A hybrid intelligent model to predict the hydrogen concentration in the producer gas from a downdraft gasifier

Roque Aguado[a,*], José-Luis Casteleiro-Roca[b], David Vera[a], José Luis Calvo-Rolle[b]

[a]*Department of Electrical Engineering, University of Jaén, E.P.S. Linares, Avda. de la Universidad s/n, 23700, Linares, Jaén, Spain*
[b]*CTC, Department of Industrial Engineering, CITIC, University of A Coruña, E.U.P. Ferrol, Avda. 19 de Febrero s/n, 15405, Ferrol, A Coruña, Spain*

Abstract

This research work presents an artificial intelligence approach to predicting the hydrogen concentration in the producer gas from biomass gasification. An experimental gasification plant consisting of an air-blown downdraft fixed-bed gasifier fueled with exhausted olive pomace pellets and a producer gas conditioning unit was used to collect the whole dataset. During an extensive experimental campaign, the producer gas volumetric composition was measured and recorded with a portable syngas analyzer at a constant time step of 10 seconds. The resulting dataset comprises nearly 75 hours of plant operation in total. A hybrid intelligent model was developed with the aim of performing fault detection in measuring the hydrogen concentration in the producer gas and still provide reliable values in the event of malfunction. The best performing hybrid model comprises six local internal submodels that combine artificial neural networks and support vector machines for regression. The results are remarkably satisfactory, with a mean absolute prediction error of only 0.134% by volume. Accordingly, the developed model could be used as a virtual sensor to support or even avoid the need for a real sensor that is specific for measuring the hydrogen concentration in the producer gas.

Keywords: Biomass gasification, Green hydrogen, Artificial intelligence (AI), Machine learning (ML), Hybrid modeling, Virtual sensor

1. Introduction

The growing energy demands and environmental awareness of the adverse effects of climate change have triggered an increasing interest in distributed cogeneration technologies from renewable energy sources such as bioenergy. Distributed cogeneration can be defined as the simultaneous generation of heat and electricity in a decentralized manner, namely, geographically distributed

*Corresponding author
Email addresses: ramolina@ujaen.es (Roque Aguado), jose.luis.casteleiro@udc.es (José-Luis Casteleiro-Roca), dvera@ujaen.es (David Vera), jose.rolle@udc.es (José Luis Calvo-Rolle)

over the area that is serviced and in close vicinity to the end consumer. Nowadays, gasification offers the best available solution in addressing Combined Heat and Power (CHP) generation from biomass resources on a distributed scale at small and medium power ranges [1–4]. In addition, biomass gasification constitutes a major potential renewable source of hydrogen [5–10], an emerging alternative form of energy storage that is expected to become essential toward a sustainable energy future [11]. Gasification is a thermochemical conversion process, whereby a carbonaceous solid feedstock such as biomass is partially oxidized and transformed into a gaseous fuel [12]. A gasifying agent is required in order to transform the carbonaceous feedstock into gaseous fuel through a series of chemical reactions releasing heat (exothermic) and requiring heat (endothermic). For biomass gasification on a distributed scale, air is preferred over pure oxygen or steam as gasifying agent. Although gasification with pure oxygen or steam avoids nitrogen dilution, leading to a gaseous product with higher calorific value known as synthesis gas or syngas, air separation units are not feasible for small-scale applications, because of the excessively high capital and operation costs offsetting any potential advantage [13]. In the particular case of gasification with air as gasifying agent, the gaseous fuel product is called producer gas and is mainly composed of carbon monoxide (CO), hydrogen (H_2), carbon dioxide (CO_2), nitrogen (N_2), methane (CH_4) and other lightweight hydrocarbons (C_nH_m), which mainly include acetylene (C_2H_2) and ethylene (C_2H_4), among others. In air-blown gasification, the feedstock supplied to the gasifier is partially oxidized in an autothermal process. This partial oxidation of the feedstock produces enough heat to sustain the high temperature atmosphere required for the endothermic reduction reactions responsible for the producer gas formation to occur [9]. The producer gas from gasification, once cooled and cleaned, can be used as fuel in internal combustion engines, micro gas turbines or fuel cells for electric and/or thermal power generation [14].

Motivated by the growing environmental concerns, biomass gasification is attracting increasing interest to produce hydrogen-rich gases for electricity generation in fuel cells or as raw material in the production of synthetic chemicals, gaseous fuels such as hydrogen, methane or ammonia,

and liquid fuels such as methanol or gasolines [15–17]. Hydrogen can be produced in large-scale industrial plants or on a smaller scale in local production facilities directly where the biomass resource is available [17]. The main advantages of large-scale hydrogen production are lower costs due to economies of scale and location close to sources of electricity and water, whereas the main disadvantage is transportation of hydrogen to the points of consumption. Nowadays, the overwhelming majority of hydrogen (>90%) is produced by means of steam reforming or partial oxidation of fossil fuels such as natural gas at unbeatable costs if the CO_2 emissions are not captured and sequestered. Alternatively, hydrogen can be entirely produced from renewable energy sources, which constitutes the so-called "green hydrogen". However, nowadays only minor amounts of green hydrogen are generated through emerging methods such as water electrolysis or biomass gasification. In particular, the cost of producing hydrogen from biomass is less, by a factor of two or more, than the cost of electrolytical hydrogen production from water using wind or photovoltaic power sources [18–21]. Moreover, among the available methods for hydrogen production from biomass, gasification provides the lowest production cost compared to other thermochemical, electrochemical or biochemical conversion technologies [21]. This makes biomass gasification the most financially advantageous process for hydrogen production from renewable energy sources. In terms of energy and exergy efficiency, biomass gasification also has a clear advantage over most other green hydrogen production methods [22].

Of all the current designs for gasification reactors, downdraft fixed-bed gasifiers are arguably the most appropriate for CHP or hydrogen production on a distributed scale, as a result of their relative simplicity of construction, low investment cost, reliability at operation, suitability for a large number of biomass feedstocks and ability to generate a producer gas with a reasonably high carbon conversion efficiency [4, 12]. In a downdraft gasifier, the solid feedstock moves downward together with air, undergoing various thermochemical reactions while proceeding through four different zones: drying, pyrolysis, combustion and reduction. Downdraft gasifiers generate a producer gas with moderate energy density (Lower Heating Value (LHV) = 4–7 MJ/Nm^3) and a low

tar content (<3 g/Nm^3) [12]. Accordingly, unlike other types of gasifiers, a simpler and less robust gas conditioning unit with lower energy consumption is required due to the lower concentration of tars in the producer gas [12].

Among the potential renewable feedstocks for downdraft gasifiers, olive residues and, in particular, exhausted olive pomace, stand out in olive oil producing regions as a result of their wide availability and low cost [13, 23–27]. Exhausted olive pomace is the main by-product of crude olive pomace oil extraction from wet olive pomace. Wet olive pomace is a waste stream from the two-phase extraction process of virgin olive oil that is produced at massive rates in olive oil mills [27]. In order to extract the remaining olive oil contained in the wet pomace, an energy-intensive drying process is required [28]. The dried pomace is eventually subjected to a solid-liquid extraction process of the crude pomace oil with an organic solvent such as hexane, leaving exhausted olive pomace as by-product. In Spain, a wide availability of this by-product from the crude olive pomace oil extraction process as feedstock to the gasification plant can be guaranteed at a cost ranging from 10 to 25 €/t [29]. In particular, Andalusia is the most representative region within Spain, accounting for roughly 75% of the national olive oil production [30, 31], and an average yearly production of exhausted olive pomace ranging between 1.20 and 1.45 million metric tons [32].

Biomass gasification is a complex process with multiple chemical species and reactions occurring simultaneously. Over the last few decades, numerous mathematical and computational models have been proposed to better explain and understand biomass gasification processes [33]. Gasification models are typically categorized into four main groups: thermodynamic equilibrium models, kinetic models, computational fluid dynamics models and artificial neural network models [12, 34]. Artificial intelligence models constitute a relatively recent approach in biomass gasification processes for accurately predicting system outputs by learning from large amounts of experimental data [33].

Artificial Intelligence (AI) techniques have proved to outperform traditional models in terms

of prediction performance. While most of the research activity has been devoted to fluidized bed gasifiers [35–37], a very scarce number of works addressing the application of AI techniques to predicting the producer gas composition from downdraft gasifiers is available in the scientific literature [38–42]. Baruah et al. [40] developed a model based on Artificial Neural Networks (ANNs) in order to predict the composition of the producer gas from fixed-bed downdraft gasifiers using experimental data from diverse biomass feedstocks. Other research works used ANN techniques [43], regression techniques [44], in addition to binary and multi-class classifiers [41] to predict the composition (H_2, CO_2, CO, CH_4) and heating value of the producer gas from a fixed-bed downdraft gasifier fed with pinecone particles and wood pellets. Mikulandrić et al. [38, 39] also used ANN modeling approaches to predict process parameters in fixed-bed gasifiers with reasonable speed and accuracy. More recently, Ozbas et al. [42] created various regression models with the aim of predicting the value of the hydrogen concentration in the producer gas and synthesis gas from a steel fixed-bed updraft gasifier fueled with olive pits, based on other parameters, such as time, temperature, heating value and concentrations of carbon monoxide, carbon dioxide, methane and oxygen. Different algorithms were used, such as Linear Regression (LR), *k*-Nearest Neighbors (*k*-NN) Regression, Support Vector Machines for Regression (SVMR) and Decision Tree Regression (DTR) algorithms, with variable performance results.

In this research work, a hybrid intelligent model was developed from experimental data in order to predict the hydrogen concentration in the producer gas from a downdraft gasifier fueled with exhausted olive pomace pellets. Unlike the usual AI algorithms, a hybrid intelligent approach based on a combination of these techniques has demonstrated to improve the prediction performance. For example, hybrid intelligent models have already been used to predict the energy demand with the aim of optimizing generation strategies and use of resources [45]. Another hybrid model was developed to predict the output temperature of a solar thermal collector [46]. This hybrid model technique allows a combination of different heterogeneous algorithms, which complement each other to achieve the best prediction, while adapting to the characteristics of the

dataset [47]. Hybrid models have also been used to accurately predict the change in the flow rate of hydrogen supplied to a Proton-Exchange Membrane Fuel Cell (PEMFC) in order to reach the desired working point [48, 49].

The main aim of the present work is to demonstrate that hybrid intelligent models are valid to reliably determine the operational parameters of a downdraft gasifier. The remainder of this article is structured as follows. Section 2 introduces the case study, describes the gasification plant and instrumentation, as well as the experimental procedure. The model approach and the procedure used to develop the hybrid intelligent model are described in Section 3, in addition to the algorithms involved, the data processing procedure and the performance metrics. Subsequently, the configuration and results of the best performing model are discussed in Section 4. Conclusions and future works are finally exposed in Section 5.

2. Case study

The present work describes a novel intelligent model that is able to predict the hydrogen concentration in the producer gas from an experimental gasification plant. The experimental methodology is divided into two subsections. Initially, the experimental gasification plant is presented and described thoroughly, including the biomass feedstock to the gasifier and the instrumentation for measuring the producer gas composition and heating value. Thereafter, the experimental procedure for plant operation and data collection is conveniently reported.

2.1. Plant description and instrumentation

As shown in Fig. 1, the experimental gasification plant consists of a downdraft fixed-bed gasifier with an open top and a producer gas conditioning unit. The downdraft gasifier (1) was manufactured of stainless steel with an internal refractory ceramic lining. As the reactor has an open top, the gasification process occurs at near ambient pressure. The producer gas is discharged from the gasifier at temperatures around 400 °C from an outlet duct for exiting gas [26]. The

hot producer gas cannot be readily used in electric power generation units or downstream gas upgrading units, as it contains traces of tar, fly ash, moisture and other inorganic impurities that can cause severe corrosion problems to all mechanical equipment. Consequently, the producer gas from gasification must be cooled down and cleaned up to particular limits set by the manufacturer of the downstream equipment. To this end, the gasification plant incorporates a producer gas conditioning unit with the following components: a cyclone (2) that separates most of the particles entrained in the producer gas; a Venturi scrubber (3) that sharply reduces the temperature of the producer gas and removes the tars formed in the gasification process by spraying a pressurized water jet; coarse filters (4), fine filters (5) and a safety filter (6) that ensure a thorough removal of most organic and inorganic contaminants from the producer gas, so that it can be eventually used as fuel for CHP generation units or downstream upgrading processes [26]. Beforehand, the flare valve (10) is fully opened in order to safely burn the producer gas in a test flare stack (11) located next to the gas control valve (7) and the gas sampling point (9). A blower (8) is used to provide the necessary suction to induce airflow through the gasifier and drive the resulting producer gas through the gas conditioning unit. Finally, a discharge valve (12) controls the flow rate of producer gas for downstream applications.

The physicochemical properties of the feedstock influence the performance of the gasification process. In this work, exhausted olive pomace pellets were used as feedstock to the gasification plant due to their high availability and low cost. The exhausted olive pomace pellets used as feedstock in this work were provided by a local biomass trading company in the Spanish municipality of Linares, Andalusia. The physicochemical properties of the exhausted olive pomace pellets are diverse, as a result of their highly heterogeneous nature. For illustrative purposes, the proximate and ultimate analyses of this biomass feedstock are provided in Table 1. A picture of a sample is shown in Fig. 2, where one can appreciate the cylindrical shape of the exhausted olive pomace pellets, with lengths ranging from 10 mm to 35 mm and diameters of about 10 mm. Exhausted olive pomace must be pelletized prior to be used as feedstock in downdraft gasifiers, because fine

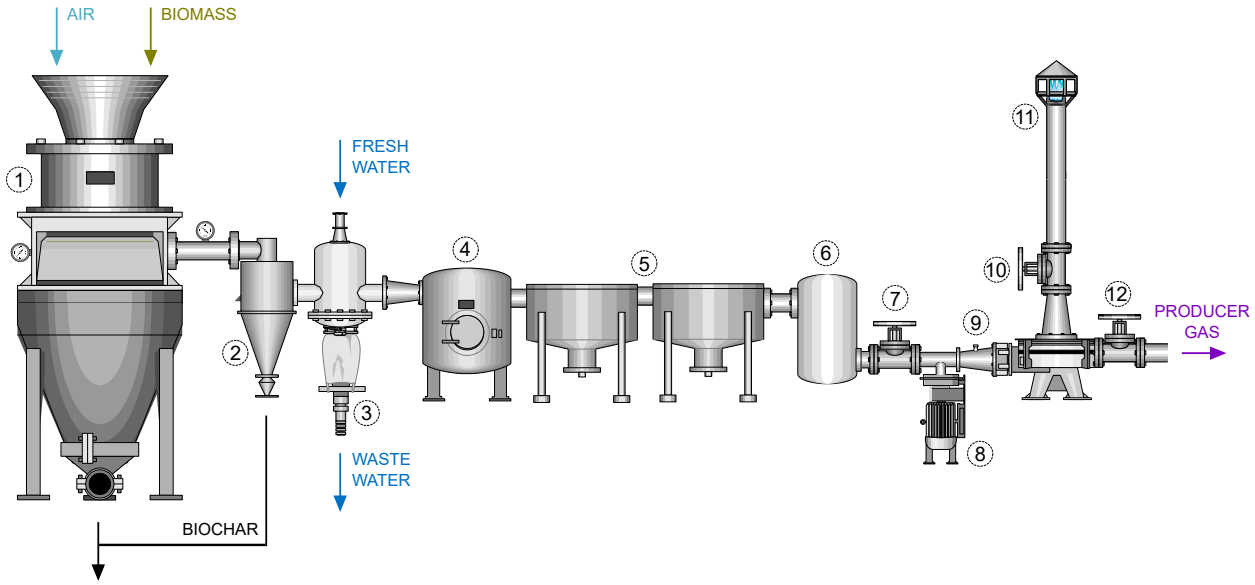

Figure 1: Layout of the experimental gasification plant. (**1**) Downdraft gasifier; (**2**) cyclone; (**3**) Venturi scrubber; (**4**) coarse particle filter; (**5**) fine particle filters; (**6**) bag filter; (**7**) control valve; (**8**) blower; (**9**) producer gas sampling point; (**10**) flare valve; (**11**) flare stack; (**12**) discharge valve.

particles in downdraft fixed-bed gasifiers (<3 mm) are conducive to the formation of ash clinkers and also disturb the optimal flow of fuel and air along the gasifier bed, leading to excessive pressure drops, abnormally low temperatures in the reduction zone and unstable operation [50].

Table 1: Proximate and ultimate analyses of the exhausted olive pomace pellets.

Proximate analysis, wt. %	Minimum	Average	Maximum
Moisture, as received	4.80	9.88	13.80
Ash, dry basis	5.40	7.20	9.20
Volatile matter, dry basis	70.24	73.53	76.67
Fixed carbon, dry basis	16.43	19.27	22.36
Ultimate analysis, wt. % dry basis	**Minimum**	**Average**	**Maximum**
Carbon	47.49	51.02	55.44
Hydrogen	5.26	5.88	6.50
Nitrogen	0.40	0.95	2.20
Sulfur	0.07	0.14	0.24
Chlorine	0.12	0.26	0.40
Oxygen	31.04	34.55	38.86

In order to monitor and record the producer gas composition and heating value, a portable syngas analyzer (Gasboard 3100P, Hubei Cubic-Ruiyi Instrument Co., Ltd., China) was used [51].

Figure 2: Exhausted olive pomace pellets used as feedstock to the downdraft gasifier.

The gasification plant is equipped with an in-line gas sampling point located at the blower outlet, where the portable syngas analyzer is connected after passing through a portable gas conditioning unit consisting of two water bubblers, an activated carbon filter and a precision filter. In order to prevent damages to the sensors of the syngas analyzer, the producer gas must be free from particles (< 1 μm) and tar traces, and its moisture content (relative humidity) must range between 5 and 95% non-condensing. These relatively strict requirements highlight the importance of the upstream gas conditioning unit.

The portable syngas analyzer was used for simultaneous measurement of the volume concentrations of up to 6 different gases in the producer gas, namely CO, CO_2, H_2, O_2, CH_4 and C_nH_m, while the N_2 balance and heating values were calculated automatically. The volume concentration of N_2 is estimated on the assumption that no other gaseous components are present in the producer gas. In practice, however, fractional amounts of other inert gases, mostly argon, are also present. The lower heating value (LHV) and higher heating value (HHV) are continuously determined from the volume concentrations (mole fractions) of the fuel gases in the producer gas (CO, H_2, CH_4 and C_nH_m), as given by Eqs. (1) and (2), respectively [51].

$$\mathrm{LHV\,(MJ/Nm^3)} = 12.620\,\mathrm{CO} + 10.777\,\mathrm{H_2} + 35.818\,\mathrm{CH_4} + 91.180\,\mathrm{C_nH_m} \quad (1)$$

$$\mathrm{HHV\ (MJ/Nm^3)} = 12.620\,\mathrm{CO} + 12.788\,\mathrm{H_2} + 39.840\,\mathrm{CH_4} + 99.220\,\mathrm{C}_n\mathrm{H}_m \quad (2)$$

The operating principle of the portable syngas analyzer is based on three types of technologies to determine the concentrations of the different gaseous constituents in the producer gas [17, 51]:

- Nondispersive Infrared (NDIR) sensor, to measure the concentrations of CO, CO_2, CH_4 and C_nH_m (optional).
- Thermal Conductivity Detector (TCD), to measure the concentration of H_2.
- Electron Capture Detector (ECD), to measure the concentration of electron-absorbing components in the producer gas, namely, O_2.

Table 2 outlines the technologies used by the portable syngas analyzer for measuring each component of the producer gas. The calibration process is specific for each gas component and only required in the event that the gas measurement error exceeds technical specifications [52].

Table 2: Measurement specifications for each component of the producer gas [51].

Component	Method	Range	Resolution	Precision
CO	NDIR	0–30%	0.01%	≤ 2% Full scale
CO_2	NDIR	0–25%	0.01%	≤ 2% Full scale
CH_4	NDIR	0–10%	0.01%	≤ 2% Full scale
C_nH_m	NDIR	0–5%	0.01%	≤ 2% Full scale
H_2	TCD	0–30%	0.01%	≤ 3% Full scale
O_2	ECD	0–25%	0.01%	≤ 3% Full scale

As the portable syngas analyzer systematically calculates the N_2 concentration by difference [17], determination of the producer gas composition is based on the three aforementioned technologies (NDIR, TCD and ECD). In particular, the portable syngas analyzer relies heavily on the NDIR spectroscopic sensor for determining the concentrations of CO, CO_2, CH_4 and C_nH_m in the producer gas with long term stability, high accuracy and low power consumption [53]. NDIR spectrometers are capable of measuring the light intensity absorbed by the carbon-containing constituents of the producer gas. These instruments consist of three basic components: an infrared

light source, an optical tube containing the producer gas and an infrared detector with a specific wavelength filter [53].

The working principle of a typical NDIR spectroscopic sensor is based on the Beer–Lambert law [53]. It is well established that most polyatomic molecular gases have a particular absorption wavelength in the infrared region. The molecules of CO, CO_2, CH_4 and C_nH_m in the producer gas consist of heterogeneous atoms and thus, have absorption spectra in the infrared wavelength region. When a beam of light with the typical absorption wavelength of a given gas is passed through the gas whose concentration is to be measured, its intensity decreases. The gas concentration determines the degree of intensity attenuation. The relationship between both parameters is governed by the Beer–Lambert law [54].

$$I = I_0 e^{-\varepsilon c l} \tag{3}$$

The absorption intensity (i) of each polyatomic molecular gas can be expressed as:

$$i = I_0 - I = I_0 \left(1 - e^{-\varepsilon c l}\right) \tag{4}$$

where:

- I_0: initial light intensity.
- I: light intensity after passing through the gaseous medium.
- ε: absorption coefficient.
- c: gas concentration.
- l: thickness of the gaseous absorption medium.

The concentrations of H_2, O_2 and N_2 cannot be determined with the NDIR sensor, as these molecules do not exhibit absorption spectra in the infrared wavelength region. Thus, specific technologies such as TCD and ECD sensors are used for measuring the volume concentrations of H_2 and O_2, respectively. Meanwhile, the concentration of N_2 is automatically estimated by

difference, assuming that no other gaseous components are present in the producer gas, apart from those measured.

2.2. Experimental procedure

The operating procedure of the experimental gasification plant for the purpose of the present work is briefly described below.

1. Prior to operation, the gasification plant is carefully checked in order to ensure that the pipeline and all the vessels are well sealed against air leakages, as the producer gas conditioning unit operates marginally below ambient pressure [52].
2. The gasifier is loaded with exhausted pomace pellets through the hopper, while the coarse and fine filtering media are filled with dry char and graded sawdust, respectively.
3. The air blower is switched on in order to induce suction of air into the gasifier and generate a continuous flow of producer gas from the gasifier toward the flare stack.
4. Initially, the gasifier requires an external heat input [26]. For this reason, the feedstock is ignited across the entire cross section with a hand torch or blow torch. As a result, the producer gas starts to form and evacuate through the flare stack. A hopper vibration mechanism driven by an electrical motor is activated at regular periods of time in order to ensure a continuous downward movement of the feedstock in the gasifier, and thus avoiding bridging and channeling phenomena through the gasifier bed.
5. Immediately after ignition, the portable syngas analyzer starts recording the producer gas composition and calculated heating value with a selected time step.
6. Approximately 30–60 minutes after the gasifier ignition, the gasification plant reaches steady state conditions and the producer gas composition and heating value tend to stabilize.
7. The producer gas is typically flared with a hand torch for a preliminary quality check. The flare must either be colorless or light blue if ignited at night, as long as the producer gas is sufficiently clean and tar-free. Since the producer gas composition and heating value are accurately measured and monitored with the in-line syngas analyzer, the main purpose of

the flare ignition is to fully oxidize air pollutants such as carbon monoxide and methane to form safer and environmentally friendlier complete combustion products such as carbon dioxide and water vapor.

The operating procedure described above was systematically repeated and data for different runs were collected in an extensive experimental campaign. The portable syngas analyzer recorded the volumetric composition and estimated heating value of the producer gas with a constant time step of 10 seconds. The resulting dataset comprises almost 75 hours of plant operation in total, under different operational conditions.

3. Model approach

A schematic data-flow diagram for the approach followed in this research work to develop the hybrid intelligent model is shown in Fig. 3. This model aims at effectively predicting the concentration of a specific component (H_2) in the producer gas from the gasification plant. The inputs to the model are the volume concentrations of the remaining major gaseous constituents in the producer gas, but only those measured with the NDIR and ECD sensors, namely CO, CO_2, CH_4, C_nH_m and O_2. As the N_2 concentration is determined by difference, it was not included as input to the model.

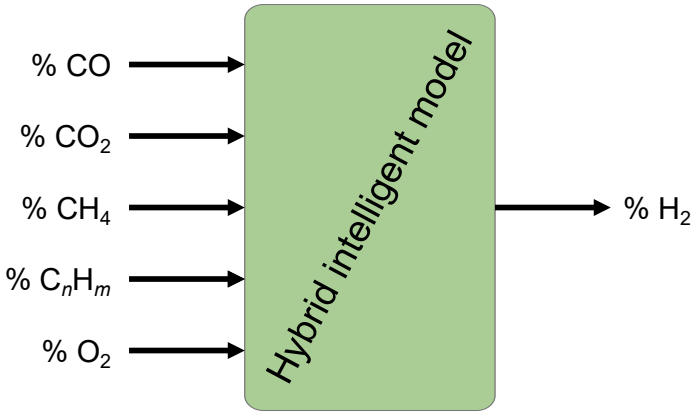

Figure 3: Simplified data-flow diagram showing the inputs and outputs of the hybrid intelligent model.

The aim of this research work is not only to trustworthily predict the hydrogen concentration

in the producer gas from a downdraft gasifier, but also to validate the procedure to create a hybrid intelligent model capable of predicting other variables in the gasification plant. The hydrogen concentration was selected as target variable mainly because it requires using a specific TCD sensor, and thus, it is possible to use the developed model for fault detection and still provide reliable results in the event of malfunction or failures in measurement. Moreover, if the number of variables to be measured increases, the model becomes more complex and produces more than one output (one per predicted variable).

A hybrid intelligent approach was selected, as hybrid models usually provide a better performance than that of a global model. As illustrated in Fig. 4, hybrid intelligent models comprise several internal submodels that are only used with a portion of the whole dataset. These local models are trained with only a cluster of the dataset. The whole dataset is divided with a clustering algorithm and then, regression models are trained for each cluster (local models).

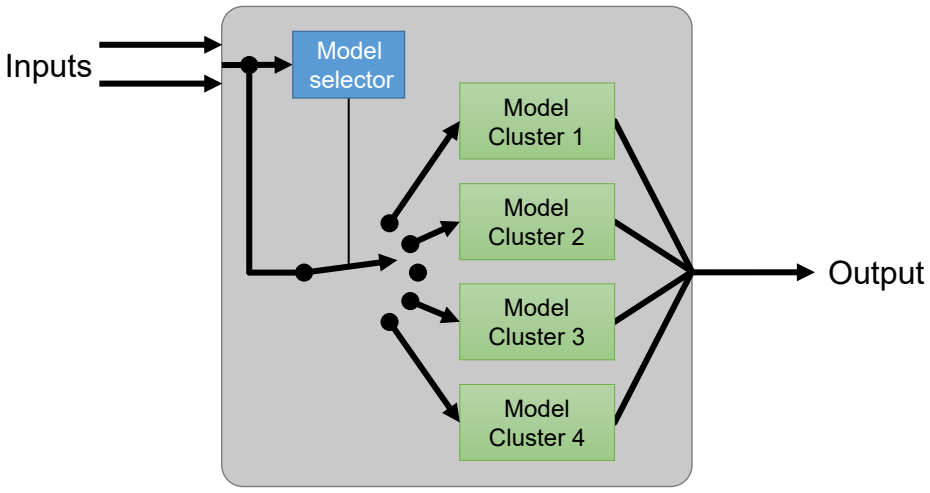

Figure 4: Internal configuration of the hybrid intelligent model.

The hybrid intelligent model represents a virtual sensor that can be used in different situations, such as when a variable lacking a real sensor is to be determined or to perform fault detection in a real sensor. In this work, the hybrid model was developed for the purpose of performing fault detection in a specific sensor of the portable syngas analyzer, namely, the TCD sensor that

measures the hydrogen concentration in the producer gas. Thereby, if the real sensor produces unreliable or none values, the virtual sensor developed in this research work could ensure provision of valid measurements all the time.

3.1. Procedure to create the hybrid intelligent model

The procedure followed in this research work to develop the hybrid model is schematically represented in Fig. 5. It consists of four consecutive steps, which are outlined below.

1. The first step is the clustering phase, in which the dataset is divided several times to create the different clusters that compose the hybrid model. As the clustering algorithm used in this work does not detect the optimal number of clusters, several sets of clusters were created.
2. After the clustering phase, several regression algorithms were trained in each cluster. Moreover, when the regression algorithm had different internal configurations, the different configurations were considered as different local models.
3. In the third step, the best regression algorithm for each cluster was identified. In this phase, k-Fold Cross Validation was used to calculate the performance of each regression algorithm. The local models were trained again with the best algorithm for each cluster using, at this time, all the data available for each cluster.
4. Finally, the hybrid configuration must be chosen. As mentioned above, since several sets of clusters were created, this phase checks the different hybrid configurations with different datasets, which were initially isolated from the rest of the training procedure. It is particularly important that this dataset should contain data from all the clusters to allow testing of all the possible configurations.

3.2. Artificial Intelligence algorithms

In this section, all the different algorithms used in this research work are described.

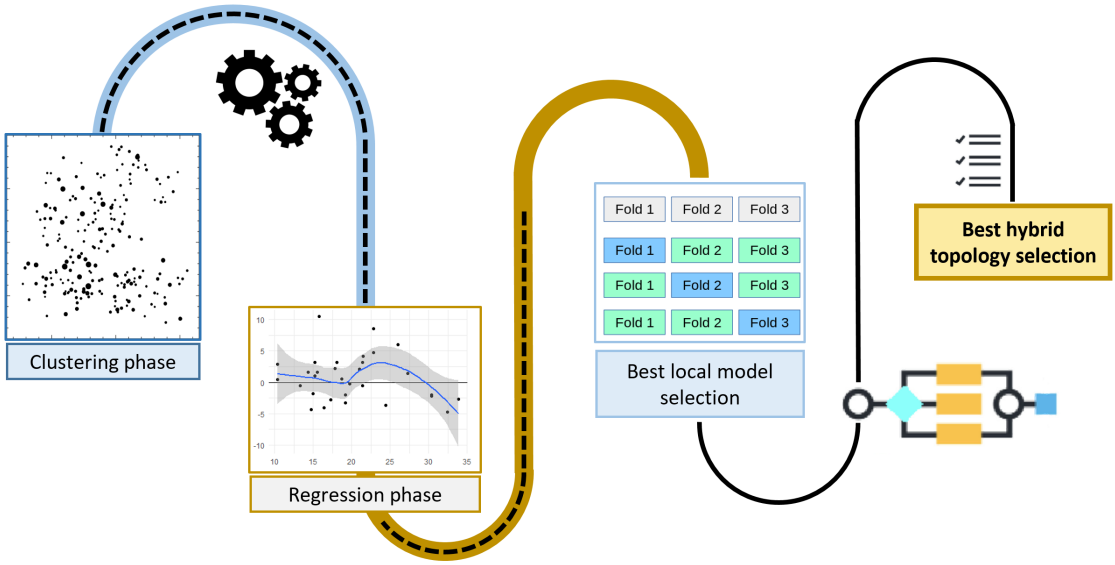

Figure 5: Flowchart to create the hybrid intelligent model.

3.2.1. k-Means

The clustering algorithm that was used to divide the dataset is the *k*-Means algorithm. This is one of the most well-known and widely used unsupervised algorithms for creating clusters. The dataset and the number of clusters (k) must be specified as inputs, while the outputs are the centroids and the clusters of data [55–57].

Each centroid is calculated as the center of each cluster. Centroids are the unique criterion necessary to assign new data to a cluster. This algorithm typically uses the euclidean distance to assign the samples to each cluster, and when the clusters are created, new samples are assigned to the cluster defined by the closest centroid [58].

The procedure to train the algorithm can be summarized in the next steps [59, 60]:

- First, the centroids are randomly chosen from the dataset; k samples are defined as centroids.
    1. All the samples are assigned to the cluster defined by its closest centroid.
    2. New centroids are calculated as the centers of each new cluster.
- The training procedure involves repeating the two previous steps until the centroids remain virtually unchanged between two consecutive iterations.

3.2.2. *Artificial Neural Networks*

A commonly used artificial intelligence algorithm to perform regression is the Artificial Neural Network (ANN). This algorithm is created with the association of several artificial neurons, that are the simplest units. Each neuron has several inputs and an activation function, that is chosen depending on the configuration and the use of the algorithm (regression or classification). ANNs provide good results when the inputs to the model are not in the range of the training set [61, 62].

Neurons are organized into three types of layers: the input layer, the hidden layer and the output layer. All neurons in the same layer have the same inputs and outputs. Each neuron adds all the weighted inputs, including a *bias*, and then, the activation function is calculated as a result of this addition. The number of neurons in the input and output layers, respectively, depends on the inputs and outputs of the model. However, the number of hidden layers and the number of neurons per hidden layer relies on the particular configuration used for each model [63–65].

In this research work, the configuration used for the ANN model was a Multi-Layer Perceptron (MLP), which is a feedforward neural network with only one hidden layer. The activation function of the internal neurons was set as the tan-sigmoid, while the output layer neuron was defined by means of a linear function. Since the model has a single output, namely, the hydrogen concentration; the output layer has only one neuron. Several models with different neurons in the hidden layer were trained.

3.2.3. *Support Vector Machines for Regression*

Support Vector Machines for Regression (SVMR) is an artificial intelligence algorithm mainly used in classification that forms an optimal hyperplane to separate data belonging to different classes. In order to use this algorithm for regression, a transformation of the data is performed to represent them in a new high-dimensional space. Then, a linear regression is trained in this new multi-dimensional space.

In this work, a modification of the original algorithm was used, which includes an auto adjust

function to calculate the weight vector (γ) and the kernel width (σ). This modification is known as Least Square-Support Vector Regression (LS-SVR) [66, 67].

3.2.4. Polynomial Regression

The last algorithm used in this research work is the Polynomial Regression. This algorithm uses simple functions to calculate the output of the model. These basic functions depend on the order of each particular algorithm and the number of inputs [68–71]. As examples, Eqs. (5) and (6) show the functions of a Polynomial Regression model with two inputs, x_1 and x_2, for the first and second order, respectively.

$$F(x) = b_0 + b_1 x_1 + b_2 x_2 \tag{5}$$

$$F(x) = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2 + b_4 x_1^2 + b_5 x_2^2 \tag{6}$$

3.3. Data processing

The dataset used in this research work was collected with a constant sampling time step of 10 seconds between consecutive samples and contains a total of 26839 samples with 8 variables each. The data include the volumetric composition (H_2, CO, CO_2, CH_4, C_nH_m, O_2, N_2) and the LHV of the producer gas. Table 3 presents the quartile distribution and average values of the dataset used to the develop the hybrid intelligent model, in addition to the N_2 concentration and the LHV. The median LHV of the producer gas was 4.62 MJ/Nm^3, while the maximum LHV reached 5.52 MJ/Nm^3. As evidenced by the minimum and maximum values of the volume percentages reported in Table 3, a number of samples from the start-up phase of the gasification plant are included, where pure air and lean producer gas compositions were recorded.

Particularly worthy of note is that hydrogen was, for most of the time, the main fuel gas in the producer gas volumetric composition, with an average H_2/CO molar ratio just over 1.35. Only during the start-up phase of the gasification plant was the volume concentration of carbon monoxide markedly higher than that of hydrogen. Fig. 6 shows a histogram of the H_2/CO molar ratio, with most of the samples (~85%) having a H_2/CO ratio between 0.8 and 1.8. In fact, the

Table 3: Quartile distribution and average values of the dataset.

Parameter	Unit	Minimum	First quartile	Median	Third quartile	Maximum	Average
H_2	vol.%	0.00	15.52	16.69	17.73	20.81	15.3030
CO	vol.%	0.00	10.35	11.89	14.33	28.28	12.3126
CO_2	vol.%	0.00	11.08	13.29	14.32	16.89	11.9922
CH_4	vol.%	0.00	2.09	2.72	3.38	4.47	2.5960
C_nH_m	vol.%	0.00	0.13	0.16	0.21	0.36	0.1675
O_2	vol.%	0.15	0.70	0.87	1.04	20.68	2.0800
N_2[a]	vol.%	47.96	52.19	53.55	54.83	81.82	55.5475
LHV[b]	MJ/Nm^3	0.00	4.46	4.62	4.71	5.52	4.2806

[a] N_2 concentration is determined by difference.

[b] LHV is estimated through a linear correlation, as indicated in Eq. (1).

producer gas or synthesis gas from biomass gasification is typically characterized by an average H_2/CO molar ratio in the range of 0.8–1.5 [72]. As observed in Fig. 6, the hydrogen concentration in the producer gas from gasification of exhausted olive pomace pellets was on average relatively high. Hydrogen could be separated from the producer gas in a downstream purification process for use as a carbon-free energy carrier or used as feedstock for renewable synthetic fuel production. However, for most chemical syntheses of carbon-based hydrogen-carrying fuels (e.g., the methanol synthesis, the Fischer-Tropsch synthesis or the methanation reaction for production of synthetic natural gas), a H_2/CO molar ratio of at least 2 is required [73]. Thus, for downstream use of the producer gas as synthesis gas, in addition to separating inerts such as nitrogen and argon, this ratio should be adjusted in a separate catalytic reactor before the synthesis reactor, where some carbon monoxide would be converted to hydrogen by means of the forward water-gas shift reaction. As the downstream processing substantially increases the investment and operation costs, the cogenerative production of electricity and heat currently appears to be the most suitable and readily available application for the producer gas.

A subset of the whole dataset containing 5% of the data was isolated at the beginning of the modeling process to perform a validation test with the final hybrid model. Accordingly, the validation dataset includes 1342 randomly chosen samples. The rest of the dataset was normalized

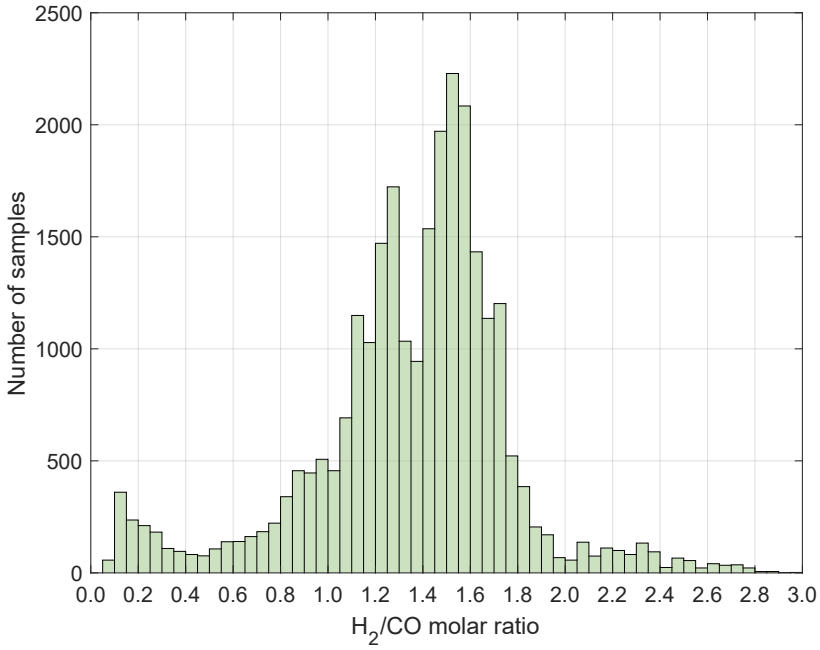

Figure 6: Histogram of the H_2/CO molar ratio in the producer gas.

in order to create the hybrid intelligent model. Moreover, once the dataset was divided in clusters, another test dataset with 5% of the data from each cluster was created for the purpose of determining the best hybrid configuration.

In Fig. 7, the Pearson linear correlation coefficients between all pairs of normalized input variables are conveniently displayed in a matrix of plots. Histograms of each variable appear along the matrix diagonal, while scatter plots of variable pairs appear in the off diagonal. The slopes of the least-squares reference lines in the scatter plots are equal to the correlation coefficients displayed in the upper right corner. It is noteworthy that the concentration of carbon monoxide (CO) is inversely correlated with the concentrations of most other gaseous components in the producer gas. A particularly high positive correlation between the concentrations of carbon dioxide (CO_2) and hydrocarbons (CH_4 and C_nH_m) is observed. This behavior can be predominantly attributed to the forward water-gas shift reaction in combination with the methanation reaction, both of which proceed in the gas phase and are moderately exothermic. The forward water-gas shift reaction yields CO_2 and H_2 from CO and H_2O, while the methanation reaction is the reverse process of the

steam methane reforming reaction.

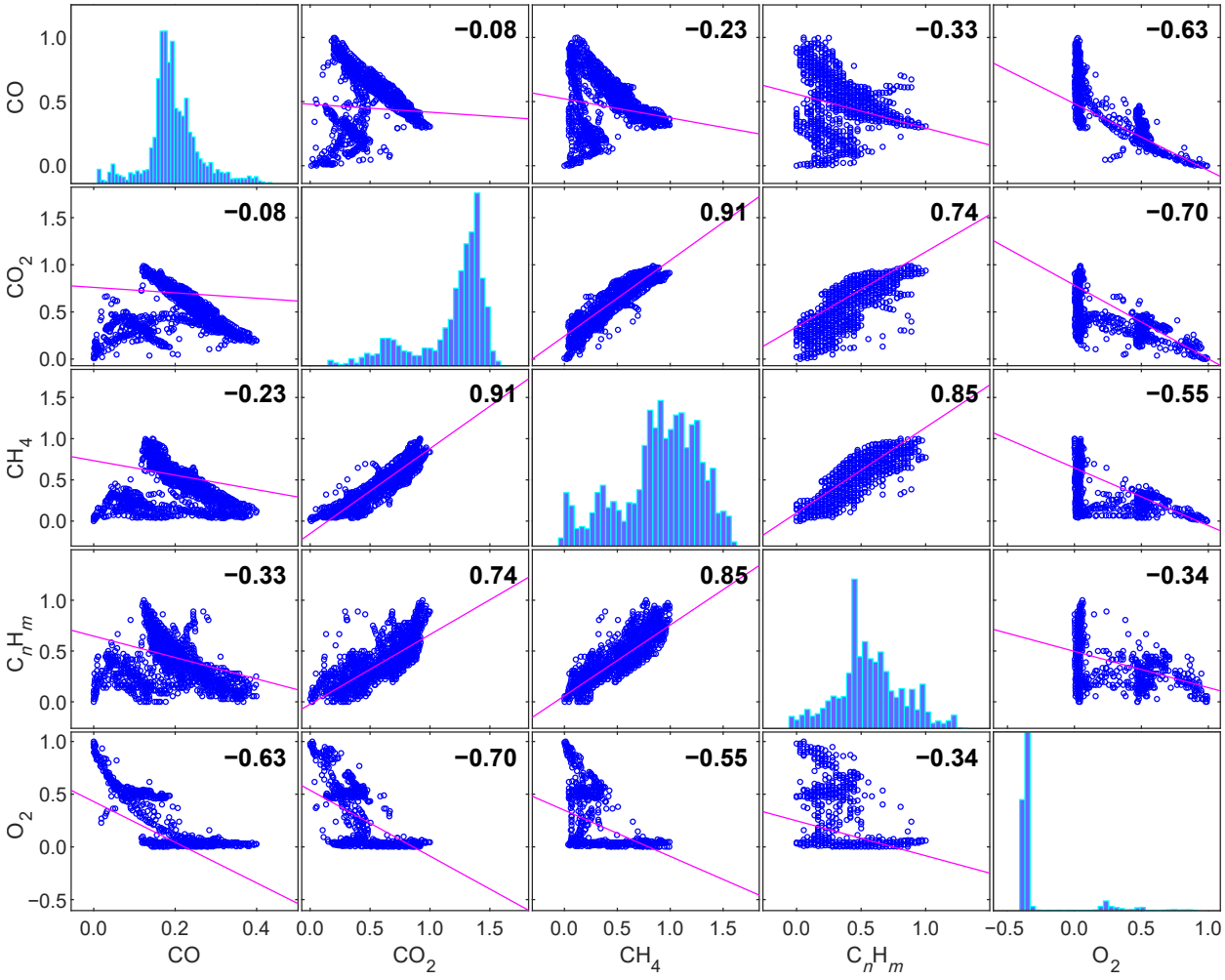

Figure 7: Pearson correlation matrix between all pairs of input variables to the hybrid intelligent model.

3.4. Performance metrics

In order to evaluate the performance of each regression model created in the clusters, *k*-Fold Cross Validation with 10 folds was used in each local model. The cross validation procedure splits the training data 10 times (10 folds), and then uses 9 of these groups to train and the last one to test the model. This procedure was repeated 10 times until all the groups created for the purpose of cross validations were used in the tests. At this point, all the samples in a cluster are used to test the model, and it is possible to evaluate the performance of each local model.

In this work, four error metrics were used; namely, the Mean Squared Error (MSE), the Mean Absolute Error (MAE), the Mean Percentage Absolute Error (MAPE) and the Normalized Mean Squared Error (NMSE). Eqs. (7), (8), (9) and (10) indicate how each error is calculated. A comprehensive review on prediction errors is provided in [74].

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{7}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\mid \hat{y}_i - y_i \mid \tag{8}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\mid \hat{y}_i - y_i \mid}{\mid y_i \mid} \tag{9}$$

$$\mathrm{NMSE} = \frac{\mathrm{MSE}}{\sigma^2} \tag{10}$$

The above equations require the following statistical parameters:

- n: total number of samples in each cluster.
- $\hat{y}_i$: predicted value.
- y_i: measured value.
- σ: variance of the error distribution.

4. Results

The results from this research work are divided into four different subsections. The first part addresses the clustering phase results. Subsequently, the regression phase results and the hybrid configuration selection are discussed. Finally, the hybrid model is tested with a specific dataset. The initial dataset, as discussed above, was divided in order to create three different datasets:

1. Initially, the validation dataset was isolated from the beginning of the modeling process.

2. The second dataset was used to choose the hybrid configuration. In order to create this dataset and to ensure that samples from all the clusters were included, this test dataset was created with 5% of each cluster data.
3. The last dataset was used to create the models in the clusters. The regression models were developed with this dataset.

4.1. Clustering phase results

As mentioned above, the *k*-Means algorithm was used to divide the dataset in several clusters. This algorithm has issues with local minima during training depending on the initial centroids. To avoid these drawbacks, the *k*-Means algorithm was used 20 times to create the same number of clusters; each time with random initial centroids.

A total of 9 hybrid configurations were created with growing numbers of clusters, ranging from 2 to 10. Accordingly, 10 different configurations were evaluated: the global model and 9 hybrid models. Table 4 indicates the number of samples included in each cluster, where each column represents a different configuration. The global configuration with one cluster is at the left, while the hybrid configurations are abbreviated as Hyb., with a number of clusters varying from 2 to 10.

Table 4: Number of samples for each cluster.

	Global	Hyb. 2	Hyb. 3	Hyb. 4	Hyb. 5	Hyb. 6	Hyb. 7	Hyb. 8	Hyb. 9	Hyb. 10
C-1	24222	4618	2681	2217	1802	341	341	345	344	344
C-2		19604	7808	2667	2666	1794	1458	1180	1165	1153
C-3			13733	9261	4514	2337	2340	1180	1176	1178
C-4				10076	6431	4515	2926	1433	1181	1179
C-5					8808	6429	4278	2928	1417	1378
C-6						8808	5468	4231	2571	1847
C-7							7413	5509	4581	2584
C-8								7417	5596	3887
C-9									6189	4944
C-10										5729

4.2. *Regression phase results*

The results from the modeling phase are reported in Tables 5 and 6. These tables show the best achieved values for the MSE and MAE in all the local models. These error values were obtained using k-Fold Cross Validation with $k = 10$. This cross validation technique divides the clusters data in k groups, and uses $k - 1$ groups to train a model, the performance of which is eventually tested against the last group. The validation procedure was repeated k times until all the groups were used as test data. At the end of this process, the error was calculated using all the data available in each cluster. This validation procedure ensures more realistic error values than those of a hold-out validation.

Table 5: MSE values for each best local model ($\times 10^{-3}$).

	Global	Hyb. 2	Hyb. 3	Hyb. 4	Hyb. 5	Hyb. 6	Hyb. 7	Hyb. 8	Hyb. 9	Hyb. 10
C-1	0.3263	0.4715	0.3940	0.6547	0.4963	0.1119	0.0953	0.1135	0.1024	0.0684
C-2		0.2109	0.5591	0.4646	0.3349	0.5123	0.5096	0.1765	0.4519	0.4883
C-3			0.1439	0.1155	0.1327	0.4444	0.4357	0.4320	0.3657	0.4372
C-4				0.4664	0.5132	0.1177	0.7369	0.4878	0.2105	0.2310
C-5					0.0804	0.4043	0.1319	0.7734	0.7786	0.7952
C-6						0.0662	0.1738	0.1182	0.0462	0.0373
C-7							0.0505	0.1893	0.1007	0.5504
C-8								0.0521	0.2733	0.0178
C-9									0.0634	0.2363
C-10										0.0451

Table 6: MAE values for each best local model.

	Global	Hyb. 2	Hyb. 3	Hyb. 4	Hyb. 5	Hyb. 6	Hyb. 7	Hyb. 8	Hyb. 9	Hyb. 10
C-1	0.0109	0.0129	0.0086	0.0170	0.0139	0.0059	0.0052	0.0061	0.0049	0.0041
C-2		0.0059	0.0156	0.0098	0.0097	0.0142	0.0142	0.0068	0.0135	0.0146
C-3			0.0040	0.0041	0.0073	0.0105	0.0101	0.0116	0.0094	0.0102
C-4				0.0113	0.0129	0.0070	0.0192	0.0140	0.0066	0.0074
C-5					0.0036	0.0110	0.0074	0.0194	0.0190	0.0182
C-6						0.0039	0.0078	0.0070	0.0036	0.0031
C-7							0.0032	0.0081	0.0057	0.0167
C-8								0.0034	0.0091	0.0027
C-9									0.0035	0.0086
C-10										0.0031

The training procedure using *k*-Fold Cross Validation was used for all the regression algorithms described previously. The three algorithms were tested with the next configurations:

- The MLPs were configured with only one hidden layer and with a tan-sigmoid activation function in all the neurons, except in the output layer, which uses a linear activation function. These regression models were adjusted using the Levenberg-Marquardt algorithm and were tested with different number of neurons in the hidden layer (from 1 to 15). These models were designated as "ANN*" in results tables, where "*" represents the number of neurons in the hidden layer, and each of these models were tested as individual models (as different algorithms).
- The LS-SVR models were trained using an auto-tune function to adjust the internal parameters. With this function, only one model is created for each cluster data, designated in tables as "LS-SVR".
- The last regression algorithm is the Polynomial Regression algorithm, which was tested in the first and second order of the polynomial model. This algorithm would have been referred to in the tables as "Poly*", where "*" indicates the order of the polynomial model, but it does not appear therein, because only the best performing algorithms are shown.

Table 7 shows the optimal regression algorithms for each local model. It is noteworthy that the tables included in this work are a summary of all the results. In order to select the best performing algorithms, 18 models were tested for each cluster: 15 artificial neural networks, 1 support vector regression and 2 polynomial regressions.

Once the best regression algorithm was selected for each local model, it was necessary to train the models once again. As *k*-Fold Cross Validation was used, 10 models were trained with the best performing algorithm; but none of them with all the cluster data in the training phase. Then, new models were trained with all the cluster data using the best performing regression algorithm, which was identified as the one with the lowest MSE.

Table 7: Best regression algorithm for each local model.

	Global	Hyb. 2	Hyb. 3	Hyb. 4	Hyb. 5	Hyb. 6	Hyb. 7	Hyb. 8	Hyb. 9	Hyb. 10
C-1	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR	ANN7	LS-SVR	ANN7	LS-SVR	LS-SVR
C-2		LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR
C-3			LS-SVR	LS-SVR	ANN15	LS-SVR	LS-SVR	LS-SVR	LS-SVR	LS-SVR
C-4				LS-SVR	LS-SVR	ANN15	ANN5	LS-SVR	LS-SVR	LS-SVR
C-5					LS-SVR	LS-SVR	ANN14	ANN15	ANN13	ANN15
C-6						LS-SVR	LS-SVR	ANN15	LS-SVR	LS-SVR
C-7							LS-SVR	LS-SVR	LS-SVR	ANN14
C-8								LS-SVR	LS-SVR	LS-SVR
C-9									LS-SVR	LS-SVR
C-10										LS-SVR

4.3. Best hybrid configuration selection

In order to establish the final internal hybrid configuration, the second dataset described at the beginning of the results section was used. These data were used as inputs for all the possible whole models (the global model and the nine hybrid models), and the performance parameters were calculated for all the configurations. Table 8 outlines the MSE for all the hybrid models, where it is possible to appreciate that the best hybrid configuration has 6 local internal submodels.

Table 8: MSE values for each hybrid configuration ($\times 10^{-4}$).

Global	Hyb. 2	Hyb. 3	Hyb. 4	Hyb. 5	Hyb. 6	Hyb. 7	Hyb. 8	Hyb. 9	Hyb. 10
3.0496	1.4504	1.2210	1.1608	1.3909	**0.9481**	1.6131	1.6305	1.1987	1.5477

4.4. Test results

Table 9 summarizes the results of the final test, which, as reported above, was performed on a dataset containing 1342 samples that was isolated at the beginning of the model creation procedure. Table 9 includes the error metrics with real values and with normalized values, both of which are reasonably good. The MSE for prediction of the hydrogen concentration in the producer gas is approximately 0.086, which corresponds to an estimation error of about 0.134% by volume, as the MAE units are the actual units of the variables.

Table 9: Performance metrics for the best hybrid configuration.

	MSE	MAE	NMSE
Hybrid configuration with 6 local models	0.0863	0.1343	0.043
(with normalized values)	$2.0027 \cdot 10^{-4}$	0.0065	

As an example, Fig. 8 presents a scatter plot of several independent log measurements of the hydrogen concentration that were used in this test. The upper subplot shows the measured values (purple circles) and the predicted values (green crosses) using the hybrid intelligent model. The lower subplot indicates the absolute error (blue circles) of each prediction, which was consistently below 2% by volume.

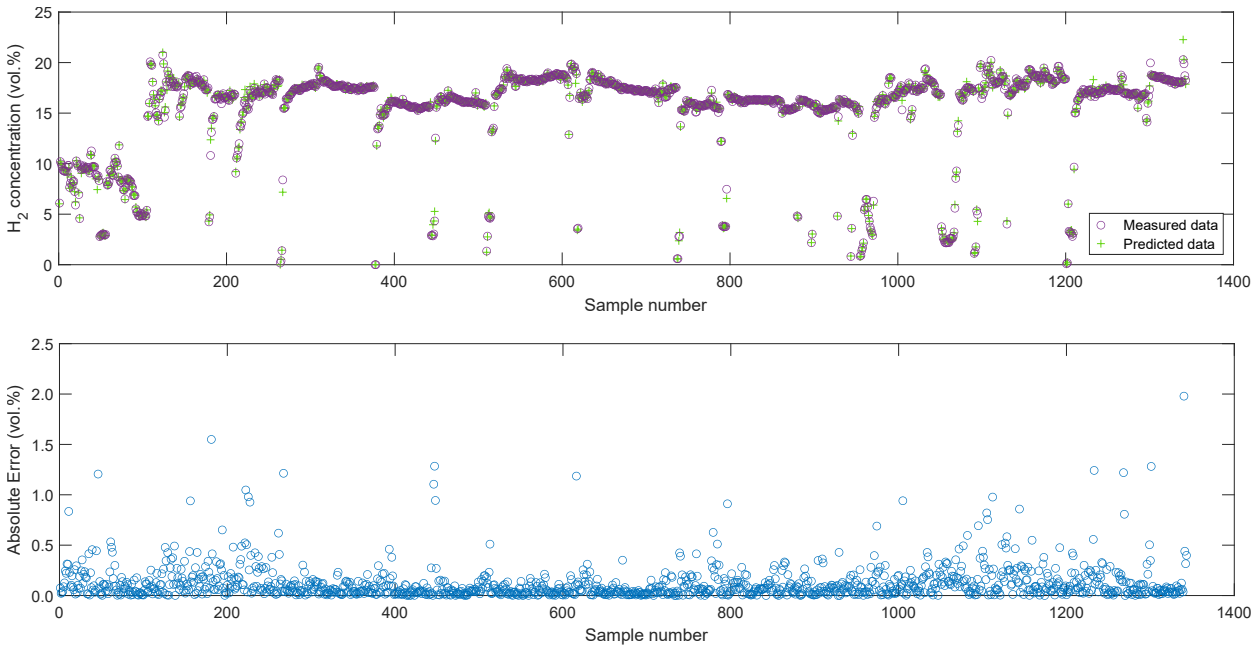

Figure 8: Scatter plots of the hydrogen concentration prediction test. Upper subplot: measured data (purple circles) and predicted data (green crosses). Lower subplot: absolute error (blue circles).

In a related work, Ozbas et al. [42] developed four detailed regression models based on supervised algorithms (LR, *k*-NN, SVMR and DTR) to predict the hydrogen concentration in the synthesis gas from pyrolysis/gasification of olive pits in a lab-scale updraft gasifier. Operational parameters such as time, temperature, heating value and concentrations of carbon monoxide, carbon dioxide, methane and oxygen were considered as input data. In all four different models, the

MAE values were below 1% by volume, which is an indication that the estimates of hydrogen concentration were very close to the actual values. In particular, the best performing algorithm was LR, with a MAE equal to 0.007%, in contrast to SVMR, which exhibited a MAE of 0.762%. In another work, George et al. [36] used a feed-forward, multi-layered artificial neural network (ANN) to predict the producer gas composition from air gasification of different biomass feedstocks (coffee husk, coconut shell, groundnut shell, sugarcane bagasse and sawdust) in a fluidized bed gasifier. The developed ANN model, which was trained by means of the Levenberg-Marquardt back-propagation algorithm, consisted of seven input variables, four output variables and one hidden layer with fifteen neurons. The performance metrics of the ANN model were within the satisfactory limit, with a MSE of 0.71%. Notwithstanding the foregoing, it should be noted that the performance metrics of these works are not strictly comparable to those of the present work, since the input variables and data to each model are different, in addition to the gasifier type and the physicochemical properties of the feedstock.

5. Conclusions and future works

This research work presents a hybrid intelligent model to predict the hydrogen concentration in the producer gas from an experimental gasification plant consisting of a downdraft gasifier fueled with exhausted olive pomace pellets and a gas conditioning unit. Very satisfactory results were obtained, with average prediction errors of only 0.134%, as it can be observed above in the hydrogen concentration prediction test. Accordingly, the developed model can be used for fault detection in measuring the hydrogen concentration in the producer gas with a real TCD sensor and still provide reliable values in the event of malfunction, or also as a virtual sensor to support or even avoid the need for such sensor in the experimental installation.

The developed hybrid intelligent model comprises six local internal submodels that combine Artificial Neural Networks and Support Vector Machines for regression. It is noteworthy that, although all the local models may have a similar configuration, they are all different, as they were

trained with different datasets.

As future works, it is worth mentioning the possibility of developing new intelligent models to predict the remaining gaseous components of the producer gas. In addition, more complex models could be developed by taking into account other inputs, such as temperature, pressure or flow rates. Thereby, the number of sensors installed in the experimental facility could be reduced and fault detection could be performed in any of them. Other future research works could be conducted in an attempt to create detailed artificial intelligence models of all the units that comprise the experimental gasification plant. These last models could be used, among other applications, for predictive maintenance of each unit or for plant performance optimization.

Nomenclature

Abbreviations

AI Artificial Intelligence

ANN Artificial Neural Network

CHP Combined Heat and Power

DTR Decision Tree Regression

ECD Electron Capture Detector

HHV Higher Heating Value

k-NN k-Nearest Neighbors

LHV Lower Heating Value

LR Linear Regression

LS-SVR Least Square-Support Vector Regression

MAE Mean Absolute Error

MAPE Mean Percentage Absolute Error

MLP Multi-Layer Perceptron

MSE Mean Squared Error

NDIR Nondispersive Infrared

NMSE Normalized Mean Squared Error

PEMFC Proton-Exchange Membrane Fuel Cell

SVMR Support Vector Machines for Regression

TCD Thermal Conductivity Detector

Funding

This research work was supported by the project with Ref. No. 1381442, co-funded by *Programa Operativo FEDER 2014-2020* and *Consejería de Economía y Conocimiento de la Junta de Andalucía* of the Spanish Government.

CITIC, as a Research Center of the University System of Galicia, is funded by *Consellería de Educación, Universidade e Formación Profesional* of the *Xunta de Galicia* through the European Regional Development Fund (ERDF) and the *Secretaría Xeral de Universidades* (Ref. ED431G 2019/01).

Roque Aguado gratefully acknowledges financial support from *Ministerio de Universidades* under the FPU Program (Ref. FPU19/00930).

Acknowledgements

The authors deeply appreciate the technical support and assistance provided by IFAPA Centro "Venta del Llano". In addition, the editor and anonymous reviewers are gratefully acknowledged for their valuable comments and suggestions, which contributed to improving the work reported in this paper.

CRediT authorship contribution statement

R. Aguado: Conceptualization, Investigation, Visualization, Writing - Original Draft, Writing - Review & Editing. **J.-L. Casteleiro-Roca:** Formal analysis, Methodology, Software, Visualization, Writing - Original Draft, Writing - Review & Editing. **D. Vera:** Funding acquisition, Investigation, Resources, Supervision. **J.L. Calvo-Rolle:** Funding acquisition, Supervision, Resources, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Dasappa, D. Subbukrishna, K. Suresh, P. Paul, G. Prabhu, Operational experience on a grid connected 100 kWe biomass gasification power plant in Karnataka, India, Energy Sustain. Dev. 15 (3) (2011) 231–239. doi:10.1016/j.esd.2011.03.004.

[2] C. R. Coronado, J. T. Yoshioka, J. L. Silveira, Electricity, hot water and cold water production from biomass. Energetic and economical analysis of the compact system of cogeneration run with woodgas from a small downdraft gasifier, Renew. Energy 36 (6) (2011) 1861–1868. doi:10.1016/j.renene.2010.11.021.

[3] D. Prando, F. Patuzzi, G. Pernigotto, A. Gasparella, M. Baratieri, Biomass gasification systems for residential application: An integrated simulation approach, Appl. Therm. Eng. 71 (1) (2014) 152–160. doi:10.1016/j.applthermaleng.2014.06.043.

[4] A. Susastriawan, H. Saptoadi, Purnomo, Small-scale downdraft gasifiers for biomass gasification: A review, Renew. Sust. Energ. Rev. 76 (2017) 989–1003. doi:10.1016/j.rser.2017.03.112.

[5] S. J. Yoon, Y.-I. Son, Y.-K. Kim, J.-G. Lee, Gasification and power generation characteristics of rice husk and rice husk pellet using a downdraft fixed-bed gasifier, Renew. Energy 42 (2012) 163–167. doi:10.1016/j.renene.2011.08.028.

[6] E. S. Aydin, O. Yucel, H. Sadikoglu, Experimental study on hydrogen-rich syngas production via gasification of pine cone particles and wood pellets in a fixed bed downdraft gasifier, Int. J. Hydrog. Energy 44 (32) (2019) 17389–17396. doi:10.1016/j.ijhydene.2019.02.175.

[7] Q. Li, G. Song, J. Xiao, T. Sun, K. Yang, Exergy analysis of biomass staged-gasification for hydrogen-rich syngas, Int. J. Hydrog. Energy 44 (5) (2019) 2569–2579. doi:10.1016/j.ijhydene.2018.11.227.

[8] M. A. Babatabar, M. Saidi, Hydrogen production via integrated configuration of steam gasification process of biomass and water-gas shift reaction: Process simulation and optimization, Int. J. Energy Res. 45 (13) (2021) 19378–19394. doi:10.1002/er.7087.

[9] M. Aziz, A. Darmawan, F. B. Juangsa, Hydrogen production from biomasses and wastes: A technological review, Int. J. Hydrog. Energy 46 (68) (2021) 33756–33781. doi:10.1016/j.ijhydene.2021.07.189.

[10] M. R. Kabli, A. M. Ali, M. Inayat, A. A. Zahrani, K. Shahzad, M. Shahbaz, S. A. Sulaiman, H_2-rich syngas production from air gasification of date palm waste: an experimental and modeling investigation, Biomass Conv. Bioref. (2022). doi:10.1007/s13399-022-02375-7.

[11] H. Balat, E. Kırtay, Hydrogen from biomass – Present scenario and future prospects, Int. J. Hydrog. Energy 35 (14) (2010) 7416–7426. doi:10.1016/j.ijhydene.2010.04.137.

[12] P. Basu, Biomass Gasification, Pyrolysis and Torrefaction, 3rd Edition, Academic Press, 2018. doi:10.1016/

C2016-0-04056-1.

[13] D. Vera, F. Jurado, K. D. Panopoulos, P. Grammelis, Modelling of biomass gasifier and microturbine for the olive oil industry, Int. J. Energy Res. 36 (3) (2012) 355–367. doi:10.1002/er.1802.

[14] S. A. Archer, R. Steinberger-Wilckens, Systematic analysis of biomass derived fuels for fuel cells, Int. J. Hydrog. Energy 43 (52) (2018) 23178–23192. doi:10.1016/j.ijhydene.2018.10.161.

[15] R. G. dos Santos, A. C. Alencar, Biomass-derived syngas production via gasification process and its catalytic conversion into fuels by fischer tropsch synthesis: A review, Int. J. Hydrog. Energy 45 (36) (2020) 18114–18132. doi:10.1016/j.ijhydene.2019.07.133.

[16] H. Gruber, P. Groß, R. Rauch, A. Reichhold, R. Zweiler, C. Aichernig, S. Müller, N. Ataimisch, H. Hofbauer, Fischer-Tropsch products from biomass-derived syngas and renewable hydrogen, Biomass Conv. Bioref. 11 (2021) 2281–2292. doi:10.1007/s13399-019-00459-5.

[17] S. Vecten, M. Wilkinson, N. Bimbo, R. Dawson, B. M. Herbert, Hydrogen-rich syngas production from biomass in a steam microwave-induced plasma gasification reactor, Bioresour. Technol. 337 (2021) 125324. doi:10.1016/j.biortech.2021.125324.

[18] R. H. Williams, E. D. Larson, R. E. Katofsky, J. Chen, Methanol and hydrogen from biomass for transportation, Energy Sustain. Dev. 1 (5) (1995) 18–34. doi:10.1016/S0973-0826(08)60083-6.

[19] P. Nikolaidis, A. Poullikkas, A comparative overview of hydrogen production processes, Renew. Sust. Energ. Rev. 67 (2017) 597–611. doi:10.1016/j.rser.2016.09.044.

[20] M. Kayfeci, A. Keçebaş, M. Bayat, Chapter 3 - hydrogen production, in: F. Calise, M. D. D'Accadia, M. Santarelli, A. Lanzini, D. Ferrero (Eds.), Solar Hydrogen Production, Academic Press, 2019, pp. 45–83. doi:10.1016/B978-0-12-814853-2.00003-5.

[21] T. Lepage, M. Kammoun, Q. Schmetz, A. Richel, Biomass-to-hydrogen: A review of main routes production, processes evaluation and techno-economical assessment, Biomass and Bioenergy 144 (2021) 105920. doi:10.1016/j.biombioe.2020.105920.

[22] C. Acar, I. Dincer, Comparative assessment of hydrogen production methods from renewable and non-renewable sources, Int. J. Hydrog. Energy 39 (1) (2014) 1–12. doi:10.1016/j.ijhydene.2013.10.060.

[23] V. Skoulou, A. Zabaniotou, G. Stavropoulos, G. Sakelaropoulos, Syngas production from olive tree cuttings and olive kernels in a downdraft fixed-bed gasifier, Int. J. Hydrog. Energy 33 (4) (2008) 1185–1194. doi:10.1016/j.ijhydene.2007.12.051.

[24] M. Dogru, Experimental results of olive pits gasification in a fixed bed downdraft gasifier system, Int. J. Green Energy 10 (4) (2013) 348–361. doi:10.1080/15435075.2012.655351.

[25] D. Vera, F. Jurado, N. K. Margaritis, P. Grammelis, Experimental and economic study of a gasification plant fuelled with olive industry wastes, Energy Sustain. Dev. 23 (2014) 247–257. doi:10.1016/j.esd.2014.09.011.

[26] R. Aguado, D. Vera, D. A. López-García, J. P. Torreglosa, F. Jurado, Techno-economic assessment of a gasification plant for distributed cogeneration in the agrifood sector, Appl. Sci. 11 (2) (2021). doi:10.3390/app11020660.

[27] R. Aguado, D. Vera, F. Jurado, G. Beltrán, An integrated gasification plant for electric power generation from wet biomass: toward a sustainable production in the olive oil industry, Biomass Conv. Bioref. (2022). doi:10.1007/s13399-021-02231-0.

[28] R. Arjona, A. García, P. Ollero, The drying of alpeorujo, a waste product of the olive oil mill industry, J. Food Eng. 41 (3) (1999) 229–234. doi:10.1016/S0260-8774(99)00104-1.

[29] Institute for Diversification and Saving of Energy (IDAE), Informe de Precios de la Biomasa para Usos Térmicos, https://www.idae.es/sites/default/files/estudios_informes_y_estadisticas/informe_precios_biomasa_usos_termicos_2t_2020.pdf (2020).

[30] Economic Affairs & Promotion Unit - International Olive Council (IOC), https://www.internationaloliveoil.org/what-we-do/economic-affairs-promotion-unit/, accessed: 2021-06-18 (2019).

[31] S. I. Patsios, K. N. Kontogiannopoulos, G. F. Banias, Environmental impact assessment in agri-production, in: Bio-Economy and Agri-production, Elsevier, 2021, pp. 83–116. doi:10.1016/b978-0-12-819774-5.00005-9.

[32] Andalusian Energy Agency, La biomasa en Andalucía, Tech. rep., Consejería de Hacienda Industria y Energía (2020).

[33] M. Puig-Arnavat, J. C. Bruno, A. Coronas, Review and analysis of biomass gasification models, Renew. Sust. Energ. Rev. 14 (9) (2010) 2841–2851. doi:10.1016/j.rser.2010.07.030.

[34] R. Bijesh, P. Arun, C. Muraleedharan, Modified stoichiometric equilibrium model for sewage sludge gasification and its validation based on experiments in a downdraft gasifier, Biomass Conv. Bioref. (2021). doi:10.1007/s13399-021-01916-w.

[35] M. Puig-Arnavat, J. A. Hernández, J. C. Bruno, A. Coronas, Artificial neural network models for biomass gasification in fluidized bed gasifiers, Biomass Bioenergy 49 (2013) 279–289. doi:10.1016/j.biombioe.2012.12.012.

[36] J. George, P. Arun, C. Muraleedharan, Assessment of producer gas composition in air gasification of biomass

using artificial neural network model, Int. J. Hydrog. Energy 43 (20) (2018) 9558–9568. doi:10.1016/j.ijhydene.2018.04.007.

[37] S. Sezer, U. Özveren, Investigation of syngas exergy value and hydrogen concentration in syngas from biomass gasification in a bubbling fluidized bed gasifier by using machine learning, Int. J. Hydrog. Energy 46 (39) (2021) 20377–20396. doi:10.1016/j.ijhydene.2021.03.184.

[38] R. Mikulandrić, D. Lončar, D. Böhning, R. Böhme, M. Beckmann, Artificial neural network modelling approach for a biomass gasification process in fixed bed gasifiers, Energy Convers. Manag. 87 (2014) 1210–1223. doi:10.1016/j.enconman.2014.03.036.

[39] R. Mikulandrić, D. Böhning, R. Böhme, L. Helsen, M. Beckmann, D. Lončar, Dynamic modelling of biomass gasification in a co-current fixed bed gasifier, Energy Convers. Manag. 125 (2016) 264–276. doi:10.1016/j.enconman.2016.04.067.

[40] D. Baruah, D. Baruah, M. Hazarika, Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers, Biomass Bioenergy 98 (2017) 264–271. doi:10.1016/j.biombioe.2017.01.029.

[41] A. Y. Mutlu, O. Yucel, An artificial intelligence based approach to predicting syngas composition for downdraft biomass gasification, Energy 165 (2018) 895–901. doi:10.1016/j.energy.2018.09.131.

[42] E. E. Ozbas, D. Aksu, A. Ongen, M. A. Aydin, H. K. Ozcan, Hydrogen production via biomass gasification, and modeling by supervised machine learning algorithms, Int. J. Hydrog. Energy 44 (32) (2019) 17260–17268. doi:10.1016/j.ijhydene.2019.02.108.

[43] O. Yucel, E. S. Aydin, H. Sadikoglu, Comparison of the different artificial neural networks in prediction of biomass gasification products, Int. J. Energy Res. 43 (11) (2019) 5992–6003. doi:10.1002/er.4682.

[44] F. Elmaz, O. Yucel, A. Y. Mutlu, Predictive modeling of biomass gasification with machine learning-based regression methods, Energy 191 (2020) 116541. doi:10.1016/j.energy.2019.116541.

[45] J.-L. Casteleiro-Roca, J. Gómez-González, J. Calvo-Rolle, E. Jove, H. Quintián, B. Gonzalez Diaz, J. Mendez Perez, Short-term energy demand forecast in hotels using hybrid intelligent modeling, Sensors 19 (11) (2019) 2485. doi:10.3390/s19112485.

[46] J.-L. Casteleiro-Roca, P. Chamoso, E. Jove, A. González-Briones, H. Quintián, M.-I. Fernández-Ibáñez, R. A. Vega Vega, A.-J. Piñón Pazos, J. A. López Vázquez, S. Torres-Álvarez, T. Pinto, J. L. Calvo-Rolle, Solar thermal collector output temperature prediction by hybrid intelligent model for smartgrid and smartbuildings applications and optimization, Applied Sciences 10 (13) (2020) 4644. doi:10.3390/app10134644.

[47] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, J. A. Méndez-Pérez, J. L. Calvo-Rolle, Anomaly detection based on intelligent techniques over a bicomponent production plant used on wind generator blades manufacturing,

Revista Iberoamericana de Automática e Informática industrial 17 (1) (2020) 84–93. doi:10.4995/riai.2019.11055.

[48] J.-L. Casteleiro-Roca, A. J. Barragán, F. S. Manzano, J. L. Calvo-Rolle, J. M. Andújar, Fuel cell hybrid model for predicting hydrogen inflow through energy demand, Electronics 8 (11) (2019). doi:10.3390/electronics8111325.

[49] H. Alaiz-Moretón, E. Jove, J.-L. Casteleiro-Roca, H. Quintián, H. López García, J. A. Benítez-Andrades, P. Novais, J. L. Calvo-Rolle, Bioinspired hybrid model to predict the hydrogen inlet fuel cell flow change of an energy storage system, Processes 7 (11) (2019). doi:10.3390/pr7110825.

[50] V. R. Patel, D. S. Upadhyay, R. N. Patel, Gasification of lignite in a fixed bed reactor: Influence of particle size on performance of downdraft gasifier, Energy 78 (2014) 323–332. doi:10.1016/j.energy.2014.10.017.

[51] Hubei Cubic-Ruiyi Instrument Co. Ltd., Portable Infrared Syngas Analyzer Gasboard 3100P User Manual, http://www.gas-analyzers.com/products/syngas-analyzer/Gasboard-3100P.html, accessed: 2021-06-18 (2017).

[52] J. Soares, A. C. Oliveira, Experimental assessment of pine wood chips gasification at steady and part-load performance, Biomass Bioenergy 139 (2020) 105625. doi:10.1016/j.biombioe.2020.105625.

[53] C.-H. Shen, J.-H. Yeah, Long Term Stable Δ-Σ NDIR Technique Based on Temperature Compensation, Appl. Sci. 9 (2) (2019). doi:10.3390/app9020309.

[54] J. Hodgkinson, R. Smith, W. O. Ho, J. R. Saffell, R. P. Tatam, Non-dispersive infra-red (NDIR) measurement of carbon dioxide at 4.2μm in a compact and optically efficient sensor, Sensors and Actuators B: Chemical 186 (2013) 580–588. doi:10.1016/j.snb.2013.06.006.

[55] E. Jove, P. Blanco-Rodríguez, J. L. Casteleiro-Roca, J. Moreno-Arboleda, J. A. López-Vázquez, F. J. de Cos Juez, J. L. Calvo-Rolle, Attempts prediction by missing data imputation in engineering degree, in: International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6–8, 2017, Proceeding, Springer, 2017, pp. 167–176.

[56] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, 1967, pp. 281–297.

[57] J. Moody, C. Darken, Fast learning in networks of locally-tuned processing units, Neural Comput. 1 (2) (1989) 281–294. doi:10.1162/neco.1989.1.2.281.

[58] J. Orallo, M. Quintana, C. Ramírez, Introducción a la minería de datos, Editorial Alhambra S.A., 2004.

[59] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, J.-A. Méndez-Pérez, J. L. Calvo-Rolle, Virtual sensor for fault detection, isolation and data recovery for bicomponent mixing machine monitoring, Informatica 30 (4) (2019)

671–687.

[60] P. Viñuela, I. León, Redes de neuronas artificiales: un enfoque práctico, Pearson Educación - Prentice Hall, 2004.

[61] M. Galipienso, M. Quevedo, O. Pardo, F. Ruiz, M. Ortega, Inteligencia artificial. Modelos, técnicas y áreas de aplicación, Editorial Paraninfo, 2003.

[62] J. González, V. Hernando, Redes neuronales artificiales: fundamentos, modelos y aplicaciones, RA-MA, 2000.

[63] A. M. C. Harston, R. Pap, Handbook of Neural Computing Applications, Elsevier Science, 2014.

[64] R. López, J. Fernández, Las Redes Neuronales Artificiales, Netbiblo, 2008.

[65] B. del Brío, A. Molina, Redes neuronales y sistemas borrosos, Ra-Ma, 2006.

[66] I. Steinwart, A. Christmann, Support vector machines, Springer Publishing Company, Incorporated, 2008.

[67] L. Wang, J. Wu, Neural network ensemble model using PPR and LS-SVR for stock et eorecasting, in: Advanced Intelligent Computing, 2012, pp. 1–8. doi:10.1007/978-3-642-24728-6_1.

[68] C. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics), Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[69] R. Heiberger, E. Neuwirth, Polynomial regression, in: R Through Excel, Use R, Springer New York, 2009, pp. 269–284. doi:10.1007/978-1-4419-0052-4_11.

[70] X. Wu, Optimal designs for segmented polynomial regression models and web-based implementation of optimal design software, State University of New York at Stony Brook, Stony Brook, NY, USA, 2007.

[71] Z. Zhang, S. Chan, On kernel selection of multivariate local polynomial modelling and its application to image smoothing and reconstruction, J. Signal Process. Syst. 64 (3) (2011) 361–374. doi:10.1007/s11265-010-0495-4.

[72] C. Pirola, A. D. Fronzo, F. Gallia, C. L. Bianchia, A. Comazzia, F. Manenti, Biosyngas conversion by Fischer – Tropsch Synthesis: experimental results and multi-scale simulation of a PBR with high Fe loaded supported catalysts, Chem. Eng. Trans. (2014). doi:10.3303/cet1437100.

[73] R. Rauch, J. Hrbek, H. Hofbauer, Biomass gasification for synthesis gas production and applications of the syngas, WIREs Energy Environ. 3 (4) (2014) 343–362. doi:10.1002/wene.97.

[74] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, A survey of forecast error measures, World Appl. Sci. J. 24 (2013) 171–176. doi:10.5829/idosi.wasj.2013.24.itmies.80032.