

RESEARCH ARTICLE

# An Approach to Improve the Performance of PM Forecasters

Paulo S. G. de Mattos Neto<sup>1\*</sup>, George D. C. Cavalcanti<sup>1</sup>, Francisco Madeiro<sup>2</sup>, Tiago A. E. Ferreira<sup>3</sup>

**1** Centro de Informática, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil, **2** Universidade Católica de Pernambuco, Recife, Pernambuco, Brazil, **3** Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil

\* [psgmn@cin.ufpe.br](mailto:psgmn@cin.ufpe.br)



## OPEN ACCESS

**Citation:** de Mattos Neto PSG, Cavalcanti GDC, Madeiro F, Ferreira TAE (2015) An Approach to Improve the Performance of PM Forecasters. PLoS ONE 10(9): e0138507. doi:10.1371/journal.pone.0138507

**Editor:** Alejandro Raul Hernandez Montoya, Universidad Veracruzana, MEXICO

**Received:** March 30, 2015

**Accepted:** August 30, 2015

**Published:** September 28, 2015

**Copyright:** © 2015 de Mattos Neto et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All PM series files were accessed from a third party source and are available to all interested parties at: <http://aqicn.org/city/helsinki/>.

**Funding:** This work was partially supported by Brazilian agencies: CNPq and Fapece.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

The particulate matter (PM) concentration has been one of the most relevant environmental concerns in recent decades due to its prejudicial effects on living beings and the earth's atmosphere. High PM concentration affects the human health in several ways leading to short and long term diseases. Thus, forecasting systems have been developed to support decisions of the organizations and governments to alert the population. Forecasting systems based on Artificial Neural Networks (ANNs) have been highlighted in the literature due to their performances. In general, three ANN-based approaches have been found for this task: ANN trained via learning algorithms, hybrid systems that combine search algorithms with ANNs, and hybrid systems that combine ANN with other forecasters. Independent of the approach, it is common to suppose that the residuals (error series), obtained from the difference between actual series and forecasting, have a white noise behavior. However, it is possible that this assumption is infringed due to: misspecification of the forecasting model, complexity of the time series or temporal patterns of the phenomenon not captured by the forecaster. This paper proposes an approach to improve the performance of PM forecasters from residuals modeling. The approach analyzes the remaining residuals recursively in search of temporal patterns. At each iteration, if there are temporal patterns in the residuals, the approach generates the forecasting of the residuals in order to improve the forecasting of the PM time series. The proposed approach can be used with either only one forecaster or by combining two or more forecasting models. In this study, the approach is used to improve the performance of a hybrid system (HS) composed by genetic algorithm (GA) and ANN from residuals modeling performed by two methods, namely, ANN and own hybrid system. Experiments were performed for PM<sub>2.5</sub> and PM<sub>10</sub> concentration series in Kallio and Vallila stations in Helsinki and evaluated from six metrics. Experimental results show that the proposed approach improves the accuracy of the forecasting method in terms of fitness function for all cases, when compared with the method without correction. The correction via HS obtained a superior performance, reaching the best results in terms of fitness function and in five out of six metrics. These results also were found when a sensitivity analysis was performed varying the proportions of the sets of training, validation and test. The proposed approach reached consistent results when compared with the forecasting method

without correction, showing that it can be an interesting tool for correction of PM forecasters.

## Introduction

Air pollution has been the focus of public concern due to its health impact on the worldwide population, mainly in the big urban centers [1, 2]. The contamination of the earth's atmosphere by biological molecules, particulates and other harmful substances causes diseases and death in humans, who are also harmed by the damage that other living organisms, such as food crops, natural vegetation and herds of animals, suffer [2].

Particulate matter (PM) concentration has been a major concern among the air pollutants as according to epidemiological studies [3–12] and several diseases have been associated with this substance [1]. The Global Monitoring Report [1] points out PM as the major urban air pollutant affecting human health. The level of damage usually depends up on the duration of exposure as well as the kind and concentration of particles in the air [2, 4, 7, 9]. In general, the short-term effects [1, 13, 14], such as irritation in the eyes, nose and throat, headaches, nausea and allergic reactions are less serious [3]. However, in some cases, the exposure to short-term air pollution can cause upper respiratory infections such as bronchitis and pneumonia and aggravate the medical conditions of individuals with asthma and emphysema [3]. The long term effects [1, 8] may include chronic respiratory disease [3], lung cancer [7], cardiovascular diseases [5], such as ischemia-reperfusion injury and atherosclerosis, and even damage to the brain [15, 16], [15, 16], liver [16, 17], or kidneys [17, 18]. Continuous exposure to air pollution [8, 9] can severely affect the health and growth of children and may aggravate medical conditions in the elderly.

The monitoring of PM concentration is a relevant issue, as it allows the governments to create public policies to prevent and warn the population regarding high levels of PM. In this scenario, Artificial Neural Networks (ANN) have been widely used for the forecasting of PM concentration [19]. A non-exhaustive search in the literature points out three general ANN-based approaches for forecasting of PM concentration: the use of an ANN itself, hybrid systems that use search algorithms for the choice of ANN parameters, and hybrid systems that combine an ANN with another forecaster. Several studies belonging to each one of the aforementioned approaches are addressed in the following.

Four different ANN models: Recurrent Network Model (RNM), Change Point Detection Model with RNM, Sequential Network Construction Model and Self Organizing Feature Model were considered by Sharma *et al.* [20] for forecasting the concentration data of seven pollutants, among them  $PM_{2.5}$  and  $PM_{10}$ , in the California area. A Multilayer Perceptron (MLP) model, a Radial Basis Function (RBF) and a Square Multilayer Perceptron (SMLP) were addressed by Ordieres *et al.* [21] for forecasting  $PM_{2.5}$  concentration in the cities of El Paso (Texas) and Ciudad Juárez (Chihuahua). Other studies also used an MLP model: Kukkonen *et al.* [22] used an MLP model with homoscedastic and heteroscedastic Gaussian noise (ANN-HeG) to forecast the  $PM_{10}$  concentration in Helsinki, Caselli *et al.* [23] compared an MLP training via backpropagation, an RBF model and a multivariate regression model to forecast daily  $PM_{10}$  in Bari, Italy, and Gennaro *et al.* [24] used the MLP model developed in [23] with a specified set of input data to forecast  $PM_{10}$  concentration in two sites in the Western Mediterranean. Forecasting the maximum average concentration of  $PM_{10}$  per day in the city of Santiago, Chile, was carried out by Perez and Reyes [25] with the use of ANN. The majority of

studies concerning PM concentration forecasting are regarded with one-step ahead forecasting. Multi-step ahead forecasting is also found in the literature. For example, Kurt and Oktay [26] used an MLP model to forecast sulfur dioxide ( $\text{SO}_2$ ), carbon monoxide (CO) and  $\text{PM}_{10}$  concentration levels for 3 days ahead for a Besiktas district (Istanbul, Turkey) and Caselli *et al.* [23] applied an MLP model to forecast  $\text{PM}_{10}$  concentrations for 1, 2 and 3 days ahead.

Intelligent hybrid systems have also been proposed through combinations of ANNs with other techniques. These techniques are generally employed for the selection of input variables and the best ANN parameters, such as number of neurons in hidden and input layers, activation function, and training algorithm among others. Examples of techniques that have been combined with ANN include: Principal Component Analysis (PCA) [27], which was used with MLP to forecast  $\text{PM}_{10}$  concentration in Thessaloniki and for selection of input variables, followed by forecasting of  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  in Thessaloniki and Helsinki via MLP and linear regression (LR) [28]; Genetic Algorithm (GA), which was applied to select the inputs for ANN to forecast  $\text{PM}_{10}$  emission in 26 Europe countries [29] and a Multi-Objective Genetic Algorithm (MOGA), which was applied to reduce the number of potential meteorological input variables, with the integration of an MLP model with a numerical weather prediction model HIRLAM (High Resolution Limited Area Model) to forecast sequential hourly time series concentrations of  $\text{PM}_{2.5}$  in Helsinki [30]; Nearest Neighbor method combined with a MLP model in the scenario of  $\text{PM}_{10}$  concentration in Santiago [31]; Wavelets with combination of an ANN ensemble to forecast the daily average concentration of  $\text{PM}_{10}$  in Warsaw, Poland [32]. Hybrid systems were also proposed in other studies: Mishra *et al.* [33] proposed a Neuro-fuzzy model for forecasting  $\text{PM}_{2.5}$  during haze conditions in Delhi, India, and Qin *et al.* [34] proposed a hybrid model based on Cuckoo search (CS) and ANN training via backpropagation to forecast PM concentration levels in the four major cities of China (Beijing, Shanghai, Guangzhou and Lanzhou). Hybrid systems are also employed for multi step ahead forecasting. For example, Ul-Saufie *et al.* [35] combined MLR and ANN with PCA to forecast daily  $\text{PM}_{10}$  concentration (one, two and three days ahead) in Negeri Sembilan, Malaysia.

Hybrid systems have been proposed assuming that one forecaster can be insufficient to model a time series [36–38], making the combination of two or more models necessary. In this context, two assumptions are adopted: (i) a time series can contain patterns that are not purely linear or nonlinear [36]; (ii) a highly nonlinear time series cannot be modeled by an ANN alone [37, 38]. In the first assumption, the use of only linear or non-linear techniques may lead to inaccurate results, making it necessary to develop models considering both the linearity and the non-linearity involved in the time series [36]. In the second assumption, the use of one ANN can be insufficient due to problems of misspecification, generating a biased or inconsistent model [37, 38]. The misspecified ANN could be generated either during structure selection or during training, causing overfitting or underfitting problems.

A particular class of hybrid systems, which uses multiforecasters, explores the error series (residuals) in an attempt to improve the forecasting. Given the forecasting of a model, the error series (residuals) is obtained from the difference between actual time series and its forecasting. Usually, it is supposed that the error series is a white noise, i.e., it consists of independent and identically distributed random shocks that are unpredictable [37, 38]. However, due to misspecification of the forecaster, or time series behavior (linear and non-linear components), or due to disturbances present in the stochastic process after the specification of the forecaster, the assumption of the white noise can be violated. Thus, the temporal patterns that still remain in the error series can be captured and used to generate the forecasting of the residuals [37, 38].

In the scenario of air pollutant concentration forecasting, hybrid systems with more than one forecaster and using ANN have been proposed. Al-Alawi *et al.* [39] proposed a hybrid method using multiple regression combined with both PCA and ANN to forecast the

concentration of ozone in Kuwait's lower atmosphere. Ettouney *et al.* [40] also used a PCA combined with two ANNs in cascade to forecast ozone concentration in two locations in Kuwait. Westerlund *et al.* [41] proposed a forecast combination (FC), using two methods: linear regression and ANN to forecast air quality in Bogota. Combinations with Autoregressive Integrated Moving Average (ARIMA) model were also proposed: Sánchez *et al.* [42] combined an Elman neural network with ARIMA for forecasting of SO<sub>2</sub> concentration registered in a control station in the vicinity of a coal-fired power station in northern Spain and Díaz-Robles *et al.* [43] proposed a hybrid ARIMA and ANN model to forecast PM<sub>10</sub> concentration levels in urban areas of Temuco, Chile.

Hybrid models have not been applied only for forecasting parameters regarding air quality. In the literature, one can observe hybrid models in other public health interest scenarios such as: ARIMA combined with generalized regression neural network (GRNN) to forecast the incidence of tuberculosis [44], ARIMA combined with nonlinear auto-regressive neural network (NARNN) to forecast incidence of hand-foot-mouth disease (HFMD) [45] and ARIMA combined with NARNN to forecast the prevalence of schistosomiasis in humans [46].

Herein, we propose a hybrid approach for improving the performance of PM concentration forecasters using residual modeling. The hybrid approach consists of a recursive modeling, which at each iteration verifies if there are temporal patterns in the remaining residuals, thus aiming to generate the forecasting of the error series. Then, the forecasting of the error series is used in the next stage of the sequence of forecasters in order to improve the forecasting. This proposed approach differs from previous multiforecaster hybrid systems based on error series in the fact that those ones use only one residual series. The proposed hybrid system uses as many residual series as necessary to obtain a residual with white noise behavior. Another highlight of the proposed approach is its generality, that is, the forecasting system is sufficiently versatile to use a set of different forecasters, e.g. an ANN followed by an ARIMA model, or an ARIMA model followed by an ANN, or a sequence of three different forecasters. Simulation results are presented for four time series: PM<sub>2.5</sub> and PM<sub>10</sub> daily concentration levels in Kallio and Vallila stations in Helsinki, Finland shown in Fig 1. More specifically, in the present study, error series are used to improve the forecasting of a recent hybrid system composed of genetic algorithm (GA) and ANN presented in the literature [19]. For residual modeling, two methods are considered in this study: an MLP neural network and our own hybrid system used in PM concentration series forecasting. The results are evaluated using a set of six well-known metrics and show that the proposed approach is capable of improving the performance of the PM forecaster considered for all cases.

## The Methodology for the Correction of Pollutant Forecasters

Fig 2 shows the architecture of the proposed approach, which is composed of two modules: (i) Forecasting and (ii) Correction. Given a univariate time series ( $\mathbf{x}$ ), the aim of the first module is to train a forecasting model ( $M_0$ ). Then, the error series  $\mathbf{e}_0$  is calculated as the difference between the time series and the output of the model,  $\mathbf{e}_0 = \mathbf{x} - M_0(\mathbf{x})$ .

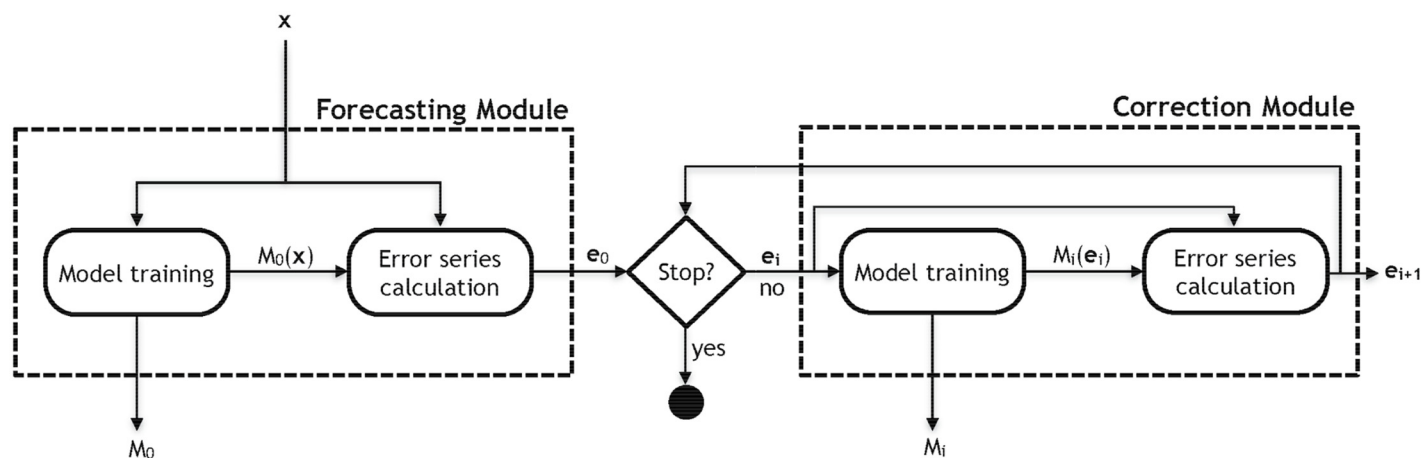
The Correction Module is only executed if the residual series ( $\mathbf{e}_0$ ) is not a white noise. In other words, a test (Fig 2—"Stop?") is performed to verify if there is enough information in the error series to improve the precision of the forecasting.

Formally, a white noise [48] is a sequence of independent and identically distributed random values with zero mean and constant variance. In general, the literature of time series supposes that the residuals generated by the forecasters have a white noise behavior, and the module (i) Forecasting of the proposed approach (Fig 2) shows this classical approach. However, we expect that the error series contains useful information that was not captured by the



**Fig 1.** Locations of the Kallio and Vallila air quality stations in Helsinki. The Vallila station is located in urban traffic (UT) environment and the Kallio station is located in urban background (UB) environment(adapted from [47]).

doi:10.1371/journal.pone.0138507.g001

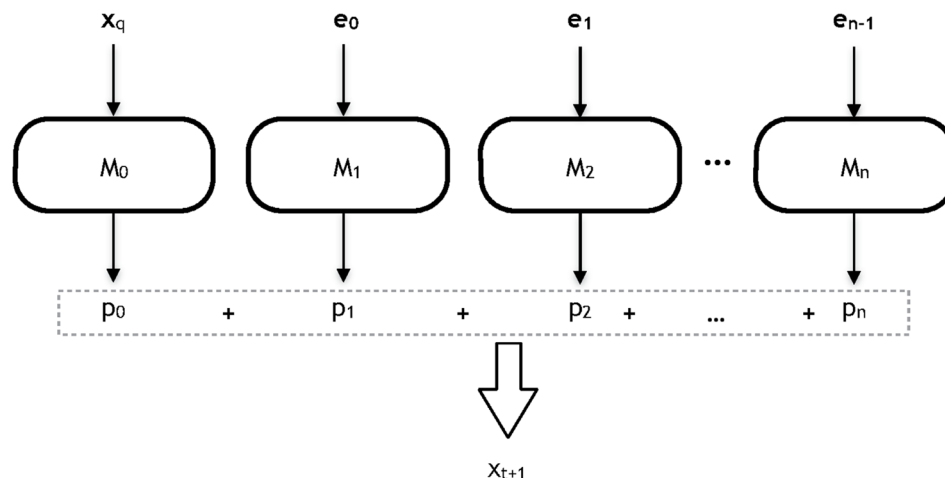


**Fig 2.** Training architecture of the proposed system.

doi:10.1371/journal.pone.0138507.g002

forecaster  $M_0$  (module (i) Forecasting). In this context, there are tests [48] to check the hypothesis that the residuals are independent and identically distributed random variables. Some of these tests are: the *portmanteau* test, spectral analysis, turning point test and autocorrelation function (ACF) [48]. The ACF test is adopted in this study and consists of the cross-correlation of a series with itself at different times, as a function of two time lags. So, ACF measures the correlation between the value of a series in the time  $t$  and  $t + k$  according to Eq (1):

$$\rho_k = \text{Corr}(y_t, y_{t+k}) = \frac{\gamma_h}{\gamma_0}, \quad (1)$$



**Fig 3. Test architecture of the proposed system.**

doi:10.1371/journal.pone.0138507.g003

where  $\text{Corr}$  is the correlation,  $\gamma_h = \text{Cov}(y_t, y_{t+h})$ , where  $\text{Cov}$  is the covariance, and  $\gamma_0$  is the sample variance of the time series.  $\rho_k$  lies in the range  $[-1, 1]$ , where 1 and  $-1$  indicate perfect correlation and perfect anti-correlation, respectively. If the values of  $\rho_k$ , where  $k = 1, \dots, n$ , are in the range  $[-2s, 2s]$  (where  $s$  is the standard deviation of data sample), then there is no correlation; otherwise, there is correlation in the series data.

Based on the assumption that it is possible to extract valuable information from the error series ( $e_0$ ) to improve the forecasting of the whole system, the Correction Module trains a model ( $M_1$ ) that aims to forecast the error of the model ( $M_0$ ). Once again, if the residual series ( $e_1$ ), where  $e_1 = e_0 - M_1(e_0)$ , of the model  $M_1$  is not a white noise, this procedure is repeated until the stop criterion is reached. So, the output of the Correction Module is given by two sets: models  $\{M_0, M_1, \dots, M_n\}$  and error series  $\{e_0, e_1, \dots, e_{n-1}\}$ .

After training, the models are used to forecast unseen patterns as shown in Fig 3. So, given a time series  $x_q = [x_1, x_2, \dots, x_t]$ , we want to forecast  $x_{t+1}$ . The predicted value of  $x_{t+1}$  is given by Eq (2).

$$\hat{x}_{t+1} = M_0(x_q) + M_1(e_0) + M_2(e_1) + \dots + M_n(e_{n-1}), \quad (2)$$

where ideally it is expected that the contribution of the  $M_{i-1}$  model is greater than that of the  $M_i$  model to the final forecasting  $\hat{x}_{t+1}$  of the proposed approach. This should occur assuming that if the remaining error series is not a white noise, at each iteration  $i$  of the proposed approach a model  $M_i$  captures temporal patterns. In this context, it is expected that the contributions of the  $M_i$  models decrease at each iteration, remaining only a signal uncorrelated in time; in other words, a white noise.

## Simulation and Results

Four time series were used for the evaluation of the proposed approach. The series correspond to daily mean concentration of particulate matter ( $\text{PM}_{2.5}$  and  $\text{PM}_{10}$ ) from Helsinki. The data set is composed of values measured between the years 2001 to 2003 from stations Kallio and Vallila [28]. Despite the two stations be located in Helsinki, they have different characteristics.



**Table 1. Metrics for forecasting assessment, where  $N$  is the size of the series,  $target_j$  is the real value at period  $j$ ,  $output_j$  is the forecasting at period  $j$  and  $target$  is the mean of the series.** In the column  $\theta$ :  $\uparrow$  means the higher the value of metric the better is the forecasting, and  $\downarrow$  means the lower the value of metric the better is the forecasting.

Metric	Equation	Range	$\theta$
Mean Squared Error (MSE)	$MSE = \frac{1}{N} \sum_{j=1}^N (target_j - output_j)^2$	$[0, +\infty)$	$\downarrow$
Mean Absolute Percentage Error (MAPE)	$MAPE = \frac{100}{N} \sum_{j=1}^N \left  \frac{target_j - output_j}{target_j} \right $	$[0, +\infty)$	$\downarrow$
U of Theil Statistics (U)	$U = \frac{\sum_{j=1}^N (target_j - output_j)^2}{\sum_{j=1}^N (output_j - output_{j+1})^2}$	$[0, +\infty)$	$\downarrow$
Average Relative Variance (ARV)	$ARV = \frac{\sum_{j=1}^N (output_j - target_j)^2}{\sum_{j=1}^N (output_j - target)^2}$	$[0, +\infty)$	$\downarrow$
Prediction of Change in Direction (POCID)	$POCID = 100 \frac{\sum_{j=1}^N D_j}{N}, D_j = \begin{cases} 1, & \text{if } (target_j - target_{j-1})(output_j - output_{j-1}) > 0. \\ 0, & \text{otherwise.} \end{cases}$	$[0, 100]$	$\uparrow$
Index of Agreement (IA)	$IA = 1 - \frac{\sum_{j=1}^N  output_j - target_j ^2}{\sum_{j=1}^N ( output_j - target  +  target_j - target )^2}$	$[0, 1]$	$\uparrow$

doi:10.1371/journal.pone.0138507.t001

The Kallio station is located in an urban background and the Vallila station is situated in an area more exposed to pollution from local traffic.

The data set is normalized to lie within the interval  $[0, 1]$  and divided into three sets: 80% for training, 10% for validation and 10% for test. For each time series, ten simulations with the proposed approach were performed and the best model was selected based on the performance (fitness value) in the validation data set. The results for all models correspond to the one step ahead forecasting from the test set. Table 1 shows the six metrics [28, 49, 50] used to evaluate the performance of the approach employed in this paper: Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), U of Theil Statistics (U), Average Relative Variance (ARV), Prediction of Change in Direction (POCID) and Index of Agreement (IA). For MSE, MAPE, U and ARV, the lower the value of those measures, the better is the forecasting of the model. U and ARV measures are used to compare the model performances with the forecasting of a Random Walk model and that of the mean of the series, respectively. If U and ARV values are equal to 1, the forecasting of the forecaster is equivalent to the random walk model ( $U = 1$ ), or to the mean of the time series ( $ARV = 1$ ), respectively. However, if U and ARV values are less than or greater than 1, the forecasting of the model is better or worse than the performance of the random walk model or mean, respectively. In case of POCID and IA, the higher the value the better is the performance of the model. The POCID can have values in the range  $[0, 100]$  and IA in the range  $[0, 1]$ . A ratio shown in Eq (3) was developed to compare the performances of the current correction with the model added and the previous correction, or the model without correction:

$$ratio = \frac{EvaluationMeasure_{correctedModel}}{EvaluationMeasure_{uncorrectedModel}}, \quad (3)$$

where  $EvaluationMeasure_{correctedModel}$  is the forecaster's performance reached after the current correction ( $n$ ) is added and  $EvaluationMeasure_{uncorrectedModel}$  is the value of forecaster's performance reached with the previous correction ( $n - 1$ , model without the current correction). For POCID, IA and fitness values, if the obtained value in Eq (3) is greater than 1, the correction

improves the forecasting. If the value is equal to or less than 1, the correction either does not add information, or worsens the previous forecasting, respectively.

The proposed approach is used to correct a hybrid system (model of the Forecasting Module—Fig 2) composed by GA and ANN of type MLP [19]. This hybrid system consists of two stages: optimization of the ANN parameters performed by GA and phase adjustment. In the first stage, the GA searches the best configuration of the following parameters of MLP: number of input neurons (relevant time lags), number of nodes in the hidden layer and the training algorithm among four candidates (Levenberg-Marquardt, Scaled Conjugated Gradient, Resilient Backpropagation (RPROP) and One Step Secant Conjugate Gradient).

The search process performed by GA is guided by fitness function, defined in Eq (4). Silva *et al.* [51] reported that the definition of an adequate fitness function is a non-trivial task. So, the used fitness function aims to aggregate different measures and the higher (as closer to 100) its value the more accurate is the forecasting model.

$$Fitness = \frac{POCID}{1 + MSE + MAPE + U + ARV}. \quad (4)$$

In the second stage, if necessary, a procedure of phase adjustment is performed aiming to minimize the difference between the forecasting performed by ANN and the PM concentration series. The hybrid system [19] was chosen because it reached superior results in terms of accuracy, when compared with literature [19, 50, 52]. Thus, if the proposed approach is capable of improving a forecaster with high accuracy, it is expected that the approach also improves the forecaster with lower accuracy.

Two models are adopted for correction (Correction Module—Fig 2): the own hybrid system and an MLP model. In both cases, the models are trained recursively until a stopping condition is reached. For ANN, the input and hidden nodes are established using ACF and trial and error method in the interval [1, 20], respectively.

The parameters used in this work are described below:

- The stopping conditions for recursive approach are: (i) the error series has a white noise behavior, or; (ii) an increase ( $\geq 5\%$ ) in the MAPE value with the addition of one correction;
- The parameters of the hybrid system are set up to: (i) Initial acceptable fitness value (1% of error); (ii) Initial maximum number of time lags (10); (iii) Maximum number of hidden units (20); (iv) Maximum number of iterations (10);
- The parameters of the GA used by the hybrid system are set up to: (i) Mutation probability of 10%; (ii) Population size of 10 individuals; (iii) Maximum number of generations equals 1000; (iv) Minimum fitness progress of  $10^{-4}$ . Three stopping conditions for the ANN training algorithms are used: (i) Maximum number of iterations equals 1000; (ii) cross-validation process with a generation loss of 5%; (iii) Progress training of  $10^{-6}$ .
- The parameters of the ANN of type multi-layer perceptron (MLP) used in the correction phase are: (i) training algorithm is Levenberg-Marquardt; (ii) Maximum number of iterations. Three stopping conditions for Levenberg-Marquardt algorithm are used: (i) Maximum number of iterations equals 1000; (ii) cross-validation process with generation loss of 5%; (iii) Progress training of  $10^{-6}$ .

## Correction via Artificial Neural Network

Table 2 shows the results in terms of the evaluation measures and fitness reached by the hybrid system without correction and proposed methodology using ANN for PM<sub>2.5</sub> and PM<sub>10</sub> time



**Table 2. Results with correction via ANN for all series.** The best result for each metric and fitness is highlighted in bold.

Station	Series	Models	MSE	POCID	U	MAPE	ARV	IA	Fitness
Kallio	PM <sub>10</sub>	HS	6.00E-04	<b>97.16</b>	0.0969	35.93	0.3077	0.950	2.60
		HS+C <sub>1</sub>	<b>5.79E-04</b>	87.04	<b>0.0847</b>	<b>31.40</b>	<b>0.1961</b>	<b>0.963</b>	<b>2.66</b>
	PM <sub>2.5</sub>	HS	3.00E-04	<b>97.19</b>	0.0596	27.41	0.1349	0.970	3.40
		HS+C <sub>1</sub>	<b>2.95E-04</b>	87.74	<b>0.0510</b>	<b>19.18</b>	<b>0.0888</b>	<b>0.981</b>	<b>4.32</b>
Vallila	PM <sub>10</sub>	HS	3.00E-04	<b>97.82</b>	<b>0.0727</b>	29.66	0.1910	0.960	3.16
		HS+C <sub>1</sub>	3.00E-04	88.57	0.0785	<b>25.06</b>	<b>0.1864</b>	<b>0.964</b>	<b>3.36</b>
	PM <sub>2.5</sub>	HS	<b>1.00E-04</b>	<b>98.11</b>	<b>0.0313</b>	21.66	<b>0.0630</b>	0.980	4.31
		HS+C <sub>1</sub>	2.01E-04	94.57	0.0450	<b>15.27</b>	0.0850	<b>0.990</b>	<b>5.76</b>

doi:10.1371/journal.pone.0138507.t002

series from Kallio and Vallila stations. The hybrid system uncorrected (without correction) is named HS. The version of the hybrid system corrected, denoted by HS+C<sub>n</sub>, corresponds to the proposed methodology, with *n* terms of corrections generated by an ANN model. For all time series analyzed here there was the need for just one correction term, suggesting that the HS was not capable of capturing all information contained in the time series. Therefore, a comparison is done between HS and HS+C<sub>1</sub>. After the first correction (HS+C<sub>1</sub>), the same stopping criterion was reached for all cases, which was the model residual for a white noise.

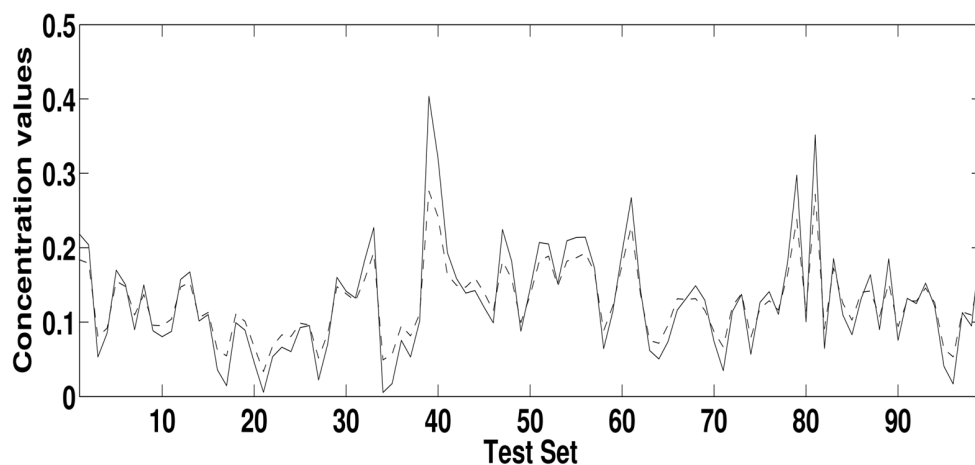
Table 2 shows that the proposed approach improved 15 out of 24 evaluation measures, achieving the best fitness for all time series. The best result for each metric and fitness is highlighted in bold. For pollutant concentration time series from Kallio station, the proposed approach improved almost all evaluation metrics, except the POCID measure. The decrease in MSE and MAPE shows that the correction was able to improve the forecasting of the uncorrected model. For Vallila station, the improvement occurred in MAPE, ARV and IA for PM<sub>10</sub> concentration time series and in MAPE and IA for PM<sub>2.5</sub> concentration time series. These results show that for all studied cases, HS+C<sub>1</sub> reached a better performance than HS with respect to the fitness function, that is the objective for the proposed method, and, as a consequence, it was also observed an improvement in the performance for other metrics.

Table 3 shows the results reached with the model HS+C<sub>1</sub> and the uncorrected model (HS) according to the ratio defined in the Eq (3). The values greater than 1 for POCID, IA and fitness show that the addition of the correction improved the forecasting of the HS. For MSE, U, MAPE and ARV the logic of the Eq (3) is inverted; values smaller than 1 show forecasting enhancement. For concentration series from Kallio station, greater improvements occurred in U, MAPE and ARV metrics. For concentration series from Vallila station, greater improvement occurred in MAPE metric. In general, the addition of corrections made the forecasting closer to the actual time series. This improvement can be seen by comparing Figs 4(a), 5(a), 6

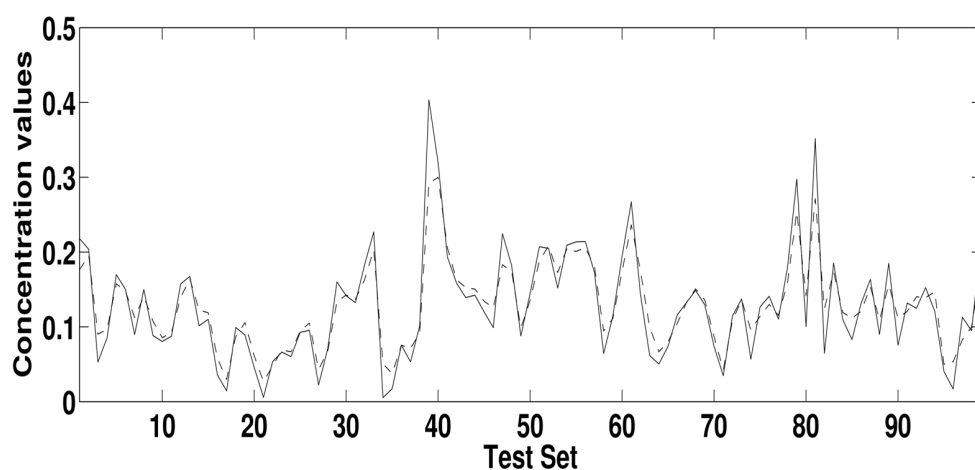
**Table 3. Comparison between subsequent corrections with ANN and regarding the model without correction measured for Eq (3) for all series.**

Series	Models	MSE	POCID	U	MAPE	ARV	IA	Fitness
Kallio—PM <sub>10</sub>	C <sub>1</sub> – C <sub>0</sub>	0.964	0.896	0.874	0.874	0.637	1.013	1.023
Kallio—PM <sub>2.5</sub>	C <sub>1</sub> – C <sub>0</sub>	0.983	0.903	0.856	0.700	0.658	1.011	1.271
Vallila—PM <sub>10</sub>	C <sub>1</sub> – C <sub>0</sub>	1.000	0.905	1.080	0.845	0.976	1.004	1.064
Vallila -PM <sub>2.5</sub>	C <sub>1</sub> – C <sub>0</sub>	2.012	0.964	1.438	0.705	1.349	1.010	1.337

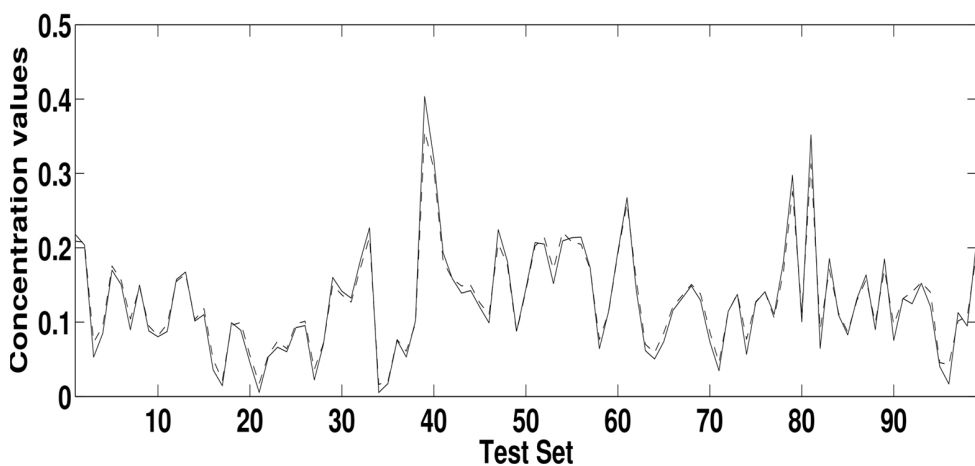
doi:10.1371/journal.pone.0138507.t003



(a) Forecasting without correction.



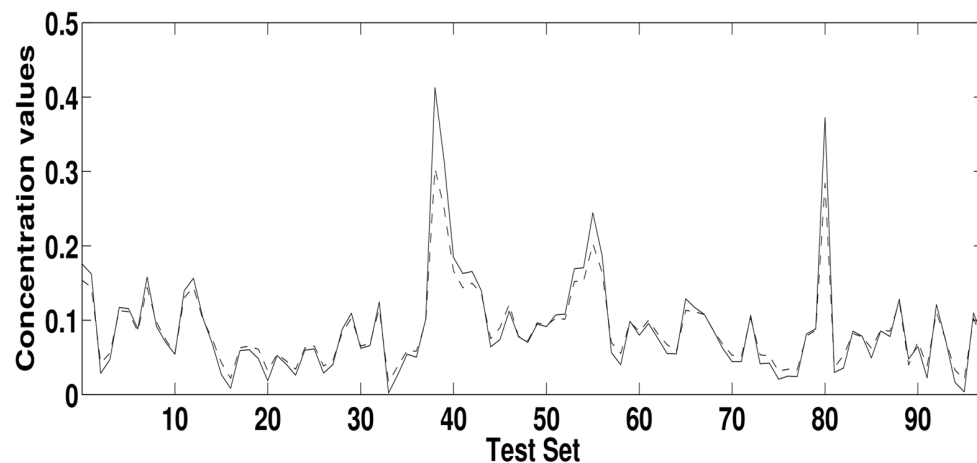
(b) Forecasting with one ANN correction ( $C_1$ ).



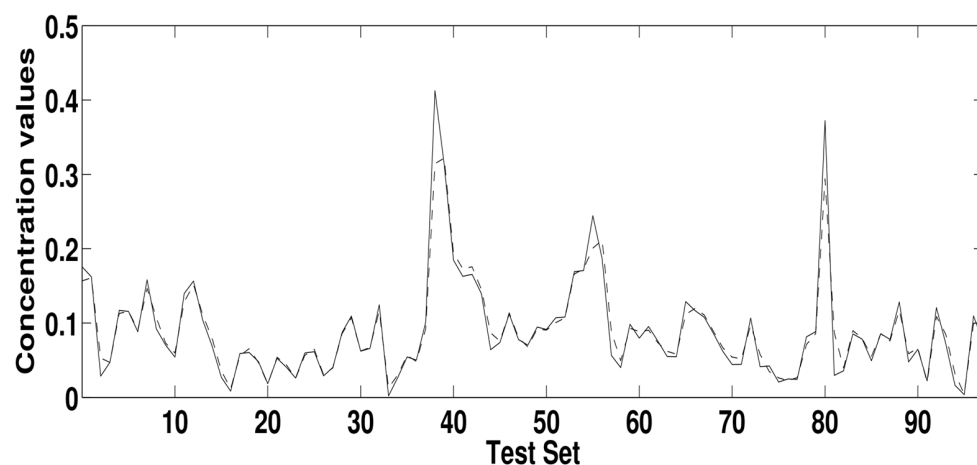
(c) Forecasting with one HS correction ( $C_1$ ).

**Fig 4. Forecasting for the  $PM_{10}$  concentration time series for Kallio Station (solid lines—actual values; dashed lines—predicted values).**

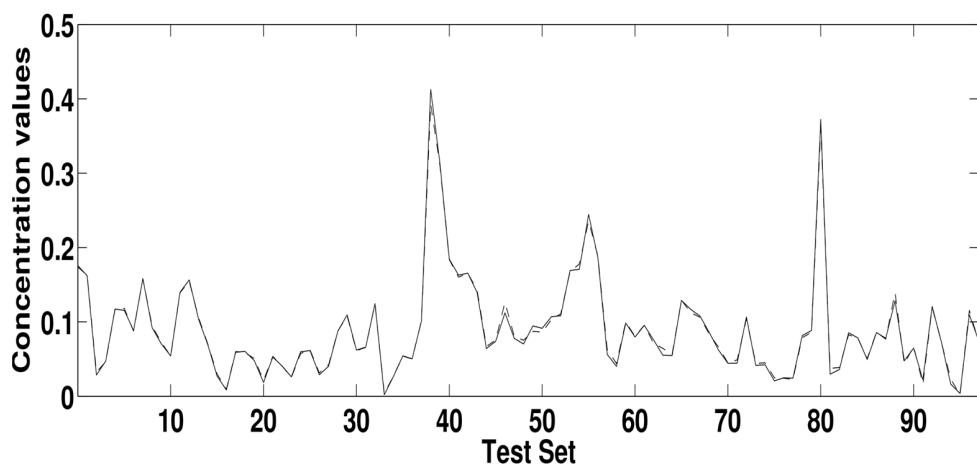
doi:10.1371/journal.pone.0138507.g004



(a) Forecasting without correction.



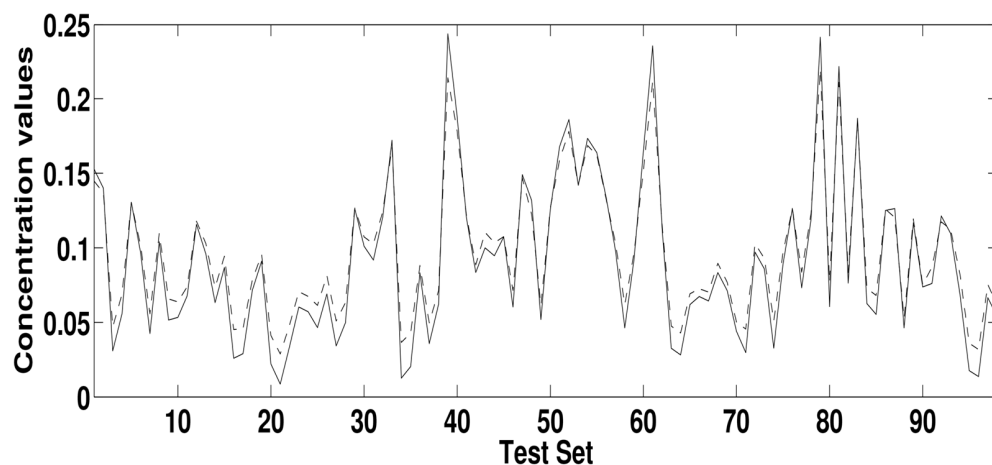
(b) Forecasting with one ANN correction ( $C_1$ ).



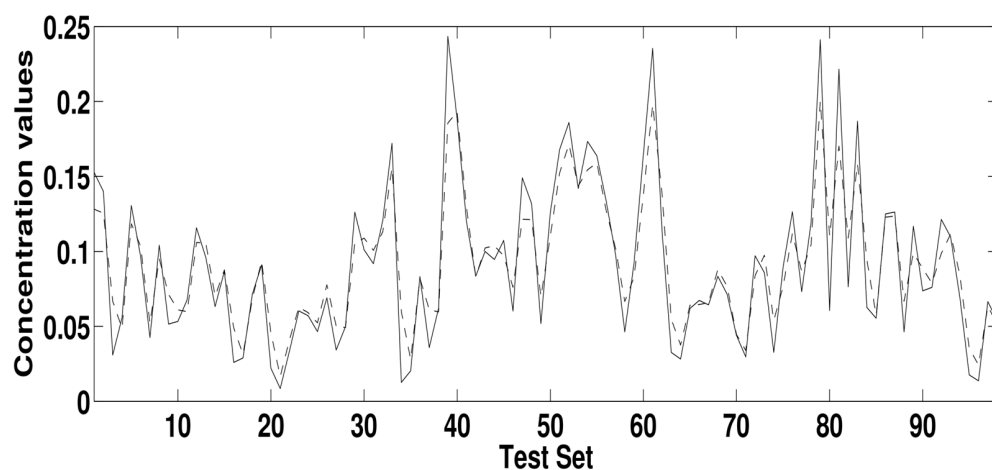
(c) Forecasting with two HS corrections ( $C_1$  and  $C_2$ ).

**Fig 5. Forecasting for the  $PM_{2.5}$  concentration time series for Kallio Station (solid lines—actual values; dashed lines—predicted values).**

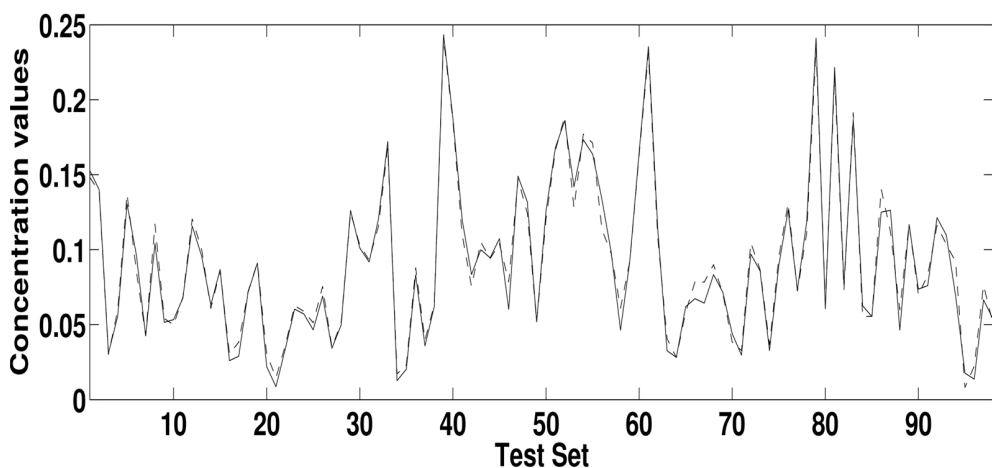
doi:10.1371/journal.pone.0138507.g005



(a) Forecasting without correction.



(b) Forecasting with one ANN correction ( $C_1$ ).



(c) Forecasting with two HS corrections ( $C_1$  and  $C_2$ ).

**Fig 6. Forecasting for the  $PM_{10}$  concentration time series for Vallila Station (solid lines—actual values; dashed lines—predicted values).**

doi:10.1371/journal.pone.0138507.g006

(a) and 7(a) with Figs 4(b), 5(b), 6(b) and 7(b), which are the forecasting for PM<sub>2.5</sub> and PM<sub>10</sub> concentration time series from Kallio and Vallila stations without and with correction, respectively.

## Correction via Hybrid System

Table 4 shows the performance of the model uncorrected (HS) and the proposed methodology using HS+C<sub>n</sub> in terms of the evaluation measures and fitness for PM<sub>2.5</sub> and PM<sub>10</sub> concentration time series from Kallio and Vallila stations. In the proposed methodology using correction via HS, the number of corrections *n* ranged from 1 to 5. The stopping condition reached was two increases in MAPE value in followed corrections, for all cases. This result suggests that the capacity of correction depends on the accuracy of the forecasting method used in this step.

Table 4 shows that the proposed methodology using correction via HS was able to correct the forecasting of the HS more than once in 3 of 4 time series.

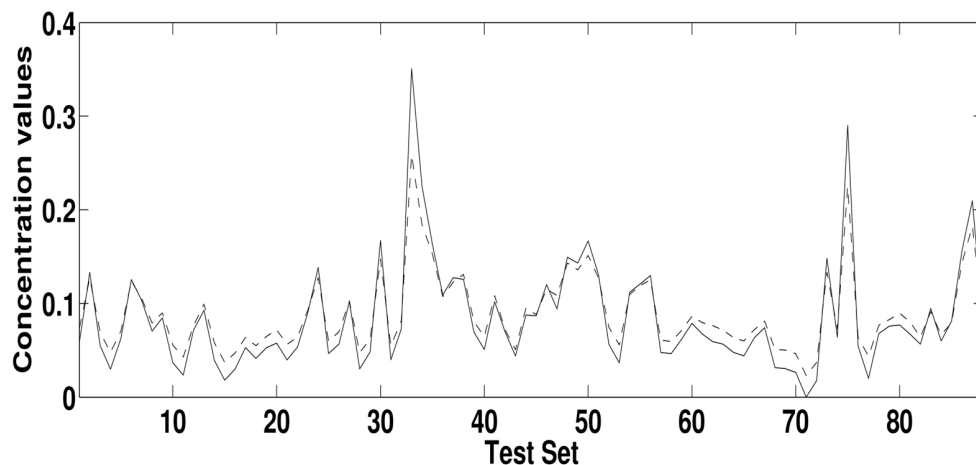
Table 4 also shows that the proposed approach improved 21 out of 24 evaluation measures, achieving the best fitness functions for all time series. The best performance for each metric and fitness is highlighted in bold. For pollutant concentration time series from Kallio station, repeating the proposed approach improved almost all evaluation metrics, except the POCID measure. For Vallila station, the improvement also occurred in almost all metrics, except in the POCID measure for PM<sub>10</sub> concentration time series. For all series, the best model considered was the last model found. For Kallio station, the best models were HS+C<sub>1</sub> and HS+C<sub>1</sub>+C<sub>2</sub> for PM<sub>10</sub> and PM<sub>2.5</sub>, respectively. For Vallila station, the best models were HS+C<sub>1</sub>+C<sub>2</sub> and HS+C<sub>1</sub>+...+C<sub>5</sub> for PM<sub>10</sub> and PM<sub>2.5</sub>, respectively. Table 4 shows that if the values of POCID and fitness are overlooked, the gain in the evaluation measures tends to be smaller at each correction, i.e., at each correction (C<sub>i</sub>), the HS aggregated less information than past correction (C<sub>i-1</sub>) in the final forecasting. Considering, for instance, Fig 8, it is observed that the MSE difference between HS+...+C<sub>n</sub> and HS (without correction) decreases as the number of corrections increases. The same behavior can be observed in Table 4 in terms of MSE, U, MAPE, ARV and IA, when more than one correction is performed.

Table 5 shows the ratio, according to Eq (3), used to compare (C<sub>n</sub> - C<sub>n-1</sub>), the model with *n* corrections with the previous model with *n* - 1 correction terms and the best model with model uncorrected (C<sub>n</sub> - C<sub>0</sub>). As in Table 4, in the Table 5, the values greater than 1 for POCID and IA, and values less than 1 for MSE, U, MAPE and ARV show that the addition of the correction terms improved the forecasting of the HS. The values of fitness greater than 1 at each iteration show that the use of HS in the correction improved the initial forecasting for all time series. The total improvement can be seen in the last lines (C<sub>n</sub> - C<sub>0</sub>) for each time series in Table 5 and by comparing the Figs 4(a), 5(a), 6(a) and 7(a) with Figs 4(c), 5(c), 6(c) and 7(c), which are the forecasting for PM<sub>2.5</sub> and PM<sub>10</sub> concentration time series from Kallio and Vallila stations without and with correction, respectively.

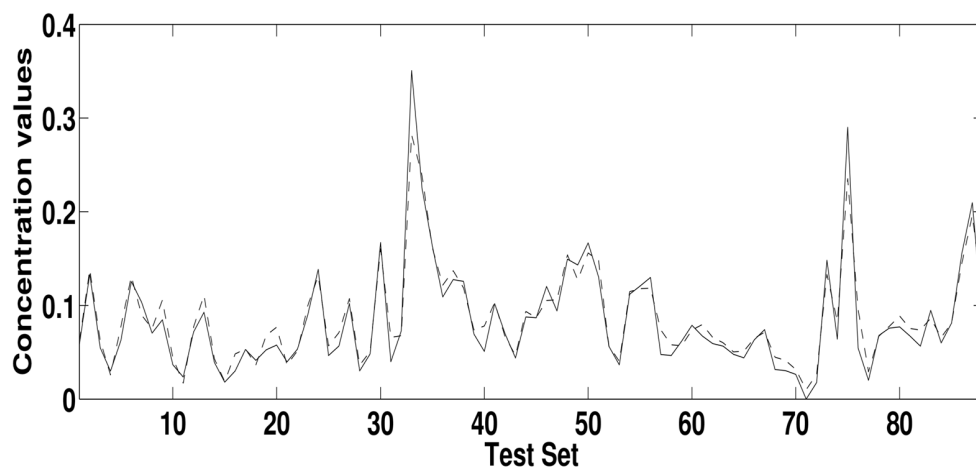
Comparing the performances of the corrections via ANN and HS from Tables 3 and 5, it can be seen that for all series, the addition of one correction (C<sub>1</sub> - C<sub>0</sub>) of HS led to better results than ANN. This fact shows that the search for HS through the combination of exploration using gradient descent algorithms and that using GA, overcomes the use of only one learning algorithm for ANN. Consequently, in cases with correction via HS, where there were more corrections, the proposed methodology improved the performance reached by HS+C<sub>1</sub>.

## Discussion

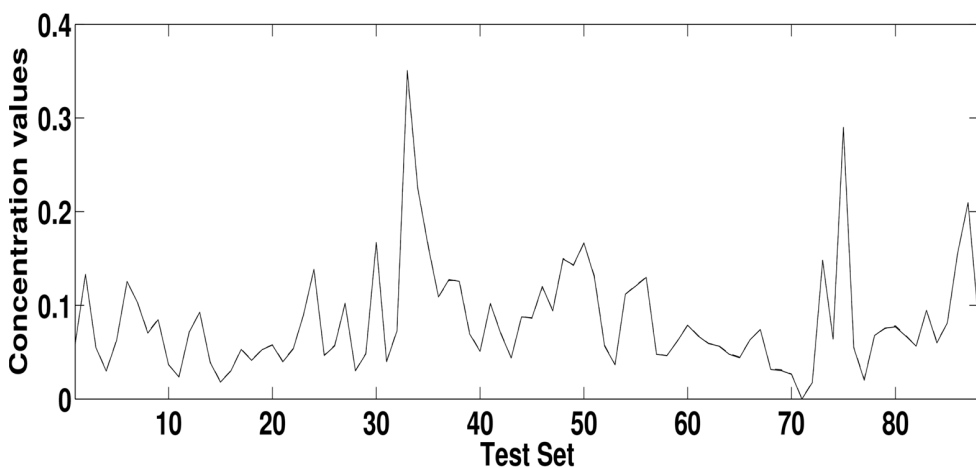
Table 6 shows for all series, both stations, all the metrics and fitness, the best performance reached via ANN correction and via HS according to Tables 2 and 4, respectively. It is possible



(a) Forecasting without correction.



(b) Forecasting with one ANN correction ( $C_1$ ).



(c) Forecasting with five HS corrections ( $C_1$  to  $C_5$ ).

**Fig 7. Forecasting for the  $PM_{2.5}$  concentration time series for Vallila Station (solid lines—actual values; dashed lines—predicted values).**

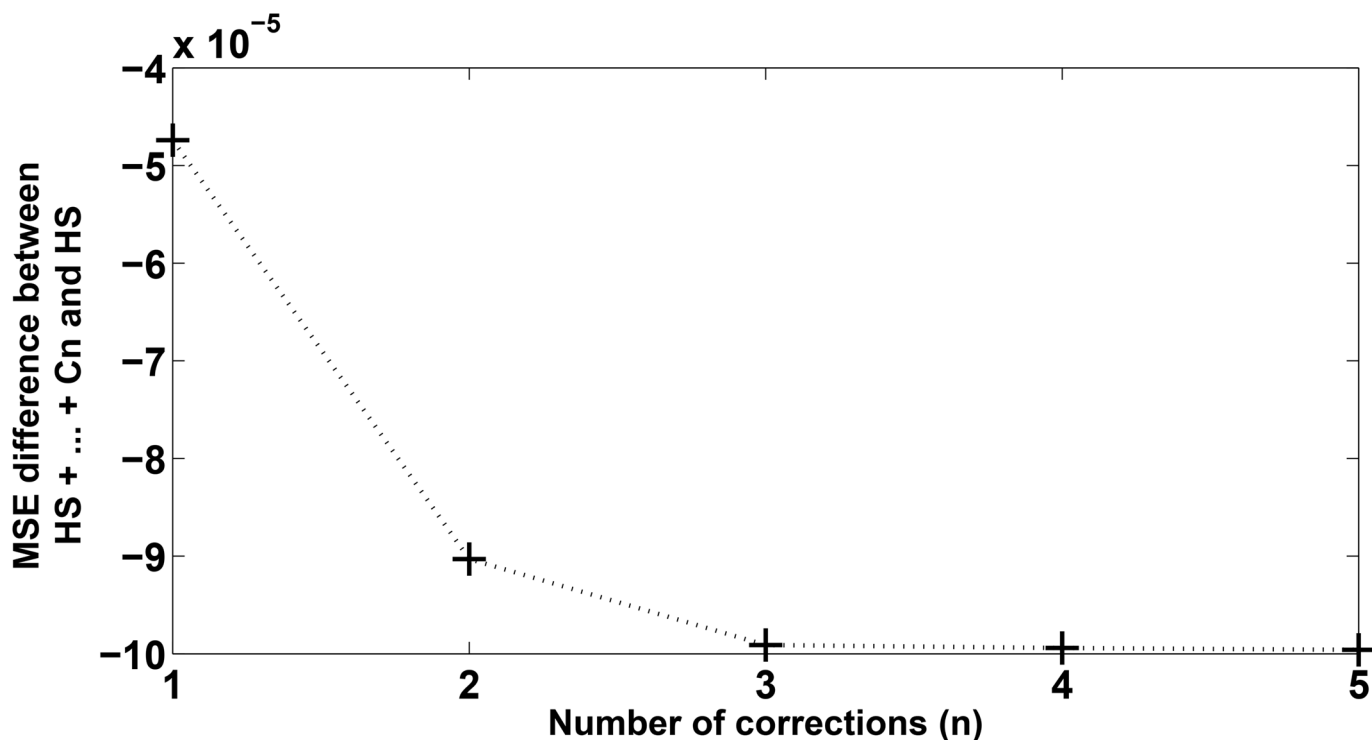
doi:10.1371/journal.pone.0138507.g007



**Table 4. Results with correction via HS for all series.** The best result for each metric and fitness is highlighted in bold.

Station	Series	Models	MSE	POCID	U	MAPE	ARV	IA	Fitness
Kallio	PM <sub>10</sub>	HS	6.00E-04	<b>97.16</b>	0.0969	35.93	0.3077	0.950	2.60
		HS+C <sub>1</sub>	<b>1.34E-04</b>	94.29	<b>0.0193</b>	<b>14.39</b>	<b>0.0339</b>	<b>0.993</b>	<b>6.11</b>
	PM <sub>2.5</sub>	HS	3.00E-04	<b>97.19</b>	0.0596	27.41	0.1349	0.970	3.40
		HS+C <sub>1</sub>	5.36E-05	97.09	0.0097	8.87	0.0135	0.997	9.82
		HS+C <sub>1</sub> +C <sub>2</sub>	<b>1.74E-05</b>	97.06	<b>0.0032</b>	<b>5.76</b>	<b>0.0041</b>	<b>0.999</b>	<b>14.34</b>
Vallila	PM <sub>10</sub>	HS	3.00E-04	<b>97.82</b>	0.0727	29.66	0.1910	0.960	3.16
		HS+C <sub>1</sub>	9.37E-05	95.19	0.0240	14.99	0.0408	0.991	5.93
		HS+C <sub>1</sub> +C <sub>2</sub>	<b>4.72E-05</b>	96.12	<b>0.0121</b>	<b>9.19</b>	<b>0.0185</b>	<b>0.996</b>	<b>9.40</b>
	PM <sub>2.5</sub>	HS	1.00E-04	98.11	0.0313	21.66	0.0630	0.980	4.31
		HS+C <sub>1</sub>	5.26E-05	97.83	0.0122	12.10	0.0210	0.995	7.45
		HS+C <sub>1</sub> +C <sub>2</sub>	9.74E-06	<b>98.90</b>	0.0022	4.98	0.0034	0.999	16.54
		HS+C <sub>1</sub> +...+C <sub>3</sub>	9.08E-07	98.89	2.09E-04	1.60	2.85E-04	0.999	38.06
		HS+C <sub>1</sub> +...+C <sub>4</sub>	6.26E-07	98.88	1.42E-04	1.27	1.96E-04	0.999	43.51
		HS+C <sub>1</sub> +...+C <sub>5</sub>	<b>4.23E-07</b>	98.86	<b>9.67E-05</b>	<b>0.85</b>	<b>1.33E-04</b>	<b>0.999</b>	<b>53.31</b>

doi:10.1371/journal.pone.0138507.t004



**Fig 8. MSE difference between HS with  $n$  corrections and HS (without correction) for PM<sub>2.5</sub> concentration series for Vallila station.**

doi:10.1371/journal.pone.0138507.g008

**Table 5. Comparison between subsequent corrections via HS and regarding the model without correction measured for Eq (3) for all series.**

Series	Models	MSE	POCID	U	MAPE	ARV	IA	Fitness
Kallio—PM <sub>10</sub>	C <sub>1</sub> – C <sub>0</sub>	0.223	0.970	0.199	0.400	0.110	1.045	2.346
Kallio—PM <sub>2.5</sub>	C <sub>1</sub> – C <sub>0</sub>	0.179	0.999	0.162	0.324	0.100	1.028	2.889
	C <sub>2</sub> – C <sub>1</sub>	0.324	1.000	0.332	0.650	0.301	1.002	1.461
	C <sub>2</sub> – C <sub>0</sub>	0.058	0.999	0.054	0.210	0.030	1.030	4.221
Vallila—PM <sub>10</sub>	C <sub>1</sub> – C <sub>0</sub>	0.312	0.973	0.329	0.505	0.213	1.032	1.874
	C <sub>2</sub> – C <sub>1</sub>	0.504	1.010	0.506	0.613	0.454	1.005	1.586
	C <sub>2</sub> – C <sub>0</sub>	0.157	0.983	0.167	0.310	0.097	1.037	2.972
Vallila -PM <sub>2.5</sub>	C <sub>1</sub> – C <sub>0</sub>	0.526	0.997	0.390	0.558	0.333	1.016	1.728
	C <sub>2</sub> – C <sub>1</sub>	0.185	1.011	0.184	0.411	0.160	1.004	2.220
	C <sub>3</sub> – C <sub>2</sub>	0.093	1.000	0.093	0.321	0.085	1.001	2.302
	C <sub>4</sub> – C <sub>3</sub>	0.690	1.000	0.682	0.796	0.689	1.000	1.143
	C <sub>5</sub> – C <sub>4</sub>	0.676	1.000	0.679	0.671	0.678	1.000	1.225
	C <sub>5</sub> – C <sub>0</sub>	0.004	1.008	0.003	0.039	0.002	1.020	12.365

doi:10.1371/journal.pone.0138507.t005

**Table 6. Comparison between the corrections via ANN and via HS for all series.** The best result for each metric and fitness is highlighted in bold.

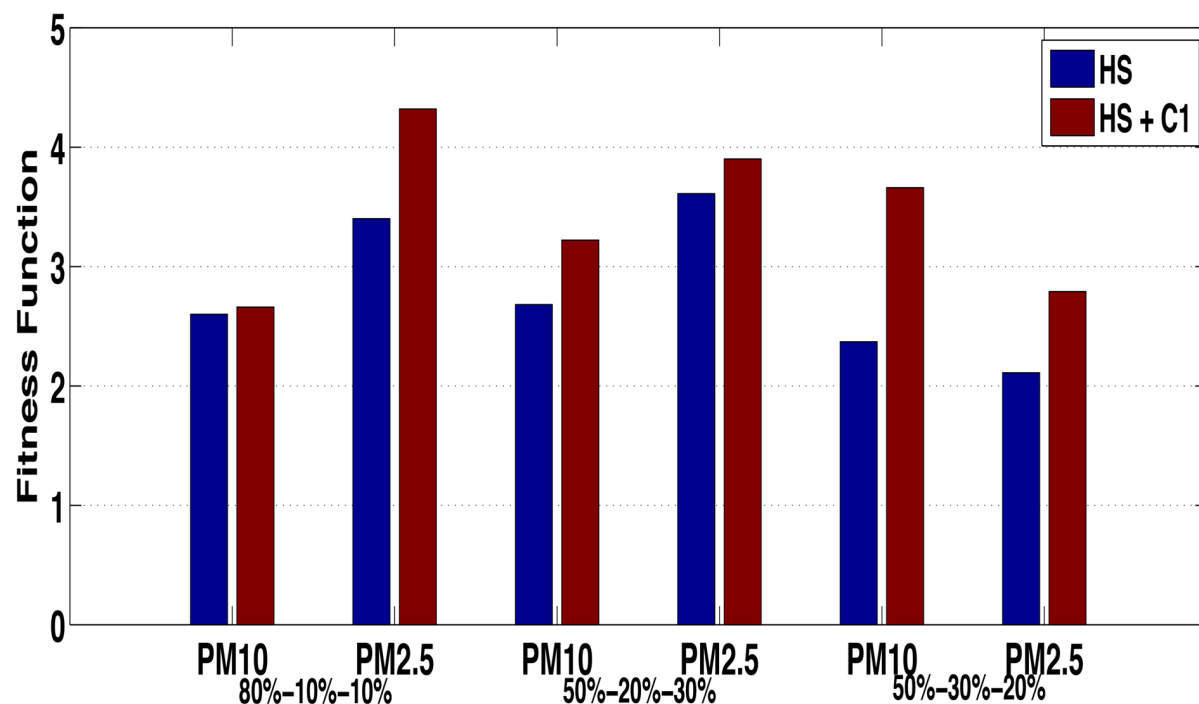
Station	Series	Correction	MSE	POCID	U	MAPE	ARV	IA	Fitness
Kallio	PM <sub>10</sub>	ANN (C <sub>1</sub> )	5.79E-04	87.04	0.0847	31.40	0.1961	0.963	2.66
		HS (C <sub>1</sub> )	<b>1.34E-04</b>	<b>94.29</b>	<b>0.0193</b>	<b>14.39</b>	<b>0.0339</b>	<b>0.993</b>	<b>5.74</b>
	PM <sub>2.5</sub>	ANN (C <sub>1</sub> )	2.95E-04	87.74	0.0510	19.18	0.0888	0.981	4.32
		HS (C <sub>1</sub> +C <sub>2</sub> )	<b>1.74E-05</b>	<b>97.06</b>	<b>0.0032</b>	<b>5.76</b>	<b>0.0041</b>	<b>0.999</b>	<b>14.34</b>
Vallila	PM <sub>10</sub>	ANN (C <sub>1</sub> )	3.00E-04	88.57	0.0785	25.06	0.1864	0.964	3.36
		HS (C <sub>1</sub> +C <sub>2</sub> )	<b>4.72E-05</b>	<b>96.12</b>	<b>0.0121</b>	<b>9.19</b>	<b>0.0185</b>	<b>0.996</b>	<b>9.40</b>
	PM <sub>2.5</sub>	ANN (C <sub>1</sub> )	2.01E-04	94.57	0.0450	15.27	0.0850	0.990	5.76
		HS (C <sub>1</sub> +...+C <sub>5</sub> )	<b>4.23E-07</b>	<b>98.86</b>	<b>9.67E-05</b>	<b>0.85</b>	<b>1.33E-04</b>	<b>0.999</b>	<b>53.31</b>

doi:10.1371/journal.pone.0138507.t006

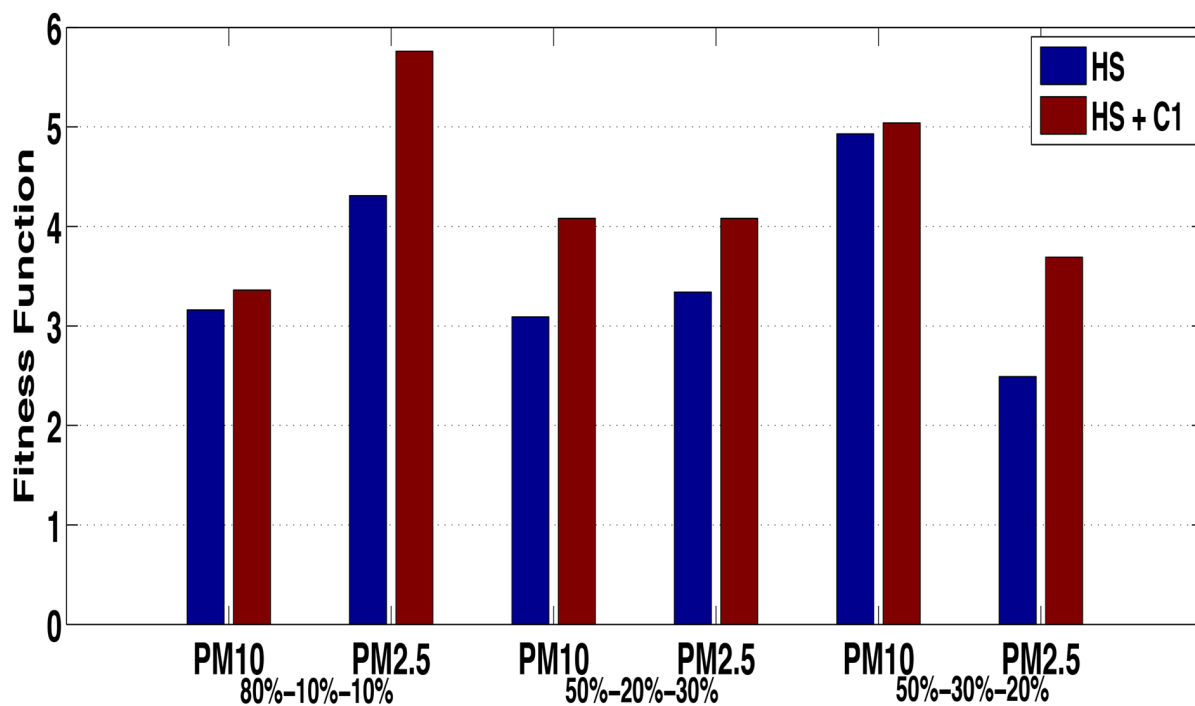
to observe, for all cases, that the correction via HS outperforms the correction via ANN. As an example, for PM<sub>10</sub> of the Vallila station the MSE and ARV results obtained by HS (C<sub>1</sub> + C<sub>2</sub>) is one order of magnitude smaller than the one obtained by ANN (C<sub>1</sub>). This result suggests that forecasting methods with high accuracy lead to better corrections. Thus, this study indicates that between the evaluated models, the more indicated for real applications is the HS corrected via HS.

A sensitivity analysis was performed with the objective to show the robustness of the proposed approach in terms of different proportions of the data set. Thus, three proportions are considered for training, validation and testing: 80% – 10% – 10%, 50% – 20% – 30% and 50% – 30% – 20%. Figs 9 and 10 show the results reached for sensitivity analysis in terms of the fitness for all series of both stations for correction via ANN and HS, respectively.

It is observed in Fig 9(a) and 9(b), that when the correction via ANN is performed the fitness value of HS + C<sub>1</sub> is greater than the HS without correction for concentration series in the Kallio and Vallila stations, respectively. For all considered cases, the same stopping criterion of the proposed approach was reached after the first correction (C<sub>1</sub>), which was the model residual for a white noise.



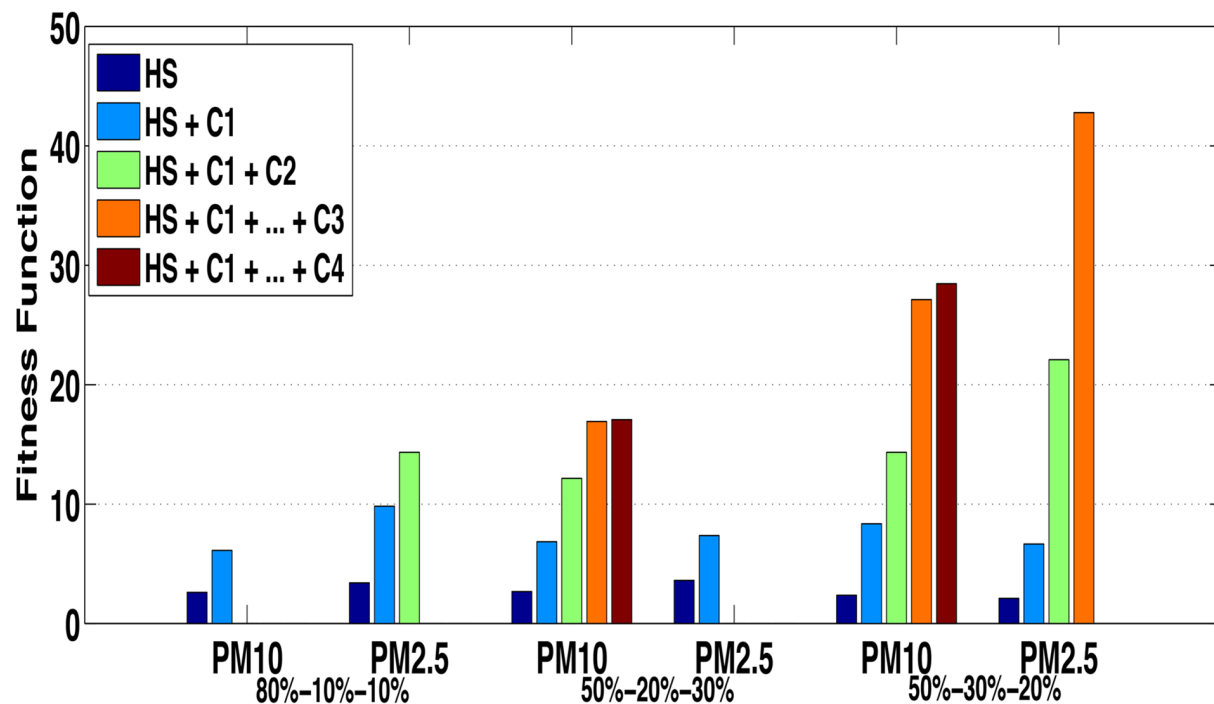
(a) Kallio station.



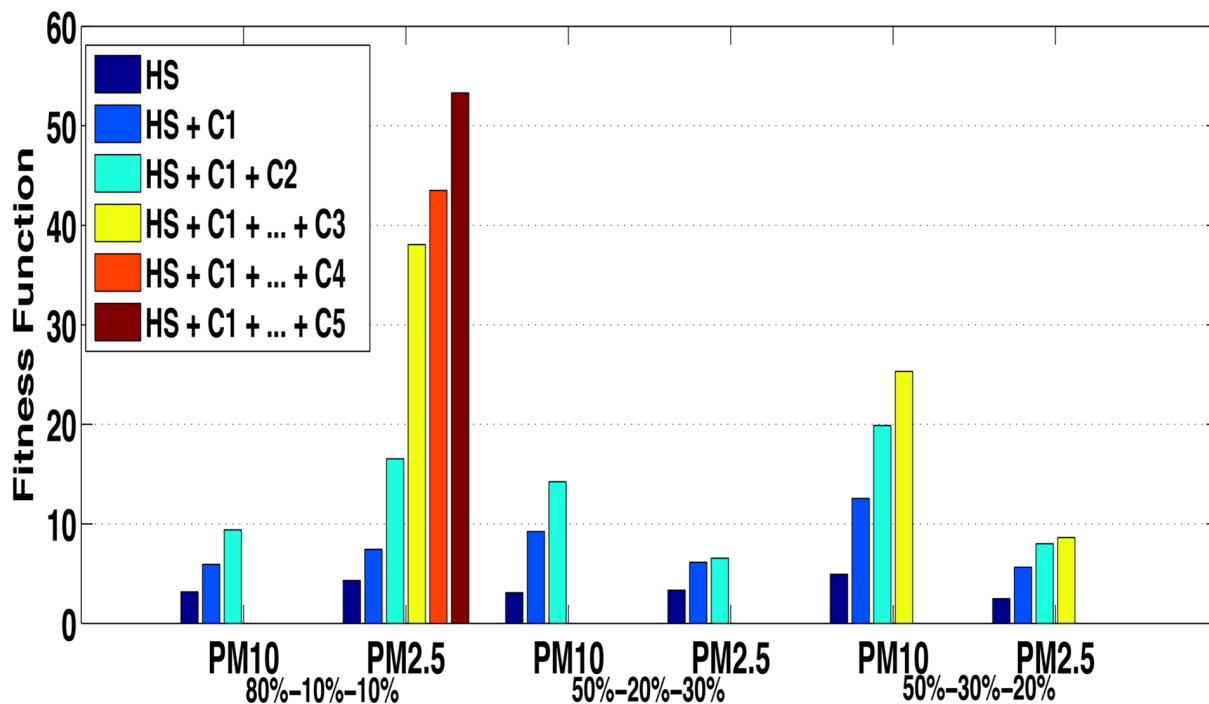
(b) Vallila station.

Fig 9. Fitness evolution for Kallio and Vallila stations with correction via ANN.

doi:10.1371/journal.pone.0138507.g009



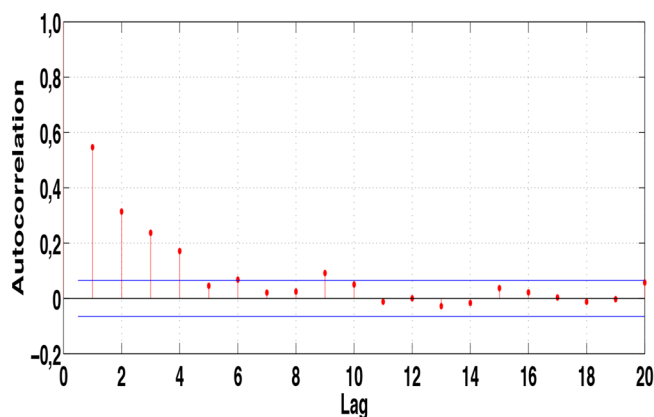
(a) Kallio station.



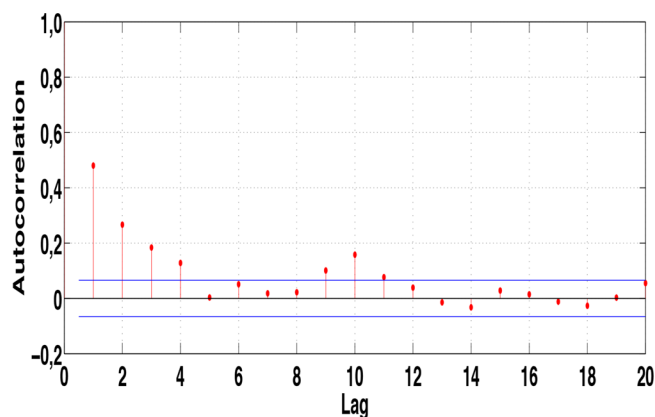
(b) Vallila station.

Fig 10. Fitness evolution for Kallio and Vallila stations with correction via HS.

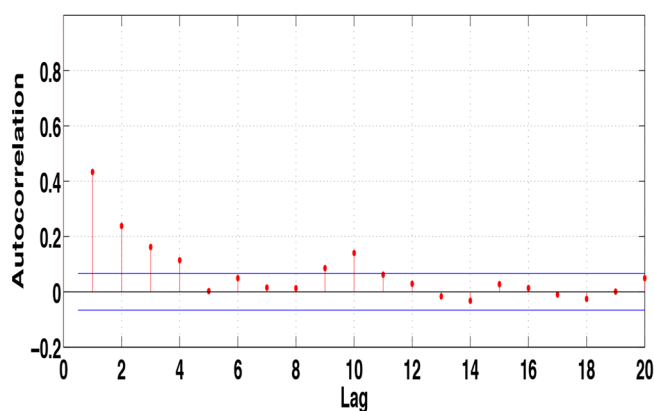
doi:10.1371/journal.pone.0138507.g010



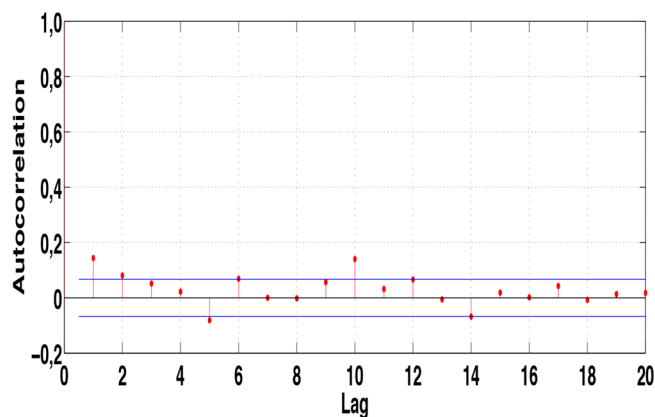
(a) ACF for HS residuals.



(b) ACF for HS +  $C_1$  residuals.



(c) ACF for HS +  $C_1$  +  $C_2$  residuals.



(d) ACF for HS +  $C_1$  +  $C_2$  +  $C_3$  residuals.

**Fig 11. Autocorrelation (ACF) for PM<sub>10</sub> concentration series for Kallio station.**

doi:10.1371/journal.pone.0138507.g011

Fig 10(a) and 10(b) show the sensitivity analysis of the fitness value with respect to the aforementioned proportions when the correction via HS is considered for Kallio and Vallila stations, respectively. For both stations, the performance at each correction (HS +  $C_n$ ) overcomes the performance of the previous correction (HS +  $C_{n-1}$ ). For all cases presented in Fig 10 the stopping criterion was the increase in the MAPE value. It is possible to observe that for most cases more than one correction via HS was performed (Fig 10), on the contrary of the correction via ANN, where only one correction was performed in all cases (Fig 9).

Fig 11 shows the ACF performed in the residuals obtained after each correction via HS for PM<sub>10</sub> series with proportion 50% – 30% – 20% for Vallila station (Fig 10(b)). It is possible to observe that none of the residuals, until the correction measured, have white noise behavior. However, Fig 11 shows that at each correction the ACF behavior tends to a white noise. In this case the stopping criteria was the increase in the MAPE value.

## Concluding Remarks

In this paper, a new approach is proposed to improve the performance of particulate matter (PM) forecasting. Important aspects of the proposed approach are highlighted as follows: it uses recursive residuals (error series) modeling; it uses as many residuals as the error series obtained so far is not assumed (by using the ACF) to be a white noise; it is quite general in the sense that it can use a set of different forecasters. Particularly, in the present work, a hybrid system (HS) [19] composed of a GA and ANN is corrected from two methods using residual modeling: an MLP neural network and own HS.

Results were presented in terms of six well-known performance metrics for time series forecasting (mean daily concentration levels) of PM<sub>2.5</sub> and PM<sub>10</sub> of two stations from Helsinki: Kallio and Vallila. Results obtained point out the benefits of using recursive residual modeling. An analysis was carried out with the purpose of evaluating the robustness of the proposed approach in terms of different proportions of the data set (training-validation-testing). It was observed that the proposed approach is robust for three different proportions: 80% – 10% – 10%, 50% – 20% – 30% and 50% – 30% – 20%. Thus, as PM concentration can damage human health, the approach presented in this paper may be used as an alternative to PM forecasting, which is a relevant issue to support decisions of organizations and governments for estimating the corresponding health risks.

With objective to investigate the performance of the proposed approach, novel studies will be performed in different scenarios with different combinations of forecasters. Future research directions include: forecasting of other contaminants (e.g. NO, NO<sub>2</sub>, NO<sub>x</sub>, CO, O<sub>3</sub>) concentrations; forecasting in the scenario of extreme events (e.g. dust-storm [53, 54]); performance evaluation considering missing data; performance assessment by using series with different times (e.g. second, minute, hour, or month); forecasting with dataset which presents measurement (or observational) errors. Furthermore, a new architecture can be developed with objective to estimate health risks from the forecasting of the proposed approach.

## Acknowledgments

This work was partially supported by Brazilian agencies: CNPq and Facepe.

## Author Contributions

Conceived and designed the experiments: PSGMN. Performed the experiments: PSGMN. Analyzed the data: PSGMN GDCC FM TAEF. Contributed reagents/materials/analysis tools: PSGMN GDCC FM TAEF. Wrote the paper: PSGMN GDCC FM TAEF.

## References

1. Bank W. Global Monitoring Report 2008: MDGs and the Environment: Agenda for Inclusive and Sustainable Development; 2008.
2. European Environment Agency. Air quality in Europe—Report 2014. European Environment Agency; 2014.
3. Wang KY, Chau TT. An Association between Air Pollution and Daily Outpatient Visits for Respiratory Disease in a Heavy Industry Area. PLoS ONE. 2013 10; 8(10):e75220. doi: [10.1371/journal.pone.0075220](https://doi.org/10.1371/journal.pone.0075220) PMID: [24204573](https://pubmed.ncbi.nlm.nih.gov/24204573/)
4. Feng J, Yang W. Effects of Particulate Air Pollution on Cardiovascular Health: A Population Health Risk Assessment. PLoS ONE. 2012 03; 7(3):e33385. doi: [10.1371/journal.pone.0033385](https://doi.org/10.1371/journal.pone.0033385) PMID: [22432017](https://pubmed.ncbi.nlm.nih.gov/22432017/)
5. Tong L, Li K, Zhou Q. Promoted Relationship of Cardiovascular Morbidity with Air Pollutants in a Typical Chinese Urban Area. PLoS ONE. 2014 09; 9(9):e108076. doi: [10.1371/journal.pone.0108076](https://doi.org/10.1371/journal.pone.0108076) PMID: [25247693](https://pubmed.ncbi.nlm.nih.gov/25247693/)
6. Block M, Calderón-Garcidueñas L. Air Pollution: Mechanisms of Neuroinflammation & CNS Disease. Trends in Neurosciences. 2009; 32(9):506–516. doi: [10.1016/j.tins.2009.05.009](https://doi.org/10.1016/j.tins.2009.05.009) PMID: [19716187](https://pubmed.ncbi.nlm.nih.gov/19716187/)



7. Fajersztajn L, Veras M, Barrozo LV, Saldiva P. Air pollution: a potentially modifiable risk factor for lung Cancer. *Nature Reviews Cancer*. 2013; 13(9):674–678. doi: [10.1038/nrc3572](https://doi.org/10.1038/nrc3572) PMID: [23924644](https://pubmed.ncbi.nlm.nih.gov/23924644/)
8. Brody H. Lung Cancer. *Nature*. 2014; 513(7517):S1. doi: [10.1038/508S1a](https://doi.org/10.1038/508S1a) PMID: [25208065](https://pubmed.ncbi.nlm.nih.gov/25208065/)
9. Dadvand P, Parker J, Bell ML, Bonzini M, Brauer M, Darrow LA, et al. Maternal Exposure to Particulate Air Pollution and Term Birth Weight: A Multi-Country Evaluation of Effect and Heterogeneity. *Environmental Health Perspectives*. 2013; 121(3):367–373. doi: [10.1289/ehp.1205575](https://doi.org/10.1289/ehp.1205575)
10. Harrison RM, Yin J. Particulate matter in the atmosphere: which particle properties are important for its effects on health? *Science of The Total Environment*. 2000; 249(1-3):85–101. doi: [10.1016/S0048-9697\(99\)00513-6](https://doi.org/10.1016/S0048-9697(99)00513-6) PMID: [10813449](https://pubmed.ncbi.nlm.nih.gov/10813449/)
11. Englert N. Fine particles and human health—a review of epidemiological studies. *Toxicology Letters*. 2004; 149(1-3):235–242. doi: [10.1016/j.toxlet.2003.12.035](https://doi.org/10.1016/j.toxlet.2003.12.035) PMID: [15093269](https://pubmed.ncbi.nlm.nih.gov/15093269/)
12. Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, et al. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Air Pollution and Health: a European Approach*. BMJ (Clinical research ed). 1997; 314(7095):1658–1663.
13. Shang Y, Sun Z, Cao J, Wang X, Zhong L, Bi X, et al. Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. *Environment International*. 2013; 54(0):100–111. doi: [10.1016/j.envint.2013.01.010](https://doi.org/10.1016/j.envint.2013.01.010) PMID: [23434817](https://pubmed.ncbi.nlm.nih.gov/23434817/)
14. Pascal M, Falq G, Wagner V, Chatignoux E, Corso M, Blanchard M, et al. Short-term impacts of particulate matter (PM<sub>10</sub>, PM<sub>10-2.5</sub>, PM<sub>2.5</sub>) on mortality in nine French cities. *Atmospheric Environment*. 2014; 95(0):175–184. doi: [10.1016/j.atmosenv.2014.06.030](https://doi.org/10.1016/j.atmosenv.2014.06.030)
15. Genc S, Zadeoglulari Z, Fuss SH, Genc K. The Adverse Effects of Air Pollution on the Nervous System. *Journal of Toxicology*. 2012; 2012(782462):23 pages. doi: [10.1155/2012/782462](https://doi.org/10.1155/2012/782462) PMID: [22523490](https://pubmed.ncbi.nlm.nih.gov/22523490/)
16. Kleinman MT, Campbell A. Central Nervous System Effects of Ambient Particulate Matter: The Role of Oxidative Stress and Inflammation. University of California and California Air Resources Board; 2014.
17. Kim JW, Park S, Lim CW, Lee K, Kim B. The Role of Air Pollutants in Initiating Liver Disease. *Toxicological Research*. 2014; 30(2):65–70. doi: [10.5487/TR.2014.30.2.065](https://doi.org/10.5487/TR.2014.30.2.065) PMID: [25071914](https://pubmed.ncbi.nlm.nih.gov/25071914/)
18. Tarantino G, Capone D, Finelli C. Exposure to ambient air particulate matter and non-alcoholic fatty liver disease. *World Journal of Gastroenterology: WJG*. 2013; 19(25):3951–3956. doi: [10.3748/wjg.v19.i25.3951](https://doi.org/10.3748/wjg.v19.i25.3951) PMID: [23840139](https://pubmed.ncbi.nlm.nih.gov/23840139/)
19. de Mattos Neto PSG, Madeiro F, Ferreira TAE, Cavalcanti GDC. Hybrid intelligent system for air quality forecasting using phase adjustment. *Engineering Applications of Artificial Intelligence*. 2014; 32(0):185–191. doi: [10.1016/j.engappai.2014.03.010](https://doi.org/10.1016/j.engappai.2014.03.010)
20. Sharma S, Barai SV, Dikshit AK. Studies of air quality predictors based on neural networks. *International Journal of Environment and Pollution*. 2003; 19(5):442–453. doi: [10.1504/IJEP.2003.004327](https://doi.org/10.1504/IJEP.2003.004327)
21. Ordieres JB, Vergara EP, Capuz RS, Salazar RE. Neural network prediction model for fine particulate matter (PM<sub>2.5</sub>) on the US-Mexico border in El Paso (Texas) and Ciudad Juárez (Chihuahua). *Environmental Modelling & Software*. 2005; 20(5):547–559. doi: [10.1016/j.envsoft.2004.03.010](https://doi.org/10.1016/j.envsoft.2004.03.010)
22. Kukkonen J, Partanen L, Karppinen A, Ruuskanen J, Junninen H, Kolehmainen M, et al. Extensive evaluation of neural network models for the prediction of NO<sub>2</sub> and PM<sub>10</sub> concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmospheric Environment*. 2003; 37(32):4539–4550. doi: [10.1016/S1352-2310\(03\)00583-1](https://doi.org/10.1016/S1352-2310(03)00583-1)
23. Caselli M, Trizio L, de Gennaro G, Ielpo P. A Simple Feedforward Neural Network for the PM<sub>10</sub> Forecasting: Comparison with a Radial Basis Function Network and a Multivariate Linear Regression Model. *Water, Air, and Soil Pollution*. 2009; 201(1-4):365–377. doi: [10.1007/s11270-008-9950-2](https://doi.org/10.1007/s11270-008-9950-2)
24. de Gennaro G, Trizio L, Gilio AD, Pey J, Pérez N, Cusack M, et al. Neural network model for the prediction of PM<sub>10</sub> daily concentrations in two sites in the Western Mediterranean. *Science of The Total Environment*. 2013; 463-464(0):875–883. doi: [10.1016/j.scitotenv.2013.06.093](https://doi.org/10.1016/j.scitotenv.2013.06.093) PMID: [23872183](https://pubmed.ncbi.nlm.nih.gov/23872183/)
25. Perez P, Reyes J. An integrated neural network model for PM<sub>10</sub> forecasting. *Atmospheric Environment*. 2006; 40(16):2845–2851. doi: [10.1016/j.atmosenv.2006.01.010](https://doi.org/10.1016/j.atmosenv.2006.01.010)
26. Kurt A, Oktay AB. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. *Expert Systems with Applications*. 2010; 37(12):7986–7992. doi: [10.1016/j.eswa.2010.05.093](https://doi.org/10.1016/j.eswa.2010.05.093)
27. Slini T, Kaprara A, Karatzas K, Moussiopoulos N. PM<sub>10</sub> Forecasting for Thessaloniki, Greece. *Environmental Modelling & Software*. 2006; 21(4):559–565. doi: [10.1016/j.envsoft.2004.06.011](https://doi.org/10.1016/j.envsoft.2004.06.011)
28. Voukantsis D, Karatzas K, Kukkonen J, Rasanen T, Karppinen A, Kolehmainen M. Intercomparison of air quality data using principal component analysis, and forecasting of PM<sub>10</sub> and PM<sub>2.5</sub> concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Science of The Total Environment*. 2011; 409(7):1266–1276. doi: [10.1016/j.scitotenv.2010.12.039](https://doi.org/10.1016/j.scitotenv.2010.12.039) PMID: [21276603](https://pubmed.ncbi.nlm.nih.gov/21276603/)

29. Antanasijević DZ, Pocajt VV, Povrenović DS, Ristić MD, Perić-Grujić AA. PM10 emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *Science of The Total Environment*. 2013; 443(0):511–519. doi: [10.1016/j.scitotenv.2012.10.110](https://doi.org/10.1016/j.scitotenv.2012.10.110) PMID: [23220141](https://pubmed.ncbi.nlm.nih.gov/23220141/)
30. Niska H, Rantamaki M, Hiltunen T, Karppinen A, Kukkonen J, Ruuskanen J, et al. Evaluation of an integrated modelling system containing a multi-layer perceptron model and the numerical weather prediction model HIRLAM for the forecasting of urban airborne pollutant concentrations. *Atmospheric Environment*. 2005; 39(35):6524–6536. doi: [10.1016/j.atmosenv.2005.07.035](https://doi.org/10.1016/j.atmosenv.2005.07.035)
31. Perez P. Combined model for PM<sub>10</sub> forecasting in a large city. *Atmospheric Environment*. 2012; 60:271–276. doi: [10.1016/j.atmosenv.2012.06.024](https://doi.org/10.1016/j.atmosenv.2012.06.024)
32. Siwek K, Osowski S. Improving the accuracy of prediction of PM<sub>10</sub> pollution by the wavelet transformation and an ensemble of neural predictors. *Engineering Applications of Artificial Intelligence*. 2012; 25(6):1246–1258. doi: [10.1016/j.engappai.2011.10.013](https://doi.org/10.1016/j.engappai.2011.10.013)
33. Mishra D, Goyal P, Upadhyay A. Artificial intelligence based approach to forecast PM2.5 during haze episodes: A case study of Delhi, India. *Atmospheric Environment*. 2015; 102(0):239–248. doi: [10.1016/j.atmosenv.2014.11.050](https://doi.org/10.1016/j.atmosenv.2014.11.050)
34. Qin S, Liu F, Wang J, Sun B. Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models. *Atmospheric Environment*. 2014; 98(0):665–675. doi: [10.1016/j.atmosenv.2014.09.046](https://doi.org/10.1016/j.atmosenv.2014.09.046)
35. Ul-Saufie AZ, Yahaya AS, Ramli NA, Rosaida N, Hamid HA. Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmospheric Environment*. 2013; 77(0):621–630. doi: [10.1016/j.atmosenv.2013.05.017](https://doi.org/10.1016/j.atmosenv.2013.05.017)
36. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003; 50(0):159–175. doi: [10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)
37. Firmino PRA, de Mattos Neto PSG, Ferreira TAE. Error modeling approach to improve time series forecasters. *Neurocomputing*. 2015; 153:242–254. doi: [10.1016/j.neucom.2014.11.030](https://doi.org/10.1016/j.neucom.2014.11.030)
38. Firmino PRA, de Mattos Neto PSG, Ferreira TAE. Correcting and combining time series forecasters. *Neural Networks*. 2014; 50(0):1–11. doi: [10.1016/j.neunet.2013.10.008](https://doi.org/10.1016/j.neunet.2013.10.008) PMID: [24239986](https://pubmed.ncbi.nlm.nih.gov/24239986/)
39. Al-Alawi SM, Abdul-Wahab SA, Bakheit CS. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environmental Modelling & Software*. 2008; 23(4):396–403. doi: [10.1016/j.envsoft.2006.08.007](https://doi.org/10.1016/j.envsoft.2006.08.007)
40. Ettouney RS, Mjalli FS, Zaki JG, El-Rifai MA, Ettouney HM. Forecasting of ozone pollution using artificial neural networks. *Management of Environmental Quality: An International Journal*. 2009; 20(6):668–683. doi: [10.1108/14777830910990843](https://doi.org/10.1108/14777830910990843)
41. Westerlund J, Urbain JP, Bonilla J. Application of air quality combination forecasting to Bogota. *Atmospheric Environment*. 2014; 89(0):22–28. doi: [10.1016/j.atmosenv.2014.02.015](https://doi.org/10.1016/j.atmosenv.2014.02.015)
42. Sánchez AB, Ordóñez C, Lasheras FS, de Cos Juez FJ, Roca-Pardiñas J. Forecasting SO<sub>2</sub> Pollution Incidents by means of Elman Artificial Neural Networks and ARIMA Model. *Abstract and Applied Analysis*. 2013; 2013(6):238259.
43. Luis A Díaz-Robles and Juan C Ortega and Joshua S Fu and Gregory D Reed and Judith C Chow and John G Watson and Juan A Moncada-Herrera. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*. 2008; 42(35):8331–8340. doi: [10.1016/j.atmosenv.2008.07.020](https://doi.org/10.1016/j.atmosenv.2008.07.020)
44. Zhang G, Huang S, Duan Q, Shu W, Hou Y, Miao X, et al. Application of a hybrid model for predicting the incidence of tuberculosis in Hubei, China. *PLOS ONE*. 2013; 8(11):1–7.
45. Yu L, Zhou L, Tan L, Jiang H, Wang Y, Wei S, et al. Application of a new hybrid model with seasonal auto-regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (ANRNN) in forecasting incidence cases of HFMD in Shenzhen, China. *PLOS ONE*. 2014; 9(6):1–9. doi: [10.1371/journal.pone.0098241](https://doi.org/10.1371/journal.pone.0098241)
46. Zhou L, Yu L, Wang Y, Lu Z, Tian L, Tan L, et al. A hybrid model for predicting the prevalence of schistosomiasis in humans of Qianjiang city, China. *PLOS ONE*. 2014; 9(8):1–12. doi: [10.1371/journal.pone.0104875](https://doi.org/10.1371/journal.pone.0104875)
47. Vlachogianni A, Kassomenos P, Karppinen A, Karakitsios S, Kukkonen J. Evaluation of a multiple regression model for the forecasting of the concentrations of NO<sub>x</sub> and PM<sub>10</sub> in Athens and Helsinki. *Science of The Total Environment*. 2011; 409(8):1559–1571. doi: [10.1016/j.scitotenv.2010.12.040](https://doi.org/10.1016/j.scitotenv.2010.12.040) PMID: [21277004](https://pubmed.ncbi.nlm.nih.gov/21277004/)
48. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*. Wiley Series in Probability and Statistics. Wiley; 2008.

49. Rodrigues ALJ, Silva DA, de Mattos Neto PSG, Ferreira TAE. An Experimental Study of Fitness Function and Time Series Forecasting Using Artificial Neural Networks. In: Genetic and Evolutionary Computation Conference (GECCO 2010). ACM; 2010. p. 2015–2018.
50. de Mattos Neto PSG, Rodrigues ALJ, Ferreira TAE, Cavalcanti GDC. An intelligent perturbative approach for the time series forecasting problem. In: IEEE World Congress on Computational Intelligence (WCCI 2010). IEEE; 2010. p. 1–8.
51. Silva DA, Alves GI, de Mattos Neto PSG, Ferreira TAE. Measurement of Fitness Function efficiency using Data Envelopment Analysis. *Expert Systems with Applications*. 2014; 41(16):7147–7160. doi: [10.1016/j.eswa.2014.06.001](https://doi.org/10.1016/j.eswa.2014.06.001)
52. Ferreira TAE, Vasconcelos GC, Adeodato PJL. A New Intelligent System Methodology for Time Series Forecasting with Artificial Neural Networks. *Neural Processing Letters*. 2008; 28(2):113–129. doi: [10.1007/s11063-008-9085-x](https://doi.org/10.1007/s11063-008-9085-x)
53. Csavina J, Field J, Félix O, Corral-Avitia AY, Sáez AE, Betterton EA. Effect of wind speed and relative humidity on atmospheric dust concentrations in semi-arid climates. *Science of The Total Environment*. 2014; 487(0):82–90. doi: [10.1016/j.scitotenv.2014.03.138](https://doi.org/10.1016/j.scitotenv.2014.03.138) PMID: [24769193](https://pubmed.ncbi.nlm.nih.gov/24769193/)
54. Wong MS, Xiao F, Nichol J, Fung J, Kim J, Campbell J, et al. A multi-scale hybrid neural network retrieval model for dust storm detection, a study in Asia. *Atmospheric Research*. 2015; 158-159(0):89–106. doi: [10.1016/j.atmosres.2015.02.006](https://doi.org/10.1016/j.atmosres.2015.02.006)