



Predictive modelling of the higher heating value in biomass torrefaction for the energy treatment process using machine-learning techniques

P. J. García Nieto¹ · E. García-Gonzalo¹ · J. P. Paredes-Sánchez² · A. Bernardo Sánchez³ · M. Menéndez Fernández³

Received: 30 January 2018 / Accepted: 9 November 2018 / Published online: 14 November 2018
© Springer-Verlag London Ltd., part of Springer Nature 2018

Abstract

Torrefaction of biomass can be described as a mild form of pyrolysis at temperatures typically ranging between 200 and 300 °C in the absence of oxygen. Common biomass reactions during torrefaction include devolatilization, depolymerization, and carbonization of hemicellulose, lignin, and cellulose. Torrefaction of biomass improves properties like moisture content as well as calorific value. The aim of this study was to obtain a predictive model able to perform an early detection of the higher heating value (HHV) in a biomass torrefaction process. This study presents a novel hybrid algorithm, based on support vector machines (SVMs) in combination with the particle swarm optimization (PSO) technique, for predicting the HHV of biomass from operation input parameters determined experimentally during the torrefaction process. Additionally, a multilayer perceptron network (MLP) and random forest (RF) were fitted to the experimental data for comparison purposes. To this end, the most important physical–chemical parameters of this industrial process are monitored and analysed. The results of the present study are two-fold. In the first place, the significance of each physical–chemical variables on the HHV is presented through the model. Secondly, several models for forecasting the calorific value of torrefied biomass are obtained. Indeed, when this hybrid PSO–SVM-based model with cubic kernel function was applied to the experimental dataset and regression with optimal hyperparameters was carried out, a coefficient of determination equal to 0.94 was obtained for the higher heating value estimation of torrefied biomass. Furthermore, the results obtained with the MLP approach and RF-based model are worse than the best obtained with the PSO–SVM-based model. The agreement between experimental data and the model confirmed the good performance of the latter. Finally, we expose the conclusions of this study.

Keywords Support vector machines (SVMs) · Particle swarm optimization (PSO) · Artificial neural networks (ANNs) · Higher heating value (HHV) prediction

1 Introduction

The world is currently facing the challenge to decrease dependence on fossil fuels and to achieve a sustainable, renewable energy supply [1]. Biomass can be an important energy source. Energy generated from biomass is taken into account carbon-neutral because the carbon dioxide released during conversion is already part of the carbon cycle [2, 3]. Increasing the use of biomass for energy can help to reduce greenhouse gas (GHG) emissions and meet the targets established in the Kyoto Protocol. Energy from biomass can be obtained from different thermochemical (combustion, gasification, and pyrolysis), biological

✉ P. J. García Nieto
lato@orion.ciencias.uniovi.es

¹ Department of Mathematics, Faculty of Sciences, University of Oviedo, 33007 Oviedo, Spain

² Department of Energy, College of Mining, Energy and Materials Engineering, University of Oviedo, 33004 Oviedo, Spain

³ Department of Mining Technology, Topography and Structures, University of León, 24071 León, Spain

(anaerobic digestion and fermentation), or chemical (esterification) processes, where direct combustion can provide short-term energy solution. The growing interest in biomass as a solid fuel takes into account combustion to produce steam for electrical power and commercial plant uses, gasification to give rise to combustible gas (large partial pressure of nitrogen and CO₂, called producer gas) as well as syngas (carbon monoxide and hydrogen with low amounts of nitrogen and CO₂).

Furthermore, some of the inherent problems of raw biomass materials compared to fossil fuel resources (low bulk density, high moisture content, hydrophilic nature, and low calorific value) make raw biomass difficult to use on a large scale. To fix ideas, nature provides a large diversity of biomass with varying characteristics. These limitations greatly impact logistics and final energy efficiency. High moisture in raw biomass is one of the primary challenges, as it decreases the efficiency of the process and increases fuel production costs. High moisture content in biomass leads to natural decomposition, turning out in loss of quality and storage issues such as off-gas emissions. In order to overcome these challenges and make biomass suitable for energy applications, the material must be pre-processed. To create highly efficient biomass-to-energy chains, torrefaction of biomass in combination with densification (pelletisation or briquetting) is a promising step to overcome logistic economics in large-scale sustainable energy solutions, i.e. make it easier to transport and store it. Pellets or briquets are lighter, drier, and more stable in storage than the biomass they derive from.

Torrefaction, which is a thermochemical pretreatment process, is a viable technology that significantly alters the physical and chemical composition of the biomass [4–6]. Several studies have defined the characteristics of the torrefaction and its process [7–9]. Torrefaction of biomass (e.g. wood or grain) is a mild form of pyrolysis at temperatures typically between 200 and 320 °C in an inert environment [10]. Torrefaction changes biomass properties to provide a much better fuel quality for combustion and gasification applications [11–13]. Furthermore, torrefaction leads to a dry product with no biological activity like rotting. Torrefaction makes the material undergo Maillard reactions [10–13]. The importance of the torrefied wood pellet manufacturing industry is increasing due to the demand of a product with an increased energy density, easy to handle and to transport, and practical to co-fire in existing coal plants. Specifically, waste biomass has been recognised as a renewable and promising source for bio-fuels via different conversion processes [14]. Throughout the Earth there are million hectares of tree plantation forests and a total annual residue of million cubic meters produced from forest harvesting, wood processing plants and pulp and paper mills. Management and utilization of

these resources is the target of both most EU governments and industries to achieve goals established by Kyoto and Paris Protocols. Thermal treatment, i.e. torrefaction, is considered as a promising method to improve waste properties for energy applications [15]. When it is applied to woody biomass, this process removes water and volatiles through the partial decomposition of hemicellulose. Figure 1 shows the typical flowchart of the biomass torrefaction process.

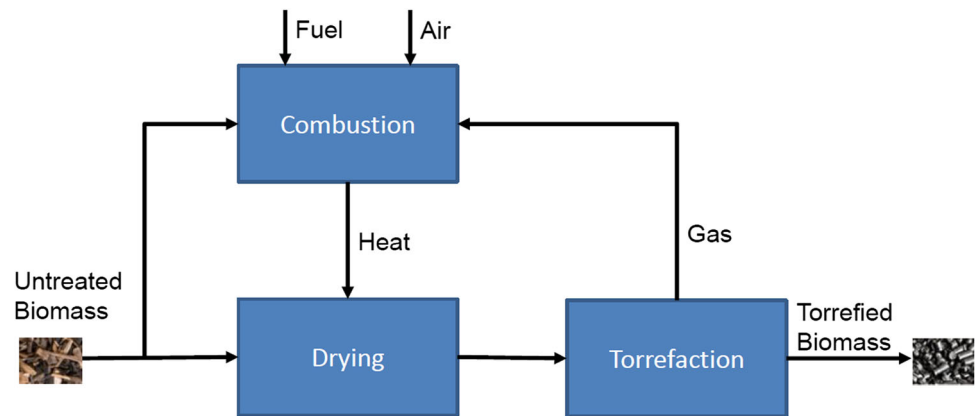
To fix ideas, the main purpose of this study is to estimate the application of the support vector machines (SVMs) approach in combination with the evolutionary optimization technique known as particle swarm optimization (PSO) as well as the multilayer perceptron (MLP) and random forest (RF) technique to identify the higher heating value (HHV) in the biomass torrefaction process, comparing the results obtained [16–24]. SVM models are based on the statistical learning theory and are a new class of techniques that can be used for predicting values from very different fields [19, 21, 23]. SVMs are associated to controlled supervised learning procedures used for grouping and regression, whose skill is well-recognized for being universal approximators of any multivariate functions to a high level of accuracy. The theory of statistical learning and minimization of structural risk are theoretical fundamentals for the learning algorithms of SVMs [20, 24].

In order to carry out the optimization stage corresponding to the kernel optimal hyperparameters setting in the SVM training, we have implemented the particle swarm optimization (PSO) technique with successful results. Indeed, particle swarm optimization (PSO) is one of the oldest swarm intelligence (SI) methods relied on bio-inspired algorithms [25]. The PSO technique is a population-based search algorithm based on the simulation of the bird flocking [26–28]. PSO uses the model of information sharing just like other SI-based evolutionary algorithms computations like Ant Colony Optimization (ACO) [29, 30], Artificial Bee Colony (ABC) technique [31–34], among others.

With this aim, hybrid PSO optimized SVM (PSO–SVM) models were used as automated learning tools, training them in order to predict the HHV of torrefied biomass from other operation parameters in a biomass torrefaction process. The SVM technique has been proven to be an effective tool to predict natural parameters. Researchers have successfully used it in a wide variety of environmental fields like forest modelling [35], estimation of solar radiation [36, 37], prediction of the air quality [38], study of water properties [39, 40] among others.

In addition, a multilayer perceptron (MLP) is a model of artificial feedforward neural network which maps sets of input data onto a set of proper outputs. The statistical learning theory relies on the MLP technique, which is a

Fig. 1 Typical flowchart of the biomass torrefaction process



new type of mathematical technique suitable to prognosticate values in quite a variety of fields [17, 41]. An MLP is made up of multiple layers of nodes in a directed diagram, with each layer fully connected to the next one. Excluding the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP uses a backpropagation, which is a supervised learning technique to train the network. MLP is a variation of the standard linear perceptron with the capacity to tell apart data that are not linearly separable from the separable ones.

Based on decision trees and combined with aggregation and bootstrap ideas, random forests (RF) technique, were introduced by Breiman [42]. They are a powerful non-parametric statistical method allowing to consider regression problems as well as two-class and multiclass classification problems, in a single and versatile framework. The RF algorithm presents several advantages [20, 42, 43]: (1) it runs efficiently on large datasets; (2) it is not sensitive to noise or over-fitting; (3) it can handle thousands of input variables without variable deletion; and (4) it has fewer parameters compared with that of other machine-learning algorithms like MLP or support vector regression (SVR).

In summary, a hybrid PSO optimized SVM (PSO–SVM) model as well as a MLP model and a RF-based model were used as automated learning tools, training them in order to foretell the HHV in a biomass torrefaction process from the operation physical–chemical input parameters measured experimentally.

The HHV is an essential property of advanced biomass fuels for the design of future thermal conversion systems. The torrefaction is a complex chemical process to upgrade biomass [44, 45]. Empirical and semi-empirical correlations are available in the literature to estimate the HHV of biomass fuels based on basic analysis data obtained from chemical analysis composition. To approximate the HHV of biomass various advanced empirical models have been developed based on basic analysis data by several researchers [46–48].

Ghugare et al. [49] showed a genetic programming (GP) for developing HHV prediction models, respectively, using the parameters of the proximate and ultimate analyses as the model’s input variables. Estiati et al. [50] and Ozveren [51] evaluate the feasibility of using artificial neural networks (ANN) and empirical correlations in HHV. However, the application of a hybrid algorithm from the operation parameters determined experimentally during the torrefaction process has been not applied to define HHV until now. This study tries to cover this gap.

This study is organized as follows. Firstly, we describe the necessary materials and methods to carry out this study. Secondly, we show and discuss the obtained results. Finally, we expose the main conclusions drawn from the results.

2 Materials and methods

The most common methods currently used to calculate the heating of biomass are the formula derived by Dulong, or experimentally, the use of a bomb calorimeter which is burdensome [52]. There have been numerous mathematical equations, which were created based on data from biomass composition, such as proximate or elemental analyses of biomass [53]. For example, Demirbas has contributed greatly in the estimation of calorific values of biomass fuels with thermal studies [54].

2.1 Experimental dataset

The biomass dataset used in this research work is a collection of experimental elemental compositions and their corresponding HHVs. The characterisation of torrefied biomass involves a great variety of raw materials. In this sense, the data have been determined according to conditions established by ASTM standard rule and taking as validation reference a research database [55]. In this sense, the data used for the SVM analysis were collected from a

certified laboratory and the total number of data processed was about 300 values. These values are obtained, in the context of previous studies, from the analysis of over one hundred samples, which include both raw and torrefied biomass. Moreover, in order to make decisions in the conversion chains for bioenergy and bioproducts from torrefaction, researchers and decision makers need models to envision the viable alternatives and evaluate their performance [56–60].

It has been considered lignocellulosic biomass in the torrefaction process from agricultural and forestry activity to define a flexible model useful for solid residual biomass. This hypothesis has permitted to define a nonconditioned model of biomass type and to get a flexible model. With respect to the input variables, we have selected the main ones controlled in the biomass torrefaction process. The dataset used for the three different models (PSO–SVM-based model, MLP-based model and RF-based model) is relied on several physical–chemical parameters. The use of biomass as a fuel in thermal applications requires knowledge of its heating value (HV) [61]. The HV reflects the energy content of a fuel in a standardised fashion and it is often expressed as the HHV [62]. Indeed, the HHV refers to the heat released by the complete combustion of a unit volume of fuel leading to the production of water vapour and its eventual condensation. The values of HHV are usually expressed on the basis of dry weight (i.e. dry basis). The experimental or direct determination of the HV of a biomass used to measure the change in enthalpy between the reactants and product by an adiabatic bomb calorimeter is very expensive. Therefore, the determination of its elemental and especially its proximate composition is a good alternative. In this way, proximate analysis is used to determine the weight percentage of moisture, volatile material (VM), fixed carbon (FC) and ash in a biomass, while elemental analysis involves the determination, again in weight percentage, of C, H, N or O [63].

The main driven input parameters in the biomass torrefaction process have been previously defined by Chen et al. [12]. Hence, the physical–chemical input variables of the model are as follows [4–13]:

- Fixed carbon (FC): it is the difference between the sum of volatile matter and ash contents by percentage.
- Volatile matter (VM): it refers to the components of coal, except for moisture, which are liberated at high temperature in the absence of air.
- Reaction temperature (RT): it is the temperature for the torrefaction process.
- Residence time (Rt) (or torrefaction process time): it is the average amount of time that a particle stays in torrefaction conditions.

- Atomic O/C ratio: it is the empirical relation between oxygen and carbon contents.
- Atomic H/C ratio: it is the empirical relation between hydrogen and carbon contents.

2.2 Support vector machine (SVM) method

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression [18–24]. Note that SVMs were originally developed for classification, and were later generalized to solve regression problems [35, 36, 40, 64]. This last method is called *support vector regression* (SVR). The regression function $y = f(\mathbf{x})$ for a given data set $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, is obtained by SVR in the linear form as follows:

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b \tag{1}$$

where \mathbf{w} and b are, respectively, the weight vector and intercept of the model, and they need to be obtained by an optimal fitting of the dataset available in D . In nonlinear cases, it is mandatory to start by mapping the input low-dimensional vectors with a nonlinear function $\Phi : \mathbb{R}^p \rightarrow F$, where F is the feature space of Φ [37–39, 65, 66]. Afterwards, it is possible to write the regression function as follows [19, 22]:

$$f(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}) + b \tag{2}$$

Next, SVR introduces the ε -insensitive loss function given by [18, 20, 23]:

$$d = |y - f(\mathbf{x})|_\varepsilon = \begin{cases} 0, & |y - f(\mathbf{x})| \leq \varepsilon \\ |y - f(\mathbf{x})| - \varepsilon & \text{otherwise} \end{cases} \tag{3}$$

which does not take into account the error if the difference between the prediction value calculated by Eq. (2) and the real value is smaller than ε . Figure 2 illustrates the ε -insensitive loss function. Moreover, the vector \mathbf{w} and

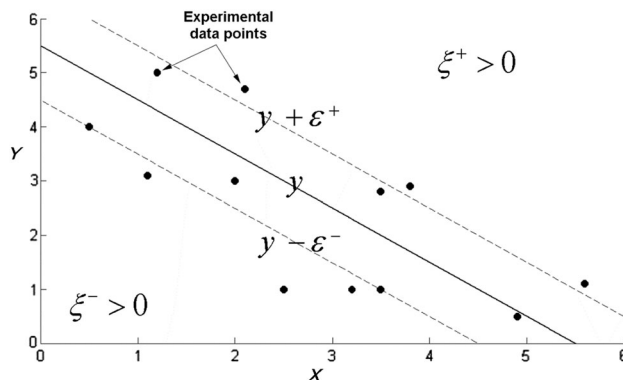


Fig. 2 Regression with ε -insensitive tube for one-dimensional problem

coefficient b can be found by solving a convex optimization problem, which balances the empirical error and the generalization ability. In SVR, the empirical error is measured by means of the ε -insensitive loss function and the generalization ability is related to the Euclidean norm of \mathbf{w} [20, 23]. Thus, the optimization problem for the identification of the regression model can be written in the following way [18–23]:

$$\begin{aligned} \text{Minimize } J(\mathbf{w}, \xi_i^+, \xi_i^-) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi_i^+ + \xi_i^-) \\ \text{s.t. } \begin{cases} y_i - \mathbf{w}^T \Phi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i^+ \\ \mathbf{w}^T \Phi(\mathbf{x}_i) + b - y_i - b \leq \varepsilon + \xi_i^- & i = 1, \dots, n \\ \xi_i^+, \xi_i^- \geq 0 \end{cases} \end{aligned} \tag{4}$$

where C expresses the *parameter of penalty or cost* between empirical and generalization errors and ξ_i^+, ξ_i^- represent the slack variables defined in Fig. 2. The solution to this quadratic optimization problem (QP) by Lagrangian dual method [23, 67] provides the numerical method to calculate the prediction value:

$$f(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}) + b = \sum_{i=1}^n (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}, \mathbf{x}_i) + b \tag{5}$$

where α_i^+, α_i^- represent the Lagrange multipliers of the dual form of the optimization problem and $K(\mathbf{x}_i, \mathbf{x}_j)$ represents the kernel function [23, 67], and can be described as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j) \tag{6}$$

The kernel function defines the space feature, where data are regressed. Hence, selecting an appropriate one during the SVM regression is essential. There is some description of the different kernel functions used here in the bibliography [21, 23, 67]:

- Radial basis function (RBF kernel):

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\sigma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \tag{7}$$

- Polynomial kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\sigma \mathbf{x}_i \cdot \mathbf{x}_j + a)^b \tag{8}$$

- Sigmoid kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\sigma \mathbf{x}_i \cdot \mathbf{x}_j + a) \tag{9}$$

where a, b and σ are parameters defining the kernel’s behaviour.

In conclusion, we have to choose a kernel and relevant hyperparameters able to map the nonlinearly separable data

into a feature space where they are linearly separable if we want to solve a regression problem for data that are not linearly separable with an SVM.

2.3 Particle swarm optimization (PSO) approach

The particle swarm optimization (PSO) approach is a mathematical optimization/search technique of metaheuristic kind [25]. Usually the PSO is used in search spaces with many dimensions. The PSO methods were originally attributed to researchers Kennedy, Eberhart and Shi [26]. At first they were conceived to elaborate models of social behaviour, such as the movement described by living organisms in a flock of birds or a bank of fish. The algorithm was then simplified and proved to be suitable to solve optimization problems. PSO allows the optimization of a mathematical problem using a population of candidate solutions, denoted as *particles*, moving throughout the search space according to mathematical rules that take into account the position and velocity of the particles. The motion of each particle is influenced by its best local position so far, as well as by the best global positions encountered by other particles as the particles travel through the search space. The theoretical basis of this performance is to make the particle cloud converge quickly to the best solutions. Furthermore, PSO is a metaheuristic technique, as it assumes no hypotheses about the problem to be optimized and can be applied in large spaces of candidate solutions.

Let S be the number of particles in the cloud, each of which has a position $\mathbf{x}_i \in \mathfrak{R}^n$, in the search space and a speed $\mathbf{v}_i \in \mathfrak{R}^n$. Similarly, we will represent the initial position of the particle as \mathbf{x}_i^0 and its velocity as \mathbf{v}_i^0 , both chosen randomly. The best positions correspond to the best values of the fitness function evaluated for each particle. Positions and velocities of each particle are updated taking into account these values as follows [27, 28, 64, 68, 69]:

$$\mathbf{v}_i^{k+1} = \omega \mathbf{v}_i^k + \phi_1 (\mathbf{g}^k - \mathbf{x}_i^k) + \phi_2 (\mathbf{I}_i^k - \mathbf{x}_i^k) \tag{10}$$

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1} \tag{11}$$

Three are the components on which the velocity of each particle, i , at iteration k , relies on: (a) the constant *inertia weight* ω , concerned by the velocity term in iteration k, \mathbf{v}_i^k ; (b) the term termed *cognitive learning*, which is the difference between the particle’s current position \mathbf{x}_i^k and the particle’s best position found up until now (called \mathbf{I}_i^k , local best); and (c) the term of *social learning*, which is the difference between the global best position found up until now in the whole swarm (called \mathbf{g}^k , global best) and the particle’s current position \mathbf{x}_i^k . The two last terms are concerned in Eq. (10) by factors $\phi_1 = c_1 r_1$ and $\phi_2 = c_2 r_2$. In

these two multipliers, c_1 and c_2 are constants, while r_1 and r_2 are random numbers distributed uniformly in the interval $[0, 1]$. Also, the Standard PSO 2011 [70] has been used in this study. This involves some improvements with respect to the preliminary implementations [26–28, 70]. Therefore, the PSO parameters are chosen here as:

$$\omega = \frac{1}{2 \ln 2} \quad \text{and} \quad c_1 = c_2 = 0.5 + \ln 2 \quad (12)$$

The swarm topology determines how particles are connected among them to interchange information with the global best. In the actual Standard PSO each particle reports only K particles, usually three, randomly chosen.

2.4 Multilayer perceptron neural network

Artificial neural networks (ANNs) are a computational model based on a large set of simple neuronal units (artificial neurons), approximately similar to the behaviour observed in axons of neurons in biological brains [17, 41]. The multilayer perceptron (MLP) is a kind of artificial neural network (ANN) made up of multiple layers that allows to solve problems that are not linearly separable. Indeed, the multilayer perceptron (MLP) consists of an input layer and an output layer and one or more hidden layers of nonlinearly activating nodes [41]. It is a modification of the standard linear perceptron so that it uses three or more layers of neurons (nodes) with nonlinear activation functions (Fig. 3).

This kind of artificial neural network implements the function $\mathbf{f} : \mathbf{X} \subset \mathbb{R}^n \rightarrow \mathbf{Y} \subset \mathbb{R}^c$, which can be written as follows [17]:

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \phi(\psi(\mathbf{x})) = (\phi \circ \psi)(\mathbf{x}) \\ \phi : \mathbf{X} \subset \mathbb{R}^n &\rightarrow \mathbf{U} \subset \mathbb{R}^m \\ \psi : \mathbf{U} \subset \mathbb{R}^m &\rightarrow \mathbf{Y} \subset \mathbb{R}^c \end{aligned} \quad (13)$$

so that \mathbf{U} is the space of hidden variables, known as the characteristics space. Based on the defined architecture, it

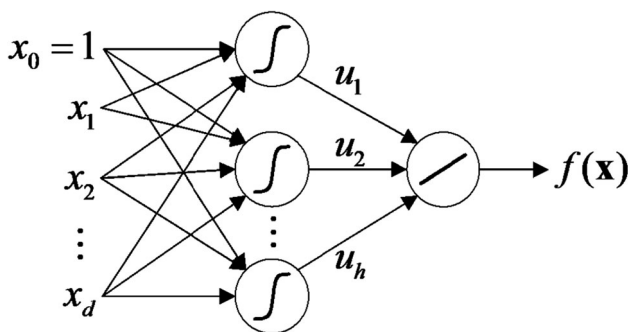


Fig. 3 Diagram of an MLP network with h neurons in the hidden layer, d neurons in the input layer and a single neuron in the output layer

is possible to write [41] that $\psi_j(\mathbf{x}) = \psi(\mathbf{w}_j^T \mathbf{x} + w_{j0})$, where ψ is the activation function for the hidden layer units; $w_{j0} \in \mathbb{R}$ is its threshold value and $\mathbf{w}_j \in \mathbb{R}^n$ is the vector of parameters for the units. The ψ function can be sigmoid, logistical or hyperbolic tangent. On the other hand, $\phi_j(\mathbf{u}) = \phi(\mathbf{c}_j^T \mathbf{u} + c_{j0})$, where ϕ is the activation function of the output layer units, $\mathbf{c}_j \in \mathbb{R}^m$ is the vector of weights for the units and $c_{j0} \in \mathbb{R}$ is its threshold value. Similarly, ϕ is the identity function, Heaviside function or any dichotomous function, normally used as the activation function.

The MLP gives places to the following implemented function [17, 41]:

$$f(\mathbf{x}) = \sum_{j=1}^m c_j \psi(\mathbf{w}_j^T \mathbf{x} + w_{j0}) + c_0 \quad (14)$$

2.5 Random forest regression algorithm

The random forest (RF) regression algorithm is an ensemble-learning algorithm that combines a large set of regression trees. A regression tree represents a set of conditions or restrictions that are hierarchically organized and successively applied from a root to a leaf of the tree [71–75]. The RF begins with many bootstrap samples that are drawn randomly with replacement from the original training dataset. A regression tree is fitted to each of the bootstrap samples. For each node per tree, a small set of input variables selected from the total set is randomly considered for binary partitioning. The regression tree splitting criterion is based on choosing the input variable with the lowest Gini Index:

$$I_G(t_{X(x_i)}) = 1 - \sum_{j=1}^m f(t_{X(x_i)}, j)^2 \quad (15)$$

where $f(t_{X(x_i)}, j)$ is the proportion of samples with the value x_i belonging to leaf j as node t [74, 75]. The predicted value of an observation is calculated by averaging over all the trees. Two parameters need to be optimized in the RF: (1) ntree: the number of regression trees (default value is 500 trees); and (2) mtry: the number of input variables per node (default value is 1/3 of the total number of variables).

2.6 Accuracy measures

In this study, in order to predict the HHV in a biomass torrefaction process using six different kernels, a SVR technique in combination with PSO approach has been built [18, 23, 67]: linear, superlinear, quadratic, cubic, sigmoid and RBF kernel functions. Additionally, a multi-layer perceptron (MLP) approach and random forest (RF) technique were also fitted with comparison purposes.

The goodness of fit and performance of the studied models for the HHV estimation in a biomass torrefaction process were evaluated and compared using four commonly used statistical indicators, i.e. coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and mean bias error (MBE). The mathematical equations of the statistical indicators are described below.

A dataset takes values t_i , each of which has an associated modelled value y_i . The former are called the observed values and the latter are often referred to as the predicted values. Variability in the dataset is calculated through different sums of squares:

- $SS_{\text{tot}} = \sum_{i=1}^n (t_i - \bar{t})^2$: the total sum of squares, proportional to the sample variance.
- $SS_{\text{reg}} = \sum_{i=1}^n (y_i - \bar{t})^2$: the regression sum of squares, also called the explained sum of squares.
- $SS_{\text{err}} = \sum_{i=1}^n (t_i - y_i)^2$: the residual sum of squares.

In the prior sums, \bar{t} is the mean of the n observed data:

$$\bar{t} = \frac{1}{n} \sum_{i=1}^n t_i \tag{16}$$

Bearing in mind the above sums, the general definition of the coefficient of determination is [76, 77]:

$$R^2 \equiv 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}} \tag{17}$$

A coefficient of determination value of 1.0 points out that the regression curve fits the data perfectly. Similarly, the mathematical expressions for the other three statistics used in this study (RMSE, MAE and MBE) are as follows [76, 77]:

$$\text{RMSE} \equiv \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - y_i)^2} \tag{18}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |t_i - y_i| \tag{19}$$

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (t_i - y_i) \tag{20}$$

Higher values of R^2 are preferred, i.e. closer to 1 means better model performance and regression line fits the data well. Conversely, the lower the RMSE, MAE and absolute MBE values are, the better the model performs.

3 Results and discussion

The operation physical–chemical input variables considered in this research work are shown in Table 1. The total number of predicting variables used to construct the hybrid PSO–SVM-based model, MLP approach and RF-based model was 6. Moreover, the output predicted variable is the HHV of torrefied biomass consequence of a previous torrefaction process. Indeed, we have built different models (specifically, the PSO–SVM-based model, the MLP approach and RF-based model) taking as dependent variable the calorific value of biomass after a torrefaction process.

Additionally, it is well known that the SVM techniques depend strongly on the SVM hyperparameters namely: the regularization factor C (see Eq. 4); the hyperparameter ε that defines the ε -insensitive tube (allowable error); a , b and σ that represents the parameters that depend on the chosen kernel function. There exist a vast body of literature in relation to the choice of hyperparameters for SVMs [18, 23]. Some methods often utilized to establish suitable hyperparameters are [23, 67]: grid search, genetic algorithms, artificial bee colony (ABC) and so on. The grid search is a brute force method and, as such, almost any optimization method increases its efficiency.

Hence, bearing this in mind, we have chosen PSO optimization technique as a suitable, efficient and simple method [27, 28] for tuning the SVR parameters and a novel hybrid PSO–SVM-based model was applied to evaluate the HHV (output variable) from the other six remaining variables (input variables) in order to predict the calorific value of torrefied biomass in a torrefaction process [5, 6, 56–60], studying their influence in order to optimize its calculation. In this way, Fig. 4 shows the flowchart of the parameter selection process used over the training dataset for the PSO–SVR-based model with cubic kernel.

The particles \mathbf{x}_i are vectors that contain the parameters to tune: for instance, $\mathbf{x}_i = (C_i, \varepsilon_i, \sigma_i, a_i, b_i)$ for the cubic kernel. In the first iteration, we initialize them randomly. According to the PSO algorithm, as described previously, the particles for the next iteration are calculated so that the objective function value for the particles is calculated in each step. Specifically, the objective function value is the minus five-fold cross-validation coefficient of determination for each particle or proposed set of solution parameters [78, 79]. The dataset has been splitted in training and testing datasets with a proportion of 80% and 20% of the samples, respectively. The training dataset is divided in 5 sets. We train the SVR with a set composed of 4 of this sets and obtain a model. This model is checked with the remaining fold of the training dataset and its coefficient of determination is obtained. This is repeated for each of the 5

Table 1 Set of operation physical–chemical input variables used in this study and their names along with their mean, standard deviation, median and MAD (median absolute deviation)

Input variables	Name of the variable	Mean	Standard deviation	Median	MAD
Fixed carbon (wt%)	FC	26.50	9.42	23.90	7.71
Volatile matter (wt%)	VM	68.32	12.47	70.40	10.09
Reaction temperature (°C)	RT	264.33	28.79	270.00	23.57
Residence time (min)	Rt	48.21	34.83	30.00	23.28
Atomic O/C ratio	O_C	0.47	0.13	0.50	0.10
Atomic H/C ratio	H_C	1.24	0.23	1.29	0.18

wt% means *weight percentage*

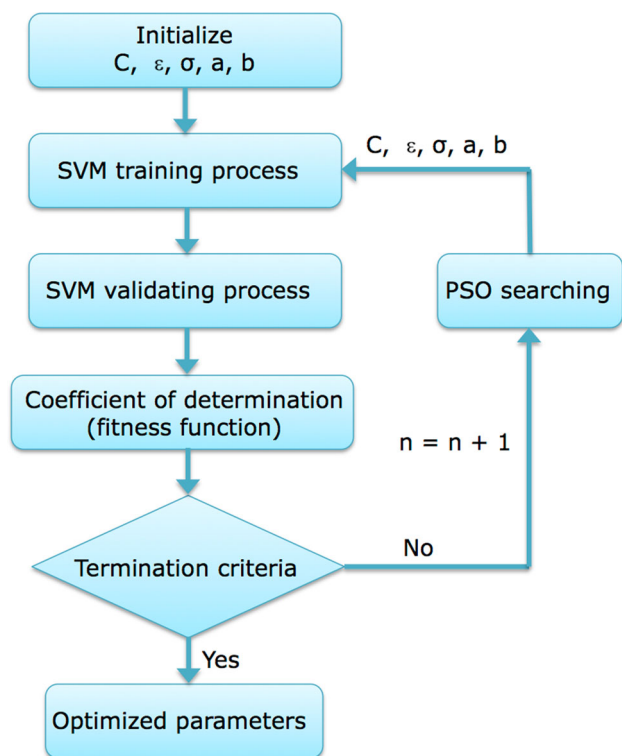


Fig. 4 Flowchart of the parameter selection process that uses the training dataset for the PSO–SVM-based model with cubic kernel

sets and the mean coefficient of determination is considered the one associated to this particle or possible solution. If the termination criteria are met, the global best \mathbf{x}_i contains the optimized parameters. Once the optimal parameters have been found we proceed to generate the model with the training data. This model is then tested with the 20% of the data that has been held out entirely of the previous process, and different goodness-of-fit indexes are obtained (R^2 , RMSE, MAE and MBE).

As per this reasoning, the parameters have been optimized with PSO, using the standard PSO 2011 version [70] and the regression modelling has been carried out with SVR- ε using the LIBSVM library [80].

The searching in the parameter space has been made keeping in mind that the SVM algorithm changes its results

significantly when its parameters increase or decrease a power of 10. For instance, we have worked with powers of ten and the searched parameters have been the exponents (except for the parameter b) being the five-dimensional search space. For instance, in case of cubic kernel $[-6, 4] \times [-10, 0] \times [-6, 1] \times [-6, 0]$. That is, C values (regularization parameter) in $[10^{-6}, 10^4]$, ε values in $[10^{-10}, 10^0]$, σ values in $[10^{-6}, 10^1]$ and a values in $[10^{-6}, 10^0]$ are utilized in the optimization phase. The stopping criterion is fulfilled if there is no improvement in the R^2 after ten iterations, along with a maximum number of iterations equal to 500.

Thus, in order to optimize the SVM parameters, the PSO module is used. The PSO looks for the best hyperparameters (i.e. C , σ , ε and a for cubic kernel) by comparing the error in every iteration. For example, the search space is organized in four dimensions for the cubic kernel function, one for each parameter. The optimal hyperparameters for the six types of kernels analysed here can be consulted in “Appendix”.

Likewise, and for comparison purposes, in order to predict its calorific value in a torrefaction process a multilayer perceptron (MLP) and RF-based model have been fitted to the experimental data corresponding to the HHV of torrefied biomass. Weka software has been used for these two methods [81, 82] and grid search was used to tune its parameters. In this sense, an artificial neural network (ANN) is typically determined by three types of parameters [17, 41]: the learning process for updating the weights of the interconnections (e.g. learning rate, searched in $[10^{-3}, 10^0]$), the interconnection pattern among different layers (tuned using from 2 to 12 layers) of neurons (see Fig. 3), the activation function that converts a neuron’s weighted input to its output activation and the momentum factor, searched in $[10^{-10}, 10^0]$, in order to avoid oscillating weight changes of the ANN. Similarly, the ANN optimal parameters for the multilayer perceptron (MLP) can be also consulted in “Appendix”. Analogously, for random forest (RF) technique, the number of input variables (mtry) tried at each split was searched between 1 and 6, and a value of 4 was the optimal value. The number of trees (ntree) in the

search phase varied between 1 and 1000 and 94 was the best value.

Moreover, Table 2 depicts the determination and correlation coefficients for the PSO–SVM-based models for the six kernels (linear, superlinear, quadratic, cubic, sigmoid and RBF kernels, respectively), multilayer perceptron and RF-based model fitted here for the HHV in a biomass torrefaction process.

According to the statistical calculations, the SVM with the cubic kernel function is the best model to evaluate the HHV of torrefied biomass, since the fitted SVM with superlinear kernel function has a coefficient of determination R^2 equals 0.9427. These results point out an important goodness of fit, that is to say, a very good agreement is achieved between our model and the observed data.

Every assessment of the importance of the variables in the prediction of an independent variable is tentative and has a relative value. As an indicator of the relative importance of the independent variables, the weights of the SVM–PSO-based model with linear kernel have been chosen. They are shown in Table 3. They have been scaled dividing them by the greatest weight in absolute value and multiplying it by one hundred and are shown in Fig. 5.

Hence, for the PSO–SVM model, the most significant variable in HHV prediction is the Volatile matter followed by *Fixed Carbon*, Atomic O/C ratio, Reaction temperature, Atomic H/C ratio and, finally, Residence time.

Volatile matter is an important component used for the quality characterization of biomass. Indeed, volatile matter is a key component of solid fuels used to measure combustion characteristics [83, 84]. The second variable in the ranking is the carbon content that is a leading parameter in the common analysis in the HHV of a conventional fuel according to Dulong formula [52]. The third most

Table 2 Coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE) and mean bias error (MBE) for the test data for the hybrid PSO–SVM-based models (with linear, superlinear, quadratic, cubic, sigmoid and RBF kernels), multilayer perceptron (MLP) approach and RF-based model fitted in this study for the higher heating value (HHV) in a biomass torrefaction process

Model	R^2	RMSE	MAE	MBE
Linear-SVM	0.8623	0.6115	0.4857	– 0.1230
Superlinear-SVM	0.8587	0.6195	0.4896	– 0.1384
Quadratic-SVM	0.7864	0.7618	0.5708	– 0.1107
Cubic-SVM	0.9427	0.3944	0.3193	0.0012
Sigmoid-SVM	0.8062	0.7257	0.5644	– 0.1597
RBF-SVM	0.9081	0.4995	0.3997	0.0455
Multilayer perceptron	0.9046	0.5092	0.4202	0.0424
Random Forest	0.5886	1.0571	0.8642	0.0671

Table 3 Weights of the variables in the fitted PSO–SVM-based model with linear kernel for the HHV of torrefied biomass

Variable	Weight
Volatile matter (wt%)	0.9892
Fixed carbon (wt%)	0.9090
Atomic O/C ratio	– 0.6259
Reaction temperature (°C)	0.1755
Atomic H/C ratio	0.1420
Residence time (min)	0.0538

wt % means weight percentage

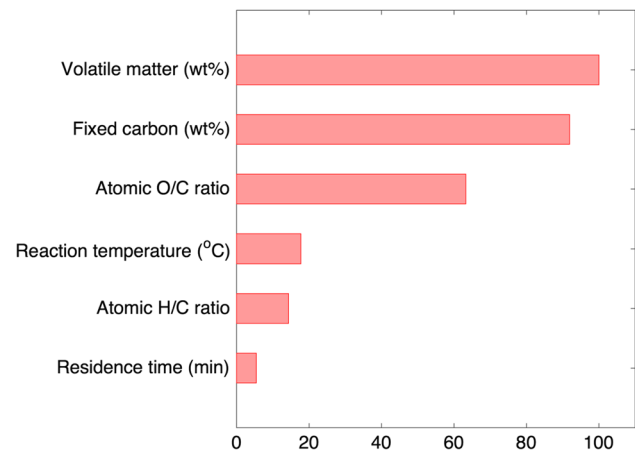


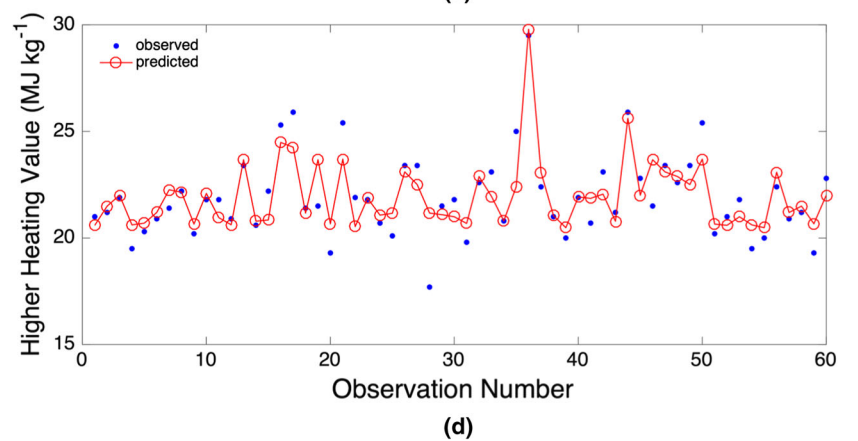
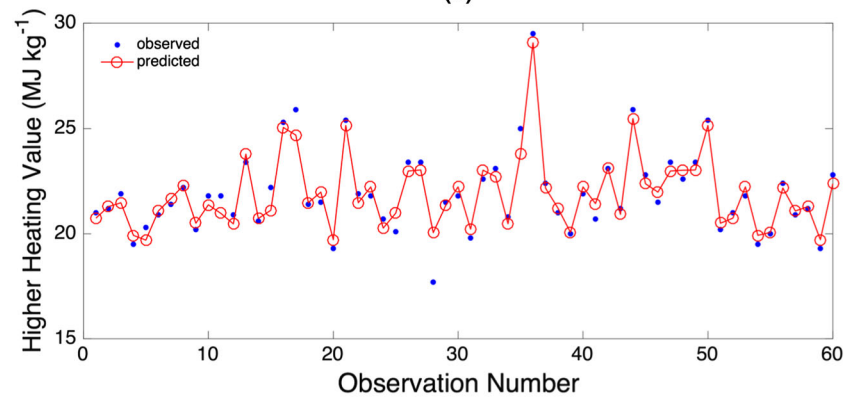
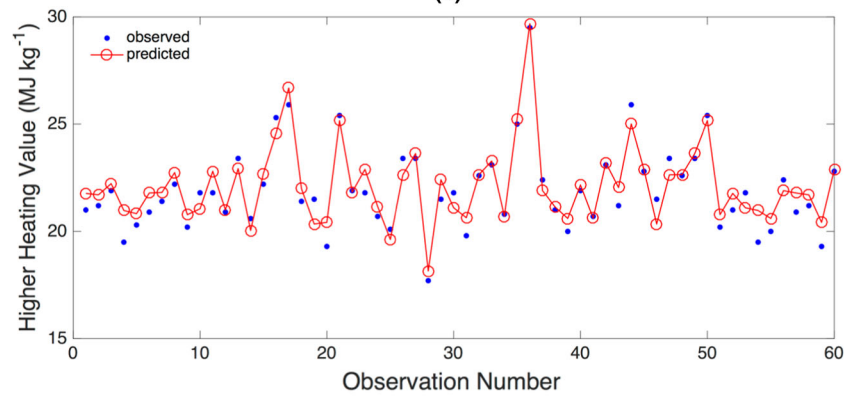
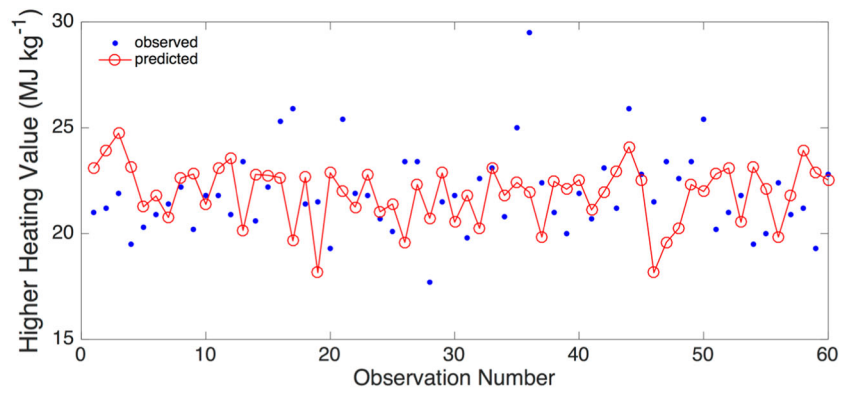
Fig. 5 Relative importance of the input variables to predict the HHV of torrefied biomass in the fitted PSO–SVM-based model with linear kernel

important variable, atomic O/C ratio, is the quotient between oxygen and carbon contents. As the carbon content in the biomass grows with the temperature and longer residence times, while hydrogen and oxygen contents decrease, the HHV increases while the corresponding O/C and H/C atomic ratios decrease.

The reaction temperature has a lower influence in the HHV because the reaction takes place within a thin range of temperatures that cannot be altered. If atomic H/C ratio grows, HHV increases slightly due to its implication in the combustion reactions. Finally, the residence time is the least important input variable. This variable is less significant because after a threshold time, the torrefaction reaction is exhausted.

Additionally and for comparison purposes, we have constructed a reduced PSO–SVM-based model with cubic kernel that uses as independent variables Fixed carbon, reaction temperature and residence time. This has allowed us to obtain a simpler and reliable model to determine the HHV for future applications since a coefficient of determination equal to 0.9 has been obtained.

Fig. 6 Comparison between higher heating values (HHVs) observed and predicted by RF-based model, the MLP approach and PSO–SVM-based model: **a** RF-based model ($R^2 = 0.59$); **b** MLP network ($R^2 = 0.90$); **c** PSO–SVM-based model with cubic kernel ($R^2 = 0.94$); and **d** reduced PSO–SVM-based model with cubic kernel ($R^2 = 0.90$)



To sum up, this research work was able to forecast the HHV of torrefied biomass in agreement to the actual experimental values observed using the PSO–SVM-based model with great accurateness and success. Indeed, Fig. 6 shows the comparison between the HHV values observed and predicted by using RF-based model (see Fig. 6a), the MLP (see Fig. 6b) and PSO–SVM-based model with cubic kernel with all the independent variables (see Fig. 6c) and with only three independent variables: Fixed carbon, reaction temperature and residence time (see Fig. 6d). It is very recommendable the use of a SVM model with cubic kernel in order to obtain the best effective approach to nonlinearities of this regression problem.

Regarding the uncertainty of the parameters of the fitted methods, it has been studied, in particular, for the SVR with the cubic kernel and the MLP. In both cases, a region around the optimal parameters (see Tables 4, 5 in “Appendix”) has been taken into account. For SVR with a cubic kernel, 1000 sets with the four parameters of the method (C , ε , σ and a) have been generated randomly in the search space. The corresponding RMSEs for the testing set have been computed for each set. With this 1000 RMSE values, a one-sample Kolmogorov–Smirnov test has been performed and the result has been that it fails to reject the null hypothesis that the data comes from a standard normal distribution, against the alternative that it does not come from such distribution, at the 5% significance level. A histogram of the RMSE values with the fitted normal density function has been plotted (Fig. 7a). The same procedure has been applied to the three parameters of the MLP method (momentum, learning rate and the number of hidden layers) with similar results, that is, the RMSEs for the testing dataset that corresponds to the 1000 set of randomly chosen parameters also follow a normal distribution (Fig. 7b).

4 Conclusions

Based on the experimental and numerical results, the main findings of this research work can be mentioned as follows:

- Firstly, torrefaction changes biomass properties to provide a much better fuel quality for combustion and gasification applications. Indeed, the resultant solid fuel has a higher energy density, improved storage and handling properties. Consequently, the development of alternative energetical diagnostic techniques is very recommendable for energy management. In this way, the new hybrid PSO–SVM-based method with a cubic kernel function utilized in this research work is a very good choice to evaluate the HHV of the torrefied biomass.

- Secondly, a hybrid PSO–SVM-based model with a cubic kernel function was successfully implemented to predict the calorific value of torrefied biomass from the other measured input operation variables, in order to lower costs in the quality assessment of these biomass higher heating values.
- Thirdly, a high coefficient of determination equal to 0.94 was achieved when this hybrid PSO–SVM-based model with a cubic kernel function was applied to the experimental dataset corresponding to the torrefied biomass HHV.
- Fourthly, the significance order of the input variables

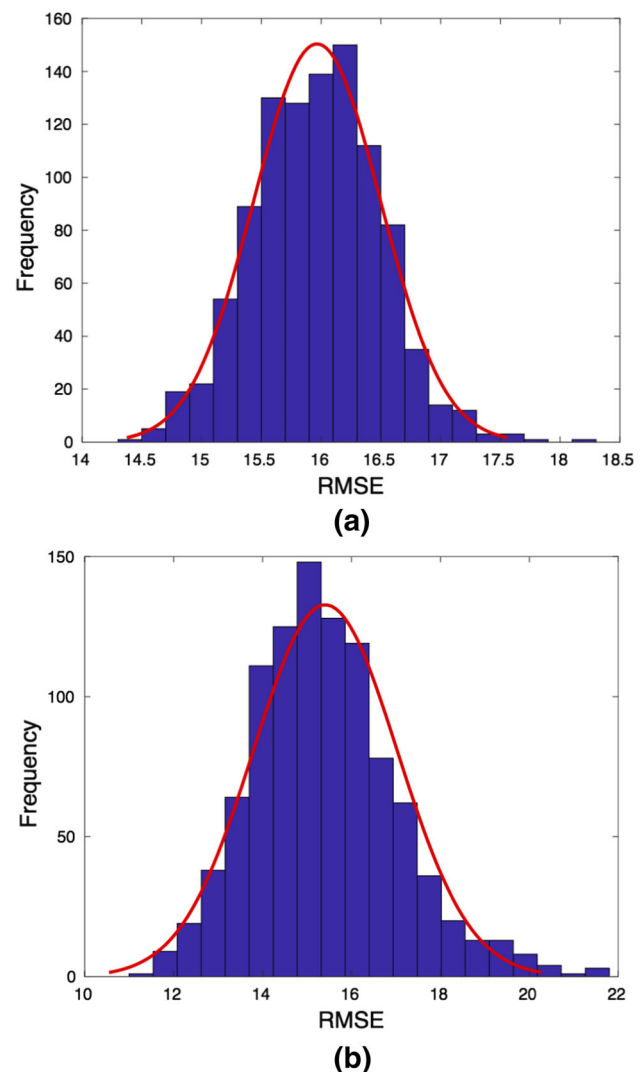


Fig. 7 Uncertainty analysis for specified parameters of the methods: **a** distribution of the RMSEs for the PSO–SVM method with cubic kernel; **b** distribution of the RMSEs for the MLP network

involved in the prediction of the calorific value of torrefied biomass was established. This is one of the main findings in this research work. Specifically, the

input operation variable volatile matter could be considered the most influential parameter in the prediction of the HHV in a biomass torrefaction process.

- Fifthly, the influence of the kernel parameters setting of the SVMs on the torrefied biomass calorific value’s regression performance was determined.
- Finally, the results indicate that the hybrid PSO–SVM regression method considerably improves the generalization capability achievable with only the SVM-based regressor.

In summary, this innovative methodology could be used in other biomass torrefaction processes with similar or different types of biomass with success, but it is always necessary to take into account the specificities of each biomass and experiment. Consequently, a PSO–SVM-based model is an effective practical solution to the problem of the determination of the torrefied biomass HHV in a biomass torrefaction process.

Acknowledgements Authors wish to acknowledge the computational support provided by the Department of Mathematics at University of Oviedo. Additionally, we would like to thank Anthony Ashworth for his revision of English grammar and spelling of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Appendix

Table 4 shows the optimal hyperparameters of the six fitted SVM-based models found with the particle swarm optimization (PSO) technique for the higher heating value (HHV) of torrefied biomass.

In this research work, the ANN optimal parameters for the multilayer perceptron (MLP) are depicted in Table 5.

Additionally, mean value and standard deviation of the coefficients of determination (R^2), corresponding to the different folds for the training data are shown in Table 6.

Table 5 ANN parameters of the fitted multilayer perceptron (MLP) of torrefied biomass in a torrefaction process for the higher heating value (HHV)

Parameters	Values
Number of hidden neurons	9
Learning rate	0.1
Momentum factor	0.00001
Activation function	Tangent sigmoid transfer function

Table 6 Standard deviation and mean value of the coefficients of determination (R^2), corresponding to the different folds for the training data for the hybrid PSO–SVM-based models (with linear, superlinear, quadratic, cubic, sigmoid and RBF kernels), multilayer perceptron (MLP) approach and RF-based model for the higher heating value (HHV) in a biomass torrefaction process

Model	Mean(R^2)	Std(R^2)
Linear-SVM	0.8706	0.0174
Superlinear-SVM	0.8671	0.0357
Quadratic-SVM	0.7792	0.0398
Cubic-SVM	0.9615	0.0280
Sigmoid-SVM	0.7889	0.0709
RBF-SVM	0.9195	0.0276
Multilayer perceptron	0.9019	0.0331
Random forest	0.5261	0.0949

Table 4 Optimal hyperparameters of the six fitted SVM-based models with linear, superlinear, quadratic, cubic, sigmoid and RBF kernels found with the PSO technique for the higher heating value (HHV) of torrefied biomass

Kernel	Values of optimal hyperparameters
Linear	Regularization factor $C = 1.1212 \times 10^1$, $\epsilon = 1.9243 \times 10^{-6}$
Superlinear	Regularization factor $C = 9.9931 \times 10^0$, $\epsilon = 2.4479 \times 10^{-7}$, $\sigma = 9.0719 \times 10^{-1}$, $a = 5.3015 \times 10^{-3}$, $b = 1.72$
Quadratic	Regularization factor $C = 7.7783 \times 10^3$, $\epsilon = 3.0514 \times 10^{-2}$, $\sigma = 5.4671 \times 10^{-4}$, $a = 6.0286 \times 10^{-1}$, $b = 2$
Cubic	Regularization factor $C = 1.6098 \times 10^2$, $\epsilon = 3.5890 \times 10^{-2}$, $\sigma = 1.2470 \times 10^{-1}$, $a = 4.5399 \times 10^{-3}$, $b = 3$
Sigmoid	Regularization factor $C = 3.0068 \times 10^2$, $\epsilon = 1.2703 \times 10^{-2}$, $\sigma = 1.0728 \times 10^{-2}$, $a = 1.0857 \times 10^{-2}$
RBF	Regularization factor $C = 1.9315 \times 10^2$, $\epsilon = 5.7783 \times 10^{-2}$, $\sigma = 1.0001 \times 10^{-1}$

References

- de Alegría M, Mancisidor I, de Basurto D, Uruga P, de Alegría M, Mancisidor I, de Arbulo R, López P (2009) European Union's renewable energy sources and energy efficiency policy review: the Spanish perspective. *Renew Sustain Energy Rev* 13(1):100–114
- Abbasi T, Abbasi SA (2010) Biomass energy and the environmental impacts associated with its production and utilization. *Renew Sustain Energy Rev* 14(3):919–937
- Kraxner F, Nordström E-M, Havlík P, Gusti M, Mosnier A, Frank S, Valina H, Fritza S, Fussa S, Kindermann G, McCalluma I, Khabarova N, Böttchera H, Seea L, Aokia K, Schmide E, Máthég L, Obersteiner M (2013) Global bioenergy scenarios: future forest development, land-use implications, and trade-offs. *Biomass Bioenergy* 57:86–96
- Shankar Tumuluru J, Sokhansanj S, Hess JR, Wright CT, Boardman RD (2011) REVIEW: a review on biomass torrefaction process and product properties for energy applications. *Ind Biotechnol* 7(5):384–401
- van der Stelt MJC, Gerhauser H, Kiel JHA, Ptasinski KJ (2011) Biomass upgrading by torrefaction for the production of biofuels: a review. *Biomass Bioenergy* 35(9):3748–3762
- Bach Q-V, Skreiberg Ø (2016) Upgrading biomass fuels via wet torrefaction: a review and comparison with dry torrefaction. *Renew Sustain Energy Rev* 54:665–677
- Prins MJ, Ptasinski KJ, Janssen FJJG (2006) Torrefaction of wood: part 1—weight loss kinetics. *J Anal Appl Pyrol* 77(1):28–34
- Chew JJ, Doshi V (2011) Recent advances in biomass pretreatment: torrefaction fundamentals and technology. *Renew Sustain Energy Rev* 15(8):4212–4222
- Bates RB, Ghoniem AF (2012) Biomass torrefaction: modeling of volatile and solid product evolution kinetics. *Biores Technol* 124:460–469
- Basu P (2013) Biomass gasification, pyrolysis and torrefaction: practical design and theory. Academic Press, New York
- Nhuchhen DR, Basu P, Acharya B (2014) A comprehensive review on biomass torrefaction. *Int J Renew Energy Biofuels* 2014:1–56
- Chen WH, Peng J, Bi XT (2015) A state-of-the-art review of biomass torrefaction, densification and applications. *Renew Sustain Energy Rev* 44:847–866
- Matali S, Rahman NA, Idris SS, Yaacob N, Alias AB (2016) Lignocellulosic biomass solid fuel properties enhancement via torrefaction. *Procedia Eng* 148:671–678
- Motghare KA, Rathod AP, Wasewar KL, Labhsetwar NK (2016) Comparative study of different waste biomass for energy application. *Waste Manag* 47:40–45
- Liu X, Wang W, Gao X, Zhou Y, Shen R (2012) Effect of thermal pretreatment on the physical and chemical properties of municipal biomass waste. *Waste Manag* 32(2):249–255
- Vapnik V (1998) Statistical learning theory. Wiley, New York
- Haykin S (1999) Neural networks: a comprehensive foundation. Pearson Education Inc., Singapore
- Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, New York
- Schölkopf B, Smola AJ, Williamson R, Bartlett P (2000) New support vector algorithms. *Neural Comput* 12(5):1207–1245
- Hastie T, Tibshirani R, Friedman J (2003) The elements of statistical learning. Springer, New York
- Hansen T, Wang CJ (2005) Support vector based battery state of charge estimator. *J Power Sources* 141:351–358
- Li X, Lord D, Zhang Y, Xie Y (2008) Predicting motor vehicle crashes using support vector machine models. *Accid Anal Prev* 40:1611–1618
- Steinwart I, Christmann A (2008) Support vector machines. Springer, New York
- Kulkarni S, Harman G (2011) An elementary introduction to statistical learning theory. Wiley, New York
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the fourth IEEE international conference on neural networks, vol 4. IEEE Publisher, Perth, pp 1942–1948
- Eberhart RC, Shi Y, Kennedy J (2001) Swarm intelligence. Morgan Kaufmann, San Francisco
- Clerc M (2006) Particle swarm optimization. Wiley-ISTE, London
- Olsson AE (2011) Particle swarm optimization: theory, techniques and applications. Nova Science Publishers, New York
- Dorigo M, Stützle T (2004) Ant colony optimization. Bradford Publisher, Cambridge
- Panigrahi BK, Shi Y, Lim M-H (2011) Handbook of swarm intelligence: concepts, principles and applications. Springer, Berlin
- Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J Glob Optim* 39(3):459–471
- Karaboga D, Akay B (2009) A survey: algorithms simulating bee swarm intelligence. *Artif Intell Rev* 31(1):68–85
- Karaboga D, Gorkemli B (2014) A quick artificial bee colony (qABC) algorithm and its performance on optimization problems. *Appl Soft Comput* 23:227–238
- Fister I, Stranad D, Yang X-S, Fister I Jr (2015) Adaptation and hybridization in nature-inspired algorithms. In: Fister I, Fister I Jr (eds) Adaptation and hybridization in computational intelligence, vol 18. Springer, New York, pp 3–50
- Shrestla NK, Shukla S (2015) Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Agric For Meteorol* 200:172–184
- Chen J-L, Li G-S, Wu S-J (2013) Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. *Energy Convers Manag* 75:311–318
- Zeng J, Qiao W (2013) Short-term solar power prediction using a support vector machine. *Renew Energy* 52:118–127
- Ortiz-García EG, Salcedo-Sanz S, Pérez-Bellido AM, Portilla-Figueras JA, Prieto L (2010) Prediction of hourly O₃ concentrations using support vector regression algorithms. *Atmos Environ* 44(35):4481–4488
- Pal M, Goel A (2007) Estimation of discharge and end depth in trapezoidal channel by support vector machines. *Water Resour Manag* 21(10):1763–1780
- Nikoo MR, Mahjouri N (2013) Water quality zoning using probabilistic support vector machines and self-organizing maps. *Water Resour Manag* 27(7):2577–2594
- Fine TL (1999) Feedforward neural networks methodology. Springer, New York
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Mitchell TM (1997) Machine learning. McGraw-Hill Company Inc, New York
- Nocquet T, Dupont C, Commandre J, Grateau M, Thiery S, Salvador S (2014) Volatile species release during torrefaction of wood and its macromolecular constituents: part 1—experimental study. *Energy* 72:180–187
- Nocquet T, Dupont C, Commandre J, Grateau M, Thiery S, Salvador S (2014) Volatile species release during torrefaction of biomass and its macromolecular constituents: part 2—modeling study. *Energy* 72:188–194

46. Bychkov AL, Denkin AI, Tikhova VD, Lomovsky OI (2017) Prediction of higher heating values of plant biomass from ultimate analysis data. *J Therm Anal Calorim* 130(3):1399–1405
47. Galhano dos Santos R, Bordado JC, Mateus MM (2018) Estimation of HHV of lignocellulosic biomass towards hierarchical cluster analysis by Euclidean's distance method. *Fuel* 221:72–77
48. Peduzzi E, Boissonnet G, Maréchal F (2016) Biomass modelling: estimating thermodynamic properties from the elemental composition. *Fuel* 181:207–217
49. Ghugare SB, Tiwary S, Elangovan V, Tambe SS (2014) Prediction of higher heating value of solid biomass fuels using artificial intelligence formalisms. *Bioenergy Res* 7(2):681–692
50. Estiati I, Freire FB, Freire JT, Aguado R, Olazar M (2016) Fitting performance of artificial neural networks and empirical correlations to estimate higher heating values of biomass. *Fuel* 180:377–383
51. Ozveren U (2017) An artificial intelligence approach to predict gross heating value of lignocellulosic fuels. *J Energy Inst* 90(3):397–407
52. Erol M, Haykiri-Acma H, Küçükbayrak S (2010) Calorific value estimation of biomass from their proximate analyses data. *Renew Energy* 35(1):170–173
53. Vargas-Moreno JM, Callejón-Ferre AJ, Pérez-Alonso J, Velázquez-Martí B (2012) A review of the mathematical models for predicting the heating value of biomass materials. *Renew Sustain Energy Rev* 16(5):3065–3083
54. Demirbas A (2004) Linear equations on thermal degradation products of wood chips in alkaline glycerol. *Energy Convers Manag* 45:983–994
55. Energy Research Centre of the Netherlands (ECN) (2018) Research database for biomass and waste. <https://www.ecn.nl/phyllis2/>. Accessed 5 July 2018
56. Chen Q, Zhou J, Liu B, Mei Q, Luo Z (2011) Influence of torrefaction pretreatment on biomass gasification technology. *Chin Sci Bull* 56(14):1449–1456
57. Phanphanich M, Mani S (2011) Impact of torrefaction on the grindability and fuel characteristics of forest biomass. *Biores Technol* 102(2):1246–1253
58. Rousset P, Aguiar C, Labbé N, Commandré JM (2011) Enhancing the combustible properties of bamboo by torrefaction. *Biores Technol* 102(17):8225–8231
59. Lu KM, Lee WJ, Chen WH, Liu SH, Lin TC (2012) Torrefaction and low temperature carbonization of oil palm fiber and eucalyptus in nitrogen and air atmospheres. *Biores Technol* 123:98–105
60. Peng JH, Bi HT, Lim CJ, Sokhansanj S (2013) Study on density, hardness, and moisture uptake of torrefied wood pellets. *Energy Fuels* 27(2):967–974
61. Callejón-Ferre AJ, Velázquez-Martí B, López-Martínez JA, Manzano-Agügliaro F (2011) Greenhouse crop residues: energy potential and models for the prediction of their higher heating value. *Renew Sustain Energy Rev* 15:948–955
62. Saidur R, Abdelaziz EA, Demirbas A, Hossain MS, Mekhilef S (2011) A review on biomass as a fuel for boilers. *Renew Sustain Energy Rev* 15(5):2262–2289
63. Yin C-Y (2011) Prediction of higher heating values of biomass from proximate and ultimate analyses. *Fuel* 90:1128–1132
64. Ziani R, Felkaoui A, Zegadi R (2017) Bearing fault diagnosis using multiclass support vector machines with binary particle swarm optimization and regularized Fisher's criterion. *J Intell Manuf* 28:405–417
65. De Leone R, Pietrini M, Giovannelli A (2015) Photovoltaic energy production forecast using support vector regression. *Neural Comput Appl* 26:1955–1962
66. de Cos Juez FJ, García Nieto PJ, Martínez Torres J, Taboada Castro J (2010) Analysis of lead times of metallic components in the aerospace industry through a supported vector machine model. *Math Comput Model* 52:1177–1184
67. Shawe-Taylor J, Cristianini N (2004) Kernel methods for pattern analysis. Cambridge University Press, New York
68. Simon D (2013) Evolutionary optimization algorithms. Wiley, New York
69. Yang X-S, Cui Z, Xiao R, Gandomi AH, Karamanoglu M (2013) Swarm intelligence and bio-inspired computation: theory and applications. Elsevier, London
70. Clerc M (2012) Standard particle swarm optimisation: from 2006 to 2011. Technical report. http://clerc.maurice.free.fr/ps/SPSO_descriptions.pdf. Accessed 23 Sept 2012
71. Breiman L, Friedman J, Olshen RA, Stone CJ (1984) Classification and regression trees. The Wadsworth statistics/probability series. Wadsworth, Belmont
72. Quinlan JR (1993) C4.5 programs for machine learning. Morgan Kaufmann, San Mateo
73. Rodriguez-Galiano V, Mendes MP, Garcia-Soldado MJ, Chica-Olmo M, Ribeiro L (2014) Predictive modeling of groundwater nitrate pollution using random forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (southern Spain). *Sci Total Environ* 476–477:189–206
74. Wang L, Zhou X, Zhu X, Dong Z, Guo W (2016) Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J* 4:212–219
75. Genuer R, Poggi J-M, Tuleau-Malot C, Villa-Vialaneix N (2017) Random forests for big data. *Big Data Res* 9:28–46
76. Wasserman L (2003) All of statistics: a concise course in statistical inference. Springer, New York
77. Freedman D, Pisani R, Purves R (2007) Statistics. W.W. Norton & Company, New York
78. Picard R, Cook D (1984) Cross-validation of regression models. *J Am Stat Assoc* 79(387):575–583
79. Efron B, Tibshirani R (1997) Improvements on cross-validation: the.632 + bootstrap method. *J Am Stat Assoc* 92(438):548–560
80. Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst and Technol* 2:1–27
81. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl* 11(1):10–18
82. Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques. Morgan Kaufmann, Amsterdam
83. Dahlquist E (2013) Biomass as energy source: resources, systems and applications. CRC Press, Boca Raton
84. Wang S, Luo Z (2016) Pyrolysis of biomass. De Gruyter Ltd, Warsaw