperature, the methods allows a fast, weather-independent and reproducible analysis of PV modules. The measurement setup and a commonly used special high voltage power supply which has also been used in our research are shown in Fig. 2. An exemplary measurement of dark IV string curves of different strings of a plant is shown in Fig. 3.

This measurement technique also shows promising results for fault detection ([7], [8]). To do so, the dark IV string curves of all strings of a plant are traced and compared among each other. Deviations from the normal dark IV curve can then indicate faults or give information about the necessity of deeper analysis. Such a deviation is also shown in Fig. 3.

Besides the detection of some faults, it would be helpful when the dark IV string curves could also give information about the bright IV curve and therefore about the power conditions of the string. This could allow—possibly in combination with other data of the PV modules—to find even more faults and to get a better overview of the condition of the PV modules.

This prediction of the MPP and of the bright IV curve of a PV string is the main topic of this project. Currently, the main approach to solve this problem is to use the *one-* or *two-diode model* of a PV module to extract parameters of the module and

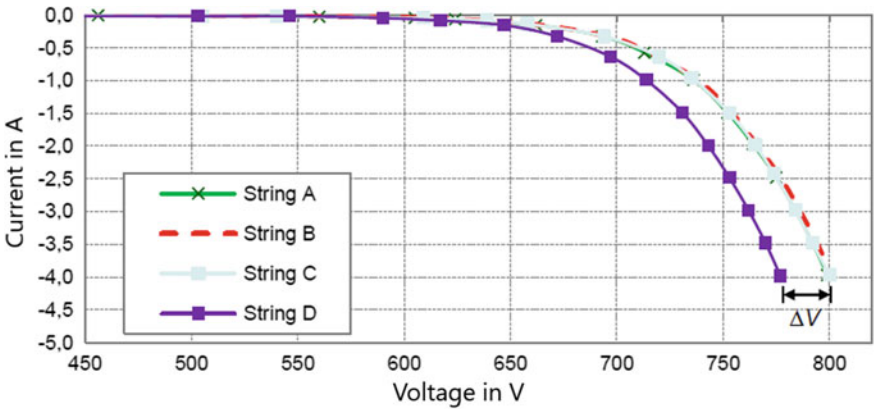


Fig. 3 Dark IV string curves of different strings. String D shows a deviation from the others strings. That could indicate a fault. Adapted from: [8]

to use them to simulate the bright IV curve. Ref. [4] examines the dark IV curve of only one module and lab-data and uses it in combination with the two-diode model, knowledge-based assumptions and hand-tuning to predict the bright IV curve. For strings, [9] relies on the one-diode model and uses lab data to simulate the bright IV curve with some deviation. As the parameters of these models derived from the dark IV curve can differ from those derived from the bright IV curve (see [1, 4, 9]), this approach is not ideal and other approaches should be researched.

In this project, we evaluate the capabilities of artificial intelligence to solve this problem of predicting the MPP and the bright IV curve using data of the dark IV string curve. Some work in progress from us on this has already been published. Firstly, we developed a User-Interface that allows PV experts to upload IV curve data [14]. It also shows first visualizations of the data but is not designed for use in practice. This step was important to get a sufficiently large data-set to train AI-models. Ref. [2] shows some work in progress with a focus on the preprocessing and cleaning of the data which has been improved over some iterations since then. A comparison of different methods to extract features from the dark IV curves to be used as input for neural networks is published in [3]. It shows that using equidistant points from the dark IV curve leads to the best prediction results of the power output conditions of a string. In [13], we present our state-of-the-art for the prediction of the MPP of a string using the dark IV string curve as input. To make our models accessible for PV experts in practice, we also developed the web-database-application *iPVModule* which stores historical data from PV modules and curve data to give the PV experts usable information and more data for better fault detection.

In the following section, we briefly describe our data-set and the preprocessing pipeline used to clean and transform the data to use it for training neural networks. We then shortly give details about our best model to predict the MPP of a module which was published in [13] and then present our new best model to predict the whole bright IV curve, evaluate its performance and illustrate it graphically. All in all, we give a compact overview about the whole project and our results by explaining and presenting our data, the preprocessing and the best models to predict the MPP and the bright IV curve given the dark IV curve. We also give a short overview about the deployment of the models for practice.

3 Data and Preprocessing

To train the machine learning models, we build on a large data-set of bright IV curves measured on-site with the peak power measurement technique and dark IV string curves. Overall we can build on 3356 bright IV string curves and 1588 dark IV string curves measured on 9 different PV plants with a total of 63 strings. The plants have modules with different module- and cell-types so that the data-set covers a wide variety of possible modules. The bright IV curves were measured during daytime in a given interval and the dark IV string curves during the night. Because of the measurements at adjacent days and nights the bright and dark IV string curves of a given string can be combined to pairs in the final data-set.

Our raw feature frame has the following features: Name and ID of the plant and the string, the number of modules connected in series for each string, the module type and the manufacturer. For the bright IV curve, we have up to 250 measuring points consisting of a current- and a voltage- component together with the module temperature and the irradiation present when taking the curve. For the dark IV curve, we have the same components of the measuring points. The data is also enriched with the *short-circuit current* (I_{SC}) and the *open-circuit voltage* (U_{OC}) data under *standard test conditions (STC)* from the data sheet.

To be able to get good results from the machine learning algorithms and to not let it learn (measurement) errors, we first have to clean and transform our data. A first version of this process has also partly been described in [2] and afterwards been iteratively improved further when new information occurred. The final preprocessing-pipeline for the bright as well as for the dark IV string curves is visually described in Fig. 4.

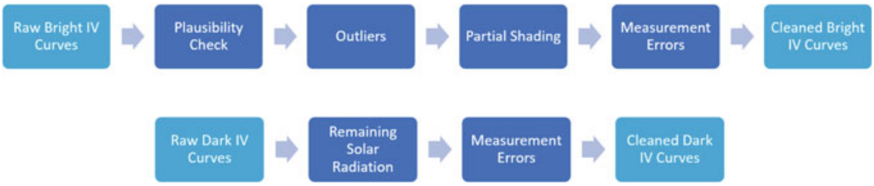


Fig. 4 Steps in the Cleaning-Pipeline for the raw bright (top) and dark (bottom) IV string curve data

The bright IV curves are first checked for basic plausibility like not having a negative or too high I_{SC} and U_{OC}. Afterwards, it is checked whether the I_{SC} of a curve deviates from the linear dependency to the irradiation by being much too high. This can only occur when the irradiation measured by the irradiation sensor is lower than the irradiation the modules actually receive. This is checked using a regression analysis. Subsequently, it is checked whether (partial) shading occurred on the PV modules while the bright IV curve was measured. These curves have to be filtered, as shading cannot be detected and therefore reproduced from the dark IV curves which are measured without irradiation at all. The pipeline finishes with some checks for measurement errors that can occur. Additionally, we drop all bright IV curves which were measured at a solar irradiation of less than 500 W/m^2 as first experiments showed that the trained models perform worst when the bright IV curves were measured under low irradiation. This is probably at least partly because the peak power measurement devices are designed for an irradiation of at least 500 W/m^2 or more. It is important to note here that we do not reduce the usefulness of our models for practical application by doing so. This specialization only means that the models are trained to predict the power capacities of a PV string for an irradiation of 500 W/m^2 or higher. This improves the performance of the models significantly but has no significant negative effect on the benefits of the results.

The dark IV curves have fewer steps in their cleaning-pipeline. It is first checked whether there was remaining sunlight when the curve was measured. In these cases, the dark IV curve loses its typical form and shifts partly into the first quadrant of the coordinate system. Therefore, these are not suitable to be used in training. Additionally, the curves are checked for some more minor (measurement) errors that can occur sometimes.

To get good results from machine learning models, it is useful to reduce the dimensionality of the input and output features and to harmonize the scales of the different attributes. While the latter is done by regular scaling techniques, we additionally divide the curves by their short-circuit current and respectively their open-circuit voltage and the number of modules which are connected in series. By doing that, we reduce the dimensionality of the input features as the string configuration and some module characteristics are now given implicitly and do not have to be passed in explicitly. The results of the model can then be easily scaled back to provide the results on the wanted scale.

As the number of measured points of the dark IV curve is not fixed, we have to extract a predefined number of features as general artificial neural networks are not capable of working with a variable number of input features. Different extraction methods have been tested and the best results come from extracting 20 points from the interpolated dark IV curves as shown in Fig. 5 (for better visualization exemplary for a bright IV curve, [3]). This interpolation also has benefits for the dimensionality of the input features. As the points are equidistant on the voltage-axis, we can remove the voltage-component from the input and just insert the current-component, as the voltage is implicitly given through the position in the input vector.

To get our final data-sets for training and testing, we split the data-set into a training-, test-, and validation-set and combine the corresponding bright and dark

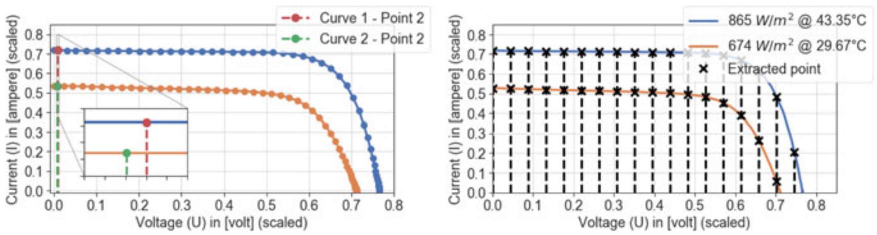


Fig. 5 Interpolation and extraction of 20 equidistant point on the voltage-axis shown for two bright IV string curves. Left: Two scaled bright IV curves. The measuring points may deviate from curve to curve. Right: 20 extracted points. Adapted from: [2]

IV string curves. We have two possibilities to do that. We can either first split the data-set and then combine the curves or vice-versa. When combining the data first, the resulting training-, test- and validation-sets would contain more data which could have a positive impact on the results of the trained models. But it is probable that dark or bright IV curves used for validation and testing have been seen by the model in training even though not in this specific combination. Contrary to the approach in [2], we split the dark and bright IV curves first and then combine the corresponding curves in the different sets. That reduces the number of usable data-pairs but we can be sure that data used for validating and testing the models were never seen in training. That can result in less overfitting and better generalization capabilities of the models. We split the data and take 25% of the data for testing, and 75% for training. From these 75% we take 20% to build a validation-set. After merging, we have 6942 dark-bright IV curve pairs for training, 427 for validation and 1205 for testing. Because of the merging, the split-percentages are shifted slightly.

4 Prediction of MPPs

Our first final model focuses on the prediction of the MPP only. Through this specialization, we had the possibility to evaluate whether the dark IV curves contain enough information to predict the bright IV curves and learn about the data.

We experimented with different architectures, activation- and optimization-functions and evaluated their influence on the results. Our best model developed for this task has been described in [13] and is visualized in Fig. 6. It has 22 input features. From these, 20 are for the current-component of the points from the interpolated dark IV curve as described in Sect. 3 and the remaining two for the module temperature and the irradiation for which the MPP should be predicted. The output layer consists of two neurons which represent the current-component (I_{MPP}) and the voltage-component (U_{MPP}) of the MPP respectively. The model has 5 fully-connected hidden layers with 6 neurons each which use the *exponential linear unit (elu)* activation function. It was trained using *ADAM*-optimizer (adaptive moment estimation, [5]) and *early stopping* to avoid overfitting.

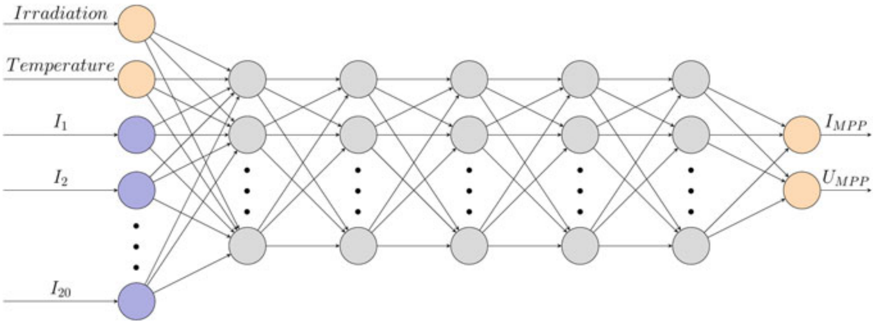


Fig. 6 Artificial neural network architecture to predict the MPP given a dark IV string curve

Table 1 Model evaluation on the test-set: Mean and standard deviation, percentiles and min. and max. deviation of Δ_{MPP} in percent [%]. Source: [13]

	mean	std. dev.	min	25%	50%	75%	max
Δ_{MPP}	2.39	1.95	0.01	0.8	1.95	3.37	8.77

To evaluate the performance of the model, we calculate the relative difference of the predicted MPP ($MPP_{prognosis,i}$) and the measured MPP ($MPP_{measured,i}$) of the i-th input vector ($i \in \mathbf{N}$) in percent [%] as

$$\Delta_{MPP,i} := \frac{|MPP_{measured,i} - MPP_{prognosis,i}|}{MPP_{measured,i}} \cdot 100 \tag{1}$$

The achieved prediction accuracy is shown in Table 1. With our model, we are able to predict the MPP of a string at a desired irradiation and module temperature given the on-site measured dark IV string curve with a mean deviation of only 2.39% and a standard deviation of 1.95% on the test-set. To get a feel for the accuracy some predictions are visualized in Fig. 7.

The achieved results are very good and show that using a suitable neural network architecture the dark IV string curves can be used to get information about the bright IV curve. This network can already help evaluating the condition of an installed PV module. Next, we want to generalize these results to predict the whole bright IV curve and compare the results.

5 Prediction of Bright IV Curves

Besides from the prediction of the MPP, PV experts will benefit even more if a machine learning model is also capable of predicting the whole bright IV curve from the dark IV curve.

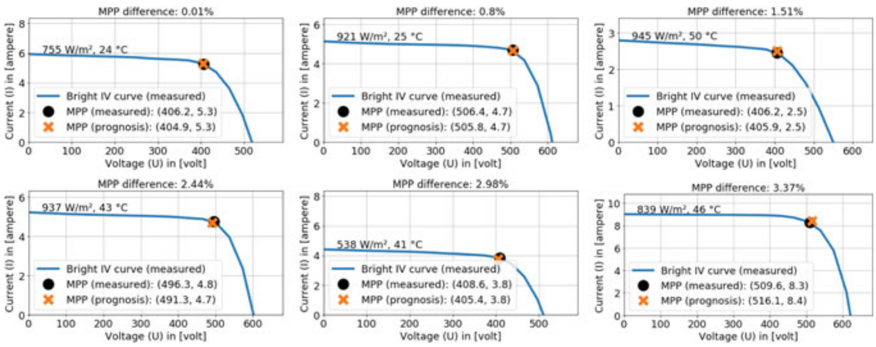


Fig. 7 Examples for the prediction accuracy of the developed model for MPP prediction: On-site measured bright IV string curves given the annotated conditions in combination with the measured and the predicted MPP. The relative difference ($\Delta_{MPP,i}$) between the MPPs is stated above each figure. *Source* [13]

With an analogous argumentation as in Sect. 3, we interpolate the bright IV curves and take 20 equidistant points from each curve. These should be predicted by the model. As the bright IV curves have a steep slope near the voltage-axis (x-axis, see for example Fig. 1), the interpolated curves also run in the fourth quadrant of the coordinate system. Because of that, we also extract points with a current-component less than zero. That means that when we plot and connect the 20 extracted points, the extracted curve fits the measured one for positive current-values but also continues for negative ones whereas the measured curve stops at zero. This is also shown in Fig. 8 (left). In short, artifacts arise with which we have to deal. This phenomenon actually occurs for the features extracted from the dark IV curves too, but as these are just used as input for the machine learning algorithms and are not seen and not interpreted by the user, we do not have to deal with them.

One approach to deal with these artifacts of the bright IV curve is to redefine all points with a current-component smaller than zero by setting them to zero. The resulting models give good results for the MPP prediction but the predicted curves sometimes get sharp edges near the x-axis as shown in Fig. 8 (right). These edges do not look good as the curve is not sufficiently fitted near the x-axis and may lead to the PV experts not accepting the results in practice. These edges are the result of redefining the negative current-values. The extraction of more points could smooth them out but would also make the model more complex.

Because of that, we decide to leave these artifacts untouched as they seem to contain useful information about the slope of the curve near the x-axis and instead cut the current-values smaller than -1. By doing so, the edges disappear and we can use the *tangens hyperbolicus (tanh)*-function as the output function of the neural network. To not distract the users in practice, we just show the current-values greater than zero of the curves.

We experimented with different neural network architectures, activation-, optimization-, and loss-functions as well as initialization strategies and evaluated

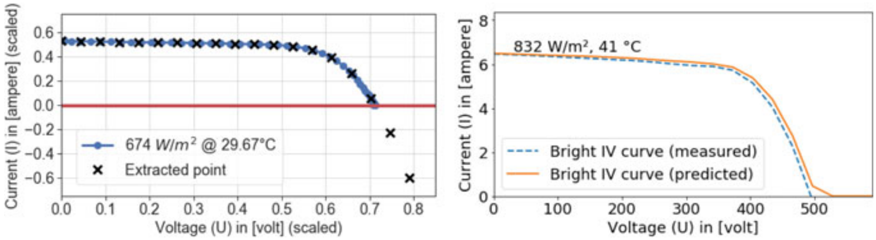


Fig. 8 Left: Visualization of the artifacts created by the interpolation of the measured curve and the extraction of 20 equidistant points. The measured curve (blue) stops at $I = 0$ (red line) whereas the extracted points (black cross) also run in the fourth quadrant. Right: Exemplary illustration of the appearing sharp edges near the x-axis when setting the negative artifacts to zero

their effect on the results. The *mean-absolute-percentage-error*-function (MAPE-function) looks reasonable on first sight but is not suitable for the evaluation and the training of the model because some points of the bright IV-curve are very close to zero and the MAPE is not defined at zero. Because of that, the models using this loss-function generate edges and the results of this function for evaluation do not reflect what the human eye considers good. Therefore, we introduce another function to evaluate the model besides the MPP difference Δ_{MPP} presented in Eq. 1. This function should be robust against relatively large but absolutely small deviations of bright IV curve current-values near zero and comprise all predicted points at once. We do that by using the following function for the i-th input vector ($i \in \mathbf{N}$)

$$\Delta_{A,i} := \frac{|\sum_{j=1}^{20} m_{i,j} - \sum_{j=1}^{20} p_{i,j}|}{\sum_{j=1}^{20} m_{i,j}} \cdot 100 \tag{2}$$

where $m_{i,j}$ and $p_{i,j}$ are the j-th current-component of the measured respectively predicted bright IV curve ($1 \leq j \leq 20$). As negative current-values are not relevant for the evaluated accuracy (see argumentation about artifacts above), we just consider points with a current-value greater than zero. This function can also be interpreted as the relative difference of the discretely approximated integrals of the measured and the predicted bright IV curves in percent [%]. This metric is not optimal but gives interpretable results. For the final evaluation and selection of the model, we use the *root-mean-square-error*-function, Δ_{MPP} (Eq. 1) and Δ_A (Eq. 2).

The best results are achieved by the model graphically represented in Fig. 9. It has 4 hidden layers with 26 neurons each. The neurons use the *rectified linear unit (relu)*-activation-function. The model was trained using the *mean-square-error*-function, *ADAM* (adaptive moment estimation, [5]) as the optimizer and *early stopping* to avoid overfitting.

The results of the model on the test-set are shown in Table 2. The model is capable of predicting the bright IV curve given the dark IV curve with a deviation of the MPP of only 2.26% on average and a standard deviation of 1.8%. This is even

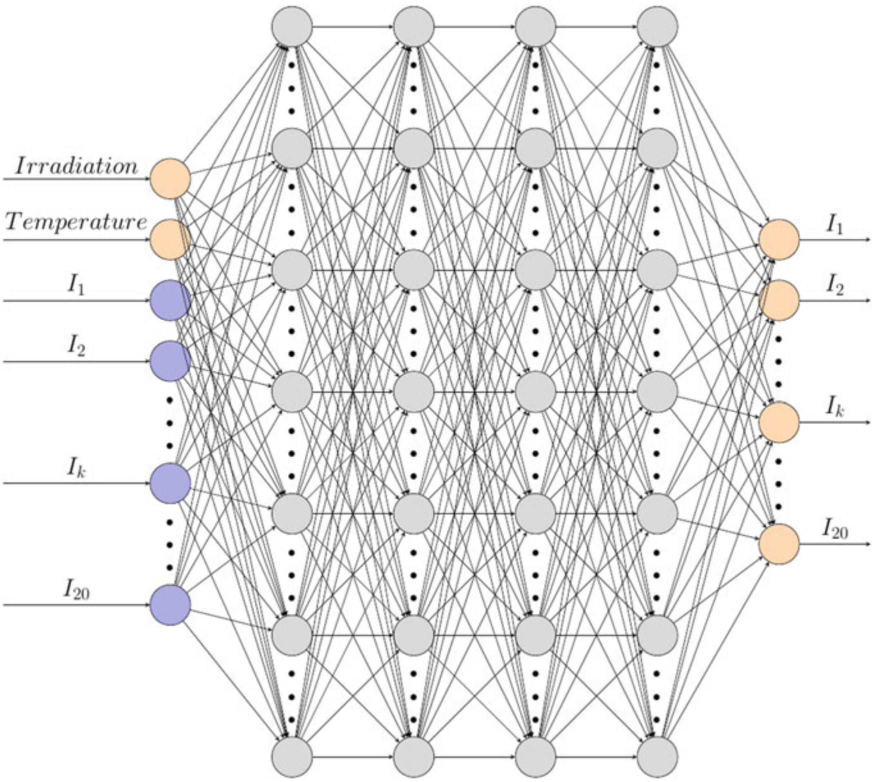


Fig. 9 Artificial neural network architecture to predict the bright IV curve given a dark IV string curve

Table 2 Evaluation of the artificial neural network to predict the bright IV curve on the test-set. Shown are mean and standard deviation, percentiles and min. and max. deviation of Δ_{MPP} and Δ_A in percent [%] as well as the *root-mean-square-error* (RMSE)

	mean	std. dev.	min	25%	50%	75%	95%	max
Δ_{MPP}	2.26	1.8	0.0	0.94	1.86	2.99	5.99	10.34
RMSE	0.04	0.02	0.0	0.02	0.03	0.05	0.08	0.14
Δ_A	2.07	1.72	0.0	0.81	1.59	2.90	5.45	10.28

slightly better than the average of the model specialized on just the MPP described in Sect. 4 though the maximum deviation is a little higher (10.34%) but still good. The results of the *root-mean-square-error- (RMSE-)* function and Δ_A are also very good. Figure 10 shows some example plots of predicted and measured bright IV curves in combination with their MPPs.

The results look very good and can give PV experts on-site useful information about the condition of the PV modules and can help in the decision-making pro-

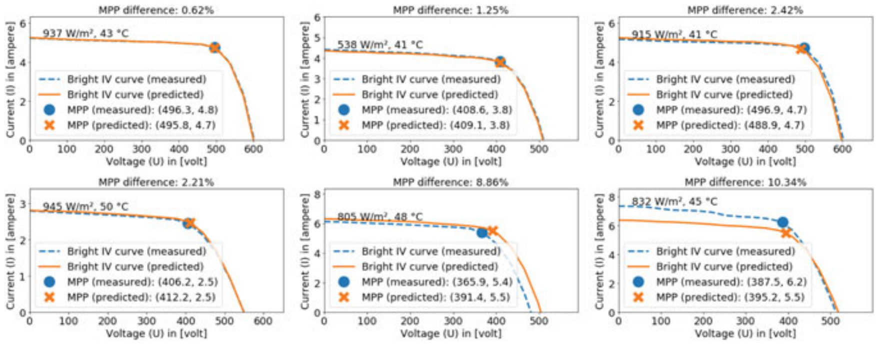


Fig. 10 Graphical illustration of the prediction accuracy of the developed model: Examples of on-site measured bright IV string curves given the annotated conditions in combination with the predicted curves. The calculated MPPs are marked and the relative difference ($\Delta_{MPP,i}$) between the MPPs is stated above each figure. On the bottom right: the largest MPP deviation. Bottom middle: The difference in the curves could indicate some form of degeneration in the string

cess. Although the model outputs more information, the MPP prediction accuracy is comparable to the model developed in Sect. 4. For an even more reliable basis for decision-making, the results of both models could be combined and compared.

To use the two models described in Sects. 4 and 5 for fault detection, the PV expert could measure the dark IV string curve regularly and use the models to predict the bright IV curve and its MPP. The curves and the MPPs can then be compared to historical data and deviations can indicate faults or be used as a starting-point for further investigations. If bright IV curves are measured too, these can also be compared to the predicted ones. In combination with monitoring data, the models can be used to check whether the predicted MPP is actually reached or if the PV string has some form of power loss.

6 Integration in the Web Database Application iPVModule

To make our models accessible and usable for PV experts, we integrated the models in our web database application *iPVModule* [10]. iPVModule is an application we developed to store historical bright and dark IV curve data of different PV modules to simplify the decision-making process of PV experts on-site, to bring PV experts and scientists together and to make scientific findings available to be used in practice. The historical data is useful because for using dark IV string curves for fault detection a comparison to other curves is needed as described in Sect. 2. On large PV plants, curves of other strings can be used but on small plants these are not available. Therefore, historical data is a necessity. It also helps on big plants, as more curves for comparison are available. The comparison of historical data is shown in Fig. 11. We also integrated a forum to let the experts share their knowledge and findings with

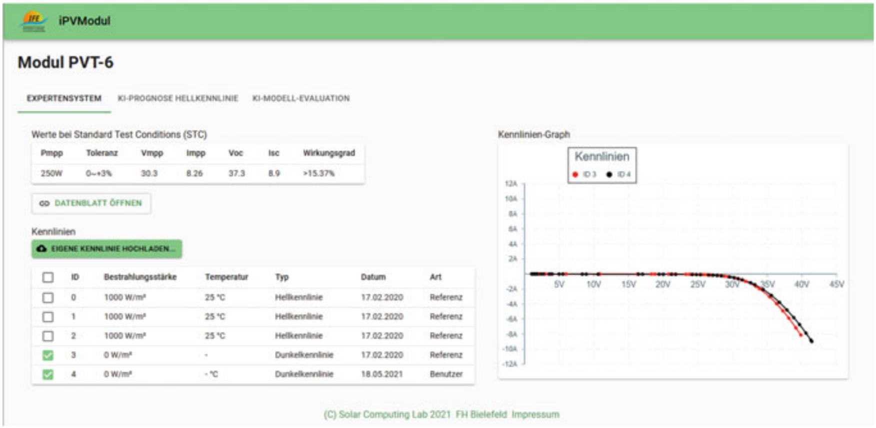


Fig. 11 Comparison of historical dark IV curve data in the web-database-application *iPVModule*

others. That helps all experts to find faults and make better and faster decisions. The applied research also benefits as it can get new labeled data from field experts which is crucial to train supervised machine learning models.

Having the goals of the application in mind, we also integrated the machine learning models into the application. That allows the experts to upload a measured dark IV curve, which can be measured fast and independent from the weather, and predict how the model should perform under irradiation. The curve can also be compared to the historical data stored in the application. This combination allows for an even better evaluation of the modules.

To slowly introduce the PV experts to the AI and to let them build up trust, we also implemented a functionality to let the experts evaluate the predictions. To do so, they can upload a measured dark IV curve in combination with a corresponding bright IV curve. The AI-model predicts the bright IV curve from the dark IV curve and compares it to the measured one. The results are graphically represented and some statistics about the deviations are shown to the user. This functionality is shown in Fig. 12. This is a very important aspect to integrate AI in the work of the experts and to make it welcomed instead of being feared and rejected.

7 Conclusion and Short Summary

In this project, we improved the dark IV string curve method used for on-site fault detection and module evaluation by training AI models to predict the MPP and the bright IV curve given the weather-independent dark IV string curve. We described the background of the work and showed how we preprocessed the data to use it to train the artificial neural networks. On the given data-set, the models reach very good

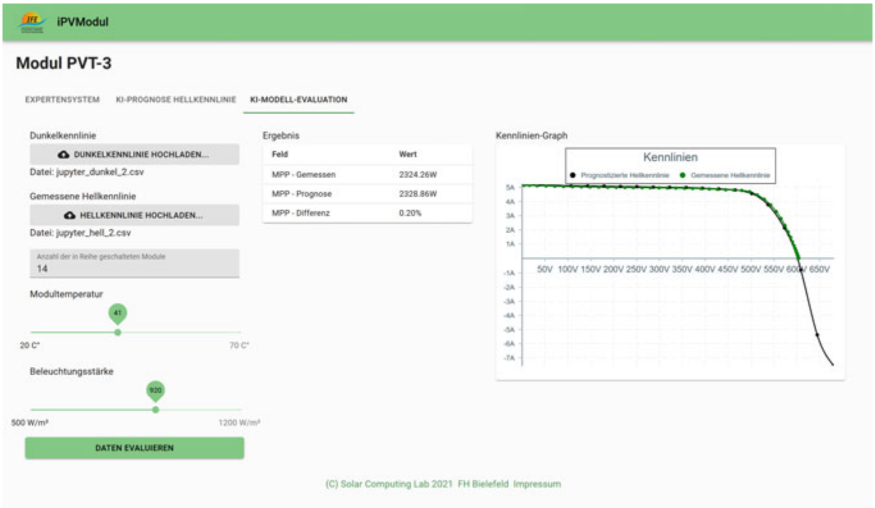


Fig. 12 Comparison of a measured and a predicted bright IV curve in *iPVModule* to show users the accuracy of the AI-models and let them build trust

results. Because of that, we also integrated them into the web-database-application *iPVModule* which stores historical IV curve data, connects experts and makes new scientific results accessible. With the developed models and iPVModule, PV experts get additional useful and interpretable information that they can use to evaluate the conditions of PV modules on-site. This makes the process faster and less error-prone and can help to make the needed check-ups of PV modules on-site faster and less expensive leading to a more efficient and high-yield usage for a longer time. Overall, the results of this project can help PV to make an even bigger contribution to accomplishing the Sustainable Development Goals of the United Nations and by doing so, to make our world a sustainable and prosperous place to live on - now and in the future.

References

1. Aberle, A., Wenham, S., Green, M.: A new method for accurate measurements of the lumped series resistance of solar cells, pp. 133–139 (1993). https://doi.org/10.1109/PVSC.1993.347065
2. Basoglu, C., Behrens, G., Mertens, K., Diehl, M.: Anlagenevaluation mittels Hell- und Dunkelkennlinien und Methoden des Maschinellen Lernens. In: 35. PV-Symposium, Bad Staffelstein (2020)
3. Basoglu, C., Behrens, G., Mertens, K., Diehl, M.: Comparison of knowledge-based feature vector extraction and geometrical parameters of photovoltaic I-V curves. In: Proceedings of the Conference "Lernen, Wissen, Daten, Analysen", CEUR Workshop Proceedings, vol. 2738, pp. 70–75. CEUR-WS.org (2020)

4. King, D., Hansen, B., Kratochvil, J., Quintana, M.: Dark current-voltage measurements on photovoltaic modules as a diagnostic or manufacturing tool, pp. 1125–1128 (1997). https://doi.org/10.1109/PVSC.1997.654286
5. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization (2017)
6. Mertens, K.: Photovoltaik—Lehrbuch zu Grundlagen, Technologie und Praxis, 5 edn. Carl Hanser Verlag München (2020)
7. Mertens, K., Arnds, A., Diehl, M.: Quick and effective plant evaluation using dark-iv string curves. In: Proceedings of 33st European Photovoltaic Solar Energy Conference, Amsterdam, pp. 2346–2348 (2017). https://doi.org/10.4229/EUPVSEC201
8. Mertens, K., Arnds, A., Diehl, M.: String-Dunkelkennlinien: Eine neue effiziente Methode zur Anlagenevaluation. In: 33. Symposium Photovoltaische Solarenergie, Staffelstein (2018)
9. Mertens, K., Gercken, L., Diehl, M., Behrens, G.: Fortschritte bei der Anlagenevaluation mittels String-Dunkelkennlinientechnik. In: 34. PV-Symposium Bad Staffelstein (2019)
10. Meyer, F., Rüter, J., Behrens, G., Diehl, M., (in press): Intelligente PV-Modul Datenbank-Applikation für Predictive Maintainance von PV-Anlagen. In: 9. Workshop Umweltinformatik zwischen Nachhaltigkeit und Wandel (UINW 2021) (2021)
11. Photovoltaikbüro Ternus und Diehl GbR: pvserve. https://photovoltaikbuero.de/en/product/inspecting-photovoltaic-systems-with-pvserve/. Accessed 18 May 2021
12. Quaschning, V.: Erneuerbare Energien und Klimaschutz, 3 edn. Carl Hanser Verlag München (2013)
13. Rüter, J., Meyer, F., Behrens, G.: Weather-independent MPP prediction of PV plants using on-site measured dark IV string curves and deep learning. In: 2021 48th IEEE Photovoltaic Specialists Conference (PVSC) (2021)
14. Teuber, M., Schlender, K., Fehring, F., Behrens, G., Mertens, K., Diehl, M.: Anlagenevaluation mittels Data-Mining für Hell- und Dunkelkennlinien. In: 34. PV-Symposium, Bad Staffelstein (2019)
15. United Nations: The 17 goals—sustainable development. https://sdgs.un.org/goals. Accessed 19 May 2021

Modelling and Simulation of Systems

Resource and Energy Efficiency Analysis in Bottle-To-Bottle Recycling Plant—Case Study

Amna Ramzy, Ahmed Elfeky, Hazem Aboulseoud, Lamia Shihata, and Volker Wohlgemuth

Abstract This research presents a practical approach towards resources and energy efficiency in production systems in small and medium enterprises (SMEs). The approach is based on mapping the production line using the modelling tool UMBERTO® Efficiency + for the representation of the material and energy flows. A case study from a bottle-to-bottle recycling plant in Egypt is selected. Material flow bottle-necks are spotted and energy hotspots are identified. Optimization is implemented in the model and the new values for savings in material, energy and carbon emissions are calculated. The optimized model estimates yearly savings by around 8% in materials and 10% in electricity consumption and consequently in costs and CO_2 emissions. The modelling tool UMBERTO Efficiency + facilitated the simple and time efficient assessment of the production sustainability as well as the prediction of the expected savings upon implementation of the modifications which is a practical and convenient approach for the industrial field.

Keywords Energy efficiency · Resource efficiency · Material flow analysis · Bottle neck analysis · Sustainability

1 Introduction

Expanded economic development and population growth created a culture of manufacturing, consumption and disposal. During the period of 2005 and 2015 global economic activity in manufacturing increased by 50%, also global gross domestic products increased by 25% [1]. Production plays a vital role in today's industrial society due to its enormous contributions in various aspects such as economics, employment, innovation and investment. The manufacturing sector is at top priority when it comes to the Egyptian government's plan for economic growth.

A. Ramzy (✉) · A. Elfeky · H. Aboulseoud · L. Shihata
Faculty of Engineering and Materials Science—German University in Cairo, Cairo, Egypt
e-mail: amna.ramzy@guc.edu.eg

V. Wohlgemuth
Hochschule Für Technik Und Wirtschaft Berlin, Berlin, Germany

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_7

Egypt participated in the Global Manufacturing and Industrial Summit in UAE, where the following initiatives were made guided by the Sustainable Development Strategy (SDS) [2]:

- Egypt vision 2030 launched in 2015.
- The industry and trade development strategy 2016–2020 launched in 2016.

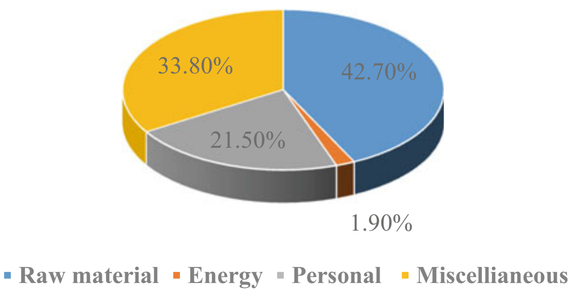
Both strategies aim for high goals when it comes to the manufacturing department including increase in the manufacturing value to the Gross Domestic Product (GDP) to 18%, and with an increase to the manufacturing growth rate with 10% [3].

This sets manufacturing as a main goal when it comes to Egypt's economy and this growth suggests initiating investment in the field of sustainability to become competitive and maximize profit while consuming less resources and energy. With the notable increase in manufacturing growth, the aim was shifted towards sustainable approach depending on resources efficiency. Resource efficiency is dispensable for sustainability, reducing environmental impact and reducing cost, as it aims to optimize the use of resources, which are material, energy, and labours, simultaneously with reducing the environmental impact and the total cost. According to the German federal statistical office, the cost of manufacturing industries is divided into four categories where only Raw material and energy are responsible of nearly 45% of the total cost [4, 5] as shown in Fig. 1.

On the other hand, small and medium-sized enterprises (SMEs) are responsible for approximately 64% of the industrial pollution in Europe [6] which makes it urgent to support this sector by energy and resources efficiency methods and tools capable of estimating their environmental impact in terms of carbon foot print [7, 8]. Generally, when SMEs conduct a resources efficiency study, they focus on the 20% contracted personal cost and explore methods to cut it down which leads to minor improvements in the overall cost savings and environmental impact and sustainability of the entity [9].

As modern manufacturing systems are becoming more complex, simulation and modelling has become essential tools required to gain insight about the behaviour of the main parameters of these complex systems. Simulation models allow data gathering, information execution and systems' testing for new operating or resource policies and new concepts or systems before their implementation and without ceasing or disturbing the actual system.

Fig. 1 Cost categories of German manufacturing industries



However, SMEs are usually working with high discretion which hinders the acquisition of essential data required to perform the simulations and obtain reliable results [10, 11].

2 Methods and Tools

2.1 Material and Energy Flow Analysis [MEFA]

Material flow and energy analysis measures materials mass in synchronization to production and emissions produced during manufacturing [12]. This is performed by considering the source, path, intermediate processes and final process of the material flow and controls the material balance of each operation in the flow process according to the law of conservation of material or energy. MEFA balances the flow of input material and that of the output material in the same process [13].

2.2 Material Flow Cost Accounting [MFCA]

MFCA focuses on the cost because of waste and material losses rather than conventional accounting, which relies on the cost of the product only, but MFCA relies on calculating the waste cost in material, energy, system, and miscellaneous. MFCA can identify problems and recognize points for improvement by studying the hotspots. The main aim of Material Flow Cost Accounting is to motivate and support the efforts of organizations to upgrade both financial and ecological performance. The optimum goal of MFCA is zero loss cost, and that can be achieved through optimizing material and energy use [14–16].

2.3 Modelling Tool

The modelling tool used in this research is UMBERTO® Efficiency + (version 10.0) software is a commercial software developed by Fraunhofer Insititute Ifu Hamburg, and commercialized by iPoint-systems GmbH in Germany. This tool allows the optimization and visualization of the production line by process mapping and the analysis using Sankey diagram representation to illustrate values of the material and energy flows, carbon footprint and cost analysis.

Accurate modelling is achieved by tailoring the equations that guide the relations between inputs and outputs at each process as shown in Fig. 2.

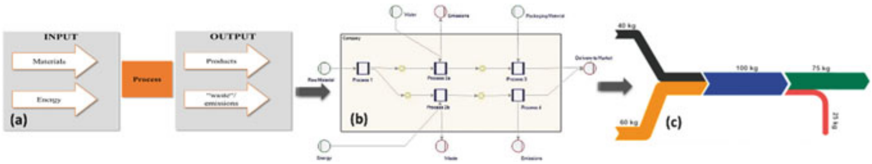


Fig. 2 Process mapping [14]; **a** Concept, **b** Process Network, **c** Sankey Diagram

2.4 *Mapping and Simulation*

The line under study recycles Polyethylene Terephthalate (PET) bottles to produce recycled PET pellets (rPET). This is the first bottle to bottle recycler in Egypt and MENA region reprocessing up to 25,000 tons which equals 1.6 billion bottles up to now. The produced food-grade quality pellets hereafter named ***QPET*** are supplied for bottle-to-bottle applications. Lower grade rPET are produced named ***APET*** and supplied to the fibres manufacturing sectors.

2.4.1 Production Line Description

- The line starts with a Mixing silo to ensure homogeneity of the flakes, followed by a metal separator to separate the metal from the PET. The flakes then go through a Hot Air dryer (HAD) where flakes are dried using hot air at the Pre-Drying Unit (PDU).
- The dried flakes enter to the processing stage in the extruder were the PET flakes are extruded into thin string shape after adding the Anti-yellow additives for colour attractiveness. Strings are cut by means of a rotating blade in the underwater pelletizer (UWP) to produce the pellets.
- Pellets enter into a centrifuge to expel water and then are directed to the heated vibrating table at the Post-crystallization unit (PCU) to adjust the
- % crystallinity of the pellets.
- Pellets enter the Pre-heaters to adjust their temperature before entering to the Solid State Polymerization Reactor (SSP).
- The SSP holds each charge for 8 h and produce food-grade pellets by evaporating the volatile organic compounds (VOC).
- Pellets are then cooled in the cooler process where Energy Recovery Kit (ERK) is installed at the cooler to direct the hot air back to the hot air dryer.

Visual Illustration of the line is shown in Fig. 3.

2.4.2 Data Acquisition and Processing

Material representative data provided belongs to year 2019 as per the Table 1 [17].

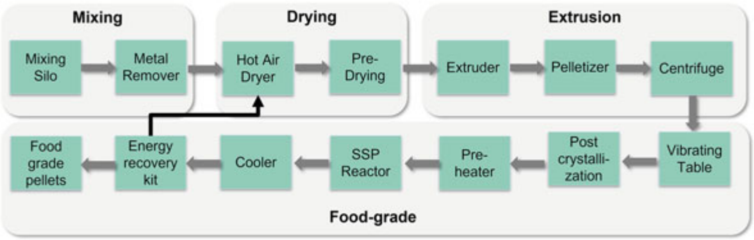


Fig. 3 rPET recycling line

Table 1 Material data (2019)

Year 2019	Total input flakes (kg)	Total output flakes (kg)	Metal flakes (kg)	Dust waste (kg)	PET waste purge (kg)
Σ_{Total}	14,192,168	13,725,444	101,346	13,557	351,821

92% of the total amount of output flakes is high quality food grade QPET and 8% is lower non-food grade APET. Metal flakes, dust and PET purge compose the process material waste stream.

Energy representative data provided belongs to year 2019 as per Table 2 [17]

Electricity in Egypt is based on Natural Gas which is a non-renewable energy source. Cost Data shown in Table 3 is calculated based on the values from [18] according to the current Egyptian electricity charges for industrial entities.

Process Inputs

(X01) Electricity consumption of the extruder: 2,407,896 Kwh.
(X02) Total extruder PET flake input: 14,192,168 Kg.
(X03) Anti-yellow: 9,357Kg.

Process Outputs

(Y00) Total extruder output: 13,725,444 Kg of PET blocks.

Table 2 Energy data (2019)

Year 2019	Total Energy consumption (kWh)	Extruder energy consumption (kWh)	Pre-heaters energy consumption (kWh)
Σ_{Total}	5,830,440	2,407,896	1,364,940

Table 3 Electricity cost data in Egypt [18]

	Amount	Emissions	Price
Electricity specifications in Egypt	1 (kWh)	0.55 kg CO_2	1.45 EGP/KW

(Y01) Total purge waste amount: 355,821 kg.

Process Parameters

EW: Purge waste which equals (X02-Y00) / X02 and initialized as a parameter with a fixed percentage equals 2.52%.
EP: Electricity consumption rate to X02 which equals X01 / X02, initialized as a parameter with a fixed percentage equals 17.06%.
AY: Anti-Yellow percentage per X02 which equals X03 / X02, initialized as a parameter with a fixed percentage.

Modelling

The main unit controlling the line is the extruder unit since it has the maximum energy consumption and is responsible of the material production and waste generation. The modelling equations for the extruder are as follows; [19]

I. **Y00 = X02*(1-EW/100)**: calculate the PET blocks output amount depending on the PET flakes amount and the Purge waste.
II. **X02 = Y00/(1-EW/100)**: inverse of 1 which is the direction of a known output amount to specify the required input amount, in this example, so this equation use PET blocks known amount to specify the PET flakes required amount.
III. **X01 = X02*(1-EP/100)**: calculate the electricity required amount depending on the known PET flakes amount and the electricity consumption rate to X02.
IV. **X03 = X02*(AY/100)**: calculate the required amount of Anti-yellow depending on the known PET flakes amount and the Anti-Yellow percentage per X02.
V. **Y01 = X02-Y00**: calculate the output waste depending on known the total PET flake input and the total extruder output.

2.4.3 Process Mapping

The processes are drawn with their inputs, intermediate products, wastes, and outputs to create a model that simulates the actual line processes as shown in Fig. 4.

Quantification of material and energy depends on the two types of processes specifications:

- *Linear specification:* used when linear relations (coefficients) between inputs and outputs are sufficient to determine all other energy and mass flows.
- *Parameterized specification:* linear specification cannot always be used to describe the relationship between inputs and outputs. For example, if defining the process depends on the waste ratio, in this case, it is easier to describe the relation by using mathematical functions and operators.

The rPET line shown in Fig. 4 is modelled over 4 main stages; Mixing stage, Drying stage, Extrusion stage and Food grade stage as described in 2.4.1.

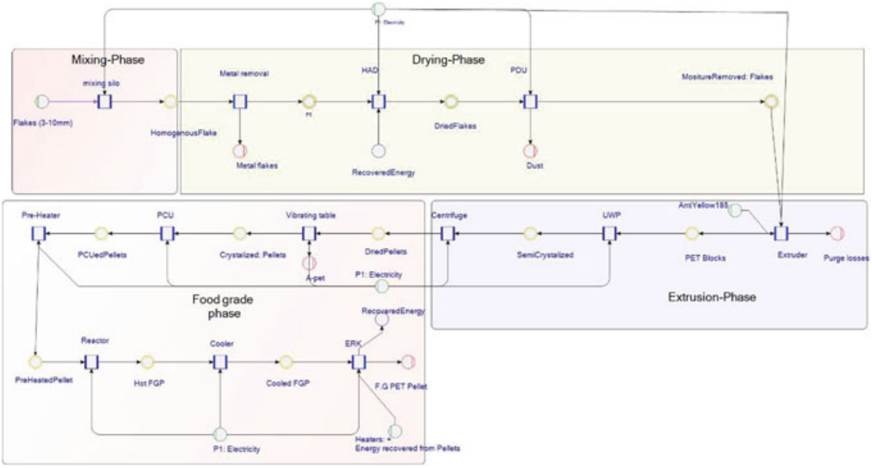


Fig. 4 rPET line mapping

Inputs and outputs (waste) are represented by green and red framed circles respectively. Processes are represented by the squares. Intermediate stages between processes are represented by the yellow circles. Each one of the four stages is enclosed in a frame for representation purposes.

The model highlights the main energy paths where electricity (P1) is provided. Another energy input is added to the HAD which is the energy recovered from energy recovery unit (ERK) at the cooling at the food grade stage.

Waste streams can be easily spotted (red circles) at metal removal, pre-drying unit (PDU) in the form of dust, purge material losses at the extruder, and APET at the vibrating table.

The storage is given by the blue circle representing the product of this process (QPET) at the end of the line.

3 Results and Analysis

3.1 Material Flow Analysis

The material flow analysis (green line in Sankey diagram) shows that for every 1 kg (PET), 0.967 QPET is produced. This makes a waste of only 3.3%, whereas based on the actual factory production, around 8% of the produced pellets are APET pellets as shown in the Sankey diagram in Fig. 5 marked Hotspot 1. This accounts as an indirect type of waste since the APET is of lower quality than QPET.

After process investigation, it was found out that there is a capacity difference between the SSP Reactor (1.618 tons) and the Extruder (1.759 tons). And since the

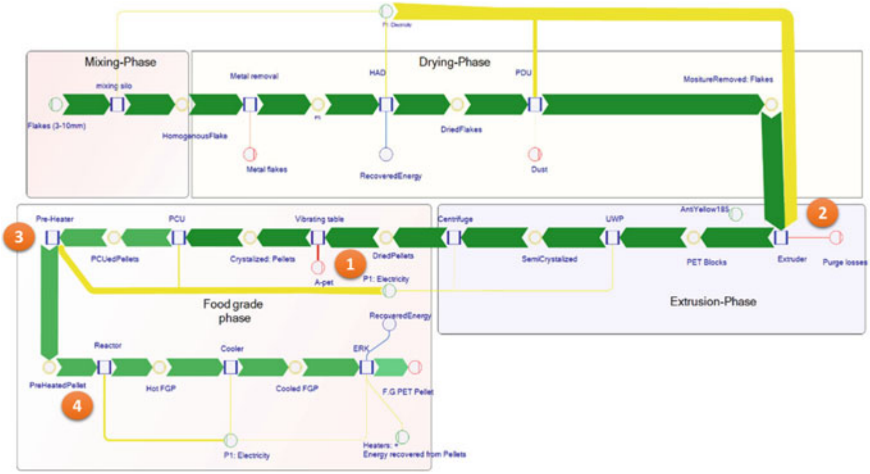


Fig. 5 Sankey Diagram of the rPET material flow identifying a bottle-neck

reactor takes a charge every 8 h, and considering a yearly operation time of 7800 h, so only 975 charges for a reactor are available. Consequently, with each charge the reactor takes there is 1126.2 kg bottleneck that will occur and hence (APET waste) is produced at the vibrating table not the Reactor itself which accounts for 1,098,045 kg APET (1126.2 kg multiplied by 975).

3.2 Energy Flow Analysis

After analysing the energy flow (yellow line in Sankey diagram), it is noticed that the main hotspots are the extruder (Hotspot 2) and pre-heater (Hotspot 3) as shown in Fig. 5. The total annual energy consumption is 5,830,440 KWh, the extruder consumption is 2,407,896 kWh representing 41.3% of the total electricity consumption. The pre-heater on the other hand, consume 23.4% of the total electricity consumption (1,364,940 kWh). The last hotspot is identified at the SSP reactor with a power consumption of 471,355 kWh. Total energy consumption calculated for all motors is 3,427,428 kWh/year representing around 58.8% of the total electricity consumption.

The high power consumption of the extruder is mainly located in the motor and the heaters, a deeper analysis in the extruder reveals the following;

- By analysing the motors' efficiency along the line, it was found out that the extruder's motor (315 kW output), is the highest consumption motor in the line with 1,496,880 Kwh/year. The motor is IE2 standard with a 95% efficiency equipped by 2 stages gearbox to increase the torque and to reduce the rpm which leads to a reduction in the net efficiency to 85%.

– The extruder is equipped by eleven ceramic band heater consuming a total of 911,016 kWh/year.

In the preheating process, pellets are heated up to 218 °C. Pellets exit the SSP reactor at 180 °C. To calculate the heat losses, the equation of heat transfer is applied; $Q = mc\Delta T$. Where m is the amount of the pellets in the reactor per hour which equals 1650 kg/h, c is the heat capacity of the material which equals 1.275 kJ/ Kg°C and ΔT is the change in the temperature which equals 218–180 = 38 °C. The heat loss = 22.265 Kw, by taking into consideration 7800 h/year operation time, a total heat loss of 173,208.75 KWh/year is estimated.

In the cooling process, the pellets enter the cooler at 180 °C and exit at 120 °C. The heat loss is estimated by 273,487.5 Kwh/year.

The calculated annual CO_2 eq. for the total power consumption is 3,206,741.99 kg.

3.3 Optimization Approaches

Based on the analysed material and energy flows and the identified hotspots, optimization approaches are implemented as follows;

Hotspot 1: Another standalone reactor with a capacity of 1 ton is to be added to the model to match the capacity of the extruder and hence increase the QPET production and decrease the APET.

For energy efficiency, four modifications are added to the model as follows;

Upgrading all line motors from high-efficiency motors (IE2) to premium efficiency motors (IE3).

Hotspot 2: New drive concept for the extruder main motor, depending on a direct drive/beltless motor [20].

Hotspot 2: Replacing the ceramic heater by Nano Infrared heaters with expected power saving of around 30% of the actual consumption [21, 22].

Hotspot 3&4: Energy recovery kit to make full use of waste heat, emitted from the reactor to the atmosphere. Expected energy saving up to 60% of the wasted energy in the reactor (103,925.25 Kwh/year) if the reactor process is supplied by an energy recovery kit.

The final optimized model estimated savings as shown in Table 4.

Table 4 Estimated annual savings

	Material (kg/year)	Electricity saved (kwh/year)	CO_2 Emissions saved (kg/year)	Cost saved in (EGP/year)
Total	1,098,045	583,440	320,674.2	845,413.8

4 Conclusion and Outlook

MEFA is a practical tool for the sustainability assessment of complex systems especially manufacturing and production lines. SMEs profit form the visualization and the time effective modelling and simulation of their processes using simulators such as UMBERTO® Efficiency + for assessment and analyzation of the manufacturing process using representative data. In this research, the PET recycling line is analyzed and process optimizations were simulated. The energy package consists of four modifications which increases the raw material resource efficiency by 8% and 10% of the actual energy consumption by 583,440 kWh/year which resembles an annual reduction of 2% of the total electricity consumed. Annual cost savings of around 845,988 Egyptian pounds (around 55 k$) are calculated. A reduction in CO_2 emissions by 320,892 kg annually is estimated.

The research also highlighted an important aspect which is the easy application of the process optimizations despite of missing data which could be retrieved from standard data sheets. As a result of the high discretion in the industrial field, it was not possible to compare the % savings estimated by the optimized model with similar ones.

Future work will focus on considering the cost of the process optimization and to calculate the feasibility of its implementation in terms of return of investment (RoI).

Acknowledgements This research is accomplished in frame of the project "SUSTAIN—Sustainable Production & Digital Transformation—Project ID: 57587884 "supported by the DAAD with funds from the German Federal Foreign Office. Appreciation for Eng. Omar Abouseada, Eng. Amr Sakr and Mr. Ahmed Nabil from BariQ for Techno and Advanced Industries for providing the technical data required for this research

References

1. Making trade an engine of growth for all: the case for trade and for policies to facilitate adjustment, International Monetary Fund, World Bank and World Trade Organization (2016)
2. The Global Manufacturing & Industrialisation Summit Outcomes Report, Paris Sorbonne University Abu Dhabi, United Arab Emirates (2017)
3. National Voluntry Review Report on the Sustainable Development Goals, High-level Political Forum (HLPF) on Sustainable Development, Egypt (2016)
4. Schmidt, M., Spieth, H., Haubach, C., Kühne, C.: 100 Pioneers in Efficient Resource Management Best practice cases from producing companies, Institute for Industrial Ecology INEC. Springer Spektrum, Germany (2018)
5. Baron, R., Alberti, K., Gerber, J., Jochem, E., Bradke, H., Dreher, C., Ott, V., Kristof, K., Liedtke, C., Acosta Fernandez, J., Bleischwitz, R., Bringezu, S., Busch, T., Kuhnt, M., Lemken, T., Ritthoff, M., Rosnow, J., Supersberger, N., Villar, A.: Study on the conception of a program for increasing material efficiency in medium-sized companies, final report (2005)
6. SMEs and the Environment in the European Union: European Commission. Publications Office of the European Union, Luxembourg (2010)
7. Edurne, L.L., Barral, M., Carlos Morla, J.: Methodology for carbon footprint calculation towards sustainable innovation in intangible assets. Vicomtech, San Sebastián, Spain (2020)

8. Olukorede, T., Adenuga, K., Mpofu, R., Innocent, B.: Energy efficiency analysis modelling system for manufacturing in the context of industry 4.0. Procedia CIRP **80**, 735–740 (2019)
9. Wohlgemuth, V., Schnackenbeck, T., Panic, D., Barling, R.-L.: Development of an open source software framework as a basis for implementing plugin-based Environmental Management Information Systems (EMIS). In: 22nd International Conference Environmental Informatics—EnviroInfo 22, Informatics for Environmental Protection Sustainable Development and Risk Management, September 10–12. Leuphana University Lüneburg. Shaker Verlag, Aachen, pp. 584–592 (2008)
10. Kleine-Moellhoff, P., Braun, A.T., Reichenberger, V., Seiter, S.: Approach to enable a material efficiency-strategy for small and medium sized manufacturing enterprises. Procedia CIRP **69**, 787–792 (2018)
11. Dimitris, M.: Simulation in the design and operation of manufacturing systems: state of the art and new trends. Int. J. Prod. Res. **58**(7), 1927–1949 (2020)
12. Czaplicka-Kolarz, K., Jerzy Korol, M., Pardała, L., Ponikiewska, K.: Material and energy flow analysis (Mefa) of the unconventional method of electricity production based on underground coal gasification. J. Sustain. Mining **13**(3), 41–47 (2014)
13. Manzhi, L., Shuai, T., Mengya, Z., Gang, H., Zhizhi, C., Zhiwei, F., Changjin, L.: Waste paper recycling decision system based on material flow analysis and life cycle assessment: a case study of waste paper recycling from China. J. Environ. Manage. **255** (2020)
14. Umberto® Efficiency+ User Manual. www.ifu.com. Accessed 15 July 2020
15. Schmidt, A., Hache, B., Herold, F., Götze, U.: Material Flow Cost Accounting with Umberto®. Chemnitz University of Technology, Chair of Management Accounting and Control (2013)
16. Material Flow Cost Accounting MFCA Case Examples, Environmental Industries Office, Industrial Science and Technology Policy and Environment Bureau, Ministry of Economy, Trade and Industry, Japan (2010)
17. Production Line Energy Consumption Data, BariQ for Techno and Advanced Industries (2019)
18. Abdallah, L., El-Shennawy, T.: Evaluation of CO_2 emission from Egypt's future power plants. Euro-Mediterr. J. Environ. Integration **49**(5) (2020)
19. Umberto® Efficiency+ software, Version10.0, ifu Hamburg, GmbH (2019)
20. Gemak Energy Saving Calculator. http://www.gamak.com/en-us/content/page/technical-support/energy-saving-calculator. Accessed 20 July 2020
21. PAMAENS Technology Co., Ltd. https://www.pamaens.com/infrared-nano-band-heater-for-extruder.html, Accessed 16 July 2020
22. Bella Zou, International business manager at Shanghai PAMAENS Technology Co., Ltd, Inquiry about Nano infrared heaters. Accessed 15 July 2020

Air Pollution Due to Central Heating of a City-Centered University Campus

Andreas Gavros and Kostas Karatzas

Abstract The aim of this study was to determine the gaseous pollutant concentrations resulting from the natural gas-powered central heating at the Aristotle University of Thessaloniki (AUTh) main campus, on the basis of limited emission and meteorological data. For this reason, a methodology was compiled addressing emissions and concentration levels as a function of a set of meteorological scenarios: Emissions were estimated based on campus operational conditions and concentration levels were calculated by employing Gaussian plume model approach via an in-house implementation in Python. The necessary climatic conditions were used to compile a sum of 1080 different weather–dispersion model scenarios. Obtained results allowed for the adequate analysis of the geospatial distribution of pollutants. Concentration levels were estimated to be below relevant limit values but dependent on the prevailing meteorological conditions.

Keywords Air pollution · Gaussian plume modelling · Meteorological scenario analysis · Python dispersion model

1 Introduction

The increased energy consumption by individuals, population growth and urbanization (more than 50% of world's population lives in cities) has affected the quality of air in urban metropolitan centers. Various air pollutants result from industrial activities, transportation, building heating needs or from energy consumption [1].

Nitrogen oxides (NO_x), Sulphur oxides (SO_x), Particulate Matter (PM), Carbon oxides (CO_x), Tropospheric Ozone (O_3) and Chlorofluorocarbons (CFCs) are among the most common air pollutants found in urban regions. They are responsible

A. Gavros · K. Karatzas (✉)
Environmental Informatics Research Group, School of Mechanical Engineering, Aristotle University, Thessaloniki, Greece
e-mail: kkara@auth.gr

A. Gavros
e-mail: andreasga@gmx.com

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_8

for a series of effects in human health and environment. Respiratory infections, Alzheimer's disease, various forms of cancer and reduced lung function have been associated with high concentrations of air pollutants. In European Union over 400,000 premature deaths associated with poor air quality are recorded each year. *Eutrophication*, as well as *acidification* of soil and water aquifers, result from excessive amounts of nitrogen oxides. Increased levels of ozone have devastating effects in the growth of plant organisms like crops and forests. Lastly, CO_2 is mainly responsible for intensifying the greenhouse effect which consequently causes climate change [2].

The city of Thessaloniki which is the subject of present study, is facing an alarming problem of deteriorating air quality in recent years. High concentration values of PM_{10} and $PM_{2.5}$ is a characteristic problem that emerges from the almost exclusive use of cars as a means of transport, in the absence of alternatives, (over 400,000 vehicles are used daily) [3]. Exceedances of limit values regarding PM_{10} pollution are often recorded: in 2019, 67 days recorded exceedances of relevant statutory limits [4]. It should be mentioned that the relative legislation limits for PM_{10} concentration are 50 $\mu g/m^3$ for daily averages (not to be exceeded more than 35 times per year) and 40 $\mu g/m^3$ for annual means [5]. Therefore, the estimation of air pollution levels in Thessaloniki is deemed necessary, especially considering the influence of meteorological conditions. For this reason, and on top of air quality monitoring stations, air pollution dispersion modeling may be used. This is a rather developed scientific field, and there are many mathematical frameworks describing transportation, dispersion and chemical transformation of pollutants [6]. Among the models that are frequently used, Gaussian Models are the most widely applied, being also easier in terms of implementation effort. Concerning models that have been used in various types of air quality studies, *AERMOD* and *CALPUFF* are among the most characteristic [7]. *CALPUFF* is a non-steady-state meteorological and air quality modeling system, which takes into consideration weather conditions, space and time. It offers the ability of analyzing various emission points, however the accuracy of its predictions seems to deteriorate regarding long-term modeling [8]. *AERMOD* is a steady-state model which provides dispersion capabilities for convective, stable or night boundary layers. It can analyze various meteorological conditions (humidity, temperature) and has been successfully applied in many case studies [9].

Building Heating (BH) is among the main polluting activities in urban areas, this being the case for a university campus located in a city center. Dispersity of fuel consumption within the campus, together with a change from heating oil to gas-powered burners, as well as the variating climatic conditions do not allow a straightforward estimation of resulting air pollution. In the frame of this research, we aim to quantify the overall air pollution burden pertained by central heating (CH) operation of AUTh main campus, considering the shift in the use of fuel (from heating oil to natural gas), as well as the need to identify the location of the maximum occurring concentrations, as a function of varying meteorological conditions. For this reason, we analytically estimate emissions per burner and fuel, and we employ the Gaussian plume dispersion approach [10] together with a number of meteorological parameter scenarios.

Fig. 1 Position of the flues abducting BH emissions in AUTh campus (each building's chimney number corresponds to numbering presented in Table 1) [13]

2 Materials and Methods

2.1 Study Area

AUTh is the largest academic institution in Greece with 75,000 students and 2,500 professors. Its main campus comprises of 33.4 hectares with a complex network of buildings and open spaces [11]. A unique characteristic of the campus is that it is located in the centre of Thessaloniki, which is the financial centre of Northern Greece, with 1,000,000 inhabitants. The city faces the sea to the south and southwest (Thermaikos Gulf), the Chortiatis Mountain to the southeast, and the Seich-Sou Forest to the northeast. The local climate is Mediterranean with hot, dry summers and mild, wet winters. Annual mean temperature is 15.6 °C, while mean relative humidity ranges from 53.2% (July) to 78% (December), with an annual mean rate of 67.3% [12].

In order to cover its BH needs, AUTh is using a system based on boilers, that are powered by natural gas burners to heat water which is then pumped into tubes and through a system of circulators, ends up in the radiator-bodies that transmit heat to campus' indoor facilities. Natural gas has been used in the last decade, replacing heating oil for both economic and environmental reasons. The burners' chimney locations and characteristics are described in Fig. 1 and Table 1 respectively.

2.2 Emissions Associated to BH

Natural gas, while much cleaner in environmental terms, compared to heating oil, does result in pollutant emissions when burned: water vapor and carbon dioxide are the main gasses emitted, which do not directly affect the environment but contribute

Table 1 Buildings of AUTh with chimney installations [17]

#	Building	Building height with its chimney (m)	Overall cross-section of chimney (m^2)
1	Management building	41.9	5.62
2	Faculty of education	47.25	0.26
3	Faculty of theology	27.68	4.54
4	Department of architecture	18.42	3.66
5	Faculty of engineering/building D	27.79	0.82
6	Faculty of engineering/building E13	18.45	0.20
7	Hydraulics building	21.68	0.21
8	University's gym	16.63	0.93
9	University student club	15	1

Table 2 Emission factors used in the current study [15]

Fuel type	PM_{10} (mg/kWh_{th})	CO_2 (gr/kWh_{th})	CO (gr/kWh_{th})	NO_x (mg/kWh_{th})	SO_x (mg/kWh_{th})
Natural gas	21	216	13	90	–

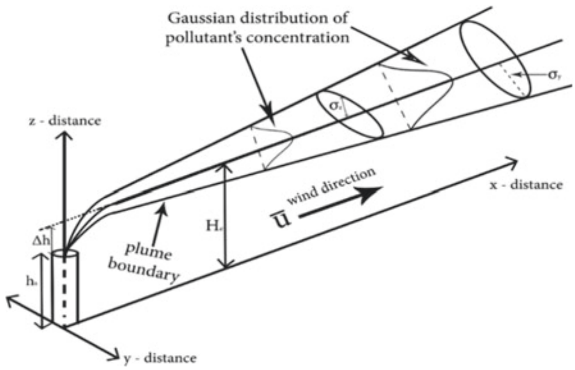
to greenhouse effect. Burning natural gas does produce nitrogen oxides, which are precursors to smog, but at much lower levels than heating oil [14]. Natural gas is considered to be the most environmentally friendly fuel and the least harmful towards public health (renewable energy sources excluded), especially when compared to heating oil [15]. There are no analytic emission data available for BH installations in AUTh campus, so relevant scientific literature was used to estimate actual emissions related to BH (Table 2).

2.3 *Gaussian Plume Model*

The air pollution modelling approach used, was the Gaussian atmospheric dispersion model, embracing Pasquill–Gifford atmospheric stability, which was selected on the basis of its simplicity of calculations, ease of use and applicability within the frame of the current study. The Gaussian approach, assumes that topography and atmospheric conditions are homogeneous throughout the dispersion area and result in the well-known Gaussian plume dispersion equation for point sources [16]:

$$C(x, y, z) = \frac{Q}{2\pi u \sigma_y \sigma_z} \cdot e^{-\frac{y^2}{2\sigma_y^2}} \cdot \left[e^{-\frac{(z-H_e)^2}{2\sigma_z^2}} + e^{-\frac{(z+H_e)^2}{2\sigma_z^2}} \right] \quad (1)$$

Fig. 2 Graphic illustration of Gaussian dispersion model



Here, C is the pollutant concentration in certain position, Q is the emission rate of the point source, u is the mean horizontal wind speed at the effective emission height, σ_y is the lateral dispersion coefficient function, σ_z is the vertical dispersion coefficient function, x is the downwind, y is the crosswind and z is the vertical direction [16], as indicated in Fig. 2.

Where h_s is the source height and Δh is the plume elevation after it is emitted from the source and H_e is the total height the plume is reaching (Δh and H_s combined). The two coefficient functions (σ_y, σ_z) are an expression of the standard deviation of dispersion and are related with downwind distance and atmospheric stability class. They were calculated with the use of Pasquill–Gifford method, which classifies atmospheric stability into six different classes (Table 3). Analysis on this case study was conducted essentially on daytime, in which the university operates, so categories F and G which correspond to night conditions were intentionally omitted. During daytime, stability depends on the intensity of solar radiation and wind speed at ground surface. Generally, when wind speed exceeds 6 m/s, atmosphere is considered neutral [18]. Coefficients k_1, k_2, k_3, k_4 and k_5, used to calculate σ_y and σ_z are presented in Table 4, while σ_y and σ_z are calculated as follows [19]:

$$\sigma_y(x) = \frac{\kappa_1 \cdot x}{\left[1 + \frac{x}{k_2}\right]^{k_5}} \sigma_z(x) = \frac{\kappa_4 \cdot x}{\left[1 + \frac{x}{k_2}\right]^{k_5}} \tag{2}$$

Table 3 Pasquill Stability classes [24]

Daytime	Surface wind speed (m/sec)	Incoming Solar Insolation		
Day		Strong	Moderate	Slight
	<2	A	A–B	B
	2–3	A–B	B	C
	3–5	B	B–C	C
	5–6	C	C–D	D
	> 6	C	D	D

Table 4 Values of coefficients k_1, k_2, k_3, k_4 and k_5 [18]

Dispersion class	k_1	k_2 (m)	k_3	k_4	k_5
A	0.25	927	0.189	0.102	−1.918
B	0.202	370	0.162	0.0962	−0.101
C	0.134	283	0.134	0.0722	0.102
D	0.0787	707	0.135	0.0475	0.465

2.4 Wind Speed Correction

Calculating the wind speed in the source point can be a critical factor on the accuracy of the model. During neutral atmospheric stratification, the mean wind speed ($\overline{U}_z$) at a height z, above the surface ground was estimated using the logarithmic wind law [20]:

$$\overline{U_z} = \frac{u_*}{\kappa} \ln\left(\frac{z_x}{z_0}\right) \tag{3}$$

where z_0 is the surface roughness length which was set to 1.3, as the urban topology of the city landscape in Thessaloniki's centre is residential with closely spaced block of buildings [21], z_x is the total height of plume in each case (H_e), κ is Karman's constant which was set to 0.4 [22] and u_* is friction velocity which can be calculated as:

$$u_* = \kappa \cdot u_w \cdot \ln \frac{z_w}{z_0} \tag{4}$$

where u_w is wind speed measured at a certain height above ground and z_w is height above the ground where measurement takes place, set at 10 m in this case study [23].

2.5 Meteorological Scenarios Examined and Fuel Type Comparison

In order to identify the maximum contribution of each CH installation to the concentration of pollutants in the Thessaloniki city center, we developed a grid of meteorological scenarios which include five different air temperatures and six different wind intensities. In addition, atmospheric stability, as dictated by incoming solar insolation, was also considered, as we added to the aforementioned meteorological conditions four different levels of sunshine and corresponding stability class: "A" stands for strong sunshine, "B" for light sunshine, "C" for slightly overcast conditions and "D" for cloudy sky.

Each one of the aforementioned scenarios was deployed for each one of the 9 university buildings (Table 1), resulting to a total of 1080 different scenarios (9 buildings × 5 different temperatures × 6 different wind speeds × 4 different levels of sunshine). The calculation of concentrations corresponding to the total grid of 1080 scenarios and the corresponding sensitivity analysis was performed with the aid of an inhouse gaussian dispersion model developed in Python programming language.

Calculations for all of these scenarios were repeated, but this time, we assumed that instead of natural gas, heating oil was used to cover the thermal needs of the university campus, in order to compare the pollutants concentration resulting from the two different fuels examined. The results of this analysis were then compared and a spatial analysis of the relevant air pollutant concentration fields (heatmaps) was performed. At all cases we assumed that wind direction was west–northwest, a common wind blowing in Thessaloniki during the winter season (known as *vardaris*).

2.6 Meteorological Scenarios Applied

As a step following the sensitivity analysis, we chose to deploy two different meteorological scenarios, which are more representative of climatic conditions in the city of Thessaloniki. Firstly, the heating period in Thessaloniki was assumed to last from November 1 until April 1. This period was further divided into 2 subperiods according to the outdoor temperature:

- Subperiod 1: December, January, February (duration: three months)
- Subperiod 2: November, March (duration: two months) [25] (Table 5)

Wind direction in urban area of Thessaloniki is primarily North and Northwest, with low wind speeds, ranging between 0.5 and 2.5 m/sec, [26] so we assumed a north-west wind direction, which is quite common during winter season for Central Macedonia. As a result of the above, the 2 scenarios chosen for each subperiod of heating were:

- *Scenario 1* for subperiod 1 (mean temperature: 5 °C & wind speed: 2.5 m/sec)
- *Scenario 2* for subperiod 2 (mean temperature: 11 °C & wind speed: 1.5 m/sec)

Table 5 Meteorological scenarios for the building examined

Scenario no	Temperature (°C)	Wind speed (m/sec)						Solar Insolation			
1	2	1.5	2.5	3.5	5.5	15	20	A	B	C	D
2	5										
3	8										
4	11										
5	14										

Concerning solar insolation categories, based on the case that Thessaloniki is experiencing a Mediterranean climate with mild winter seasons, we assumed sunny conditions for both scenarios (B for scenario 1 and A for scenario 2). Concentration of pollutants was calculated in proportion of total concentration derived from gaussian plume model for each of the fuel types, given the emission factors for each pollutant [27, 28].

2.7 Software Development and Sensitivity Analysis

In this case study we decided to avoid using one of the already available dispersion models, as the needs of this research were highly specialized. To avoid the high computational cost of investigating over 1000 different climatic scenarios, we decided to approach this problem by implementing a model designed from zero, to fully fit our needs. This approach provided us with the option of constructing a very dense network of receptors and analyzing a very big number of meteorological conditions in a fairly limited computational cost, that would not have been easily achievable with other models. As air pollutant dispersion modeling is a highly complex problem and each region has peculiarities that are hard to be fully adopted by general models, developing our own mathematical simulation would provide with the advantage of case-tailored calculations. In order to perform the numerical analysis of the data obtained from each one of the university buildings, we deployed several data analysis libraries such as *Pandas 1.0.3* and scientific computing libraries like *NumPy 1.18.2*.

In order to obtain the heatmaps that visualize the spatial distribution of the pollutants in color variation where each color stands for areas having similar concentrations, we used *Folium 0.10.1* which is a data visualization library in Python, built primarily to visualize geospatial data. Sensitivity analysis was performed by computing the pollutants' concentration at a spatial step which was set to be 10 m, moving away from the source point (building under examination). At each iteration we performed a further sensitivity analysis, calculating the concentration as we distance from the x-axis, at a distance step which was set as 50% of the standard deviation for the y-axis each time. The analysis was conducted until we reached a location with a distance 4 times greater than the standard deviation in y—distance. Analysis was further continued until we reached a point that the distance was 20 times greater than the point where the maximum concentration occurred, which gave an excellent overview of the spatial distribution of air pollutant concentrations.

Table 6 Emission factors for pollutants

Fuel	CO_2 (kg/GJ)	CO (g/GJ)	NO_x (g/GJ)	VOC (g/GJ)	SO_x (g/GJ)	PM (g/GJ)
Natural gas	56.06	30	40	2	0.3	0.9
Heating oil	74	130	942	50	48	60

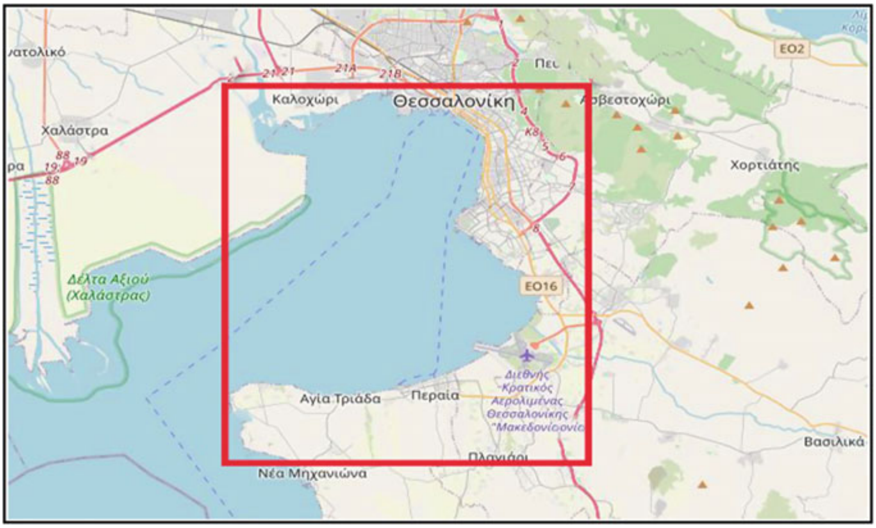


Fig. 3 Analysis area (delimited in the red rectangle)

We chose and set-up an analysis area, starting from a standard northwest point (40.645443 N, 23.006214 E) located just outside of the urban district, northern of the city's historic centre. We divided the analysis space into square kernels with dimensions of 20 m per side (20 × 20). We then scanned the whole analysis space with this kernel, starting from the standard point and reaching a southeast point (40.467621 N, 22.811892 E) at the outskirts of town (Fig. 3). In each kernel's centre we set up a hypothetical receptor that estimated the concentration of pollutants resulting from each one of the nine point-sources (buildings) separately. Additionally, each receptor has the ability to provide us with the sum of the concentrations of each building that is located inside the kernel.

In this project, we made use of *OpenStreetMap* (OSM) in order to visualize our geospatial data, resulting from the sensitivity analysis. OSM is an open data source which can be used freely, as it is a volunteered and collaborative geographic information (VGI) project, in which millions of citizens and scientists, have contributed into creating an online vector map of the world [29].

The developed simulation model is freely provided for scientific/academical research.[1] The majority of Python models developed are available in order to fit specific needs of future researches.

[1] https://github.com/andreasgav/Novel-Gaussian-Dispersion-Model-for-urban-climate-studies.

3 Results

3.1 Calculation of Maximum Concentration Levels

Results from the sensitivity analysis of the 1080 different climatic scenarios examined, indicated that most of minimum values of concentrations were estimated, as expected, for the meteorological scenarios with the lowest temperatures and wind speeds, in addition with cloudy weather (Solar Insolation category C). On the other hand, the highest concentrations of pollutants were estimated for windy conditions, strong sunshine and higher temperatures, as it can be seen in Table 7.

As expected, concentrations of pollutants resulting from heating oil combustion are relatively higher, compared to the concentrations resulting from natural gas when used as fuel. In Table 8 it is evident that pollutant emissions abstain by a large margin given a comparison between the two different fuels, examined in this case study. Heating oil emissions are far greater, compared to these of natural gas, in all the types of pollutants, especially concerning these with the greater impact on human health and urban environment/climate (NO_x, PM, etc.).

Table 7 Weather conditions for highest concentrations, for each building

Building		T (°C)	Wind Speed (m/sec)	Solar Insolation
Management building	min	2	1.5	C
	max	14	20	A
Faculty of Education	min	2	20	D
	max	14	2.5	A
Faculty of Theology	min	2	15	C
	max	14	20	A
Department of Architecture	min	2	1.5	C
	max	14	20	A
Building D	min	2	1.5	C
	max	14	15	A
Building E13	min	2	1.5	C
	max	14	5.5	C
Hydraulics Building	min	2	1.5	C
	max	14	20	A
University's gym	min	2	1.5	C
	max	14	20	A
University Student Club	min	2	1.5	C
	max	14	20	A

Table 8 Highest concentration levels ($\mu g/m^3$) for building/fuel (NG: Natural Gas, HO: Heating Oil), regarding worst case scenario

Building	Fuel	Distance of max (m)	CO_2	CO	NO_x	VOC	SO_x	PM
Management building	NG	418.5	343.8	0.182	0.24	0.012	0.018	0.005
	HO		489.5	0.86	6.23	0.44	0.318	0.397
Faculty of Education	NG	371.6	407.8	0.216	0.29	0.015	0.022	0.007
	HO		580.8	1.022	7.393	0.518	0.377	0.471
Faculty of Theology	NG	333.5	272.3	0.145	0.194	0.01	0.015	0.004
	HO		387.9	0.683	4.937	0.346	0.252	0.315
Department of Architecture	NG	319.7	27.5	0.015	0.020	0.001	0.001	0.0004
	HO		39.2	0.069	0.499	0.035	0.026	0.032
Building D	NG	265.4	154.1	0.082	0.110	0.006	0.008	0.002
	HO		219.5	0.386	2.794	0.196	0.143	0.178
Building E13	NG	419.4	2553	1.356	1.820	0.091	0.137	0.041
	HO		3636	6.399	6.281	3.243	2.363	2.948
Hydraulics Building	NG	368.8	210.5	0.112 lePara>	0.150	0.008	0.011	0.003
	HO		299.7	0.527	3.815	0.267	0.195	0.243
University's gym	NG	208.2	76.9	0.041	0.055	0.003	0.004	0.001
	HO		109.5	0.193	1.394	0.098	0.071	0.089
University Student Club	NG	336.4	31.7	0.017	0.023	0.001	0.002	0.0005
	HO		45.2	0.080	0.575	0.040	0.029	0.036

3.2 Results of Representative Climatic Scenarios

Concentration levels reported in Sect. 3.1 correspond to "extreme" weather conditions, which are unlikely or rarely seen, so we deployed *scenario 1* and *2*, in order to obtain results that are more realistic and can be recorded on a regular winter season day. As a result, we obtained a new set of maximum values for the pollutants concentrations, in which distances from the source points are similar for both scenarios examined here. This can be explained by the fact that the climatic parameters entered into the model, are almost identical, especially the wind speed velocities, which have a greater impact on Gaussian distribution.

In order to achieve better reproducibility of the results shown in this work, we have chosen one of the pollutants examined to be set as the base pollutant. This was NO_x, which is among the most common pollutants in urban environments. We then examined how the pollutants' concentration distribution vary, according to 2 different parameters. The first was how the concentration of the pollutant fluctuates, while it distances from the source point in the central distribution axis (x–distance), before and after the maximum point. Figure 5 demonstrates that maximum concentrations are located approximately 1000 m. from source points. After the point of maximum concentration, values are gradually reducing and almost zeroed at distances about 3000 m.

The second was an analysis of concentration's reduction as the distance from main propagation axis rises. Concentration reached almost zero levels at a distance equal with *3*standard deviation distance*, according to normal Gaussian distribution. Examples of these results are presented in Figs. 4 and 5, also reflecting the significant difference between concentration levels related to the two fuel types examined.

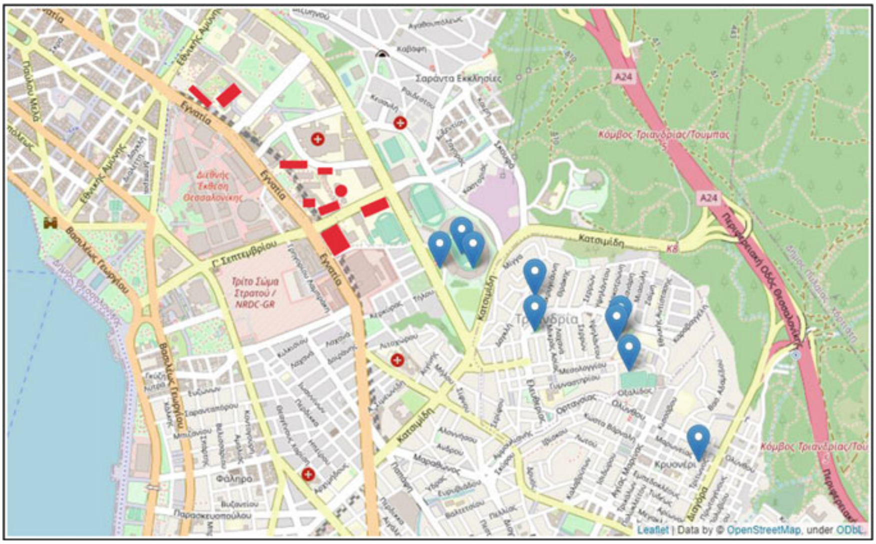


Fig. 4 Points in Thessaloniki were maximum values of concentrations occur (scenario 1). Buildings examined are highlighted in red

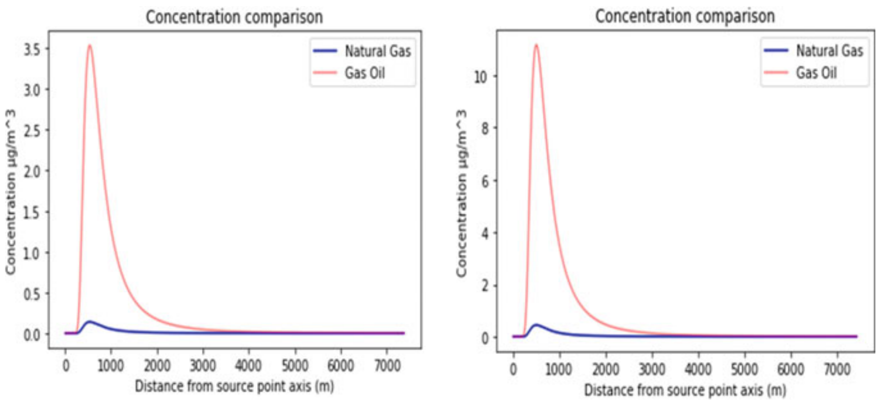


Fig. 5 Fluctuation of NOx's concentration as the distance from source point increases for scenario 2 (left: Hydraulics Building, right: Faculty of Education)

3.3 Heatmaps of Pollutants Concentrations

We constructed heatmaps for the concentrations resulting from the emissions of all the buildings aforementioned, in order to investigate their spatial distribution. It is evident from the results that the higher concentrations are located in a circular region between 1000 and 2000 m from the point source, (yellow region in natural gas map and light red in heating oil map). As the distance increases from these regions, it becomes obvious that the concentration levels are gradually being reduced, as expected. The analysis was repeated for each campus building separately in order to produce a more targeted visualization of the pollutants' concentration (Figs. 7 and 8).

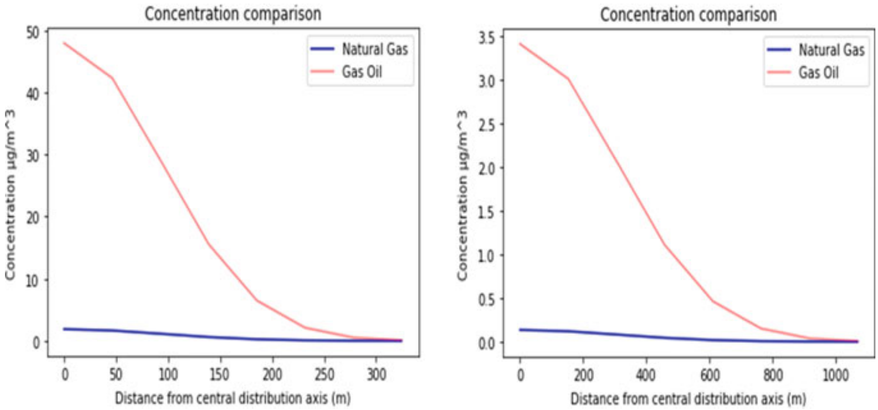


Fig. 6 Fluctuation of NOx's concentration as we distance from central axis for scenario 1 (left: Management building, right: Building E13)

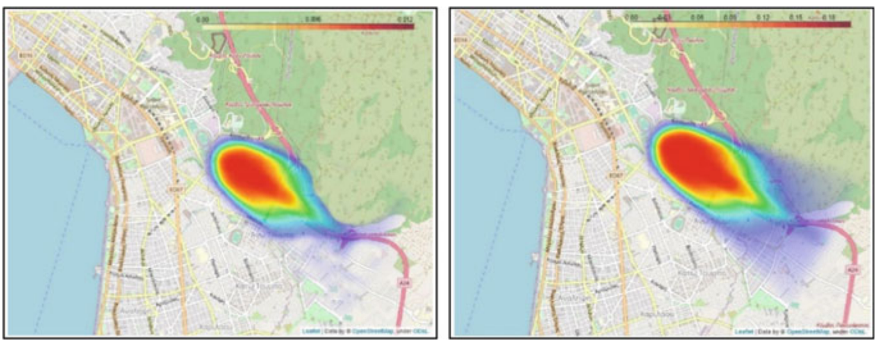


Fig. 7 Heatmap of NO_X concentration for Department of Architecture in scenario 2 (left: natural gas emissions, right: heating oil emissions). Differences in focal point of high concentrations are located in variations of red color which displays the maximum values. In the case of heating oil, we have darker shades of red compared to the ones of thematic map of natural gas

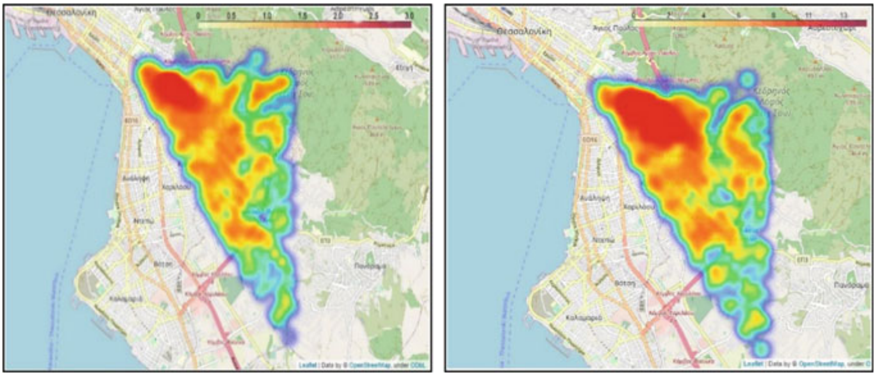


Fig. 8 Heatmap for NOx concentrations (emissions of all buildings) for scenario 1 (left: natural gas, right: heating oil). Color scale indicates the variations in pollutants' concentrations (Blue and green shades indicate low NO_x concentrations. Yellow and red shades are indicating higher values of concentration. Darker tones of red, stand for the maximum values that were computed). Legend scale corresponds to concentration values of the heatmaps

Borderlines of study area can easily be recognized in heatmaps of Fig. 8, explaining the linearity of distribution observed in the borders of maps concerning all source points. In this case, the pollutants form a practically straight border line in eastern part of the map, while concentrations in the vicinity of the studied area are relatively small.

Finally, we created heatmaps for both scenarios and fuel types analyzed, in which the difference between natural gas and heating oil is again obvious and can easily be observed. Apparently, NO_x concentrations resulting from heating oil were significantly higher, compared to these of natural gas, and this difference is also depicted in heatmaps of Fig. 8, where the area of pollutants' spatial expansion is bigger in the case of heating oil and concentrations are again relatively higher.

The highest values of NO_x concentrations, resulting from natural gas emissions, were calculated at a distance range of 400–1540 m, with the maximum NO_x concentration calculated overall, reaching 2.806 $\mu g/m^3$. Faculty of Education had the nearest-to-source highest concentration value (0.439 $\mu g/m^3$) at a distance of 510 m. The lowest concentrations were calculated for the buildings of Architecture and University Student Club (0.008 $\mu g/m^3$), while the highest concentration was observed for Building 13 (1.885 $\mu g/m^3$) which was located in a distance of 400 m. from the source.

4 Discussion

The use of natural gas as a fuel in order to meet the heating needs of such a large building area as the Aristotle University campus, has multiple environmental benefits.

Its use puts a lower burden on the urban environment surrounding the university in terms of air pollution, when compared to heating oil powered building heating. Concentrations of pollutants are significantly less when the two fuels are compared, sometimes two orders of magnitude, for pollutants like SO_x.

Based on the meteorological scenarios examined in this study, we observed that the urban areas of *Triandria* and *Toumpa* (both located in the Northwest region of the city) are facing the larger environmental burden, coming from the heating operations of AUTh. However, this result is based on the hypothesis that a northwestern wind is blowing in the city. If we consider a slightly different prevailing wind direction (northern wind), the urban areas mainly affected will be *Charilaou* and *Kalamaria* (two urban districts that are located in the eastern Thessaloniki and are densely constructed and populated). Concerning the spatial distribution of the pollution, it should be mentioned that the straight border lines observed in heatmap depicted in Fig. 8 could be softened by expanding the analysis area, but given the sensitivity analysis step which was set to 20 m, this would have a great impact on the computational cost of this research. So as a result, we chose to use the initial borders of the study, in favor of a more flexible computational model.

The resulting environmental burden is relatively low and it does not significantly contribute to air pollution of the city of Thessaloniki. Nevertheless, as the city in facing new challenges related to existing, as well as recently added emission sources (the latter due to biomass usage for household heating purposes), it is important to investigate the meteorological conditions that could lead to an increasing air pollution problem. On this basis, the method developed in this case study can be deployed in similar case studies to precisely identify the overall environmental burden as the spatial distribution of pollutants may cause non-negligible problems within the urban area.

It should be noted that the predicted values by our model indicate that AUTh is not significantly contributing in the air pollution profile of Thessaloniki's city centre. Air quality measurements by two monitoring stations located in the city centre (*Aristotelous* and *Agia Sofia* stations) support this argument. Indicatively, the hourly averages regarding PM_{10} (one of the pollutants addressed in this study) for the period of January, February, December (subperiod 1) are 47.64 $\mu g/m^3$ in *Agia Sofia* station and 36.3 $\mu g/m^3$ for *Aristotelous* station. Regarding November and March (subperiod 2) the hourly averages are 45.1 $\mu g/m^3$ (*Agia Sofia*) and 39.69 $\mu g/m^3$ (*Aristotelous*). These measurements (recorded in 2018) [30] suggest that although, high air pollution concentration values are present, the pollutants resulting from AUTh heating system (in the case of PM_{10} less than 1 $\mu g/m^3$), are not significantly contributing to the existing problem.

5 Conclusions

The maximum concentration values for natural gas emissions in comparison to heating oil emissions depict the importance of using more environmentally friendly forms of energy to meet heating or cooling needs, in urban areas. The significantly higher concentrations of pollutants when heating oil is used as a fuel, highlights the need to replace heating oil for BH purposes. Overall concentration levels in both cases are low, especially when compared to existing ones resulting from air quality measurements.

A major limitation of the modelling approach used is related with the assumption that the pollutants addressed are considered to be chemically inactive and that no additional pollution sources and sinks are considered. Moreover, the terrain elevation and its relevant land uses are not included in the analysis. Nevertheless, the applied method in this paper can easily address multiple meteorological scenarios and may lead to results that can serve as the basis for additional investigations. Of course, a research based on data coming from various sources (transportation etc.), could give us a more accurate profile of the environmental impact of the university on the city centre. As a result, future research should focus on gathering various sources of information/data in order to produce results of higher spatial and temporal accuracy.

References

1. Changotra, R., Rajput. H., Dhir. A.: Comparative study of air pollution modeling techniques from point source(s) of thermal power plant. Environ. Model. Assess. **25**, 531–543 (2020)
2. European Environment Agency: Air quality in Europe—2019 report. Publications Office of the European Union, Luxembourg (2019)
3. Kyriakidis, I., Karatzas, K., Papadourakis, G., Ware, A., Kukkonen, J.: Investigation and forecasting of the common air quality index in Thessaloniki, Greece. Artif. Intel. Appl. Innov. 390–400 (2013)
4. European Commission: Air quality: Commission decides to refer GREECE to the Court of Justice over poor air quality. European Commission—Press release (2020)
5. European Union: Directive 2008/50/ec of the European parliament and of the council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official J. Eur. Union (2008)
6. Khan, S., Hassan, Q.: Review of developments in air quality modelling and air quality dispersion models. J. Environ. Eng. Sci. **16**, 1–10 (2021)
7. Nebenzal, A., Fishbain, B., Kendler, S.: Model-based dense air pollution maps from sparse sensing in multi-source scenarios. Environ. Model. Softw. **128**, 104701 (2020)
8. Holnicki, P., Kałuszko, A., Trapp, W.: An urban scale application and validation of the CALPUFF model. Atmos. Pollut. Res. **7**, 393–402 (2016)
9. Kakosimos, K., Assael, M., Katsarou, A.: Application and evaluation of AERMOD on the assessment of particulate matter pollution caused by industrial activities in the Greater Thessaloniki Area. Environ. Technol. **32**, 593–608 (2011)
10. Brusca, S., Famoso, F., Lanzafame, R., Mauro, S., Marino Cugno Garrano, A., Monforte, P.: Theoretical and experimental study of gaussian plume model in small scale system. Energy Procedia **101**, 58–65 (2016)
11. Pitsiava-Latinopoulou, M., Basbas, S., Gavanas, N.: Implementation of alter-native transport networks in university campuses: the case of the Aristotle University of Thessaloniki, Greece. Int. J. Sustain. High. Educ. **14**, 310–323 (2013)
12. Slini, T., Papakostas, K.: 30 Years Air Temperature Data Analysis in Athens and Thessaloniki, Greece. Green Energy and Technology. Springer International Publishing, pp. 21–33 (2015)
13. Gavros A., Karatzas, K.: Microclimate profile and contribution of air conditioning to local heat island effects at the Aristotle University of Thessaloniki main campus. In: IOP Conference Series: Earth and Environmental Science, vol. 410, p. 012003 (2020)

14. Bruce, N.: Adair-Rohani, Puzzolo, E: WHO Indoor Air Quality Guidelines: Household Fuel Combustion Review 4: Health effects of household air pollution (HAP) exposure, 1st edn. WHO Document Production Services, Geneva (2014)
15. Paulrud, S., Gustafsson: T. Emission factors and emissions from residential biomass combustion in Sweden. Swedish Environmental Protection Agency, Norrköping (2010)
16. Leelossy, A., Molnar, F., Izsak, F., Havasi, A., Lagzi, I., Meszaros, R.: Dispersion modeling of air pollutants in the atmosphere: a review. Central Eur. J. Geosci. **6**, 257–278 (2014)
17. Gavros, A.: Environmental implications from the operation of the central university campus of AUTh. Master Thesis, Aristotle University of Thessaloniki (2018)
18. Aggelidis, D.I.: Numerical Study of Wind Flow Characteristics and Pollutant Dispersion in a Typical Urban Street Canyon of Athens. PhD Thesis, National Technical University of Athens (2010)
19. Isaksson, M., Rääf, C.: Environmental Radioactivity and Emergency Preparedness, 1st edn. CRC Press (2017)
20. Kent, C., Grimmond, S., Gatey, D.: Aerodynamic roughness parameters in cities: inclusion of vegetation. J. Wind Eng. Ind. Aerodyn. **169**, 168–176 (2017)
21. Grimmond, C.S.B., Oke, T.R.: Aerodynamic properties of urban areas derived from analysis of surface form. J. Appl. Meteorol. **38**, 1262–1292 (1998)
22. Howard, T., Clark, P.: Correction and downscaling of NWP wind speed forecasts. Meteorol. Appl. **14**, 105–116 (2007)
23. De Visscher, A.: Air Dispersion Modeling: Foundations and Applications, 1st edn. Wiley (2013)
24. Vallero, D.: Fundamentals of Air Pollution, 4th edn. Academic Press, New York (2008)
25. Karagiannidis, A., Karatzas, K., Moussiopoulos, N., Sahm, P.: The role of temporal operation profiles on the study of SO_2 pollution from domestic heating in Thessaloniki, Greece. Trans. Ecol. Environ. **6**, 31–40 (1995)
26. Kantzioura, E., Kosmopoulos, P., Dimoudi, A., Zoras, S.: Experimental investigation of microclimatic conditions in relation to the built environment in a central urban area in Thessaloniki (Northern Greece): a case study. Sustain. Cities Soc. **19**, 331–340 (2015)
27. European Environment Agency: EMEP/EEA Air Pollutant Emission Inventory Guidebook, 1st edn. Publications Office of the European Union, Luxembourg (2019)
28. Nielsen, M., Nielsen, O.K., Plejdrup, M.S.: Danish emission inventories for stationary combustion plants. Inventories until 2015. Aarhus University, DCE – Danish Centre for Environment and Energy. Scientific Report from DCE—Danish Centre for Environment and Energy No. 279 (2018)
29. Fonte, C., Lopes, P., See, L., Bechtel. B.: Using OpenStreetMap (OSM) to enhance the classification of local climate zones in the framework of WUDAPT. Urban Climate **28**, 100456 (2019)
30. European Environment Agency—European Union: Download of air quality data. https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm. Accessed 16 July 2021

EpiDesktop—A Spatial Decision Support System for Simulating Epidemic Spread and Human Mobility Trends Under Different Scenarios

Ahmed Derdouri and Toshihiro Osaragi

Abstract Human mobility has been recognized as one of the critical factors determining the spread of contagious diseases, such as SARS-CoV-2, a highly contagious and elusive virus. This virus disrupts the normal lives of more than half of the global population in one way or another, claiming the lives of millions. In such cases, mobility should be managed via the imposition of certain policies. This proposed study presents a newly developed spatial platform aimed at simulating and mapping the spread of infectious diseases and mobility patterns under different scenarios based on different epidemiological models. In addition to the "business as usual" scenario, other response scenarios can be defined to reflect real-world situations, taking into consideration various parameters, including the daily rise in infections and deaths, among others. The developed system provides insights to decision-makers about strategies to be implemented and measures for controlling the spread of the virus.

Keywords Agent-based modeling · GIS-based decision support system · COVID-19 · Visual analytics

1 Introduction

Throughout human history, epidemics have killed millions of people and have affected the lives of millions. Until now, epidemics have been among the biggest threats facing humanity despite modern advancements in medicine and scientific research. The ongoing COVID-19 pandemic is a pertinent example. It has been linked to more than 196 million confirmed cases and 4.19 million confirmed fatalities as of July 30, 2021 [1]. The tremendous economic losses and negative social impacts resulting from it make it one of the most severe pandemics in history.

A. Derdouri (✉) · T. Osaragi
School of Environment and Society, Tokyo Institute of Technology, 2-12-1-M1-25 Ookayama, Meguro-Ku, Tokyo 152-8550, Japan
e-mail: derdouri.a.aa@m.titech.ac.jp

T. Osaragi
e-mail: osaragi.t.aa@m.titech.ac.jp

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_9

The fast spread of infectious airborne diseases has been unavoidable in recent decades for many reasons. One reason is the fast human population growth. The world population jumped from 3 billion in 1960 to over 7.5 billion in 2021 [2]. Another reason has been the dramatic shift in population from rural to urban areas. This mass migration has led to an increase in densely populated regions. In 2020, 4.4 billion people were residing in urban settings, making cities home to over half of the world's population (56%) [3]. Moreover, over 40% of the population in 26 nations and territories lives in cities with populations of at least one million [3]. These densely populated areas are well connected by dense public transportation networks of different means, including flights, trains, and buses. The mobility of people through these networks leads to the rapid spread of infectious airborne diseases, such as COVID-19.

As a result of extensive investigations, many researchers have singled out human mobility for its crucial role in the spatiotemporal propagation of infectious airborne diseases. For instance, Hadjidemetriou et al. [4] found that a reduction in human movement had a substantial effect in reducing COVID-19-related mortality in the UK. In the same vein, Zhou et al. [5] noted that limiting human movement in the city of Shenzhen, China, had a significant impact on managing the COVID-19 epidemic. Few studies, however, have focused on developing geographic information system (GIS)-based platforms for simulating infection spread based on mobility trends that are designed to help decision-makers gather insights and make informed decisions under different scenarios. This project fills this gap. In this paper, we introduce EpiDesktop, a newly developed spatial decision support system (SDSS) for simulating infectious disease spread and mobility patterns under different response scenarios.

The remainder of this article is organized as follows. Section 2 provides a comprehensive literature review highlighting other GIS-based decision support systems. Section 3 describes the fundamental concepts behind the developed system. Section 4 details its architectural design. The implementation of the application is discussed in Sect. 5, and in addition to conclusions, Sect. 6 provides a discussion on future directions for this project.

2 Past Work

Several GIS-based applications for epidemic simulation have been developed in the last two decades, mainly due to the progress that has been made in GIS capabilities, hardware performance, etc. Table 1 lists some of the major simulation platforms that have been developed in the last twenty years, including a short description, their presented case studies, and their coverage areas. A detailed overview of each study is given in the following paragraphs.

One of the first attempts to develop an epidemiology-related SDSS dates back to 2005 when Chu et al. [6] developed a multi-criteria decision system integrating spatial and non-spatiotemporal variables to monitor disease spread. The developed

Table 1 List of developed GIS-based epidemic simulation platforms

Study	Description	Case study	Coverage
[6]	Multi-Criteria SDSS Architecture	The spread of SARS in 2003 in Hong Kong	Country (Hong Kong)
[7]	Visual analytics-based SDSS for epidemic modeling and response evaluation	Pandemic Influenza Rift Valley Fever	Country (US)
[8]	GLEAMviz: Epidemic evolutions simulator at the global scale	An example of the mitigation of an emerging influenza pandemic in North America and Western Europe	Worldwide
[9]	PanViz: A pandemic influenza modeling and visualization tool	Pandemic spread originating in Chicago and its effects on the US	Country (US)
[10]	EpiSimS: An epidemic simulation system	Mosquito-borne disease (500,000 persons affected)	City (Washington DC, US)
[11]	ID-Viewer: A visual analytics framework for infectious disease monitoring and response control	Hemorrhagic fever outbreak in Lahore city in 2011	City (Lahore, Pakistan)
[12]	PandemCap: A decision support system for managing the epidemic spread	Influenza outbreak in the French region of FR22	Region (FR22 region, France)
[13]	A WebGIS application for visualization of COVID-19 simulation	The spread of COVID-19 in Wuhan city	City (Wuhan, China)

system was tested via an analysis and display of the spatial data of Hong Kong's severe acute respiratory syndrome (SARS) in 2003.

Afzal et al. [7] proposed a decision support environment with which the impacts of numerous mitigation strategies starting at different points over time were compared. The results can be compared either through a spatiotemporal model view or a decision history view. While the former offers the possibility to investigate the effects of the mitigation actions spatiotemporally, the latter allows the user to add decision points to a timeline and subsequently explore the impacts of different countermeasures by assessing the number of saved and lost lives.

Another example of such developed systems is GLEAMviz [8], a freely accessible software that simulates epidemic spread based on stochastic mathematical models and population and mobility data. It enables users to create compartmental models that include mitigation measures. The produced model may be sent to the GLEAMViz servers, which run hundreds of simulations. The result is an interactive map and a collection of charts that statistically illustrate the disease's spatiotemporal progression. The program is set up as a client–server architecture. Users may install the program on their local workstations to create simulations run on the server side, obviating the need for substantial computing capabilities on the user's end.

PanViz [9] is another visualization and modeling tool for supporting public health decision-making. It is an influenza modeling platform that employs spatiotemporal views to aid public health authorities to comprehend how decisions affect the transmission of the disease.

Bryan et al. [10] developed an agent-based model for disease spread within the United States. They used the modeling of a mosquito-borne illness, with around 500,000 people affected, in the Washington, DC region as a case study.

Ali et al. [11] developed a system called ID-viewer capable of detecting epidemic outbreaks in Pakistan by processing real-time streaming data from emergency department visits, emergency calls, and medicine purchases.

PandemCap, developed by Yaez et al. [12], is another decision-making tool that public health officials may use to make better, well-informed decisions during pandemics. PandemCap employs visual analytics and epidemic modeling techniques to provide decision-makers in public health with an interactive and adaptable application. Furthermore, the tool enables the analysis of the effects of different treatments or control measures, such as vaccinations, antivirals, hospital beds, and ventilators.

In the context of the ongoing COVID-19 pandemic, to the extent of our knowledge, few standalone spatial applications have been developed to simulate its propagation on a local scale (e.g., within cities). The majority of available GIS-based applications consist of web applications in the form of dashboards that monitor cases/fatalities on a global scale, such as the famous COVID-19 Map developed and updated by Johns Hopkins University [1]. Simulation-wise, several applications or packages have been developed, such as COVID-ABS [14], which is an agent-based model designed to predict health and economic impacts based on seven social distancing scenarios (e.g., lockdown, use of masks, etc.). However, most of these developed applications do not include a spatial component to visually describe the spread. Few attempts have been made in this regard. For instance, in [13], Lipeng et al. developed a COVID-19 simulation webGIS platform to simulate and visualize epidemic propagation based on intra-urban travel intensity. A case study on Wuhan was provided. This application, however, lacks an analysis of the spread under different scenarios. In summary, aside from simple COVID-19-related platforms that are designed to monitor and display basic information on a global scale, the developed simulation applications present at least one of the following two limitations: (i) they lack a spatial component to visualize the spread and/or (ii) simulation is based on a single scenario. The developed prototype described in this paper seeks to fill these gaps by providing a GIS-based platform that is designed to simulate the spread on a local scale based on mobility patterns under different scenarios flexibly defined by the user.

3 Fundamental Concepts

3.1 Agent-Based Modelling

Agent-based modeling (ABM) is a computational approach for simulating the behavior of autonomous agents and their interactions. ABM simulations aim to replicate different real-world scenarios and assess their effects on a given system or a spatial environment. Each agent has its own set of distinguishing characteristics, and behavioral traits govern what an agent will do in certain situations [15].

Agent-based models may be used in large urban-scape simulations based on queuing theory, which may provide information on more intricate levels, thanks to improved computer performance [16]. In recent years, a rising number of agent-based models have been developed for a range of application fields. Modeling agent behavior in stock markets, supply chains, and battle simulations, as well as anticipating the spread of diseases, biological warfare threats, and real-world business challenges, are all examples of applications [17]. Obtained results from such applications are vital in the decision-making process, especially for effective crisis response management.

3.2 Epidemic Modelling

The spread of infectious diseases through a population is defined by epidemic models, which are simplistic depictions of reality [18]. These models permit the monitoring of the progression of an epidemic, measuring the effects of mitigating interventions, and aiding in emergency preparedness and risk management [12].

Among the most popular models that have been used in previous works are SIR (Susceptible, Infected, and Recovered), SEIR (Susceptible, Exposed, Infected, and Recovered), and SEIRD (Susceptible, Exposed, Infected, Recovered, and Dead). The SIR model was introduced by Ronald Ross and William Hamer and was initially applied by Kermack and McKendrick in 1927 [19]. The SIR has been employed for a variety of infectious viruses, such as measles and airborne diseases. The SEIR model was developed by Aron and Schwartz. It has been utilized for a wide range of diseases, including HIV and, more recently, COVID-19 [19].

These models assume that in a given population, individuals are divided into several compartments based on their disease-transmitting dynamics as follows:

- Susceptible (S) people are yet to be exposed and infected.
- Exposed (E) people have contracted the virus but are not yet infectious.
- Infected (I) represents an infectious group of persons.
- The Recovered (R) fraction is made up of those who have recovered from their illness and are now immune to it.
- The Dead (D) fraction represents people who have died from the disease.

Models estimate the number of individuals in every group with the assumption that population size (N) does not change over time.

3.3 Spatial Decision Support Systems

SDSSs are expert- or knowledge-based systems designed to incorporate human-like reasoning to develop useful scenarios utilizing computing power [20]. They are often meant to assist decision-makers in solving complex challenges of spatial nature by integrating several aspects [20]: (i) analytical and spatial modeling skills, (ii) spatial and nonspatial data handling, (iii) domain knowledge, (iv) spatial display capabilities, and (v) reporting capabilities.

A variety of SDSSs from a wide range of domains, including hydrology [21]; tourism [22]; agriculture and farming [23]; and risk assessment [24], have been developed in the last several decades.

4 Architectural Design

The developed platform consists mainly of three components, as illustrated in Fig. 1: (i) input module; (ii) core (i.e., simulation engine and database); and (iii) visual analytics and spatial distribution viewer. Details on each component are provided in the following.

The first component of the system is the input module. Through this module, a user can define the simulation parameters, such as key dates, percentage of infected population, and the mathematical epidemic model. This component also offers the possibility of creating a customized response scenario (RS) via the definition of three elements: response triggering condition (RTC), response strategy countermeasures

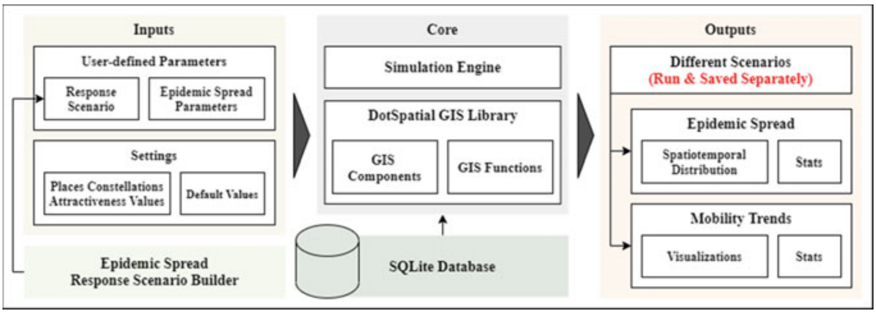


Fig. 1 Framework design

(RSCs), and response easiness condition (REC). More detail is provided on these aspects in Sect. 5.4.

The second component of the system consists of the core part. This component is composed of three essential elements: the engine responsible for running the simulation; the GIS module with capabilities of querying spatial data and executing spatial operations; and the spatial database that contains agents' records.

The last component of the system consists of panels for viewing the visual analytics and statistics related to infection spread and mobility trends as well as a GIS interactive map for viewing the spatial distribution of different agents. It should be noted that this version of the system does not support simultaneous RSs. Results of each RS need to be saved and compared with previously run ones.

5 Implementation

For the implementation of the system, which included the integration of system components and the creation of graphical user interfaces (GUIs), we utilized Microsoft Visual Studio as the integrated development environment (IDE) and C# as the programming language. SQLite was selected as the relational database management system. The open-source GIS library Dotspatial[1] was employed to implement spatial operations (e.g., buffer) and to embed GIS components within the platform and subsequently visualize and interact with geographic data. In the following sections, we provide an overview of the ABM model, simulation framework, database structure, and the main interface and functionalities.

5.1 ABM Model: Description and Setup

In the developed ABM model, the study area is divided into several small zones with unique ID numbers. Agents are individuals and places. Each individual is characterized by their geographical location (i.e., geographical coordinates and the current small zone ID where they are located) and infection status, which are updated hourly. Additionally, an individual is associated with a magnification ratio that reflects how many persons they represent in the real world. This concept is adopted to optimize the performance of the model and to minimize the simulation execution time. The movement of an individual is controlled by a mobility schedule that contains the small zones in which they are located every hour.

Interaction between two individuals is based on the geographical distance separating them. This interaction mainly affects the infection status of one of the individuals. An illustration of such an interaction based on the SEIR model is shown in Fig. 2 At T_0, healthy but "Susceptible" Person A is outside the infection buffer

[1] https://github.com/DotSpatial/DotSpatial.

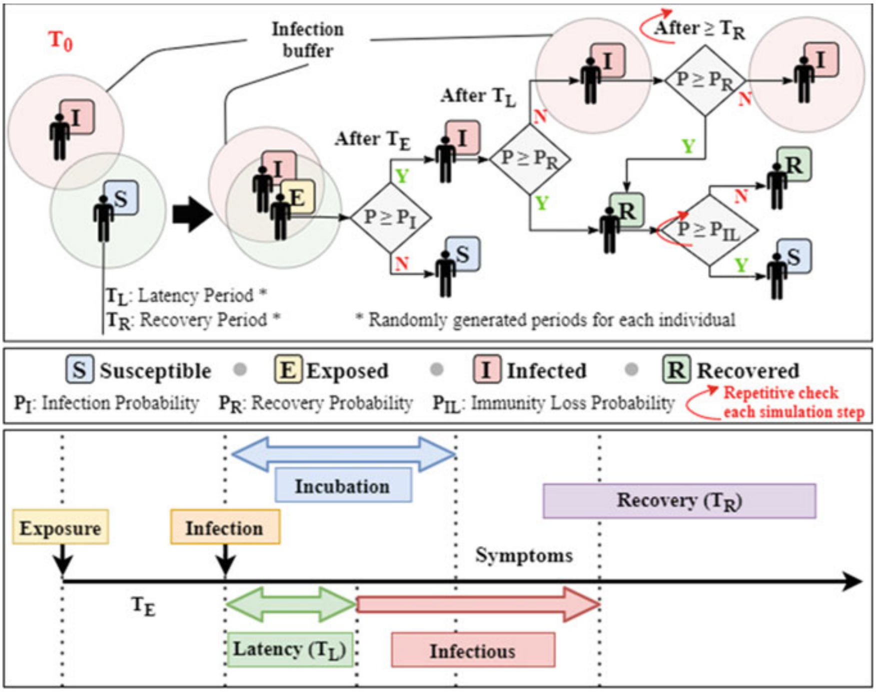


Fig. 2 (Top) Example of SEIR-based infection status alterations as a result of interactions between two individuals (a healthy person and another infected and infectious) (middle) legend of different status and probabilities (bottom) main infection status of an individual

zone—defined based on a predefined infection distance—of infectious Person B. In this case, the infection status of A does not change. However, when the two persons come closer to each other until person A is within the infection buffer zone of person B, A becomes "Exposed" but not necessarily "Infected." The infection status of A will be determined later on by comparing a randomly generated probability (P) with a user-defined infection probability (P_I). If P is greater than or equal to P_I, A becomes "Infected" but not yet contagious. Otherwise, A becomes susceptible again. After a latency period (T_L), it will be decided whether A becomes infectious, meaning capable of spreading the disease to other persons. Once the T_L has passed, a randomly generated probability (P) is compared with a user-defined recovery probability (P_R). Person A recovers if $P \geq P_R$; otherwise, they become contagious. After a recovery period (T_R), in the following simulation steps, a repetitive check is conducted based on a random probability value to check if person A can recover or not. If A recovers, they will be subject to another repetitive check to determine if they can lose the immunity gained after recovery. For that, another probability value is randomly generated to compare against a user-defined immunity loss probability (P_R). In the case where $P \geq P_{IL}$, person A stays "Recovered" or immune; otherwise, they lose their immunity and become "Susceptible" again. The process is repeated as long as the simulation is underway. It should be noted that T_L and T_R are specific to each person, as they are generated randomly when agents are created. The same process is applied under SIR and SEIRD models with minor adjustments. The SEIRD model is an extension of the SEIR model, as another status is added, which is "Deceased." A person can die as a result of their infection if a randomly generated probability is greater than or equal to a user-defined death probability (P_D). The SIR model is more simplistic, as "Exposed" and "Deceased" statuses are not recognized, so a person can become infected directly.

Places are categorized into six groups similar to the Google mobility reports (GMR) published by Google [25]: (i) grocery and pharmacy, (ii) parks, (iii) transit stations, (vi) retail and recreation, (v) residential, and (vi) workplaces. Each place has a maximum capacity calculated using its floor count and footprint area. These places are also characterized by an attractiveness index, which changes depending on the time of the day. Place attractiveness reflects the capacity of a place to attract individuals to enter. Depending on the type of the place, approximate opening and closing times are additional characteristics that users of applications can adjust. It is worth noting that places of type "residential" are always open.

For the sake of performance, places within the same small zone are grouped as a "place constellation." Consequently, places of the same category within the same zone are managed as one place in terms of attractiveness. The capacity of a place constellation is the sum of capacities of all places within it. Individuals can access these place constellations within pre-defined business hours based on their capacities and attractiveness.

5.2 Simulation Framework Overview

Figure 3 illustrates the general framework of the application. We distinguish three main steps to perform the simulation: (1) data loading and spatial environment creation; (2) simulation parameters input and RS defining; and (3) simulation results display, including visual analytics, statistics, and spatial distribution.

Once the application is launched, it creates the spatial environment, which refers to the target area split into small zones. Then, place constellations are created as well. A place constellation is defined as a collection of places of the same GMR category within the same small zone. Next, the population is generated based on individuals' mobility data and magnification ratios. These steps summarize the first step of the framework.

In the second step, the user is required to input the simulation parameters. These adjustable parameters are included under five general categories detailed in the following (the corresponding interface is presented in Sub-section "Simulation Parameters Interface" of Sect. 5.4):

- Key dates: a set of dates and times related to the simulation start and end in addition to those of the first infection.

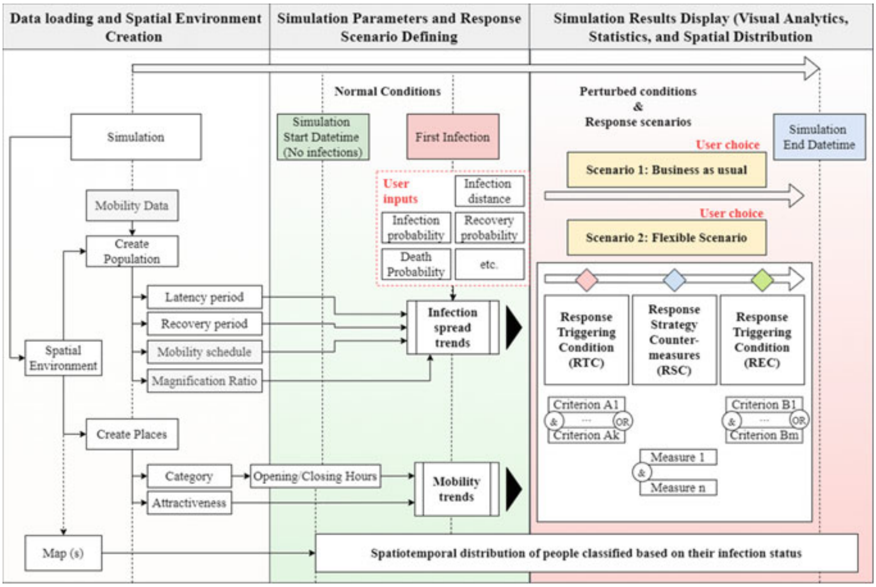


Fig. 3 Flowchart of the main simulation steps

- Population: Parameters under this category include the total population to consider in the simulation. The percentages of the infected and immune populations can be adjusted as well.
- Epidemic model: the mathematical model (SIR, SEIR, or SEIRD) to consider in addition to probabilities of transmission, recovery, immunity, and death. Other variables can be adjusted as well, including the maximum distance of disease transmission and latency and recovery periods.
- RS: the strategy to follow to respond to the infection spread. The user can choose one of two scenarios: (1) business as usual, meaning no action will be taken, or (2) a flexible scenario through which the user can select the response condition, response strategy, and/or easiness conditions. In Sub-section "RS Builder" of Sect. 5.4, we present the corresponding interface and a more detailed explanation of the steps to define an RS.
- Mapping: The user can visualize the geolocation of individuals with their respective infection status and places while the simulation is running. However, this might affect the performance and lengthen the execution time.

Once all simulation parameters are adjusted, the last step of the analysis includes the display of the results in the form of charts, statistics, and the spatial distribution of agents. These results are updated hourly.

Table 2 Individuals' characteristics and mobility schedules

Individual ID	Homestayer	3–4	4–5	…	1–2	2–3	Magnification ratio
Individual #1	True	$Z1_1$	$Z1_2$	…	$Z1_{23}$	$Z1_{24}$	150
…	…	…	…	…	…	…	…
Individual #n	False	Zn_1	Zn_2	…	Zn_{23}	Zn_{24}	25

Note Zi_j refers to the small zone where individual i is located

Table 3 Place constellation attributes

Place Constellation ID	Small Zone ID	WKT Geometry	GMR Category	Total Levels	Total Area
Place Constellation #1	Small Zone #1	MULTIPOLYGON (((…)), …, ((…)))	PARKS	150	5,000
…	…	…	…	…	…
Place Constellation #m	Small Zone #m	MULTIPOLYGON (((…)), …,((…)))	WORKPLACES	25	2,700

5.3 Database

SQLite is used to create the database component of the system. The database is composed of two main tables: individuals (Table 2) and place constellations (Table 3). The former contains records of individuals with their visited small zones throughout the 24 h of the day. It lists whether the individual is a homestayer and an approximate estimation of how many people they represent in the real world (i.e., magnification ratio). The concept of "magnification ratio" has been added to optimize the performance and to minimize the simulation execution time. The latter table lists the place constellations identified by an ID and associated with the following set of information: a foreign key pointing to the ID of the small zone where place constellations are located; a well-known text (WKT) representation of geometry; the GMR category of the place constellation; the levels count, which is the sum of the number of floors of all places forming the constellation; and the total area of the constellation.

5.4 Main Interfaces and Functionalities

EpiDesktop was designed with a simple yet attractive interface that can be used by users regardless of their technical knowledge. In this section, the essential interfaces of the platform and functionalities of the system are presented.

Main Interface

Figure 4 shows a screenshot of the main GUI of the developed system. EpiDesktop

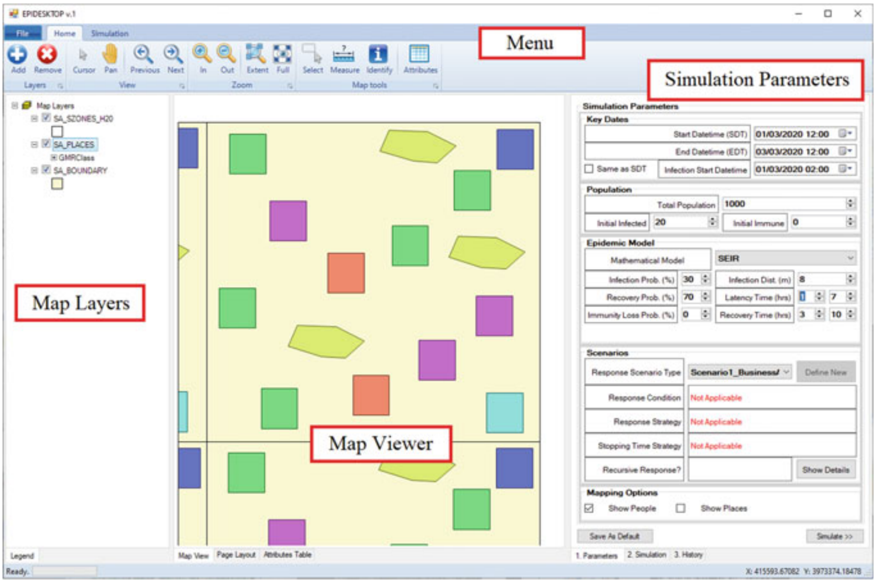


Fig. 4 EpiDesktop main interface

offers a basic interface comparable to standard desktop GIS platforms with the inclusion of a simulation parameters form and a results' viewer panel. A ribbon-style main menu, map viewer, and associated map layer viewer are all included in the main GUI. The major graphical element is the map, while the rest of the interface has tools for finding, acquiring, and managing data.

GIS Functionalities Menu

Handling and interaction with GIS data are assured by the open-source DotSpatial GIS components. The main GIS functionalities are accessible through the first tab of the application ribbon, as illustrated in the close-up in Fig. 5(1). New layers, including vectors and rasters, can be added to the map viewer or removed if selected from the "Map Layers (ML)" list located on the left side of the main interface. Through the "View" and "Zoom" submenus, the user can control the displayed view in the map viewer by zooming to a certain spatial feature or showing the full extent of the target area. Additional tools in the "Map tools" provide the ability to "Select" a given spatial feature (e.g., building) from the map and show its attributes (e.g., perimeter, area) using "Identify". Furthermore, the distance between features can be calculated using the "Measure" option. In addition, for each vector layer in the ML, the associated attributes table can be displayed. Layer-specific options, such as removing, adding labels to features of the layer, and changing the layer color scheme, are also available, as shown in Fig. 5(2).

Simulation Parameter Interface

The right panel of the main interface illustrated in Fig. 4 shows the simulation parameters panel. These parameters are required inputs prior to starting the simulation. They can be categorized into five groups: (i) key dates, (ii) population, (iii) epidemic model, (vi) RS, and (v) mapping options. These parameters and their corresponding descriptions are listed in Table 4.

RS Builder

One of the most crucial components of the developed application is the one used to define customized RSs. Figure 6 shows the interface through which the user can define a flexible RS, which is defined by an RTC, RSC, and an optional REC.

The user defines the RTC—the clause of conditions that trigger the RS. To avoid complexity related to nested conditions, the interface helps build the whole clause in an organized way using a TreeView and two logical operators (AND OR) in addition to SINGLE, in which the user chooses only one condition. At present, the following indicators are considered: The number of Susceptibles, Exposed, Infections, Recoveries, Deaths, Daily Rise of Susceptibles, Daily Rise of Exposed, Daily Rise of Infections, Daily Rise of Recoveries, Daily Rise of Deaths.

Next, the user can specify the response strategy countermeasures, labeled as RSC, to counter the spread of the disease (i.e., RTC is met). For that, the same logic as defining the response triggering condition is employed to specify the remedy steps (only the logical operator AND is used). We consider two variables: modifying the capacity of some types of place constellations (workplace, etc.) or changing the opening and/or closing hours of some types of place constellations. It should be noted that the user is required to control how long the above measures will remain in effect by specifying a stopping time strategy. Three choices are given to the user: the first is a limited-time option where days/hours can be specified. Second, the user can apply the RSC until the RTC is not met. The last option allows the user to specify a response easiness condition (REC), which cancels the RSC once met. Once all RS requirements are set, the user can specify whether this response should be recursive, meaning it repeats every time the RTC is met. The flowchart in Fig. 7 sheds more light on the steps of the algorithm behind managing the RS.

When RS is in "idle" mode (Fig. 7I), the algorithm checks the RTC condition regularly. If met, it forces the application of the RSC and RS status changes to "underway" mode (Fig. 7II). While in this mode, the algorithm checks the status of RTC, RSC, and REC regularly. Depending on the RS stopping time strategy (limited time, RTC is not met, or REC is met), RSC can be stopped or continued to be applied. If the RSC is finally aborted, the RS status changes to "Completed" (Fig. 7III). The algorithm then checks whether the RS is recursive; if true, it keeps checking the status of RTC to repeat the whole process once again when met.

Results Panel

Once the selection of stimulation parameters is completed, the results of the simulation while running can be viewed through multiple charts, statistics, and map viewers,

Table 4 Summary of simulation parameters

Categories	Parameter	Description
Key Dates	Start Datetime (SDT)	Start date and time of the simulation
	End Datetime (EDT)	End date and time of the simulation
	Infection Start Datetime (IDT)	Start date and time of the infection spread
	Same as SDT	Whether SDT and IDT are equal
Population	Total population	The total population to consider for the simulation
	Initial infected (%)	The percentage of initial infected people
	Initial immune (%)	The percentage of initial immune people
Epidemic Model	Mathematical model	The mathematical epidemic model (SIR, SEI, or SEIRD)
	Infection probability (%)	The probability of transmission of the disease
	Recovery probability (%)	The probability of recovery from the disease
	Immunity loss probability (%)	The probability of immunity loss after recovery
	Death probability (%)	The probability of death due to the disease
	Infection distance (m)	Maximum distance from infected people to get exposed
	Latency time (hrs)	Period before becoming infectious
	Recovery time (hrs)	Required time before recovery from infection
Response Scenario	Response scenario type	Response scenario type (business as usual or flexible scenario)
	Response condition	Required condition(s) to start the response strategy measures
	Response strategy	The response strategy measures to encounter the infection spread
	Stopping time strategy	When the response strategy should be ended
	Recursive response	Whether or not the response scenario should be activated whenever the condition is met
Mapping Options	Show people	Show the spatial distribution of people during simulation
	Show places	Show the spatial distribution of places during simulation

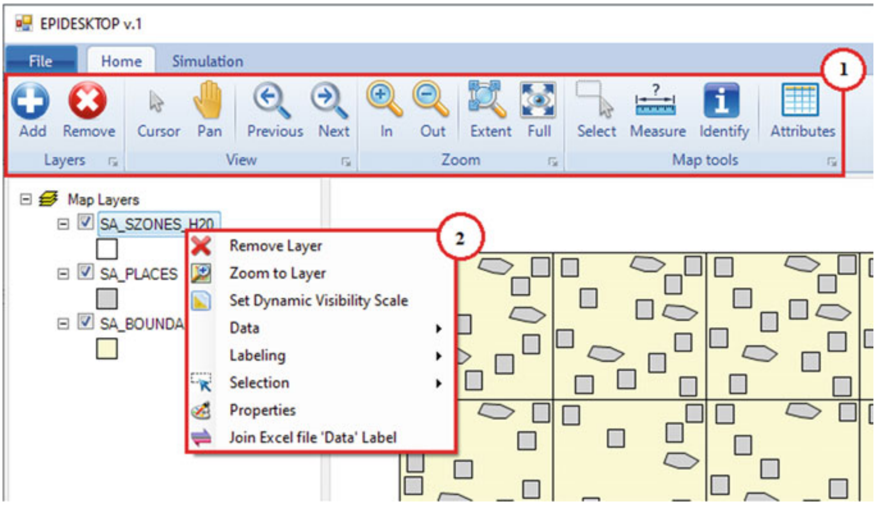


Fig. 5 (1) Main GIS functionalities to control GIS data in the map viewer. (2) The main options to handle added layers

as illustrated in Fig. 8. Additionally, the user will be able to visualize hourly results after the simulation is completed using the simulation timeline that is placed under the ribbon to the left.

The results panel shows the progress of the simulation through a progress bar (Fig. 8(1)). Just below it, the current simulation date and time and the considered population are shown (Fig. 8(2)). The percentage of people who are located indoors and outdoors is shown as well (Fig. 8(4)). The main charts in the results panels show (i) the trends of human mobility to different place categories (Fig. 8(3)), (ii) the trends of epidemic spread (Fig. 8(6)), and (iii) the daily/daily changes in epidemic spread (Fig. 8(7)). Corresponding statistics are shown on the right side of these charts. Regarding the RS, the statuses of RTC, RSC, and REC are presented in Fig. 5. The background color of the boxes indicates whether conditions or strategy countermeasures are activated (rose) or not (green).

6 Conclusions and Future Directions

This paper introduced EpiDesktop, a newly developed SDSS for simulating epidemic spread and human mobility trends under different scenarios. We presented its architectural design, implementation, interfaces, and main functionalities. Comparable to standard GIS desktop applications, the developed system provides users with a wide range of tunable simulation parameters, including customized response strategies for infectious disease spread based on mobility patterns.

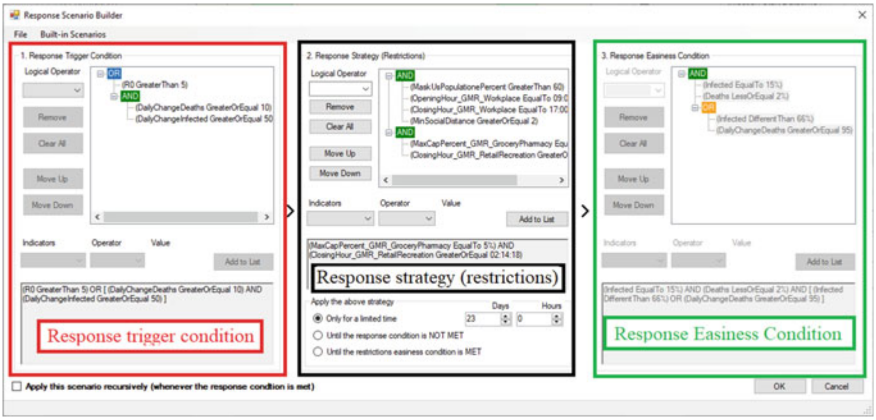


Fig. 6 Response scenario builder interface

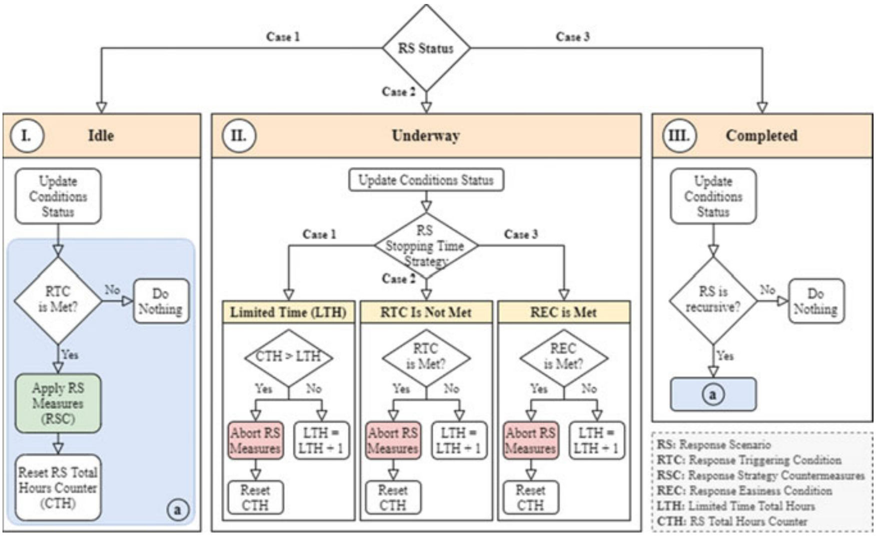


Fig. 7 Response strategy implementation flowchart

The early testing of the application shows that EpiDesktop has the potential to be a versatile tool for monitoring the spread of infectious disease and exploring customized RSs defined by users. Future directions of this project include running a simulation based on real-world settings and comparing mobility patterns with GMR records published by Google. We also plan to include other response strategy measures to counter the spread of the disease, such as social distancing rules and mask-wearing.

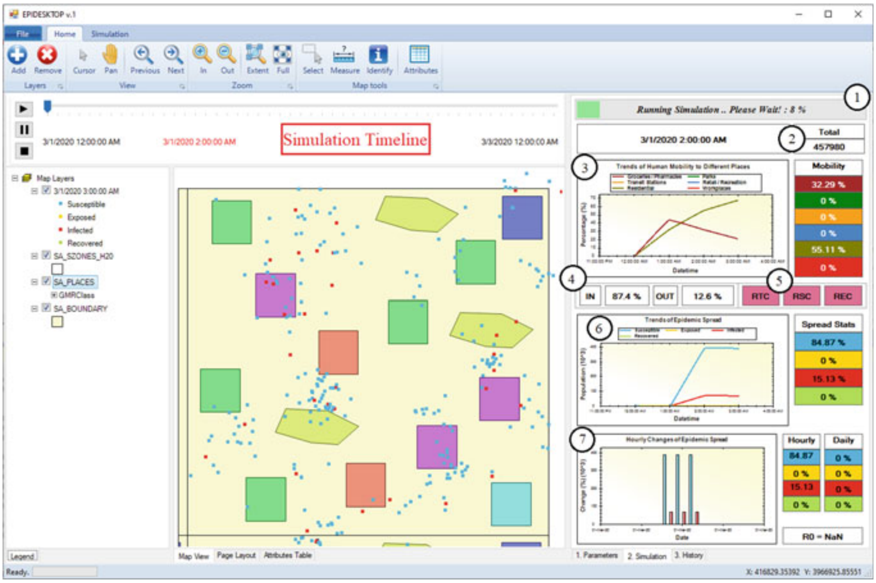


Fig. 8 Results panel

References

1. Dong, E., Du, H., Gardner, L.: An interactive web-based dashboard to track COVID-19 in real time. Lancet. Infect. Dis. **20**, 533–534 (2020). https://doi.org/10.1016/S1473-3099(20)30120-1
2. World Bank: Population, total | Data, https://data.worldbank.org/indicator/SP.POP.TOTL. Accessed 28 April 2021
3. Population Reference Bureau: 2020 World Population Data Sheet (2020)
4. Hadjidemetriou, G.M., Sasidharan, M., Kouyialis, G., Parlikad, A.K.: The impact of government measures and human mobility trend on COVID-19 related deaths in the UK. Transp. Res. Interdisc. Perspect. **6**, 100167 (2020).https://doi.org/10.1016/j.trip.2020.100167
5. Zhou, Y., Xu, R., Hu, D., Yue, Y., Li, Q., Xia, J.: Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. The Lancet Digital Health. **2**, e417–e424 (2020). https://doi.org/10.1016/S2589-7500(20)30165-5
6. Chu, S., Gao, S., Sundaram, D.: A flexible multi-criteria spatial decision support architecture: design and evaluation by healthcare test cases. J. Inf. Technol. Healthc. **3**, 5–20 (2005)
7. Afzal, S., Maciejewski, R., Ebert, D.S.: Visual analytics decision support environment for epidemic modeling and response evaluation. In: VAST 2011—IEEE Conference on Visual Analytics Science and Technology 2011, Proceedings (2011). https://doi.org/10.1109/VAST.2011.6102457
8. Broeck, W.V.D., Gioannini, C., Gonçalves, B., Quaggiotto, M., Colizza, V., Vespignani, A.: The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. BMC Infect. Dis. **11**, (2011). https://doi.org/10.1186/1471-2334-11-37
9. Maciejewski, R., Livengood, P., Rudolph, S., Collins, T.F., Ebert, D.S., Brigantic, R.T., Corley, C.D., Muller, G.A., Sanders, S.W.: A pandemic influenza modeling and visualization tool. J. Vis. Lang. Comput. **22**, 268–278 (2011). https://doi.org/10.1016/j.jvlc.2011.04.002

10. Bryan, C., Wu, X., Mniszewski, S., Ma, K.-L.: Integrating predictive analytics into a spatiotemporal epidemic simulation. In: VAST 2015— IEEE Conference on Visual Analytics Science and Technology, Proceedings (2015). https://doi.org/10.1109/VAST.2015.7347626
11. Ali, M.A., Ahsan, Z., Amin, M., Latif, S., Ayyaz, A., Ayyaz, M.N.: ID-Viewer: a visual analytics architecture for infectious diseases surveillance and response management in Pakistan. Public Health. **134**, 72–85 (2016). https://doi.org/10.1016/j.puhe.2016.01.006
12. Yañez, A., Duggan, J., Hayes, C., Jilani, M., Connolly, M.: PandemCap: Decision support tool for epidemic management. In: VAHC 2017— IEEE Workshop on Visual Analytics in Healthcare, Proceedings (2018). https://doi.org/10.1109/VAHC.2017.8387497
13. Lipeng, J., Xuedong, Z., Jianqin, Z., Zhijie, X., Shaocun, D.: Design and Development of a Visualization System for COVID-19 Simulation Based on WebGIS. In: 2020 International Conference on Public Health and Data Science (ICPHDS), pp. 278–282 (2020). https://doi.org/10.1109/ICPHDS51617.2020.00061
14. Silva, P.C.L., Batista, P.V.C., Lima, H.S., Alves, M.A., Guimarães, F.G., Silva, R.C.P.: COVID-ABS: an agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. Chaos, Solitons & Fractals **139**, 110088 (2020). https://doi.org/10.1016/j.chaos.2020.110088
15. Macal, C.M., North, M.J.: Agent-based modeling and simulation. In: Proceedings of the 2009 Winter Simulation Conference (WSC), pp. 86–98 (2009). https://doi.org/10.1109/WSC.2009.5429318
16. Kashiyama, T., Pang, Y., Sekimoto, Y.: Open PFLOW: creation and evaluation of an open dataset for typical people mass movement in urban areas. Transp. Res. Part C: Emerg. Technol. **85**, 249–267 (2017). https://doi.org/10.1016/j.trc.2017.09.016
17. Sharma, S., Ogunlana, K., Scribner, D., Grynovicki, J.: Modeling human behavior during emergency evacuation using intelligent agents: a multi-agent simulation approach. Inf. Syst. Front. **20**, 741–757 (2018). https://doi.org/10.1007/s10796-017-9791-x
18. Vynnycky, E., White, R.: An Introduction to Infectious Disease Modelling. Oxford University Press, Oxford, New York (2010)
19. Aggarwal, L., Goswami, P., Sachdeva, S.: Multi-criterion intelligent decision support system for COVID-19. Appl. Soft Comput. **101**, 107056 (2021).https://doi.org/10.1016/j.asoc.2020.107056
20. Sugumaran, R., Degroote, J.: Spatial Decision Support Systems: Principles and Practices. CRC Press, Boca Raton (2010)
21. Ames, D.P., Horsburgh, J.S., Cao, Y., Kadlec, J., Whiteaker, T., Valentine, D.: HydroDesktop: Web services based software for hydrologic data discovery, download, visualization, and analysis. Environ. Model. Softw. 37, 146–156 (2012). https://doi.org/10.1016/j.envsoft.2012.03.013
22. Lazoglou, M., Angelides, D.C.: Development of a spatial decision support system for land-use suitability assessment: the case of complex tourism accommodation in Greece. Res. Globalization **2**, 100022 (2020).https://doi.org/10.1016/j.resglo.2020.100022
23. Rupnik, R., Kukar, M., Vračar, P., Košir, D., Pevec, D., Bosnić, Z.: AgroDSS: a decision support system for agriculture and farming. Comput. Electron. Agric. **161**, 260–271 (2019). https://doi.org/10.1016/j.compag.2018.04.001
24. Sermet, Y., Demir, I., Muste, M.: A serious gaming framework for decision support on hydrological hazards. Sci. Total Environ. **728**, 138895 (2020).https://doi.org/10.1016/j.scitotenv.2020.138895
25. Google LLC: Google COVID-19 Community Mobility Reports. https://www.google.com/covid19/mobility/. Accessed 12 June 2021

Urban Environments and Information Systems

Human Behavior Model in Public Pedestrian-Only Space Estimated Using High-Precision Trajectory Data

Toshihiro Osaragi, Arisa Homma, and Hiroyuki Kaneko

Abstract A human behavior model describing the actions of pedestrians in public pedestrian-only space was constructed on the basis of high-precision trajectory data gathered using a laser scanner sensor system. First, a route selection model that could be used to describe macroscopic trajectories was constructed. Next, a walking model that includes the psychological stress imposed by the presence of walls, columns, and other hindrances to motion. These models were then combined to create the human behavior model. Next, using laser scanner sensors, highly precise measurements were taken of the trajectories of pedestrians in the reception area of a hospital. The observational data were employed to estimate unknown parameters in the human behavior model, and to note the characteristics (sex, patient or staff, mobility aid usage) of the pedestrians as they varied with pedestrian attributes. Finally, we proposed a procedure for evaluating the comfort and efficiency of public pedestrian-only space. Using simulation analysis, we demonstrated that it is possible to uniquely estimate the spatial distribution of comfort and efficiency at any given location from the frequency of passages at that location.

Keywords Human behavior model · Laser scanner sensor · Behavior monitoring · Psychological stress

1 Introduction

There is growing interest in evaluating the comfort and efficiency of public pedestrian-only spaces by closely observing user activities and then quantitatively analyzing those results as an aid to the design and planning of such spaces. This is particularly true of locations where widely varying numbers of users are expected to pass through or even linger for some time. Accordingly, a vast number of studies have

T. Osaragi (✉) · A. Homma
School of Environment and Society, Tokyo Institute of Technology, Tokyo, Japan
e-mail: osaragi.t.aa@m.titech.ac.jp

H. Kaneko
Kajima Technical Research Institute, Tokyo, Japan

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_10

attempted to examine how pedestrian behaviors are affected by pedestrian spaces, and the capabilities of numerous different technologies and a wide variety of methods have been used to observe and measure their behaviors in actual pedestrian spaces.

For example, Fudamoto et al. [1] made assessments of pedestrian-only spaces by measuring and otherwise observing a unique set of visual characteristics associated with pedestrians (such as holding hands, smiling, and positions on pathways) using a stress scanner and a video camera, while Sano et al. [2] gathered pedestrian trajectory data using a video camera and added time information to display their findings in three-dimensional (3D) graphic form, after which they proposed a procedure that can be used to obtain an intuitive understanding of the degree of crowding and the flows of groups in pedestrian spaces.

The social force model outlined in Helbing et al. [3] is an advanced concept that examined how pedestrians mutually influence each other. More specifically, that study provides a simple model for the complicated flows within a group based on an equation of motion that describes the interactions between pedestrians. In another study, Osaragi [4] constructed a (walking) model describing pedestrian motions by estimating the psychological resistance (psychological stress) exerted by pedestrians other than oneself (other people). He derived and validated his model using trajectory data estimated from the video camera imagery.

However, a comprehensive examination of the technologies used for observing pedestrian trajectories is expected to reveal more advanced research methods than the conventional approaches of polling and visual surveys. Among such newer methods are wireless detection (Kawano et al. [5], Matsushita et al. [6]) such as image analysis (Konno et al. [7], Tahara and Osaragi [8]), radio frequency identifiers (RFID), "ultra-wideband" (UWB) radio, and ultrasound. In addition, movement monitoring with laser scanner sensors (laser measurements) has received significant attention as a way to obtain highly precise data on pedestrian trajectories while avoiding privacy infringements. Results obtained using this method have been reported for several studies on train station concourse areas and office spaces [9–13].

In this study, a model describing macroscopic pedestrian trajectories (route selection model) was first constructed with reference to the findings of previous studies. Next, a walking model was constructed that incorporated obstacle stress (OS), which is the sense of oppression felt by persons in the vicinity of columns, walls, or other objects, into a pre-existing walking model These two models were then combined to establish our human behavior model. Next, a behavior monitoring survey in which pedestrian trajectory data were gathered was carried out using laser scanner sensors in the reception area of a hospital [Total floor area: 30,273 m^2, Number of beds: 535 beds, Number of outpatients per day: 1,182 persons (2015)] in the Tokyo region. Those data were then analyzed in order to assign values to previously unknown parameters in the human behavior model. Finally, in an attempt to validate the practicality of this model, we performed a comparative analysis using behavior monitoring data.

2 Human Behavior Model

2.1 Model Overview

A walking model was expanded using psychological stress concepts to describe the behaviors of pedestrians intermixing at varying speeds and while traveling in different directions. The chief points of expansion were, first, that macroscopic descriptions of pedestrian motions were allowed and, second, that the effects of objects present, as well as of other people, were incorporated in the form of variables representing each kind of stress. Figure 1a shows details of the improved model.

When considering the behaviours of pedestrians as they make their way toward a destination, these may be characterized by two processes: (1) they decide on the approximate paths they should take in order to reach their destinations efficiently, and (2) as they proceed along their selected paths, they make constant "course corrections" to adjust to the dynamically changing circumstances, the distances to nearby people, and the movements of those other people. Therefore, for the present model, pedestrian behaviors were classified into two types. Corresponding to case 1 above, macroscopic walking behavior is described by the route selection model (see Sect. 2.2 for details).

This model selects the estimated path that yields the minimal cost of the motion to the destination (the weighted motion distance). For the case 2 walking model described above (see Sect. 2.3 for details), pedestrian walking behavior is described

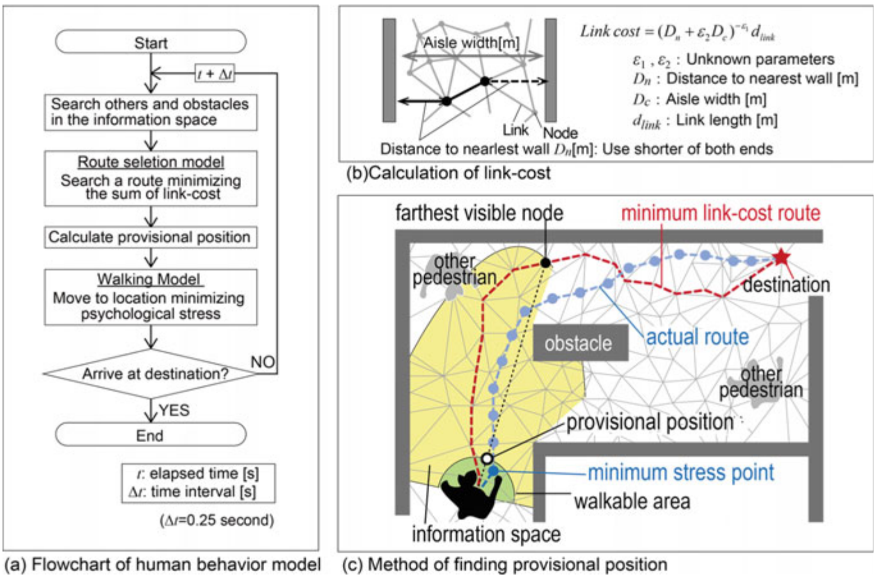


Fig. 1 Outline of human behavior model

in a microscopic context in terms of pedestrian stress (PS), which is caused by the presence of other people, and the OS caused by the presence of physical hindrances (pillars, walls, chairs, etc.) to pedestrian freedom of movement. These stresses arise from destination stress (DS), which is defined as the necessity of avoiding other people and any objects that would divert a person from his/her intended path.

2.2 Route Selection Model

A macroscopic route selection model that avoids pillars, walls, and other obstacles is created to the destination.

(1) A set of points is generated at random within the walkable region, and a triangulated irregular network (TIN) consisting of these nodes is constructed.
(2) In order to express a pedestrian's avoidance of obstacles while proceeding toward the destination, a link cost is determined for every link making up the TIN. These costs vary with corridor width and the distances between pillars and walls (Fig. 1b).
(3) The path with the minimal summed link costs from the starting point to the destination is found using Dijkstra's Algorithm, which is a shortest-path algorithm (Fig. 1c). The actual path, more specifically, the macroscopic path, is defined by the above method as the pedestrian travels around either the right or left side of each obstacle.

2.3 Walking Model

It is assumed that a pedestrian will proceed by a path that will be as comfortable (low-stress) as possible. In other words, a pedestrian is assumed to walk a path that poses the lowest degree of stress due to the presence of other people, objects, and any other hindrances to his/her freedom of movement. In this study, the walking model based on psychological stress was developed to account for three factors: (1) PS due to the presence of other pedestrians, (2) OS caused by hindrances to the freedom of movement such as walls and pillars, and (3) DS due to diversions from the direct path to the intended destination. This model assumes that each pedestrian moves toward a final destination through locations chosen to minimize the combined total stress due to these factors (Fig. 2).

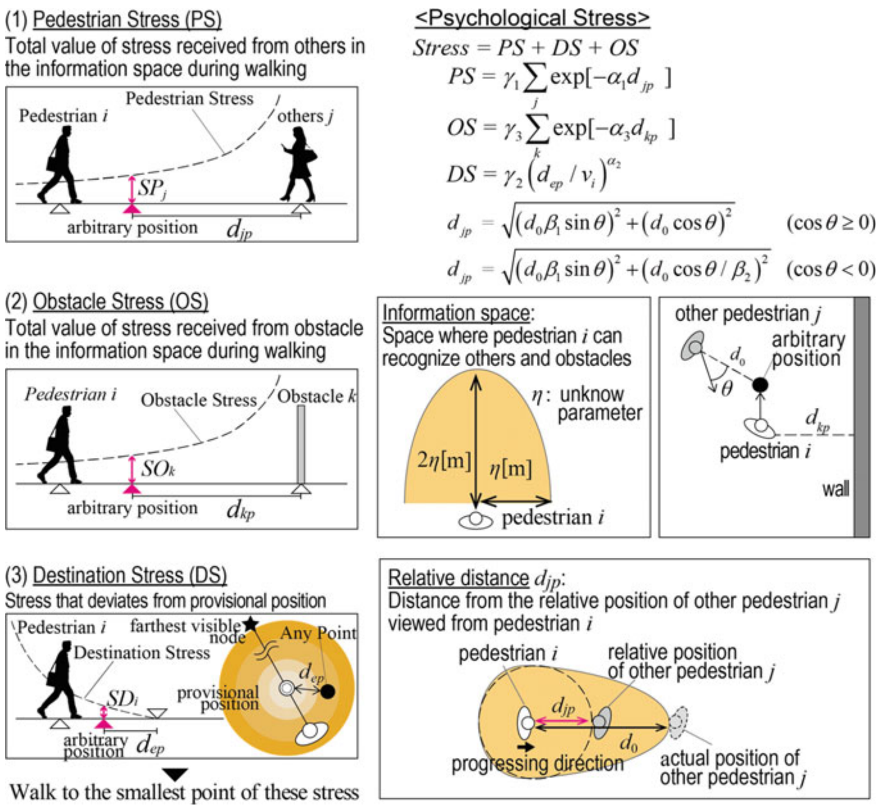


Fig. 2 Definition of psychological stress

3 Creation and Validation of Human Behavior Model

3.1 Laser Scanner Sensor-Based and Visual Behavior Monitoring Survey

In the first survey (Survey 1 in Table 1) of this study, eight laser scanner sensors were employed to take measurements in the reception area of the hospital in the Tokyo area (Fig. 3). Pedestrian trajectories were observed with high precision using these sensors [13], which determine the target's distance and motion direction from laser beam reflection return times. A visual survey (Survey 2 in Table 1) was conducted simultaneously to record the clock times and the attributes (sex, patient or staff, use of mobility aids) of all the pedestrians entering the survey area (Fig. 4). The time data and location data were employed as keys when creating the measured trajectory record [A], which was used for analyses relating pedestrian attributes to their trajectories.

Table 1 Outline of surveys

	Survey 1: Behavior monitoring survey by laser scanner sensor	Survey 2: Survey of pedestrian attributes by visual observation
Date	Wednesday, October 21, 2015	
Time	6:50–17:00	7:30–9:30, 10:30–12:20, 14:30–15:30
Survey method	The standing/walking positions within the survey area were measured by eight laser scanner sensors	The entry-time and attributes of people entering the survey area were visually surveyed at entrances
Observed data	Behavior monitoring data of walking trajectory	Attribute data of people entering the survey area
The number of observed persons	07:30–07:45 160 persons 10:30–10:45 279 persons 12:00–12:15 336 persons Total 775 persons	07:30–07:45 141 persons 10:30–10:45 219 persons 12:00–12:15 272 persons Total 632 persons

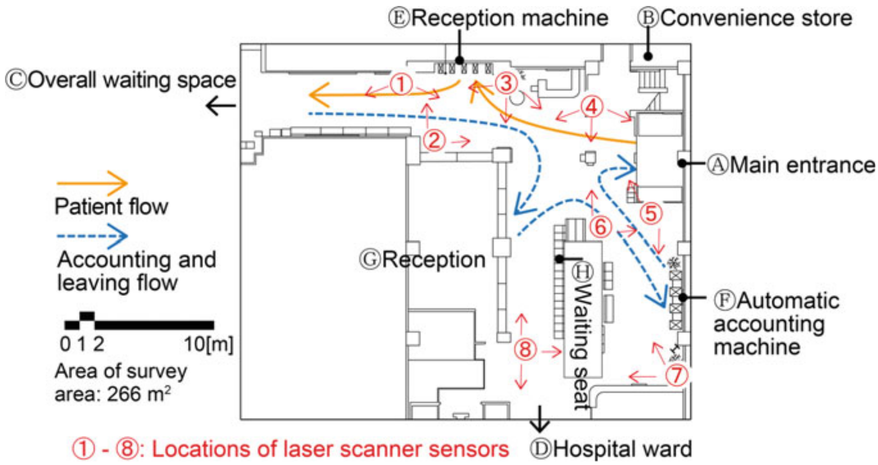
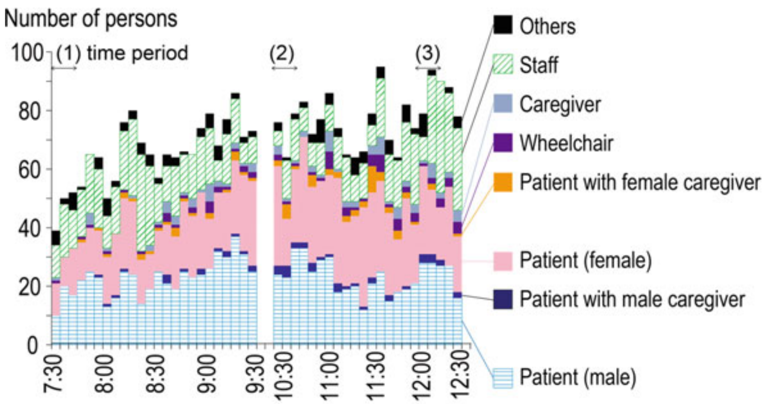


Fig. 3 Plan of reception area of hospital and locations of laser scanner sensors

3.2 Method for Estimating Parameters

Route Selection Model. Parameters ε_1 and ε_2 of the route selection model were estimated using the gradient method while minimizing the trajectory error of observed Trajectory [A] from the ideal trajectory connecting the starting point to the intended target point (interim destination) (Fig. 5). The data used for the estimation process were all of the observed trajectories of the people within the pedestrian-accessible area.



(1) 7:30 - 7:45	Before starting reception, there are many visitors
(2) 10:30 - 10:45	Visitors and returnees are mixed and crowded
(3) 12:00 - 12:15	Reception area is crowded and many staff move on lunch break

Fig. 4 Changes in the number of visitors to and from the survey area

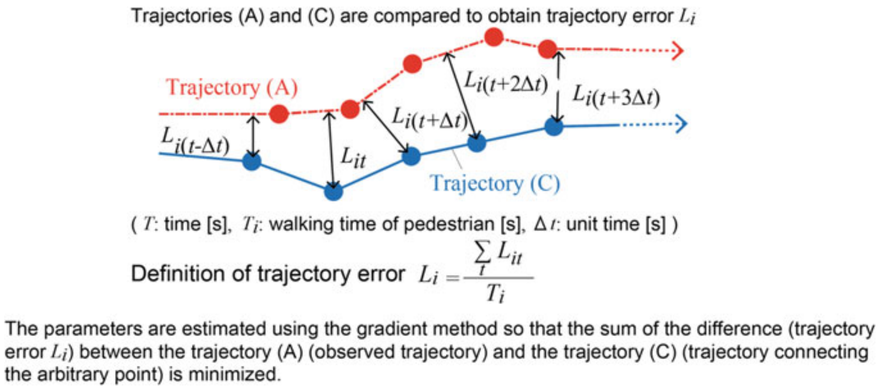


Fig. 5 Estimation method of parameters of route selection model

Walking Model. The linear distance D_i^t (trajectory error) perpendicular from each point-to-point chord of the ideal trajectory P_i^t with minimized stress to the actual trajectory Q_i^t for pedestrian i at time point t was calculated. The parameters (α_1, α_2, α_3, β_1, β_2, γ_1, γ_2, and γ_3) of the walking model were then estimated (Fig. 6) using the gradient method so as to minimize the mean error D found for all pedestrians of any given set of attributes at any given time point. In the process of creating this model, it was felt that using the observed trajectories of pedestrians who never actually encountered other people while travelling would generate parameter instabilities. Therefore, the data were sampled to select individuals who actually had passed close to other people or objects in order to generate the above parameters (Fig. 7a).

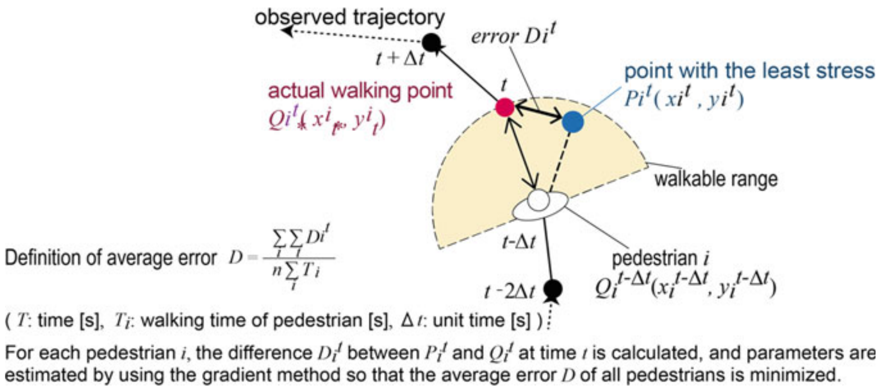


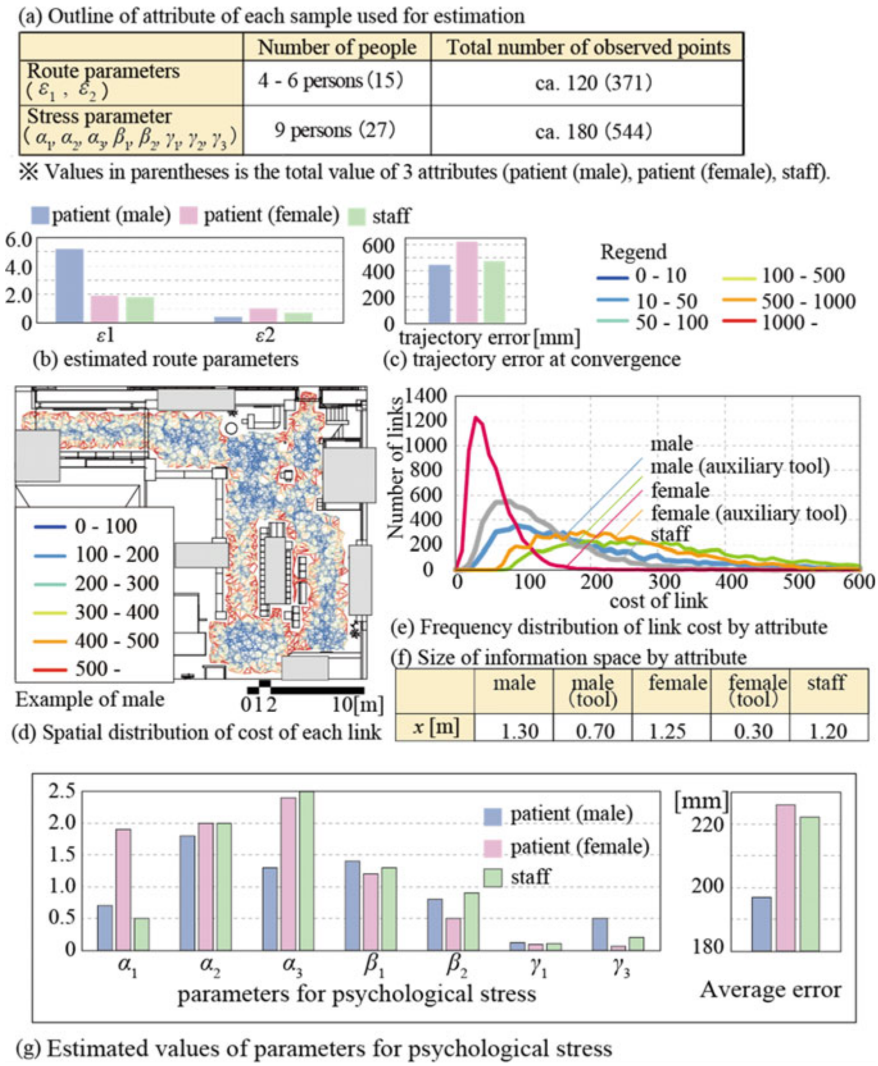
Fig. 6 Estimation method of parameters of walking model

3.3 Estimated Parameters

Route Selection Model. The trajectory errors converged to approximately 400–600 mm for all sets of attributes (Fig. 7c). An examination of the link costs in the TIN calculated using the estimated parameters (Fig. 7b) reveals that men had a higher decay parameter ε_1 than women or staff members, thereby indicating that there was little influence from OS if the obstacles were a certain distance away. In other words, men tended to show increasing link cost with distance from the centerline of an aisle, approaching a wall (Fig. 7d). In contrast, women and staff showed a lower tendency for this behavior, and the data indicated that their link costs were high when they passed through narrow aisles, so they tended to select wide aisles. Individuals of all attributes showed quite high link costs when they passed through narrow aisles, which was manifested as a tendency to avoid such locations.

Walking Model. Figure 8 presents the estimated values for the parameters for each set of pedestrian models. Parameter α_1, which represents PS decay, was highest in women, thus indicating a relatively low sensitivity to the visible presence of other people if they are some distance away. In contrast, men showed low values for parameter α_3, which represents OS decay, and higher values for γ_3, which represents OS magnitude. Men also manifested a tendency to pass through locations farther from objects (such as wide aisles). On the other hand, the data also shows many pedestrians moving in groups of several individuals. It is reasonable to infer low levels of PS when the individual knows the other people around him/her, but this factor was not considered in this model. Accordingly, we believe that models created in the future must account for the occurrence of pedestrians moving in groups of acquaintances.

The values of unknown parameters estimated using the high-accuracy but limited number of observation data should be further more investigated, since they are somewhat not stable if the data used for the estimation are exchanged.



	Number of people	Total number of observed points
Route parameters (ε_1, ε_2)	4 - 6 persons (15)	ca. 120 (371)
Stress parameter ($\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3$)	9 persons (27)	ca. 180 (544)

	male	male (tool)	female	female (tool)	staff
x [m]	1.30	0.70	1.25	0.30	1.20

Fig. 7 Estimated unknown parameters of route selection model

3.4 Validation of Human Behavior Model

When all the pedestrians are validated at the same time, minute differences in the behaviors of a few individuals will influence others, giving rise to widening "ripples" of behavioral changes that render validation of the model impossible. Therefore, it was initially decided to validate whether the model would correctly predict the trajectories of a limited number of individuals. To validate the accuracy (properly speaking, the trajectory error) of the estimated trajectory for pedestrian *i*, the observed

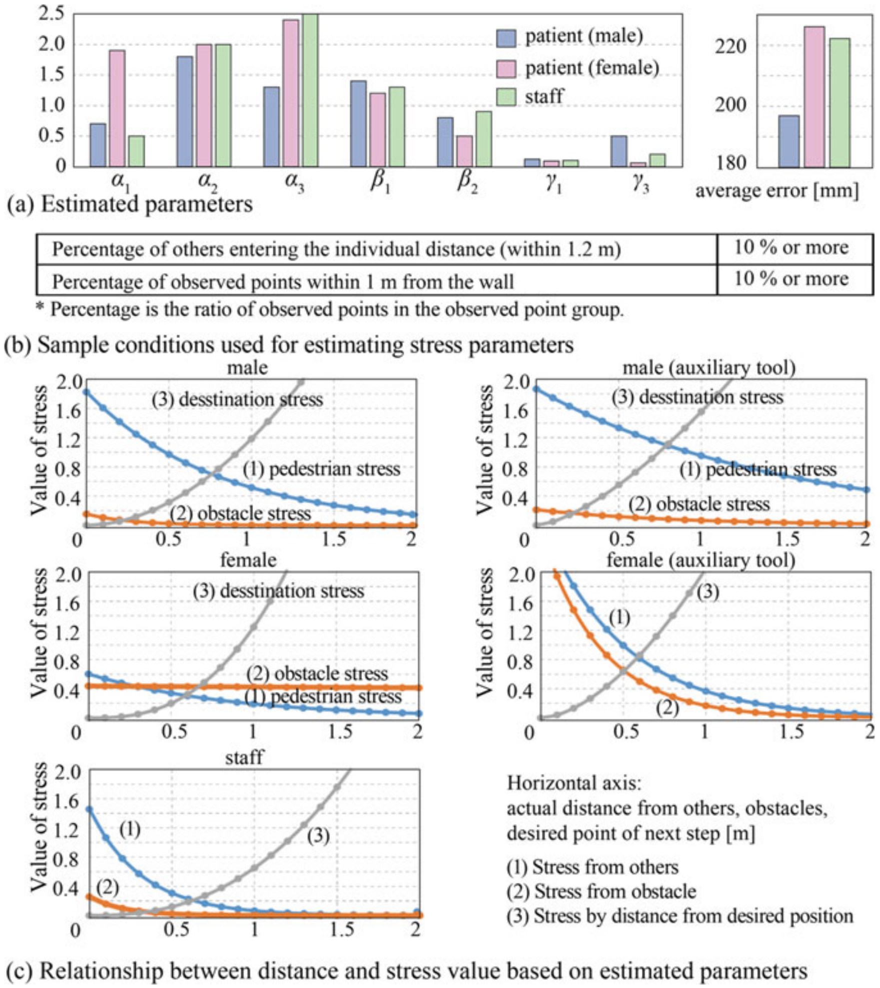


Fig. 8 Estimated parameters of walking model

trajectories of pedestrians present in the same time frame other than pedestrian *i* were employed “as-is”.

The reader can see that the model provided quite accurate predictions. Specifically, the trajectories of over 80% of the pedestrians of all attributes were predicted correctly within 1 m (Fig. 9c). We can see that the trajectory error of females is within a range of slightly smaller error than that of males and staff. However, it cannot be said that the difference is statistically significant. In Fig. 9b, which compares the estimated walking distances with those actually observed for those individuals, the model again provided good accuracy from the viewpoint of walking distance. Nonetheless, there remained a group of individuals whose actual trajectories diverged widely from

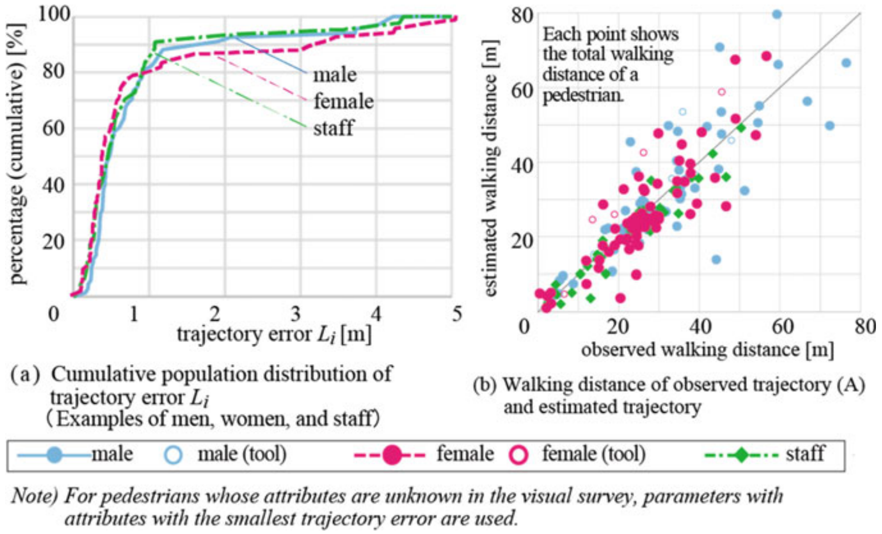


Fig. 9 Validation of human behavior model

the predictions. Close examinations of their routes showed that they had chosen to stay in the reception area or had diverged from the routes to their destinations at branching locations such as pillars and walls. These behaviors indicate the need for improvements to the model. Specifically, it will be necessary to incorporate the static behavior of the former group, who were waiting for medical appointments, while another solution will have to be discovered for the behavior of the latter group because it cannot be expressed by minimizing the motion cost.

4 Space Assessment According to Observed Data

4.1 Method for Space Assessment

The system described here makes it possible to make quantitative tests of the walking efficiency and stress experienced by a pedestrian as he/she walks from the start point toward his/her destination. By extension, the magnitudes of the PS and OS also provide an indication of the comfort of the pedestrian space, which varies with the time and the number of people occupying the space. The magnitude of DS and the distance of deviation from the path to the intended interim destination allow the efficiency of movement to the destination to be evaluated.

The above assessments can be made in the following procedure. The value for the stress experienced on the locus of a given pedestrian is calculated using the

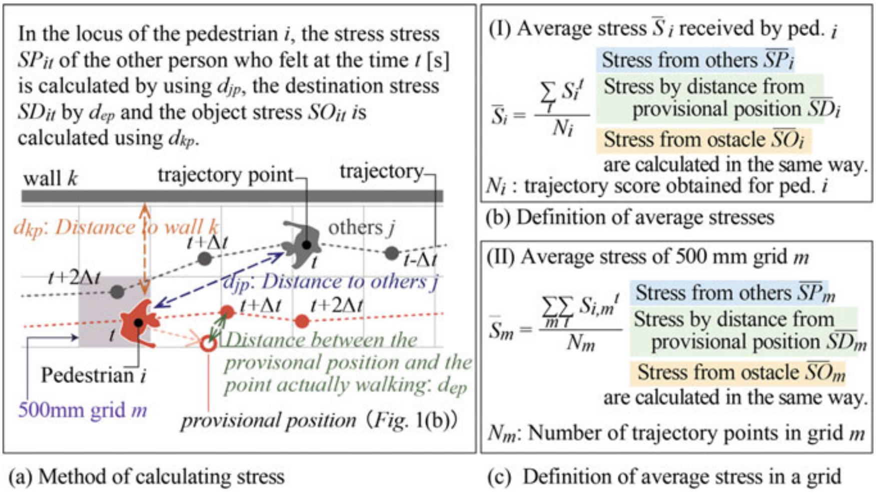


Fig. 10 Spatial evaluation method based on observed trajectory

trajectory obtained from the observed trajectory and from the estimated parameters (Fig. 10a). The level of comfort for the original pedestrian space was assessed using the stress values during motion. First, the mean S_i was found for each person (Fig. 10b), and then, the mean S_m for each grid element used by pedestrians was found (Fig. 10c). In the next section, in addition to the previous indices found directly from the observed data (such as the frequency of use by pedestrians, walking speed, and walking distance), the comfort of the observed area described in Sect. 3 is reassessed using the indices shown in Fig. 10b,c.

4.2 *Example of Space Assessment*

This subsection presents a comparative analysis of the PS levels at different timeframes, when the number of pedestrians and their attributes varied, as shown in Periods (1) to (3) of Fig. 4. A comparison of the mean PS_i values for different pedestrians revealed that most people show low PS_i in time Period (1), 07:30–07:45, thereby indicating a high comfort level of the observed area at this time of the morning (Fig. 11a). The distributions of PS_i for the pedestrians in Periods (2) and (3), 10:30–10:45 and 12:00–12:15, respectively, resembled each other closely. However, upon an examination of the spatial distributions at these times, it could be seen that they showed different locations (Fig. 11b). More specifically, the high PS_m observed at locations (i) and (ii) during Period (3) differed from the high PS_m location (iii) observed during Period (2). Focusing on Period (3), it is notable that PS_m is high at location (ii) (Fig. 11d), even though a low number of people passed through (Fig. 11c). This was thought to be due to the high number of people occupying this

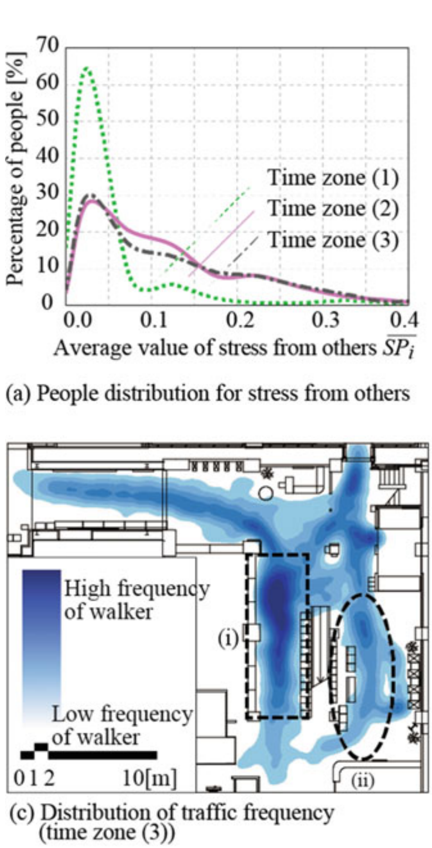


Fig. 11 Spatial evaluation results based on observed data

location (people standing and waiting or standing in line). Thus, we can see that this method is able to identify spatial comfort level distributions that are not detectable in the spatial distributions of pass-through rates that are based on conventional kernel density functions.

5 Summary and Conclusions

In this study, a behavior model describing the actions of pedestrians was constructed on the basis of detailed pedestrian trajectory data gathered using a laser scanner sensor system. Analysis of the collected data showed the model was effective for evaluating existing spaces and as a tool for floor plan design.

The study began with the construction of a route selection model that could be used to describe macroscopic trajectories that minimize the motion cost of travel to selected destinations. This model was then expanded to include the well-established concept of psychological stress that is imposed by the presence of walls, columns, and other hindrances to motion. These results were then combined to create the human behavior model.

Next, using laser scanner sensors, highly precise measurements were taken of the trajectories of pedestrians in a behavior monitoring survey in the reception area of hospital. The observational data were employed to estimate unknown parameters in the human behavior model, and to note the characteristics (sex, patient or staff, mobility aid usage) of the pedestrians as they varied with pedestrian attributes. The observational data demonstrated that the human behavior model showed a good level of accuracy.

Subsequently, a procedure for evaluating the comfort and efficiency of a pedestrian space from the viewpoint of psychological stress was proposed, and values for the stress experienced by pedestrians, which are generally difficult to estimate, were calculated using the human behavior model. The distribution of the stress levels at differing time frames and locations were compared, and the results obtained indicate that it is possible to uniquely estimate the spatial distribution of comfort and efficiency at any given location from the frequency of passages at that location.

This study included observational surveys in a hospital and successfully described the walking characteristics of adult pedestrians, including those who use mobility aids. However, it is unknown whether this model could be applied in institutions serving large numbers children of elementary school age or younger, who tend to walk in a much more cheerful and energetic manner than typical hospital clients. Accordingly, the construction of walking behavior models that are suitable for a wide variety of institutions will be a matter for future research. Furthermore, since this model is also unable to describe the behavior of groups of acquaintances, that aspect will be addressed in the next stage of our research, along with behavior descriptions of people who linger within a particular area.

Acknowledgements A portion of this paper was presented at the EDRA 2021. The authors would like to give their special thanks to Ms Miho Fukui (Kajima Technical Research Institute) who cooperated in our surveys at the hospital.

References

1. Fudamoto, T., Kojima, A., Kubota, H.: Study on evaluation method for street space focusing on pedestrian behavior. J. Jpn. Soc. Civ. Eng., Ser. D3 (Infrastruct. Plan. Manag.) **67**(5), I_919–67_I_927 (2011)
2. Sano, T., Takayanagi, H., Watanabe, H.: The evaluation of cognition in pedestrian space with space-time diagram model. J. Arch. Plan. Environ. Eng. (Trans. AIJ) **67**(555), 191–197 (2002)
3. Helbing, D., Monar, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**(5), 4282–4286 (1995)
4. Osaragi, T.: Modeling of pedestrian behavior and its applications to spatial evaluation. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'04), vol. 2, pp. 834–841. IEEE (2004)

5. Kawano, N., Yukawa, K., Munemoto, J. Matsushita, D. et al.: Method of the Investigation on Behavior at Working places and Making Database: Investigation on Working Behavior observed using UWB Sensor Network Part1, In: Summaries of Technical Papers of Annual Meeting, **E-1**, pp. 785–786. Architectural Institute of Japan (2009).
6. Matsushita, D., Odawara, R., Isogawa, E., Yamashita, and T. Kumagawa, T.: Development of Nursing Activities Measurement Method with Ultrasound Positioning:−Investigative report in acute care ward of Showa University Hospital−, AIJ Journal of Technology and Design, **19**(43), pp. 1079–1084 (2013).
7. Konno, H., Saji, M., Matsumoto, Y., Naka, R., Yamaguchi, S.: A study of supportive tools for behavioral observation in office environments, In: Johoshisutemu-riyo-gijutsushinpojiumu-ronbunshu, pp. 73–78. Architectural Institute of Japan (2007)
8. Tahara, S., Osaragi, T.: Trajectory detection and behavior analysis of visitors in exhibition space. In: Papers and Proceedings of the Geographic Information Systems Association (CD-ROM), **19**. GIS Association of Japan (2010)
9. Nakamura, K., Zhao, H., Shibasaki, R., Sakamoto, K., Ooga, T., Suzuki, N.: Tracking Pedestrians Using Multiple Single-Row Laser Range Scanners and Its Reliability Evaluation, Trans. Inst. Electron., Inf. Commun. Eng. **J88-D-II**(7), 1143–1152 (2005)
10. Zhao, H., Shibasaki, R.: A novel system for tracking pedestrians using multiple single-row laser-range scanners. IEEE Trans. Syst., Man, Cybern. Part A **35**(2) (2005)
11. Sakamoto, K, Shibasaki, R., Zhao, H., Nakamura, K.: A passengers flow analysis with laser scanner sensors in railway stations. In: Summaries of Technical Papers of Annual Meeting, **E-1**, pp. 895–896. Architectural Institute of Japan (2006)
12. Le, J.H., Km, Y.S., Kwata, H., Ohya, H., Yuta, S.: A People Tracking and Counting Method with Multi SOKUIKI Sensors, Gakujutsukoenkai, The Robotics Society of Japan, 24th (2006)
13. Kaneko, H., Osaragi, T.: Extraction of The Spatio-temporal Activity Patterns Using Laser-scanner Trajectory Data **1**, 9 (2020). https://doi.org/10.5194/agile-giss-1-9-2020.AGILEGIScienceSer

Approaches to Assessing the Vulnerability of Large City Population to Natural and Man-Made Hazards Using Mobile Operators Data (Case Study of Moscow, Russia)

S. Badina, R. Babkin, and A. Mikhaylov

Abstract The complex nature of threats to large cities residents requires a rethinking of the existing methods for the vulnerability assessment of the population to various kinds of them. Moscow concentrates about 9% of the Russian population, and with the Moscow oblast—about 12%. The high level of spatial concentration of the population and its active dynamics determine the increased level of natural and man-made risk in the city. The purpose of this study is to assess the vulnerability of the Moscow population to natural and man-made hazards, taking into account the actual population size and its movement in the city within different time cycles (daily and weekly-seasonal). Trying to find an appropriate solution, authors used alternative sources of statistical information—mobile operators data. The use of these data made it possible to obtain more detailed information on the state of socio-geographical systems, to overcome barriers associated with the incompleteness and 'static' nature of official statistical information (data provided by Rosstat). Mobile operators data allow obtaining more reliable depictions of the localization of its users at a certain point in time, which made it possible to adjust and clarify the current ideas about the distribution of the population in Moscow. As a result of the study, it was shown that for the center of Moscow and for New Moscow, the population vulnerability level is much higher than reflected in official documents. On the contrary, in the peripheral areas of Old Moscow, the potential risks are reduced because the real population density is significantly lower than it is estimated in the calculations provided by Rosstat.

Keywords Population vulnerability · Natural and man-made risks · Moscow · Urban real-time population density dynamics · Data of mobile operators

S. Badina (✉) · R. Babkin · A. Mikhaylov
Plekhanov Russian University of Economics, Stremyanny lane, 36, Moscow 117997, Russia
e-mail: bad412@yandex.ru

S. Badina
Faculty of Geography, Lomonosov Moscow State University, GSP-1, Leninskiye Gory, 1, Moscow 119991, Russia

A. Mikhaylov
Institute of Geography RAS, Staromonetny lane, 29, Moscow 119017, Russia

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_11

1 Introduction

The increase in the frequency of natural and man-made emergencies, as well as the scale of their consequences (the number of victims, damage, long-term negative effects affecting socio-economic development) require new scientific approaches and methods to assess the risks and vulnerability of citizens. Large cities and urban agglomerations are areas of increased danger due to the high concentration of the population. The use of so-called 'classical' methods of statistical data analysis does not allow obtaining a clear and representative picture of the population spatial distribution across the city's territory during different times of a day; nevertheless, it is critically important for vulnerability assessments and development of security measures corresponding to the real population numbers.

The studies of natural and man-made risks of socio-economic development have become widespread in the last decades due to the high degree of their relevance, and especially because of the increase in the occurrence frequency of natural and man-made hazards on the one hand and the magnitude of human casualties and material damage on the other, which is typical for all countries and regions in the world [28, 49]. According to the modern interdisciplinary approach, the notion of a 'risk' is defined as a probability of damage and its potential scale due to the emergence of a dangerous event. UN-Habitat defines risk as 'the effect of uncertainty on objectives'.[1] The key problem, constituted by the phenomenon of risk, is finding a way to avoid decisive mistakes which can entail this damage [42]. Natural risk studies are carried out at the intersection of physical and socio-economic geography, since risk manifests itself only when there are potential direct and indirect damages to society and the economy. Generally, the majority of modern researches in the field of studying the socio-economic aspects of the natural risk geographies come down to:

- Vulnerability assessments of the population and the local economy to natural hazards, damage level predictions [11, 34, 40, 46];
- Assessments of the natural disasters' impact on the particular territory, its economic structure and growth in general [20, 25, 48];
- Identification of economic mechanisms and assessments the effectiveness of investments in damage reduction from natural hazards and studying insurance as a tool for reduction of negative impacts in various territories [27, 30, 43];
- Analysis of public perceptions of natural disasters and risk reduction in different countries and regions [15, 26, 31];
- Analysis of regional risk management features, aspects of risk reduction in strategic planning [18, 50] and adaptation to risk [17, 38];
- It is also necessary to highlight the researches related to the study of man-made and natural-man-made risks [2, 22].

Some special attention is paid to the analysis and forecasting of natural risk for the world's largest cities and urban agglomerations. Research works on this topic focus

[1] https://unhabitat.org/un-habitat-enterprise-risk-management-implementation-guidelines (accessed 15 July 2021).

on issues of urban growth as a factor of risk increase [21, 24] and vulnerability of urban residents to natural hazards [36, 37, 44]. In the city-level studies, social risks and risks to the population, are more prioritized oftentimes than economic aspects, whereas in research works conducted at state and regional levels the situation is directly opposite.

Methods for natural risk assessment have not been unified yet. They depend on specific research tasks, the place-specific features of the study area and the set of particular threats taken into account. This study proposes an approach to vulnerability assessment for the population of a large city, based on determining the most probable number of people at a certain time interval in a certain place on a relatively large scale (0.25 sq. km cells). The data of mobile operators is used as the main source of information, making it possible to trace the population movement and obtain the spatial distribution of the population in the city which corresponds to the reality better than in case of use of the 'classic' data sources. Over the past 5–10 years, significant numbers of scientific papers using data of mobile operators have been written about pulsation processes in the settlement system of the Moscow metropolitan area [16, 32, 33, 41]. It is also important to note that the similar data sources were used in studies of structural and functional shifts in settlement system due to the influence of various infrastructural projects [33].

The data of mobile operators has been used in the analysis of settlement systems since the mid-2000s. By now, considerable research experience has been accumulated in such countries as the USA, France, Great Britain, Belgium, Estonia and many others [1, 19, 23]. There are some attempts being made in order to integrate the data of mobile operators into official statistics. A study commissioned by Eurostat was carried out in 2012, aiming to assess the possibility of using mobile data to obtain a variety of statistical information [4]. Comparison of mobile operators' data with census statistics and population register showed their high correlation, while it was noted that mobile data is much more relevant in the analysis of dynamic socio-economic processes. While speaking about the largest international projects using this kind of data in studying the consequences of natural disasters and emergencies, it is necessary to outline the analysis of the earthquake consequences in the Republic of Haiti in 2011, as well as the paper of an international team of scientists dedicated the spread of malaria in Kenya [14, 47]. It should be noted that nowadays the studies carried out using the data of mobile operators are pioneering for Russia, and yet are at the early stages of inclusion into the contemporary Russian scholar discourse.

2 Methods and Data

The methodological approach adopted in this study is based on the 'risk' concept, which is a likelihood function of an emergency and the magnitude of the potential consequences for the population, urban economy and infrastructure (material damage and the number of people affected). In other words, it is a combination of danger and vulnerability. Therefore, to a decisive extent, the Moscow metropolitan area[2] is characterized by an increased risk level, due to the extremely high density and concentration of the population in Russian context. The majority of risk assessment models are based on the calculation of population numbers in the zone, potentially affected by the threat (the higher the concentration of the population—the higher the risks). The population number is a key parameter, in addition to which such variables as the age and sex population structure, the proportion of people under or over working age, the share of citizens with special needs in population, can also be used. The population size of a particular area is a key characteristic of social vulnerability to natural and man-made emergencies used by the Ministry of Emergency Situations of Russia accordingly to normative assessments of the needs for basic necessities in case of emergencies, as well as in calculations of financial and material resources reserved in case of a potential hazard. Therefore, it is important to understand the dynamics of actual spatial population distribution across the city of Moscow, and not to rely on 'static' official statistical information.

The official statistics on the spatial distribution of the population in Moscow are very approximate nowadays. Moreover, within the framework of the daily, weekly and seasonal mobility cycles, the indicator of the present population number undergoes significant transformations, which are practically not taken into account by official statistics. The development of transport infrastructure, a significant expansion of the daily commuting zone in the Moscow metropolitan area, the changes of the phenomena of extended cycles of labor commuting—semi-migrants and seasonal migrants significantly complicate the dynamics of the population mobility within the existing urban and suburban spatial structure. Accordingly, in the area of a probable natural or man-made emergency, at a certain point of time there may be a significantly larger number of people than it is usually expected, which can significantly complicate rescue measurements, cause a shortage of material resources and forces, and increase the risks of deaths or sufferings for a large number of people. So far studying the settlement system and the features of its spatio-temporal functioning becomes especially vital.

Temporal differences of population numbers Moscow reach 40% during a year. The most dramatic ones can be observed between a winter weekday and a summer weekend night. The first one is the time of maximal concentration of people in Moscow, especially in its most attractive districts with a variety of jobs, situation of educational, shopping and entertainment functions. For example, the population of the capital on a weekday winter day reaches 13.3 million people. In the summer the population inside the city limits is about 7 million people. The intraday population density varies for 10–15% (Fig. 1). A more detailed analysis of population pulsations with high spatial and temporal detailing requires the use of fundamentally new types of analysis. For example, the methods of neural networks have been already included in practice of some European countries [29].

[2] The metropolitan area of Moscow consists of the Moscow City and the Moscow Oblast (administratively, they are two separate regions of the Russian Federation).

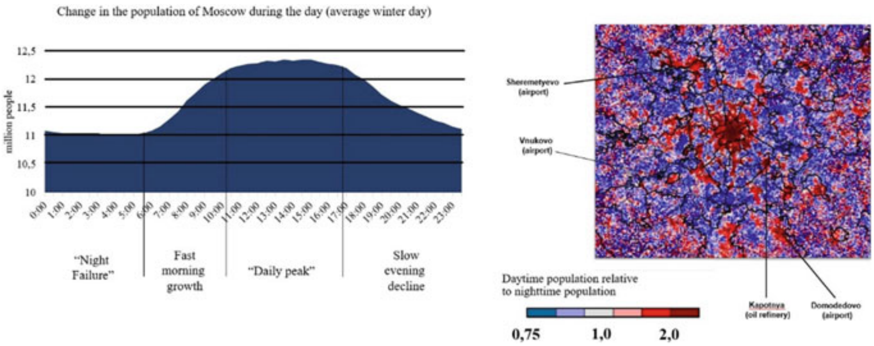


Fig. 1 Intraday changes in the population of Moscow (*Source* compiled by the authors based on data from mobile operators)

The key vulnerability parameter in the study is the density of the present population. In author's previous papers [6–12], the rationality of use of the density characteristics with the vulnerability parameters was substantiated and the relationship between them was empirically confirmed. According to the probability multiplication theorem, the higher the concentration of population and economic activity is in the area, the higher amount of damage in case of natural and (or) man-made hazards there can be. This pattern is especially clear for hazards that have an 'areal' impact and affect large territories.

The assessment of the vulnerability of Moscow residents was made taking into account the real number and movement of the population in different time periods. In addition to official statistical sources, mobile operators' data is also involved, characterizing the localization of users at a certain time moment, which can clarify the currently existing 'portrait' of spatial distribution of the population across Moscow, as well as provide better understanding for the nature of peak concentrations in the key elements of urban infrastructure and the likely risks of 'shock' population fluctuations due to potential emergencies.

Since June 1, 2021, the national standard "Safety in emergencies. Planning of measures for evacuation and dispersal of the population in case of a threat and occurrence of emergency situations"[3] has come in operation in Russia. According to this standard, the regions of Russia must develop scenarios for evacuation and spatial dispersal of the population in case of emergency. Planning evacuation scenarios is extremely important for Moscow, as it helps to avoid overloading of the transportation system (for example, serious traffic jam have practically paralyzed the city after the terrorist attacks in the Moscow Metro in 2010) and life support systems of the city. McKinsey named the transportation system of Moscow as one of the most efficient in the world: Moscow achieved 5th place in the ranking of urban mobility out of

[3] https://docs.cntd.ru/document/566430562 (accessed 15 July 2021).

24 considered global cities.[4] However, despite the fact that since the beginning of the 2010s the urban road network is continuously modernized, the traffic congestion level remains high. This fact can play a negative role in case of massive spontaneous evacuation of urban residents.

The study is based on the anonymized data of mobile operators, depicting localization of users for 2019, provided by the Moscow Department of Information Technologies. The data of mobile operators represent information on the location of network users during the day (with a time fraction of 30 min, and spatial—0.25 sq. km cells). Location determination is conducted by measurement of the distance of the person's distance from three nearest cellular stations by the signal strength of his cell phone. At the same time, the system registers only those subscribers who make calls, accumulating information about their movements during the day (measurements are carried out each 30 min). Further, calls are depersonalized and the sample is cleared of signals from modems, tablets, as well as phones used for data transmission via the Internet, and data of the subscribers who use two or more SIM cards. In total, over 36 million measurements were analyzed in more than 7 thousand time slices. The use of data from cellular operators involves the analysis of changes in the real population in Moscow that characterizes the daily and seasonal fluctuations due to the influence of commuting patterns in domains of housing, labor and leisure.

2.1 A Brief Overview of Natural and Man-Made Hazards: Case of Moscow

Areas of maximum risk are determined by the ratio of the vulnerability parameters for the existing population and the maximum danger parameters. The area of Moscow is characterized by an increased likelihood of a wide range of emergencies, both natural and man-made. Additionally, there are higher risks of terrorist attacks. The statistical base for the study included the open sources data, regulatory documents, state reports "On the state of protection of the population and territories of the Russian Federation from natural and man-made emergencies" for 1999–2020 [for example, 36], the General plan of Moscow city. The key natural hazards on the city territory include dangerous hydrometeorological phenomena [3], hazardous engineering and geological processes and phenomena (including flooding) [39], smoke pollution of vast areas of the city due to massive forest and peat fires in the surrounding Moscow Region etc.

There are some key man-made hazards which should be especially highlighted among the others in the context of Moscow. They include man-made fires, accidents at electric power facilities, accidents at railway and road transport with the release of hazardous substances and the emergence of vast areas of fire, collapse of structural

[4] https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Sustainability/Our%20Insights/Elements%20of%20success%20Urban%20transportation%20systems%20of%2024%20global%20cities/Urban-transportation-systems_e-versions.ashx (accessed 15 July 2021).

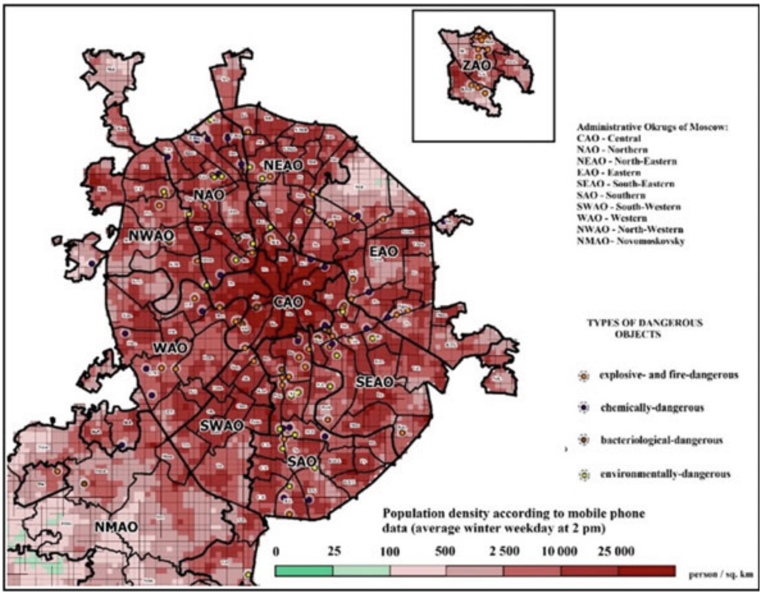


Fig. 2 Localization of hazardous industrial facilities in Moscow (*Source* compiled by the authors based on data of mobile operators)

elements of buildings and structures of transport communications, etc.), the emergence of floods due to the potential destruction of water-limiting devices on Moskva river and adjacent water reservoirs. Particular attention should be paid to the dangers associated with potentially hazardous industrial facilities (explosive, chemical, radiation and fire hazardous),[5] institutions working with pathogens of high pathogenicity. There are 117 potentially hazardous facilities on the territory of the city, including 17 radiation hazardous, 40 chemically hazardous, 6 biologically (epidemiologically) hazardous, 54 explosive and fire hazardous (see Fig. 2).[6]

Chemically hazardous facilities in Moscow (the likelihood of accidental emissions of chlorine, ammonia, acids) are primarily represented by food industry enterprises that use ammonia as a refrigerant and waterworks that use chlorine for water disinfection. Even though Mosvodokanal (Moscow Water Supplement Company) has now almost completely switched from the use of liquid chlorine to sodium hypochlorite, some enterprises still have supplies of chlorine in case of potential emergencies. The situation is similar with the use of ammonia. The transition to freon use is gradually being carried out—however, freon refrigeration equipment requires 2.5 times more

[5] Hazardous production facilities in accordance with the Russian Federal Law of 21.07.1997 N 116-FZ (as amended on 08.12.2020) "On industrial safety of hazardous production facilities".

[6] According to the resolution of the Government of Moscow of September 23, 2011 N 443-PP on the approval of the state program of the city of Moscow "Safe City" (as amended by resolutions of the Government of Moscow 2012–2019).

electricity, which is a significant barrier, and freon also cannot be considered absolutely safe. The group of chemically hazardous facilities also includes enterprises of the key and most dynamically developing sub-sectors of the Moscow chemical industry—pharmaceuticals, perfumery and cosmetics, and their production cycles include a wide range of potentially hazardous chemicals.

The radiation situation in Moscow is stable, but there is a likelihood of accidents at research reactors and equipment pieces using radioactive substances, which can cause the formation of zones of radiation contamination [45]. Additional factors of vulnerability increase can also be attributed to the reduction of sanitary protection zones around potentially hazardous facilities due to the lack of land resources (especially because of their partial residential development, stimulated with high housing demand levels). Thus, the established sanitary protection zones are actually smaller than the calculated ones in terms of area and do not fully or partially cover the impact areas around many of potentially hazardous objects (according to the analysis of data from the Integrated Automated Information System for ensuring urban planning activities in the city of Moscow). An important factor in vulnerability increase is the physical aging of fixed assets in industry and the life support sector, a decrease in working discipline control and a growth of the number of deviations from the established technological regimes, the emergence of a large numbers of small enterprises lacking the necessary technological supervision, partially insufficient equipment of industrial enterprises and facilities with modern protection systems. So far, according to the General Plan of Moscow, about 1.8 thousand hectares of residential areas with a total number of residents accounting for 93.2 thousand people are located within the sanitary protection zones. Therefore, significant financial resources which are aimed for prevention of natural and man-made emergencies in Moscow can reduce potential risk levels considerably, but not to zero.

As shown in the charts (see Figs. 3 and 4), the greatest threat to the population in Moscow consists of man-made hazards, and, in contrast to natural hazards, their sources are more clearly localized in the city space.

556 emergencies have occurred in Moscow during the period from 1999 to 2020, with only 6 of them natural, 11 biological-social. About 900 people died, and about 4000 suffered because of these emergencies. A significant part of man-made threats

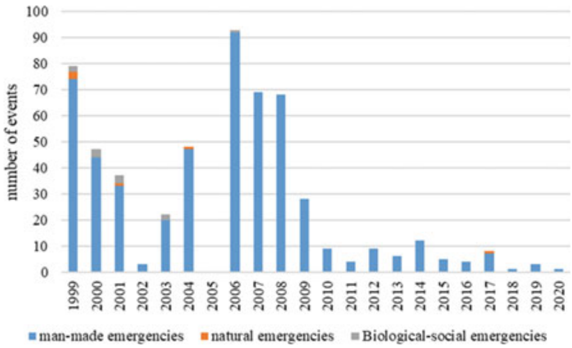


Fig. 3 The number of emergencies in Moscow, 1999–2020 (*Source* compiled by the authors based on the data of state reports "On the state of protection of the population and territories of the Russian Federation from natural and man-made emergencies" for 1999–2020)

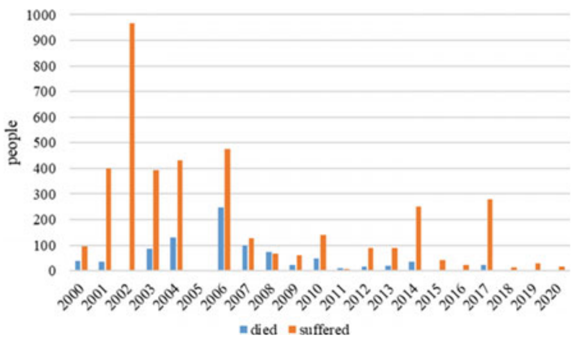


Fig. 4 The number of victims (dead and suffered) in emergencies in Moscow, 1999–2020 (*Source* compiled by the authors based on the data of state reports "On the state of protection of the population and territories of the Russian Federation from natural and man-made emergencies" for 1999–2020)

is associated with fires, collapses of buildings and structures, as well as accidents at transport facilities. The decrease in the emergency numbers in the last decade is associated with an improvement in the organization of the work of the Ministry of Emergency Situations of Russia [13], an increase in protective and preventive measures funding (the city of Moscow accounted for a third part of financial resources reserved for the emergency elimination, and 16% of all other material resources among the regions of Russia in 2020) including measurements dedicated to prevention of terrorist attacks. Terrorist attacks constituted a dramatic share in the structure of emergency situations in Moscow until 2010 and, despite the increased security levels, the likelihood of their occurrence still remains at a high level.

2.2 Limitations of Official Statistics

The Moscow metropolitan area is a complex dynamic system characterized by continuous population movement and shows specific daily, weekly and seasonal patterns. The complexity of the research object (settlement system) and the subject (population vulnerability to natural and man-made hazards) require new research approaches based on data with higher spatial and temporal detailization. For a long time, the most accurate source of Russian demographic statistics was the Rosstat (Russian Federal Statistics Service) data, including current registration and census population datasets).

Two main distorting factors need to be highlighted in the case of Moscow: the imperfect technology of census organization and the high dynamism of the population inside the urban agglomeration. Earlier studies have often highlighted the imperfection of census technology. As a result, the discrepancy between the official statistics and the real situation has been articulated [5]. One of the clearest examples of this discrepancy shows the systematic inconsistencies between the data of the population censuses and the current population registration: according to the population censuses of 2002 and 2010, there were significantly more residents in Moscow than it was estimated during the census dates [35].

As a result, census data overestimate the size of the largest urban centers such as Moscow, while current census data, on the contrary, underestimate it. At the same time, the scale of the existing distortions was not clear for a long time and was the subject of numerous speculations. It was often mentioned in the media sources that much more people live in Moscow and its satellite towns than is stated by official statistics. However, no representative data source was provided to test this hypothesis, although within the framework of various design works there were attempts to calculate the real numbers of Moscow population using so-called 'indirect indicators' of metro traffic, waste volumes, housing construction or food consumption, which provided only approximate results.

Moreover, there is practically no commuting data available in Russian statistics. This fact interferes any research on the mobility dynamics. At the same time, every day Moscow receives hundreds of thousands of commuters, which radically change the density of population distribution across its territory. The high population mobility levels are associated with weekday-weekend and winter-summer cycles, which also creates a specific imprint on the density indicators data.

As a result, it is impossible to present a reliable picture of population spatial distribution using methods of analysis based on official statistics. Nevertheless, these data depicting the differentiation of urban space according to the vulnerability degree is vital for correct operation of the Ministry of Emergency Situations and related organizations in terms of the rational distribution of forces and resources throughout the city. This situation would be considerably improved by application of dynamic data methods and technologies. Moreover, it is necessary the development of long-term plans for the urban spatial development of Moscow aiming to reduce potential vulnerability by a more even distribution of housing construction and creation of working places. The goal of Sendai Framework for Disaster Risk Reduction 2015–2030[7] is to "prevent new and reduce existing disaster risk through the implementation of integrated and inclusive economic, structural, legal, social, health, cultural, educational, environmental, technological, political and institutional measures that prevent and reduce hazard exposure and vulnerability to disaster". In our opinion, in addition to these key measures, special emphasis should be placed on measures for territorial planning.

3 Results and Discussion

The majority of the population distribution assessments across the Moscow are static and do not reflect the real dynamics of the city's population observed in the framework of labor, educational, consumer, cultural, leisure and recreational mobility. The data of mobile operators shows that density as a characteristic of the potential vulnerability of urban areas for a weekday winter day (the time period when the maximal share

[7] https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf (accessed 15 July 2021).

of people is situated at their working places) undergoes significant transformations in comparison with a weekday winter night, when most of the city's residents are staying at home.

The spatial structure of Moscow is transformed the same way within the framework of weekly-seasonal cycles. On weekends approximately 0.7–1.0 million inhabitants leave the city limits. The tendency reaches its peak during the summer periods, when the city loses up about a third part of its population (3.0–3.5 million people).

At the municipal level the magnitude and direction of inconsistencies vary widely. Depending on these characteristics, it is possible to distinguish areas in which population density is overestimated or underestimated in comparison with the official statistics (see Fig. 5).

Let us consider two of the most indicative time sections characterizing the population distribution at night and daytime hours for the winter season, when the majority of residents prefer to stay in the city.

Generally, three groups of municipalities with particular specifics of density characteristics can be distinguished in the city: 1. Highly attractive areas of the Center

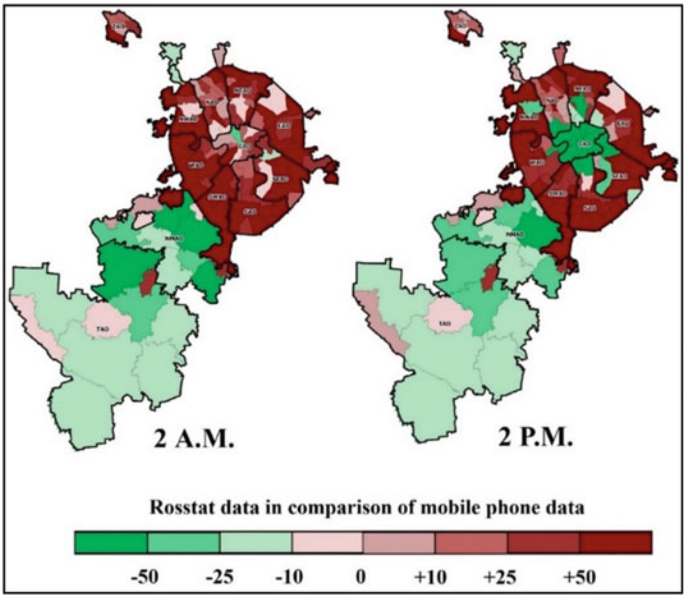


Fig. 5 Comparison of population density calculations according to data from mobile operators and Rosstat (for beginning of 2020) (*Source* compiled by the authors based on data of mobile operators)

of Moscow; 2. Peripheral areas of Old Moscow (so-called 'sleeping areas'[8]); 3. Territories of 'New Moscow'.[9]

There is a relatively high concurrence between the night population density values and the official data in the city center. However, in the daytime the population density in municipalities of Arbat, Tverskoy, Meshchansky and others usually increases by 3–4 times.

On the contrary, there is an overcounting of the number of residents living here by the official statistics in the "sleeping areas" of the city. The official statistics, in comparison with the mobile operators' data, averagely overestimates the population of the residential areas of Moscow by 10%. A significant part of the population living in these municipalities commutes to work in the city center, so far there is a significant excess of the night population over the daytime on weekdays. There are four municipalities with significant levels of underestimation of the inhabitants by Rosstat: Metrogorodok, Nekrasovka, Molzhaninovsky and Kurkino. All of them are characterized by a relatively low absolute population size, which leads to a so-called 'low-base' effect caused by active housing construction.

For the rest of Moscow municipalities, the distortion levels are constituted by the influence of several factors. Firstly, some of Moscow residents registered in their dachas and second suburban[10] houses (more than half a million owners of SIM cards registered in Moscow stay out of the city limits on a weeknight). Secondly, the situation is caused by shortcomings of the population censuses, which serve as reference points for the subsequent current registration. Finally, significant underestimations of the population density by official data are observed in New Moscow (in which the general distortion sometimes reaches 2.5 times), as well as in areas similar to it in terms of the urban spatial structure (for example, in the aforementioned districts of new development: Nekrasovka, Kurkino and Molzhaninovsky).

Therefore, we would like to stress the need to adjust some of the population estimation methods of the Russian Ministry of Emergency Situations basing of the results obtained. For example, according to the methodological recommendations of the Ministry of Emergency Situations, the regional forecasts for potential emergencies are made using the population size (each forecast is based on the size of the population and the affected population), and the need for priority supplement reserves for population are calculated. There are strict standards for providing the population with everything necessary in such cases (organization of the protection from damaging effects of a man-made accident, provision of personal protective equipment, basic supplements, standards of medical support, medicines, etc.). All these reserves are created according to the requirements of each territory. However, the calculations

[8] The 'Sleeping areas' are the zones of standardized residential buildings, which were formed mainly in the peripheral territories of the city during the 1960-1980s.

[9] 'New Moscow' is a peripheral part of the city to the southwest from the center, administratively included in the city limits in 2010. 'Old Moscow' is an unofficial name of the other parts of the city, mostly situated inside the Moscow Circle Highway (MKAD).

[10] The following tendency is often observed in Moscow metropolitan area: residents buy cheaper (including the cost of utilities) housing in the Moscow Oblast and provide their apartments in the city for rent to tenants.

take into account not the actual but the statistically registered population. Authors recommend to adjust it according to the actual population distribution, using the maximum possible population size in each area as a benchmark for calculations of reserves.

4 Conclusions

The complexity of the Moscow spatio-temporal structure and its interactions with other regions and countries predetermines constant changes in one of the most important socio-economic indicators—population density. The level of population vulnerability also changes: in a certain location, the number of potential victims dead and suffered in an emergency during one time period may be several times less than in another.

The discrepancies in the distribution density of city residents revealed in this paper make it possible to distinguish three groups of Moscow municipalities with different variations in the distortions between the official Rosstat data and the data of mobile operators. The first group—highly attractive areas of the urban center, they are characterized by a significant discrepancy with the official statistics of the daytime population. During the daytime, about 2 million residents of other parts of the urban agglomeration come to work there, which greatly increases the population density and leads to increased risks in case of a potential emergency. The second group is the peripheral areas of the city, for which the data of mobile operators show smaller numbers of night and day population than the current accounting data provides. Finally, the third group—areas of new development, especially areas of New Moscow. There is a twofold underestimation of the population by official statistics due to active housing construction. Contradicting to the cases of the first group, the night and day populations in these municipalities are comparable with each other.

According to calculations based on data of mobile operators, there are 850 thousand residents of Moscow in the immediate vicinity of potentially dangerous objects (within a radius of 1000 m) during the cold season of the year (in the summer months there are less numbers of them—600–650 thousand people). During the daytime, this number increases by one and a half times to 1.3 million people (in the summer it reaches to about 1.0 million people). Thus, about 9–12% of the city's population is regularly situated in a high-risk zone in the case of a man-made emergency. The residents of the Southern, South-Eastern and Northern Administrative Districts of the capital are in the greatest danger. There are a significant number of hazardous objects of all considered groups (explosive and fire hazardous; chemically hazardous; enterprises working with causative agents of dangerous diseases; environmentally hazardous) on their territory. This situation leads to the need of development of comprehensive specific measures to prevent potential risks. The least dangerous facilities are located in the South-West Administrative District, as well as in New Moscow. The rest of the districts are characterized by an average number of such enterprises.

In further research works related to this field, the indicated discrepancies can be used both in order to improve the methodology for vulnerability analysis of the population to potential emergencies, and in order to optimize the spatial structure of the Moscow Region in terms of risk level reduction, as well as in order to improve the quality of risk management (organization of preventive measures, measures for elimination of negative effects, etc.). The goal of Sendai Framework for Disaster Risk Reduction 2015–2030[11] is to "prevent new and reduce existing disaster risk through the implementation of integrated and inclusive economic, structural, legal, social, health, cultural, educational, environmental, technological, political and institutional measures that prevent and reduce hazard exposure and vulnerability to disaster". In our opinion, in addition to these key measures, special emphasis should be placed on measures for territorial planning. A reliable picture of the population distribution throughout the city in different time periods makes it possible to form a list of necessary measures aiming to prevent excessive concentrations of people in particular areas. Authors also see the practical possibility to adjust some methods of the Ministry of Emergency Situations basing on results obtained—recalculation of reserves needed in case of emergencies, assessment of the capacity of infrastructure facilities for population protection etc.

Acknowledgements The research was funded by Russian Foundation for Basic Research and Moscow city Government according to the project 21-35-70004.

References

1. Ahas, R., Silm, S., Järv, O., et al.: Using mobile positioning data to model locations meaningful to users of mobile phones. J. Urban Technol. **1**(17), 3–27 (2010)
2. Akimov, V.A.: Assessment of the state of science in the Russian Federation on the study of technogenic threats. Civ. Secur. Technol. **15**, **1**(55), 4–9 (2018) (in Russian)
3. Akimov, V.A., Durnev, R.A., Sokolov, Y.I.: Dangerous hydrometeorological phenomena on the territory of Russia. FGU VNII GOChS (FC), Moscow (2009).(in Russian)
4. A Study on Urban Mobility and Dynamic Populationи Estimation by Using Aggregate Mobile Phone Sources. CSIS Discussion Paper, 115 (2014), http://www.csis.utokyo.ac.jp/dp/115.pdf, last accessed 2021/05/11
5. Babkin, R.A.: Estimation of the population of municipalities in the Moscow metropolitan region according to the data of mobile operators. Vestnik Moskovskogo Universiteta, Seriya 5: Geografiya 4, 116–121 (2020) (in Russian)
6. Baburin, V.L., Badina, S.V.: Evaluation of the social-economic potential of natural hazard-subjected territories. Vestnik Moskovskogo Universiteta, Seriya 5: Geografiya 5, 9–16 (2015) (in Russian)
7. Baburin V.L., Badina S.V.: Forecasting of probable socio-economic losses from natural hazards (case study of the "Northern Caucasus resorts" tourist cluster). Vestnik Moskovskogo universiteta. Seriya 5, Geografiya 2, 25–34 (2021) (in Russian)

[11] https://www.preventionweb.net/files/43291_sendaiframeworkfordrren.pdf (accessed 15 July 2021).

8. Baburin, V.L., Badina, S.V., Derkacheva, A.A. et al.: Economic assessment of debris flow risk (Case study of the Siberian federal district). Vestnik Moskovskogo Universiteta, Seriya 5: Geografiya 4, 3–14 (2019) (in Russian).
9. Badina, S.V.: Assessment of probable economic and social damage from hazardous natural processes in the North-West and Central Caucasus. INTERCARTO/INTERGIS **25**, 219–228 (2019). (in Russian)
10. Badina, S.V.: Estimation of the value of buildings and structures in the context of permafrost degradation: The case of the Russian Arctic. Polar Sci. (2021). https://doi.org/10.1016/j.polar.2021.100730
11. Badina, S.V.: Prediction of socioeconomic risks in the cryolithic zone of the Russian Arctic in the context of upcoming climate changes. Stud. Russ. Econ. Dev. **31**(4), 396–403 (2020)
12. Badina, S.: Socio-economic potential of municipalities in the context of natural risk (case study—Southern Siberian regions). IOP Conf. Ser.: Earth Environ. Sci. **190**, 1–7 (2018)
13. Badina, S., Porfiriev, B.: Natural Hazards Governance in Russia. Oxford University Press, United Kingdom (2019)
14. Bengtsson, L., Lu, X., Holme, P.: Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. **29**(109), 11576–11581 (2012)
15. Bird, D.K.: The use of questionnaires for acquiring information on public perception of natural hazards and risk mitigation—a review of current knowledge and practice. Nat. Hazards Earth Syst. Sci. **9**(4), 1307–1325 (2009)
16. Bogorov, V.G., Novikov, A.V., Serova, E.I.: Self-knowledge of the city. Archeology of the periphery. Materials of the Moscow urban forum, pp. 380–405. Meganom, Strelka Institute, Moscow (2013) (in Russian)
17. Botzen, W. J. W. et al.: Integrated disaster risk management and adaptation. Loss and damage from climate change, pp. 287–315. Springer, Cham (2019)
18. Bründl, M., et al.: The risk concept and its application in natural hazard risk management in Switzerland. Nat. Hazards Earth Syst. Sci. **9**(3), 801–813 (2009)
19. Calabrese, F., Diao, M., Lorenzo, D., et al.: Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. Transp. Res. Part C: Emerg. Technol. **26**, 301–313 (2013)
20. Cavallo, E., Noy, I.: The Economics of Natural Disasters. A Survey. Inter-American Development Bank, Washington (2010)
21. Choi, C.: Does economic growth really reduce disaster damages? Index decomposition analysis for the relationship between disaster damages, urbanization and economic growth and its implications. Int. J. Urban Sci. **20**(2), 188–205 (2016)
22. Cruz, A.M., Okada, N.: Consideration of natural hazards in the design and risk management of industrial facilities. Nat. Hazards **44**(2), 213–227 (2008)
23. ESSnet Big Data.: European Commission, http://www.ec.europa.eu/eurostat/cros/content/essnet-big-data_en (2020) last accessed 2021/05/01
24. Garschagen, M., Romero-Lankao, P.: Exploring the relationships between urbanization trends and climate change vulnerability. Clim. Change **133**(1), 37–52 (2015)
25. Hanewinkel, M., Hummel, S., Albrecht, A.: Assessing natural hazards in forestry for risk management: a review. Eur. J. Forest Res. **130**(3), 329–351 (2011)
26. Haynes, K., Barclay, J., Pidgeon, N.: Whose reality counts? Factors affecting the perception of volcanic risk. Volcanol. Geoth. Res. **172**, 259–272 (2008)
27. Insurance against emergencies. Under total. ed. S.I. Voronova. FGBU VNII GOChS (FC), Moscow (2016) (in Russian).
28. IPCC, 2014: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, USA (2014)
29. Kim, D.Y., Song, H.Y.: Method of predicting human mobility patterns using deep learning. Neurocomputing **280**, 56–64 (2018)
30. Kunreuther, H.: Mitigation and financial risk management for natural hazards. Geneva Pap. Risk Insur.-Issues Pract. **26**(2), 277–296 (2001)

31. Lujala, P., Lein, H., Rød, J.K.: Climate change, natural hazards, and risk perception: The role of proximity and personal experience. Local Environ. **20**(4), 489–509 (2015)
32. Makarova, E.A., Porfiriev, B.N.: Economic assessment of damage from natural disasters and catastrophes. Bull. Russ. Acad. Sci. **84**, 12, 1059–1072 (2014) (in Russian)
33. Makhrova, A.G., Babkin, R.A., Kazakov, E.E.: Dynamics of the daytime and nighttime population as an indicator of structural and functional changes in the territory of the city in the zone of influence of the Moscow Central Ring using data from mobile operators. The contours of global transformations: politics, economics, law, 13, 1, 159–179 (2020) (in Russian)
34. Makhrova, A.G., Kirillov, P.L.: Seasonal pulsation of settlement in the Moscow agglomeration under the influence of dacha and labor pendulum migration: approaches to the study and assessment. Reg. Stud. **1** (47), 117–125 (2015) (in Russian)
35. Mkrtchyan, N.V.: Population dynamics of Russian regions and the role of migration: a critical assessment based on the 2002 and 2010 censuses. Bull. Russ. Acad. Sci.: Geogr. **5**, 28–41 (2011). (in Russian)
36. Orttung, R.W., Anisimov, O., Badina, S. et al.: Measuring the sustainability of Russia's arctic cities. Ambio, 2020.
37. Osipov, V.I.: Urbanization and natural hazards Tasks to be solved. Geoecology **1**, 3–9 (2007). (in Russian)
38. Osipov, V.I., Aksyutin, O.E., Ishkov, A.G. et al.: Adaptation is the most important technology for the development of the Russian subarctic territories. Bull. Russ. Acad. Sci. **89**, 1, 56–63 (2019) (in Russian)
39. Osipov, V.I., Burova, V.N., Zaikanov, V.G., et al.: Map of large-scale (detailed) engineering-geological zoning of the territory of Moscow. Geoecology **4**, 306–318 (2011). (in Russian)
40. Osipov, V.I., Burova, V.N., Zaikanov, V.G., et al.: Basics for assessing the vulnerability of territories to hazardous natural processes that determine emergency situations (principles and methodological approaches). Geoecology **3**, 195–203 (2015). (in Russian)
41. Popov, A.A., Kuricheva, E.K.: Development of housing construction in the 2010s as a factor in the transformation of the Moscow agglomeration. Regional Studies, 1 (47), 104–116 (2015) (in Russian)
42. Porfiriev, B.N.: Nature and economics: risks of interaction. (Ecological and economic essays). Edited by Academician V.V. Ivanter. Ankil, Moscow (2011) (in Russian)
43. Porfiriev, B.N.: Economic consequences of the catastrophic flood in the Far East in 2013. Bull. Russ. Acad. Sci., **85**, 22, 128–137 (2015) (in Russian)
44. Roncancio, D.J., Nardocci, A.C.: Social vulnerability to natural hazards in São Paulo Brazil. Nat. Hazards **84**(2), 1367–1383 (2016)
45. Shaposhnikov, A.S.: Analysis of the effectiveness of monitoring and forecasting systems for natural and man-made emergencies using the example of Moscow. Civ. Secur. Technol. **6**, 3–4, 210–215 (2009) (in Russian)
46. Streletskiy, D., Shiklomanov, N., Suter, L.: Assessment of the cost of climate change impacts on critical infrastructure in the circumpolar Arctic. Polar Geogr. **42**, 267–286 (2019)
47. Tizzoni, M., Bajardi, P., Decuyper, A., et al.: On the use of human mobility proxies for modeling epidemics. PLoS Comput. Biol. **7**(10), 1–35 (2014)
48. Tokareva, E.A.: Organization of financing of the consequences of natural disasters in foreign countries. Edited by L.I. Tsvetkova. Ankil, Moscow (2015) (in Russian).
49. UNDRR (2019): Global Assessment Report on Disaster Risk Reduction. United Nations Office for Disaster Risk Reduction (UNDRR), Geneva, Switzerland (2019).
50. Zhou, H., et al.: Resilience to natural hazards: a geographic perspective. Nat. Hazards **53**(1), 21–41 (2010)

CITADINE: A Platform to Provide the Experiences of Survivors of Natural Disasters as Open Educational Resources for Risk Communication

Michael Klafft, Ivana Harari, Ricardo Gacitua Bustos, Agnieszka Dudzińska-Jarmolińska, and Solhanlle Bonilla Duarte

Abstract This paper presents the concept, and the prototype of a platform that enables citizens from local communities to share their experiences with historic disasters online as open educational resources for teachers, community builders, urban planners, or volunteers. Key focus of the platform is to highlight the impact of disasters on citizens' everyday lives. Based upon an overview of previous research, a requirements analysis, personas and related use cases, we present the insights into the prototypical implementation of the platform. We also provide a brief outlook upon possible practical uses of the platform, and its materials.

Keywords Risk communication · Emergency preparedness · Open educational resources · Disaster impact · UN Sendai Framework · Stakeholders knowledge sharing

1 Introduction

A key challenge in disaster risk communication is to motivate the population to be prepared for a possible disaster, and to take the required protective actions in time. Initially, practitioners and researchers believed that a key reason why citizens failed to

M. Klafft (✉)
Jade University of Applied Sciences, Wilhelmshaven, Germany
e-mail: Michael.Klafft@jade-hs.de

Fraunhofer Institute for Open Communication Systems, Berlin, Germany

I. Harari
Universidad Nacional de La Plata, La Plata, Argentina

R. G. Bustos
Universidad de La Frontera, Temuco, Chile

A. Dudzińska-Jarmolińska
Warsaw University, Warsaw, Poland

S. B. Duarte
Instituto Tecnologico de Santo Domingo INTEC, Santo Domingo, Dominican Republic

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_12

protect themselves was lack of information [1]. However, the discussions around this information deficit approach of risk communication has shown that simply informing the population about risks is often insufficient to cause people to act [2]. Protection motivation theory explains this by pointing out two additional cognitive steps that individuals have to pass before taking protective action: threat appraisal and coping appraisal. This means that citizens will only act if they believe that a threat applies to them, and that they can successfully prevent or mitigate its impact [3]. As a consequence, risk communication is a complex process comprising the following steps (adapted from [4], based on previous works such as [5, 6]):

- Production and dissemination of the risk communication intervention,
- Receiving, noticing, and understanding the content (on the side of the individual citizen—at this point, the information deficit has been covered),
- Verification of the content by the individual, and personalization to his or her own situation (threat appraisal),
- Belief in the suitability of the proposed protective action, and perceived ability to conduct this action successfully (coping appraisal).

While reaching out to the population with information has been successfully achieved in many cases, threat appraisal, and coping appraisal, remain two key obstacles that have to be overcome by risk communicators. This is particularly relevant in case of natural disasters that are severe but rare, so that many citizens in a certain area never personally experienced such a disaster. To close this gap in the risk communication process, active community participation in risk communication interventions has been proposed, and it was shown that this approach is very effective in increasing disaster preparedness [2]. This paper's aim is to propose, and present a tool and a concept how to involve the community more actively in the risk communication process. The approach is centered around a web-based platform, which allows citizens to upload and share multimedia materials (e.g. survivor interviews, photos, or videos) on historic natural disasters from their region. Goal of the platform is to provide this community-generated content in a structured way as open educational resources. Risk communicators and multipliers such as teachers or volunteers from disaster management organizations shall be able to search for content of the disasters' impacts on citizens' everyday lives. By showing the concrete impact of previous disasters on citizens' lives in the region, the threat appraisal process is strongly supported. Discussion of successful coping strategies in the uploaded survivor interviews will also enhance coping appraisal, so that communication interventions using materials from the platform can support citizens' compliance with recommended protective measures.

2 Current State of the Art

Using websites on historic natural disasters as a tool for risk communication has been established practice for several years. One example is the website of the city of

Hamburg, which provides a disaster map, as well as photos, and texts with survivors' experiences about the storm surge disaster in 1962 [7]. Similarly, concepts on how to use digital storytelling on such websites have been designed [8], and websites about historic natural disasters which enhance disaster learning by appealing to the visitor's emotions have been proposed and implemented [9]. However, none of these sites offers the possibility to upload user-generated content on regional disasters, and to tag the content in a structured way. Furthermore, these sites do not foresee any search and download functionalities, which makes it difficult to retrieve and extract suitable materials for subsequent risk communication activities (and, in many cases, the materials are copyrighted contents which are not suitable for royalty-free use outside of the site). One exception is the Aceh Digital Disaster Archive, which explicitly focuses on freely accessible content on the 2004 tsunami disaster, including some multimedia materials created in cooperation with the local population. The platform also encourages integration of its contents into social media. However, again, no targeted tagging, search and download functionalities are presented, and the utilization concept of its materials focuses on disaster-related "dark tourism" [10]. While this approach might be suitable in areas with a high (economic) dependency on tourism, we propose to focus more on traditional multiplication approaches involving schools, universities, and volunteer organizations. Furthermore, we propose to develop a platform that covers different types of disasters, (a) to provide a single point of information on all locally relevant disasters, but also (b) in order to support learning about disasters that are currently occurring elsewhere, but which are likely to gain increased local relevance in the future due to phenomena such as climate change.

3 Background: The CITADINE Project

The work presented here is part of the ongoing international research project CITADINE (Citizen Science and Nature-based Solutions for Improved Disaster Preparedness) involving partners from Germany, Poland, Argentina, Chile, and the Dominican Republic. Goal of this project is to develop "an ICT tool and an organizational concept for collecting existing knowledge in the population on low-frequency, high impact disasters, and to make this knowledge and experience usable for risk and crisis communication, urban planning, and discussions with relevant decision makers" [11]. The project places a focus on the qualitative impact of disasters on different dimensions of everyday's lives: economy, health, infrastructure, social well-being and cohesion, work life or education—all of which have the potential to support threat appraisal. In order to support coping appraisal, good practices for disaster preparedness are presented, either as part of the survivor interviews, or by providing good examples of cost-effective and easy-to-implement nature-based solutions in gardening, urban planning and building construction that can help to mitigate the impact of natural disasters (heat waves and floods in particular). The transcontinental dimension of the project with actors from Europe, South America, and the Carribean shall also support mutual information exchange and learning, and shall enable the transfer (and possible reuse) of good practices and examples for risk communication across the participating regions.

4 Methodology

In conceptual design and implementation, the project followed a multi-method approach comprising the following elements:

1. **Semi-structured interviews with potential users to discuss platform requirements** (and to assess initial concepts of the user interface at a later stage). Interviews were conducted with several disaster suvivors, seven teachers, three representatives of not-for profit volunteer organization, and the head of an extracurricular learning place, all in their function as possible multipliers.
2. A **workshop with urban planners and architects** to discuss the selection and presentation of nature-based solutions for disaster preparedness on the platform.

Results from steps 1 and 2 were not only used for requirements engineering, but also for defining the personas that should be used as guidance during platform design and development. Personas are “an alternative method for representing and communicating customer needs” and provide “product designers with a vivid representation of the design target” [12]. Finally, a **usability assessment** of platform mock-ups was conducted with experts from the project (one accessibility counsellor, a usability experts, two risk-communication specialist, one expert on nature-based solutions). Mock-ups were revised according to the results of this assessment before starting with platform implementation.

5 Personas and Use Cases

Early in the design phase, the following personas were defined as representations of typical user groups who will likely use the platform for these specified use cases:

- A **teacher** who uses the platform to receive information, photos, interviews and videos from contemporary disasters witnesses to share them as examples in her lectures with students. The examples shall demonstrate disaster impact, but should also support theoretical findings addressed in the course (e.g. on disaster-related phenomena). She would also upload her own educational materials to the platform in order to share them with colleagues.
- An **urban planner** who wants to use the platform to be able to explain disaster related planning decisions to citizens in a better way (e.g. refusals of building permits in areas at risk, or protective construction measures taken). He needs the materials in order to support risk appraisal by highlighting that disaster risks are

real. A second use case relates to possible exchanges with other urban planners about best practices on nature-based solutions for disaster preparedness.

- A **community builder** who wants to bring the local community together in order to facilitate the exchange of information on and enhance the preparedness for disasters. She uses materials from the platform as stimuli for community events, but would also plan risk communication events with the community, and in cooperation with survivors, to create new materials for the platform.
- A **data collector** who helps others to share their experiences on disasters, e.g. by recording interviews with eyewitnesses, by collecting disaster photographs or videos, and by uploading the materials to the platform. Data collectors can be family members of disaster survivors, or independent citizens who realize the importance of disaster preparedness (e.g. because they or their beloved ones have been personally affected by a disaster), and are therefore motivated to support risk communication activities.
- A **disaster survivor** who is willing to share experiences on historic disasters which potentially happened long ago. As an elderly person, survivors are often not well versed in the use of IT, and of mobile devices in particular. They may need help in producing materials for the platform.
- An **experienced volunteer** who works for a not-for-profit disaster management organization such as the Red Cross and who wants to use materials from the platform as stimuli for disaster preparedness trainings with other local volunteers.

6 Requirements

Semi-structured interviews with representatives of different target groups led to the following key requirements of the platform[1]:

- Upload of contributions to the platform should be possible using desktop computers (with mobile uploads being a desirable second option).
- Metadata on publications should include at least disaster type, time, and location, and it should be possible to filter publications according to these criteria.
- It is desirable if the content could further be specified, whether it is a contribution on how disasters appeared, how people prepared for a disasters, what the consequences were, or how it was possible to cope with these consequences.
- Search for different media types should be possible (disaster videos, transcripts, photos, eyewitness interviews).
- It should be possible to download materials from different countries in order to compare how they are affected by / can cope with disasters.
- Video contributions from eyewitnesses or disaster videos should be rather short (recommended durations ranged from approx. 2 to 5 min). In case of longer films and interviews, these should be cut into shorter segments, which address specific aspects of the disaster, and should be tagged accordingly.

[1] Please note that this is just a selection of key some key requirements.

- During uploads, contributors have to confirm that rights will be granted to place the material in the public domain (e.g. CC-BY-NC [13] or a similar license).
- This requires a registration process for contributors, all other persons shall be able to use the platform without registration.
- The platform should provide advice or examples on how to use the different types of materials in practice, including possible tasks for pupils / students.
- It would be good if teachers can upload educational materials that they produced with input from the platform, or project results that their pupils / students have produced.
- It would be beneficial if links to additional materials, e.g. official guidelines from disaster management agencies, could be provided.
- Regarding the presentations of (nature-based) solutions for disaster preparedness, the material should be easy to understand, but with additional references to background information for experts.

7 Prototypical Implementation

Technically, the developed platform is based upon WordPress [14] and MySQL [15]. FFmpeg [16] was used as a tool to cut (longer) videos and interviews into different topical segments before upload. Key reason for this selection of technologies was the aim to make instances of the platform available free of charge to interested institutions with low budgets, such as schools, or volunteer organisations. Therefore, the prototypical implementation had to be based on open source components that can be used without license fees.

The landing page is displayed in Fig. 1. It provides some background information on the project, and provides users with links

- for sharing their own experiences,
- for searching for disaster-related materials,
- for informing themselves about nature-based solutions for disaster prevention, and
- for sharing educational materials as part of school communities.

Standard search functionalities enable users to filter contributions by country, disaster type, time of disaster, and language of the contribution. The advanced search, which is currently being implemented, will also allow for more precise location-based searches on a map, and for tag-related searches (e.g. for contributions addressing emotional impacts, economic impacts, or health impacts of a disaster). Figure 2 presents an example of one specific contribution, here a survivor interview of someone who experienced a severe thunderstorm in a small town in Germany. In this case, the interview is available in German, with English subtitles. The location of the disaster addressed is shown on a map below the contribution.

A separate section of the platform addresses measures for disaster prevention, with a specific focus on easy-to-implement nature-based solutions. These solutions

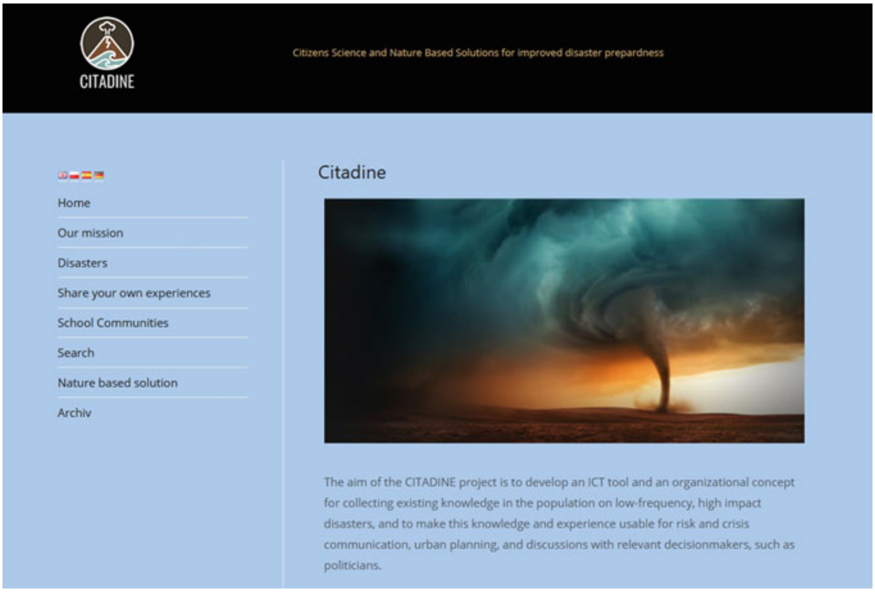


Fig. 1 Landing page of the CITADINE platform

will be linked to the corresponding contributions on historic disasters which they could have prevented or mitigated. Figure 3 provides an overview of the different types of nature-based solutions which are being presented (or will be presented).

As shown, the CITADINE platform is also multilingual. At the moment, key functionalities of the platform have been implemented, but it still needs to be filled with more content, as some contributions need to be completed or translated into different languages.

8 Conclusion and Outlook

In this paper, we have outlined the requirements and presented the concept for a platform, which provides the experiences of survivors of natural disasters as open educational resources for risk communication. The prototype of the platform is currently undergoing beta testing at Universidad Nacional de La Plata, where it is used as part of a class on mobile programming. Students have the task to create mobile games or mobile risk communication applications using materials from the platform. The next steps will be to conduct a similar course in autumn 2021 in Germany, and to use materials from the platform to produce risk communication interventions (events, exhibitions, community events) in the Dominican Republic, and in Argentina. A direct success criterion for the platform will be the number of projects, lectures, courses, and risk communication interventions, which are using materials from the

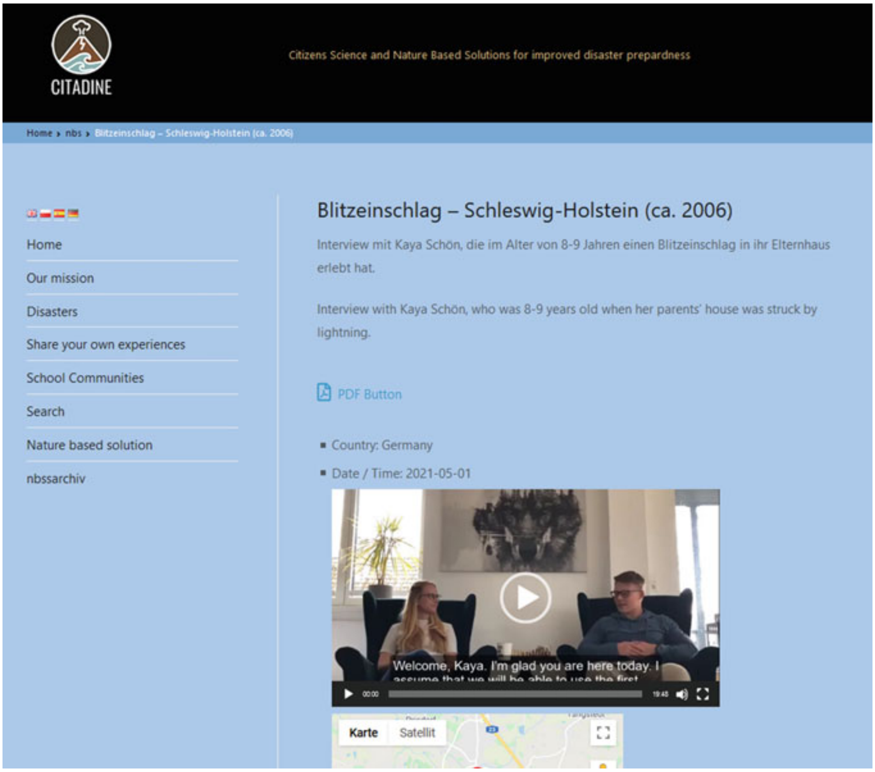


Fig. 2 Example of a survivor interview

Fig. 3 Overview of Nature-based solutions explained on the platform

platform (and to what extent materials are used). Another success criterion will be the perceived usefulness and perceived ease of use of the platform among the addressed user groups (in accordance with the technology acceptance model by Davis [17]).

Indirectly, success criteria will be (a) the number of citizens actively involved in platform-related events, (b) the number of visitors / persons reached, and (c) a better threat appraisal and a better coping appraisal of those who have been reached by the communication interventions. However, this would only measure a combination of the quality of the event itself, and the quality of the platform and its materials.

Beyond the applications proposed in this paper, the approach and concept also have the potential to be developed further into a methodology which makes historic disaster information available for ad-hoc consideration in decision-making during an emerging or ongoing disaster, which is another issue for future research.

Acknowledgements This work is funded by the German Federal Ministry of Education and Research (BMBF), the Chilean National Commission for Scientific Investigation and Technology (CONICYT), the Dominican Ministry of Higher Education, Science and Technology (MESCyT), and the Polish National Centre for Research and Development through their respective CITADINE project grants as part of the ERANet-LAC program. We would like to thank all participating students and citizens for their support, and we thank UNLP for their participation without funding. We also thank three anonymous reviewers for their valuable comments.

References

1. United Nations (Ed.): Sendai Framework for Disaster Risk Reduction 2015–2013. The United Nations Office for Disaster Risk Reduction, Geneva. https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030, last accessed Jul 23rd, 2021
2. Abunyewah, M., Gajendran, T., Maund, K., Okyere, S. A.: Strengthening the information deficit model for disaster preparedness: Mediating and moderating effects of community participation. Int. J. Disaster Risk Reduct. **46**, 101492 (2020)
3. Neuwirth, K., Dunwoody, S., Griffin, R.J.: Protection motivation and risk communication. Risk Anal. **20**, 721–734 (2000)
4. Klafft, M.: Die Warnung der Bevölkerung im Katastrophenfall. In: Sicherheitskritische Mensch-Computer-Interaktion, 2nd edn, pp. 341–362. Springer Vieweg, Wiesbaden (2021)
5. Jagtman, H.M.: Cell broadcast trials in the Netherlands: using mobile phone technology for citizens alarming. Reliab. Eng. Syst. Saf. **95**(1), 18–28 (2010)
6. United Nations (Ed.): Global survey of early warning systems—an assessment of capacities, gaps and opportunities toward building a comprehensive global early warning system for all natural hazards. Report prepared at the request of the Secretary-General of the United Nations, UN/ISDR (2006)
7. Stadt Hamburg: Gedenken an die Katastrophe—Hamburger Sturmflut von 1962. https://www.hamburg.de/sturmflut-1962/, last accessed 2021/06/14
8. Wittich, S: Digitales Storytelling. Entwicklung einer Website zur Hamburger Sturmflut von 1962. HAW Hamburg: Bachelor Thesis (2016)
9. Klafft, M., Strangmann, L., Fianke, M.: Communicating storm surge risks via risk communication websites: a novel approach. Mensch und Computer 2019-Workshopband, pp. 291–296, Gesellschaft für Informatik e.V., Bonn (2019)

10. Watanave, H.: Lesson learned from the Aceh Tsunami of 2004: a digital multimedia display of the Aceh archive using an open-source platform for sustainable disaster risk reduction and global information. Dejitaru Akaibu Gakkaishi **2**(2), 124–127 (2018)
11. Klafft, M., Dudzińska-Jarmolińska, A., Harari, I., Bustos, R.G., Duarte, S.B., Morrobel, T.: A citizen science approach using information systems to provide qualitative information on historic natural disasters to risk communicators and the general public. Umweltinformationssysteme 2018 Umweltdaten-in allen Dimensionen und zu jeder Zeit? pp. 197–208 (2018)
12. Miaskiewicz, T., Kozar, K.A.: Personas and user-centered design: how can personas benefit product design processes? Des. Stud. **32**(5), 417–430 (2018)
13. Creative Commons (Ed.): About The Licenses. https://creativecommons.org/licenses/?lang=en, last accessed July 23rd, 2021
14. Wordpress.org (Ed.): WordPress Coding Standards. https://make.wordpress.org/core/handbook/best-practices/coding-standards/, last accessed July 23rd, 2021
15. Oracle (Ed.): MySQL documentation. https://dev.mysql.com/doc/, last accessed July 23rd, 2021
16. FFmpeg.org (Ed.): FFmpeg—a complete, cross-platform solution to record, convert and stream audio and video. http://www.ffmpeg.org/, last accessed July 23rd, 2021
17. Davis, F. D.: A technology acceptance model for empirically testing new end-user information systems: theory and results, Doctoral dissertation, Massachusetts Institute of Technology, 1985

Sustainable Mobility

Evaluation of a Sustainable Crowd Logistics Concept for the Last Mile Based on Electric Cargo Bikes

Richard Schulte, Mattes Leibenath, Lars Wöltjen, Uta Kühne, and Benjamin Wagner vom Berg

Abstract During to the recent growth of e-commerce and as the number of shipments is rising last mile services are facing many challenges and hurdles especially in relation to sustainable action. Within the project NaCl measures were tested to address and to positively impact the three dimensions of sustainability on the last mile. Based on the usage of electric cargo bikes and a crowd logistics approach shipments of a regional logistics service provider and beyond of regional stationary retailers were realized during a three-month pilot phase. The project's activities and objects were evaluated regarding positive effects on social and ecological sustainability while including the economic aspects of the transportation system. The aim of the evaluation was to be able to make a fundamental statement regarding the usability of the crowd logistics approach and the sustainable Customer Relationship Management approach and, if necessary, to identify optimization potential and concrete suggestions for improvement. Another objective of the evaluation was to determine the needs of different participants such as retailers and shippers as well as of receivers and deliverers while considering sustainable goals. Various data collection and analysis methods were adopted to evaluate the project objectives in quantitative and qualitative ways. The evaluation audited the approach's ability to have positive impacts on different dimensions of sustainability and showed further possibilities for improvement.

Keywords Sustainability · Electromobility · Last mile · Crowd logistics · SusCRM · Cargo bikes

1 Introduction

The "NaCl—Nachhaltige Crowdlogistik (engl.: 'sustainable crowd logistics')" project pursues the development of an innovative sustainable logistics concept based on electric cargo bikes and a crowd logistics approach. Project participants are the

R. Schulte (✉) · M. Leibenath · L. Wöltjen · U. Kühne · B. Wagner vom Berg
Hochschule Bremerhaven, An der Karlstadt 8, 27568 Bremerhaven, Germany
e-mail: rschulte@hs-bremerhaven.de

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_13

Bremerhaven University of Applied Sciences, the industrial manufacturer of cargo bikes Rytle GmbH and the local transport service provider Weser Eilboten GmbH (WEB). In close cooperation, the aim is to design a sustainable logistics system for the last mile. The system consists mainly of electric cargo bikes and a digital infrastructure. The basic objectives of the project are the positive effects on the ecological and social dimensions of sustainability, while considering the economic aspects. Other participants in the project are the crowdworkers, the B2B (Business-To-Business) and B2C (Business-To-Consumers) customers respectively the retailers and recipients.

In order to test the effectiveness of the project-related measures, a three-month pilot phase was planned in a test field close to the city center in Bremerhaven. As part of the piloting, an evaluation of the logistical concept and the project objectives was carried out. The evaluation realized within the framework of a bachelor thesis includes a systematic analysis and evaluation of the project, project-related activities, and the pilot phase, considering the interests of stakeholders. The evaluation is based on an empirical collection of project-related data and information. The results of the evaluation should help to provide information on the achievement of the objectives and to identify weaknesses of the NaCl concept as well as to contribute to the further development and optimization of the concept.

The evaluation procedure first requires the definition of the project-related objectives of the NaCl approach. Subsequently, the evaluation objects of the project are determined. A concept is developed from the underlying objectives and subjects of the evaluation, which includes a selection of appropriate methods for data collection and analysis. After the collection and analysis of the corresponding data, these are interpreted and evaluated. The findings must then be documented and reported to the project partners. Finally, the findings can be used for optimization and as a basis for decision-making.

2 State of Art

2.1 Fundamentals

Sustainability. In the context of sustainability, there has been increasing talk of "sustainable development" since the Brundtland report in 1987. Sustainable development is described there as a process that guarantees that today's generations can meet their needs without limiting the opportunities for future generations to meet theirs. Since then, sustainability has been seen as a three-pillar model that combines the three dimensions of ecology, sociology, and economics, with the core idea often being the equal consideration of the three dimensions, while other approaches go a step further and place ecology above sociology and this in turn above economics [1].

The ecological dimension of sustainability includes efficient and ongoing resource management to ensure that the rate of degradation does not exceed the rate of regeneration of renewable natural resources. At the same time, we must prevent the extraction of non-renewable resources, or use non-renewable resources only to the extent that equivalent compensation can be created in renewable raw materials. Nature's regulatory and supportive functions must also be preserved, which means that emissions into the environment must be adapted to nature's material absorption capacity [2].

The agenda 21 and the Brundtland report describe social sustainability based on intergenerational and intragenerational justice, i.e., that future generations have the same opportunities as today and that all people living on Earth have the same opportunities [1, 3]. Social sustainability is closely linked to the ecological dimension, since a balance between human needs and the potentials and resources of nature must be sought [4].

According to Heins, social acceptance for sustainable development, social and health protection, social stability and a fair distribution of wealth and equal opportunities are necessary to solve these problems [5]. According to Zimmermann, another crucial level is participation in the form of involvement and co-decision-making of all social classes [4].

The economic dimension of sustainability combines the principles of welfare economics and sustainability. Sustainable development is directly influenced by the way people operate and from the management of natural energy and material resources. Material inputs into the environment that arise in the economic process, such as waste and emissions, illustrate the direct relationship between the ecological and economic dimension [4]. Intelligent management of the economy can reduce entries into nature, and it is necessary to adapt these entries to the absorption capacity of the environment. A look at the distribution problem, which also deals with the possibilities of basic services for all people, shows that the economic dimension is also strongly linked to social sustainability.

Crowd Logistics. Crowd logistics is an approach that is currently increasingly found in specialist literature and is being used in practice to meet the difficulties on the last mile. The term "crowd logistics" can be traced back to the term "crowdsourcing", which was first named by Howe in Wired Magazine [6]. Crowdsourcing, for its part, is a word composition of the words "crowd", which describes a defined mass of people—in the sense of crowd logistics: potential delivery drivers—and "outsourcing", which means the transfer of processes and functions—such as delivery services—to third parties [7]. Crowdsourcing is often described in literature in relation to a digital, internet-based way of working [8, 9].

According to Leimeister and Zogaj, the multitude of existing definitions around the term crowdsourcing underlies the following three elements: the client, the crowd and the process. The client or crowdsourcer and simultaneous initiator starts the process of crowdsourcing by outsourcing internal tasks, e.g. logistical activities such as parcel delivery, to contractors or the crowd. The crowd describes a previously defined or undefined large mass of people or so-called crowdworkers, whereby the size of the crowd depends on parameters of the task to be completed [8, 10].

Blohm et al. state to supplement by a fourth main component regarding internet-based crowdsourcing: the crowdsourcing platform, which is provided internally or through an external intermediary [7]. The platform acts as an interface between the crowdworkers and the crowdsourcer by providing the tasks that can be accepted by the crowdworkers and then be executed.

Electric Cargo Bikes. Cargo bikes are bicycles designed to transport loads and is basically not a novelty, but new designs and innovations in battery technology have ensured a revival of these means of transport.

As a result, cargo bikes are now found not only as manually driven vehicles, but also as electrically assisted vehicles, referred to as electric cargo bikes or e-cargo bikes. E-cargo bikes make use of the electric drive technology, so that electric cargo bikes can be assigned to the electromobility sector. According to Scheurenbrand et al., electromobility is a highly networked industry and aims to meet mobility needs with the inclusion of sustainable aspects. The vehicles will be equipped with portable energy sources such as batteries and an electric drive [11].

Growing environmental awareness and new developments in modern mobility have led to new perceptions of the benefits of the internal combustion engine. The fact that fossil fuels such as oil are finite and that mobility-related environmental pressures have increased and continue to increase has led to a rethink of the dominant drive. Guidelines on sustainable development and an increasing degree of urbanization promote technological change from the internal combustion engine to electric driving concepts [12].

2.2 Challenges and Opportunities

The Last Mile is currently described as one of the most expensive, less efficient and most polluting sections of the supply chain [13]. Up to 75% [13, 14] of the total logistics costs are attributed to the last part of the supply chain, that can be connected to the high inefficiencies and poor environmental performance due to e.g. high stop densities and low stop factors or the absence of the recipients [15].

An ever-increasing volume of transport and the associated traffic volume [16], especially in inner cities, call for urgent action to relieve society, the economy and, above all, the environment. Consequences of increased traffic include congestion, noise, greenhouse gas and pollutant emissions, as well as safety risks such as accidents. It is therefore not surprising that the last mile has negative effects on the environment and quality of life of people. According to a study by the World Economic Forum, developments on the last mile will lead to increases in CO_2 emissions from delivery traffic of 32 percent and around 21 percent more traffic obstructions until 2030 [17].

Another major problem of urban logistics is the conflict that consumers want a fast and flexible delivery of their goods, especially in online retail, while at the same time low or no delivery costs are required [18]. This discrepancy, together with the difficult to calculate order structure and the atomization of shipment sizes, have a significant impact on the shipping costs of courier, express and parcel service providers (CEP service providers), as well as the heterogeneity of transport packaging and the handling of goods, which is determined by the type of goods such as products to be cooled [19].

Overcoming the Last Mile presents logistics service providers with many complex environmental and social challenges but offers several opportunities to generate competitive advantages.

Compared to 1990, a 40% reduction in greenhouse gas emissions in Germany should be achieved by 2020. In 2018, the reduction compared to 1990 was 27.5%. According to a recent report by the Federal Environment Agency, with an absolute value of 163 million tons of CO_2 equivalent, 20.2% of greenhouse gas emissions in 2019 were in the transport sector and 254 million tons of CO_2 equivalent accounted for 31.5% in the energy sector [20].

Between 1995 and 2018, CO_2 emissions from road freight transport increased by 22% [21]. Despite improvements in efficiency in transport, there has been little success in reducing greenhouse gas emissions, which can be explained by the increase in traffic volumes and transport services, especially in freight and transit transport [22]. The volume of CEP shipments in Germany has increased by 108% from 1.69 billion in 2000 to 3.52 billion in 2018 and the trend is increasing [16]. A comparison of the trends in the volume of shipments and of goods carried by road transport shows an unequal increase in the two key figures and suggests that a trend towards lighter, but more shipments in terms of quantities has developed. This trend, coupled with emissions developments, poses major challenges and changes in the transport sector. Especially on the last mile, these tendencies are leading to major issues for the CEP sector and solutions must be found to address this situation. At the same time, opportunities arise to distinguish themselves from competition and generate competitive advantages. In connection with the increasing environmental impact, drive technologies based on electromobility are exceedingly popular.

In a study by the Fraunhofer Institute on climate balances of different drive technologies is stated, that electric cars have clearly the best climate balance for both 2019 and forecasted for 2030 [23]. The study used Well-to-Wheel analysis so that the entire impact chain was considered. It can also be observed that cargo bikes are used in the logistics sector with an increasing tendency. A study by the "Ich entlaste Städte (engl.: 'I relieve cities')" initiative shows that companies tend to use cargo bikes as an alternative to delivery vehicles with internal combustion engines on the last mile. The initiative lends cargo bikes to companies for one euro per day for a maximum period of three months. A survey of this activity showed that after the test, one in five users bought a cargo bike or tended to buy one [24]. According to a study of the "Cycle-Logistics "-project, cargo bikes have the potential to replace 50% of all urban transport-related journeys and have a positive impact on air quality and quality of life as well as safety in the city [25].

Sustainability in customer relationship management is understood by the abbreviation SusCRM that is a composition of the word "sustainable" and the abbreviation CRM, that stands for customer relationship management. SusCRM is a novel approach to conveying sustainability in the sense of all three dimensions within companies, in the customer relationship and to third parties. The aim is to increase participation in sustainable processes and to attract a loyal clientele for a sustainable business model [26].

An idea for shortening delivery routes per shipment and thus saving emissions on the last mile is bundling and can be designed differently. One approach is the bundling of different products and different service providers, as well as pick-ups and deliveries. This can achieve positive environmental impacts through better use of cargo capacity and the reduction of transport routes. Involving local retailers can increase their competitiveness over online retail, making a positive contribution to social sustainability [26].

Modern logistics concepts use approaches to crowd logistics to achieve more flexibility and withstand the increasing cost pressure. Companies such as Uber Freight and Delivery Hero are well-known examples that have already successfully applied this new way of working and have seen great growth. In addition to the advantages of crowd logistics, such as flexibility, efficiency improvements and cost reductions, many problems are also related to these concepts [27].

On the social dimension of sustainability, challenges related to the privacy of crowdworkers can be identified, as they share sensitive data such as their (live) location. Mechanisms must be found to gain the confidence of the crowdworkers at the same time, to create transparency and to maintain privacy through appropriate framework conditions [27].

While crowdworkers are offered a flexible working environment, this approach is associated with risks in terms of working conditions, social security, as well as income and job security [28]. Furthermore, serious changes in the traditional logistics industry can be identified, as numerous startups build logistics concepts based on low-cost structures and on-demand platforms and thus take the market shares of the long-established logistics providers. Crowdworkers are employed when a need arises, creating the potential for efficiency gains for costs and time, but this process also offers the possibility of replacing permanent employees [27].

Advancing digitization and automation in the field of crowd logistics can be supportive, but also disruptive, replacing workers in their working environment [28]. Many novel approaches in the field of the last mile offer opportunities to meet the new challenges. Crowd logistics can help to make better use of transport capacities and to make logistics systems on the last mile more flexible and efficient. Electromobility has the potential to make the supply chain more environmentally friendly. However, the use of these technological instruments must not obscure social sustainability.

3 Project NaCl

The research project NaCl takes up the challenges and opportunities within the last mile and develops a logistical solution system for the problems in delivery logistics on the last mile. The basic idea of the logistics concept described here is the development and piloted application of a novel and sustainable logistics concept for optimizing the last mile. The aim is to generate positive effects on the three dimensions of sustainability.

The sustainable logistics concept is essentially based on the use of electric cargo bikes, a crowd logistics approach and sustainable customer relationship management (SusCRM). The implementation of electric cargo bikes should above all have positive impacts on environmental sustainability and the regional transport system, as electric cargo bikes are locally emission-free and traffic-friendly. The use of this mode of transport compared to vehicles based on internal combustion engines, which are the current standard in the CEP sector, has further potentials, e.g. for the economic dimension in personnel flexibility, since the driving of electrically driven cargo bikes does not require a driving license. Further elasticity is to be achieved through a crowdsourcing approach, in which a crowdworker pool consisting of students from the Bremerhaven University of Applied Sciences is hired by the local logistics service provider Weser Eilboten GmbH while the delivery is taken over with the help of the cargo bikes of the Rytle GmbH. The aim is to provide crowdworkers with a flexible and voluntary working relationship. The use of crowdworkers is intended to offer potentials for the relief of the traditional delivery staff and not to replace them. Student crowdworkers are to be used at times of peak loads to avoid overtime for regular drivers, to create regular working hours and to reduce work stress.

A bundling of transports in relation to the goods delivered as well as to different contracting entities was carried out to achieve cost reductions, productivity increases and further environmental relief. Empty journeys and incompletely filled boxes should be avoided and collective deliveries of various products and customers should be implemented. The inclusion of pick-up and delivery orders should be dynamic and flexibly plannable into existing tours.

The project uses a SusCRM approach to sensitize recipients and companies, as well as crowdworkers, to the use and participation in sustainable crowd logistics. This approach is implemented through the provision of information on delivery methods and sustainability.

Various project partners are involved in the presented research project. This is on the one hand the Bremerhaven University of Applied Sciences, which is under the management of the project, Rytle GmbH, which provides the digital and technical infrastructure, and Weser Eilboten GmbH, which acts as a transport service provider. Other actors are the crowdworkers, who are under contract with the Weser Eilboten GmbH and handle the operation of the cargo bikes and the delivery, as well as B2C and B2B customers that act as shippers and recipients.

4 Evaluation of the NaCl Pilot Phase

4.1 Evaluation Planning and Design

At the beginning of the evaluation, the evaluator determined the evaluation objectives and objects of the logistics concept with the involvement of the project partners. These objectives have been broken down into a target hierarchy. The objective of positive effects on the sustainability of the last mile was identified as the overall guiding objective. The main objectives were the positive effects on the environmental, social and economic dimensions of sustainability.

Based on the evaluation objects and objectives, the participants and stakeholders of the evaluation were identified. These actors included, on the one hand, the NaCl project partners and, on the other hand, the target groups of the evaluation objects and objectives such as end consumers or recipients, shippers (stationary retail), crowd-workers and the project partners. Criteria and indicators have been defined in relation to the individual evaluation objectives to measure the achievement of the objectives set. The criteria and indicators provided the basis for deciding which data collection methods are used to study the objects of evaluation. Stakeholders and participants acted as data providing of the survey methods, while the cargo bikes telematic systems in conjunction with Rytle's delivery data system acted as data providing of the quantitative data collection.

Table 1 shows the different evaluation objects, their associated objectives and the data collecting and analysis methods that have been applied to them. The data collection methods included different methods of qualitative and quantitative research. Qualitative methods, such as surveys and the focus group were used to determine

Table 1 Evaluation design of the NaCl project

Object	Objective	Data collecting	Data analysis
Crowd logistics	Fair working conditions	Survey	Qualitative content analysis, descriptive statistics
Pickup and delivery	Strengthening of regional Retail	Survey, focus group, counting (delivery data)	Qualitative content analysis, descriptive statistics
Bundling	Reduction of transport ways and stops	Counting, measurement (delivery data)	Descriptive statistics, inductive statistics
E-Cargo-bikes	Reduction of noise and GHG emissions	Measurement (delivery data), literature review	Qualitative content analysis, descriptive statistics, Well-To-Wheel-Analysis
SusCRM	Participation in sustainable logistics	Survey, focus group	Qualitative content analysis, descriptive statistics

subjective assessments of affected and involved stakeholders. Quantitative methods, such as counting and measurement, were used to collect numerical data from the evaluation. After extensive preparation of the collected data, these were examined using various methods of analysis corresponding to the purpose. The data collected from the surveys were evaluated using qualitative content analyses and the application of descriptive statistics. During the counts and measurements, delivery data such as time stamp and parcel number of deliveries as well as process errors in relation to the project's resources were collected. The analysis of the delivery data was carried out by descriptive statistical evaluations. Delivery data were classified into numbers of delivery tours, stops and packages to build a basis for further analysis. As a result, process times and key figures were determined for the assessment of the evaluation objects. This data was used in the further procedure to calculate measures of the achievement of the target. For the determination of e.g. caused emissions or costs during the pilot phase, further analyses were carried out.

4.2 Evaluation Objects and Objectives

The Evaluation objects and linked objectives of the NaCl project are addressing different dimensions of sustainability. Figure 1 shows the allocation of the various objects to the three dimensions.

Electric Cargo Bikes. The use of electric cargo bikes should have a positive impact on the environmental and social dimension of sustainability. In this context the aim was to reduce exhaust and noise emissions and to reduce congestion in the city center. In addition, the goal of practicality in the delivery process is to be achieved in a process-related manner and thus a good working environment for the crowdworkers. These effects were tested by comparing the noise and CO_2-emissions produced using cargo bikes and conventional diesel vehicles. The potential to replace conventional diesel vehicles should be identified. The relief of inner-city traffic should be achieved and investigated through the use of electric cargo bikes.

Crowd Logistics Approach. The crowd logistics approach is linked to the objectives of improving the working conditions of delivery drivers as well as fair and flexible working conditions for the crowdworkers. Positive impacts of the approach are contributed to the economic and social dimension of sustainability. The improvement in the working conditions could not be determined due to a lack of data. Also, the shortness of the pilot phase did not provide a significant assessment. Instead, a potential of the delivery concept was identified to contribute to the relief of delivery drivers. The potential was determined by the limits of the delivery concept, which is characterized by the cargo bikes and the crowdworking approach. For the two cargo bikes used, MovR and Triliner from Rytle GmbH, the limitations were determined via the maximum payload, the load volume, the maximum package sizes that can be stored and the ranges of the cargo bikes. It must be said that a conventional diesel delivery vehicle exceeds all the limits of electric cargo bikes mentioned above. The

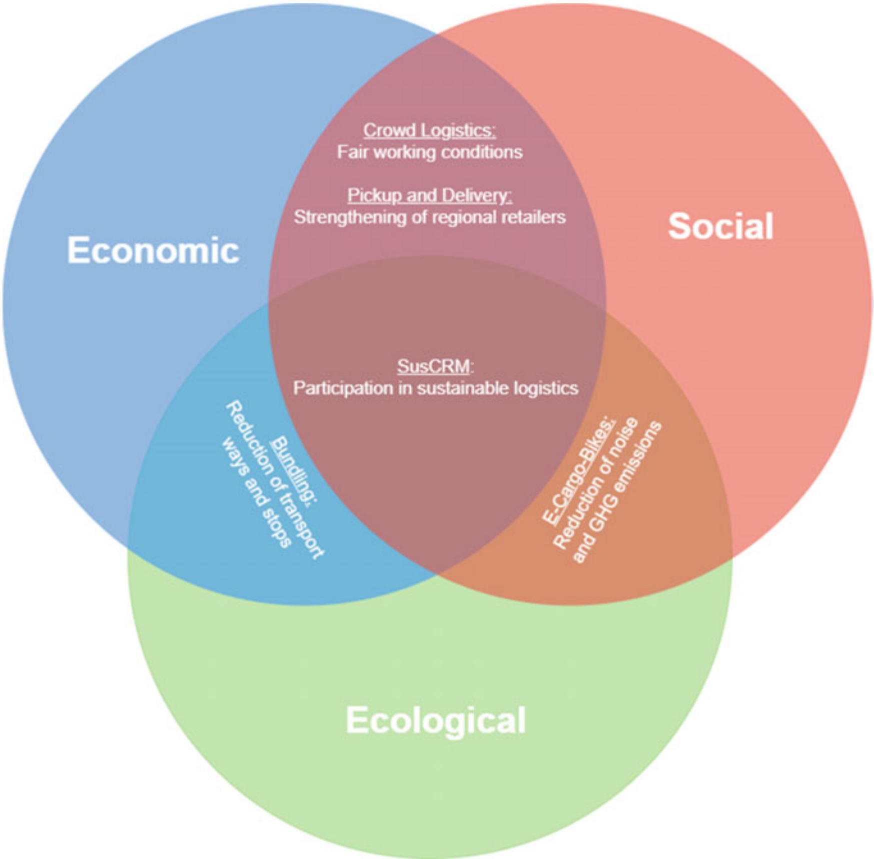


Fig. 1 Evaluation objectives and objects in conclusion under allocation of the dimensions of sustainability

approach studied is primarily geared towards relieving employees during peak times in terms of social sustainability, without replacing them, and is therefore specifically geared towards flexibility. Going further, the potential is limited by the number of available cargo bikes, the size of the crowdworker pool, and the availability of crowdworkers.

SusCRM Approach. The SusCRM was intended to promote participation in the sustainable logistics solution at various levels, such as the B2C customers, the B2B customers, but also of the crowdworkers hired. The SusCRM approach addresses all three dimensions of sustainability.

The study of the SusCRM of the NaCl project has provided insights into the interest and needs of B2B and B2C customers as well as the crowdworker for participation in a sustainable delivery solution for the last mile.

Bundling. The evaluation object bundling addresses the economic and ecological dimension of sustainability and is linked to the objectives of reducing exhaust and noise emissions as well as the relief of inner cities. The process-related goal is the successful bundling of supply orders from different customers.

The bundling of consignments and contracting entities should help to reduce exhaust and noise emissions and ease the burden on inner cities by shortening delivery routes and reducing the number of delivery stops. During the test phase, the bundling measure was not implemented to the planned extent, so it was not possible to assess it according to the criteria and indicators set out in the evaluation design. During the three-month pilot phase, however, bundling was carried out regarding orders from WEB (local logistics service provider) and retailers from the city center.

Pickup and Delivery. Another evaluation object of the project NaCl was the implementation of pickup and deliveries oof the stationary retail. In addition to the logistics partner's deliveries, interested parties from the local retail trade were offered the opportunity to commission pickups and deliveries via a mobile customer app, which were either bundled by the crowdworkers in an existing tour or executed separately. The project NaCl guaranteed same day delivery until 14:00 [29].

4.3 Evaluation Results

Electric Cargo Bikes. The evaluation shows that the use of electric cargo bikes on the last mile has positive effects through CO_2-savings compared to diesel vehicles. With the help of the Well-to- Wheel analysis, the CO_2 emissions generated using the two cargo bikes Triliner and MovR of Rytle GmbH and the diesel-powered delivery vehicles Fiat Doblo and Ducato were determined and subsequently compared. According to the DIN EN 16258 standard [30], a Well-To-Wheel (WTW) analysis is a summary of the Well-To-Tank (WTT) and Tank-To-Wheel (TTW) or indirect and direct emissions. The WTT analysis includes the indirect emissions of the fuel supply from the source or extraction of the raw materials to the vehicle tank or energy storage. The TTW analysis includes all direct emissions from vehicle operation, e.g., from the combustion of diesel fuel. The emission values of the diesel vehicles were 15 to 28 times higher than those of the electric delivery bikes [29]. Furthermore, the emission values determined for the vehicles were used to calculate the emissions for sample tours from the NaCl pilot phase. The tours were carried out with the cargo bikes to determine the level of emissions that would have been generated if the diesel vehicles had been used and the level of emissions that could be saved by using the electric cargo bikes. The determined values show that significant savings in CO_2 emissions can be achieved by using electric cargo bikes. The emission values per kilometer of the electric bikes are significantly lower than the values of the diesel transporters. In addition, apart from one studied tour, the electric bikes required a shorter route to complete the tours. This results from the route planning, which determined the fastest route for the respective means of transport in each case, whereby the highway was often calculated as a partial route for the transporters [29].

The study of the use of the electric cargo bikes showed that these vehicles are suitable as a means of working the last mile. The evaluation of the cargo bikes by the surveyed crowdworkers, as users of these work tools, is positive. The results of the Well-To-Wheel analysis show a positive CO_2 balance in the operation of the cargo bikes compared to the diesel vans. Already on one kilometer, up to 270 g of the greenhouse gas can be saved [29]. It should be noted that although the determination of CO_2 emissions via a Well-to-Wheel analysis takes a comprehensive look at the emissions caused by the provision and consumption of the drive energies, it does not consider the emissions generated during the production of the vehicles and upstream processes. It is also important to note that a last-mile emissions comparison should always be calculated on a per-package basis to provide a basis for comparison that can be evaluated. At this point, it should be mentioned that the comparison in this paper was made for last mile tours that were carried out with a cargo bike during the pilot phase. A van has much higher loading capacities and can therefore deliver more parcels on one tour, which reduces the emission values per parcel. Nevertheless, the environmental balance in terms of CO_2 emissions per kilometer is clearly in favor of the electric cargo bikes. Positive effects on the ecological dimension of sustainability can thus be identified [29].

The determination of noise emissions is based on a literature search. To determine the specific noise emissions precisely, measurements should have been made which were omitted due to lack of time and resources. In the case of cars, the drive noise up to 25 km/h is the largest source of noise [31]. This corresponds approximately to the average speed at which transporters move in parcel delivery [15]. According to studies, electric vehicles have significantly lower noise volumes in the low speeds range than e.g. diesel vehicles, but this advantage is no longer represented at higher speeds, as the noise source shifts from the drive to the unrolling and wind resistance of the vehicles [32]. Rolling noises of the vehicles depend on the type of vehicle and the vehicle construction, the road surface and the speed, wind noise due to the shape and size of the vehicle and the speed. Here it must be assumed that the cargo bikes cause less noise due to their smaller size and the lower maximum speed of 25 km per hour. Since the cargo bikes and courier vehicles usually operate in a speed range of less than 30 km/h, it is assumed that the noise emissions of the cargo bikes by the electric drive are lower for this area of application.

In a survey of crowdworkers after completion of the test phase, 100% of the respondents stated that they believe the delivery concept can contribute to the relief of inner cities. These answers were justified by the size of the cargo bikes, which is smaller than that of vans, so that they could be better overtaken by other road users and that cargo bikes could use paths that would not be passable by cars. It has also been mentioned that volume and CO_2-emissions would be reduced. The flexibility of the bikes has been mentioned several times. [29]

In addition, advantages such as the possibility to maneuver in narrow streets and pedestrian only zones due to cargo bikes' smaller size, the possibility to park on sideways and therefor save time and money as well as more safety due to less severe collision in comparison to delivery trucks can be mentioned [33].

Crowd Logistics Approach. During the NaCl pilot phase, one MovR and two Triliners were in use. The worker pool consisted of eight active drivers. There is a further restriction due to the working time limit of the student crowdworkers, who work on a 450€ basis. The maximum working time per month for them is 40 h [29].

In a survey, during the pilot phase, it was investigated on which days of the week and at which times the student crowdworkers are free to work. It was found that the availability of the crowdworkers apart from one day in a period of three hours was always sufficient to manage two tours in parallel. It was also found that for much of the time, competition between drivers can arise when the pool of available drivers exceeds the pool of cargo bikes and thus parallel tours [29].

Availability was further determined by the time that elapsed between order request and order acceptance. Based on the delivery data from the pilot phase, these times were determined as an average of 55 min for the individual packages and 65 min for the tours. The time between order acceptance and parcel pickup was 132 min for the individual parcels and 129 min for the tours [29].

To measure the target of fair and flexible working conditions for the crowdworkers, they were questioned during different times of the pilot phase. More than 80% of the crowdworkers rated the net wage of €11.13 as appropriate for the job. More than 70% of the crowdworkers also described crowdworking as a good side job. The reasons for this were that the activity takes place in the fresh air and is flexible. Nearly 60% of crowdworkers would do the job again. The evaluation of typical crowd logistics characteristics, such as flexibility, self-responsibility, voluntariness, compensation, training, communication, and participation were largely rated positively by the crowdworkers. The goal of providing fair and flexible working conditions for crowdworkers has been achieved [29].

SusCRM Approach. It has been noted that there is a fundamental interest on all sides for sustainable parcel delivery. Most respondents to B2C customers—both those identified as recipients and potential customers—have shown a clear tendency to decide for a sustainable mode of transport when they choose. There has also been a signal of a willingness to pay more for such a form of transport. More than 50% of those who showed this willingness would agree to pay between €1 and €2.50 more per shipment. Overall, support for the relevance of environmental challenges and sustainable action was high among B2C customers. They clearly showed understanding for the importance of environmental issues and improvement measures [29].

The survey of the crowdworkers showed that for a large part of the crowdworkers the motivation for the project participation was from the environmental issue. Overall, the weighting and approval of environmental issues was rather high in the applied surveys [29].

B2B customers have also shown an interest in the sustainable concept. Some of the retailers cited environmental and innovation aspects as well as regionality as the main motivation for participating in the project. They also considered the importance of changes in purchasing behavior to reduce the burden on the environment. During the acquisition, a great deal of interest and enthusiasm of the dealers for the sustainable orientation of the project could be noted [29].

Bundling. The analysis of the delivery data showed that out of the total of tours carried out for regional retailers in a quota of 50% successful bundling could be applied. The remaining orders of retailers were handled in separate tours. Moreover, it could be stated that savings in delivery routes were achieved with every bundled delivery tour. For this scenario, it was assumed that the orders would be divided into individual tours for each client. In order to determine the bundling effect, the kilometers driven by the individual tours were then subtracted from the kilometers driven by the bundled tour [29].

The reduction in delivery stops could not be compared due to non-existent data from the WEB regarding stops per tour and packages per stop. Bundling, which leads to the reduction of delivery stops, requires that orders from different customers (logistics service provider and retail) are delivered in one tour to the same customer or customers who live in the same house or proximity. This fact could not be measured within the evaluation. The probability of this situation occurred during the pilot phase is extremely low, as it requires a large volume of deliveries. To reduce delivery stops by bundling, a check of delivery addresses must be carried out when orders are created in order to determine that the addresses of existing orders and new orders match. In the case of same-day deliveries, the merging of the same delivery addresses into a tour is made even more difficult since the order acceptance and execution by the crowdworker takes place at short notice and here the probability of a match between two delivery addresses with the accepting crowdworker is low. The bundling of the same delivery addresses therefore requires a large delivery volume and the involvement of many customers for short-term implementation. Also, a corresponding order may only be sent to the crowdworker whose existing tour contains the same delivery address, in order to avoid another driver accepting the tour. Long-term planning of delivery orders makes it easier to bundle the same delivery addresses, but this minimizes the likelihood of same day deliveries [29].

Pickup and delivery. For the retail trade, pick-ups of goods and deliveries to customers were to be realized, also in the form of same day deliveries. The aim is to strengthen brick-and-mortar retail. The verification of the targets was verified by comparing the requested pickups and the deliveries made. Same day deliveries were measured by the variance of order date and delivery date [29].

In principle, the goal of the realization of pickups and deliveries for the retail trade was achieved and the promise of delivery on the same day could be complied with when ordering by 14:00. Provided that failed delivery attempts due to absence of the recipients were not considered, the same-day delivery could be fulfilled in any case during the pilot phase. Pickup and delivery, however, requires process design to be as simple as possible to keep the implementation effort down for the participating dealers and thus to promote participation. It turned out that the integration into a trading platform with corresponding functions for marketing, billing, etc. as well as an automation of the underlying logistics processes has a high potential for a more successful marketing of the delivery offers as well as the participation of the dealers. [29]

5 Conclusion and Future Outlook

In summary, it can be stated that the crowd logistics approach presented, and the logistics system used, which is based on electric cargo bikes, address all three dimensions of sustainability. Potentials for improvement could be observed on the ecological as well as on the social level. A compatibility of these potentials with the economy of the logistics concept can be assumed but requires further adjustments of the system by adapting to a mode of operation based on electric cargo bikes.

The reduction of emissions in comparison to traditional delivery vehicles is achieved by using light electric vehicles such as electric cargo bikes and micro depots, while from an economic point of view this combination can be more economical than conventional delivery methods [34]. As stated, it is possible to map almost the entire city logistics with such a system. The area of use of the electric cargo bikes plays a decisive role for the economic efficiency of the logistics concept and should be optimized in disposition and route planning.

Significantly more efficient and environmentally friendly logistics can be achieved by bundling transports, in terms of the products delivered and the logistics service providers commissioning them. Furthermore, delivery and pick up are bundled in one tour. The implementation of the idea of a singular regional logistics service provider achieves maximum bundling effects since optimization can be achieved across the entire flow of goods.

The logistics concept drives the strengthening of regional retail through regional logistics service providers. It offers a possibility of same day delivery for stationary retail. Regional retail can be reinforced against the e-commerce companies and gains competitiveness by combining the advantages of stationary retail such as proximity to customers and the advantages of e-commerce such as convenience of “shopping on the couch”. Combined with a regional e-commerce platform and a SusCRM approach, retail can potentially regain significant market share. Regional added value can be promoted in a targeted manner and transport distances can be shortened enormously. In the context of a regional retail platform the creation of a critical mass of participants from stationary retail and end consumers is of great importance. The research project "R3 - Resilient, Regional, Retail in der Metropolregion Nordwest" of the Bremerhaven University of Applied Sciences addresses this challenge. It has started in June 2021.

By addressing the social dimension within SusCRM, crowdworkers, end consumers and shippers can be motivated to participate in a sustainable supply chain e.g., by using a more sustainable delivery method and supporting a fair labor organization.

Potential goals of improving the working conditions of permanent employees through flexibly deployable crowdworkers can be achieved. This reduces fluctuation as well as associated costs and establishes long-term employment relationships with the regional logistics service provider. Efficiency gains associated with this and bundling, may enable higher wage levels and can help put an end to the current, sometimes precarious, employment relationships in the industry. For a successful implementation of a crowd logistics approach the limitations of the applied system must be considered.

Through the promotion of political restructuring processes, such an approach should be supported e.g., this includes instruments such as taxes on emissions, infrastructure realignments such as more bike lanes and less road space, the establishment of car restricted zones in inner cities as well as the provision of space in inner-city areas for micro-hubs, cargo bikes, and charging stations.

Many aspects can be identified and further developed based on the concept evaluated, that has the potential to make a major contribution to a more sustainable economy and society while relieving the burdens on environment and climate.

Acknowledgements The NaCl project is funded by EFRE and the state of Bremen within the program "Applied Environmental Research (AUF)" of the city of Bremen.

References

1. Hauff, V.: Unsere gemeinsame Zukunft. [der Brundtland-Bericht] (1987)
2. Rogall, H.: Ökologische Ökonomie. Eine Einführung, 2nd edn. VS-Verl. für Sozialwiss, Wiesbaden (2008)
3. Agenda 21 (Übersetzung). Konferenz der Vereinten Nationen für Umwelt und Entwicklung, Bonn. https://www.bmu.de/fileadmin/bmu-import/files/pdfs/allgemein/application/pdf/agenda21.pdf (1997). Accessed 23 Oct 2020
4. Zimmermann, F.M.: Nachhaltigkeit wofür? Springer, Berlin Heidelberg, Berlin, Heidelberg (2016)
5. Heins, B.: Soziale Nachhaltigkeit, 1st edn. Analytica, Berlin (1998)
6. Howe, J.: The Rise of Crowdsourcing. Wired Magazine (2006)
7. Blohm, I., Jan Marco, L., Zogaj, S.: Crowdsourcing und Crowd Work—ein Zukunftsmodell der IT-gestützten Arbeitsorganisation? In: Brenner, W., Hess, T. (eds.) Wirtschaftsinformatik in Wissenschaft und Praxis. Business Engineering, pp. 51–64. Springer Berlin Heidelberg, Berlin, Heidelberg (2014)
8. Leimeister, J.M., Zogaj, S.: Neue Arbeitsorganisation durch Crowdsourcing. Arbeitspapier Nr. 287. Mitbestimmungs-, Forschungs- und Studienförderungswerk des DGB, Düsseldorf. https://www.boeckler.de/pdf/p_arbp_287.pdf (2013). Accessed 5 Nov 2020
9. Vukovic, M., Lopez, M., Laredo, J.: People Cloud for the Globally Integrated Enterprise. https://www.academia.edu/42952170/PeopleCloud_for_the_Globally_Integrated_Enterprise. Accessed 6 Nov 2020
10. Estellés-Arolas, E., González-Ladrón-de-Guevara, F.: Towards an integrated crowdsourcing definition. J. Inf. Sci. (2012). https://doi.org/10.1177/0165551512437638
11. Scheurenbrand, J., Engel, C., Peters, F., Kühl, N.: Holistically Defining E-Mobility: A Modern Approach to Systematic Literature Reviews. Karlsruhe Institute of Technology (2016)

12. Hüttl, R.F., Pischetsrieder, B., Spath, D.: Elektromobilität. Springer, Berlin Heidelberg, Berlin, Heidelberg (2010)
13. Gevaers, R., van Voorde, E. de, Vanelslander, T.: Characteristics and Typology of Last-Mile Logistics from an Innovation Perspective in an Urban Context. Chapters (2011)
14. Voß, P.H.: Logistik—eine Industrie, die (sich) bewegt. Springer Fachmedien Wiesbaden, Wiesbaden (2015)
15. Brabänder, C.: Die Letzte Meile. Springer Fachmedien Wiesbaden, Wiesbaden (2020)
16. Esser, K., Kurte, J.: KEP-Studie 2020. Analyse des Marktes in Deutschland. KE-CONSULT Kurte&Esser GbR, Köln. https://www.biek.de/files/biek/downloads/papiere/BIEK_KEP-Studie_2020.pdf (2020). Accessed 6 Nov 2020
17. World Economic Forum: The Future of the Last-Mile Ecosystem. Cologny, Geneva. http://www3.weforum.org/docs/WEF_Future_of_the_last_mile_ecosystem.pdf (2020). Accessed 1 June 2021
18. Allen, J., Thorne, G., Browne, M.: Good Practice Guide on Urban Freight Transport. European Comission DG Transport and Energy, Rijswijk. http://www.bestufs.net/download/BESTUFS_II/good_practice/English_BESTUFS_Guide.pdf (2008). Accessed 6 Nov 2020
19. Schnedlitz, P., Lienbacher, E., Waldegg-Lindl, B., Waldegg-Lindl, M.: Last Mile: Die letzten—und teuersten—Meter zum Kunden im B2C ECommerce. In: Crockford, G., Ritschel, F., Schmieder, U.-M. (eds.) Handel in Theorie und Praxis, pp. 249–273. Springer Fachmedien Wiesbaden, Wiesbaden (2013)
20. Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU): Klimaschutz in Zahlen: Fakten, Trends und Impulse deutscher Klimapolitik Ausgabe 2020, Berlin (2020)
21. Umweltbundesamt: Emissionen des Verkehrs. https://www.umweltbundesamt.de/daten/verkehr/emissionen-des-verkehrs#pkw-fahren-heute-klima-und-umweltvertraglicher (2020). Accessed 10 Nov 2020
22. Statistisches Bundesamt: Beförderungsmenge und Beförderungsleistung nach Verkehrsträgern. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Transport-Verkehr/Gueterverkehr/Tabellen/gueterbefoerderung-lr.html (2020). Accessed 9 Nov 2020
23. Martin Wietschel et al.: Klimabilanz, Kosten und Potenziale verschiedener Kraftstoffarten und Antriebssysteme für Pkw und Lkw. Fraunhofer-Institut für System- und Innovationsforschung, Karlsruhe. https://www.isi.fraunhofer.de/content/dam/isi/dokumente/cce/2019/klimabilanz-kosten-potenziale-antriebe-pkw-lkw.pdf (2019). Accessed 11 Nov 2020
24. DLR: DLR-Forschungsprojekt “Ich entlaste Städte” präsentiert Zwischenbilanz mit ersten Ergebnissen. https://www.dlr.de/content/de/artikel/news/2018/4/20181126_projekt-ich-entlaste-staedte-zwischenbilanz.html (2020). Accessed 10 Nov 2020
25. City Changer Cargo Bike. http://cyclelogistics.eu/about (2020). Accessed 10 Nov 2020
26. Kühne, U., Leibenath, M., Rau, C., Schulte, R., Wöltjen, L., vom Berg, B.W., Schopka, K., Krüger, L.: Sustainable Processes on the last mile—case study within the project ‘NaCl’. In: Kamilaris, A., Wohlgemuth, V., Karatzas, K., Athanasiadis, I.N. (eds.) Advances and New Trends in Environmental Informatics. Digital twins for. Progress in IS, pp. 91–108. Springer Nature, [S.l.] (2021)
27. Buldeo Rai, H., Verlinde, S., Merckx, J., Macharis, C.: Crowd logistics: an opportunity for more sustainable urban freight transport? Eur. Transp. Res. Rev. (2017). https://doi.org/10.1007/s12544-017-0256-6
28. Groen, W.P. de, Lenaerts, K., Bosc, R., Paquier, F.: Impact of digitalisation and the on-demand economy on labour markets and the consequences for employment and industrial relations. European Economic and Social Commitee (2017)
29. Schulte, R.: Bachelorthesis: Evaluation eines nachhaltigen Logistik-Konzepts auf Basis von elektromobilen Lastenrädern (unveröffentlicht), Hochschule Bremerhaven (2021)
30. DIN EN 16258:2013-03, Methode zur Berechnung und Deklaration des Energieverbrauchs und der Treibhausgasemissionen bei Transportdienstleistungen (Güter- und Personenverkehr); Deutsche Fassung EN_16258:2012. Beuth Verlag GmbH, Berlin
31. Umweltbundesamt: Kurzfristig kaum Lärmminderung durch Elektroautos. https://www.umweltbundesamt.de/sites/default/files/medien/377/dokumente/position_kurzfristig_kaum_laerminderung_im_verkehr.pdf (2013). Accessed 15 Jan 2021

32. Karl Otto Schallaboeck: Ueberlegungen zu Laerm und Schadstoffen im Zusammenhang mit dem Betrieb von Elektrofahrzeugen (2012)
33. Sheth, M., Butrina, P., Goodchild, A., McCormack, E.: Measuring delivery route cost trade-offs between electric-assist cargo bicycles and delivery trucks in dense urban areas. Eur. Transp. Res. Rev. (2019). https://doi.org/10.1186/s12544-019-0349-5
34. IHK: Die Ladezone im Blickpunkt. Anforderungen an die Güterversorgung in Köln und Leverkusen. https://www.ihk-koeln.de/upload/IHK_Studie_Ladezone_Web_66706.pdf (2018). Accessed 2 June 2021

Mobility as a Service and the Avoid-Shift-Improve Approach

Tina Ringenson and Anna Kramers

Abstract During the last few years, "Mobility as a Service" (MaaS) has been conceptualized and researched as a platform for integrated, mixed-mode mobility. While some hope it will lead to environmental benefits, its real effects are still unclear. Here, we explore how MaaS is related to, and can be combined with, the established "Avoid-Shift-Improve" transport planning approach (ASI). We see that the MaaS concept described in research does not support "Avoid"-ing unnecessary transport. We combine learnings from MaaS research with learnings from a living lab, where mobility services can be booked in combination with a local co-working hub for commuters. In both literature and living lab, we especially examine the role of public authorities for ASI in MaaS. We conclude that more research is needed on how MaaS can be guided by ASI, and suggest that non-travel accessibility services, such as co-working hubs, could be part of the MaaS concept to support "Avoid"-ing unnecessary transport. Furthermore, we suggest that urban form needs to be considered in MaaS research. We also see that public authorities have an important role to play in ensuring that MaaS serves ASI and sustainable mobility.

Keywords Mobility as a service · Sustainable mobility · Co-working hubs · Accessibility as a service · Avoid-shift-improve · Urban planning

1 Introduction

During the last few years, new mobility services enabled by digital technologies have started to appear in cities all over the world. Some of these services are taxi-like, such as Uber and Lyft. Other services offer access to vehicles, such as e-scooters, free-floating carpools, and bike-schemes. These services are often considered to be part of smart mobility and are sometimes suggested to be important parts of future sustainable transport systems. For example, the European Environment Agency [1] mentions in its report *Towards clean and smart mobility* that "[s]mart mobility

T. Ringenson · A. Kramers (✉)
Department of Sustainable Development, Environmental Science and Engineering (SEED), KTH Royal Institute of Technology, Teknikringen 10B, 100 44 Stockholm, Sweden
e-mail: kramers@kth.se

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_14

can combine different modes and options (public transport, car-sharing, car rental services, taxis and a bicycle system) to cater for mobility needs by using IT, apps, and smart invoicing" (p. 60).

Mobility as a Service (MaaS) is a relatively new smart mobility concept that combines different transport modes to offer mobility services as a package [2, 3]. Exactly what this entails is contested, as both theoretical and practical MaaS concepts vary significantly between sources, including both grey literature and academic sources. Some researchers [4] describe MaaS as "a new transport paradigm", where the whole transport sector will be a cooperative, interconnected ecosystem of seamlessly combinable transport services that reflect customer needs. Other researchers mean that MaaS may not lead to a paradigm shift [5], but is rather a service model that may impact travel behavior. However, one of MaaS' core characteristics is that it integrates different transport modes for point-to-point journeys and offers common payment options for the services, and that planning and payment are done through an online platform [2]. In this article, we use the phrase Mobility as a Service (MaaS) to describe an online platform that offers and integrates different transport modes and services, often from different service providers, and enables planning and paying for journeys. Furthermore, we suggest that this platform could also offer non-travel alternatives for such journeys by e.g., suggesting video meetings, online shopping, local workspaces for remote work, or other services.

On a global basis, the transport sector is a crucial sector to address for the transition to lower carbon dioxide emissions and energy use. Data from the International Energy Agency shows that transportation is responsible for 24% of direct CO_2 emissions from fuel combustion, whereof road vehicles account for nearly three-quarters [6]. In the EU, 25% of the GHG emissions are caused by the transport sector, whereof road transport stood for 70% in 2014 [7]. Apart from GHG emissions, traffic has also been connected to other environmental issues such as acidification and health problems due to particles and noise pollution, congestion, and problematic land use [8, 9].

According to the European Commission's low-emission mobility strategy the efficiency of the transport system should be increased in part through use of digital technologies [10]. It has also been suggested that digital technology can be used to decrease demand for transport [11, 12] and promote more efficient travel modes, such as walking, bicycling, and public transport through digital tools for trip planning [13].

The "Avoid-Shift-Improve" (ASI) approach has been formulated to support sustainable transport in urban planning and policy making [14, 15]. Drawing from [14–17], the ASI approach suggests to:

1. *Avoid* (reduce) inefficient or unnecessary travel or transport where appropriate. Travel does not take place and/or distances are shortened, e.g., with the help of urban planning for compact, multimodal communities and telework.
2. *Shift* to more environmentally friendly modes: from low-occupancy cars to active modes such as walking or cycling, and public transport. Congestion

pricing and public transport system improvements are examples of measures that could possibly facilitate this.

3. *Improve* the environmental performance of transport through higher energy efficiency. This can be achieved by technological improvements as well as higher occupancy of both public and private transport. Improved public transport performance, reducing energy intensity of each mode, and reducing the CO2 content of fuels can also be part of this.

Although the ASI approach has mostly been used in policy making and planning [14], its overall message is also in line with academic sources [8], which describes non-travel as the least unsustainable, followed by walking and bicycling and thereafter public transport. A number of indicators to determine the potential for a digital service to contribute to sustainable transitions of the transport system has been suggested [18]. In this article, not only mobility services are included, but also digital accessibility services, such as online meetings and local work spaces (sometimes referred to as "Accessibility as a Service", see [18]). The concept of "telework cottages" or "telecommuting centers" in residential areas has been around since the 1970's, but rarely lived past pilot stage due to a lack of utilization by users and the unwillingness of employers to pay rent for them [19]. During the 2010s, the market for commercial co-working spaces expanded significantly, with e.g. We Work establishing over 800 workspaces in 120 cities globally [20]. It should however be mentioned that current co-working spaces tend to target mainly start-up companies and be located in central areas and thereby not necessarily shorten commuting trips [21].

The real environmental effects of MaaS are yet unclear and have mainly been discussed on a hypothetical level [9, 22]. On one hand, as its basic idea is to provide the user with transport that does not require car ownership, there is hope that MaaS would disrupt car society and promote more sustainable transport modes (*shift*). There is also hope that MaaS can lead to a more efficient use of available transport infrastructure including vehicles (*improve*). On the other hand, it has been questioned on the basis of possibly undermining public service and leading to an increased demand of transport [9].

As MaaS per definition involves many different actors, and since it is expected to change transport significantly, public actors, such as municipalities, public transport actors, and national governmental bodies, need to decide how to act in relation to MaaS. Much of the work on MaaS consists of think pieces and theoretical discussions, while fewer focus on empirical studies. Part of the reason why empirical studies have been difficult to carry out is that there are few pilots and that much of the benefits, or problems, of MaaS will not show until the concept has gained a larger ground. Theoretical studies are therefore important. However, to also gain empirical knowledge, early or mock-up versions of services and policies can be tried out. "Designerly living labs" have been suggested as a useful setup for exploring and understanding more about emerging technologies and lifestyles [23]. As ASI is acknowledged as an approach for policies and planning for a sustainable transport system, it could be useful for steering MaaS towards supporting sustainable mobility. So far, there has been little exploration of the combination of ASI and MaaS on the basis of both theoretical and empirical learnings.

1.1 Aim and Research Questions

The aim of this article is to explore how the Avoid-Shift-Improve approach can guide MaaS towards better environmental performance of urban mobility. We take on this aim through three research questions:

- What is the relationship between ASI and current understandings of MaaS?
- How do public authorities currently act with regards to MaaS and its possible ASI effects, especially reducing unnecessary transport (avoid)?
- How can the ASI approach be used to improve the environmental effect of MaaS?

1.2 Scope and Limitations

Discussions about Mobility as a Service and smart mobility often include discussions about autonomous vehicles (AVs) (see e.g. [24]). However, the current MaaS pilots and implementations do not, and AVs have their own legal, technical and ethical questions to deal with. This article mainly focuses on the MaaS structure in itself, with the technologies available today at its center, and does not deal with AVs.

Furthermore, this article departs from environmental concerns. Those are often interlinked with questions of e.g. social equity, and such questions are central in many of the discussions regarding MaaS. Since environmental and social issues are so interlinked, social factors will appear in the text, although to a lesser extent.

Non-travel options, such as neighborhood telecommuting centers and online shopping, may have additional energy impacts apart from that from changed travel patterns. For example, in order for neighborhood telecommuting centers to lead to net energy savings, there needs to be considerations of the energy use of the main offices of the commuters as well as of the neighborhood telecommuting center [25]. Although important, these perspectives are not included in this article.

Last, this paper is based on data that was collected before the outbreak of the COVID-19 pandemic. It is possible that some of the statements by the interviewees, as well as some of the literary sources, could have been different with consideration of the pandemic. A forthcoming research paper will further explore the concept of work hubs in the light of the pandemic [26]. While we have not been able to include a follow up study during and after COVID-19 in this paper, we briefly discuss the possible impact of the pandemic on the paper's issues in the discussion section.

2 Methodology

This article uses two main methods: a small set of interviews in the context of a living lab, and a literature review. We first conducted the interviews to learn about the possible role of public authorities to ensuring environmental benefits from MaaS and AaaS, as described in Sect. 2.1. These interviews lay the foundation for a literature review directed towards providing a better understanding of the current relationships between public actors, MaaS and ASI, as described in Sect. 2.2. Citations and learnings from the interviews are combined with the literature review in order to add empirical and practical knowledge as well as to strengthen the avoid/reduce perspective in the paper.

2.1 Interviews

The interviews were conducted within the context of a "Designerly" Living Lab (DLL) [23]. The term "living lab" has been used for a number of different approaches, loosely tied together by a common purpose of developing knowledge through practice [27]. Commonly, living labs invite different stakeholders for collaboration and experimentation, often through testing and piloting innovations [23]. In governance contexts, the term "Urban Living Lab" has emerged to describe a type of experimental governance, where new technologies and lifestyles can be tested out and developed by urban stakeholders [28]. While the Urban Living Lab is characterized by its placement in a geographical area [29], the DLL is more focused on exploring (e.g. more sustainable practices) through interventions and learning from participants' reflections, as well as from the experience of creating the living lab in itself [23]. Co-creation of the lab with participants and research partners is important. DLLs are holistic in the sense that they include the full complexity of reality, both within the lab and in the surrounding society (ibid.). Although design methods and designers play a central part in DLLs, we have here taken advantage of the DLLs holistic characteristic to also harvest knowledge for non-design issues.

The DLL that sets the stage for the research in this paper is called "Jobba nära, resa smart"—"Work Closer, Travel Smarter" (hereon referred to as Work Closer). Work Closer is run by a team of researchers and PhD students from KTH Royal Institute of Technology for the purpose of creating knowledge about travel and non-travel service options, when linked together and offered together. The set-up of the living lab was created in a set of workshops together with representatives from partners in the research program. Some of them also acted as public and employer participants in the lab. A suburban residential area in Botkyrka municipality was chosen as a main node for the living lab. The area was recommended by the municipality in the co-creation process of formulating the living lab. The recommendation was motivated with the relatively high car-ownership and numbers of commuters in the area. Most of the participants of the living lab are local office workers that normally have long commutes and participate in the living lab as users. In line with the holistic perspective of the DLL and its inclusion of experiences of the "real world" in which it is set, both users', employers', and public authorities' perspectives are being explored. This article focuses on public authorities [21]. Building on these ideas, the non-travel part of the lab consists of a neighborhood telecommuting center (NTC). At the time of writing, the COVID-19 pandemic has led to a significant drop of usage, but before that, the NTC had around 60 participants and 44 active users. Its 14 desks, 3 phone boots, and one meeting room are bookable through a web application, and offered in combination with travel-related services such as trip planning, booking and paying for public transport, and access to two electric bicycles and one electric cargo bicycle. Support for video meetings has also been offered to the participants. Thus, Work Closer is not just a living lab for trying out a neighborhood telecommuting center, but rather for integrating non- and low-travel accessibility with mobility offers.

Altogether, four interviews were conducted (see Table 1). Three of the interviews were with representatives from the two municipalities that took part in the living lab, Botkyrka and Stockholm, respectively. We also interviewed an expert on public transport and special passenger transport within the Swedish Association of Local Authorities and Regions (SALAR). SALAR is an organization that organizes and represents Sweden's municipalities and regions [30]. The interviewees are

Table 1 Interview details

Ref	Role	Organization	Interviewer	Duration of interview
A	Area strategist Tullinge, Department of community development	Botkyrka Municipality (part of living lab)	Tina Ringenson	1 h 15 min
B	Overview planning and strategic location and district development issues as well as external analysis Development leader urban planning, Department of Sustainable Societal Development	Botkyrka Municipality (part of living lab)	Anna Kramers and Tina Ringenson	1.5 h
C	Expert in public transport and special passenger transport Department of growth and the built environment	SKR—Swedish Association of Local Authorities and Regions (not part of living lab)	Anna Kramers and Tina Ringenson	46 min
D	Innovation office, Stockholm Executive Office	City of Stockholm (part of living lab)	Anna Kramers and Tina Ringenson	48 min

further presented in Table 1. The interviews were semi-structured and qualitative [31], and carried out in August and September 2019. Questions were formulated to capture knowledge about possible responsibilities for mobility and accessibility services and for a digital MaaS platform. As previously mentioned, the interviews motivated making a literature review, which is now the focus of this article. The interviews complement the literature review with real-life perspectives, rather than being analyzed on their own. All the interviews were made in Swedish and have been translated to English by the authors. When the interviewees are quoted, the translations are as close to the original Swedish as possible.

2.2 Literature

After the interviews, we saw a need for a literature study focused on empirical insights [32] through a synthesis of what was known in research regarding the relationship between public actors, MaaS and ASI. We were also interested in the literature's relevance for real-world applications [32]. Van Wee and Banister [32] explain the added value of relevance for real-world applications as "a discussion or synthesis of how useful the literature is for real-world applications" (p. 281). We understand the studied literature as inherently focused on real-world application, and further discuss how the results could aid a real-world application.

MaaS has been frequently discussed in whitepapers and consultancy reports as a possible new business opportunity and as an expected game changer for transport planning [2]. However, our experience of working with MaaS has shown us that such reports tend to be optimistic about the effects of MaaS, and seldom deal with the possible needs for public authorities to handle any negative effects of it. Here, we want to learn what peer-reviewed research has asked and understood about how MaaS can or should be governed. We used Scopus and ScienceDirect to search for research articles and book chapters written in English on MaaS and governance, especially regarding ASI effects, by using the search terms *"mobility as a service" AND governance*, *"mobility as a service" AND policy* and *"mobility as a service" AND avoid-shift-improve* in keywords, abstract, and/or title. The papers we found were published between 2014 and 2020, with the vast majority having been published between 2018 and 2020. We also searched for *"mobility as a service" AND "co-working"* as well as *"mobility as a service" AND "non-travel"*, but none of these searches generated relevant results. It could also be noted that *"mobility as a service" AND avoid-shift-improve* only resulted in three hits. We went through the abstracts and excluded articles and book chapters that were either mainly focused on AVs or air travel, which are outside the scope of this article, or that only mentioned public actors in passing but focused on user studies or business policy. These exclusions left a first selection of 56 research articles and 13 book chapters. Then we searched for "MaaS" and "mobility as a service" in the texts. Papers where these keywords only appeared once or twice in passing were sorted out, leaving 34 papers to be analyzed. During the review process, 6 of the articles turned out to be only briefly mentioning

MaaS and/or public authorities' role in relation to them, and were therefore sorted out during the review process. The final number of papers in the literature review is 28.

We then read through the literature with focus on the research questions. The analysis of the data was inspired by deductive thematic analysis as described by Braun and Clarke [33]. Initial topics were formulated as follows:

- non-travel service mentions in the literature;
- environmental effects of MaaS,
- current actions taken by public authorities with regards to MaaS,
- the literature's conclusions and recommendations of actions from public authorities regarding MaaS.

The first data collection focused on instances in the literature that had some relevance to any of the above stated topics. After that, these instances were divided into themes. These themes were not determined through a specific quantitative rule of occurrence, but rather to whether they captured something important in relation to the question [33]. Tentative themes were formulated as we went, and any instance could hold several different themes. However, as the intention of this article was not to deduce patterns in the literature, but to explore the relationship between ASI and MaaS, these themes have not been explicitly written into the article. Instead, they served as soft guidelines for structuring the text before analyzing it from an ASI point of view. We understood mentions of decreased travel as related to *avoid*, mentions of MaaS enabling exchanging one mode of transport to another as related to *shift*, and mentions of efficiency in the transport system as related to *improve*, based on our understanding of ASI as described in Sect. 1.

3 Results

This section presents results from literature and interviews. The first section describes findings regarding the relationship between MaaS and ASI, specifically avoiding (Sect. 3.1). Then this is widened to how public authorities are acting regarding MaaS and ASI (Sect. 3.2) and then in what ways they can influence MaaS in an ASI direction (Sect. 3.3).

3.1 The Relationship Between ASI and MaaS

Very few of the reviewed texts included non-travel accessibility options as an alternative to travel options in their understandings of MaaS and its potentials [34] Instead, their suggestion is that the MaaS platform could include event and entertainment services and store discount coupons as rewards for choosing more sustainable transport modes. Here, the offers that are not mobility-centered are an incentive for shifting to more sustainable options. In a Delphi study with MaaS and transport experts, "[s]everal experts expressed the need to expand the circle of stakeholders beyond transportation and to include other sectors, such as tourism, large companies with many employees, public social service organizations (health, education) and real estate and residential developers [35]. The inclusion of services, such as home delivery and children pickup services, can also reduce the need for a private car and adds value and importance to MaaS". This formulation centers the reduction of needing to part with the private car (shift) rather than on avoiding travel overall. However, we do not find mentions of using MaaS to avoid unnecessary travel. One article even state that "in the future*, it is not about* reducing mobility as such since individual mobility is the prerequisite for social participation, progress, growth, and self-realization" [36] (p. 216, emphasis added).

Only one paper mentions the ASI approach [37]. The paper describes a multimodal integration project called "Seamless Transportation for Kochi" that was recently launched by Kochi's local metro rail venture, which is co-owned by the Government of India and the Government of Kerala. According to this paper, the project follows the ASI framework [15]. However, the ASI framework addresses more than transit-based transportation and the focus of the article is mainly the "*shift* from a mass transit dominated system to a larger and diverse MaaS ecosystem" [37] (p. 222, emphasis added). Note that this shift does not directly refer to shifting from cars to more sustainable transport modes, but rather to connected modes.

In the interviews, it was generally difficult to talk about non-mobility accessibility services and MaaS as connected and part of the same concept, even though Work Closer includes both mobility and NTC parts. For example, when asked about what possible benefits an integrated mobility and accessibility platform could have for municipal inhabitants, interviewees A and D only mentioned positive aspects of the NTC—"Suburbs will be more vivid during daytime", "Users will support the local economy when they shop for groceries, eat lunch, and run errands close by", and "There will be more adults in the local area during the hours between the end of the school day and when people normally get home from work". Reminders that the NTC was part of a DLL that also included mobility services were needed to get answers concerning mobility. Interviewee A then said that "travelling is a question for the region. Travel packages are more on regional level [than the municipal level]". However, interviewee C—the only interviewee who was not involved in the living lab—gave answers regarding both NTC and mobility services when asked questions that did not specify which of the service type we considered.

A high number of papers mention that environmental benefits of MaaS are so far unclear [9, 22, 35, 36, 38–46].

The reasons for this unclarity vary. For example, a reversed avoid effect where they suggest that the smart mobility rhetoric promises both a mobility system that can reduce demand, and one that create new demand [9]. The reasons they state for this are MaaS' unboundedness and easy access and the mobility service provider's incentive to generate more mobility in order to earn larger profits. Interviewees in one study mention the risk of congestion due to growth in ride hailing services, which would mean a reversed improve effect [47]. One text [48] warns that easier transactions for taxi- and ride hailing services might lead to a reversed shift effect in the form of increased car travel at the expense of public transport. On the other hand another text [49] present results that show how the participants in a MaaS pilot in Gothenburg ended up using car-based services even less than they had assumed when signing up, and shifted to other modes of transport. However, the same text also state that this created problems with the revenue model of MaaS because of the lower revenues than expected due to the low car use.

"[T]he general idea of MaaS is that the digital integration of the totality of mobility service within a territory /.../ the explicit aim of this transformation of the mobility market is to steer travel practices toward more sustainable forms based on combinations of public transport and shared transport modes put together to meet individual needs, and to discourage car ownership" [49], (p. 156). Here, MaaS is understood as inheritably orchestrated by public authorities aiming towards shifting and possibly improving transport by making the transport system more efficient. MaaS can improve the match between demand and supply of transport and thereby "contribute to an improvement in the sustainability of the transport system by enhancing its efficiency [50]. Additionally, the hyper-convenient transport service in MaaS is likely to increase public transport patronage and attracting car drivers away from their personal motorized vehicle. Hence, MaaS would reduce the negative impacts of the transport system on the environment and decrease the demand for road and parking spaces" (p. 51). Here, then, we see both expected improve effects in the form of a more efficient transport system, and shift effects from privately owned car to public transport.

There is a warning that a transport system based on shared vehicles may lead to car-based services overtaking public transport and resulting in increased traffic with lower efficiency—in other words, a negative effect for shift and improve measures [24]. For more societal benefits from MaaS, it will likely be necessary to uphold a certain amount of traditional public transport [39]. This, could be done through government funding models, where a MaaS subscription price should include road user charges for car use [24]. Similarly, another suggestion is that while economic incentives for more sustainable travelling may have limited effects on their own, MaaS can be priced to reflect the real costs of different trips, and therefore "offers a 'big bang' approach in allowing an easily implementable digital system which can price [transport] according to time of day/../, geography /.../, and modal /.../ efficiency." (p. 16) [51]. This, Wong et al. [51] A suggestion, to achieve a much more efficient and low-GHG transport system by making car users pay the real cost of the car, rewarding shifting transport modes, and improving the transport system overall by managing the transport flow [52]. However, the same text also suggest that it is possible that MaaS will be a niche product, in which case economic interventions introduced via MaaS will be less effective in changing the transport system to the better.

The dilemma of the sometimes-conflicting interests between traditional public service, specifically rail-bound services, and point-to-point mobility services can be connected to one of the interview findings. Interviewee A described some recent dilemmas in the intersection between different transport interests. A fairly newly built residential area in Tullinge is currently lacking in public transport and highly reliant on car traffic. The current inhabitants of the new residential area want buses that go straight to a close-by larger business and service centrum area where many of the resident's work, but which belongs to a different municipality. Meanwhile, Botkyrka municipality has an interest in seeing these inhabitants instead go by bus to the local commuter train station and then take the train the one stop it takes to get to the other centrum area. This is because an increased flow to the local commuter train station would motivate the public transport company to improve the station building, which would benefit the municipality.

3.2 Public Authorities' Current Actions for MaaS and ASI

Looking at how public authorities have dealt with MaaS so far, [5, 22, 40, 41] all mention legal change having been made by public authorities to in some way deal with the emergence of MaaS. Most commonly mentioned was "the Finnish transport code", which has made it mandatory for any transport provider to provide open information and ticketing. This is considered important for the transport brokers of MaaS if they want to include public transport ticketing as part of their MaaS packages [5, 22, 40, 41]. In New South Wales, Australia, previous regulations for taxis and hire cars have been replaced by the point-to-point act and regulation [5, 53, 54]. This legalized ride-sourcing, removed license requirements on booked-only hire vehicles, and moved the responsibility of vetting drivers from the government to companies (ibid.)—all which could be considered enabling mobility services, with a focus on car-based services. Requiring public transport to enable public transport trips to be sold through MaaS can be seen as a way to encourage shifting to public transport through MaaS.

There are also mentions of public actors supporting and/or participating in MaaS pilots and explorative projects [5, 50, 55]. Work Closer can also be considered such an instance. In the words of interviewee A, Botkyrka works with "two types of measures—operative measures, that are carried out right now; and strategy to learn what can be good in the future. [Work Closer] is part of both operative measures and strategy, because it can both solve problems with car traffic and teach the municipality how to operate in the future."

One article that takes its starting point in existing policy rather than MaaS-adjacent projects analyses the discourse around MaaS in strategies, reports etc. from Swedish public actors [38]. According to this article, the overall message from these documents is that MaaS and digitalized mobility will eventually happen and probably lead to positive change, and that the role of the public sector should mainly be to enable this transition, although without being too passive.

The interviewees had different understandings of possible roles for public actors in MaaS. Interviewee B mentioned that the municipal comprehensive plan can set as a goal to build public transport (supporting shift), and that municipalities' climate strategy documents can state mobility as an important area to invest in for urban development programs. According to interviewee A, the need for the municipality to act on the basis of public benefit meant that the municipality should not manage an NTC: "As a municipal actor you have to think on the basis of public benefit. And then you want to be able to say what type of public benefit something is /…/ It could be a system with a private actor and something like a gym card." However, interviewee A mentioned that a NTC could be mentioned in an urban planning program, development plan or comprehensive plan, which would make it easier for a commercial actor to establish it in the area. Interviewee D did not share the view that organizing a NTC was not within the responsibilities of the municipality: "This is obviously within the municipality's mission! /…/ With the help from local NTCs you can satisfy the traffic and communication ecosystem of the future." However, according to interviewee D, the current legal framework and state-region-municipality organization are barriers for implementing NTCs.

3.3 Possible Municipal and Regional Policy and Actions for ASI Effects of MaaS

Several of the texts clearly state that authorities cannot count on environmental benefits from MaaS unless they make efforts in that direction [56]. Transport needs to be properly priced to better reflect the societal costs it imposes. Modern technology means this could be done through some form of charging mechanism based on a range of factors such as time of travel, location, vehicle occupancy, emissions generated, and level of demand for the infrastructure, as well as on individual traveler attributes /…/ [G]overnment needs to consider a desirable direction for the development and uptake of new passenger transport and related technologies /…/ perhaps by incentivizing these technologies /…/ either financially or via regulatory mechanisms" (p. 66) [57, 58, 22] expressed that climate policies should formulate the need for adoption of MaaS schemes, while at the same time ensuring that they lead to positive change.

MaaS is not an objective in itself, but could be a part of a larger strategy of long-term visions of public goals [5]. However, the same authors state that "measures that increase the relative attractiveness of mobility services (e.g., through urban design and transport subsidies) and facilitate the decline of the current private car regime (including viable exit alternatives) are essential for the future development of MaaS" (p. 59), and that the public sector should "transform its general approach to public innovation, in terms of increased external influence to goal-setting" (p. 59). In our interviews, when asked about the municipalities' possibilities in stimulating sustainable transport use among their inhabitants, interviewees B and C both pressed the importance of municipalities to create larger visions. Interviewee B said that it is not just about strategic documents, but about creating good public environments: "It will be the best outcome if the entire organization is directed towards the goals it wants to achieve." Interviewee C underlined that to achieve a change in behavior, one cannot just implement e.g. certain mobility services and expect change; it is important to see what needs arise with time: "You have to look upon it in a bigger picture: How do we switch to the sustainable transport system?" Both interviewees A and D expressed expectations that a NTC would lead to increased local shopping and eating lunches locally. As walking distance services are part of the avoid strategy [16], it also seems that if these expectations on NTCs and their impact on the local economy are true, NTCs will also lead to avoid on a larger scale. In the case of Work Closer, [25].

Land use and urban planning are also essential parts of the ASI framework, especially avoid and share [15, 17]. Several of the interviewees brought up the relationship between traditional, "physical" urban planning, and MaaS. According to interviewee A, the municipality can "Make land allocation agreements, development contract contents, and try to get in a builder that is interested in carpools. But that is easier for some municipalities than others. It is easier if you have high demand for your land". Interviewee B mentioned that "the urban planning can work with better bicycle connections, and make it easier to find with the help of signs". Interviewee C said that "If one thinks through what one wants, one can influence [MaaS developments and effects], but then the municipalities need also have political courage and dare lowering the parking spot requirements, for example /…/ One has to see it in a larger picture: How do we transition into the sustainable transport system?". These thoughts are in line with research on attitudes towards mobility as a service which shows that although users may be more open to try a new service if it is offered via MaaS, their willingness to use certain sustainable modes in a MaaS service are also connected to the physical conditions for using these modes [57]. Safety was a major barrier, especially for cycling, and the conclusion is that investments in good cycling infrastructure is important for encouraging bicycle-centered mobility services. Similar conclusions were drawn in a living lab where the participants were given direct economic incentives in the form of payment for bicycling for bicycling, but safe, well-kept, and easy to orient bicycle routes were still more important factors [42]. A MaaS app can also be used to help users find safer bicycle routes [57].

4 Discussion

There are mentions of rebound effects of increased travel due to increased availability in the literature, but this is countered with shift-related measures to avoid increased car travel rather than avoid options. As mentioned in Sect. 3.1., it was difficult to establish the connection between non-travel access and MaaS in the interviews as well. Instead, the interviewees discussed NTCs by itself, and MaaS by itself, even as we tried to ask them about them as connected in one package. Although this effect could perhaps have been avoided with a different setup or phrasing of the interview questions or how we talk about the living lab, the combined results of the literature review and the interviews still suggest that the current MaaS discourse has a strong focus on travelling, with little to no considerations of non-travel accessibility as an alternative to travelling.

Furthermore, some of the results raise concerns about the possibility of including non-travel accessibility services to MaaS. First of all, there are uncertainties around e.g. NTCs in themselves—who should run them and how, and what the legal aspects are. Although our interviewees are optimistic about the NTCs performance, various versions of "telework cottages" have been around as a concept since decades, but have never caught on, as mentioned in the introduction—although centrally located NTCs targeting start-ups have. The COVID-19 pandemic may impact the perceived usefulness of NTC's in unexpected ways. It could be argued that many have been able to work successfully from home, but working from home also presents challenges that could be solved by the NTC. At the same time, the pandemic may lead to a more generous view on remote work, which could support NTC's as an alternative to working from home. This is further discussed in our forthcoming paper [26]. Second, the already complicated situation regarding legal, policy, and payment aspects of MaaS would be even further complicated with additional types of services. As mentioned in Sect. 3.1., the interviewees that were involved in Work Closer initially gave answers that specifically had to do with its NTC aspects. However, we cannot be sure if that has to do with the newness of the concept of integrated mobility and non-travel accessibility options, that the NTC was a physical space they knew of while the mobility options were not as visible to them, or if this simply could have been prevented through a different setup of the interview.

On the other hand, the literature also raises concerns that MaaS could lead to increased travelling, often car-based. From this perspective, incorporating non-travel accessibility services into MaaS could be useful to further the avoid aspect of the ASI approach.

Both our interviews and the literature indicate that travel patterns and public space will both impact and be impacted from MaaS. This goes for service availability as well as the quality of e.g. bicycle infrastructure. From the perspective of environmental impact, it seems that traditional investments in sustainable transport such as safe bicycle lanes can also help support sustainability in MaaS.

The point-to-point idea of transport overlooks the importance of urban form as a dwelling space, and how these transport options are experienced, not only from a time and efficiency perspective but also from perspectives of safety and enjoyment. There are also other questions that need to be asked. For example, the comment from interviewee A regarding the wish from residents to be able to travel more efficiently to a nearby larger business and service area conflicting with the municipalities' goals of getting a larger flow to the local center and commuter train station, so that these can be improved. There are connections between transport patterns and the vividness and functionality of the city. We will need to further examine the connections between MaaS and urban form, and an ASI perspective can both provide an urban form perspective and safeguard those environmental concerns are not forgotten in that examination.

Public authorities need to make efforts not only to implement MaaS, but to ensure that MaaS actually achieves environmental benefits, according to our literature review. However, to the extent that the interviews cover these issues, none of the interviewees seems sure of how this should happen and what role the public should play. This is also in line with results from the literature, pointing to an uncertainty among public actors concerning how to act. Furthermore, aside from some texts mentioning the risks of purely commercial actors undermining the public transport system and/or steering towards more car travel, less attention is given to exactly how MaaS should be governed to actually achieve environmental benefits.

5 Conclusions

Connecting back to the paper's aim to understand how the Avoid-Shift-Improve approach can guide MaaS towards better environmental performance of urban mobility, much of the literature point towards a need for authorities to ensure a positive environmental performance of MaaS, but few examples are given of how this should be done, although regulations and/or financial incitements are mentioned. Therefore, more research needs to be done that focuses specifically on this.

As for the "avoid" part specifically, we found that few texts mentioned non-travel options. However, as several authors discuss risks of increased motorized travelling through MaaS, we conclude that non-travel options should be considered in the MaaS discussions and their possibilities further investigated. As urban form continues to be important for how we move in cities—and our travel patterns continue to be important for urban form—this aspect should also be considered in MaaS discussions. Using an ASI approach to MaaS can be a way to add that perspective.

Although a lot of the literature calls for more action from the public authorities in order to ensure a sustainable MaaS transition, the interviewees we spoke to give a picture of public authorities lacking both knowledge and resources for doing so. Furthermore, as so much research seems to assume that a MaaS or Maas-adjacent future is ahead, and this view also seems to be widespread among public actors, we align ourselves with researchers pointing to the need for public actors to familiarize themselves with the MaaS concept. This should not only entail an optimistic view that if public actors just get out of the way, urban transport systems will become more economically viable as well as more sustainable from a climate perspective. Instead, a critical view of risks as well as opportunities need to be adapted, where environmental goals are prioritized. As the ASI approach is originally formulated for sustainable transport planning on a general level, we believe that it can guide MaaS towards better environmental performance of urban mobility by being a framework for consideration by both public and private actors.

Acknowledgements We want to thank the interviewees for their time and contributions.

This work was funded by Mistra, the Swedish foundation for strategic environmental research. The authors have no competing interests to declare.

Glossary

AaaS Accessibility as a Service.
ASI Avoid-Shift-Improve.
DLL Designerly Living Lab.
MaaS Mobility as a Service.

References

1. EEA.: EEA Signals 2016: towards clean and smart mobility—transport and environment in Europe (2016)
2. Jittrapirom, P., Caiati, V., Feneri, A.-M., et al.: Mobility as a service: a critical review of definitions, assessments of schemes, and key challenges. Urban Plan. **2**, 13 (2017). https://doi.org/10.17645/up.v2i2.931
3. Hietanen, S.: 'Mobility as a Service'—the new transport model ? Eurotransport (2014)
4. König, D., Eckhardt, J., Aapaoja, A., et al.: Business and operator models for MaaS (2016)
5. Smith, G., Hensher, D.A.: Towards a framework for Mobility-as-a-Service policies. Transp. Policy **89**, 54–65 (2020). https://doi.org/10.1016/j.tranpol.2020.02.004
6. International Energy Agency.: Tracking Transport 2019. Paris (2019)
7. European Commission Transport emissions: A European Strategy for low-emission mobility. https://ec.europa.eu/clima/policies/transport_en. Accessed 15 Jun 2020
8. Banister, D.: The sustainable mobility paradigm. Transp. Policy **15**, 73–80 (2008). https://doi.org/10.1016/j.tranpol.2007.10.005
9. Docherty, I., Marsden, G., Anable, J.: The governance of smart mobility. Transp. Res. Part A: Policy Pract. **115**, 114–125 (2018). https://doi.org/10.1016/j.tra.2017.09.012
10. European Environment Agency (EEA).: Transitions towards a more sustainable mobility system (2016)
11. Mitchell, W.J.: E-topia: Urban life, Jim--but not as we know it (2000)
12. Hilty, L.M., Arnfalk, P., Erdmann, L., et al.: The relevance of information and communication technologies for environmental sustainability—a prospective simulation study. Environ. Model. Softw. (2006). https://doi.org/10.1016/j.envsoft.2006.05.007
13. Kramers, A., Höjer, M., Lövehagen, N., Wangel, J.: Smart sustainable cities—exploring ICT solutions for reduced energy use in cities. Environ. Model. Softw. **56**, 52–62 (2014). https://doi.org/10.1016/j.envsoft.2013.12.019
14. United Nations.: Mobilizing Sustainable Transport for Development: Analysis and Policy Recommendations from the United Nations Secretary-General's High Level Advisory Group on Sustainable Transport. Technology (2016)
15. Dalkmann, H., Brannigan, C.: Sustainable Transport: A Sourcebook for Policy-makers in Developing Cities (2007). http://www.gtkp.com/assets/uploads/20091123-095443-1692-5e_TCC.pdf. https://doi.org/10.1007/978-3-319-02165-2
16. Tiwari, R., Cervero, R., Schipper, L.: Driving CO_2 reduction by Integrating Transport and Urban Design strategies. Cities **28**, 394–405 (2011). https://doi.org/10.1016/j.cities.2011.05.005
17. Lah, O.: Sustainable development synergies and their ability to create coalitions for low-carbon transport measures. In: Transportation Research Procedia (2017)
18. Ringenson, T., Arnfalk, P., Kramers, A., Sopjani, L.: Indicators for promising accessibility and mobility services. Sustainability **10** (2018). https://doi.org/10.3390/su10082836
19. Feitelson, E., Salomon, I.: The Political Economy of Transport Innovations. In: Beuthe, M., Himanen, V., Reggiani, A., Zamparini, L. (eds.) Transport Developments and Innovations in an Evolving World, pp. 11–26. Springer, Berlin, Germany (2004)

20. WeWork Här finns WeWork. https://www.wework.com/sv-SE/locations. Accessed 16 Jul 2020
21. Kramers, A., Höjer, M., Nyberg, M., Söderholm, M.: Work hubs: Location considerations and opportunities for reduced travel. Proc. Enviroinfo ICT Sustain. **2015**, 126–135 (2015)
22. Audouin, M., Finger, M.: Empower or Thwart? Insights from Vienna and Helsinki regarding the role of public authorities in the development of MaaS schemes. Transp. Res. Procedia **41**, 6–16 (2019). https://doi.org/10.1016/j.trpro.2019.09.003
23. Sjöman M, Hesselgren M Designerly Living Labs : Early-stage exploration of future sustainable concepts. In: The Design Research Society 2020 International Conference, BRISBANE, 11–14 AUG. Brisbane
24. Hensher, D.A.: Tackling road congestion—what might it look like in the future under a collaborative and connected mobility model? Transport Policy **66**(2018).https://doi.org/10.1016/j.tranpol.2018.02.007
25. Vaddadi, B., Bieser, J., Pohl, J., Kramers, A.: Towards a conceptual framework of direct and indirect environmental effects of co-working. In: ICT4S'20, June, Bristol, UK. Bristol (2020)
26. Ringenson Tina VBSMHMKA Do they work? Exploring possible potentials of Neighbourhood Telecommuting Centresin supporting sustainable travel
27. Cheyne, C.: Cities and low carbon transitions. Urban Policy Res. (2013). https://doi.org/10.1080/08111146.2013.845133
28. Bulkeley, H., Castán Broto, V.: Government by experiment? Global cities and the governing of climate change. Trans. Inst. Br. Geogr. (2013). https://doi.org/10.1111/j.1475-5661.2012.00535.x
29. Palgan, Y.V., McCormick, K., Evans, J.: Urban Living Labs. In: Urban Living Labs (2018)
30. Swedish Association of Local Authorities and Regions Swedish Association of Local Authorities and Regions. https://skr.se/tjanster/englishpages.411.html. Accessed 28 Apr 2020
31. Edwards, R., Holland, J.: What is Qualitative Interviewing? Bloomsbury Academic, London (2013)
32. Van, W.B., Banister, D.: How to write a literature review paper? Transp. Rev. (2016). https://doi.org/10.1080/01441647.2015.1065456
33. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
34. Polydoropoulou, A., Pagoni, I., Tsirimpa, A., et al.: Prototype business models for Mobility-as-a-Service. Transp. Res. Part A: Policy Pract. **131**, 149–162 (2020). https://doi.org/10.1016/j.tra.2019.09.035
35. Jittrapirom P, Marchau V, van der Heijden R, Meurs H (2018) Future implementation of mobility as a service (MaaS): Results of an international Delphi study. Travel. Behav. Soc., 1–14https://doi.org/10.1016/j.tbs.2018.12.004
36. Spickermann, A., Grienitz, V., Von Der Gracht, H.A.: Heading towards a multimodal city of the future: multi-stakeholder scenarios for urban mobility. Technol. Forecast. Soc. Chang. **89**, 201–221 (2014). https://doi.org/10.1016/j.techfore.2013.08.036
37. Singh M (2020) India's shift from mass transit to MaaS transit: Insights from Kochi. Transportation Research Part A: Policy and Practice 131:219–227
38. Henriksson, M., Witzell, J., Isaksson, K.: All change or business as usual? The discursive framing of digitalized smart accessibility in Sweden. Transp. Res. Procedia **41**, 625–636 (2019). https://doi.org/10.1016/j.trpro.2019.09.112
39. Hensher, D.A.: Tackling road congestion – What might it look like in the future under a collaborative and connected mobility model? Transp. Policy **66**, A1–A8 (2018). https://doi.org/10.1016/j.tranpol.2018.02.007
40. Karlsson, I.C.M., Mukhtar-Landgren, D., Smith, G., et al.: Development and implementation of Mobility-as-a-Service – A qualitative study of barriers and enabling factors. Transp. Res. Part A: Policy Pract. **131**, 283–295 (2020). https://doi.org/10.1016/j.tra.2019.09.028
41. Pangbourne, K., Mladenović, M.N., Stead, D., Milakis, D.: Questioning mobility as a service: Unanticipated implications for society and governance. Transp. Res. Part A: Policy Pract. (2019). https://doi.org/10.1016/j.tra.2019.09.033

42. Sjöman, M., Ringenson, T., Kramers, A.: Exploring everyday mobility in a living lab based on economic interventions. Eur. Transp. Res. Rev. **12**(2020). https://doi.org/10.1186/s12544-019-0392-2
43. Smith, G., Sochor, J., Karlsson, I.C.M.A.: Mobility as a Service: Development scenarios and implications for public transport. Res. Transp. Econ. **69**, 592–599 (2018). https://doi.org/10.1016/j.retrec.2018.04.001
44. Alonso-González, M.J., Hoogendoorn-Lanser, S., van Oort, N., et al.: Drivers and barriers in adopting Mobility as a Service (MaaS) – A latent class cluster analysis of attitudes. Transp. Res. Part A: Policy Pract. **132**, 378–401 (2020). https://doi.org/10.1016/j.tra.2019.11.022
45. Noussan, M., Tagliapietra, S.: The effect of digitalization in the energy consumption of passenger transport: An analysis of future scenarios for Europe. J. Clean. Prod. **258**, 120926 (2020). https://doi.org/10.1016/j.jclepro.2020.120926
46. Haglund, N., Mladenović, M.N., Kujala, R., et al.: Where did Kutsuplus drive us? Ex post evaluation of on-demand micro-transit pilot in the Helsinki capital region. Res. Transp. Bus. Manag. **32**, 100390 (2019). https://doi.org/10.1016/j.rtbm.2019.100390
47. Hirschhorn, F., Paulsson, A., Sørensen, C.H., Veeneman, W.: Public transport regimes and mobility as a service: Governance approaches in Amsterdam, Birmingham, and Helsinki. Transp. Res. Part A: Policy Pract. **130**, 178–191 (2019). https://doi.org/10.1016/j.tra.2019.09.016
48. Pangbourne, K., Mladenović, M.N., Stead, D., Milakis, D.: Questioning mobility as a service: unanticipated implications for society and governance. Transp. Res. Part A: Policy Pract. **131**,(2020). https://doi.org/10.1016/j.tra.2019.09.033
49. Aguilera, A., Boutueil, V.: Implications for public policy. Elsevier Inc. (2018)
50. Jittrapirom, P., Marchau, V., van der Heijden, R., Meurs, H.: Dynamic adaptive policymaking for implementing Mobility-as-a Service (MaaS). Res. Transp. Bus. Manag. **27**, 46–55 (2018). https://doi.org/10.1016/j.rtbm.2018.07.001
51 Wong YZ, Hensher DA, Mulley C (2020) Mobility as a service (MaaS): Charting a future context. Transportation Research Part A: Policy and Practice **131**, 5–19. https://doi.org/10.1016/j.tra.2019.09.030
52. Wong, Y.Z., Hensher, D.A., Mulley, C.: Mobility as a service (MaaS): Charting a future context. Transp. Res. Part A: Policy Pract. **131** (2020). https://doi.org/10.1016/j.tra.2019.09.030
53. NSW Government (2019) Point to point transport (taxis and hire vehicles) act 2016 No 34. https://www.legislation.nsw.gov.au/#/view/act/2016/34. Accessed 30 Apr 2020
54. NSW Government (2019) Point to Point Transport (Taxis and Hire Vehicles) Regulation 2017. https://www.legislation.nsw.gov.au/#/view/regulation/2017/424. Accessed 30 Apr 2020
55. Sochor, J., Strömberg, H., Karlsson, I.C.M.: Implementing mobility as a service: Challenges in integrating user, commercial, and societal perspectives. Transp. Res. Rec. **2536**, 1–9 (2015). https://doi.org/10.3141/2536-01
56. Enoch, M.P., Cross, R., Potter, N., et al.: Future local passenger transport system scenarios and implications for policy and practice. Transp. Policy **90**, 52–67 (2020). https://doi.org/10.1016/j.tranpol.2020.02.009
57. Matyas, M.: Opportunities and barriers to multimodal cities: lessons learned from in-depth interviews about attitudes towards mobility as a service. Eur. Transp. Res. Rev. **12**https://doi.org/10.1186/s12544-020-0395-z
58. Hoerler, R., Stünzi, A., Patt, A., Del Duce, A.: What are the factors and needs promoting mobility-as-a-service? Findings from the Swiss Household Energy Demand Survey (SHEDS). Eur. Transp. Res. Rev.**12** (2020)https://doi.org/10.1186/s12544-020-00412-y

A Framework for Assessing Impacts of Information and Communication Technology on Passenger Transport and Greenhouse Gas Emissions

Jan C. T. Bieser and Mattias Höjer

Abstract Information and communication technology (ICT) provides unprecedented opportunities to reduce greenhouse gas (GHG) emissions from passenger transport by avoiding, shifting or improving transport. Research on climate protection through ICT applications in passenger transport mainly focuses on theoretical potentials, is assuming that digital mobility services replace GHG-intensive transport modes (e.g. car travel), and does not specify the conditions under which decarbonization potentials will materialize. It is known that digital mobility services can also take a complementary (as opposed to substituting) role in travel or replace non-motorized travel, which can increase GHG emissions. Based on existing literature, we develop a conceptual framework to guide qualitative and quantitative assessments of the relationship between ICT use, passenger transport and GHG emissions. The framework distinguishes three types of effects: (1) First-order effects, GHG impacts of producing, operating and disposing the ICT hardware and software, (2) second-order effects, impacts of ICT on properties of transport modes, transport mode choice and travel demand, and (3) third-order effects, long-term structural changes due to ICT use (e.g. residential relocation). We qualitatively demonstrate the framework at the example of automated driving and discuss methodological challenges in assessments of ICT impacts on passenger transport such as the definition of system boundaries, consideration of socio-demographic characteristics of individuals and the inference of causality. The framework supports researchers in scoping assessments, designing suitable assessment methods and correctly interpreting the results, which is essential to put digitalization in passenger transport at the service of climate protection.

Keywords ICT · Digitalization · Climate · Travel · Passenger transport · Mobility

J. C. T. Bieser (✉) · M. Höjer
Department of Sustainable Development, Environmental Science and Engineering (SEED), KTH Royal Institute of Technology, Stockholm, Sweden
e-mail: bieser@kth.se

M. Höjer
e-mail: hojer@kth.se

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_15

1 Introduction

Global greenhouse gas (GHG) emissions from transport increased from 2.9 Gt CO_2e in 1970 to 7.1 Gt CO_2e in 2010, mostly driven by increasing GHG emissions from road transport [1]. In the EU-28 countries, transport caused 27% of total GHG emissions in 2017; 72% of transport emissions came from road transport, 44% of which came from passenger cars [2].

The “Avoid-Shift-Improve”-Approach for sustainable transport suggests to meet future transport demand by avoiding unnecessary travel, shifting travel to more environmentally-friendly transport modes and improving the environmental performance of transport modes [3]. Information and communication technologies (ICT) provide new possibilities to avoid, shift and improve transport and can thereby reduce passenger transport-related GHG emissions. For example, virtual mobility (e.g. videoconferences) can avoid transport, multimodal mobility-as-a-service (MaaS) platforms can lead to a modal shift from motorized individual transport to public transport, and automated driving can increase fuel efficiency in car transport through platooning [4].

Existing research in the field of ICT, passenger transport and GHG emissions can be clustered into two camps. On the one side, various studies have assessed the climate protection potential of ICT applications in transport [5–8]. These studies usually agree that ICT provides and unprecedented potential for reducing GHG emissions in passenger transport, e.g. by substituting physical with virtual mobility or by increasing utilization of transport modes and reducing vehicle kilometers traveled through car or ride sharing. On the other side, various transportation researchers that have investigated the conceptual relationship between ICT use and travel argue that ICT can not only substitute or reduce physical travel, but also induce additional travel [9, 10]. Additionally, ICT applications can lead to a shift to more GHG-intensive transport modes. For example, e-scooters are often considered climate-friendly, but existing e-scooter services often substitute non-motorized or public transport and increase GHG emissions [11].

Thus, even though large theoretical potentials to reduce passenger transport GHG emissions through ICT exist, existing studies do not provide indications about the actual impacts in a real-life setting because they apply too narrow system boundaries, take overoptimistic assumptions or neglect potential increasing effects of ICT on GHG emissions from passenger transport. In order to put ICT applications in passenger transport at the service of climate protection, a thorough understanding of the interrelation between ICT applications, the transport system and GHG emissions is required.

In this article, we develop a conceptual framework of the relationship between ICT applications in passenger transport and GHG emissions. To do so, we first provide an overview of drivers of GHG emissions caused by passenger transport, literature on the relationship between ICT use and passenger transport and literature on GHG impacts of ICT. Based on this overview, we develop a conceptual framework of the relationship between ICT applications in passenger transport, the transport system and GHG emissions. Finally, we demonstrate the applicability of the framework at the example of automated driving, discuss methodological challenges in assessments and derive recommendations for future work.

2 ICT Impacts on Passenger Transport and GHG Emissions

In the following, we describe drivers of passenger transport GHG emissions, research on impacts of ICT on passenger transport, and research on impacts of ICT on GHG emissions.

2.1 Drivers of Passenger Transport GHG Emissions

The IPCC [1] summarizes that yearly GHG emissions caused by (passenger) transport mainly depend on:

- Total distances traveled per year (passenger km per year, pkm_{total})
- Modal choice of passengers (pkm_{mode}/pkm_{total})
- Energy intensity (MJ/pkm) by transport mode
- The fuel carbon intensity (t CO_2e/MJ) by transport mode.

While the IPCC focuses on direct GHG emissions associated with operation of vehicles, passenger transport is also associated with GHG emissions embedded in vehicles, infrastructures and fuels, i.e. caused by vehicle and infrastructure production, maintenance and disposal as well as fuel provisioning.

Figure 1 provides an overview of the amount of GHG emissions by pkm and transport modes in Switzerland considering direct and embedded GHG emissions [12]. It shows that vehicles with internal combustion engines (e.g. gasoline/diesel cars and autobuses) cause significantly more GHG emissions per pkm than vehicles with electric drivetrains. At the same time, public transport causes on average less GHG emissions per pkm than motorized individual transport as vehicles are shared among more passengers and/or are electrified.

In the case of electricity-powered transport modes, GHG emissions from operation of vehicles and fuel provisioning depend on the energy carriers used to generate electricity. Thus, in countries whose electricity generation is mainly based on fossil fuels, GHG emissions caused by electric vehicles are higher than in Switzerland.

As of today, passenger cars (most of which are still powered with diesel or gasoline engines) are the prevalent transport mode in the EU (83% of all pkm in 2018), followed by motor coaches, buses, trolleybuses (9%) and trains (8%) [13]. In future, electric drivetrains will significantly increase the GHG efficiency in passenger transport. However, their GHG reducing effect is limited to some extent, because

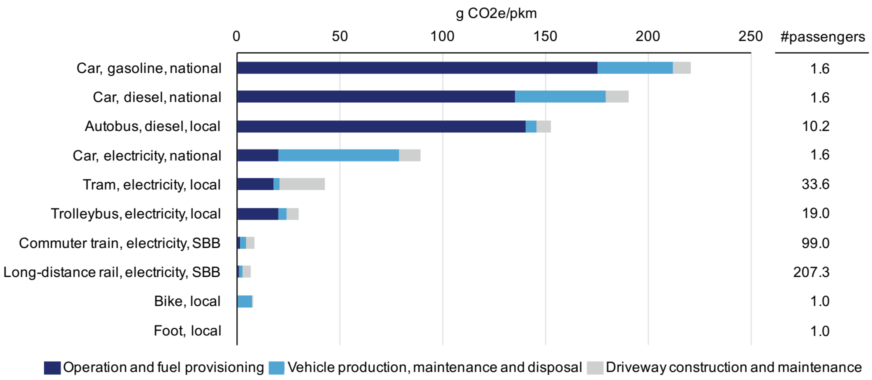


Fig. 1 GHG emissions per passenger kilometer by transport mode in Switzerland, based on [12]. "#passengers" is the average number of passengers on board the vehicle. Driveways are for example roads and rails. "Local" means trips up to 5 km, "national" between 5 and 200 km. For commuter and long-distance trains, the company specific electricity mix of the Swiss national railway company SBB is used to estimate GHG impacts. No impacts were allocated to walking (foot)

(renewable) electricity generation also causes GHG emissions [12]. Plus, motorized individual transport with cars is inefficient in terms of utilization of vehicles and land use [12, 14, 15]. Thus, further strategies to reduce GHG emissions (and other environmental impacts) from passenger transport are required. These can aim at a total reduction of pkm (e.g. through telecommuting or videoconferencing), a shift to more energy- or GHG-efficient transport modes (e.g. through multimodal mobility platforms), improving the energy efficiency of transport modes (e.g. by increasing utilization of transport vehicles and infrastructure), or at reducing GHG emissions embedded in transport infrastructure, vehicles and fuels.

2.2 *Impacts of ICT on Passenger Transport*

Since the late 1980s researchers have discussed the conceptual relationship between ICT (or telecommunications) and passenger transport, suggesting both, reducing and increasing effects of ICT use on passenger transport [10, 16, 17]. In a recent study, Pawlak et al. [9] review and conceptualize literature on the relationship between ICT and physical mobility. They provide a taxonomy of types of interactions between ICT and mobility (Fig. 2) that distinguishes "first- (lower) order relationships" (changes to specific journeys and activities) and "second- (higher) order relationships" (changes to other aspects of life that can have consequences for mobility).

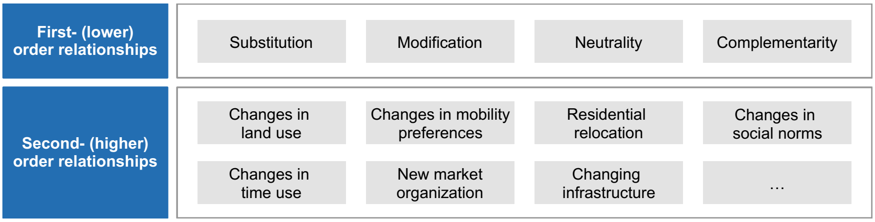


Fig. 2 Different levels of interaction between ICT and physical mobility, based on [9]

First- (lower) order relationships are:

- *Substitution*: Using ICT instead of making a trip (e.g. through videoconferencing)
- *Modification*: Changes to trip timing, destination, transport modes and transport mode choice, range and productivity of activities during trips (e.g. through multimodal mobility platforms)
- *Complementarity*: Generation of additional trips that would not take place without ICT (e.g. social media can create the desire to travel [16])
- *Neutrality*: None of the three effects above. Pure neutrality means that no effect occurs, net neutrality means that effects cancel each other out.

Second- (higher) order relationships are for example ICT-induced changes to land-use and locational decisions (e.g. residential), changes in mobility preferences, in lifestyles, time use and social norms (e.g. remote work), or new business models (e.g. car sharing) that have consequences for mobility patterns. These usually occur over a longer time horizon and are more difficult to assess.

Pawlak et al. [9] also identify challenges in studies on the relationship between ICT and mobility. First, with omnipresent ICTs the relationship between ICT and travel might be bidirectional (ICTs impacting travel, travel motivating ICT use), which is an underexplored hypothesis. Second, most studies use cross-sectional data which is less suitable than time-series data for assessing changes in ICT adoption and travel behavior over time. Third, the rapid development of ICTs limits the generalizability of results, specifically into the future. Third, various metrics to measure ICT use exist (e.g. time spent on the Internet, on social media, mobile phone usage) whose relevance can change over time. Finally, most studies focus on a specific context (e.g. economic situation, occupations, geographical scope, culture) which limits the transferability of results to other contexts.

Studies on the climate protection potential of ICT applications in passenger transport are also based on assumptions about the impacts of ICT on passenger transport. As stated in the introduction, many of these studies have several limitations and do not indicate under what conditions and to what extent these potentials can be exploited in a real-life setting because of several reasons:

First, many of these studies are based on overoptimistic assumptions about the GHG reducing effect of ICT applications. For example, GeSI [6] uses an estimate from the Gartner Hype Cycle that says that by 2030 50% of people living in OECD countries will adopt ride sharing; an assumption that is not in line with existing studies on ride sharing adoption [18]. Also, many studies assume that digital mobility services substitute conventional transport (e.g. videoconferencing replacing car travel) [4]. As stated above, various studies have already warned about complementary effects of ICT on transport [16, 17], which has been observed in the case of Uber [19]. Even if substitution effects outweigh complementary effects, the GHG impact of a digital mobility service depends on which transport mode it replaces. Third, ICT use has "diverse and complex impact patterns" that can lead to unintended effects such as rebound effects (see also Sect. 2.3) that occur if increases in energy efficiency lead to an increase in energy use [20, p. 826]. For example, telecommuting is often assumed to reduce emissions due to reduced commuting. However, it can also lead to increased leisure travel that can compensate for the commute-related GHG savings [21, 22].

Thus, even though large theoretical potentials to reduce passenger transport GHG emissions through ICT exist, there is a lot of uncertainty about the actual impacts in a real-life setting and the conditions under which potentials can be exploited. Empirical studies on the relationship usually focus on a specific ICT application (e.g. telecommuting or car sharing) in a specific context (e.g. geographic scope, prevalent mobility patterns), which is why their generalizability is limited [9].

2.3 Impacts of ICT on GHG Emissions

Various researchers have discussed GHG impacts of ICT, or environmental impacts of ICT in general [5, 23–26]. One of the most common taxonomies of environmental effects of ICT was developed by Hilty and Aebischer [23]. The framework (Fig. 3) distinguishes first-, second- and third-order effects, also called direct, enabling and systemic effects, that either reduce or increase environmental loads.

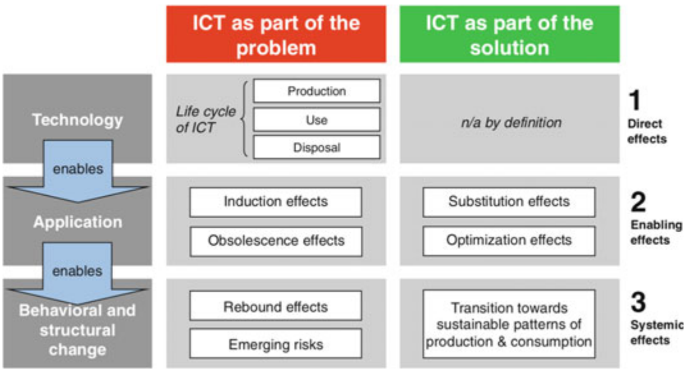


Fig. 3 A matrix of ICT effects [23]

Direct effects refer to the resources and energy requirements as well as emissions caused throughout the life cycle of ICT hardware, i.e. production, use and disposal. Direct effects are by definition unfavorable for the environment and are usually assessed with life cycle assessments (LCA) [27]. Enabling effects are the effects of applying ICT, which can be favorable (e.g. optimizing or substituting resource use) or unfavorable (e.g. inducing additional resource use) for the environment. Systemic effects describe the "long-term reaction of the dynamic socio-economic system to the availability of ICT services, including behavioral change (life styles) and economic structural change" such as rebound effects (p. 25). ICT rebound effects have received special attention in literature [28–30]. These occur if an ICT-induced efficiency improvement leads to an increase in demand such that the actual savings of an input factor are lower than the theoretical savings with constant demand [4, 31]. As ICTs are general purpose technologies that lead to efficiency increases (e.g. in terms of energy, cost or time expenditure) in many sectors, they are subject to rebound effects [30]. Income or time rebound effects, which occur if ICT leads to an increase in cost or time efficiency that allows for additional consumption of the same or other goods and services [32, 33], seem to be specifically relevant for ICT applications in passenger transport. For example, telecommuting, multimodal mobility platforms and navigation services can all reduce time and money spent on transport, which can be spent on travel for other purposes or other goods, services and activities [5]. While some empirical studies on rebound effects exist [21, 34, 35], they are highly context-dependent and the uncertainty in estimations of rebound effects is very high [36]. Still, in order to assess net impacts of ICT use in passenger transport on GHG emissions, all relevant effects (direct, enabling and systemic effects) have to be considered.

If we compare the taxonomy of Hilty and Aebischer to the taxonomy of Pawlak et al. [9] described in Sect. 2.2, it becomes apparent that the two frameworks have different perspectives. Hilty and Aebischer focus on ICT impacts in general, whereas Pawlak et al. focus on ICT applications in passenger transport. Thus, the first-order (direct) effects in Hilty and Aebischer's framework (the life cycle effects of ICT) are not part of the framework by Pawlak et al. The second-order effects (effects of ICT application) and third-order effects (systemic effects) in Hilty and Aebischer's framework roughly correspond to the first-order and second-order effects in the framework by Pawlak et al. respectively.

3 An Integrated Framework of ICT Impacts on Passenger Transport and GHG Emissions

Based on existing work on the relationship between ICT, transport and GHG emissions, we develop a conceptual framework that comprehensively describes the relationship between ICT applications in passenger transport and GHG emissions. We distinguish first-, second- and third-order effects of ICT use in passenger transport on GHG emissions in line with the taxonomy by Hilty and Aebischer. Even though we focus on GHG emissions, the framework could also be applied to further environmental impacts of passenger transport, e.g. particulate matter.

3.1 First-Order Effects

First-order effects describe the GHG impacts of providing the ICT hardware and software required for an ICT application in passenger transport. These include the entire life cycle, i.e. production/development of ICT hardware and software, their operation and disposal. By definition, the life cycle impacts of ICT hardware and software are unfavorable as they increase GHG emissions [23]. For example, videoconferencing equipment requires materials and energy during production, electricity for operating the equipment and data transmission, and energy for waste treatment at its end-of-life (e.g. for recycling). Warland et al. [37] estimated that a 4 h videoconference between Zurich and Paris is associated with 1.2 kg CO_2e considering the whole life cycle of the equipment.

3.2 Second-Order Effects

Second-order effects describe the impacts of applying ICT, i.e. ICT impacts on specific trips, and the consequences for GHG emissions. Figure 4 illustrates the relationship between ICT application in passenger transport and GHG emissions of passenger transport, which is explained in the following and based on [1, 9].

GHG emissions caused by passenger transport depend on the GHG intensity (CO_2e/pkm) of transport modes and the distances driven (pkm) by transport mode.

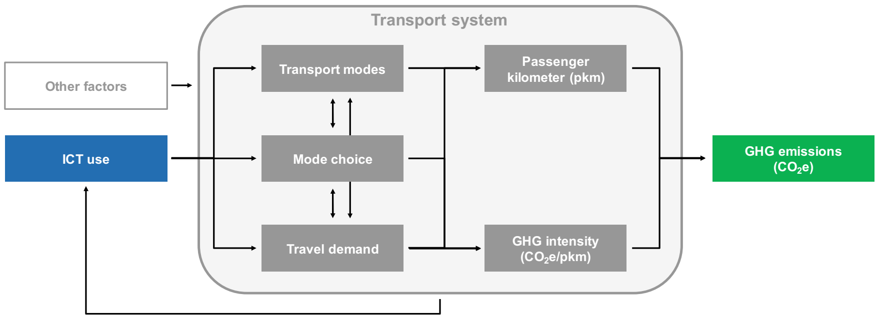


Fig. 4 Framework of the second-order relationship between ICT use, the transport system and GHG emissions. Various further representations of the transport system exist and ours is simplified and specified to the context of this article

GHG emissions per pkm depend on the energy intensity of transport modes (MJ/pkm) and the fuel carbon intensity (g CO_2e/MJ). The energy intensity can be reduced by reducing energy consumption of transport modes or increasing their utilization. The fuel carbon intensity can be reduced by using more GHG efficient energy carriers (e.g. electricity generated with renewable energies instead of diesel or gasoline).

Pkm by transport mode and average GHG emissions per pkm depend on total travel demand, availability and properties of transport modes and mode choice of individuals, who consider various transport mode-specific criteria, such as travel time, cost, comfort, reliability or hygiene [38]. Transport modes, travel demand and mode choice all interact, e.g. transport mode choice depends on availability of transport modes and travel demand.

ICT use impacts transport modes, mode choice and travel demand.

- *ICT impacts on transport modes*: Some ICT applications (e.g. automated driving, navigations systems, traffic control systems) change management and properties of transport modes (e.g. GHG emissions per pkm, travel time, price and comfort). ICTs also provide fully virtual transport modes (e.g. videoconferences) that directly compete with conventional transport modes in transport mode choice (e.g. virtual meetings are usually more time-efficient but less interactive than personal meetings).
- *ICT impacts on mode choice*: ICTs change access to transport modes (e.g. digital trip planners) and thereby influence transport mode choice. If ICT changes properties of transport modes (e.g. travel time, price), this also influences transport mode choice.
- *ICT impacts on travel demand*: ICT impacts travel demand, e.g. with respect to trip purposes, number and timing of trips. For example, automated driving can reduce cost per passenger kilometer and lead to additional car travel [39]. Virtual communication (e.g. via social media or videoconferences) can create the desire to travel [16].

The transport system also affects ICT services and their use, i.e. various ICT solutions are specifically provided to increase transport efficiencies (e.g. with respect to time or energy efficiency), creating a feedback loop [9, 16].

Additionally, other factors impact the transport system and thus the impact of ICT on the transport system. These include, for example, other technologies (e.g. e-mobility), policies (e.g. lockdowns during pandemics, congestion pricing), socio-economic and cultural aspects (e.g. car ownership, family status) [9, 40] cited from [41].

3.3 Third-Order Effects

Third-order effects of ICT use in passenger transport on GHG emissions depend on interactions with variables in the broader system and thus are only observable from a broader system perspective [33]. They describe long-term changes due to

ICT use, which are difficult to predict, but can have significant consequences for GHG emissions [4].

As stated by Pawlak et al. [9], ICT use can lead to broader structural changes, e.g. in lifestyles, time use, social norms, land-use, location decisions, mobility preferences, which have consequences for passenger transport and GHG emissions. These effects can also lead to changes in demand, construction and maintenance of transport infrastructures (e.g. streets, train stations), which are associated with GHG emissions.

While, rebound effects of ICT applications that lead to additional passenger transport are already captured within the second-order effects, indirect rebound effects of ICT use in passenger transport exist. For example, ICT use in passenger transport can also lead to an increase in consumption of other, non-travel, goods and services that cause GHG emissions. Income and time rebound effects seem to be relevant, especially if virtual mobility substitutes physical mobility. I.e. attending activities virtually can yield significant reductions in money and time expenditure for traveling [33], a phenomenon that has been observed during the ongoing COVID-19 pandemic [42]. Saved income and time will be spent on other goods, service and activities, that are also associated with GHG emissions [33]. Even though, income and time rebound effects seem to be specifically large for substituting effects of virtual transport modes, they also apply to other ICT applications in passenger transport that yield cost and time savings. Plus further rebound effects, e.g. macroeconomic rebound effects exist [29].

It depends on the size of each individual impact on all three layers, whether an ICT application or ICT use in general leads to a reduction or increase in passenger transport GHG emissions. Considering all impacts in the assessment of GHG impacts of ICT use is challenging as these are not unidirectional and many variables interact.

4 Example Application of the Framework

One of the major trends in the automotive sector is the development of driving automation systems supported by digital ICT. In the following we demonstrate the framework by qualitatively assessing first-, second- and third-order impacts of automated driving (AD) on GHG emissions.

4.1 First-Order Effects

A driving automation system is the "hardware and software that are collectively capable of performing part or all of the [dynamic driving task] on a sustained basis" [43, p. 3], such as LiDAR components to capture information on the surroundings of vehicles. In a life cycle assessment of a connected and automated vehicle, Gawron et al. [44] estimate that systems required for connectivity and automation (e.g. radars, sonars, computer) weigh 22.4 kg and cause 1,300 kg CO_2e throughout their life cycle.

If AD requires installing additional ICT equipment in roadside infrastructure (e.g. on traffic lights) and data exchange between vehicles and a central unit (e.g. exchange of data captured via sensors to analyze live traffic and as input for improving AD algorithms), the associated GHG emissions also need to be considered.

4.2 Second-Order Effects

In Table 1, we provide an overview of selected second-order effects of AD on transport modes, mode choice and travel demand, as well as their potential consequences for pkm and the GHG intensity of transport modes.

In general, AD can increase the attractiveness of car transport for several reasons (e.g. no parking spot required, other activities can be done during driving) and thus can lead to an increase in pkm of car transport. The GHG impacts of this effect depend on whether it is a substitution effect, i.e. car transport replaces transport with other modes, or if it is a complementary effect, i.e. it induces additional trips and thus increases total travel demand.

In Table 1, the effects are illustrated separately. In reality many effects interact and these interactions have to be considered in GHG assessments of AD. For example, platooning allows to reduce distances between driving vehicles and increases fuel economy due to slipstreams, but also increases capacity of roads, can reduce congestion and thus increase attractiveness of car transport. The assessment of such interactions requires systems thinking and complex systems modeling [33].

Table 1 Second-order effects of automated driving on transport modes, mode choice and travel demand incl. consequences for pkm and the GHG intensity

Effect	Effect on	Consequences for	
		pkm	GHG intensity
Lower fuel consumption due weight reductions of cars (due to higher safety) [47]	Transport mode	No change	Car: reduction
Increase in fuel economy of cars through platooning [47]	Transport mode	No change	Car: reduction
AD facilitates vehicle electrification (e.g. driving mode according to battery life) [49]	Transport mode	No change	Car: reduction[1]
Increased access for underserved population (e.g. for children or elderly) [50]	Travel demand/mode choice	Car: increase Other modes: potential reduction[2]	No change
Increased attractiveness of car as no parking spots are required [47]	Travel demand/mode choice	Car: increase Other modes: potential reduction[2]	No change
Increased capacity of roads due to shorter distance between vehicles [47]	Travel demand/mode choice	Car: increase Other modes: potential reduction[2]	No change
Reduced travel time cost because other activities can be done in parallel [47]	Travel demand/mode choice	Car: increase Other modes: potential reduction[2]	No change
Increase in attractiveness of public transport if AD is used to improve public transport (e.g. automated on-call buses) [51]	Travel demand/mode choice	Car: potential reduction[2] Public transport: increase	Public transport: reduction if utilization or fuel economy of vehicles increases

[1] Depends on the GHG intensity of the electricity mix.

[2] Depends whether car transport replaces other transport modes (or vice versa) or whether it leads to additional travel demand.

4.3 Third-Order Effects

Additionally, AD can lead to long-term structural changes, that have consequences for GHG emissions. In the following we summarize several effects:

- AD, electrification, right sizing of vehicles, shared mobility and increased availability of public transport can reduce costs of vehicle, infrastructure and transportation system operations by more than 40% [45]. Cost reductions of transport can induce additional consumption of other goods and services that are associated with GHG emissions (income rebound effect) [46].
- If AD reduces time spent in transport, e.g. if it reduces traffic congestion, saved time can be spent on other activities that cause GHG emissions (time rebound effect) [46].
- As AD increases the capacity of roads due to shorter distances between vehicles, this can reduce demand for built out of road infrastructure, whose construction and maintenance causes GHG emissions, and lead to changes in land use [46].
- AD can increase living density in city centers if less parking space is required [47]. In contrast, as AD reduces travel time cost, it can lead to more dispersed settlements if people will move further away from city centers or employer offices [47]. Both effects would impact transport-related GHG emissions (e.g. different mode choice and travel distances [33, 48]) and GHG emissions from other sectors (e.g. buildings).

Such third-order effects are long-term consequences of AD that can impact GHG emissions, but are difficult to predict as they depend on many variables in the broad socio-economic system. In order to assess the net impacts of AD on GHG emissions all direct and indirect impacts of this technology need to be considered.

5 Methodological Challenges of Assessments

Even when being aware of all potential direct and indirect effects of ICT on passenger transport and GHG emissions, researchers face various methodological challenges in such types of studies. In the following, we discuss important challenges based on our framework and existing research [9, 52].

First, while first-order effects have a relatively clear system boundary (the lifecycle of the ICT hardware and software required), defining the system boundaries for second- and specifically third-order effects is more challenging as they depend on the interaction of various variables in the transport and broader socio-economic system. In general, there is a difference between studies that focus on ICT use in general and on specific ICT applications. For example, studies on telecommuting often focus on substitution of commuting with telecommuting and do not capture the complementary effects that might occur due to ICT use in general (e.g. increase in collaboration across country borders that might induce travel). Thus, to correctly interpret the results of an assessment, the applied system boundaries, i.e. the effects and variables in and out of scope of the assessment, need to be considered.

Second, when assessing the impact of an ICT solution, it is important to consider that "ICT typically does not induce efficiency on its own, but only in a suitable technological, political or organizational context" [53, p. 138]. It can be challenging to assess ICT effects isolated from effects caused by new policies or other technologies [36, 54]. For example, automated driving and electric drivetrains are technologies that are developed in parallel and automated driving technology is mostly applied to electric vehicles. It is difficult to investigate the impact of automated driving technology in world without e-mobility.

Third, to assess the impact of an ICT application, a scenario with the use of the application needs to be compared to a reference scenario without the use of the application (often called baseline scenario) [36, 55]. However, "as soon as an ICT service is introduced, it is impossible to know how the reference activity would have developed without it" [55, p. 39]. The choices made when defining and estimating the reference scenario influence the results of the overall assessment. For example, the baseline "can be fixed as a base value for the moment at which an ICT service was introduced" or "be defined based on projections on how the situation without the ICT service would have evolved into the future" [55, p. 40].

Fourth, as Pawlak et al. [9, p. 15] illustrate, because ICTs rapidly evolve and provide new capabilities, the question arises to what extent "yesterday's research results are valid for use as a reference for future investigations". Also, different indicators can be used to measure the degree of ICT use [9], which differ among studies of ICT use in general (e.g. total time spent online, number of social media accounts) and studies of specific ICT applications (e.g. number of home office days for telecommuting, penetration of a specific app for ride sharing). The selection of indicators influences the results and their significance.

Fifth, the behavioral response of individuals to ICT use depends on their socio-economic background (e.g. having a child, cohabiting) and the policy context [9, 52]. Assessments need to identify and account for these characteristics. Still, as most studies take place in a specific context (e.g. region, types of occupation), the generalizability of study results is often limited [9].

Sixth, the relationship between ICT use and passenger transport is bidirectional, i.e. ICT changes passenger transport and passenger transport influences ICT use. Assumptions about the direction of causality need to be considered when designing assessments and interpreting their results [9].

Seventh, many studies use cross-sectional data [9] which is useful for investigating between-person or within-person differences [52]. Time-series data is required to assess changes in mobility behavior and consequences for GHG emissions over time, i.e. by comparing mobility patterns of individuals before and during the adoption of a specific ICT application or during times with different degrees of ICT use [52]. To include uncertain third-order effects in the assessment, longitudinal studies (e.g. over several years) are required [9].

Finally, while our framework focuses on GHG emissions, further environmental impacts of passenger transport exist which have to be considered as well and weighted against each other [52].

6 Conclusion

To date, transportation researchers have investigated the conceptual relationship between ICT and passenger transport. Researchers in the field of environmental effects of ICT have investigated the climate protection potential of ICT applications in passenger transport; however, they mainly focus on theoretical potentials and do not yet consider the complexity of interactions between ICT use, transport mode choice, properties of transport modes and travel demand. Based on existing literature from both fields, we developed a framework that provides a comprehensive understanding of the relationship between ICT use, passenger transport and GHG emissions and discuss methodological challenges of assessments of ICT impacts on passenger transport and GHG emissions. The framework distinguishes three types of effects:

- *First-order effects*: GHG impacts of producing, operating and disposing the ICT hardware and software required
- *Second-order effects*: Impacts of ICT use on specific trips, specifically on properties of transport modes, transport mode choice and travel demand and the consequences for pkm travelled and the GHG intensity of transport modes
- *Third-order effects*: Long-term structural changes due to ICT use (e.g. residential relocation, land use) incl. indirect rebound effects that have consequences for passenger transport and GHG emissions.

Methodological challenges in assessments of GHG impacts of ICT applications in passenger transport include the definition of system boundaries, the isolation of ICT effects from other effects, the estimation of reference scenarios without the introduction of an ICT application, the identification of suitable indicators to measure ICT use, the consideration of economic and socio-demographic characteristics of individuals, trade-offs between various environmental impacts, and inference of causality which can be bidirectional and is specifically challenging with cross-sectional data. In general, assessments of the impact of an ICT application always occur in a specific context that limits the generalizability of results to other contexts and in the future.

When conducting assessments, researcher should be aware of these potential effects and challenges in order to set a suitable assessment scope and design suitable assessments methods. Even if some effects are not in scope of an assessment or cannot be quantified, it is important to be aware of these in order to correctly interpret the results. We hope this article supports researchers in identifying relevant, direct and indirect, impacts of ICT on passenger transport on GHG emissions. Only if all relevant effects are considered, we can derive strategies for ICT use in passenger transport that lead to a net reduction of emissions, and put digitalization at the service of climate protection.

Acknowledgements This research is supported by the research program Sustainable Accessibility and Mobility Services—Mistra SAMS, funded by the Swedish Foundation for Strategic Environmental Research, Mistra.

References

1. Sims, R., Schaeffer, F., Creutzig, X., Cruz-Núñez, M.D., Dimitriu, D., Figueroa Meza, M.J., Fulton, L., Kobayashi, S., Lah, O., McKinnon, A., Newman, P., Ouyang, M., Schauer, J.J., Sperling, D., Tiwari, G.: Transport. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (2014)
2. European Environment Agency: Greenhouse gas emissions from transport in Europe (2021). https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12, last accessed 2021/04/12
3. United Nations: Mobilizing sustainable transport for development (2016)
4. Bieser, J., Hintemann, R., Beucker, S., Schramm, S., Hilty, L.: Klimaschutz durch digitale Technologien: Chancen und Risiken. Bitkom, Boderstep Institute for Innovation and Sustainability, University of Zurich, Berlin, Germany (2020). https://doi.org/10.5167/uzh-190091
5. Hilty, L., Bieser, J.: Opportunities and Risks of Digitalization for Climate Protection in Switzerland. University of Zurich, Zurich (2017). https://doi.org/10.5167/uzh-141128
6. GeSI: #SMARTer2030. ICT Solutions for 21st Century Challenges. Brussels, Belgium (2015)
7. WWF Sweden: The potential global CO2 reductions from ICT use Identifying and assessing the opportunities to reduce the first billion tonnes of CO2, (2008)
8. Malmodin, J., Bergmark, P.: Exploring the effect of ICT solutions on GHG emissions in 2030. In: Proceedings of EnviroInfo and ICT for Sustainability 2015. pp. 37–46. , Copenhagen, Denmark (2015)
9. Pawlak, J., Le Vine, S., Polak, J., Kopp, J.: ICT and physical mobility: State of knowledge and future outlook. Institute for Mobility research (ifmo), Imperial College London (2015)
10. Mokhtarian, P.: Telecommunications and Travel: The Case for Complementarity. J. Ind. Ecol. **6**, 43–57 (2002). https://doi.org/10.1162/108819802763471771
11. Moreau, H., de Jamblinne de Meux, L., Zeller, V., D'Ans, P., Ruwet, C., Achten, W.M.J.: Dockless E-Scooter: A Green Solution for Mobility? Comparative Case Study between Dockless E-Scooters, Displaced Transport, and Personal E-Scooters. Sustainability. 12, 1803 (2020) https://doi.org/10.3390/su12051803
12. Mobitool: Umweltdaten & Emissionsfaktoren von mobitool: Treibhauspotenzial (Update 2020, v2.1). (2020)
13. Eurostat: Modal split of passenger transport, (2021). https://ec.europa.eu/eurostat/databrowser/view/tran_hv_psmod/default/table?lang=en, last accessed 2021/04/12
14. Bates, J., Leibling, D.: Spaced out: perspectives on parking policy. RAC Foundation (2012)
15. Rodrigue, J.-P.: Transportation Modes, Modal Competition and Modal Shift. In: The Geography of Transport Systems. Routledge, New York (2020)
16. Mokhtarian, P.: A typology of relationships between telecommunications and transportation. Transportation Research Part A: General. **24**, 231–242 (1990). https://doi.org/10.1016/0191-2607(90)90060-J
17. Salomon, I.: Telecommunications and travel relationships: a review. Transportation Research Part A: General. **20**, 223–238 (1986). https://doi.org/10.1016/0191-2607(86)90096-8

18. Gerte, R., Konduri, K.C., Eluru, N.: Is There a Limit to Adoption of Dynamic Ridesharing Systems? Evidence from Analysis of Uber Demand Data from New York City. Transp. Res. Rec. **2672**, 127–136 (2018). https://doi.org/10.1177/0361198118788462
19. Tirachini, A.: Ride-hailing, travel behaviour and sustainable mobility: an international review. Transportation **47**, 2011–2047 (2020). https://doi.org/10.1007/s11116-019-10070-2
20. Erdmann, L., Hilty, L.: Scenario Analysis: Exploring the Macroeconomic Impacts of Information and Communication Technologies on Greenhouse Gas Emissions. J. Ind. Ecol. **14**, 824–841 (2010)
21. Bieser, J.C.T., Vaddadi, B., Kramers, A., Höjer, M., Hilty, L.M.: Impacts of telecommuting on time use and travel: A case study of a neighborhood telecommuting center in Stockholm. Travel Behaviour and Society. **23**, 157–165 (2021). https://doi.org/10.1016/j.tbs.2020.12.001
22. Vaddadi, B., Pohl, J., Bieser, J., Kramers, A.: Towards a conceptual framework of direct and indirect environmental effects of co-working. In: Proceedings of ICT4S 2020 – 7th International Conference on ICT for Sustainability. ACM, Virtual Conference. (2020). https://doi.org/10.1145/3401335.3401619
23. Hilty, L., Aebischer, B.: ICT for Sustainability: An Emerging Research Field. In: Hilty, L. and Aebischer, B. (eds.) ICT Innovations for Sustainability. pp. 3–36. Springer International Publishing (2015). https://doi.org/10.1007/978-3-319-09228-7_1
24. Berkhout, F., Hertin, J.: Impacts of Information and Communication Technologies on Environmental Sustainability: speculations and evidence. , Brighton, United Kingdom (2001)
25. Pohl, J., Hilty, L., Finkbeiner, M.: How LCA contributes to the environmental assessment of higher order effects of ICT application: A review of different approaches. Journal of Cleaner Production. **219**, 698–712 (2019). https://doi.org/10.1016/j.jclepro.2019.02.018
26. Bieser, J., Coroamă, V.: Direkte und indirekte Umwelteffekte der Informations- und Kommunikationstechnologie. Sustainability Management Forum, NachhaltigkeitsManagementForum. **29**(1), 1–11 (2021). https://doi.org/10.1007/s00550-020-00502-4
27. Itten, R., Hischier, R., Andrae, A.S.G., Bieser, J.C.T., Cabernard, L., Falke, A., Ferreboeuf, H., Hilty, L.M., Keller, R.L., Lees-Perasso, E., Preist, C., Stucki, M.: Digital transformation—life cycle assessment of digital services, multifunctional devices and cloud computing. Int. J. Life Cycle Assess. **25**, 2093–2098 (2020). https://doi.org/10.1007/s11367-020-01801-0
28. Börjesson Rivera, M., Håkansson, C., Svenfelt, Å., Finnveden, G.: Including second order effects in environmental assessments of ICT. Environmental Modelling & Software. **56**, 105–115 (2014)https://doi.org/10.1016/j.envsoft.2014.02.005
29. Coroamă, V., Mattern, F.: Digital Rebound - Why Digitalization Will not Redeem us our Environmental Sins. In: ICT4S2019. 6th International Conference on Information and Communication Technology for Sustainability. CEUR Workshop Proceedings, Lappeenranta, Finland (2019). http://ceur-ws.org/Vol-2382/
30. Gossart, C.: Rebound Effects and ICT: A Review of the Literature. In: ICT Innovations for Sustainability. pp. 435–448. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-09228-7_26
31. Madlener, R., Alcott, B.: Herausforderungen für eine technisch-ökonomische Entkopplung von Naturverbrauch und Wirtschaftswachstum. Enquete-Kommission "Wachstum, Wohlstand, Lebensqualität" des Deutschen Bundestages (2011)
32. Bieser, J., Haas, D., Hilty, L.: VETUS - Visual Exploration of Time Use Data to Support Environmental Assessment of Lifestyles. In: ICT4S2019. 6th International Conference on Information and Communication Technology for Sustainability. CEUR Workshop Proceedings, Lappeenranta, Finland (2019). http://ceur-ws.org/Vol-2382/
33. Bieser, J., Hilty, L.: Conceptualizing the Impact of Information and Communication Technology on Individual Time and Energy Use. Telematics and Informatics. 101375 (2020). https://doi.org/10.1016/j.tele.2020.101375
34. Amatuni, L., Ottelin, J., Steubing, B., Mogollón, J.M.: Does car sharing reduce greenhouse gas emissions? Assessing the modal shift and lifetime shift rebound effects from a life cycle perspective. Journal of Cleaner Production. **266**, 121869 (2020). https://doi.org/10.1016/j.jclepro.2020.121869

35. Llorca, M., Jamasb, T.: Energy efficiency and rebound effect in European road freight transport. Transportation Research Part A: Policy and Practice. **101**, 98–110 (2017)
36. Bieser, J., Hilty, L.: Indirect Effects of the Digital Transformation on Environmental Sustainability: Methodological Challenges in Assessing the Greenhouse Gas Abatement Potential of ICT. In: Penzenstadler, B., Easterbrook, S., Venters, C., and Ahmed, S.I. (eds.) ICT4S2018. 5th International Conference on Information and Communication Technology for Sustainability. pp. 68–81. EasyChair, Toronto, Canada (2018). https://doi.org/10.29007/lx7q
37. Warland, L., Hilty, L.: Factsheet: Business Travel. University of Zurich (2016)
38. Muro-Rodríguez, A.I., Perez-Jiménez, I.R., Gutiérrez-Broncano, S.: Consumer Behavior in the Choice of Mode of Transport: A Case Study in the Toledo-Madrid Corridor. Front Psychol. 8, (2017). https://doi.org/10.3389/fpsyg.2017.01011
39. Hörl, S., Becker, F., Dubernet, T., Axhausen, K.W.: Induzierter Verkehr durch autonome Fahrzeuge: Eine Abschätzung. Schweizerische Vereinigung der Verkehrsingenieure und Verkehrsexperten (SVI), Bundesamt für Strassen (2019)
40. Johansen, L.: Lectures on Macroeconomic Planning. North-Holland Publishing Company, Amsterdam (1978)
41. Höjer, M.: What is the Point of IT? : Backcasting urban transport and land-use futures, http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-3016, (2000)
42. Our World in Data: COVID-19: Google Mobility Trends, https://ourworldindata.org/covid-google-mobility-trends, last accessed 2021/04/12
43. SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. (2016)
44. Gawron, J.H., Keoleian, G.A., De Kleine, R.D., Wallington, T.J., Kim, H.C.: Life Cycle Assessment of Connected and Automated Vehicles: Sensing and Computing Subsystem and Vehicle Level Effects. Environ. Sci. Technol. **52**, 3249–3256 (2018). https://doi.org/10.1021/acs.est.7b04576
45. Fulton, L., Mason, J., Meroux, D.: Three Revolutions in Urban Transportation. UC Davis, ITDP (2017)
46. Bieser, J., Salieri, B., Hischier, R., Hilty, L.: Next generation mobile networks: Problem or opportunity for climate protection? University of Zurich, Empa, Swisscom, Swisscleantech, Zurich, St. Gallen, Switzerland (2020). https://doi.org/10.5167/uzh-191299
47. Anderson, J.M., Kalra, N., Stanley, K.D., Sorensen, P., Samaras, C., Oluwatola, T.A.: Autonomous Vehicle Technology: A Guide for Policymakers. Rand Corporation (2016)
48. Mokhtarian, P.: If telecommunication is such a good substitute for travel, why does congestion continue to get worse? Transportation Letters. **1**, 1–17 (2009). https://doi.org/10.3328/TL.2009.01.01.1-17
49. Taiebat, M., Brown, A.L., Safford, H.R., Qu, S., Xu, M.: A Review on Energy, Environmental, and Sustainability Implications of Connected and Automated Vehicles. Environ. Sci. Technol. **52**, 11449–11465 (2018). https://doi.org/10.1021/acs.est.8b00127
50. Harper, C.D., Hendrickson, C.T., Mangones, S., Samaras, C.: Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. Transportation Research Part C: Emerging Technologies. **72**, 1–9 (2016). https://doi.org/10.1016/j.trc.2016.09.003
51. Perret, F., Fischer, R., Bruns, F., Abegg, C., de Haan, P., Straumann, R., Hofer, M., Raymann, L.: Einsatz automatisierter Fahrzeuge im Alltag - Denkbare Anwendungen und Effekte in der Schweiz. EBP, BaslerFonds, Schweizerischer Städteverband, Zürich (2018)
52. Bieser, J.C.T.: A time-use approach to assess indirect environmental effects of information and communication technology: time rebound effects of telecommuting, https://www.zora.uzh.ch/id/eprint/191486/, (2020). https://doi.org/10.5167/uzh-191486
53. Coroamă, V., Schien, D., Preist, C., Hilty, L.: The Energy Intensity of the Internet: Home and Access Networks. In: ICT Innovations for Sustainability. pp. 137–155. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-09228-7_8

54. Bergmark, P., Coroamă, V., Höjer, M., Donovan, C.: A Methodology for Assessing the Environmental Effects Induced by ICT Services. Part II: Multiple Services and Companies. In: Proceedings of ICT4S 2020 – 7th International Conference on ICT for Sustainability. ACM, Virtual Conference. (2020). https://doi.org/10.1145/3401335.3401711
55. Coroamă, V., Bergmark, P., Höjer, M., Malmodin, J.: A Methodology for Assessing the Environmental Effects Induced by ICT Services. Part I: Single services. In: Proceedings of ICT4S 2020 – 7th International Conference on ICT for Sustainability. ACM, Virtual Conference. (2020). https://doi.org/10.1145/3401335.3401716

Sustainability Impacts and Indicators

Shareable Goods and Impacts on Consumption; The Case of Digital Sharing Platforms

Maria J. Pouri

Abstract Digital platforms promote the shared consumption of a variety of material and immaterial goods. As they grow and appeal to more consumers, the new practices and patterns of sharing can have both desirable and undesirable impacts from a sustainability point of view. The present paper takes a closer look at the various impacts of shared consumption practiced in the Digital Sharing Economy (DSE) with the aim to propose a guideline for assessing the DSE's sustainability impacts. The guideline builds on a typology of shareable goods and a classification of different positive and negative environmental impacts of sharing. By considering different consumption scenarios (based on consumers' behavioral responses, where applicable), the study develops a conceptual framework for the DSE's impacts on the environment. This is also an extension of the general impact categories of Information and Communication Technology (ICT) which already exist in the literature. In addition to the impact of consumption, the concepts of the impact of provision, impact of access, and impact of maintenance are introduced as the four main evaluations comprising the sustainability assessment of sharing platforms. The considerations addressed can be helpful in delineating further evaluations of the sustainability of various instances of digital sharing systems.

Keywords Shared consumption · Digital sharing economy · DSE · Sharing platforms · Sustainability impacts · Sustainable consumption · Information and communication technology · ICT

1 Introduction

Sharing has become a ubiquitous behavior for today's consumer [1, 2]. In comparison with its traditional form, sharing has now substantially enhanced access to

M. J. Pouri (✉)
University of Zurich, 8050 Zurich, Switzerland
e-mail: pouri@ifi.uzh.ch

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_16

resources[1] as it has grown beyond the limits of local and social proximity. Sharing via digital platforms is now possible among people within and across widely distributed communities without constraints of time, place, communication, and coordination [3].

Digital Information and Communication Technology (ICT) promotes shared consumption of an expansive variety of material and immaterial resources by connecting providers of goods and those who seek them in large-scale (peer) communities. Digital platforms (i.e., applications and websites) coordinate sharing processes and optimize resource allocation by effectively matching supply and demand [4, 5] with very low transaction costs [4, 6, 7]. Platform-enabled services have increased choice options for consumers [8] as they are typically affordable, more efficient, convenient, and accessible than their equivalents in conventional marketplaces (compare, e.g., Airbnb offerings with hotel stays) [3]. As a result, sharing has gained traction over other modes of consumption, i.e. ownership and conventional renting [9, 10]. In some cases, this has been to the point where patterns of consumption and production have developed more rapidly, have been influenced [11–13], and have sometimes significantly changed, for example in hospitality and transportation [14, 15].

In contrast to industrial societies—which are driven by maximizing production—and in alignment with the post-industrial phase—or "second machine age" [16]—and the service economy where "almost any activity, asset, and skill can be bought on competitive markets" [17, p. 22], the sharing economy has the potential to promote sustainable growth and development [18, 19]. Sustainable development is a crucial concept. It expresses that the needs of people are met without overshooting Earth's ecological ceiling [20]. Given that the global overuse of natural resources for production and consumption processes is driving the planet's environmental systems to their breaking point, if not beyond, it is important that these boundaries are not breached; otherwise, humankind would risk adversely affecting ecosystems, the economy, and society [21]. Therefore, it is a relevant effort to reflect on the potential implications of the digital sharing economy (DSE) for sustainability [11] as it has changed, at its core, the way we consume goods.

Although the emergence of the sharing economy was accompanied by promises of environmental benefits and sustainable patterns of consumption [22, 23], it does have far-reaching impacts on these areas [11, 24]. One way to develop a critical perspective on the environmental impacts of shared consumption can be based on firstly identifying the type of the resources that are shared and secondly investigating how sharing those resources can affect the sustainability of their consumption as well as, where relevant, other consumption practices and systems that they may promote [25]. Focusing on these questions, the present paper aims to propose a guideline for assessing the sustainability impacts of digital sharing platforms. In the course of the analysis, the category of ICT impacts on the environment—which already exists in

[1] Throughout this paper, the words "resource" and "good" have been used almost interchangeably to refer to assets or things that are being shared via platforms in the DSE. "Natural resources" as addressed in sustainability and environmental studies are clearly specified as such.

the literature [26]—is expanded in the context of the DSE as a special application of ICT. Finally, a guideline for the activities that need to be incorporated into the assessment of impacts is derived.

2 The Digital Sharing Economy

The sharing economy has been described in various ways (see [27–29] for a review of the definitions). The current literature lays out a valuable body of knowledge that includes myriad interpretations and conceptualizations of the phenomenon. The DSE explains the recent changes in sharing practices and patterns over the course of the digital transition of society, or digitalization. In particular, the DSE allows for transitional modifications—which have not been typically associated with the traditional notion of sharing, and may therefore seem counterintuitive—to capture the nature and purpose of sharing via digital platforms. For example, it allows for both for-profit and non-profit as well as free and price-based sharing. To illustrate, platforms such as Uber, Airbnb, Couchsurfing, Zipcar, BlaBlaCar, hOurworld, SwapHomes, EatWith, Peerby and many others can all fall within the definitional domain of the DSE although they are highly diverse in their business models.

By definition, the DSE "*is a class of resource allocation systems based on sharing practices which are coordinated by digital online platforms and performed by individuals and possibly (non-) commercial organizations with the aim to provide access to material or immaterial resources. Digital sharing systems operate in the space between traditional sharing and the formal market economy*." [3, p.130]. This definition rests on three main dimensions:

- The technical aspect of sharing refers to a resource's capacity or property of serving multiple consumers, independent of the mechanism that makes this shared use possible. This aspect discusses the *shareability* of resources.
- The social aspect of sharing is about the patterns of social interactions and economic transactions through which the act of sharing takes place. This aspect addresses *sharing practices.*
- The coordination aspect addresses the intermediation required for sharing. Digital platforms enable the implementation of coordination mechanisms that organize the social and technical aspects of sharing with exceptional efficiency. This aspect deals with the *operationalization and scalability* of sharing.

In the DSE, as in any economy, people make purposeful choices about their resources and interact with others when they make these choices. Digital platforms have enhanced these choices both technically and socially. In other words, they have enabled incorporating a great variety of shareable resources in the technical sense and realizing sharing beyond the limits of small groups and personal relationships in the social sense [3].

The present paper focuses on the technical aspect of sharing and investigates how it affects consumption from a sustainability point of view. Analyzing the diversity of shareable resources is important for determining what is shareable in principle and what can be offered as a service via sharing platforms. This analysis can then be supplemented by examining the impacts of shared consumption.

3 Shareable Goods

In the context of sharing, what is relevant is that there is slack capacity of (privately) owned goods relative to the need of their owner and that sharing will clear this surplus [30]. Sharing via digital platforms enhances this clearance since digital coordination allows for more participants to be involved, larger pools of resources to be created, and a greater diversity of them to be included [3].

Generally, goods can be either material (tangible, physical) or immaterial (intangible, non-physical) [31]. In relation to consumption—where sharing is itself a mode of consumption [9, 32, 33]—the durability or non-durability of goods is also relevant. A durable good is one that can be used repeatedly or continuously over time, whereas a non-durable (or consumable) good is a single-use good to meet a need (e.g., food, drink, or firewood that can be used only once to satisfy hunger, thirst, or a need for heating) [34].

Concerning shared consumption, resources can be grouped into four broad categories: durable material goods, consumable (i.e., non-durable) material goods, durable immaterial goods, and consumable immaterial goods [35]. Pouri and Hilty [3] further distinguish between the types in each category: durable material goods with *idle* or *free* capacity, consumable material goods of *sufficient* or *abundant* availability, *durable information* and *competence* as durable immaterial goods, and *consumable information* and *time* as consumable immaterial goods.

Sharing puts the unused capacity of a durable material good into productive use. When a durable material good is shared, its unused capacity serves additional demands either sequentially (using *idle* capacity) or concurrently (using *free* capacity). Making a distinction between idle and free capacity is meaningful because the two can have different sustainability impacts when used. Take for example, the two technical options for sharing a car as a durable physical resource: activating the car's idle capacity via an Uber standard service, or occupying its free capacity (i.e., empty seats) via BlaBlaCar's carpooling service. Although both options can optimize resource utilization by increasing the car's lifetime efficiency—which is defined as the total number of functional units delivered over the service life of a technical resource (the car) that itself is connected to the inflow of natural resources required for producing it [36]—Uber rides, in particular, can result in unfavorable effects such as rebound effects too (see Sect. 4).

Unlike durable material goods, consumable material goods (such as food) cannot be accessed repeatedly because they are depleted once they are used. Nevertheless, the act of consuming them can be shared (such as shared consumption of food in groups). Consumable material goods are shared either to be sufficient or there is an abundance of them that exceeds the demand and this excess is shared, for example, to avoid waste.[2]

Immaterial goods can be described in terms of their reliance on the person possessing them. That is, they can be either detached from their bearer or they are inseparable from them. The former refers to *information goods*, and the latter subsumes *competence* and *time*. Competence—defined as a person's underlying characteristic which results in effective and/or superior performance of a task [39]—and time are inextricably bound to the individual who possesses them. In other words, both types require the presence of the provider in order for the act of sharing to materialize (consider the skill of fixing a tool for competence, and the time spent on doing a particular activity that require the provider to be involved in sharing). Nonetheless, competence is imperishable in the sense that it does not deplete or decay—at least not under normal conditions—while time is spendable in the sense that the time spent executing an activity is gone and cannot be re-accessed or re-gained (as opposed to a durable good).

Information goods are immaterial goods that are person-dependent. They can be further divided based on whether they are durable or non-durable. Durable information goods are information products with timeless content such as software products, music, and scientific content. Consumable information goods are volatile pieces of information that lose timely value quickly, such as information on the traffic situation.

Table 1 depicts the different types of shareable resources along with their property of being shareable, or "shareability."

[2] In economics and social sciences, the concepts of sufficiency and abundance are discussed as two status of resources influencing human interaction and motivating exchange behaviors [37, 38]. There is also scarcity as the third status, which is not relevant to sharing.

Table 1 A typology of shareable goods (source [25])

Broad classification of resources	Further classification	Type of "shareability"	Description	Example of sharing service and platform
Material	Durable	Free capacity	Sharing the free capacity of a durable, material resource enhances its utilization by serving multiple demands *at the same time*	Occupying free seats of a car via a carpooling service offered by platforms such as BlaBlaCar
		Idle capacity	Sharing the idle capacity of a durable, material resource reduces its idle time—i.e., the time when the resource is not being used—and enhances its utilization through serving more demands *in sequence*	Sharing a household tool via platforms such as Peerby
	Consumable	Sufficient availability	Sharing consumable material goods in order to be sufficient for a group of consumers	Food sharing via platforms such as OLIO and EatWith[a]
		Abundant availability	Sharing consumable material goods that are of abundant availability (e.g., to avoid wasting the surplus)	

(continued)

Table 1 (continued)

Broad classification of resources	Further classification	Type of "shareability"	Description	Example of sharing service and platform
Immaterial	Durable	Durable information goods	Refers to information products with timeless content—such as software products, music, or educational content—that allow for repeated use without rapid degradation in value, applicability, effectiveness, etc.	Providing access to durable information goods in a sharing community[b]
		Competence	Sharing competence which is the sufficiency of knowledge, skill, etc. to perform a task effectively	Helping others with tasks that require special skills. TaskRabbit is a familiar example of this category
	Consumable	Consumable information goods	Refers to volatile information—i.e., information that loses value rapidly over time—such as real-time traffic information	Platforms on which information is provided through community-based participation[c]
		Time	Sharing time as a resource which cannot be stored. This category refers to time as the only resource required to be shared in performing a task (no expertise or skill is required)	Platforms that enhance connections in neighborhoods, e.g., through encouraging collaborative shopping. Pool.farm is a platform of this type that supports shopping tasks for a group of people

[a]Distinguishing between the two types may not be possible without knowing the intention of the provider to share food with others
[b]Prospective model
[c]Platforms such as YouTube are not considered instances of sharing in the context of the DSE

4 Impacts of Shared Consumption

The environmental impacts of shared consumption in the DSE can be described as follows [25]:

Optimization effect: the case where sharing optimizes the consumption of a resource through enhancing its utilization. In general, the optimization effect of sharing is based on increasing the lifetime efficiency of a resource—i.e., the total number of functional units delivered per life-cycle-wide resource use of all product/service systems involved in the consumption of a particular resource [36]. The optimization effect can be subdivided depending on whether it occurs with respect to the same or another resource/service system:

Direct optimization effect: when the number of functional units (as used in the Life Cycle Assessment (LCA) method[3]) produced by a good and its complementary consumption increases. Optimization might take place across the lifespan of a physical (durable) resource or per service it delivers. Employing Uber's standard ride service or BlaBlaCar's carpooling rides to activate a car's idle capacity and occupy its free capacity, respectively, are relevant instances for optimization through sharing. For ride services the number of functional units increases only over the service life of the car (because it does not serve multiple riders simultaneously), whereas carpooling allows multiple riders to travel in the same car at the same time, whereby the number of passenger kilometers increases both per service and throughout the car's service life.

Indirect optimization effect: the situation in which sharing a resource results in optimizing the production of another resource. For example, utilizing a shared space or community farm for raising crops and preparing food can optimize food production processes [41]. Another example is shared local three-dimensional (3D) printer facilities, which have the potential for sustainable, resource-efficient, and optimized production [41, 42].

Overall, putting underutilized assets to use can alleviate pressure on production and consumption [43]. While the optimization effect of sharing is undeniably beneficial and characterized as a positive effect, its efficiency improvements and gains can be partially or even totally compensated by behavioral changes in consumption, as in the case of rebound effects [44], which will be examined further below.

Induction effect: Sharing a resource stimulates the consumption of additional resources that are linked to it (or complementary resources). For example, in the case of car sharing—in which a car is the primary resource for shared use—the usage of fuel and transportation infrastructure as supplemental consumptions for driving a car increases with sharing (see [45, 46]).

[3] LCA uses a bottom-up modeling approach that incorporates environmental assessment from the life cycle perspective. It requires the collection of data on input materials, energy carriers, land resources, and related emissions and wastes for each small process unit, and is better for evaluating direct impacts [40].

Induction effects are not directly associated with efficiency gains; they are about an increase in consumption [26]. Although induced consumption is generally considered a negative environmental effect [40, 47], the assessment of sharing models should include the marginal consumption of complementary resources per functional unit. If this marginal consumption decreases, sharing can still have a resource-saving benefit [36].

Rebound effect: the rebound effect emerges as a result of increased efficiency. The efficiency gains (including time and money) from using shared goods may stimulate consumption [11, 36, 48, 49] and increase demand in a way that outweighs the initial efficiencies. The rebound effect can be differentiated into two types of behavioral and economic responses to improved efficiency:

– **Direct rebound effect**: the case of rebound effects occurring within the same resource or service system. For example, Airbnb's affordable accommodation enables guests to use more of the platform's services.
– **Indirect rebound effect** (also called savings-induced consumption [40]): the case of the rebound effect occurring outside the resource or service system. Indirect effects are secondary increases in consumption of other goods and services when the residual savings from the consumption of primary goods and services are made available. Given that sharing typically requires less investment [50] and can even generate income for the provider of the good [44], the indirect rebound effects of sharing are noticeable [49, 51, 52].

In particular, attention should be given to time rebound effects when considering the sustainability of technology-based efficiency gains [53]. Time rebound comes from a decline in the time needed to acquire and use a service, lowering the explicit and implicit costs of time [54]. This saved time may be spent on consuming more of the same service or on some other high-resource-intensity activity.

There are, however, beneficial effects as discussed by Santarius and Soland [55] that outweigh rebound effects. Beneficial effects are explained through the concepts of "improved control over frugal use" of an efficiency-improved technology and "increased responsibility" for environmental protection as a result of the adoption of an efficiency-improved technology [55]. Since they are identified in relation to rebound effects, beneficial effects here are not considered as primary impacts of shared consumption; nevertheless, they are worth consideration.

Substitution effect: When the shared resource or the sharing service replaces the use of another resource or service, that amounts to substitution. Two sub-categories can be considered for the substitution effect:

– **Direct substitution effect**: substitution occurs in the same resource/service system. For example, if car sharing offered by Zipcar replaces the conventional car rental services provided by businesses in established industries.
– **Indirect substitution effect** (also called cross-activity substitution effect [56]): substitution that occurs in regard to other service systems. For example, when a shared mobility scheme (Zipcar, Uber, bike sharing, etc.) replaces low-impact

and green modes of transport such as public transportation, private bicycles, and walking [57, 58]. Depending on whether the new consumption patterns substitute for option with greater or smaller environmental impact, substitution effects can have different implications for sustainability.

Degradation effect (or accelerated asset depreciation (shortened lifetime) [23], or premature deterioration [57]): the condition whereby the intensified utilization of a shared resource leads to its more rapid degradation [59].

For example, in a case study of dockless e-scooters Moreau and colleagues [57] highlight the relevance of e-scooters' lifespan in measuring their environmental effects. They found that dockless e-scooters are not yet environmentally beneficial. To make significant improvements, the incidence of vandalism must be reduced and the lifespan of e-scooters should reach at least 9.5 months in order to be considered a green transportation option [57]. This is despite the fact that, according to an examination of an e-scooter sharing system [60], the scooters' average lifespan was only 28.8 days.

Figure 1 structures the above effects into the typology of the environmental effects of the DSE.

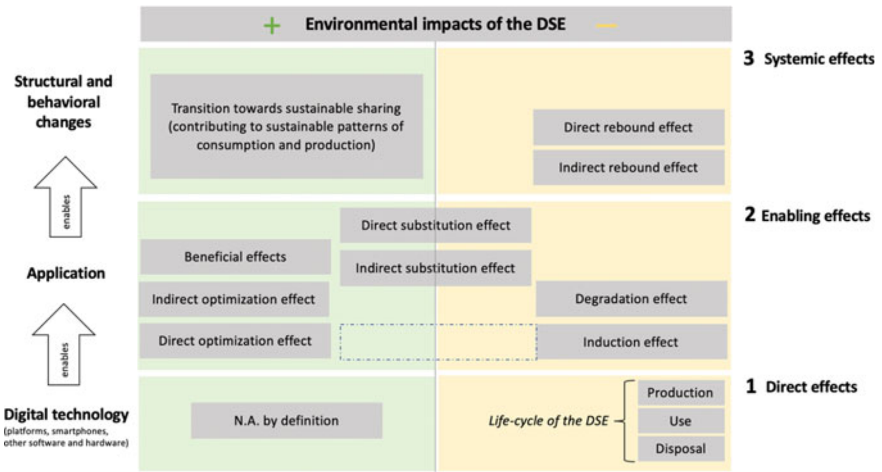


Fig. 1 A conceptual framework for the environmental effects of the DSE (*Source* adapted from the matrix of ICT effects in [26]). Direct effects address the sustainability impacts of production, use and disposal of ICT hardware and software. Enabling effects refer to both favorable and unfavorable impacts of actions enabled by the application of ICT in production and consumption processes. Systemic effects are the long-term impacts of ICT on life styles, value systems, and social and economic structures

Although the most important impacts of the DSE are found at Level 2 and 3 of the model in Fig. 1, the first-order impacts should also be included to be comprehensive. In general, the first-order impacts of ICT are the direct environmental impacts of the production, use, and disposal of ICT, including the consumption of material and energy throughout the entire life cycle of the hardware and software.

Concerning the DSE—as an example of the application of ICT—it can be inferred that the total life-cycle impacts of a digital sharing platform is trivial compared to the impacts of the resources being shared, be they positive or negative. For example, compared to the life-cycle impact of a single car—which generates as much CO_2 for each 200–300 km driven as a smartphone over its entire life cycle [61, 62]—the smartphones and servers needed to organizing car sharing have a small life-cycle impact [11]. Furthermore, only a small part of a smartphone's or other end-user device's life-cycle impact can be attributed to participation in sharing platforms because a smartphone is used for a variety of purposes and applications.

Nevertheless, there are cases where first-order effects can become noticeable. Some sharing schemes require the deployment of ICT devices as a resource modification so that it may be utilized in a sharing system. These are often seen in business-to-consumer sharing models. For example, Zipcar has installed digital card readers on its vehicles to lock and unlock them. To the best of the author's knowledge, no studies on the life-cycle assessment and environmental impacts of the production of such devices in the context of their usage in the DSE have yet been reported in the literature. However, the studies on the life cycle of ICT hardware can be consulted as relevant sources with reliable approximation (e.g., [63, 64]).

5 Discussion: An Assessment Guideline

The environmental impacts of the DSE are not caused only by the consumption of shareable resources. There are additional effects that are not directly linked to the act of sharing, but are identified in activities and processes that are necessary for sharing and keeping the sharing system in operation. These activities pertain to the provision of shared goods, access to shared goods, and maintenance of shared goods.

Impact of provision: Although the mantra of the sharing economy is to make better use of underutilized assets, investing in assets that will be used in a sharing system is viable too. This is in particular relevant to the provision of durable physical assets (though it also applies to consumables), regarding whether the supply comes from *existing* or *invested* resources [36]. Existing resources are resources owned by a (private) provider that already makes them available, and upon entering into the sharing system, they will become accessible also to others. Invested resources refer to a status in which a resource is provided to be used only in a sharing system and have no further functions for the owner. The important difference between the two modes of provision is that the latter can involve *new* production and manufacturing for the purpose of sharing (such as the production of e-scooters and e-bikes for sharing), hence those products will generate production-related impacts [57, 65] and will often be subject to faster degradation due to intensified use and less careful behavior by the user [60, 66]. Overall, with respect to production and improving material efficiency, sharing business models should be examined for increasing unnecessary production and generating inefficient overcapacity of goods that can counter their sustainability potential [67].

Impact of access: In addition to provision, access to a resource can also be a source of environmental impacts. In a study of consumer-to-consumer asset sharing, Martin et al. [58] identify logistics and transportation impacts that arise from transporting goods from the provider to the user (or vice versa) based on the transportation mode—high-impact or low-impact transportation method (private car, bus, biking, and walking)—and the distance between provider and user. Transportation in this case can be generally referred to as "access". For example, in order to help someone out with repairing a tool—which is a positive practice (sharing the skill of repairing) because it extends the lifetime of that tool—the provider may choose a high-impact mode of consumption that can equal or outweigh the benefit of the sharing service. Such activities are worth consideration as they are directly related to the way providers and consumers interact to grant and obtain access to a resource for shared use.

Impact of maintenance: Another source of impact is maintenance activities. Maintenance is the key instrument for prolonging the lifespan of durable physical assets [57]. It can affect sustainability directly through the execution of maintenance activities, and indirectly through production processes and final product quality as a result of the efficiency/inefficiency of maintenance [68]. Therefore, "maintenance concepts should improve the level of sustainability in manufacturing through innovative management practices, based on the wider perspective of asset life cycle, the extensive adoption of predictive measures and the further exploitation of potentials available from new enabling technologies (such as ICT)." [68, p. 92]. In an LCA study on shared dockless e-scooters, Hollingsworth et al. [65] found that without efforts to modify and improve distribution, collection, use, and charging strategies, the integration of e-scooter sharing schemes in urban transportation systems can even exacerbate emissions and environmental impacts.

To conclude, an assessment of the environmental impacts of the DSE should consider the following components:

- identifying the primary resource for shared use (Table 1),
- identifying complementary resource(s),
- identifying other resource/service systems that may be affected by sharing,
- evaluating the impacts of the provision of shared resources,
- identifying and evaluating the impacts of the consumption of shared resources (Fig. 1),
- evaluating the impacts of access to shared resources, and
- evaluating the impacts of the maintenance of shared resources.

This set of assessment activities emphasizes considering the different types and sources of the DSE's environmental impacts when conducting sustainability assessment as they might be overlooked or insufficiently addressed by conventional approaches. The aim is to draw attention to the need to integrate them in any sustainability evaluation approach used. Obviously, not all of the four impacts might be relevant to all platforms and sharing systems; the degree of their significance may not be equal, either.

6 Conclusion and Outlook

The deployment of digital platforms has enabled sharing to evolve remarkably in its technical sense, that is, in the diversity and volume of shareable goods that can be supplied for common use. Thanks to platforms, we can now share literally anything that is shareable, including food, time, space, skills, tools, accommodation, cars, etc. Shared consumption appeals to consumers because of the efficiencies provided by digital platforms. Yet the actual sustainability impacts of digitally enabled sharing are more complicated, and research is limited.

This study presented a typology of goods that can be offered for shared used via platforms (Table 1). It further elaborated on the potential sustainability impacts of shared consumption (Fig. 1). Thereafter, the implications of shared consumption, provision of resources, access to resources, and their maintenance were discussed and consolidated in the context of an assessment guideline for analyzing the environmental impacts of the DSE (Sects. 4 and 5).

It was also explained that the impacts of sharing can vary greatly from platform to platform. Likewise, not all the impacts (consumption, provision, access, and maintenance) are equally sensitive and remarkable across all sharing models. In addition, consumers' behavioral responses and changes are important in shaping and actualizing the positive or negative impacts of sharing.

Last but not least, individual providers and consumers of shared goods can make a difference through (changing) their consumption choices and improving their environmental awareness. In addition to the economic benefits of sharing and its opportunities for developing social relationships, participants may also value sustainability to a high extent. Platforms need to incorporate sustainability values in their services in order to have targeted the needs and values of environmentally conscious consumers. More importantly, policies should be directed toward minimizing the negative implications of digitally enabled sharing while encouraging the implementation of sharing models that can promote environmentally responsible production and consumption patterns together with life-enhancing social and economic opportunities in sharing communities.

References

1. Sarlay, S., Neuhofer, B.: Sharing economy disrupting aviation: travelers' willingness to pay. Tour. Rev. (2020)
2. Davidson, A., Habibi, M.R., Laroche, M.: Materialism and the sharing economy: a cross-cultural study of American and Indian consumers. J. Bus. Res. **82**, 364–372 (2018)
3. Pouri, M.J., Hilty, L.M.: The digital sharing economy: A confluence of technical and social sharing. Environ. Innov. Soc. Trans. **38**, 127–139 (2021)
4. Wirtz, J., So, K.K.F., Mody, M.A., Liu, S.Q., Chun, H.H.: Platforms in the peer-to-peer sharing economy. J. Serv. Manag. (2019)
5. Sutherland, W., Jarrahi, M.H.: The sharing economy and digital platforms: a review and research agenda. Int. J. Inf. Manage. **43**, 328–341 (2018)
6. Puschmann, T., Alt, R.: Sharing economy. Bus. Inf. Syst. Eng. **58**(1), 93–99 (2016)
7. Felländer, A., Ingram, C., Teigland, R.: Sharing economy. In: Embracing Change with Caution. Näringspolitiskt Forum Rapport, No. 11 (2015)
8. Eckhardt, G.M., Houston, M.B., Jiang, B., Lamberton, C., Rindfleisch, A., Zervas, G.: Marketing in the sharing economy. J. Mark. **83**(5), 5–27 (2019)
9. Matzler, K., Veider, V., Kathan, W.: Adapting to the sharing economy. MIT Sloan Manag. Rev. **56**(2), 71 (2015)
10. Barbu, C.M., Florea, D.L, Ogarcă, R.F., Barbu, M.C.R.: From Ownership to Access: How the Sharing Economy is Changing the Consumer Behavior. Amfiteatru Economic **20**(48), 373–387 2018
11. Pouri, M.J., Hilty, L.M.: Conceptualizing the digital sharing economy in the context of sustainability. Sustainability **10**(12), 4453 (2018)
12. Liu, X., Chen, H.: Sharing economy: promote its potential to sustainability by regulation. Sustainability **12**(3), 919 (2020)
13. Niezgoda, A., Kowalska, K.: Sharing economy and lifestyle changes, as exemplified by the tourism market. Sustainability **12**(13), 5351 (2020)
14. Guttentag, D.: Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. Curr. Issue Tour. **18**(12), 1192–1217 (2015)
15. Willis, G., Tranos, E.: Using 'Big Data' to understand the impacts of Uber on taxis in New York City. Travel Behaviour and Society **22**, 94–107 (2021)
16. Brynjolfsson, E., McAffe, A.: The Second Machine Age. Work, Progress and Prosperity in Time of Brilliant Technologies. W.W. Norton and Co, New York (2016)
17. Wirtz, J., Ehret, M.: Asset-based strategies for capturing value in the service economy. In Handbook of Service Science, vol. II, pp. 785–810. Springer, Cham (2019)
18. Stepnov, I.: Value creation by the sharing economy in the post-industrial society. In: Post-Industrial Society, pp.123–136. Palgrave Macmillan, Cham (2020)
19. Mi, Z., Coffman, D.M.: The sharing economy promotes sustainable societies. Nat. Commun. **10**(1), 1–3 (2019)
20. Raworth, K.: Doughnut economics: seven ways to think like a 21st-century economist. Chelsea Green Publishing (2017)
21. Cohen, M.J., 2016. *The future of consumer society: Prospects for sustainability in the new economy*. Oxford University Press.
22. Heinrichs, H.: Sharing economy: a potential new pathway to sustainability. Gaia **22**, 228–231 (2013)
23. Frenken, K.: Sustainability perspectives on the sharing economy. Environ. Innov. Soc. Trans. **23**, 1–2 (2017)
24. Martin, C.J.: The sharing economy: a pathway to sustainability or a nightmarish form of neoliberal capitalism? Ecol. Econ. **121**, 149–159 (2016)
25. Pouri, M.: Eight impacts of the digital sharing economy on resource consumption. Resour. Conserv. Recycl. **168**, 105434–105434 (2021)
26. Hilty, L.M., Aebischer, B.: ICT for sustainability: An emerging research field. In: ICT Innovations for Sustainability, pp. 3–36. Springer, Cham (2015)

27. Görög, G.: The definitions of sharing economy: a systematic literature review. Management (18544223), 13(2) (2018)
28. Ranjbari, M., Morales-Alonso, G., Carrasco-Gallego, R.: Conceptualizing the sharing economy through presenting a comprehensive framework. Sustainability **10**(7), 2336 (2018)
29. Hossain, M., 2020. Sharing economy: A comprehensive literature review. *International Journal of Hospitality Management*, *87*, p.102470.
30. Benkler, Y.: Sharing nicely: on shareable goods and the emergence of sharing as a modality of economic production. Yale LJ **114**, 273 (2004)
31. Diefenbach, T.: Intangible resources: a categorial system of knowledge and other intangible assets. J. Intellect. Cap. (2006)
32. Huang, C.Y.: File sharing as a form of music consumption. Int. J. Electron. Commer. **9**(4), 37–55 (2005)
33. Park, H., Armstrong, C.M.J.: Collaborative apparel consumption in the digital sharing economy: an agenda for academic inquiry. Int. J. Consum. Stud. **41**(5), 465–474 (2017)
34. UN.: Classification by Broad Economic Categories, Revision 5. United Nations Department of Economic and Social Affairs Statistics Division. Statistical Commission, 8–11 March 2016. https://unstats.un.org/unsd/statcom/47th-session/documents/BG-2016-11-Manual-of-the-Fifth-Revision-of-the-BEC-E.pdf [last accessed 2 June 2021].
35. Pouri, M.J., Hilty, L.M.: A Typology of Digital Sharing Business Models: A Design Science Research Approach. In International Conference on Design Science Research in Information Systems and Technology, pp. 297–308. Springer, Cham (2020)
36. Pouri, M.J., Hilty, L.M.: The relevance of digital sharing business models for sustainability. In: Proceedings of the 7th International Conference on ICT for Sustainability, pp. 77–87 (2020)
37. Bååth, J., Daoud, A.: Extending Social Resource Exchange to Events of Abundance and Sufficiency (2020). arXiv preprint arXiv:2010.02658
38. Daoud, A.: Scarcity, Abundance and Sufficiency: Contributions to Social and Economic Theory. Department of Sociology; Sociologiska Institutionen (2011)
39. Vathanophas, V.: Competency requirements for effective job performance in Thai public sector. Contemp. Manag. Res. **3**(1), 45–45 (2007)
40. Plepys, A., Singh, J.: Evaluating the sustainability impacts of the sharing economy using input-output analysis. Edward Elgar Publishing, A Research Agenda for Sustainable Consumption Governance (2019)
41. Cohen, B., Muñoz, P.: The emergence of the urban entrepreneur: how the growth of cities and the sharing economy are driving a new breed of innovators. ABC-CLIO (2016)
42. Ford, S., Despeisse, M.: Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. J. Clean. Prod. **137**, 1573–1587 (2016)
43. Botsman, R., Rogers, R: What's mine is yours. The rise of collaborative consumption (2010)
44. Gössling, S., Michael Hall, C.: Sharing versus collaborative economy: how to align ICT developments and the SDGs in tourism? J. Sustain. Tour. **27**(1), 74–96 (2019)
45. Erhardt, G.D., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J.: Do transportation network companies decrease or increase congestion? Sci. Adv. **5**, 5 (2019). https://doi.org/10.1126/sciadv.aau2670
46. Schaller, B.: Unsustainable? The growth of app-based ride services and traffic, travel and the future of New York City. Schaller Consulting (2017). http://schallerconsult.com/rideservices/unsustainable.htm [last accessed 2 May 2021]
47. Cheng, M., Chen, G., Wiedmann, T., Hadjikakou, M., Xu, L., Wang, Y.: The sharing economy and sustainability—assessing Airbnb's direct, indirect and induced carbon footprint in Sydney. J. Sustain. Tour. **28**(8), 1083–1099 (2020)
48. Palgan, Y.V., Zvolska, L., Mont, O.: Sustainability framings of accommodation sharing. Environ. Innov. Soc. Trans. **23**, 70–83 (2017)
49. Frenken, K., Schor, J.: Putting the sharing economy into perspective. Edward Elgar Publishing, A Research Agenda for Sustainable Consumption Governance (2019)
50. Dyllick, T., Rost, Z.: Towards true product sustainability. J. Clean. Prod. **162**, 346–360 (2017)

51. Tussyadiah, I.P., Sigala, M.: Shareable tourism: Tourism marketing in the sharing economy. J. Travel Tour. Mark. **35**(1), 1–4 (2018). https://doi.org/10.1080/10548408.2018.1410938
52. Verboven, H., Vanherck, L.: The sustainability paradox of the sharing economy. uwf UmweltWirtschaftsForum, 24(4), 303–314 (2016)
53. Binswanger, M.: Technological progress and sustainable development: what about the rebound effect? Ecol. Econ. **36**(1), 119–132 (2001)
54. Coroamă, V.C., Mattern, F.: Digital Rebound–Why Digitalization Will Not Redeem Us Our Environmental Sins (2019). http://www.vs.inf.ethz.ch/publ/papers/CoroamaMattern2019-DigitalRebound.pdf [last accessed 2 July 2021]
55. Santarius, T., Soland, M.: How technological efficiency improvements change consumer preferences: towards a psychological theory of rebound effects. Ecol. Econ. **146**, 414–424 (2018)
56. Salomon, I., Mokhtarian, P.L.: Can telecommunications help solve transportation problems? A decade later: are the prospects any better? Emerald Group Publishing Limited, Handbook of Transport Modelling (2008)
57. Moreau, H., de Jamblinne de Meux, L., Zeller, V., D'Ans, P., Ruwet, C., Achten, W.M.: Dockless e-scooter: A green solution for mobility? A comparative case study between dockless e-scooters, displaced transport, and personal e-scooters. Sustainability **12**(5), 1803 (2020)
58. Martin, M., Lazarevic, D., Gullström, C.: Assessing the environmental potential of collaborative consumption: peer-to-peer product sharing in Hammarby Sjöstad Sweden. Sustainability **11**(1), 190 (2019)
59. Weber, T.A.: The dynamics of asset sharing and private use. In: Proceedings of the 51st Annual Hawaii International Conference on System Sciences (HICSS), Waikoloa Village, HI, USA, 3–6 January 2018, pp. 5205–5211 (2018)
60. Griswold, A.: Shared scooters don't last long (2019). https://qz.com/1561654/how-long-does-a-scooter-last-less-than-a-month-louisville-data-suggests/ [last accessed 2 July 2021]
61. Warland, L., Hilty, L., Küng, J., Reinhard, J.: Fact Sheet: Business Travel. Sustainability Team, University of Zurich, Zurich, Switzerland (2016)
62. Louis-Philippe, P.V.C., Jacquemotte, Q.E., Hilty, L.M.: Sources of variation in life cycle assessments of smartphones and tablet computers. Environ. Impact Assess. Rev. **84**, p.106416 (2020)
63. Osibanjo, O., Nnorom, I.C.: Material flows of mobile phones and accessories in Nigeria: Environmental implications and sound end-of-life management options. Environ. Impact Assess. Rev. **28**, 198–213 (2008)
64. Böni, H., Schluep, M., Widmer, R.: Recycling of ICT equipment in industrialized and developing countries. In: ICT Innovations for Sustainability, Advances in Intelligent Systems and Computing, pp. 223–241. Springer International Publishing, Cham (2015)
65. Hollingsworth, J., Copeland, B., Johnson, J.X.: Are e-scooters polluters? The environmental impacts of shared dockless electric scooters. Environ. Res. Lett. **14**(8), 084031 (2019)
66. Tukker, A.: Product services for a resource-efficient and circular economy—a review. J. Clean. Prod. **97**, 76–91 (2015)
67. Curtis, S.K., Mont, O.: Sharing economy business models for sustainability. J. Clean. Prod. **266**, 121519 (2020)
68. Garetti, M., Taisch, M.: Sustainable manufacturing: trends and research challenges. Prod. Plan. & Control. **23**(2–3), 83–104 (2012)

A Review on Key Performance Indicators for Climate Change

Jelise Schokker, Andreas Kamilaris, and Savvas Karatsiolis

Abstract Climate change is one of the biggest threats to humanity in the near future. Almost all different scenarios of climate change involve large-scale disasters and hazards. In order to define goals to cities, regions and countries in regards to mitigating climatic change, we first need to understand which the important key performance indicators (KPIs) are, how they can be measured and which values they take. Then, each country can calculate its performance based on these KPIs, setting realistic goals for better performance in the near future. This paper performs a large survey to identify and list 63 relevant KPIs, together with suggested units and metrics associated with them, divided in eight different thematic areas. It can be considered as an important contribution in the global efforts to understand climatic change, shaping policies and setting goals associated with it.

Keywords Climate change · Key performance indicators · Metrics · Units · Countries · Performance.

1 Introduction

In the past decades, climate change has been recognized as an undeniable high-risk problem for humanity and its future on this planet. The year 2020 was tied with 2016 as the warmest years on record [1]. The average temperature of the earth has risen over 1.2 degrees Celsius since the late nineteenth century. Climate change and the environmental issues it causes affect the life of humans around the world significantly. As temperatures keep rising, an increasing amount of climate-related

J. Schokker · A. Kamilaris (✉)
Department of Computer Science, University of Twente, Enschede, The Netherlands
e-mail: a.kamilaris@utwente.nl; a.kamilaris@cyens.org.cy

J. Schokker
e-mail: n.j.schokker@student.utwente.nl

A. Kamilaris · S. Karatsiolis
CYENS Center of Excellence, Nicosia, Cyprus
e-mail: s.karatsiolis@cyens.org.cy

V. Wohlgemuth et al. (eds.), *Advances and New Trends in Environmental Informatics*, Progress in IS, https://doi.org/10.1007/978-3-030-88063-7_17

risks arise. Up to now, natural ecosystems have proven to be extremely resilient to change [2]. However, we do not know where the tipping point is and we definitely do not want to find out.

To define realistic and measurable goals for different cities, regions and countries around the world for avoiding or at least mitigating climatic change, we first need to understand which the important *key performance indicators* (KPIs) are, how they can be measured and which values they take. KPIs are defined as measures that are used to assess essential factors related to a given objective, such as reducing the effects of climate change [3]. The effectiveness of a country or city in achieving these objectives can be determined by these factors. Each country can thus calculate its performance based on these KPIs, setting realistic goals for better performance in the near future, empowering its citizens to work collaboratively in this direction, shaping environmental policies accordingly.

The contribution of this paper is to perform a survey to identify and list relevant, important and measurable KPIs, together with units and metrics associated with them. These KPIs and measurement units are based on how different countries around the world have been measuring those KPIs to date. Although some efforts have already been made in this direction (e.g. [4–6]—see Sect. 2), we argue that they do not capture the complete spectrum of important and relevant KPIs, which should be included in such estimations and calculations (see Sects. 3 and 4), nor they link those KPIs with their actual use by countries around the world (see Sect. 5). Our goal is to stimulate the discussion and motivate more effort in quantifying climatic change, environmental performance and sustainability at local and global level in a more *complete, inclusive, transparent and fair manner*. The rest of the paper is organized as follows: Sect. 2 lists related work, Sect. 3 describes the methodology used, Sect. 4 presents our findings, Sect. 5 discusses the overall results and finally, Sect. 6 concludes this paper.

2 Related Work

Only few research works tried to compile lists of KPIs related to the protection of the environment, environmental sustainability and performance at local level (i.e. city, country) in relation to climatic change. We describe these few efforts here. An effort to link sustainability aspects and company strategies via an appropriate set of KPIs is studied in [3]. The study identifies suitable KPIs that affect company performance, based on 82 papers identified, and proposes a new perspective on how to integrate sustainability issues in company strategies. KPIs are defined as environmental performance indicators (EPIs) in [7], used to measure an organization's impact on the environment, including ecosystems, land, air and water.

Moreover, a methodological framework for the selection of KPIs to assess smart city solutions is proposed in [8]. The framework proposed led to the development of a repository of 75 KPIs categorized in six dimensions: technical, environmental, economic, social, ICT and legal. To help cities in their choice, the work in [9] compared seven recently published indicator standards for smart sustainable cities, identifying 413 indicators which were classified based on the following taxonomy: five conceptual urban focuses (types of urban sustainability and smartness), ten sectoral application domains (energy, transport, ICT, economy, etc.) and five indicator types (input, process, output, outcome, impact). From the 413 indicators, only 45% were typical for sustainability assessment.

It seems that numerous KPIs have been proposed to deal with aspects of sustainability and sustainable development in relation to companies and organizations [3, 7] or smart cities [8, 9]. However, these proposals do not directly relate to the performance of cities and countries in relation to climate change.

Worth mentioning is the effort of the Intergovernmental Panel on Climate Change (IPCC), the United Nations group tasked with assessing climate science for policy makers, which proposed ways to measure and monitor emissions [5], as well as land use [10]. Further, the European Environment Agency (EEA) analyzed trends and projections of around 40 indicators, with a focus on key climate variables and on impacts of climate change on health, environment and economy [4]. Moreover, the performance of EU Member States on environmental innovations is measured by the *Eco-Innovation index*,[1] which is a composite indicator obtained by taking an unweighted average of 16 indicators included in the measurement framework adopted by the European Commission.

This absence of relevant metrics and criteria motivated organizations such as Germanwatch, to propose relevant indexes, like the *Global Climate Risk Index* (GCRI) [11] and the *Climate Change Performance Index* (CCPI) [6]. GCRI analyses to what extent countries and regions have been affected by impacts of weather-related loss events (storms, floods, heat waves etc.). On the other hand, CCPI acts as an independent monitoring tool for tracking the climate protection performance of 57 countries and the EU. The CCPI aims to enhance transparency in international climate politics and it enables comparison of individual countries' climate protection efforts and progress. CCPI assesses countries based on the following: (a) GHG Emissions (40% of overall score), (b) Renewable Energy (20% of overall score), (c) Energy Use (20% of overall score) and d) Climate Policy (20% of overall score). It is remarkable that still no country has made it to the top three ranks, based on CCPI.

Finally, the United Nations' 2030 Agenda for Sustainable Development [12] is an effort worth mentioning. The agenda is a roadmap for prosperity, peace and stability for humanity and for the planet. It consists of 17 *Sustainable Development Goals* (SDGs), 169 targets and 231 unique indicators that can be used to achieve sustainable development. Economic, social and environmental sustainable development are the three main groups of indicators. The latter group, in particular, is the one more relevant to our study.

[1] European Commission, Eco-innovation index. https://ec.europa.eu/environment/ecoap/indicators/index_en.

3 Methodology

This paper aims to fill the aforementioned gaps in the literature, by addressing the following research questions:

1. Which are the KPIs used by different countries and cities around the world to measure climate change?
2. Which units, metrics and ranges of values are associated with these KPIs?
3. Which countries actually use these KPIs, assigning values and setting goals for mitigating climate change based on these KPIs?

It is important first to properly define *climate change*, in the context of this survey. Many people relate climate change to changes in climate that result from greenhouse gas forcing of the climate system [13]. Under the FCCC, the term is defined as a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability over comparable time periods. On the other hand, the IPCC states that climate change is any change in climate over time whether due to natural variability or as a result of human activity [14]. We adopt the IPCC definition of climate change here.

Search for related work was performed through the Web scientific indexing services *Web of Science* and *Google Scholar*. The following query was used:

["Key Performance Indicators" OR "Metrics" OR "Units"] AND ["Climate" OR "Climate change" OR "Environment]

Twenty three (23) scientific papers and reports were found via this approach. To increase the range of our bibliography, a search of the related work as appeared in these 23 papers was also performed. This effort allowed to increase the number of papers discovered to 52. Each of these papers was studied in detail, aiming to address the research questions as listed above. We tried to record KPIs, understand their semantics and how they have been used in different regions of the world. From the 83 KPIs identified, 63 KPIs were included to our survey, based on the following criteria:

- This KPI has been adopted by some city or country.
- This KPI is measurable based on some quantitative metric.
- This KPI has some unique value based on the quantitative metric associated with it.

4 Results

Based on the 63 KPIs selected, eight general clusters were created, placing the KPIs inside those clusters. These clusters are the following: (a) pollution; (b) resource use; (c) climate hazards; (d) biodiversity; (e) transport; (f) land use; (g) health and

(h) other. These clusters are described in Sect. 4.1 below, while the KPIs discovered, associated with these clusters, are listed in Sect. 4.2.

4.1 Clusters of KPIs

The 8 general clusters are described as follows:

4.1.1 Pollution

Pollution includes air pollution and emissions, plastic pollution, waste, especially solid waste, water and soil contamination, as well as aspects of recycling. One of the most popular types of emissions recorded is greenhouse gases (GHG), which are gases that absorb and emit infrared radiation. Water vapor, carbon dioxide, methane, nitrous oxide and ozone are the GHG present in the Earth's atmosphere [15]. The earths greenhouse effect has increased significantly since the industrial revolution [16]. Therefore, human activity is a key element in reducing major greenhouse gas emissions [17].

Plastics and micro-plastics that harm marine ecosystems are another form of pollution. Plastic particles have been discovered in the intestines of dead aquatic animals, demonstrating that plastic has caused catastrophic damage to living organisms [18]. The production of plastic has been growing rapidly in the past and is expected to increase further in upcoming years [19]. It is important to properly manage this plastic to prevent it from ending up in the oceans.

Regarding waste, because of inappropriate treatment and transportation, it can cause pollution of the air, water, and soil, causing numerous environmental and health risks. Thus, solid waste management is critical [20]. Solid waste disposal into landfills is still a common method of disposal, despite the health and pollution hazards linked to this disposal method [21]. Other ways of waste disposal, such as recycling, incineration, or pyrolysis, are used to manage solid waste pollution and its harmful consequences [20]. Recycling is a popular method of waste disposal. It entails reusing certain waste components and it therefore saves resources, reduces the manufacturing of new resources, and minimizes pollution [20]. However, recycling is not yet a sufficiently widespread practice, as a considerable amount of waste still ends up in landfills or in nature. Since 1950, only nine percent of plastic waste has been properly recycled [22].

Finally, in respect to water and soil contamination, this occurs as a consequence of increased industrialisation and urbanisation. Polluted soil is harmful to humans and animals, since contaminating materials enter their body through the food chain [23]. Polluted water is harmful to marine life, plants, humans and the environment [24]. While companies and treatment plants make use of different wastewater treatment procedures, some industries continue to dispose untreated wastewater into water bodies and nearby soils. This causes pollutants such as heavy metals and organic pollutants to penetrate the soil and the water. Dyes, pharmaceuticals, product waste, and petroleum organic pollutants are several of the organic pollutants that have posed a significant threat to aquatic species, plants, and people.

4.1.2 Resource Use

Human demand as well as the availability of natural resources are dispersed unevenly across the world [25]. Water is a critical resource for every person on earth. However, over 700 million people do not have access to basic drinking-water services. On top of that, half of the world's population is expected to be living in water-stressed areas by 2025 [26]. A crucial aspect of climate change mitigation is to ensure that everyone will have access to reliable water services. It is important that countries and cities are aware of current and future risks to their water security. Threats can include increased water stress or scarcity, droughts, water-borne diseases and other events that affect the available water supply [27]. The implementation of smart water resource management strategies can lead to an increased resilience against the impacts of climate change, especially in developing countries [28].

Regarding energy as a resource, the consumption of fossil fuels has grown substantially during the last 50 years [29]. Between 1980 and 2019, global fossil fuel consumption almost doubled [30], despite the fact that 16% percent of primary energy came from low-carbon sources in 2019.

Moreover, reducing the amount of food wasted is of great importance. In many regions around the world, it is expected that these reductions can result in a considerable gain in water and food security [31]. According to [32], the world is currently not on course to achieve the hunger and malnutrition targets that had been set for 2030. Although the number of people suffering from hunger had been declining for decades, this is not the case anymore.

4.1.3 Climate Hazards

Climate hazards are an important aspect of climate change. Extreme events will most likely change in frequency and severity as a result of climate change [33]. One example is North America, where there are increasingly more and worse wildfires [34]. Alternatively, climate change can also cause changes in precipitation. This in turn can result in more precipitation-induced flooding [35]. It is important for cities and countries to know what climate hazards may pose a threat, either today or in the future.

Moreover, public health systems are crucial for the quality of living of citizens, but they can also face several risks related to climate change. These risks include diseases and pandemics as a result of climate change, threats to food security and safety, and many more [27].

4.1.4 Biodiversity

Climate change has an impact not only on people but also on physical ecosystems [36]. Even a minor shift in climate can result in the extinction of some vulnerable and endangered animal species, as well as birds and fish. The Biodiversity Intactness Index (BII) shows how native terrestrial species' average abundance compares to their abundance before human intervention [37]. It is necessary to use such indexes to comprehend the interactions between plants, animals and biodiversity, implementing strategies to improve biodiversity based on that understanding.

4.1.5 Transport

Transport is an important part of everyday life of billions of people. Over 16% of all GHG emissions in 2016 can be attributed to transportation [38]. Almost 12% of emissions were caused by road transport [38]. This means that the electrification of road transport could make a significant impact on GHG emissions, combined with alternative ways to produce electricity [39]. The implementation of good public transit networks is also important [40, 41]. If these networks are of good quality, being accessible to as many people as possible, they can reduce the use of private transport [41].

4.1.6 Land Use

An important metric related to land use is the deforestation rate. This is critical in the earth's largest natural reserves, such as the Amazon rainforest. Brazil has large deforestation rates, which are rising [42]. Related to biodiversity as well, a relevant metric here is the number of living organisms lost per square kilometer of cut-down forest [43].

Further, agriculture emits an especially high amount of GHG. It contributes 17% to global emissions directly through agricultural operations and 7–14% indirectly via land use changes [44]. Agriculture not only contributes to climate change, but it is also particularly vulnerable to it. Most of the risks associated with agriculture are caused by adverse climatic conditions and climate variability, with climate change posing an additional concern [45], The food chain contributes largely to emissions too. Food often goes a long way before it is consumed: it is produced, processed, packed, shipped, and prepared. Every step causes GHG emissions to be released into the earth's atmosphere [46].

4.1.7 Health

Health relates mostly to humans, especially in terms of mortality rates and illnesses caused by or being correlated to climatic change. It seems that a large percentage of such illnesses relates to respiratory diseases caused by air pollution. Finally, public health seems to constitute a critical indicator, since climate change might stress the public health systems in case of climatic disasters occurring.

4.1.8 Others

Finally, a more general cluster involves metrics relating to global temperature and sea level, which are hard to be attributed to different countries and regions, but relate to the whole planet. The mean global temperature increase may have a variety of negative consequences, including an increase in the likelihood of flooding, droughts, and heat waves [47]. Sea-level rise is associated with global warming, and constitutes a global phenomenon with potentially massive consequences, including coastal erosion and marine habitat destruction [48]. Countries in low-elevation areas, and small islands in particular, are expected to suffer from significant losses of land or even disappear, e.g. the case of Maldives.

4.2 *List of KPIs*

Table 1 lists the KPIs identified and recorded through this study, organized into the 8 clusters introduce in Sect. 4.1. Each KPI is associated with a topic, a unique name, a short description, the unit mostly used to measure it, as well as the range of values it takes, from a negative/bad/low value (−) to a positive/good/high one (+). The column *SDG target* shows the corresponding SDG and SDG target for each KPI, wherever available. The first number refers to the SDG, and the second shows the specific target. For example, the KPI #1 falls under the general SDG goal 11 (i.e. *"Make cities and human settlements inclusive, safe, resilient and sustainable"*) and target 11.6 (i.e. *"By 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management"*). All SDGs and SDG targets are described in the United Nations' 2030 Agenda for Sustainable Development [12], for the reader's reference.

5 Discussion

As illustrated in Sect. 4, a large number of KPIs is used to measure a country's or city's performance in relation to climate change. We have identified 63 such KPIs, dividing them into eight categories, based on the criteria set in Sect. 3.

Some of the KPIs mentioned above (e.g. #1, 2, 5, 6, 12, 17, 22–26, 27, 28, 32, 44, 45, 47, 48—Table 1) have already been used to measure and quantify sustainability and environmental performance, as well as to set targets in organizational settings [3, 7] or in smart city solutions [8, 9]. However, many other KPIs (e.g. those relating to

Table 1 List of key performance indicators related to climate change, as recorded through this study

Category	Topic	#	KPI	Unit	Description	SDG target	+	–
1 Pollution	1.1 Air Pollution	1	Air Quality Index	#	Ranking of cities/countries based on annual average PM2.5 concentration ($\mu g/m^3$)	11.6	0–50	>100
	1.2 Greenhouse Gases	2	Existence of emissions inventory system	Yes/No	Inventory which includes emissions that are within the city boundary	13.2	Yes	No
		3	Existence of GHG emissions reduction target	Yes/No	Existence of a target to reduce greenhouse gas emissions	13.2	Yes	No
		4	Emissions generated by government operations	tCO_2e	Scope 1, 2 and 3 emissions as a result of government operations	13.2	0	–
		5	Emissions generated by community activities	tCO_2e	Scope 1, 2 and 3 emissions as a result of community activities	13.2	0	–
	1.3 Plastic Pollution	6	Existence of plastic policies	Yes/No	Existence of regulations on the use and disposal of (single-use) plastics	N/A	Yes	No
		7	Percentage of mismanaged plastic waste	%	The percentage of total waste that is not properly disposed of	14.1	0%	100%
		8	Amount of plastics currently in the oceans	Metric tonnes	Amount of macro- and micro-plastics currently in the oceans	14.1	0	–
	1.4 Solid Waste	9	Annual solid waste generation	tonnes/year	Total solid waste generation per year	11.6	0	–
		10	Solid waste disposed to landfill or incineration	tonnes/year	Solid waste that is disposed of in landfills or that is incinerated	11.6	–	–
	1.5 Recycling	11	Solid waste diverted away from landfill or incineration	%	Solid waste that does not end up in landfills or incineration because of recycling	12.5	100%	0%
		12	Percentage of population with access to recycling	%	Percentage of the population that have access to a recycling point	12.5	100%	0%
	1.6 Soil Pollution	13	Percentage of land that is polluted	%	Percentage of surface area that is affected by soil pollution	15.3	0%	–
	1.7 Water Pollution	14	Percentage of heavy metal concentration in river and lake water bodies	%	Heavy metal concentration in global river and lake water bodies as a cause of water pollution	6.3	0%	100%
		15	Biochemical Oxygen Demand	mg/L	Measurement of non toxic organics in water	6.3	<1 mg/L	>8 mg/L
		16	Chemical Oxygen Demand	mg/L	Measurement of total toxic and non toxic organics in water	6.3	–	–

(continued)

Table 1 (continued)

Category	Topic	#	KPI	Unit	Description	SDG target	+	–
2 Resource Use	2.1 Water	17	Water consumption	L/person/day	The average liters of water used by one person in one day	6.4	–	–
		18	Water stress level	Low-high	The ability to meet a region's demand for water	6.4	Low	High
		19	Existence of a public Water Resource Management strategy	Yes/No	Existence of a plan for dealing with water use and resources	6.5	Yes	No
		20	Existence of any current or future risks to the city's water security	Yes/No	Existence of (climate change related) risks that will decrease the city's water security	N/A	No	Yes
		21	Percentage of population with access to potable water supply	%	Percentage of people that have access to clean and safe drinkwater	1.4	100%	0%
	2.2 Energy	22	Energy consumption	kWh/household/year	The average amount of energy consumed by one household per year	7.3	0 kWh	–
		23	Share of renewable energy sources	%	The share of a city's energy mix that consists of renewable sources	7.2	100%	0%
		24	Existence of renewable energy or electricity target	Yes/No	Existence of a target to increase the use of renewable energy	N/A	Yes	No
		25	Existence of target to increase energy efficiency	Yes/No	Existence of a target to use energy more efficiently and eliminating energy waste	7.3	Yes	No
		26	Percentage of energy grid that is zero carbon	%	Zero carbon includes solar, wind, hydro, biomass and geothermal as the source to produce electricity	N/A	100%	0%
	2.3 Food	27	Annual food waste	tonnes/year	Amount of food that is wasted each year	12.3	0 tonnes/year	–
		28	Ecological footprint of consumption per person	gha/person	The Ecological Footprint per person is a measure of the rates of consumption and the total population of a country	12.2	<1.6	>5
3 Climate Hazards	3.1 Climate Hazards	29	Global Climate Risk Index	#	The Global Climate Risk Index shows the level of exposure and vulnerability to extreme weather events	13.1	>100	0–50
		30	Existence of inventory of relevant climate hazards	Yes/No	Existence of an inventory to keep track of relevant current or future climate hazards in a city	13.1	Yes	No
		31	Most significant climate hazards faced by the city	N/A	Identification of the most important climate hazard a city faces or will face	13.1	No risks	many risks or no identification
		32	Existence of a climate change risk and vulnerability assessment	Yes/No	Existence of an assessment of current or future risks and the city's vulnerability	13.1	Yes	No

(continued)

Table 1 (continued)

Category	Topic	#	KPI	Unit	Description	SDG target	+	–
4 Biodiversity	4.1 Biodiversity Intactness	33	Biodiversity intactness index	%	The Biodiversity Intactness Index (BII) shows how native terrestrial species' average abundance compares to their abundance before human intervention	15.5	100%	>60%
	4.2 Terrestrial Animal Diversity	34	Percentage of known terrestrial species that are threatened	%	Percentage of terrestrial species that are threatened according to the IUCN Red List	15.5	0%	100%
		35	Terrestrial protected land area as percentage of total land area	%	Percentage of land surface area that is protected	15.1	–	0%
	4.3 Marine Animal Diversity	36	Percentage of known marine species that are threatened	%	Percentage of marine species that are threatened according to the IUCN Red List	15.5	0%	100%
		37	Marine protected land area as percentage of total land area	%	Percentage of marine surface area that is protected	14.5	–	0%
		38	Existence of policies for commercial fishing	Yes/No	Existence of policies to curb commercial fishing rates	14.4	Yes	No
		39	Commercial fishing rates	nr	The amount of fish caught for commercial purposes	14.4	Low	High
		40	Bycatch rates	nr	The amount of marine animals that are caught unintentionally while fishing for other animals	14.4	High	Low
5 Transport	5.1 Public Transport	41	Percentage of population living within 500m of a mass transit station	%	The amount of people that live within 500 m of a mass transit station and have access to public transportation	9.1	100%	0%
		42	Quality of public transport	%	The quality rating given by inhabitants to a city's public transport system	9.1	100%	0%
		43	GHG emissions caused by public transport	tCO2e	The total amount of GHG emissions caused by public transport	9.4	0	–
	5.2 Private Transport	44	Percentage of population that owns a private car	%	Describes how many people own a private car	N/A	0%	100%
		45	GHG emissions caused by private transport	tCO2e	The amount of GHG emissions caused by private transport	N/A	0	–
		46	Existence of a zero- or low-emission zone in the city	Yes/No	The existence of an area in the city where only zero- or low-emission vehicles are allowed	9.4	Yes	no

(continued)

Table 1 (continued)

Category	Topic	#	KPI	Unit	Description	SDG target	+	–
	5.3 Electric Vehicles	47	Percentage of private cars that are electric	%	The percentage of total private cars that are electric	N/A	100%	0%
		48	Public access EV charging points per capita	nr	The number of charging points for electric vehicles per capita	N/A	>1	0
	5.4 Aviation	49	GHG emissions caused by aviation	tCO2e	The amount of GHG emissions caused by air travel	N/A	0	–
		50	Per capita emissions from domestic and international flights	kg	The total combined emissions caused by domestic and international flights per capita	N/A	0	>500
6 Land Use	6.1 Deforestation	51	Deforestation rate	Mha/year	The total forest surface area that is cut down each year	15.2	0	–
		52	Percentage of global land cover that is tree covered	%	The percentage of total land area that is covered by trees	15.1	>30	0
	6.2 Agriculture	53	GHG emissions caused by agriculture	tCO2e	The amount of GHG emissions caused by agriculture	N/A	0	–
		54	Percentage of land used for agriculture	%	The percentage of total land area used for agriculture	N/A	–	–
		55	Surface area of potential agricultural spaces in a city	km2	The total land area that has the potential to be turned into agricultural space	N/A	–	–
		56	Vulnerability to climate change related agricultural risks	N/A	The vulnerability to climate related agricultural risks such as droughts	1.5	Not vulnerable	Very vulnerable
7 Health	7.1 Public Health	57	Identification of risks to public health or health systems associated with climate change	Yes/No	Identification of risks to the public health or health systems of a city	11.5	No risks	Many risks or no identification
	7.2 Mortality	58	Excess mortality caused by extreme heat	%	Addresses the number of people that die from extreme heat	N/A	0%	100%
		59	Deaths caused by air pollution	deaths per 100.000 people	Mortality rate linked to household and ambient air pollution	3.9	0	>1
	7.3 Illnesses	60	Number of heat related illnesses	nr	Identifies the amount of illnesses that were caused by extreme heat	N/A	0	>1
		61	Number of respiratory diseases caused by increased air pollution	nr	Identifies the amount of respiratory diseases that were caused by an increase in air pollution	3.9	0	>1
8 Other	8.1 Global Temperature	62	Annual rise in global temperature	°C/year	The rise in global temperature per year	N/A	<1.5	>2
	8.2 Sea Level	63	Annual sea level rise	mm/year	The rise of the sea level per year	N/A	0	3

climate hazards, biodiversity, land use, other—Table 1) have not directly been linked to companies and smart cities yet.

Moreover, some of the most prominent global efforts to define and measure climate change through various KPIs, such as IPCC [5, 10] and EEA [4], focus only on a subset of the indicators/KPIs identified and proposed in this study. We have counted 35 KPIs discussed by IPCC and 40 indicators proposed by EEA. While the KPIs of IPCC relate mostly to pollution, resource use, biodiversity and land use, the indicators of EEA focus on impacts of climate change on health, environment and economy. Here, it is worth mentioning that through our study we have not identified any economy-related KPIs which have been adopted by any city or country around the world and have been somehow measured.

One of the most mature and complete efforts to date in the direction of measuring the climate change-related performance of countries, based on various KPIs proposed, is the work of Germanwatch and the proposed Climate Change Performance Index (CCPI) [6]. CCPI embraces a wide list of KPIs, but we argue that this list could have been extended with additional ones (e.g. ones related to biodiversity, transport, land use—Table 1).

A more thorough comparison is shown in Table 2. The table lists the CCPI indicators and the corresponding KPIs presented in this paper. One insight that can be concluded from this is that only six KPIs are also covered in the CCPI, namely numbers 3 (Existence of GHG emissions reduction target), 4 (Emissions generated by government operations), 5 (Emissions generated by community activities), 22 (Energy consumption), 23 (Share of renewable energy sources), and 24 (Existence of renewable energy or electricity target). This means that the large majority of the KPIs listed in this paper are not included in the CCPI.

As stated in Sect. 2, the United Nations' 2030 Agenda for Sustainable Development is also an important effort to consider. Many of the KPIs in Table 1 have a corresponding SDG target. However, when comparing the findings of this paper with the UN SDGs, two categories are notably not sufficiently covered by the targets: transport and agriculture.

We are very much in favour of a globally accepted index, which combines a wealthy and diverse list of KPIs, comparing the performance of cities and countries in respect to climate change. On one hand, this can trigger a healthy competition, creating public pressure to governments and local actors to act and, on the other hand, it will facilitate the work of policy-makers in terms of creating environmentally friendly laws, policies and directives. While individual policies focusing on specific KPIs can lead to better monitored and assessed results, at the same time, embracing a diverse list of KPIs in unison can also give a complete and comprehensive representation of countries' efforts to mitigate climatic change.

The European Union seems to work towards the right direction in some clusters, e.g. resource use and energy [50], land use and forestry [51], as well as biodiversity [52]. For example, in relation to the share of renewable energy sources, EU plans to have a renewable energy share of at least 32%, plus using energy more efficiently [53], with a target of 32.5% savings by 2030, in comparison to 2020. In relation to biodiversity and land use, EU builds strategies to protect nature and reverse the

Table 2 Comparison of the CCPI and the KPIs presented in this thesis

CCPI indicator [49]	Corresponding KPI(s)	Comparison with this paper
Current Level of GHG Emissions per Capita	Emissions generated by government operations (4) and emissions generated by community activities (5)	The CCPI focuses on GHG emissions in general, while this index separates it into government and community emissions.
Past Trend of GHG Emissions per Capita	Emissions generated by government operations (4) and emissions generated by community activities (5)	Unlike the CCPI, this index does not have separate indicators for current and past emissions.
Current Level of GHG Emissions per Capita compared to a well-below-2 °C compatible pathway	N/A	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons.
GHG Emissions Reduction 2030 Target compared to a well-below-2 °C compatible pathway	Existence of GHG emissions reduction target (3)	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons. However, targets for GHG emissions are covered in the index.
Current Share of Renewables per Total Primary Energy Supply (TPES)	Share of renewable energy sources (23)	Both indexes include a similar KPI on the share of renewables.
Development of Energy Supply from Renewable Energy Sources	N/A	No specific KPI on the development of renewables is included in the index.
Current Share of Renewables per TPES compared to a well-below-2 °C compatible pathway	N/A	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons.
Renewable Energy 2030 Target compared to a well-below-2 °C compatible pathway	Existence of renewable energy or electricity target (24)	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons. However, targets for renewable energy are covered in the index.
Current Level of Energy Use (TPES/Capita)	Energy consumption (22)	Both indexes include a similar KPI on energy use.
Past Trend of TPES/Capita	Energy consumption (22)	Unlike the CCPI, this index does not have separate indicators for current and past emissions.
Current Level of TPES/Capita compared to a well-below-2 °C compatible pathway	N/A	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons.
TPES/Capita 2039 Target compared to a well-below-2 °C compatible pathway	N/A	This index does not include comparisons in the KPIs, it solely states the KPIs that can be used for comparisons.
National Climate Policy	N/A	No KPI for general climate policy is included in this index.
International Climate Policy	N/A	No KPI for general climate policy is included in this index.

degradation of ecosystems. We expect and hope that these strategies and initiatives will be complete, realistic, fair for member countries, with proper incentives and aids in order to reach the goals proposed. The EU initiatives could become exemplar for the rest of the world.

5.1 Countries and Use of KPIs

In Fig. 1, the use of the 63 KPIs in 193 countries is displayed as a categorical heatmap. The KPIs that were used for the graph are those that are implemented at a country level and have recent data available (#1, 6, 7, 9, 10, 11, 17, 18, 21, 22, 23, 24, 25, 27, 28, 29, 35, 37, 38, 50, 51, 52, 53, 54, 59—Table 1). This means that a maximum number of 25 (KPIs) can be scored by each country. These 25 KPIs cover the whole spectrum of categories as defined in this paper, except from transport (i.e. only aviation is included). Furthermore, out of the 25 KPIs used in the heatmap, twelve are used by more than 150 out of 193 countries included in the heatmap. These indicators are: 9 (Annual solid waste generation), 17 (Water consumption), 18 (Water

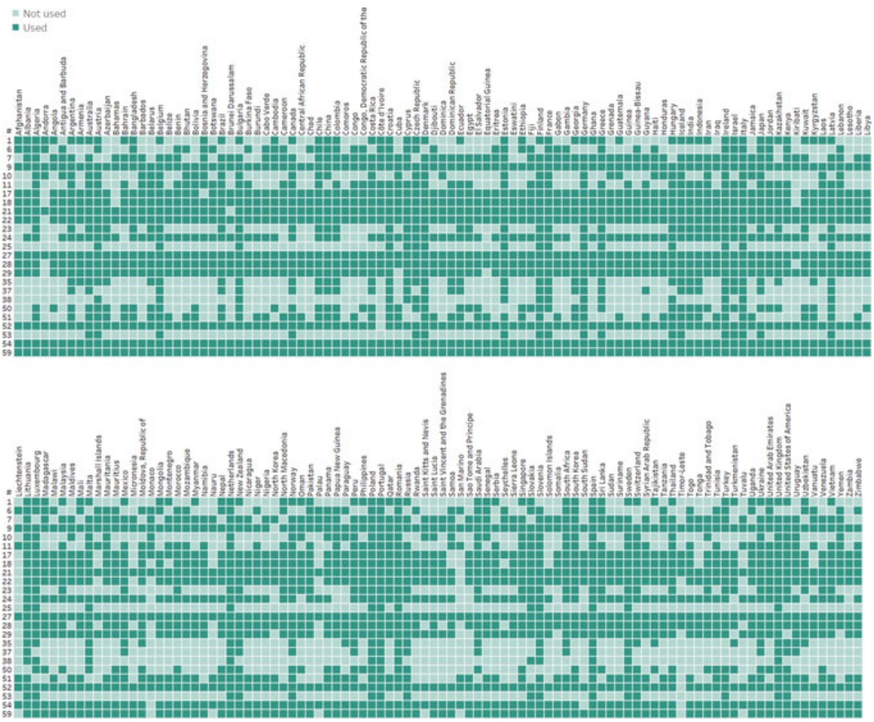


Fig. 1 Heatmap showing the use of KPIs per country [54–73]

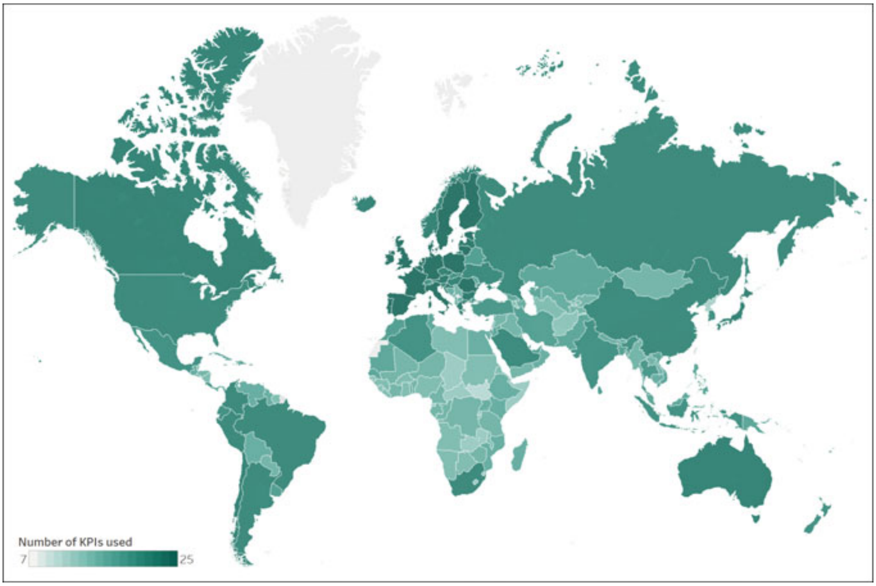


Fig. 2 Choropleth map showing the use of KPIs per country [54–73]

stress level), 21 (Percentage of population with access to potable water supply), 22 (Energy consumption), 24 (Existence of renewable energy or electricity target), 27 (Annual Food Waste), 28 (Ecological footprint of consumption per person), 29 (Global Climate Risk Index), 52 (Percentage of global land cover that is tree covered), 54 (Percentage of land used for agriculture) and 59 (Deaths caused by air pollution). The only KPI that is used by all countries and also has available data for each of them is 27 (Annual food waste). The countries that employ all 25 KPIs are all European ones. Six countries use all KPIs: Ireland, Italy, the Netherlands, Poland, Portugal and Spain. Additionally, there are seven countries that use all KPIs but one. These countries include: Denmark, Estonia, France, Lithuania, Romania, Sweden and the United Kingdom. Figure 2 depicts the aggregated number of KPIs adopted by countries around the world. As can be seen in the map, there are some clear regional differences in the use of the KPIs. Many countries in Africa and some countries in central Asia use significantly fewer KPIs than the European ones. There is a clear gap between developed and developing countries in respect to the use of KPIs to measure climate change. Both visualizations, presented in Figs. 1 and 2, can be better viewed via this link.[2]

[2] Visualizations of this project: https://superworld.cyens.org.cy/demo2.html.

5.2 Limitations and Future Work

As a limitation of this work, we note again that the KPIs recorded and listed in our survey are based on the criterion that they have been adopted by a country or city around the world. This might leave out some relevant KPIs which have not yet been in use (e.g. circular economy, eco-innovations, biodiversity of insects and plants, burned areas due to wildfires). Further, identifying and developing new KPIs that are presently not covered by the survey is outside the scope of this research. Another important limitation of this study is the data available. As stated before, all countries that adopt all KPIs are European ones. Data for other countries might still exist, but not be available in English (i.e. main language used during bibliography review). This can affect the accuracy of the information presented in Fig. 2.

As future work, we are preparing a complete, elaborate climate change index, similar to the CCPI [6], which includes all the KPIs listed in Table 1. We aim to examine how cities and countries perform, in relation to an index that combines all or at least a large percentage of these KPIs, giving a unique score to each country, facilitating and promoting comparisons and competitions. Further, we aim to improve our results considering the limitations mentioned in the previous paragraph. It is important to observe whether countries located in different parts of the world give priority or focus only on specific KPIs, which are linked to their most significant and/or urgent climate change risks.

6 Conclusion

This paper performed a large survey to identify and list key performance indicators, used by different cities and countries around the world to measure and quantify climate change. Sixty three different KPIs have been identified, listed together with suggested units and metrics associated with them, divided in eight different thematic areas: pollution, resource use, climate hazards, biodiversity, transport, land use, health and other. As climate change is one of the biggest threats to humanity, we need to understand, embrace and use such KPIs, in order to define tangible and fair goals for cities, regions and countries around the world, in regards to mitigating climatic change. Each country can calculate its performance based on these KPIs, setting realistic objectives for better performance in the near future, exchanging knowledge and best practices with other countries and regions. This paper can be considered a small but important contribution in the global efforts to understand and measure climatic change, to define properly and fairly policies and to set goals associated with it. It is of crucial importance that countries form appropriate policies to combat climate change and the problems that arise, as the clock is ticking against us as species.

References

1. NASA Global Climate Change.: Vital signs of the planet. https://climate.nasa.gov/vital-signs/global-temperature/. Accessed 25 Dec 2019 (2018)
2. Côté, I.M., Darling, E.S.: Rethinking ecosystem resilience in the face of climate change. PLoS Biol **8**(7), e1000438 (2010)
3. Hristov, Ivo, Chirico, Antonio: The role of sustainability key performance indicators (kpis) in implementing sustainable strategies. Sustainability **11**(20), 5742 (2019)
4. European Environment Agency. Climate change, impacts and vulnerability in Europe 2016—an indicator-based report. EEA Report No. 1/2017 (2016)
5. Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., et al.: Mitigation pathways compatible with 1.5 c in the context of sustainable development. In: Global warming of 1.5° C, pp. 93–174. Intergovernmental Panel on Climate Change (IPCC) (2018)
6. Burck, Jan, Hagen, Ursula, Höhne, Niklas, Nascimento, Leonardo, Bals, Christoph: Climate change performance index 2021. Germanwatch, Bonn (2020)
7. Jamous, N., Müller, K.: Environmental performance indicators. In: Organizations' Environmental Performance Indicators, pp. 3–18. Springer (2013)
8. Angelakoglou, Komninos, Nikolopoulos, Nikos, Giourka, Paraskevi, Svensson, Inger-Lise., Tsarchopoulos, Panagiotis, Tryferidis, Athanasios, Tzovaras, Dimitrios: A methodological framework for the selection of key performance indicators to assess smart city solutions. Smart Cities **2**(2), 269–306 (2019)
9. Huovila, Aapo, Bosch, Peter, Airaksinen, Miimu: Comparative analysis of standardized indicators for smart sustainable cities: what indicators and standards to use and when? Cities **89**, 141–153 (2019)
10. Shukla, P.R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.O., Roberts, D.C., Zhai, P., Slade, R., Connors, S., Van Diemen, R., et al.: Ipcc9: Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (2019)
11. Eckstein, David, Künzel, Vera, Schäfer, Laura, Winges, Maik: Global climate risk index 2020. Germanwatch, Bonn (2019)
12. United Nations. Transforming our world: the 2030 Agenda for Sustainable Development (2015)
13. Pielke, R.A., Misdefining, Jr.: "climate change": consequences for science and action. Environ. Sci. Policy **8**(6), 548–561 (2005)
14. Houghton, Ed.: Climate change 1995: The science of climate change: contribution of working group I to the second assessment report of the Intergovernmental Panel on Climate Change, vol. 2. Cambridge University Press (1996)
15. Easterbrook, D.J.: Greenhouse gases. In: Evidence-Based Climate Science: Data Opposing CO2 Emissions as the Primary Source of Global Warming: Second Edition, pp. 163–173. Elsevier (2016)
16. Houghton, J.: Global Warming The Complete Briefing, 3rd edn. Cambridge University Press (2004)
17. El Zein, Ahmad L.: The effect of greenhouse gases on earth's temperature. Int J Environ Monit Anal **3**(2), 74 (2015)
18. Budi Kurniawan, S., Sheikh Abdullah, S.R., Imron, M.F., Ismail N.: Current state of marine plastic pollution and its technology for more eminent evidence: a review. J Clean Prod **278**, 123537 (2021)
19. Billard, G., Boucher, J.: The challenges of measuring plastic pollution. Field Actions Sci Rep 2019 (Special Issue 19), 68–75 (2019)
20. Chadar, S., Keerti, C.: Solid waste pollution: a hazard to environment. Recent Adv Petrochem Sci**2**(3) (2017)
21. Hamer, G.: Solid waste treatment and disposal: effects on public health and environmental safety. Biotechnol Adv **22**(1–2), 71–79 (12 2003)

22. Woldemar d'Ambrières. Plastics recycling worldwide: current overview and desirable changes. Technical Report Special Issue 19 (2019)
23. Popa, Cristina, Petrus, Mioara: Heavy metals impact at plants using photoacoustic spectroscopy technology with tunable CO2 laser in the quantification of gaseous molecules. Microchem J **134**, 390–399 (2017)
24. Inyinbor, A.A., Adebesin, B.O., Abimbola, O.P., Adelani-Akande, T.A., Adewumi, D.O., Oreofe, T.A.: Water pollution: effects, prevention, and climatic impact. In: Water Challenges of an Urbanizing World. InTech (3 2018)
25. Almond, R.E.A., Grooten, M., Petersen, T.: WWF Living Planet Report 2020. Technical Report 2 (2020)
26. World Health Organization. Drinking Water Fact Sheet (2019)
27. CDP. Cities 2020 Questionnaire - CDP Cities disclosure cycle 2020, 2020
28. Soto-Montes, Gloria, Herrera-Pantoja, Marina: Implications of climate change on water resource management in megacities in developing countries: Mexico city case study. Environ Manag Sustain Dev **5**(1), 47 (2015)
29. Stephenson, M.: Energy and climate change: an introduction to geological controls, interventions and mitigations. In: Energy and Climate Change: An Introduction to Geological Controls, Interventions and Mitigations, pp. 175–178. Elsevier (2018)
30. Ritchie, H., Roser, M.: Energy. Our World in Data (2020)
31. Kummu, M., de Moel, H., Porkka, M., Siebert, S., Varis, O., Ward, P.J.: Lost food, wasted resources: global food supply chain losses and their impacts on freshwater, cropland, and Fertiliser use. Sci Total Environ **438**, 477–489 (2012)
32. FAO, IFAD, UNICEF, WFP, and WHO. The State of Food Security and Nutrition in the World 2020. Technical report (2020)
33. Banholzer, S., Kossin, J., Donner, S.: The impact of climate change on natural disasters. In: Reducing Disaster: Early Warning Systems for Climate Change, vol. 9789401785, pp. 21–49. Springer (2014)
34. Schoennagel, T., Balch, J.K., Brenkert-Smith, H., Dennison, P.E., Harvey, B.J., Krawchuk, M.A., Mietkiewicz, N., Morgan, P., Moritz, M.A., Rasker, R., Turner, M.G., Whitlock, C.: Adapt tomore wildfire in western North American forests as climate changes (2017)
35. Hore, K., Kelman, I., Mercer, J., Gaillard, J.C.: Climate Change and Disasters. In: Handbook of Disaster Research, pp. 145–159. Springer International Publishing (2018)
36. Rahtore, A., Jasrai, Y.T.: Biodiversity : importance and climate change impacts. Int J Sci Res Publ **3**(3), 5 (2013)
37. Tim Newbold et al. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. Science **353**(6296), 288–291 (7 2016)
38. Ritchie, H., Roser, M.: CO2 and Greenhouse Gas Emissions. In: Our World in Data (2020)
39. Nanaki, Evanthia A., Koroneos, Christopher J.: Climate change mitigation and deployment of electric vehicles in urban areas. Renew Energy **99**, 1153–1160 (2016)
40. American Public Transportation Association. Public Transportation Reduces Greenhouse Gases and Conserves Energy (2008)
41. Banister, David: Cities, mobility and climate change. Journal of Transport Geography **19**(6), 1538–1546 (2011)
42. Silva Junior, C.H.L., Pessôa, A.C.M., Carvalho, N.S., Reis, J.B.C., Anderson, L.O., Aragão, L.E.O.C.: The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. Nat Ecol Evol **5**(2), 144–145 (2 2021)
43. Vieira, I.C.G., Toledo, P.M., Silva, J.M.C., Higuchi, H.: Deforestation and threats to the biodiversity of Amazonia. Braz J Biol **68**(4 suppl):949–956 (11 2008)
44. Organisation for Economic Co-operation and Development. Agriculture and Climate Change: Towards Sustainable, Productive and Climate-Friendly Agricultural Systems (2016)
45. Selvaraju, R.: Climate risk assessment and management in agriculture. In Building resilience for adaptation to climate change in the agriculture sector, pp. 71–90 (2012)
46. European Environment Agency. Agriculture and climate change (2015)

47. Arnell, N.W., Lowe, J.A., Challinor, A.J., Osborn, T.J.: Global and regional impacts of climate change at different levels of global temperature increase. Climatic Change **155**(3), 377–391 (2019)
48. Mimura, N.: Sea-level rise caused by climate change and its implications for society (2013)
49. Burck, J., Ursula, H., Christoph, B., Niklas, H., Leonardo, N.: Climate Change Performance Index. Technical report, Germanwatch, Bonn (2020)
50. European Commission. 2030 Climate & Energy Framework (2018)
51. European Commission. Land use and forestry regulation for 2021–2030 (2020)
52. European Commission. Biodiversity strategy for 2030 (2021)
53. European Commission. Energy Efficiency Targets (2014)
54. IQAir. World Air Quality Report. 2020 World Air Quality Report, (August):1–35 (2020)
55. United Nations Environmental Programme (UNEP). Legal Limits on Single-Use Plastics and Microplastics. Technical report (2018)
56. Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Lavender Law, K.: Plastic waste inputs from land into the ocean. Science **347**(6223), 768–771 (2015)
57. The World Bank. What a Waste Global Database (2018)
58. FAO. AQUASTAT Core Database
59. Ritchie, H.: Energy Production and Consumption. In: Our World in Data (2019)
60. Ritchie, H., Roser, M.: Renewable Energy. In: Our World in Data (2020)
61. Kieffer, G., Couture, T.D.: Renewable Energy Target Setting. Technical Report (June 2015)
62. Commission, European: 2019 assessment of the progress made by Member States towards the national energy efficiency targets for 2020. Technical report, Brussels (2020)
63. United Nations Environment Programme. Food Waste Index Report 2021. Technical report (2021)
64. Global Footprint Network. Ecological Footprint per Person (2019)
65. David, E., Vera, K., Laura, S.: Global climate risk index 2021. Technical report, Germanwatch, Bonn (2021)
66. Organisation for Economic Co-operation and Development. Protected Areas. In: OECD.Stat
67. European Commission. Fishing Quotas (2021)
68. Ritchie, H.: Where in the world do people have the highest CO2 emissions from flying? In Our World in Data (2020)
69. Ritchie, H.: Deforestation and Forest Loss. In: Our World in Data
70. Ritchie, H.: Forest area. In: Our World in Data
71. Organisation for Economic Co-operation and Development. Greenhouse gas emissions
72. Hannah Ritchie and Max Roser. Land Use. In: Our World in Data (2019)
73. Hannah Ritchie and Max Roser. Air pollution. In: Our World in Data (2017)