Article

FOCUSED–Short-Term Wind Speed Forecast Correction Algorithm Based on Successive NWP Forecasts for Use in Traffic Control Decision Support Systems

Zdravko Kunić ^{1,2,*} , Bernard Ženko ³ and Biljana Mileva Boshkoska ^{1,3} ¹ Faculty of Information Studies in Novo Mesto, 8000 Novo Mesto, Slovenia; biljana.mileva@fis.unm.si² Department for Information Systems and Business Analytics, Algebra University College, 10000 Zagreb, Croatia³ “Jožef Stefan” Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia; bernard.zenko@ijs.si

* Correspondence: zdravko.kunic@algebra.hr; Tel.: +385-98-205-393

Abstract: In this paper, we propose a new algorithm, called FOCUSED (FOrecast Correction Using Successive prEDictions), for forecast correction of short-term wind speed predictions. We developed FOCUSED with the aim of improving the forecast of bora gusts, which frequently result in high-speed wind situations dangerous for traffic. The motivation arises from occasionally ambiguous results of the currently deployed decision support system, which aids traffic management in strong and gusty wind conditions at the coast of Croatia. The proposed correction algorithm uses characteristics of numerical weather prediction models to iteratively forecast the wind speed multiple times for the same future window. We use these iterative predictions as input features of the FOCUSED algorithm and get the corrected predictions as the output. We compared the proposed algorithm with artificial neural networks, random forests, support vector machines, and linear regression to demonstrate the superiority of the algorithm’s performance on a data set comprising five years of real data measurements at the Croatian bridge “Krak” and complementary historical forecasts by ALADIN (Aire Limitée Adaptation dynamique Développement InterNational) numerical weather prediction model.

Keywords: traffic management; wind speed prediction; forecast correction; neural networks; successive forecasts

Citation: Kunić, Z.; Ženko, B.; Boshkoska, B.M. FOCUSED–Short-Term Wind Speed Forecast Correction Algorithm Based on Successive NWP Forecasts for Use in Traffic Control Decision Support Systems. *Sensors* **2021**, *21*, 3405. <https://doi.org/10.3390/s21103405>

Academic Editor: Dimitrios Moshou

Received: 31 March 2021

Accepted: 10 May 2021

Published: 13 May 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bora (Bura in Croatian) is a cold, strong and gusty wind that blows along the Eastern Adriatic coast and islands. Each winter, several damaging bora storms hit the coastal region of Croatia, strongly affecting sea, air and road transport safety, and life in general [1]. In particular, bora can have a severe negative impact on road traffic security [2]. In critical situations, road managers use wind forecasting models as an aid in making decisions to open or close a road for traffic or to define new speed limits. Three parameters define a critical wind situation in the context of traffic control in the Eastern Adriatic coast: wind direction, 10-min mean wind speed and 10-min maximum wind speed. For each location, these parameters define critical wind situations that vary depending on a specific terrain configuration and pavement condition (dry/wet/icy) [2].

According to the time-horizon classification [3], in the process of wind modelling and wind forecast correction, we distinguish three different forecasting time scales: immediate short-term (8 h ahead), short-term (a day ahead) and long-term (multiple days ahead) forecasts. The ALADIN [4] forecast is used by an Anemo-Alarm [5] traffic control decision support system (DSS) twice a day. Due to long time intervals between two consecutive forecasts, traffic managers occasionally are presented with contradictory measurements (e.g., current wind speed) and forecasts from ALADIN. Hence, road managers have to deal

with boundary cases in which they have to decide about decreasing maximum allowed speed limits or even closing a road based on inconsistent or ambiguous forecasts and measurements. The most evident problem of the existing decision support system is its inability to suggest a right decision when there are significant differences between forecasted and measured wind speed trends, and especially when the trends are opposite. In such cases, road managers may select a decision which can lead to potentially disastrous consequences. In particular, the decision to close a road due to strong wind may lead to substantial loss of drivers' time, as well as a financial loss for drivers and companies, and consequently, the country's economy. On the other hand, a decision of leaving a road open may lead to significant material damage or even more catastrophic scenarios, including life losses. It is, therefore, important to provide the best possible wind forecasts to support decision makers, and in particular, it is important to provide them with wind forecasts between two consecutive ALADIN forecasts.

To address this problem, we propose a correction algorithm, called FOCUSED, that uses existing forecasts from ALADIN as inputs to an ANN (Artificial Neural Network) and outputs corrected predictions. The proposed algorithm aims to improve the immediate-short-term and short-term (1 to 12 h in advance) wind forecast accuracy during the period when a newer ALADIN forecast is not available and supports road managers in the process of choosing the most appropriate action based on timely and relevant information. In this research we show that the latest successive forecast is not always the best forecast and that it is possible to improve the wind speed forecast accuracy using only consecutive forecasts of the existing physical forecast models (i.e., ALADIN), without any additional mathematical or physical input features.

FOCUSED can be used with any data-mining model, therefore we compared its usage with artificial neural networks, random forests, support vector machines, and linear regression to demonstrate the superiority of the algorithm's performance on a data set comprising five years of real data measurements at the Croatian bridge "Krk" and complementary historical forecasts by ALADIN numerical weather prediction model.

The motivation for the proposed algorithm comes from the challenges posed by the existing forecasting system at the "Krk" bridge in Croatia. In particular, our objective is to improve prediction of short-burst critical high-speed wind situations that are dangerous for traffic. Our solution can easily be integrated into the existing decision support system and is computationally inexpensive. However, usability of the proposed method is not limited to the above scenario and can be also used for improving forecasts in general, for example, for forecasting wind power plant production.

2. Related Work

In meteorology, the improvement of the forecasts is performed by employing post-processing methods using local measurements and weather prediction model outputs [6]. Model improvements using correction algorithms represent an acceptable option because they improve the inaccurate representation of atmospheric state by NWP (Numerical Weather Prediction) models and decrease observation errors [7]. In general, a correction algorithm feeds the input features into a "correction" model that outputs corrected original predictions. Input features can include data not utilized by the original NWP model, for instance, orography roughness optimization (considering more detailed topographic relief), historical data, real-time measured data and calculated errors. The correction model itself can be a result of a broad spectrum of data-mining and post-processing methods, as we describe below. It can also be based on the latest results obtained by the NWP models or unused features in the original NWP models. Many correction algorithms have their origins in Kalman filtering [8–10], neural networks [11] and their combination [12], as well as other statistical methods which combine different post-processing methods to reduce errors of physical models [10].

Kalman filtering is one of the popular post-processing methods [8–10]. According to [13], when dealing with limited-area atmospheric models with different options and

capabilities of horizontal resolution, it leads to the elimination of possible systematic errors, even in lower resolution cases, contributing further to significant reduction of the required CPU time. In [14], the error forecast models are based on a support vector machine or extreme learning machine. In [15], authors propose wind power forecast correction by subtracting the biased mean from the wind power forecast error.

A variational method [7] for correcting non-systematic NWP forecast errors is based on previous numerical forecasts assuming that the error is linearly dependent on some combination of the forecast fields. Using the Single Value Decomposition (SVD) of the covariance matrix between the forecast and forecasting they obtain the inverse mapping from flow space to the error space during the training period. To avoid the difficulty in solving the inverse matrix, they reduced the background covariance matrix to a simple diagonal matrix.

Errors of wind speed forecasts can also be reduced by searching for an optimal combination of post-processing methods, as described in [10] where adaptive approaches to post-processing wind speed forecasts are discussed and compared using automatic methods for combining forecast streams.

In the syntheses-correcting forecasting model [16], the wind speed forecasting bias correction method is based on the Empirical Orthogonal Function (EOF) and regression analysis.

In [6], the correction models use polynomial neural networks for modelling real complex systems. This method also revises forecasts according to the corrective function that depends on real observations.

Combined forecasting methods, proposed in [14], investigate the correlation relationships of forecast errors of the autoregressive model, the persistence method and the support vector machine model in various forecasting modes. Authors proposed a strategy for selecting the input variables and defined the range of input variables according to the results of the correlation analysis.

In [17], the authors aim to improve the performance of the real-time decomposition-based forecasting method after they uncover the factors attributed to its unsatisfactory performance. They decomposed the raw wind speed time series into a different subseries. To reduce the disturbance of illusive components, they used (a) kernel density estimation-based Kullback-Leibler divergence and (b) energy measure. Finally, the hybrid of least squares support vector machine and generalized autoregressive conditionally heteroscedastic model is introduced to correct the resulting error component if its inherent correlation and heteroscedasticity cannot be neglected [17].

Related to the problem of short-time wind speed forecast correction with the aim of improving decision support systems for traffic control in dangerous wind situation is also the problem of wind farm power prediction. For example, in the study [18] about wind farm NWP wind speed correction methods, measured time series were decomposed into different bands by wavelet multi-resolution analysis. Correction premise was verified using the Pearson product-moment correlation coefficient, and then the linear correction method was used to correct the low-frequency stationary NWP wind speed.

We started the research with the hypothesis that the last NWP forecast is not always the most accurate existing forecast. None of the abovementioned research and correction algorithms take into consideration previous forecasts for the same future time period, and they also do not differentiate continuous wind speed/direction measurements from critical wind speed/direction situations for traffic.

The most related algorithm to FOCUSED is a forecast error correction method in numerical weather prediction using recent multiple-time evolution data [19]. It also falls into the category of those who act as if the last forecast is the best forecast. It compares the first part of the forecasted values with actual measured values to find a function that describes the forecast error and uses this function to the rest of the forecast to correct the errors. The model error is expressed as a Lagrange interpolation polynomial, while the coefficients of the polynomial are determined by past model performance. Both algorithms,

FOCUSED and recent multiple-time evolution data divide forecasted time series data into two parts: past and future forecasted data, with respect to the observation point (which is in fact the present time). Both algorithms need some initial time after receiving new forecasts to calculate initial errors. The differences start with the scope and evaluation data set—the scope of the recent multiple-time evolution algorithm is general NWP correction based on vectors of multiple various meteorological dimensions and evaluation is based on an artificially created data set, while our scope is only wind speed correction without using other variables, and evaluation is based on a real data set. To generate corrected time series with the best possible congruence between forecasted and measured data recent multiple-time evolution searches for an error function between two time series (one forecast and one measurement) using Lagrangean interpolation polynomial, we combine multiple time series (multiple forecasts and one measurement) using ML recent multiple-time evolution use flow of last n samples of vectors (of one recent forecast), and we use groups of samples for the same forecasted period/interval (multiple time series data of multiple recent forecasts).

Therefore, we propose a novel algorithm in this paper that differs from existing algorithms by taking into consideration multiple successive forecasts instead of only the last one and a single bora duration as the time horizon for training the data set instead of historical data for the observed location.

3. Materials and Methods

3.1. Forecasting Setup

The FOCUSED algorithm is a correction algorithm which uses previous forecasts from the NWP model as its inputs (Figure 1). Forecasts that share particular outcome periods are grouped together. Generally, any NWP model whose outputs are to be used in FOCUSED needs to calculate a forecast for the exact number of hours ahead F_p (forecast period), to be scheduled to run in regular time intervals T_s (time between two consecutive NWP model runs), and it may spend some processing time T_p (time of processing) to acquire new data and make the output of a physical model available to the DSS. T_{NWP} denote the point in time when NWP has created the forecasts in a successive group, and T_{DSS} denotes the point in time when the forecast group becomes available to DSS. The data set used in this research consists of two parts: hourly wind speed/direction forecasts for 72 h periods calculated every 12 h, and actual wind speed/direction data measured every 10 min.

Figure 1. A group of successive forecasts.

Ideally, the correction algorithm should get all forecasts at time T_{NWP} . However, a delay of T_p can occur between the end time of forecast modelling and the time when DSS gets the data, which is a result of a data synchronization process between the modelling system and decision support system. It may last from several seconds to several hours, depending on the infrastructural and business relationships between the NWP forecast provider and decision support system user.

Knowing the forecast period F_p , the time gap T_p and the time T_s between two consecutive runs of NWP, we can calculate the number of overlapping forecasts, which we denote as a number of relevant previous forecasts (N_{pf}). We calculate the number of relevant previous forecasts as an input parameter to the proposed algorithm as:

$$N_{pf} = A \frac{(F_p - T_p)}{T_s} \quad (1)$$

Ideally, the N_{pf} should be equal to the number of consecutive forecasts in order to use all whole available previous forecasts, however due to T_p , this number may decline because we will lose an hour of the oldest forecast for each hour of T_p .

3.2. The FOCUSED Algorithm

The FOCUSED algorithm consists of four main steps and is presented in Figure 2.

Figure 2. Steps of the FOCUSED algorithm.

We used multiple models to test the performance of the algorithm: Artificial Neural Networks (ANN), Random Forests (RF), Support Vector Machines (SVM) and Linear Regression (LR). Since there is no difference in steps of the algorithm when different data-

mining methods are used, we will thoroughly walkthrough using ANN, the model that in this research statistically performed the best.

The first step is data pre-processing (Figure 3). We used two types of input data: successive overlapping forecasts for a future 12 h period, and actual measured data up to the observation moment. In the pre-processing step, all input data are normalized to the same time granulation level (in our case, 1 h) and it is ensured that there are no missing or extreme values in the data set. Additionally, one should be aware that accuracy of corrected wind speed obtained by running the FOCUSED algorithm is also subject to input data integrity issues. In this research we used validated data sets, but in order to offer a universal guide we should point out the need to address integrity issues as well. They can arise from two main directions: (a) unintentionally unreliable measurements (hardware or software errors or malfunctions) or (b) intentional data disruption—e.g., cyber-attacks related to data integrity that are expected to harm the performances of forecasting systems [20], or false data injection attacks against wind power deterministic and probabilistic forecasting [21].

Figure 3. Data pre-processing steps.

The next step is the filtering of non-critical wind situations from the data set. This step is important because the algorithm is intended to support decisions regarding road closure or speed limitations during strong wind situations, hence the need to isolate only such conditions. To distinguish critical from non-critical wind situations, we defined the starting and the ending point of the single bora period. For the purpose of this study which uses continuous data from five consecutive years, we identified a critical bora situation as the wind that blows from any direction between north and east ($0\text{--}90^\circ$), whose average hourly speed remains above wind speed threshold for at least 6 h and lasts until the speed drops below the silence threshold for at least 5 h. In Croatia, a wind speed that is higher than 17 m/s represents one of national weather alert criteria [22]. Therefore, the upper wind speed threshold is set to that speed, and for the same reason the silence threshold is set to 5 m/s, just below the lower boundary during icy road condition periods. In Figure 4, we show an example of an hourly average measurements during the five years with upper and lower boundaries used to identify critical wind situations.

Step 3 of the FOCUSED algorithm comprises filtering out too short critical wind situations. In particular, we decided to exclude critical winds shorter than 24 h from the research, which is a period that we divided to a minimum of 12 h of data for training the model (training data set) and 12 h of data for testing the model's performance (testing data set), as shown in Table 1. In windy situations which last longer than 24 h, we created different lengths of a training set, while the length of the test data set was always 12 h. In this research, one hour of time series wind data is aggregated and is denoted as one data sample. Each strong wind situation with the length of n_t (total number of samples) was repeatedly tested n_r (number of algorithm runs) times:

Figure 4. Hourly average wind speed measurements.

$$n_r = n_t - (24 - 1) = n_t - 23. \quad (2)$$

Table 1. Training and test set samples for four algorithms' runs.

Algorithm Run Number	Training Set Samples	Test Samples
1	1–12	13–24
2	1–13	14–25
3	1–14	15–26
4	1–15	16–27

For example, if the time interval of Bura is 27 h, resulting in $n_t = 27$ samples, then $n_r = 27 - (24 - 1) = 4$ algorithm runs.

In the last algorithm step, Step 4, we performed training of the model for each particular strong wind situation, defined in Step 3, to get the initial prediction FC (Forecast Corrected). We defined the training period as a minimum of 12 h after detecting a strong wind situation and created an input vector of successive wind speed time series.

The correction between two consecutive NWP forecasts (Figure 5) is based on a forecast group that starts at t_1 and lasts for n 12 h time slots. The oldest forecast FD_1 (Forecasted Data 1) starts at t_1 and contributes to the model with its fifth (training/test data set) and sixth (evaluation data set) time slots. The next forecast FD_2 starts at t_2 and contributes to the model with its fourth and fifth time slots, and so on until the last forecast FD_6 starts at t_{n-1} and contributes to the model with its first two time slots.

Figure 5. Data sets immediately after receiving a new NWP forecast.

There are n consecutive runs of the NWP model, resulting in n forecasts (FD_1 – FD_n) for the respective periods F_p , denoted in yellow. The white boxes denote the forecasts used as the data-mining model's input variables and the blue box denotes actual measured wind speed used as a target variable. Gray colors denote successive forecasts predicted by the NWP model for the observed future 12 h period.

Consider that F_p is 72 h and T_s is 12 h. At the time t_1 , the NWP model has created a forecast for the next 72 h, denoted as FD_1 . At the time t_2 , the NWP model makes a new forecast for the next 72 h denoted as FD_2 , and so on. During one F_p period, the NWP model makes six forecasts whose lengths are equal to the T_s period. In application where the proposed algorithm is used as a part of real-time decision support system, the modelling process should repeat each hour, immediately after receiving new measured hourly wind-speed data. When the critical wind situation lasts longer than F_p , the training set becomes larger than F_p (Figure 6). Every new NWP forecast becomes FD_n . Previous FD_n becomes FD_{n-1} , and at the end FD_2 becomes FD_1 , hence the training set could grow indefinitely.

Figure 6. ANN data sets during long critical wind situations.

We used a group of forecasts from the time slot between t_{0+12h} (12 h after start of critical wind condition) and t_{obs} (observation point) as input vectors to train the model. The model's performance was tested on time-slot data between t_{obs} and $t_{obs+12h}$.

To test the algorithm's performance, the data was split into training and test subsets according to the length of the bora situation and actual observation point. The training subset consisted of 12 h data or more, depending on the length of critical bora situation, calculated with:

$$X_{length} = B_{length} - n, \forall n : T_s \leq n \leq (B_{length} - T_s) \quad (3)$$

where X_{length} is the length of training subset, B_{length} is the length of the whole observed bora situation, and T_s is the time between two consecutive NWP model runs.

The final part of the fourth step of the algorithm is vertical shift of the corrected curve, so that first corrected wind speed data point is aligned with the last measured data point (Figure 7). The end result of the proposed algorithm is FCA (FC aligned) which keeps the shape of the FC, with its beginning point aligned to the last measured value.

Figure 7. Alignment of corrected forecast to the last measured value.

3.3. Training the Model (ANN Example)

Among available ANN methods, for our correction model we used a MLP (Multi-Layer Perceptron) regressor with three hidden layers of 500 neurons and ReLU (Rectified Linear Unit) activation function. Multi-layer perceptron represents a type of ANN that consists of at least three layers of artificial neurons: input, hidden and output (Figure 8).

Figure 8. Input vectors of ANN with n neurons in hidden layer.

Neurons between input/hidden and hidden/output layers are connected by weighted connections that produce output based on the feature vectors, weights of the connections and non-linear activation function (Figure 9).

Figure 9. Artificial neuron.

The learning process starts with random initialization of weights and calculation of the ANN output based on the first feature vector from the training data set as an input. Given the input values X , weights w , and activation function f , the ANN algorithm calculates the output y of each neuron using [23]:

$$y = f(X_1w_1 + X_2w_2 + \dots + X_nw_n) \quad (4)$$

The activation function f introduces non-linearity to the weighted sum of the neuron's connections to better fit the real-world data. Some commonly used activation functions are sigmoid, tanh and ReLU [24]:

1. sigmoid

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (5)$$

2. tanh

$$f(x) = 2\sigma(2x) - 1 \quad (6)$$

3. ReLU

$$f(x) = \max(0, x) \quad (7)$$

We used ReLU because of performance superiority and also to avoid the vanishing gradient problem present with sigmoid and tanh functions [25]. Output o of each neuron i calculated using ReLU activation function is given by:

$$o_i = \max\left(0, \sum_{k=1}^n w_{ki}o_k\right). \quad (8)$$

The ANN algorithm compares the final output of neural network o , for respective feature vector i , with the target value $t(i)$ giving the degree of error err :

$$err_i = t_i - o_i \quad (9)$$

Depending on the difference between target and ANN output value, the algorithm updates the weights through a back-propagation process [23] of minimizing the squared error ε in the output of each neuron i , which is given by:

$$\varepsilon = \sum_i err_i^2, \quad (10)$$

During the learning process, MLP iteratively adjusts the weights of these connections using the back-propagation process in order to find weights, such that the output value for every input vector in the training set yields the closest value to the target value. The algorithm steers the refinement of the weights by partial derivatives [23] used to decrease the error ε gradually:

$$w'_i = w_i - \eta \frac{\partial \varepsilon_{total}}{\partial w_i}, \quad (11)$$

where w'_i is the new weight of the i -th neuron, while η represents the learning rate.

The back-propagation process repeats until one of three conditions are met: (a) a predefined number of iterations is achieved, (b) error below a predefined value is minimized, or (c) weights do not change significantly between iterations given a chosen threshold value. After training the MLP, the model is ready to predict the outputs for new, previously unseen data input feature vectors.

4. Results

As the FOCUSED algorithm substantially differs from existing wind-speed correction algorithms (described in Section 2: Related Work), we conducted three heterogeneous tests to show its potential. We provide evaluation that brings (a) statistical evaluation that confirms the potential of using previous successive forecasts as input features, (b) comparison with the widely used Autoregressive Moving Average model (ARMA) based on modelling residual errors to correct predictions, and (c) empirical evaluation that shows insights from the traffic manager's point of view, who is the main stakeholder which will benefit from the use of this algorithm.

4.1. Statistical Evaluation of Using Successive Forecasts Instead of the Last Forecast

We tested the algorithm using four data-mining models: ANN, RF, SVM and LR. We compared and evaluated their respective error distributions with Mann-Whitney U Tests [26].

The Mann-Whitney U Test is a nonparametric statistical procedure for comparing two independent, non-related samples. In our case, the last NWP model's FD_6 prediction was compared with each of the four data-mining models predictions, respectively. Input samples were created as lists of percentages of forecast/correction errors compared to actual measured wind speed, calculated as:

$$MW_1 = \frac{FD_6 - sp_{mean}}{sp_{mean}} \quad MW_2 = \frac{MLM - sp_{mean}}{sp_{mean}}, \quad (12)$$

where MW_1 and MW_2 are Mann-Whitney input values, FD_6 is the last NWP model's forecast, MLM is the machine-learning model's corrected forecast value, and sp_{mean} is actual measured wind speed.

Two samples were then combined and sorted in order to determine if the values from the two samples were clustered or randomly mixed. Mann-Whitney U Test statistic for each of the samples was determined by the following formula:

$$U_i = n_1 n_2 + \frac{n_i(n_i + 1)}{2} - \sum R_i \quad (13)$$

where U_i is the test statistic for respective samples, n_1 and n_2 are the length of the MW_1 and MW_2 samples, and R_i represents the sum of ranks for each sample.

The next step was the examination for significance, for which we stated the null and research hypotheses as:

- H_0 (null hypothesis): There is no tendency for ranks of corrected forecasts based on tested models to be significantly different than ranks of the original NWP (FD_6).
- H_1 (research hypothesis): The ranks of corrected forecasts are significantly different than those of the FD_6 .

Our research hypothesis was directional because it indicates the expectation of lower values for the sample related to corrections calculated by FOCUSED algorithm, and higher values for the sample related to NWP FD_6 predictions. We calculated U_1 for FD_6 predictions and U_2 for respective data-mining corrections.

The existence of clustered groups would indicate that there is significant difference between samples, and a random rank order would mean that there is no difference between samples.

After computing the U statistic, we calculated the z-score.

$$z = \frac{U - m_U}{\sigma_U}, \quad (14)$$

where m_U is mean deviation of U given by:

$$m_U = \frac{n_1 n_2}{2}, \quad (15)$$

and σ_U is standard deviation of U given by:

$$\sigma_U = \sqrt{\frac{n_1 n_2}{12} \left((n+1) - \sum_{i=1}^k \frac{t_i^3 - t_i}{n(n+1)} \right)}, \quad (16)$$

where $n = n_1 + n_2$, t_i is the number of subjects sharing rank i , and k is the number of distinct ranks.

The results are shown in Table 2.

Table 2. Results of Mann-Whitney U Test.

Tested Model	<i>p</i> -Value	U ₁ -FOCUSED	U ₂ -FD ₆
FCA_ANN	0.0015	109,164	135,861
FCA_RFR	0.0017	109,366	135,659
FCA_SVM	0.0322	114,195	130,830
FCA_LRM	0.0348	114,348	130,677

Input vectors are expressed as percentage error differences between actual and predicted values. Therefore, lower U-values are considered better.

4.2. Statistical Comparison with Autoregressive Moving Average (ARMA) Model

The Autoregression (AR) method models the next step in the sequence as a linear function of the observations at prior time steps. The Moving Average (MA) method models the next step in the sequence as a linear function of the residual errors from a mean process at prior time steps. The Autoregressive Moving Average (ARMA) method models the next step in the sequence as a linear function of the observations and residual errors at prior time steps. It combines both Autoregression (AR) and Moving Average (MA) models.

The results of the Mann-Whitney U Test comparison of FOCUSED and ARMA are shown in Table 3.

Table 3. Results of Mann-Whitney U test comparison of FOCUSED and ARMA models.

Tested Algorithm	<i>p</i> -Value	U ₁ -FOCUSED	U ₂ -ARMA
FOCUSED	0.0015	109,164	135,861
ARMA	0.0139	113,076	132,940

Input vectors are expressed the same way as in the statistical evaluation.

Finally, we rejected the null hypothesis because statistical significance of the algorithm's performance in combination with all of tested data-mining models was confirmed by *p*-values lower than 0.05. The rejection of the null hypothesis and the differences between U₁ and U₂ values indicate that errors of FOCUSED corrections are systematically ranked lower than FD₆ forecasts and ARMA correction method applied to the last received forecast.

4.3. Empirical Evaluation

Here we used past hourly averaged data for five consecutive years (2011–2016) with the following parameter values:

$$T_s = 12 \text{ h}, F_p = 72 \text{ h}, T_p = 6 \text{ h}$$

To measure the difference between the original and corrected forecasts, we used RMSE (Root Mean Square Error):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (v_c(t) - v_m(t))^2}{n}}, \quad (17)$$

where n is the length of prediction, $v_c(t)$ is the corrected wind speed at time t , and $v_m(t)$ is the measured wind speed at time t .

To demonstrate the algorithm, we observed 53 wind situations considered dangerous for the road traffic over a period of five years. The modelling results for each sample are categorized according to the calculated RMSE values, as shown in Table 4.

Table 4. Algorithm’s performance categories.

Category	Description
Better	RMSE FCA lower than any RMSE of original NWP forecasts
Comparable	RMSE FCA between lowest and highest RMSE of original NWP forecasts
Worse	RMSE FCA higher than any RMSE of original NWP forecasts

The granularity of the time series was one hour. The number of respective samples per strong wind situation varied depending on its duration, resulting in total of 1691 modelling samples (hours) distributed as shown in Figure 10.

Figure 10. Distribution of the results of the correction algorithm applied to a real five-year data set from Croatian bridge Krk.

Figure 11 shows an example of the “Better” category sample. We present the first 20 h of six consecutive forecasts annotated with grey lines and increasing numbers. These forecasts are the source for the ANN. The oldest forecast is marked as FD_1 and presented with the light grey line. It refers to the forecast created 72 h ago. The next one, created 60 h ago, is marked FD_2 and is a bit darker than FD_1 . The last (most shaded grey) forecast, marked as FD_6 , represents the newest result of the ALADIN model. FC represents the direct output of the ANN model. A challenge has arisen as the prediction of the model revealed the most probable shape of the final forecast, but the overall wind speed was often shifted across ordinate of the graph, causing increased RMSE. To correct that, we aligned the curve’s wind speed starting point with the last measured data point. The FCA represents the aligned (final) corrected forecast.

Figure 11. An example of the corrected 12 h forecast.

Figures 11 and 12 show that FCA follows the actual speed curve more tightly than any of the previous FD_1 – FD_6 forecasts.

Figure 12. An example of improved accuracy: the final correction (RMSE_FCA) has better accuracy than any of the previous forecasts.

Consequently, the algorithm outperformed the last ALADIN forecast as expected. The improvement of RMSE in the above example is confirmed through comparison of actual speed with corrected forecast (Figure 12), showing that the result of the proposed algorithm surpasses the accuracy of FD_6 (and, in this case, all previous forecasts).

This example is also interesting because the oldest prediction (72 h old, denoted as $RMSE_{FD1}$) is more accurate than the newest one (denoted as $RMSE_{FD6}$). The comparison of RMSE in the example of bora (Figure 11) shows the strength of the proposed algorithm to reduce the overall error of original forecasts, regardless of their order.

Figure 13 shows an example of the “Worse” category. Such significant discrepancies sometimes occur during the first hours after a sudden large jump of measured values in a short time, or when a strong wind situation occurs immediately after a long period of constant slow wind speed. Shortly after that sudden wind speed change (typically 1–3 h), results of the correction algorithm return to the “Better” or “Comparable” category again. We expect to minimize these discrepancies in further research by comparing forecasted wind speed curves shape with actual wind speed curve shape and minimizing the shape

difference errors using time series transformations available in the REFII (Raise-Equal-Fall model Version II) methodology for holistic time series analysis, based on a time series transformation model. [27,28].

Figure 13. An example of the “Worse” category of the corrected forecast at the moment of a sudden large speed jump.

When we visualize the last NWP forecast (FD_6) and the corrected forecast (FCA) during a longer period, we can spot better prevailing accuracy of FCA for one year. Figure 14 shows an example for year 2016—the comparison of FD_6 (the newest and considered the most relevant before correction) and FCA for all training samples longer than or equal to 12 h. Comparison is made for all observation points (hours) in the whole data set, using only data from the 12th hour after the bora start (minimal training sample size is 12). The figure also shows relatively rare “Worse” situations, represented as high peaks on the graph. Those periods are typically short and followed by fast stabilization, resulting in further improvement of the model’s accuracy.

Figure 14. Comparison of hourly RMSE between corrected and uncorrected forecasts for one year. A reliable algorithm should fit the FCA curve inside FD_6 (grey) area most of the time.

The comparison of the results achieved with the FD_1 – FD_6 and FCA forecasts shows that the FCA outperforms all original forecasts (FD_1 – FD_6), as presented in Figure 15.

Figure 15. In the far majority of cases, FCA is the best compared to successive ALADIN forecasts.

5. Conclusions and Future Work

In this paper, we present a short-term wind speed forecast correction algorithm. It uses previous successive forecasts as input features to an ANN and outputs the corrected value. We have shown that our algorithm outperforms all of the successive original forecasts in the majority of cases in a total of 1691 observed samples. In 49.97% of cases, our algorithm outperformed all of the original forecasts. In 31.82% of cases, the results were inside the range of the original forecasts. This research also shows that the newest forecast created with the NWP model is not always the best one in comparison with forecasts generated previously by the same model for the same prediction interval of 12 h, even though in practice it is used as if it were.

However, empirical evaluation of the algorithm also exposed significant degradations of the forecast accuracy in 18.21% cases, mostly after an abrupt change of measured values or when strong gusts occur immediately after a long period of relatively low wind speed. Since the correction of multiple previous forecasts for a specific time may depend on the congruence between the shape of the forecasted and measured data, we expect to solve these exceptional cases in further research by combining various segments of the time series transformation model REFII. As part of the future work, the presented algorithm will be further improved by adding non-meteorological input parameters, avoiding utilization of any input variables used during the creation of the original series of forecasts in order to keep the correction as non-biased as possible.

Author Contributions: Conceptualization, Z.K.; Formal analysis, Z.K.; Methodology, Z.K., B.Ž. and B.M.B.; Software, Z.K.; Supervision, B.M.B.; Validation, B.Ž. and B.M.B.; Writing—original draft, Z.K.; Writing—review & editing, Z.K., B.Ž. and B.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: The work of the last authors was partially supported by the Slovenian Research Agency (ARRS) core research programmes (P1-0383 and P2-0103).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from Croatian Meteorological and Hydrological Service. Restrictions apply to the availability of these data, which were used under license for this study. The procedure and request form to access weather data are available at https://meteo.hr/proizvodi_e.php?section=proizvodi_usluge¶m=services, accessed on 31 March 2021.

Acknowledgments: Authors would like to thank the Croatian Meteorological and Hydrological Service, and Croatian Motorways Ltd. for providing the data sets for this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bajić, A. Bora wind and road traffic safety. In Proceedings of the Fourth Croatian Road Maintenance Conference; Gospodarsko interesno udruženje trgovačkih društava za održavanje cesta Hrvatski cestar: Šibenik, Croatia, 2009; pp. 307–312.

2. Bajić, A.; Ivatek-Šahdan, S.; Žibrat, Z. Anemo-alarm operational use of wind speed and direction forecast. In Proceedings of the GIU Hrvatski cestar Zagreb, Third Croatian Road Maintenance Conference, Šibenik, Croatia, 20–22 October 2008; pp. 109–114.
3. Wang, X.; Guo, P.; Huang, X. A review of wind power forecasting models. *Energy Procedia* **2011**, *12*, 770–778. [[CrossRef](#)]
4. ALADIN International Team The ALADIN Project: Mesoscale modelling seen as a basic tool for weather forecasting and atmospheric research. *WMO Bull.* **1997**, *46*, 317–324.
5. Žibrat, Z.; Tomšić, D.; Jakopović, Z.; Kunić, Z. Meteorological measuring systems and software in the network of automatic weather stations in Meteorological and hydrological service of the Republic of Croatia. *Hrvat. Meteoroloski Cas.* **2012**, *46*, 69–84.
6. Zjavka, L. Wind speed forecast correction models using polynomial neural networks. *Renew. Energy* **2015**, *83*, 998–1006. [[CrossRef](#)]
7. Aime, S.; Shuang, X.I.; Chongjian, Q.I.U. A variational method for correcting non-systematic errors in numerical weather prediction. *Sci. China Ser. D Earth Sci.* **2009**, *52*, 1650–1660.
8. Welch, G.; Bishop, G. *An Introduction to the Kalman Filter*; UNC: Chapel Hill, NC, USA, 2006; Volume 7, pp. 1–16.
9. Libonati, R.; Trigo, I.; DaCamara, C.C. Correction of 2 m-temperature forecasts using Kalman Filtering technique. *Atmos. Res.* **2008**, *87*, 183–197. [[CrossRef](#)]
10. Sweeney, C.P.; Lynch, P.; Nolan, P. Reducing errors of wind speed forecasts by an optimal combination of post-processing methods. *Meteorol. Appl.* **2013**, *20*, 32–40. [[CrossRef](#)]
11. Zjavka, L. “Aladin” weather model local revisions using the differential polynomial neural network. *Neural Netw. World* **2014**, *24*, 143–156. [[CrossRef](#)]
12. Shukur, O.B.; Lee, M.H. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew. Energy* **2015**, *76*, 637–647. [[CrossRef](#)]
13. Louka, P.; Galanis, G.; Siebert, N.; Kariniotakis, G.; Katsafados, P.; Pytharoulis, I.; Kallos, G. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J. Wind Eng. Ind. Aerodyn.* **2008**, *96*, 2348–2362. [[CrossRef](#)]
14. Liang, Z.; Liang, J.; Wang, C.; Dong, X.; Miao, X. Short-term wind power combined forecasting based on error forecast correction. *Energy Convers. Manag.* **2016**, *119*, 215–226. [[CrossRef](#)]
15. Ko, W.; Hur, D.; Park, J. Journal of International Council on Electrical Engineering Correction of wind power forecasting by considering wind speed forecast error. *J. Int. Counc. Electr. Eng.* **2015**, *5*, 47–50. [[CrossRef](#)]
16. Nan, X.; Li, Q.; Qiu, D.; Zhao, Y.; Guo, X. Short-term wind speed syntheses correcting forecasting model and its application. *Int. J. Electr. Power Energy Syst.* **2013**, *49*, 264–268. [[CrossRef](#)]
17. Jiang, Y.; Huang, G. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. *Energy Convers. Manag.* **2017**, *144*, 340–350. [[CrossRef](#)]
18. Dong, L.; Ren, L.; Gao, S.; Gao, Y.; Liao, X. Studies on wind farms ultra-short term NWP wind speed correction methods. In Proceedings of the 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, 25–27 May 2013; pp. 1576–1579.
19. Xue, H.L.; Shen, X.S.; Chou, J.F. A forecast error correction method in numerical weather prediction by using recent multiple-time evolution data. *Adv. Atmos. Sci.* **2013**, *30*, 1249–1259. [[CrossRef](#)]
20. Luo, J.; Hong, T.; Fang, S.-C. Benchmarking robustness of load forecasting models under data integrity attacks. *Int. J. Forecast.* **2018**, *34*, 89–104. [[CrossRef](#)]
21. Zhang, Y.; Lin, F.; Wang, K. Robustness of Short-Term Wind Power Forecasting against False Data Injection Attacks. *Energies* **2020**, *13*, 3780. [[CrossRef](#)]
22. Bajić, A. Gale-force wind in Croatia. In Proceedings of the Zbornik Radova s 2. Konferencije Hrvatske Platforme za Smanjenje Rizika od Katastrofa; Državna Uprava Za Zaštitu i Spašavanje: Zagreb, Croatia, 2011; pp. 141–147.
23. da Silva, I.N.; Hernane Spatti, D.; Andrade Flauzino, R.; Liboni, L.H.B.; dos Reis Alves, S.F. *Artificial Neural Networks*; Springer International Publishing: Cham, Switzerland, 2017; ISBN 978-3-319-43161-1.
24. Sharma, S.; Sharma, S.; Anidhya, A. Understanding Activation Functions in Neural Networks. *Int. J. Eng. Appl. Sci. Technol.* **2017**, *4*, 310–316.
25. Zhao, G.; Zhang, Z.; Guan, H.; Tang, P.; Wang, J. Rethinking ReLU to Train Better CNNs. In Proceedings of the 2018 24th the International Conference Pattern Recognition, Beijing, China, 20–24 August 2018; pp. 603–608.
26. Mann, H.B.; Whitney, D.R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [[CrossRef](#)]
27. Klepac, G. Otkrivanje Zakonitosti Temeljem Jedinstvenoga Modela Transformacije Vremenske Serije, Doktorska Disertacija. Ph.D. Thesis, Sveučilište u Zagrebu, Fakultet Organizacije i Informatike Varaždin, Varaždin, Croatia, 2005.
28. Klepac, G.; Kopal, R.; Mršić, L. *REFII Model as a Base for Data Mining Techniques Hybridization with Purpose of Time Series Pattern Recognition*; Springer: New Delhi, India, 2016; pp. 237–270.