Applied Soft Computing 107 (2021) 107497

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Multivariate influence through neural networks ensemble: Study of Saharan dust intrusion in the Canary Islands

D. Gonzalez-Calvo [a,*], R.M. Aguilar [b], C. Criado-Hernandez [c], L.A. Gonzalez-Mendoza [d]

[a] University of La Laguna, 38200 La Laguna (Tenerife), Spain
[b] Department of Computer and Systems Engineering, University of La Laguna, 38200 La Laguna (Tenerife), Spain
[c] Department of Geography and History, University of La Laguna, 38200 La Laguna (Tenerife), Spain
[d] Department of Chemistry, University of La Laguna, 38200 La Laguna (Tenerife), Spain

ARTICLE INFO

Article history:
Received 13 July 2019
Received in revised form 3 May 2021
Accepted 6 May 2021
Available online 21 May 2021

Keywords:
Neural networks
Ensemble methods
Relative importance
Calima
Saharan dust
Canary Islands

ABSTRACT

Analyzing and predicting the concentration of airborne dust is vital to the economic activity and to the health of the population. In this study, we use a set of artificial neural networks that we structure through ensemble learning to yield a complex variable, such as the concentration of dust, based on actual data such as air temperature, relative humidity, atmospheric pressure and wind speed. The statistical performance indices obtained, show the effectiveness of the proposed approach through the application of a cross-validation committee. It is thus vital to have a reliable calculation method for determining relative importances that can be applied to this type of ensemble architecture by way of artificial neural networks.

Unlike other relative importance methods, where calculations are done based directly on the weights in the artificial neural network and whose results in ensemble sets exhibit high dispersion, we propose our own procedure, which selectively chooses the variation in the inputs to readjust the architecture of the neural network. This allows us to measure those variables with the greatest effect on the target variable, thus obtaining the multivariate influence on the surface dust concentration through a computational model.

This method thus provides a real alternative for calculating and estimating relative importance that can be generalized to any type of problem for multivariate systems modeled using artificial neural networks for both, a simple configuration, and an ensemble architecture.

1. Introduction

In an effort to efficiently protect human health and economic activity on the islands, real-time information on airborne dust concentrations is a key component to determining the air quality in the archipelago. We are interested in determining if the behavior of one variable (the concentration of dust on the surface) can be determined from one or several variables in order to better develop our computational procedure using ensemble methods for deep learning and machine learning neural networks so that said approach can be used to understand and improve the forecasts and estimates in any other industrial fouling scenario through the use of artificial neural networks (ANNs). We use the supervised learning approach, resulting from applying this architecture in an effort to study and identify those variables that have the most impact on the sand values.

Accordingly, the main contributions of this paper are summarized below:

(1) Estimate the concentration value of calima, which is the local term used in the Canary Islands, based on the basic variables measured, such as ambient temperature, relative humidity, wind speed and atmospheric pressure. To this end, we developed an ANN model where we apply a cross-validated committee (CVC) ensemble architecture.
(2) Design a hybrid system to calculate the relative importances, Selective Importance Measure (SIM), which combines a sensitivity analysis that eliminates inputs selectively by means of a new adjustment to the network layers in the architecture, and by setting the same initial weights for each new training run.
(3) Validate the strength of the SIM method by estimating which variables measured within the climate environment of the Canary Islands have the greatest impact on the concentration of calima and justifying these results with the related literature.

* Corresponding author.
E-mail addresses: alu0100311273@ull.edu.es (D. Gonzalez-Calvo), raguilar@ull.edu.es (R.M. Aguilar), ccriado@ull.edu.es (C. Criado-Hernandez), lagonmen@ull.edu.es (L.A. Gonzalez-Mendoza).

https://doi.org/10.1016/j.asoc.2021.107497

(4) Verify the stability of the SIM procedure with learning conducted in an ANN with an ensemble architecture, checking the dispersion of the results and comparing it to another related classical method where the importances are calculated directly based on the weights of the ANN.

We intend to use non-experimental data that are collected by passively observing the real world, such that the data are not the result of controlled or empirical experiments. The experimental data are often gathered in laboratory settings, as happens with the natural sciences; specifically, we will resort to time series, which are data collected from observing a variable over time, such as observing certain variables in the air, like temperature, humidity, atmospheric pressure or wind speed. The chronological order of the observations yields potentially important information. Most atmospheric series are related to their recent history.

Calima is the local term used in the Canary Islands to refer to dust from the Sahara Desert. The arrival of Saharan dust was important throughout the Pleistocene and remains so to this day [1,2]. The arrival of airborne dust is made evident by the increased haze in the air, as visibility decreases, and the sky takes on a reddish hue as the density of the dust rises. The entrainment of the air in the areas of origin – located within the Sahara Desert – can occur both due to trade winds (*Harmattan*) and to convective phenomena (*Haboobs*) [3]. In either case, the dust can arrive in the Canarian archipelago with easterly to southerly winds. Due to the proximity of the African continent, calima events are more frequent in Lanzarote, Fuerteventura and Gran Canaria (see Fig. 1).

The influx of Saharan air takes place primarily in fall and winter, with these seasons accounting for 65% of all calima events, versus 35% in the spring and summer [2,4].

Part of the Saharan dust that reaches the Canaries settles via dry deposition, though this is not a dominant phenomenon in the Canaries, with the wet deposition or "blood rain", being particularly striking [5–7]. Rains after winter events can remove a significant amount of dust from the air.

The transfer of Saharan minerals to the Canaries is particularly important, with Saharan minerals like quartz and mica being found in soil and sediments formed during the Quaternary period [8]. Inputs of phosphates and iron appear to account for the fertility of the archipelago's soils and, but calima events also cause grave damage to the economy (by affecting agriculture) and to the health of the population of the Canary Islands [4,5,9]. These events are almost always accompanied by high temperatures and a significant drop in relative humidity, which turns these events, especially in the summer, into times of high forest fire risk [10]. Moreover, the presence of dust, which reaches an average of 127.30 $\mu g/m^3$ versus the 27.76 $\mu g/m^3$ with sea air, noticeably reduces visibility and causes respiratory problems, especially in at-risk groups like young children, the elderly and those with respiratory conditions [2,11]. The presence of dust can also affect the industry and must be considered during the maintenance cycle of rotating industrial components.

The main concept behind an ensemble learning model is the simple intuitive idea of a committee of experts working together to solve a problem. All members contribute their own experience and initiatives and the group as a whole can choose to uphold or to reject a new idea on its own merits [12].

In line with Saviozzi, the ensemble-averaging method is usually implemented to achieve more accurate results than a single ANN [13]. The main idea of this method is to train different networks and combining their outputs in order to have a better prediction. Weng, shows in their experiment with ensemble methods, an improvement of on average 30% in test performance metrics compared to a single two-layer neural network with the same characteristics and features, including number of neurons [14]. Other recent studies conducted that use a combination of multiple artificial neural networks through an ensemble method have exhibited better performance and yielded improved results. This can be useful to deal with real world applications [15–18].

The most relevant ensemble architectures include the bagging method, developed by Breiman [19]. The bagging or bootstrap aggregation method can be used to create and train each new classifier or regression model of the ensemble with a bootstrap sample of the available original training dataset. This yielded results with a lower variance compared to the use of an simple ANN [20].

Studies based on a combination of ensemble methods and deep learning also reached the conclusion that the best option is to group ANNs in an ensemble configuration to solve regression and classification problems; in this case, for example, by means of a bagging-based ensemble design that relies on new, semi-supervised learning, versus classical supervised learning. This new method relies on a parallel, unsupervised learning phase to add new extracted features through a kernel learning regression model [21].

The work undertaken by Liu is expanded with another case that compares models based on kernel learning, whose improved performance and prediction reliability again stem from an ensemble-based model [22].

However, it is important to note that in an ensemble set, not every ANN trained will yield good results. Locating the ANN that exhibit low performance in the final model is a complex task [23]. Along these lines, Parmanto and Murrugarra proposed and verified the efficiency, in terms of the model's performance, by designing an ensemble structure based on the distribution given by the cross-validation index, since the divisions that are made in this method can be used to exclude some training sub-sets with each iteration [24,25].

In order to analyze the relative importance (RI) of an ANN, some researchers have used methods in an effort to obtain the individual contributions from the input variables of an ANN, such as the sensitivity analysis methods that use input perturbation or calculate partial derivatives [26–30]. For example, Dimopoulos and Zeng reached the conclusion that the partial derivative algorithm and input perturbation algorithm performed relatively well [27,28].

Maosen in his research, notes the instability problem of calculations based on sensitivity analysis, which have not paid enough attention to the limitations of this type of method, which is determined by the trained structure of the ANN [31].

Other algorithms, such as the one proposed by Garson, and subsequently modified by Goh, called Modified Garson's algorithm (MGA), yield an approximation of importances for ANNs by doing calculations involving different synaptic weights for the input, hidden and output layers [32].

However, the studies by De Oña et al. indicated that this method of connecting weights to measure the RI with the MGA yielded high variability when applied to a neural network ensemble architecture that has been trained with different initial weights [33].

The paper is organized as follows. Section 2 describes how the data are structured and the architecture is designed, using a committee machine model based on ensemble-averaging with *k-fold* partitioning and multi-layered feedforward neural networks, to estimate the concentration of Saharan dust in the Canary Islands using actual humidity, atmospheric pressure, temperature and wind speed data collected from weather stations in the south of the island of Tenerife. Section 3 presents our results, which are compared with actual data for Saharan dust concentrations taken

Fig. 1. Effect of Saharan dust concentration in the Canary Islands.

from the public database and the BSC-DREAM8b v2.0 model, the goal being to verify the strength of the neural models and to apply various methods intended to measure the relative importance (RI). The paper concludes with a presentation of our findings and a comparison of the methods applied.

2. Methods

The general computation procedure in this study is shown in Fig. 2, where, after a data normalization and structuring phase, the **Saharan dust concentration** is estimated within an ANN ensemble architecture, based on the partitioning generated by the cross-validation method through a supervised training process of four environmental measurements and a calculated variable for the input layer (Ambient temperature; atmospheric pressure; wind speed; relative humidity; month number).

Once the statistical control phase is complete, we apply our SIM method, developed to estimate the RI of the final model obtained, in order to compare our methodology with a classical method to calculate the RI of an ANN.

2.1. Data structure

The data set is divided into two main blocks. The first consists of an initial data set (DS1) containing the input and output data from July 2016 to July 2018, with one sample per minute for the environmental input data. For the actual output and reference values to be forecast, we took the concentration of Saharan dust on the surface of the island of Tenerife, obtained from v2.0 of the BSC-DREAM8b model (Fig. 3), using linear interpolation to adapt its six-hour samples to the one-minute sampling frequency used in the input data.

The second set of independent data (DS2), not used to train the neural network, is from the July 2018 to December 2018 period and also contains input and output data. It is intended exclusively for the testing and results verification phase. In other words, we will forecast the dust concentration using only the environmental input data from DS2 and compare the results statistically to see if they conform to the actual outputs for this period.

As part of preparing the data for the neural network, and in an effort to facilitate the learning of the training algorithm, the data in time series DS1 and DS2 were smoothed (Fig. 4). The data were processed by means of a simple moving average (SMA) with a time period of three days. The SMA is the unweighted mean of the previous N datum points [34].

Wind speed and direction data were extracted and shown graphically in a wind rose (Fig. 5), which displays the predominant directions from the first and second quadrants that are typical in the Canary Islands [1]. The variable wind direction was not considered when building the model, however, given its variability and normalization complexity.

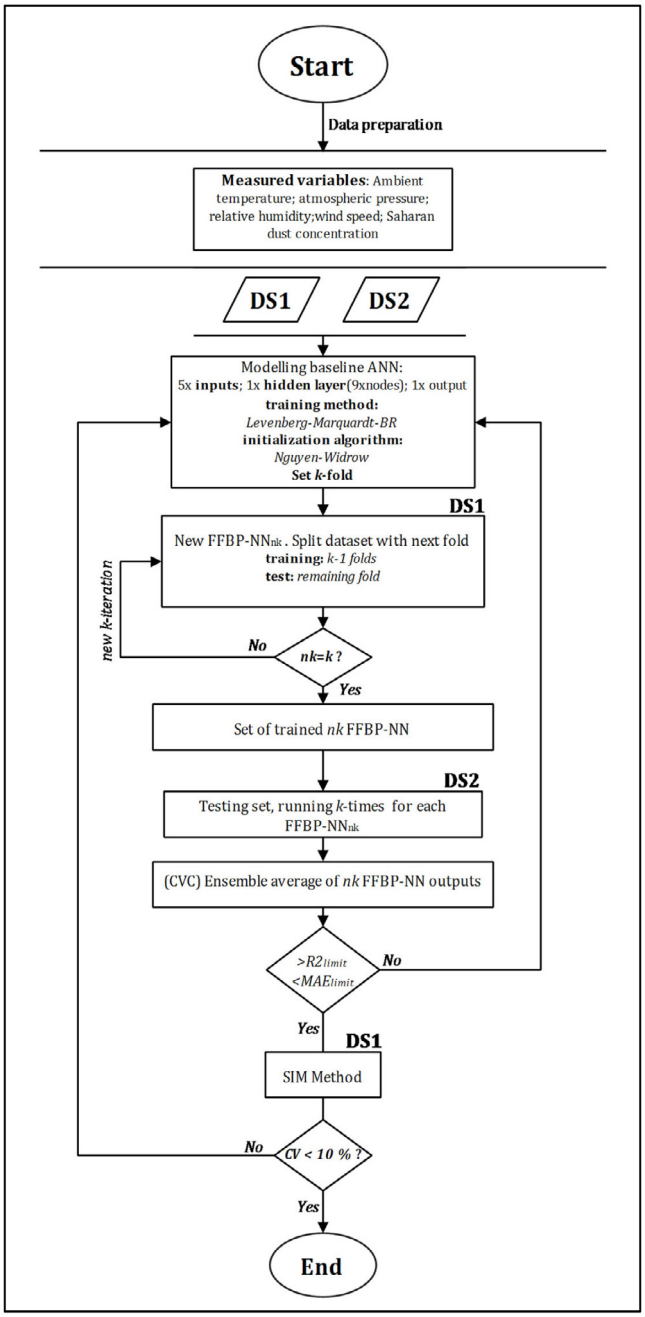

Fig. 2. Flowchart of the main procedure used in this work. Sections 2.1 to 2.4.

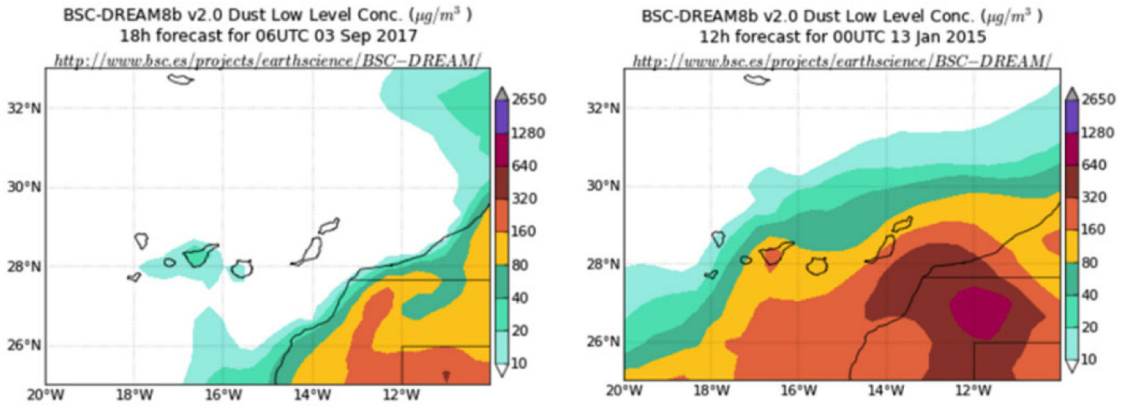

Fig. 3. Data and images from the BSC-DREAM8b-Dust model for Canary Islands, operated by the Barcelona Supercomputing Center. *(*http://www.bsc.es/ESS/bsc-dust-daily-forecast*)*.

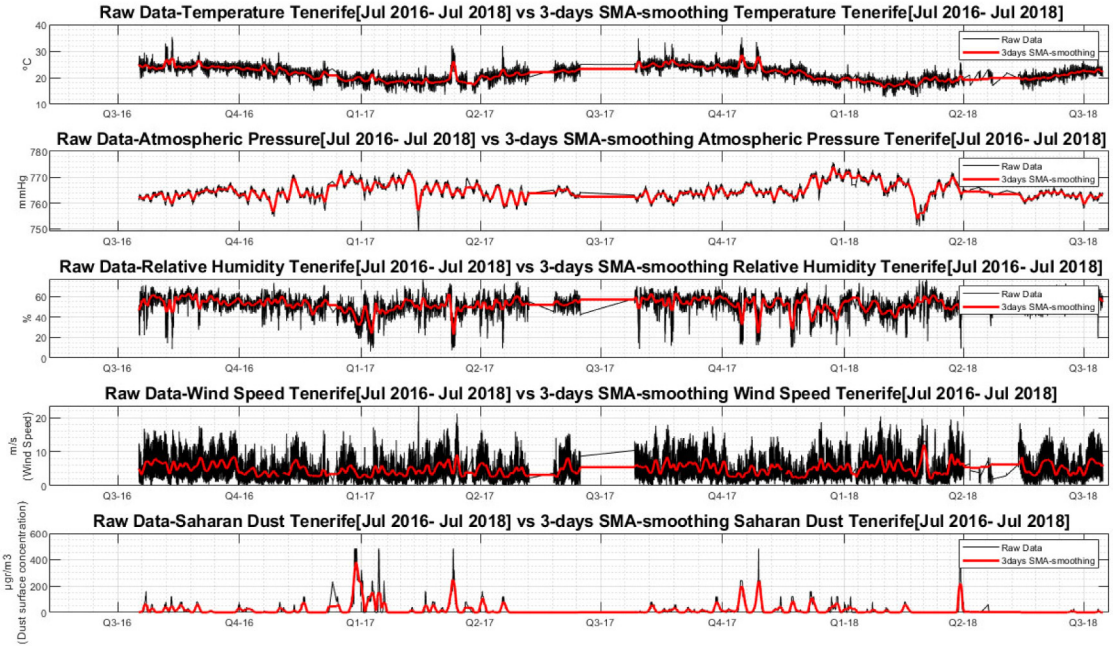

Fig. 4. Raw Data vs 5-day SMA; January, February and March (Q1); April, May and June (Q2); July, August and September (Q3); and October, November and December (Q4).

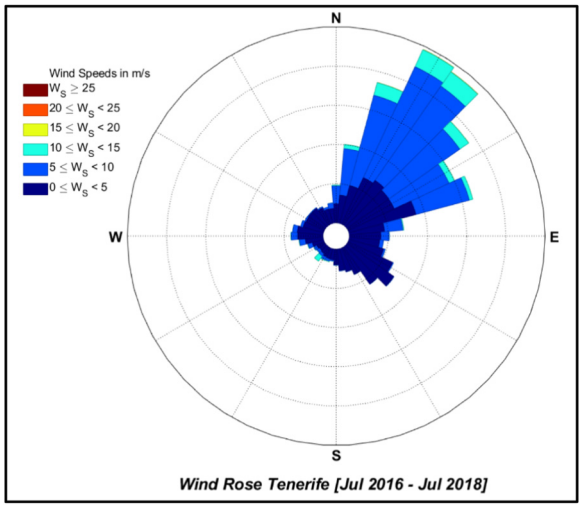

Fig. 5. Wind rose with DS1 in the south of the island of Tenerife.

2.2. Neural network architecture

Artificial neural networks draw their inspiration from biological neural networks and attempt to imitate the brain's behavior to engage in certain actions, such as pattern recognition or object classification [35].

Haykin defines neural networks as follows [36]:

"A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

1. Knowledge is acquired by the network from its environment through a learning process
2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge".

The type of neural network presented in this paper is defined as a Feedforward Back-propagation Neural Network (**FFBP-NN**). Starting below, we propose the main architecture of the baseline model.

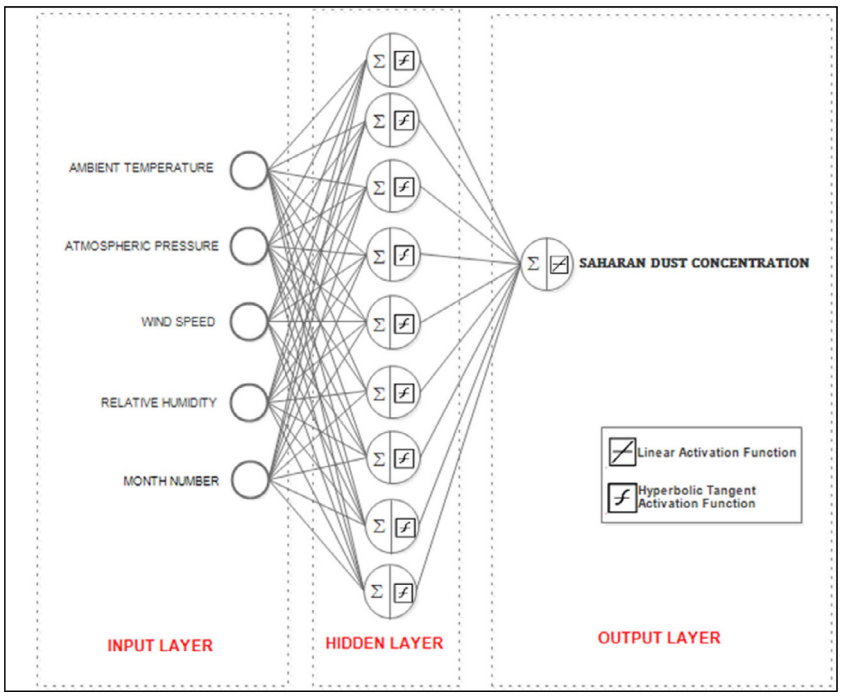

Fig. 6. Topology of the main structure of the neural network, designed with five inputs, nine nodes in the hidden layer with a hyperbolic tangent activation function and one node in the output layer with a linear activation function.

Its basic architecture consists of five inputs (Ambient temperature; atmospheric pressure; wind speed; relative humidity; month number), a single hidden layer with nine nodes with an activation function that uses the hyperbolic tangent, and an output layer with a linear activation function (**Saharan dust concentration**) (Fig. 6). The number of nodes in the hidden layer was adjusted by adapting the method proposed by Huang, which relates the number of inputs to the output variables using artificial neural models in a hidden layer [37].

The statistical performance indices used were the determination coefficient (R^2) and the mean absolute error (MAE). According to Willmott, the MAE offers a natural measure of the mean error (unlike RMSE) [38].

A supervised learning method was implemented and, to improve the efficiency of the multilayer perceptron, we chose a backpropagation algorithm with an adaptive initialization of the weights, determined using the Nguyen–Widrow method, to shorten the training phase [39]. Specifically, we used the Levenberg–Marquardt optimization method, modified through Bayesian regularization, designed to minimize sum-of-square error functions of nonlinear systems [40,41].

The Bayesian regularization makes a slight modification to the optimization method, incorporating Bayes' theorem into the algorithm and compensating for potential overtraining of the neural network [42].

The data re-sampling, cross-validation in *k*-iterations or *k*-fold, technique is used to partition the training data set into *k* subsets with equal dimensions, such that the final training partition is constructed using *k*-1 subsets, and the rest is defined as the validation, or test, set where 90% of each subset will be used in the training phase and the remaining 10% for testing, such that the entire data set is represented [43].

This paper uses a standard value of $k = 10$, since according to Kohavi, the best method to use for model selection is ten-fold stratified cross validation, even if computation power allows using more folds [44]. Other studies indicate that simulations done using ten-fold validations repeatedly yield the best statistical performance [25,45].

2.3. Ensemble learning

For this study, we used the *k*-fold partitioning or cross-validated committees (CVC) ensemble method, similar to the ensemble structure used by Parmanto, Xia, Dong and Murrugarra, where the set in DS1 was partitioned with the same dimensions [24,25,46,47].

Each fold set will be used to build a FFBP-NN_{nk}. In addition, the initial synaptic weights will be different for each partition *nk* and the final ensemble value will be the average of the results obtained with each. We will also obtain a preliminary MAE_0 and R_0^2 (Fig. 7).

The final resulting value will be evaluated using the average value of the combination of DS2 and each FFBP-NN_{nk} trained in the previous step with DS1, until the final statistical performance exceeds the specifications in Eqs. (1) and (2) in terms of MAE and R^2.

In the final validation stage for the models in this paper, the training iteration cycle ends when an average final MAE is obtained that is below a control limit (Eq. (2)). This limit is the result of adapting the Shewart c-chart, statistical quality control method called the control chart for nonconformities and a value of R^2 that exhibits a strong degree of correlation (Eq. (1)), [48]. If the specified limit is not exceeded, the model will readjust the initial weights and the nodes of the hidden layer in each FFBP-NN_{nk} (Fig. 7).

$$R^2_{limit} > \mathbf{0.7} \tag{1}$$

$$MAE_{limit} = \bar{c} + \mathbf{3}\sqrt{\bar{c}} \tag{2}$$

where $\bar{c}$ represents the average value of the DS1 set for the target output variable.

The possibility exists that the limit values specified in Eqs. (1) and (2) will not be exceeded. In this case, consideration must be given to reviewing the normalization of the data sets or adding a new variable that might help improve the learning phase.

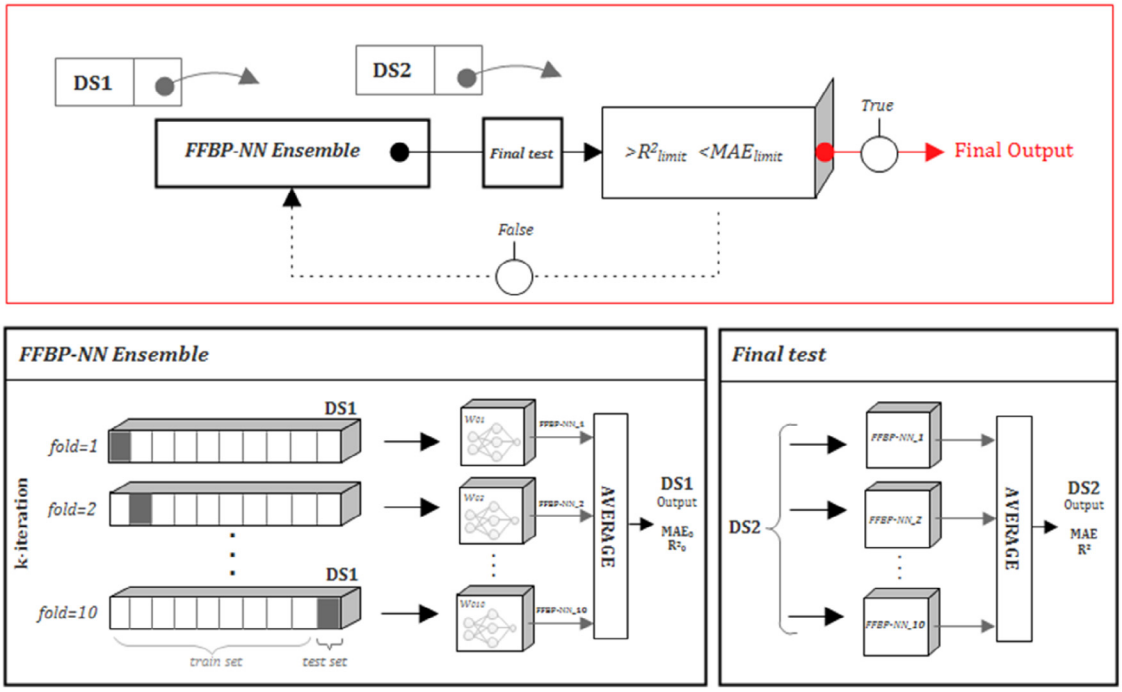

Fig. 7. Main diagram of the training algorithm used, with detailed views of the Ensemble Learning k-fold partitioning and Final Test blocks.

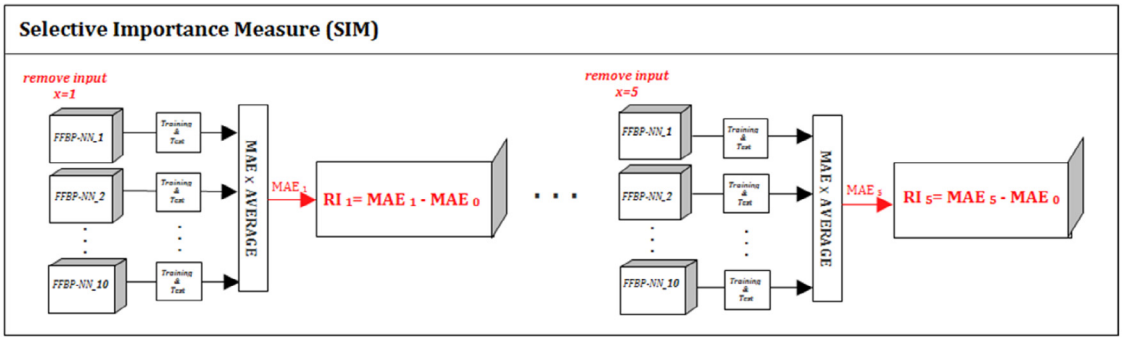

Fig. 8. Main diagram of the SIM procedure.

2.4. Relative importance calculation

Next, we describe the methods proposed for measuring the degree of relative importance of the various variables that make up the model for our environmental system. We include our own procedure, which combines the alteration of the inputs and new training of the ANN to adjust the new architecture, and a classical method based on the weights of an ANN.

2.4.1. Selective importance measure

We designed our own computational procedure for this work that offers a method for studying and directly measuring the RI calculation, following the training of a single artificial neural network or a "committee" of networks.

$$nn\,(MAE_x) = \frac{\sum_{nk=1}^{k}\left(MAE_{(m-1)nk}\right)}{k} \tag{3}$$

$$nn\,(MAE_0) = \frac{\sum_{nk=1}^{k}\left(MAE_{0nk}\right)}{k} \tag{4}$$

where x indicates the input variable that is evaluated based on the average MAE_x, obtained from the ensemble of all the k-fold during the new training of the FFBP-NN_{nk} with DS1.

Iterations(k) are carried out with new training after eliminating an input variable with each cycle. This means that for each iteration carried out in each FFBP-NN_{nk}, the new training configuration will always have $m-1$ inputs, where m is the total number of inputs used in the system being modeled (Eq. (3)).

Finally, to obtain the final RI value for each input variable (Eq. (5)), the new training $nn\,(MAE_x)$, it is compared with the average performance $nn(MAE_0)$, resulting from the final original ensemble model with DS1 (Eq. (4)).

$$\boldsymbol{SIM}_x = nn\,(MAE_x) - nn\,(MAE_0) \tag{5}$$

The selection process for eliminating input variables implies deleting them individually, along with their set of weights, in order to again calculate the average mean absolute error resulting from the new training of the ensemble $nn\,(MAE_x)$, to analyze the result of altering the model (Fig. 8).

For each of the final trained FFBP-NN_{nk} of the k-fold partitioning ensemble, all of their initial input, output and bias synaptic weights, w_{i0xh}, w_{o0h}, b_{i0h} and b_{o0} respectively, will be stored. In the expression, h is the neuron of the associated hidden layer, i the input layer, o the output layer and the value 0 indicates the initial state.

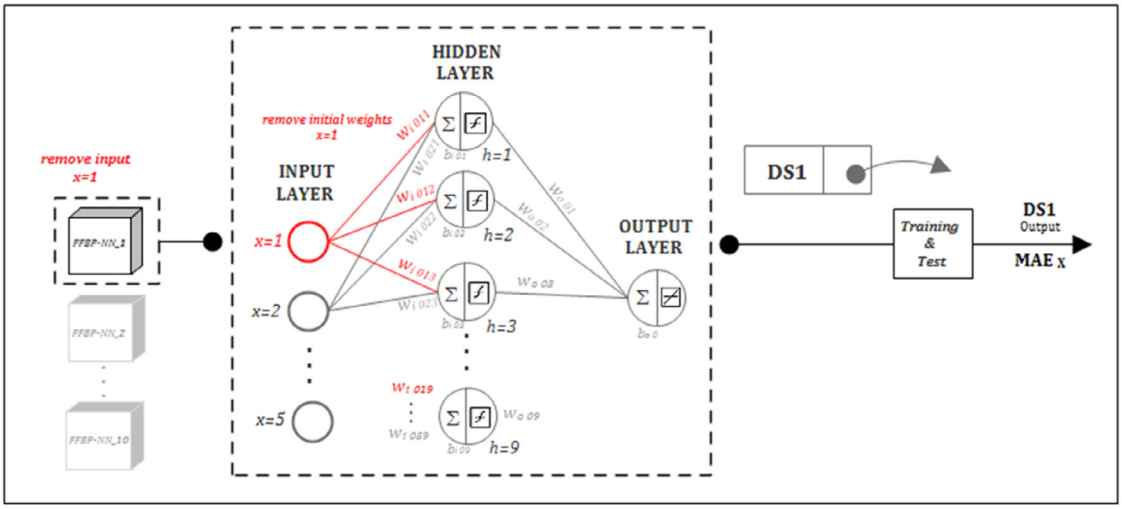

Fig. 9. Expanded diagram of the SIM procedure for eliminating one input variable and its corresponding weights toward the hidden layer.

These stored weight and bias variables will be used in the initial configuration of the training for the new ensemble set, $nn\,(MAE_x)$ after eliminating the initial input weights w_{i0x}, which correspond to the x variable that has been removed (Fig. 9). The process concludes by retraining each FFBP-NN_{nk} with the new configuration and repeating the cycle until all of the model's inputs are processed.

Therefore, for this procedure to be valid, a final control condition is required that somehow verifies the degree of dispersion of the results (Eq. (6)). To this end, a limit value for the coefficient of variation (CV) is set at 10% of the resulting final set. Below this value, the results are considered to have a low variability [49,50].

The expression for the CV is defined as:

$$\boldsymbol{CV}_x(\%) = \frac{\sigma_{nn(MAEx)}}{\mu_{nn(MAEx)}} \tag{6}$$

where μ indicates the average of the $nn\,(MAEx)$ ensemble set, obtained for each input, and σ is the average standard deviation for that same set.

If this limit is exceeded for any of the input variables and the resulting degree of variation is high, this might indicate that specific variable does not contribute significantly to the set, allowing us to discard it, filter it out of training or revise the normalization of its data set.

Lastly, the resulting algorithm is shown using the following pseudocode:

Step 1: If the final trained model of the average ensemble with k-partitioning (DS2) MAE & R^2 < MAE_{LIM} & $R^2{}_{LIM}$. Set the FFBP-NN_{nk} (DS1).
Step 2: Store initial weights and biases in the constants $w_{i\,0xh}$, $w_{o\,0h}$, $b_{i\,0h}$ and $b_{o\,0}$ of each FFBP-NN_{nk} (DS1).
Step 3: Eliminate input x of FFBP-NN_{nk} and in the matrix with the initial input weights, eliminate the corresponding $w_{i\,0xh}$ weights.
Step 4: Set up the properties of the new FFBP-NN_{nk} with the resulting weights $w_{i\,0xh}$, $w_{o\,0h}$, $b_{i\,0h}$ and $b_{o\,0}$ from Step 3.
Step 5: Start training the new FFBP-NN_{nk} (DS1).
Step 6: Analyze MAE_x of result from step 5. Store result.
Step 7: Load initial weights from Step 2 and repeat steps 3 to 6 for the entire set of input variables m.
Step 8: Repeat steps 1 to 7, k times until the average result of the ensemble set $nn(MAE_x)$ is obtained for all m inputs.
Step 9: Calculate SIM_X for the m inputs.
Step 10: Check variability of the results by verifying CV_X < CV_{LIMIT} for the entire SIM_X set obtained for all m input variables.

2.4.2. Modified Garson's algorithm

This method for the synaptic weights of an ANN is derived from the first algorithm proposed by Garson and later modified by Goh, represented as follows [32,51]:

$$\boldsymbol{MGA}_{index} = \sum_{h=1}^{j} \frac{|w_{ih} \cdot w_{ho}|}{\sum_{i=1}^{m} |w_{ih} \cdot w_{ho}|} \tag{7}$$

where m and j indicate the number of inputs and hidden nodes, respectively. w_{ih} are the synaptic weights between input i and the neuron in the hidden layer h, and w_{ho} are the weights between the hidden layer h and output o.

3. Results

3.1. Performance analysis for the ensemble model

In total, for this work we will selectively analyze a total of 50x FFBP-NN; that is, 10x FFBP-NN for each input variable x. This means that each variable will yield a set of 10 results in terms of the MAE.

For the first data prediction with DS1 data, whose output was known beforehand by the FFBP-NN, the statistical performance indices obtained were $\mathbf{R^2}$≈0.93 and a mean absolute error of **MAE≈ 12.57** μgr/m^3 (Fig. 10), with a concentrated error distribution close to zero (Fig. 11).

For the forecast with data from DS2 data, the output of which was unknown and whose input data were not used to train the FFBP-NN, the statistical performance indices obtained were $\mathbf{R^2}$≈**0.87** and a mean absolute error of **MAE≈12.71** μgr/m^3 (Fig. 12). The error distribution was slightly scattered around the zero central value but limited to low error values (Fig. 13).

3.2. Measures of relative importance for the ensemble model

Having successfully completed the evaluation phase with new DS2 data, we know that we have a model that has correctly learned the system dynamics. Therefore, for this new phase we only want to check each variable's importance to the system, and we are interested in doing so with the best performing data set, in this case, DS1.

Next, we present the results for each of the methods described in Section 2.4, that measure the level of RI in terms of percent; that is, the RI value that each input variable in the set represents with respect to the total set (Fig. 14). As described in Section 2.4.1, a control method was set up for the SIM method, in terms of the CV (Fig. 15).

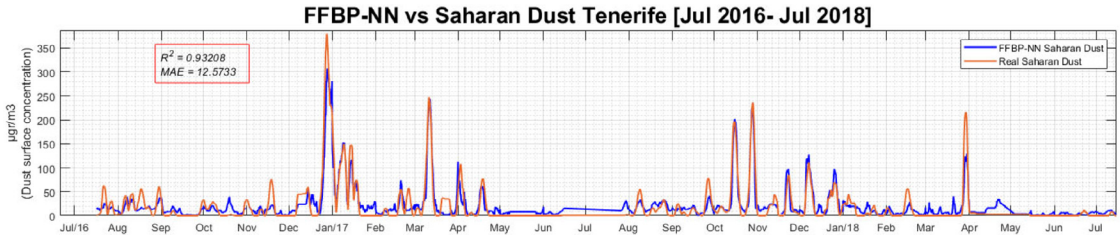

Fig. 10. Predicted results and statistical comparison for the data in DS1. July 2016 to July 2018.

Table 1
Average results obtained for each input variable in the SIM, MGA calculation to the FFBP-NN Ensemble with k-fold partitioning (DS1).

Input variable	SIM (%)	SIM Ranking	SIM CV (%)	MGA (%)	MGA Ranking	MGA CV (%)
Temperature	34.63	**1**	2.6	19.9	**3**	34
Atmospheric Pressure	9.84	**4**	5.6	22.46	**2**	31.74
Wind speed	4.96	**5**	4.5	12.69	**5**	41.92
Relative humidity	26.60	**2**	3.4	19.70	**4**	16.61
Month number	23.96	**3**	2.8	25.13	**1**	30.68

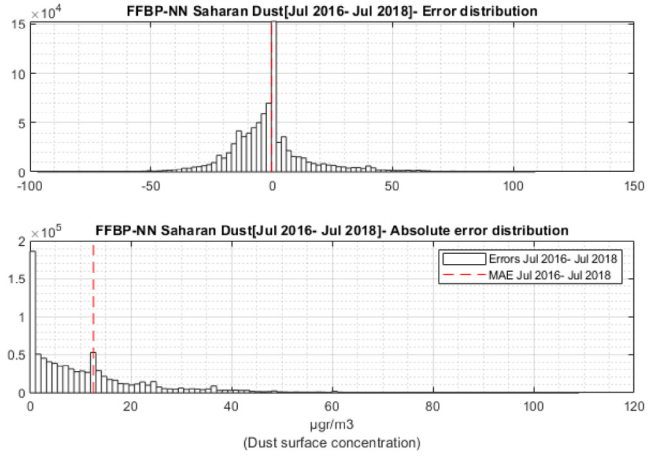

Fig. 11. Error distribution histogram for the forecast with DS1.

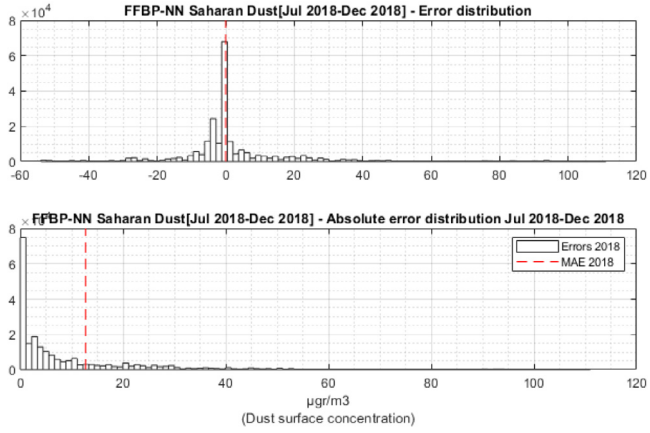

Fig. 13. Error distribution histogram for the forecast with DS2.

The main results of the average RI calculations, ranked by relevance, and measures of CV obtained for our ensemble models, are shown in Table 1, which compares our proposed SIM method to the MGA based directly on the weights of an ANN.

The average results for our proposed SIM method (Fig. 14) agree with observations made in the Canary Islands during the arrival of continental tropical air masses from the Sahara Desert, commonly referred to as "southern weather", which features high dust concentrations in mainly high-temperature, low relative humidity conditions and month number [10,52,53]. The changes in relative humidity are key to the formation of dust [10,54–56].

The wind speed variable ranks last and does not exhibit as much importance compared to the other variables, which have more of an effect on the sand concentration measured on the island of Tenerife.

In their research, Tao and Jamalizadeh discussed the irregular contribution of the wind speed variable in different scenarios, and found some cases where there was no direct correlation between this variable and the presence of sand storms [57,58].

The literature reviewed describes the problem of randomness when re-training a neural network due to the initialization of the weights. However, our solution takes this into account, which is why the algorithm, as input variables are eliminated in each new training run, always uses the same initial weights that were

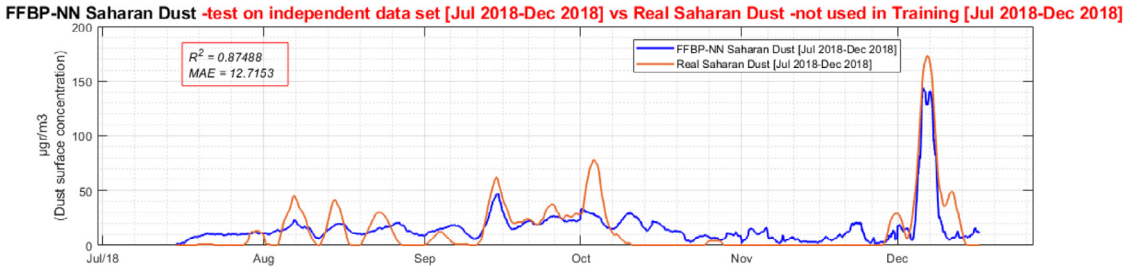

Fig. 12. Results of the forecast and statistical comparison for the data in DS2.

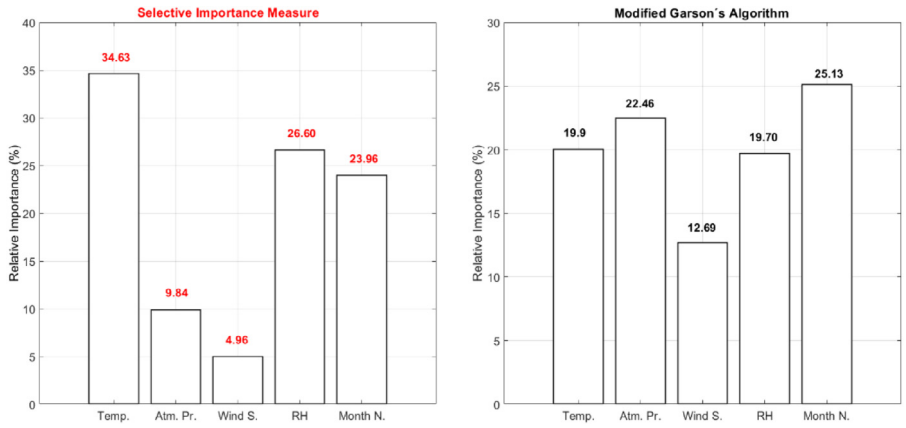

Fig. 14. Degree of RI obtained for the SIM & MGA methods, for each input variable to the FFBP-NN Ensemble (DS1); Temperature (Temp.); Atmospheric Pressure (Atm. Pr.); Wind speed (Wind S.); Relative humidity (RH.); Month Number (Month N.).

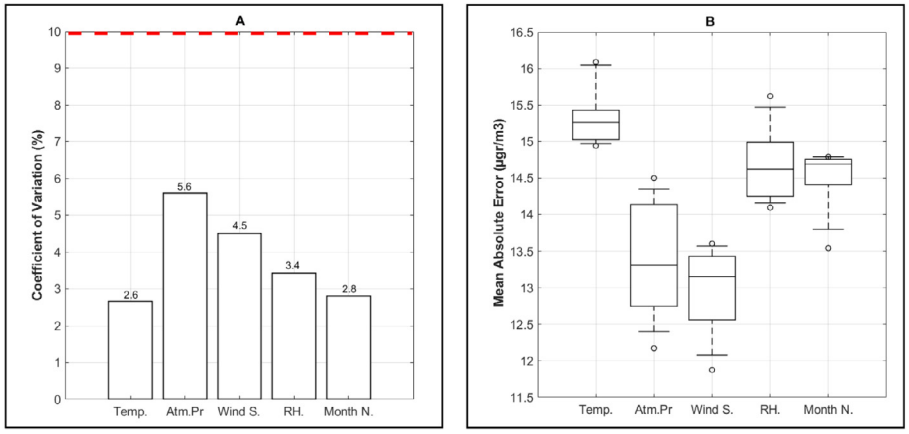

Fig. 15. Coefficients of Variation (CV) obtained for each input variable of the SIM procedure (A) and boxplots representing the amount of MAE spread for each input, k times with SIM method (B).

used in the first original trained network. Therefore, the initial conditions of the new training run $m-1$, are the same for each iteration for the different inputs.

In their research, Maosen et al. noted, in addition to the problems involving the sensitivity analyses for RI, the lack of determination when defining the number of neurons in the hidden layer [31]. To address this, our CVC procedure, and as part of the initial design of the ANNs, utilizes the same method to select the ideal number of hidden layers, which is set based on the number of inputs and outputs [37].

Findings with the MGA reflect a high degree of output variability (Table 1). Some of the average results are similar to those calculated with the SIM method in terms of relative humidity, month number and wind speed, though somewhat more scattered with respect to the air temperature and atmospheric pressure variables.

The values of RI with the MGA obtained high variability when is applied to an FFBP-NN ensemble architecture that has been trained with different initial weights. Therefore, does not guarantee the validity of the results of RI [33].

Lastly, note that the results satisfy the quality control we established, which thus validates the SIM method based on the degree of variability (Fig. 15), where the results lie within the level defined as "low" (<10%) in the resulting final set. This shows that the final 10xFFBP-NN for each of the five input variables used in this model are relevant and indicate the correct distribution and good performance of the entire data set utilized in DS1.

4. Conclusion

We managed to obtain a model that can be used to estimate a more complex variable, like the calima concentration on the surface, from a basic time series with online data measured on the island of Tenerife. Its statistical performance indices show the effectiveness of the FFBP-NN ensemble proposed, and structured using k-fold partitioning, with $k = 10$ and initial different weights in each fold. The results were combined based on an average value after having been validated using a statistical performance control limit. Essential to the results was the preliminary normalization and smoothing of the data set by means of an SMA.

The results of the CVC procedure improve the individual performance by combining the average resulting from all the trained ANNs. It is thus essential to have a reliable calculation method to determine relative importances that can be applied to this type of ensemble architecture with ANNs.

Unlike other RI methods, where calculations are done based directly on the weights in the artificial neural network and whose results as ensemble sets exhibit significant dispersion and large randomness in each new ANN trained, in our procedure we propose a hybrid SIM system, in which we opted to analyze the result of a new training method that selectively relies on the input values and that also readjusts the ANN architecture by setting the initial weights in order to assess its training in terms of the mean absolute error.

We also managed to achieve good stability in the results obtained for all the iterations, which closely approximated the environmental reality studied. This allowed us to apply a calculation methodology to differentiate the values with the greatest effect.

The degree of RI calculated using our own SIM method correctly indicated two of the key variables that affect sand concentration: relative humidity (26.6%) and air temperature (34.65%). This is in contrast to the MGA calculation, where temperature ranks third at 19.7%, and relative humidity fourth at 19.9%, with the month and atmospheric pressure leading the ranking.

However, the statistical dispersion control carried out for the MGA case based on a calculation of the weights yields highly variable results, confirming the findings obtained in other research, as discussed in this paper.

SIM method thus provides a real alternative for calculating and estimating RI that can be generalized to any type of problem for multivariate systems modeled using ANNs, both for a simple configuration and for an ensemble architecture.

Having a model trained with historical data spanning several years, and having learned the relationship between the output variable and the basic environmental variables, will allow us to continue with a new stage of simulations involving the sand output variable in any environmental scenario we need to study.

Moreover, the information that is obtained by calculating the RI using SIM can be used to improve the predictions or estimates of an actual model, thereby eliminating irrelevant variables. This might occur, for example, when studying a case of industrial fouling resulting from an environmental phenomenon, providing an insight into the phenomenon's dynamics over time and thus minimize any associated costs.

Because of this, and as a continuation of this work, we will use the sand concentration estimate from our model to include it in a new ensemble neural architecture – trained using the SIM method – associated with a rotational industrial component. This will help us to study the effect that the various variables involved in the process and in the local environment have on the performance of an industrial machine, and to develop our own methodology for data-driven maintenance, applied to the compressor of a power generation gas turbine.

In this particular case, having a model showing the local dust concentration as a function of basic environmental variables, will allow us to know in advance the evolution of the thermal performance of the machine, by using the dust variable as part of the input variables of a new CVC model. This will provide a medium-term estimation to be able to plan the next cleaning maintenance shutdown and ensure in real time the correct operating status of the gas turbine's production.

CRediT authorship contribution statement

D. Gonzalez-Calvo: Investigation, Writing, Methodology, Programming/coding. **R.M. Aguilar:** Writing - review & editing. **C. Criado-Hernandez:** Supervision, Writing. **L.A. Gonzalez-Mendoza:** Supervision, Writing - review & editing.

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] C. Criado, et al., Intercalaciones de polvo sahariano en paleodunas bioclásticas de fuerteventura (Islas Canarias), Cuaternario Geomorfol. 26 (1–2) (2012) 73–88.

[2] P. Dorta, et al., Frecuencia, estacionalidad y tendencias de las advecciones de aire sahariano en Canarias (1976–2003), Invest. Geogr. (38) (1976).

[3] K. Pye, Aeolian Dust and Dust Deposits, Academic Press, 1987.

[4] P. Dorta, et al., Algunas consideraciones sobre la importancia del polvo de origen sahariano en el clima del archipiélago canario y su aporte a las aguas superficiales oceánicas: El episodio de abril de 2002, en el agua y el clima, in: III Congreso de la Asociación Española de Climatología, 2002, pp. 13–24.

[5] M.E. Torres-Padrón, et al., Variability of dust inputs on the canigo zone, Deep-sea Res. II 49 (2002) 3455–3464, http://dx.doi.org/10.1016/S0967-0645(02)00091-7.

[6] C.A. Friese, et al., Environmental factors controlling the seasonal variability in particle size distribution of modern Saharan dust deposited off Cape blanc, Aeolian Res. 22 (2016) 165–179, http://dx.doi.org/10.1016/j.aeolia.2016.04.005.

[7] C. Criado, P. Dorta, An unusual blood rain over canary islands (Spain). The storm of january 1999, J. Arid Environ. 55 (1999) 765–783, http://dx.doi.org/10.1016/S0140-1963(02)00320-8.

[8] C.H. Mizota, Y. Matsuhisa, Isotopic evidence for the eolian origin of quartz and mica in soils developed on volcanic materials in the canary archipelago, Geoderma 66 (3–4) (1995) 313–320, http://dx.doi.org/10.1016/0016-7061(95)00004-8.

[9] N.J. Middleton, Desert dust hazards: a global review, Aeolian Res. 24 (2017) 53–63, http://dx.doi.org/10.1016/j.aeolia.2016.12.001.

[10] P. Dorta, Aproximación a la influencia de las advecciones de aire sahariano en la propagación de los incendios forestales en la provincia de santa cruz de tenerife, in: XVII Congreso De Geógrafos Españoles, 2001.

[11] J. García, et al., Invasión de viento Sahariano y su impacto en la asistencia sanitaria urgente, Emergencias 13 (2001) 372–376.

[12] B. Baruque, E. Corchado, Fusion methods for unsupervised learning ensembles, Stud. Comput. Intell. 322 (2011) http://dx.doi.org/10.1007/978-3-642-16205-3.

[13] M. Saviozzi, S. Massucco, F. Silvestro, Implementation of advanced functionalities for distribution management systems: load forecasting and modeling through artificial neural networks ensembles, Electr. Power Syst. Res. 167 (2019) 230–239, http://dx.doi.org/10.1016/j.epsr.2018.10.036.

[14] Weng, et al., Predicting short-term stock prices using ensemble methods and online data sources, Expert Syst. Appl. 112 (2018) 258–273, http://dx.doi.org/10.1016/j.eswa.2018.06.016.

[15] M. González, et al., Increase attractor capacity using an ensembled neural network, Expert Syst. Appl. 71 (2017) 206–215, http://dx.doi.org/10.1016/j.eswa.2016.11.035.

[16] H. Ashtawy, N. Mahapatra, Bgn-score and bsn-score: bagging and boosting based ensemble neural networks scoring functions for accurate binding affinity prediction of protein-ligand complexes, BMC Bioinformatics 16 (4) (2015) http://dx.doi.org/10.1186/1471-2105-16-S4-S8.

[17] N. Alami, M. Meknassi, N. En-nahnahi, Enhancing unsupervised neural networks-based text summarization with word embedding and ensemble learning, Expert Syst. Appl. 123 (2019) 195–211, http://dx.doi.org/10.1016/j.eswa.2019.01.037.

[18] Z. Wang, C. Lu, B. Zhou, Fault diagnosis for rotary machinery with selective ensemble neural networks, Mech. Syst. Signal Process. 113 (2018) 112–130, http://dx.doi.org/10.1016/j.ymssp.2017.03.051.

[19] L. Breiman, Bagging predictors, Mach. Learn. 24 (1996) 123–140, http://dx.doi.org/10.1023/A:1018054314350.

[20] A.S. Khwaja, et al., Improved short-term load forecasting using bagged neural networks, Electr. Power Syst. Res. 125 (2015) 109–115, http://dx.doi.org/10.1016/j.epsr.2015.03.027.

[21] Y. Liu, et al., Ensemble deep kernel learning with application to quality prediction in industrial polymerization processes, Chemometr. Intell. Lab. Syst. 174 (2018) 15–21, http://dx.doi.org/10.1016/j.chemolab.2018.01.008.

[22] Y. Liu, Z. Zhang, J. Chen, Ensemble local kernel learning for online prediction of distributed product outputs in chemical processes, Chem. Eng. Sci. 137 (2015) 140–151, http://dx.doi.org/10.1016/j.ces.2015.06.005.

[23] Z. Zhou, J. Wu, W. Tang, Ensembling neural networks: Many could be better than all, Artificial Intelligence 137 (1–2) (2002) 239–263, http://dx.doi.org/10.1016/S0004-3702(02)00190-X.

[24] B. Parmanto, P. Munro, H. Doyle, Improving committee diagnosis with resampling techniques, in: Proceedings of the 8th International Conference on Neural Information Processing Systems, 1996, pp. 882–888.

[25] N.E. Murrugarra Llerena, L. Berton, A.d. A. Lopes, Graph-based cross-validated committees ensembles, in: International Conference on Computational Aspects of Social Networks, 2012.

[26] M.I. Gevrey, Y. Dimopoulos, S. Leka, Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecol. Modell. 160 (3) (2003) 249–264, http://dx.doi.org/10.1016/S0304-3800(02)00257-0.

[27] Y. Dimopoulos, P. Bourret, S. Lek, Use of some sensitivity criteria for choosing networks with good generalization ability, Neural Process. Lett. 2 (1995) 1–4, http://dx.doi.org/10.1007/BF02309007.

[28] X. Zeng, D. Yeung, A Quantified Sensitivity Measure for Multilayer Perceptron To Input Perturbation, Vol. 15, (1) MIT Press, 2003, pp. 183–212, http://dx.doi.org/10.1162/089976603321043757.
[29] W. Wang, P. Jones, D. Partridge, Assessing the impact of input features in a feedforward neural network, Neural Comput. Appl. 9 (2000) 101–112, http://dx.doi.org/10.1007/PL00009895.
[30] J. Montaño, A. Palmer, Numeric sensitivity analysis applied to feedforward neural networks, Neural Comput. Appl. 12 (2003) 119–125, http://dx.doi.org/10.1007/s00521-003-0377-9.
[31] M. Cao, P. Qiao, Neural network committee-based sensitivity analysis strategy for geotechnical engineering problems, Neural Comput. Appl. 17 (5–6) (2008) 509–519, http://dx.doi.org/10.1007/s00521-007-0143-5.
[32] A.T.C. Goh, Back-propagation neural networks for modeling complex systems, Artif. Intell. Eng. 9 (1995) 143–151, http://dx.doi.org/10.1016/0954-1810(94)00011-S.
[33] J. Oña, C. Garrido, Extracting the contribution of independent variables in neural network models: a new approach to handle instability, Neural Comput. Appl. 25 (3–4) (2014) 859–869, http://dx.doi.org/10.1007/s00521-014-1573-5.
[34] Y. Chou, Statistical analysis: With business and economic applications, 1975.
[35] J.P. Jesan, D.M. Lauro, Human brain and neural network behavior a comparison, Ubiquity 4 (37) (2003) 12–18, http://dx.doi.org/10.1145/962068.958078.
[36] S. Haykin, Introduction, in Neural Networks and Learning Machines, 1999, pp. 33–34.
[37] G. Huang, Learning capability and storage capacity of two-hidden-layer feedforward networks, IEEE Trans. Neural Netw. 14 (2) (2003) http://dx.doi.org/10.1109/TNN.2003.809401.
[38] C.J. Willmott, K. Matsuura, Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance, Clim. Res. 30 (2005) 79–82, http://dx.doi.org/10.3354/cr030079.
[39] D. Nguyen, B. Widrow, Neural networks for self-learning control systems, IEEE Control Syst. Mag. 10 (3) (1990) 18–23, http://dx.doi.org/10.1109/37.55119.
[40] S. Sapna, Backpropagation learning algorithm based on levenberg marquardt algorithms, in: The Fourth International Workshop on Computer Networks and Communications, 2012.
[41] J. Cervantes, et al., Comparative analysis of the techniques used in a recognition system of plant leaves, RIAI Rev. Iberoam. Autom. Inf. Dnd. 14 (1) (2017) 104–114, http://dx.doi.org/10.1016/j.riai.2016.09.005.
[42] F. Burden, D. Winkler, Bayesian regularization of neural networks, Lit. Rev. Methods Mol. Biol. 458 (2009) 23–42, http://dx.doi.org/10.1007/978-1-60327-101-1_3.
[43] I. Berrar, W. Dubitzky, Avoiding model selection bias in small sample genomic datasets, Bioinformatics 22 (19) (2006) 2453, http://dx.doi.org/10.1093/bioinformatics/btl435.
[44] R. Kohavi, A study of cross validation and bootstrapfor accuracy estimation and model selection, in: Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, pp. 1137–1143.
[45] S. Borra, A. Di ciaccio, Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods, Comput. Statist. Data Anal. 54 (12) (2010) 2976–2989, http://dx.doi.org/10.1016/j.csda.2010.03.004.
[46] R. Xia, C. Zong, S. Li, Ensemble of feature sets and classification algorithms for sentiment classification, Inform. Sci. 181 (6) (2011) 1138–1152, http://dx.doi.org/10.1016/j.ins.2010.11.023.
[47] Y. Dong, K. Han, A comparison of several ensemble methods for text categorization, in: IEEE International Conference on Services Computing, 2004, pp. 419–422.
[48] D. Montgomery, Introduction To Statistical Quality Control, sixth ed., 2008, p. 289.
[49] M.A. Braz, et al., Classification of the coefficient of variation to variables in beef cattle experiments, Ciencia Rural 47 (2017) 11, http://dx.doi.org/10.1590/0103-8478cr20160946.
[50] F.P. Gomes, Curso de Estatística Experimental, 1985.
[51] G.D. Garson, Interpreting neural network connection weights, AI Expert 6 (1991) 47–51.
[52] L. Cana, The Saharan dust episode of 26 february 2000 over the canary archipelago: a synoptic overview, Weather 57 (10) (2000) 385–389, http://dx.doi.org/10.1256/wea.275.01.
[53] K. Gyan, et al., African dust clouds are associated with increased paediatric asthma accident and emergency admissions on the Caribbean island of trinidad, Int. J. Biometeorol. 49 (2005) 371–376, http://dx.doi.org/10.1007/s00484-005-0257-3.
[54] J. Csavina, et al., Effect of wind speed and relative humidity on atmospheric dust concentrations in semi-arid climates, Sci. Total Environ. 487 (1) (2014) 82–90, http://dx.doi.org/10.1016/j.scitotenv.2014.03.138.
[55] A.M. Al-salihi, T.H. Mohammed, The effect of dust storms on some meteorological elements over baghdad, Iraq: study cases, Iosr J. Appl. Phys. (Iosr-Jap) 7 (2) (2015) 1, http://dx.doi.org/10.9790/4861-07220107.
[56] J.L. Diaz-hernandez, A. Sanchez-Navas, Saharan dust outbreak sand iberulite episodes, JGR Atmos. 121 (2016) 7064–7078, http://dx.doi.org/10.1002/2016JD024913.
[57] T. Gao, et al., Impacts of climate abnormality on remarkable dust storm increase of the Hunshdak Sandy Lands in northern China during 2001–2008, Meteorol. Atmos. Sci. 19 (3) (2001) 265–278, http://dx.doi.org/10.1002/met.251.
[58] P. Jamshid, A. Vahid, Dust storm prediction using ANNs technique: A case study-Zabol City, World Acad. Sci. Eng. Technol. Int. J. Civil Environ. Eng. 2 (7) (2008) http://dx.doi.org/10.5281/zenodo.1061318.