Short Time Wind Forecasting with Uncertainty

Petros Karvelis, Stefanos Petsios, George Georgoulas and Chrysostomos Stylios, *Member, IEEE*

Abstract— **Forecasting the weather and especially the wind is important for a number of applications like wind farms or for maritime operations. Nowadays machine learning techniques are becoming more reliable and robust for forecasting due to the fact that a plethora of available datasets exist. However, forecasts for shorter time horizon less than two hour is not reliable due to the frequent wind fluctuations. Nevertheless, the need for algorithms that can have a small memory and cpu footprint is needed for hardware e.g. microcontrollers that are on board of vessels. In this manuscript a method for short time wind forecasting is proposed and scaled for a microcontroller. The method also computes prediction intervals with a certain probability. Our method was tested using real data recorded from a weather station on board of a ship conducting trips across the Aegean Sea (Greece).**

Keywords—weather forecasting, regression, multiple linear regression, prediction intervals.

I. INTRODUCTION

Maritime plays an important role not only for the Europe economy but also for the whole world. More and more goods are transferred by vessels. However, one of the factors that influence maritime operations is the weather. On the other hand, wind speed is another factor with high influence on weather and the reason for that is that it affects the surface water and waves [1], [2].

Wind speed can be predicted using different time horizons e.g. next few minutes, hours, or days. Four different horizons are used for the for weather forecast [3]: a) now cast: where we try to predict the current state of the weather, b) short-term: which involve 72 hours predictions, c) medium-range – which involve a period of 3-7 days and finally long-term – where weather predictions are made for a longer period e.g. one week to months.

Machine learning techniques can be applied to solve the wind speed forecasting problem [4], [5]. Such weather forecasting techniques either depend on the utilization of generative numerical methods [6], or using time-series analysis such as Auto Regressive Integrated Moving Average (ARIMA) models [7]. Other approaches use Artificial Neural Networks (ANNs) [8], [9] or Support Vector Machines [10], [11]. Nowadays the effective use of approaches relying on Deep Learning for a number of domains have proved their utilization for climate and weather prediction. However, moving from simple models like the ARIMA ones to the Neural Networks a main drawback is the complexity of the model that is used to predict different weather parameters.

The LINCOLN[1] project aims to propose a number of services utilizing an IoT (Internet of Things) platform, which intends to provide predictions to various maritime actors. Specifically, the IoT platform consists of a number of IoT microcontrollers and thus one must pay careful attention on how we chose the model because of the limited computer resources (CPU, memory and storage) available for the implementation of the method.

In this work, we implemented a model that could be embedded in the LINCOLN IoT platform with an adaptive learning process in order to predict as accurately as possible the weather parameters. The adaptive learning process is based on the Stochastic Gradient Descent (SGD) method which pay attention only on a few training data before changing the model, and is known to generalize well to various unseen data [13]. Besides the estimate of the wind speed, the uncertainty associated with the prediction is quantified and upper/lower confidence bounds [14] are also provided which could be used as an alert generator for the skipper.

The manuscript is organized as follows: Section II presents the a) regression model b) the training algorithm of the model and c) the prediction interval computation. Section III, describes the data collection process and in the next section Results of our method are provided. Finally, in the last Section IV some concluding remarks and future work are presented.

II. REGRESSION

When it comes to forecasting meteorological parameters like wind speed, special care must be taken in order to tackle the problem of the local environment and the complexity of these parameters that can seriously affect the performance of the forecasting methods.

On the other hand, regression is a well-known machine learning problem. Mainly, it is used for the prediction of the value of a variable (output) provided the values of a number of variables (inputs). We are provided with a set of training points $(x_1, \mathrm{y}_1), (x_2, \mathrm{y}_2), ..., (x_n, \mathrm{y}_n)$ where $x_i \in \Re^m$, $y_i \in \Re$, $1 \le i \le n$ and the task is to estimate a function which best fits the data subject to an error function. Once such a function is found we can predict the output value based on the values of the unknown point.

In this study we have used as input to our regression algorithm the values of the Temperature, Dew point temperature, Humidity, Wind direction, Pressure,

P. Karvelis, S. Petsios and C. Stylios are with Computer Technology Institute & Press "Diophantus" Patras, Greece, (e-mail:{pkarvelis@gmail.com, stefanos.petsios@gmail.com, stylios@cti.gr).

G. Georgoulas is with the Robotics Group, Control Engineering Division of the Department of Computer, Electrical and Space Engineering, Lulea University of Technology, Lulea 97187, Sweden (e-mail:georgoul@ltu.se).

[1] LINCOLN: Lean Innovative Connected Vessels Project www.lincolnproject.eu. Horizon 2020 research and innovation program

Precipitation and the goal is to predict the value of the Wind Speed.

A. Multiple Linear Regression

Multiple Linear Regression (MLR), is a linear model [16] where we seek to predict the output value (wind speed) as a linear combination of the inputs (Temperature, Dew point temperature, Humidity, Wind direction, Pressure, Precipitation):

$$h_\theta(x_i) = \hat{y}_i = \theta_0 + \theta_1 x_1^1 + \ldots + \theta_m x_1^m = \sum_{j=0}^{m} \theta_i x_i^j \ , \qquad (1)$$

where $x_0 = 1$, $\boldsymbol{\theta} = \{\theta_{\boldsymbol{i}} \mid \boldsymbol{i} = 0,\ldots,\mathrm{m}\}$ are the parameters of the linear function and $\hat{y}_i$ the predicted value.

In order to learn the values of the parameters of the linear function an error function must be defined between the output $h_\theta(x)$ and the actual value y_i. A common function that is used is the least square error function defined as:

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{n}\left(h_\theta(x_i) - y_i\right)^2 , \qquad (2)$$

B. Stochastic Gradient Descent

One of the most widely used methods for iterative minimization would be the Gradient Descent (GD) method. The parameters $\boldsymbol{\theta} = \{\theta_{\boldsymbol{i}} \mid \boldsymbol{i} = 0,\ldots,\mathrm{m}\}$ are initialized at some random values and then the GD method repeatedly changes these values in order to reach a minimum value for the error function $J(\theta)$. The following update is employed for the parameters:

$$\theta_j^{new} = \theta_j^{old} - l \cdot \frac{\partial J(\theta)}{\partial \theta_j},\ 1 \le j \le m \qquad (3)$$

where θ_j^{new}, θ_j^{old} is the new and old value respectively of the j-th parameter, l is a learning parameter which control the rate or convergence of the algorithm and is usually set to values $10^{-3} \le l \le 10^{-2}$. Lower values of this hyperparameter leads to slower convergence.

The update on the parameters is simultaneously performed for all values of $j = 1,\ldots,m$ using the whole training dataset $i = 1,\ldots,\mathrm{n}$. However, special care must be taken for the case of embedded environments or microcontrollers as in our case. All this hardware offers minimum processing power and memory and thus makes the work of deploying machine learning algorithms extremely difficult.

A way to overcome all these limitations would be to update the parameters $\boldsymbol{\theta} = \{\theta_{\boldsymbol{i}} \mid \boldsymbol{i} = 0,\ldots,\mathrm{m}\}$ as the new data are received from the weather station. This type of training is called online training and the algorithm that were used is called Stochastic Gradient Descent [17].

SGD computes the gradient based only on a few number of training points e.g. for a pair $(x_i.y_i)$ from the training set and update the parameters based on:

$$\theta_j^{new} = \theta_j^{old} - l \cdot \frac{\partial J(\theta; x_i.y_i)}{\partial \theta_j} , \qquad (4)$$

where the learning parameter l is set to a smaller value than in the case of GD algorithm.

C. Prediction Intervals

For a prognostic task, besides the point predictions, it is useful to provide upper and lower confidence bounds of the predicted value. A way to compute these prediction intervals is through quantile regression which aims to assess the conditional quantiles from the observed data [18].

Suppose a new input value x_0 and the predicted value by the MLR model will be $\hat{y}_0$. The $1-a$ prediction interval of $\hat{y}_0$ is therefore [16]:

$$\hat{y}_0 \pm t_{crit} \cdot \sqrt{S\left(1 + x_0^T \left(X^T X\right)^{-1} x_0\right)} \qquad (5)$$

where t_{crit} is the critical value of the t-distribution with $df = n - m - 1$ degrees of freedom with significance level $a/2$ and

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^2 & \ldots & x_1^m & 1 \\ x_2^1 & x_2^2 & \ldots & x_2^m & 1 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ x_n^1 & x_n^2 & \ldots & x_n^m & 1 \end{bmatrix}, \qquad (6)$$

$$S = \frac{1}{n-m}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 . \qquad (7)$$

III. Dataset

The source of the data gathering was a weather station installed on a Tug boat [19]. The boat was chosen as it was frequently conducting sea trips (currently in the Aegean Sea) so that the method is tested on real sea conditions. In the Tug boat we have installed the weather station Airmar 220WX [20] and we have collected data using the NMEA 2000 channel [21]. The NMEA 2000 channel is a single network cable which accommodates navigation equipment, electrical power generation and distribution systems, engines and other machinery, piloting and steering systems, fire and other alarms, and controls. Commands data as also status of various sensors share the same cable at speed which is almost 26 times greater than the classic serial cable. There is no need for a master controller and extra equipment or sensors can be hooked up or even removed without shutting down the network. A figure of the weather stations is presented below.

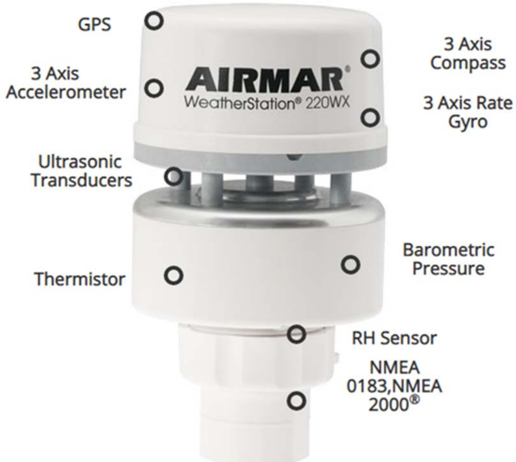

Figure 1. The Airmar 220WX weather Station.

The weather station can measure not only the apparent wind speed and direction but also the true wind speed and direction. Virtually all mechanical and ultrasonic anemometers report apparent wind speed and direction. The Airmar WX Series is unique because it calculates both theoretical and apparent wind speed and direction. These wind readings are the same if the unit is mounted in a fixed location. The weather station receives various weather parameters which are described in the following Table. The experimental data set consists of 1-hour records.

TABLE I. METEOROLOGICAL PARAMETERS WHICH ARE PROVIDED BY THE WEATHER STATION

Measurement	**Unit of measure**	**Data source**	**Frequency**	**Typical operation range**
True Wind Speed	m/s	NMEA 2000	1 hour	[0, 60] m/s
True Wind Direction	deg	NMEA 2000	1 hour	[0, 360°]
Apparent Wind speed	m/s	NMEA 2000	1 hour	[0, 60] m/s
Apparent Wind Direction	deg	NMEA 2000	1 hour	[0, 360°]
Air temperature	oC	NMEA 2000	1 hour	[-30, +60°C]
Barometric Pressure	kPa	NMEA 2000	1 hour	[540, 1100] hPa
Relative humidity	%	NMEA 2000	1 hour	[0%, 100%]
Global Positioning System (GPS)	Degrees	NMEA 2000	1 hour	
Time / Timestamp from GPS	time unit	NMEA 2000	1 hour	-

The data we collected were recorded from the tug boat from 22/1/2018 – 30/3/2018. We have divided the initial dataset into two datasets namely DataSet A: a trip from Chios island (Greece) to Athens and the second Dataset B: a trip from Athens to Chios island and back. The reason for doing this is to present the results of our method for two different sea trips of the Tug boat. The profile of the two Datasets/Trips are described in the following Table.

TABLE II. THE PROFILE OF THE TRIPS TRAVVLED BY THE TUG BOAT

	Date	**Days**	**Distance (Km)**
Dataset A	22/1/2018 to 9/2/2018	18	563
Dataset B	9/2/2018 to 30/3/2018	50	818

IV. RESULTS

In order to train our method for each dataset we have used the first two days data for each dataset (e.g. 48 recordings) and for each new sample received from the weather station, the method predicted the wind speed for the next 1-hour. Finally, we have used the Median Absolute Error (MAE) and the Root Mean Square Error ($RMSE$) which are defined below as [22], [23] in order to test our method:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| , \tag{8}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2} , \tag{9}$$

where N is the number of the testing samples, $\hat{y}_i$ is the predicted value and y_i the true wind speed value.

In the following Table we present our results with respect to the error rates for the two Datasets.

TABLE III. MAE AND $RMSE$ ERROR MEASURES FOR THE TWO DATASETS.

	MAE	$RMSE$
Dataset A	2.19	2.04
Dataset B	1.70	1.50

The following two figures display the predicted and the real value of the wind for 90% prediction intervals. As one can easily observe all the predicted values fall between the 90% prediction interval.

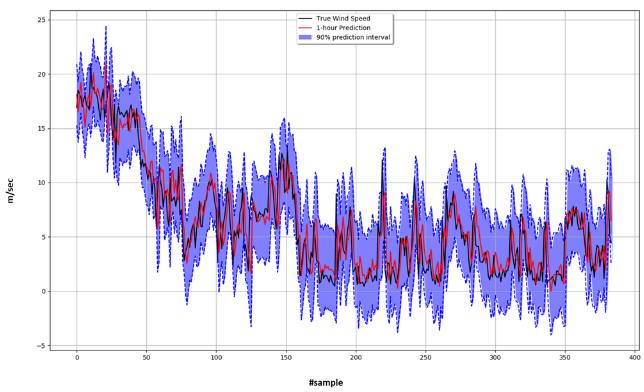

Figure 2. Performance of our method over test data (Dataset A) for 1-hour prediction.

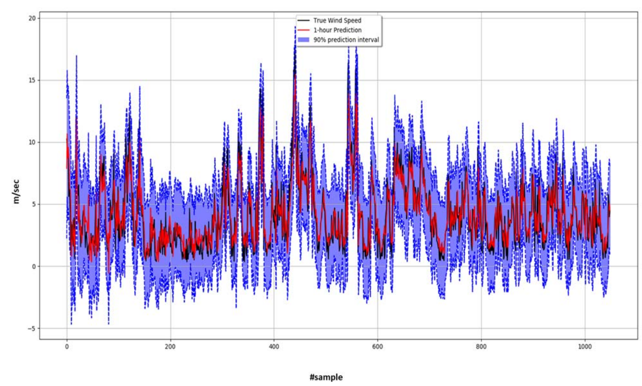

Figure 3. Performance of our method over test data (Dataset B) for 1-hour prediction.

V. Conclusion

A method for the prediction of the wind speed has been presented and evaluated for real weather data acquired in the Aegean Sea. The training method of the regression model is based on the stochastic gradient descent algorithm. The choice of the SGD algorithm makes the deployment of the method for microcontrollers suitable since only a few data points are needed and kept in the memory of the microcontroller. The computation of the prediction interval is advantageous for regression problems since the prediction values are provided based on a predefined probability threshold. As for feature work we plan to test our method bigger datasets and evaluated for different time lags.

Acknowledgment

This study is supported by the National funds allocated by "LINCOLN" (Lean Innovative Connected Vessels) Project (www.lincolnproject.eu) Horizon 2020 research and innovation program (Grant Agreement: 727982).

References

[1] A. K. Gupta, S. K. Gupta and R. Patil, "Environmental management plan for port and harbour projects," Clean Technologies and Environmental Policy, vol.7(2), pp. 133–141, 2005.

[2] Guide to Wave Analysis and Forecasting, World Meteorological Organization (WMO), available online (https://library.wmo.int/index.php?lvl=notice_display&id=7700#.XLMwgOgza70)

[3] M.O Mandal, P. Soman, S. Zareipour , "A review of wind power and wind speed forecasting methods with different time horizons." In: Proceedings of the 2010 north american power symposium (NAPS). 2010, pp. 1–8.

[4] P. Karvelis, S. Kolios, G. Georgoulas, and C. D. Stylios, "Ensemble Learning for Forecasting Main Meteorological Parameters," In: Proceedings of 2017 IEEE International Conference on Systems, Man and Cybernetics (SMC)), 5-8 October 2017, Banff Center, Banff, Canada, pp. 3711-3714 (2017).

[5] G. Georgoulas, P. Karvelis, S. Kolios and C. D. Stylios, "Examining nominal and ordinal classifiers for forecasting wind speed," In: Proceedings of 2016 IEEE 8th International Conference on Intelligent Systems (IS'16), 3-5 September 2016, Sofia, Bulgaria. pp. 504-509 (2016).

[6] G. Marchuk. Numerical methods in weather prediction. Elsevier, 2012.

[7] H. Saima, J. Jaafar, S. Belhaouari, T. Perak and T.A. Jillani, "Intelligent methods for weather forecasting: A review," In Proceedings of the National Postgraduate Conference, (Tronoh Perak, Malaysia) Sept. 19-20, 2011.

[8] V. Krasnopolsky, and M. FoxRabinovitz. "Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction." Neural Networks, vol. 19, pp. 122-134 , (2006).

[9] L. Lai, Lei, et al. "Intelligent weather forecast," Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on. Vol. 7. IEEE, 2004.

[10] N. I. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey," Computational Intelligence Magazine, IEEE, vol. 4, pp. 24–38, 2009.

[11] Y. Radhika and M. Shashi "Atmospheric temperature prediction using support vector machines," International Journal of Computer Theory and Engineering, vol. 1, pp. 1793–8201, 2009

[12] J. N. K Liu, Y. Hu, Y. He, P. W. Chan and I. Lai. "Deep Neural Network Modeling for Big Data Weather Forecasting," Information Granularity, Big Data, and Computational Intelligence. Springer International Publishing, pp. 389-408, 2015.

[13] L. Bottou, F. Curtis and Jorge Nocedal. Optimization methods for large-scale machine learning. SIAM Review, 60(2):223–311, 2018.

[14] R. Koenker and H. Hallock, "Quantile regression," J. Econ Perspect, 5: 143–156, 2001.

[15] J. Friedman, T. Hastie and R. Tibshirani, The Elements of Statistical Learning, Data Mining, Inference and Prediction, 2nd Ed, Elsevier, 2009."

[16] G. Seber and A. Lee, Linear Regression Analysis, Whilley, 2003.

[17] L. Bottou and LeCun Yann, "Large Scale Online Learning, Advances in Neural Information Processing Systems, 16, 2016.

[18] N. Meinshausen, Quantile Regression Forests, Journal of Machine Learning Research, vol. 7, pp. 983–999, 2006.

[19] Tug Boat, https://en.wikipedia.org/wiki/Tugboat, (accessed June 2019).

[20] Airmar Technology Corporation, http://www.airmartechnology.com, (accessed June 2019).

[21] NMEA 2000, https://www.nmea.org, (accessed June 2019).

[22] A. Saxena, J. Celaya, B. Saha,S. Saha and K. Goebel, "Metrics for offlineevaluation of prognostic performance," International Journal of Prognostics and Health Management, vol. 1, no. 1, pp.4-23, 2010.

[23] S. Makridakis, S. Wheelwright and R. Hyndman, Forecasting Methods and Applications, 3rd Edition, Wiley, 2012.