Online Bayesian Learning with Natural Sequential Prior Distribution Used for Wind Speed Prediction

Nawal Cheggaga

Department of electronics, Faculty of Technology, Saad Dahlab University, Blida, Algeria,
Laboratory of Electrical Systems and Remote control. P O Box 270 route de Soumaa ,Blida
09000, Algeria
cheggaga@yahoo.fr

Abstract

Predicting wind speed is one of the most important and critic tasks in a wind farm. All approaches, which directly describe the stochastic dynamics of the meteorological data are facing problems related to the nature of its non-Gaussian statistics and the presence of seasonal effects .In this paper, Online Bayesian learning has been successfully applied to online learning for three-layer perceptron's used for wind speed prediction. First a conventional transition model based on the squared norm of the difference between the current parameter vector and the previous parameter vector has been used. We noticed that the transition model does not adequately consider the difference between the current and the previous wind speed measurement. To adequately consider this difference, we use a natural sequential prior. The proposed transition model uses a Fisher information matrix to consider the difference between the observation models more naturally. The obtained results showed a good agreement between both series, measured and predicted. The mean relative error over the whole data set is not exceeding 5 %.

Keywords: artificial neural network, Bayesian learning, Fisher information, online learning , wind speed forecasting.

Introduction

Due to the high penetration of wind power in the electricity system, the forecast accuracy of wind power prediction systems becomes increasingly important. In recent years, many scholars have done a lot of research on wind power prediction. The forecast accuracy has improved constantly, and it can be expected that intense research and development efforts are already on track.

Many researches were achieved in order to predict wind behaviour. However, it is still one of the most difficult quantities to forecast [1], namely due its stochastic nature. The actual state of the art includes five main families of methods: persistence Method [2], physical Methods[3], spatial Correlation Models[5], artificial Intelligence Methods[6] and hybrid Methods[7].However, there will always be an inherent and irreducible uncertainty in every prediction.

Bossanyi [8] used a Kalman Filter with the last 6 values as input and got up to 10% improvement in the RMS error over persistence for 1-min averaged data for the prediction of the next time step. This improvement decreased for longer averages, and disappeared completely for 1-hourly averages.

Dambrosio and Fortunato [9] used a one-step-ahead adaptive control by means of a recursive

least squares algorithm for the electrical part of the turbine. They show a fast and reliable response to a step in the wind.

Nogaret et al [10] reported that for the control system of a medium size island system, persistent forecasting is best with an average of the last 2 or 3 values, ie 20-30 minutes.   Tantareanu [11] found that ARMA models can perform up to 30% better than persistence for 3-10 steps ahead in 4-sec averages of 2.5Hz-sampled data. Dutton et al [12] used a linear autoregressive model and an adaptive fuzzy logic based model for the cases of Crete and Shetland. They found minor improvements over persistence for a forecasting horizon of 2 hours, but up to 20% in RMS error improvement for 8 hours horizon.   However, for longer horizons, the 95% confidence band contained most of the likely wind speed values, and therefore a meteorological-based approach was deemed more promising on this time scale.

In the same team, Kariniotakis et al [13,14] were testing various methods of forecasting for the Greek island of Crete. These included adaptive linear models, adaptive fuzzy logic models and wavelet based models. Adaptive fuzzy logic based models were installed for on-line operation in the frame of the Joule II project CARE (JOR3-CT96-0119).

Fukuda et al [15] worked on an Auto Regressive model for blade angle optimisation. Using data mining, they found that the use of additional variables was helpful only in December, but not in June.   Hunt and Nason [16] used an analysis of principal components of wavelets derived from wind speed time series for a measure-correlate-predict technique.   In [17] comprehensive comparison study on the application of different artificial neural networks in 1-h-ahead wind speed forecasting is presented. Three types of typical neural networks, namely, adaptive linear element, back propagation, and radial basis function, are investigated. The wind data used are the hourly mean wind speed. The results show that even for the same wind dataset, no single neural network model outperforms others universally in terms of all evaluation metrics. Moreover, the selection of the type of neural networks for best performance is also dependent upon the data sources. Different network structures, learning rates, and inputs are believed to result in different forecast accuracies. Among the optimal models obtained, the relative difference in terms of one particular evaluation metric can be as much as 20%. This indicates the need of generating a single robust and reliable forecast by applying a post-processing method.

For this purpose in this paper, the variation in wind data is formulated as a system identification problem, where the input of the system is the past values (z(t - 1), z (t - 2), z(t - 3), …) of a time series and its desired output zt is the future of one value. So data is generated sequentially from unknown time-varying systems, the problem of approximating the target system by a parameterized probabilistic model via the data sequence is called an online learning problem. A reasonable approach to such problems is online Bayesian approach considered with an observation model P(y(t) |θt) parameterized by time-varying parameter θt ($\in$ IRk) for the behavior of the observation data zt.

Used Data
The dataset of these parameters were recorded in (Algeria), the collected experimental data is the hourly average wind speed. Data have been collected during 2013 at a height of 60m**.**
Figure 1 shows the evolution of the different wind speed time series during April 2013. Therefore, the objective of this work is to predict these data based on the past observations.

Results
There is a lot of scope in meteorological time series prediction and what is needed is the accuracy of forecast. First the performances of a three-layer perceptron in combination with an online Bayesian learning, was comprehensively investigated for forecast value of next hourly average wind speed.
The three-layer perceptron defined as follows is used as the function f (·)

$$f(x_t; \theta_t) = a_{0,t} + \sum_{i=1}^{n_h} a_{i,t} \sigma(x_t; b_{i,t})$$

where the integer nh (> 0) denotes the number of hidden units, and the integer nx (> 0) is the dimension of the input variable xt . The parameter a0,t ($\in$ IR) is the output bias, the parameter ai,t ($\in$ IR)(i = 1,..., nh) describes a weight parameter of the i th hidden unit, and the parameter vector bi,t($\in$ IRnx+1)(i = 1,..., nh) denotes a weight parameter vector connecting from the inputs xt and the input bias to the i th hidden unit.

The observation model use P(z$_t$|θ$_t$) parameterized by time-varying parameter θt ($\in$ IRk) for the behavior of the observation data zt at time t , and a transition model (prior distribution) P(θt |θt−1) to describe the behavior of the time-varying parameter θt .

Transition models P(θt |θt−1) based on the squared norm of differences between the current parameter θt and the previous parameter θt−1, which can be written as

$$P\left(\theta_t | \theta_{t-1}\right) \propto \exp\left(-\frac{\gamma_t}{2}\|\theta_t - \theta_{t-1}\|^2\right) \qquad (1)$$

have been considered [18]–[20]. Here, the variable γt is often called the hyperparameter, which controls the scale of the transition. However, the squared norm ||θt−θt−1||2 does not adequately relate to the difference between the current observation model P(zt |θt) and the previous observation model P(zt−1|θt−1), and the changes between the observation models P(zt |θt) and P(zt−1|θt−1) are highly dependent on the values of θt−1.

In figure 1 The neural network cannot explore or predicts the speed below 1.5m / s or up to 11m / s because the change between the observation models is extremely small at θt−1 = θa, whereas the change between the observation models is large at θt−1 = θb with the transition model (1). We may consider that this phenomenon can happen when the eigenvalues of the Fisher information matrix are small at θt−1 = θa and when the eigenvalues of the Fisher information matrix are large at θt−1 = θb. This phenomenon of the transition model (1) has affect the tracking of the system or data zt .

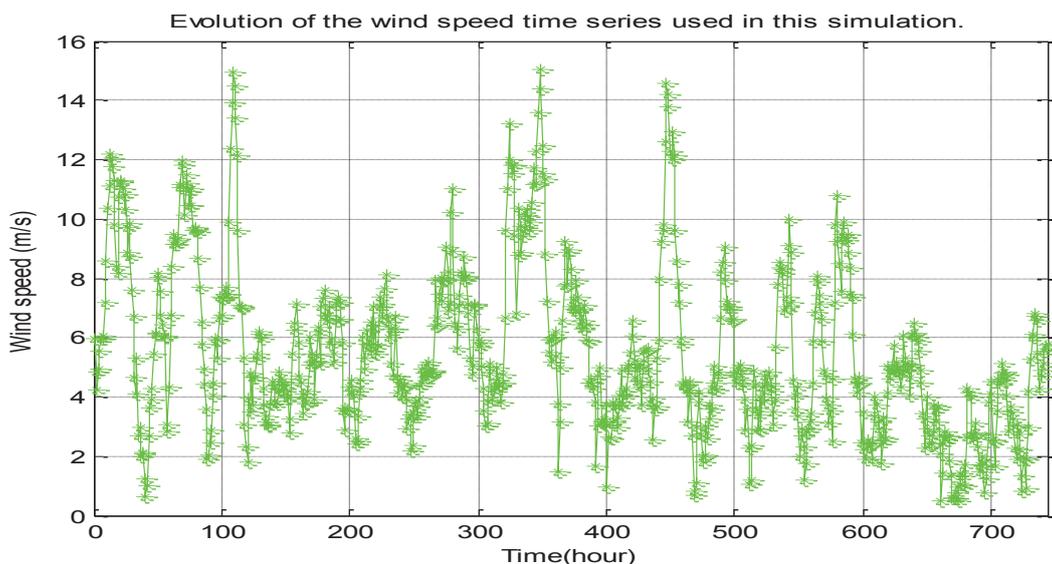Evolution of the wind speed time series used in this simulation.

Fig. 1: Evolution of the wind speed time series used in this simulation.

To break this bridge we will use the same structure of artificial neural network with an online Bayesian learning. the observation data zt generated by an input–output system. The observation model P(zt |θt ,βt)describing the observation data zt will be:

$$P(y_t | x_t, \theta_t, \beta_t) = \frac{1}{Z(\beta_t)} \exp\left(-\frac{\beta_t}{2}(y_t - f(x_t; \theta_t))^2\right) \quad (20)$$

where the variables yt (∈ IR) and xt (∈ IRnx) stand for the output variable and the input variable at time t , respectively, and the observation data zt is considered as the set of these variables xt and yt , i.e., zt = (yt , xt). The function P(xt) describes a probability distribution of the input variable xt. And the online Bayesian approach use natural sequential prior transition model [21]:
to adequately consider the difference between the observation models P(zt |θt) and P(zt−1|θt−1), and to avoid influences on the learning and/or tracking of the target system. Here,θt= θt − θt−1,and T stands for the transpose operator of vectors and matrices. The natural sequential prior uses a Fisherinformation matrix Ft−1

P(z|θt−1)(∂ log P(z|θt−1)/∂θt−1)(∂ log P(z|θt−1)/∂θt−1)T dz derived from the observation model P(zt−1|θt−1), in order to consider the difference between the observation models more naturally. In information geometry approaches, this Fisher information matrix Ft−1(∈ IRk×k) is defined as a metric matrix of a model manifold corresponding to the observation model P(zt−1|θt−1) [11]. The natural sequential prior is based on the norm θtT Ft−1θt, which considers such a metric of the model manifold. Given the observation model in (19), the Fisher information matrix Ft can be written as

$$F_t = \iint P(x_t) P(y_t | x_t, \theta_t, \beta_t)$$
$$\times \frac{\partial \log P(y_t | x_t, \theta_t, \beta_t)}{\partial \theta_t} \frac{\partial \log P(y_t | x_t, \theta_t, \beta_t)}{\partial \theta_t}^T dx_t \, dy_t. \quad (24)$$

by assuming that the variable yt contains Gaussian noise. The variable βt ∈ (0,∞) stands for a hyperparameter describing the inverted variance of the Gaussian noise at time t .

Figures 2 and 3   present the simulation results for generating a single-step (one hour) prediction with natural sequential prior transition model.
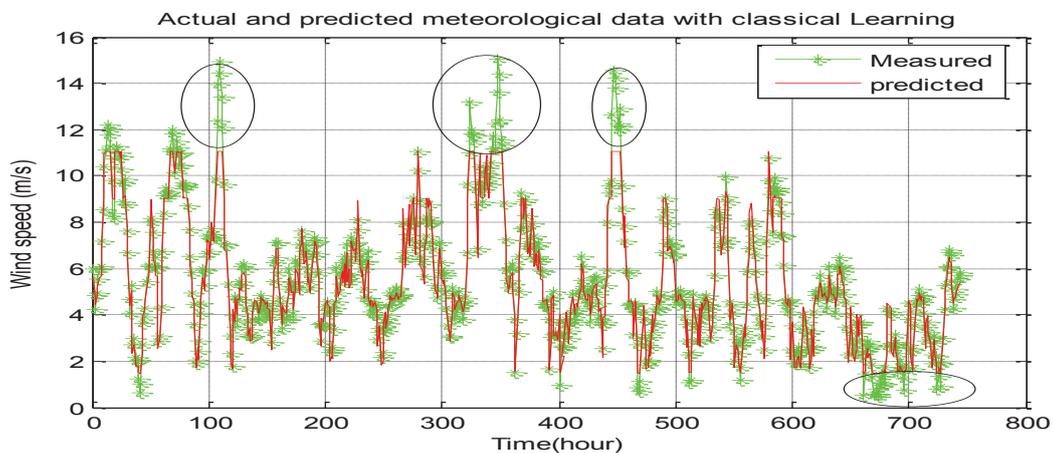

Fig. 2. Actual and predicted wind speed data with classical learning.
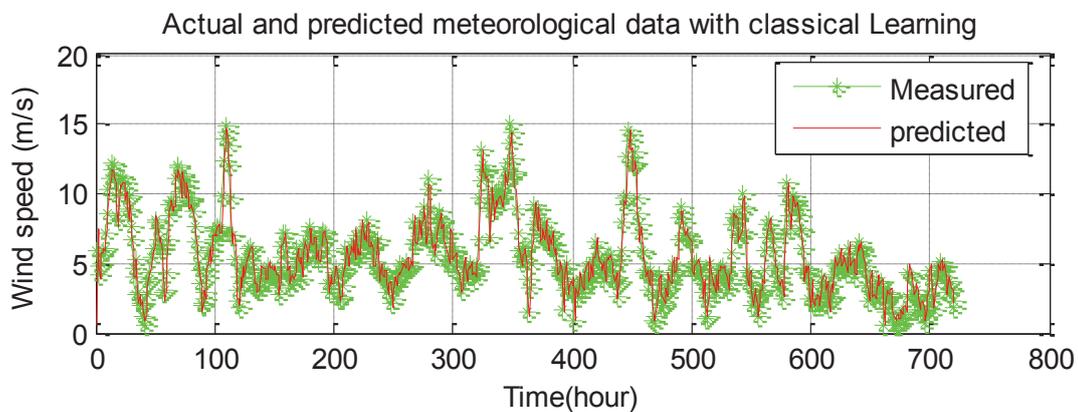

Fig.3. Actual and predicted wind speed data with our approach.

In this work we performed experiments with a step time of one hour. The input network is a sequence of hourly mean values prior. The target output is made of future values.

Root Mean Square the Error (RMS) differences between observed and estimated values, by chosen models, were used to evaluate the performance of models RMSE were computed by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_t - \hat{y}_t)^2} \qquad (9)$$

Where y_t *the original time series,* $\hat{y}t$ *is* is the computed.

The RMS error and Mean relative error has been calculated for the different estimation of the wind speeds and compared at real measured data wind speed. The resultants are given in table1. According to this table it should be noted that the MAE is improved from 21% to 3% and the RMSE is improved from 1.6102 to 0.1892, the coefficient of correlation (r) is improved at 0.996.

Table 1: Comparison between observed and predicted meteorological data

a *models*	*E (%)*	*RMSE*	*r*
Our algorithm predictor	*03%*	*0.1892*	*0.9986*
classical learning	*21%*	*1.6102*	*0.9433*

CONCLUSION

In this paper, a natural sequential prior which uses a Fisher information matrix was applied to the online learning of three-layer perceptron used for wind speed

prediction purpose. As a result, the model consider the difference between the observation models more naturally and confirmed the superiority of the proposed model in terms of errors. One of the most important aspects of future work will be to generalize and/or extend the proposed model to apply it to other renewable energy sources such as solar irradiation and air temperature.

REFERENCES

[1] Lazić L, Pejanović G, Živković M, Wind forecasts for wind power generation using the Eta model, Renewable Energy, Volume 35, Issue 6, June 2010, Pages 1236-1243, ISSN 0960-1481

[2] Zhao, X., Wang, S.X. and Li, T. (2011) Review of Evaluation Criteria and Main Methods of Wind Power Forecasting.

[3] Ma, L., Luan, S.Y., Jiang, C.W., Liu, H L. and Zhang, Y. (2009) A Review on the Forecasting of Wind Speed and Generated Power. Renewable and Sustainable Energy Reviews, 13, 915-920. http://dx.doi.org/10.1016/j.rser.2008.02.002

[4] Saima, H.; Jaafar, J.; Belhaouari, S.; Jillani, T. A., "Intelligent methods for weather forecasting: A review," National Postgraduate Conference (NPC), 2011 , vol., no., pp.1,6, 19-20 Sept. 2011 Energy Procedia, 12, 761-769. http://dx.doi.org/10.1016/j.egypro.2011.10.102

[5] Ma, L., Luan, S.Y., Jiang, C.W., Liu, H L. and Zhang, Y. (2009) A Review on the Forecasting of Wind Speed and Generated Power. Renewable and Sustainable Energy Reviews, 13, 915-920. http://dx.doi.org/10.1016/j.rser.2008.02.002

[6] Ma Lei, Luan Shiyan, Jiang Chuanwen, Liu Hongling, Zhang Yan, A review on the forecasting of wind speed and generated power, Renewable and Sustainable Energy Reviews, Volume 13, Issue 4, May 2009, Pages 915-920, ISSN 1364-0321

[7] G. Jianming Hu, Jianzhou Wang, Guowei Zeng, A hybrid forecasting approach applied to wind speed time series, Renewable Energy, Volume 60, December 2013, Pages 185-194, ISSN 0960-1481

[8] Bossanyi, E.A.: Short-Term Wind Prediction Using Kalman Filters. Wind Engineering 9(1), pp. 1-8 (1985)

[9] Dambrosio, L, and D. Fortunato: One-step-ahead control of a wind-driven, synchronous generator system. Energy 24, pp. 9-20 (1999), doi:10.1016/S0360-5442(98)00067-X

[10] Nogaret, E., G. Stavrakakis, J.C. Bonin, G. Kariniotakis, B. Papadias, G. Contaxis, M. Papadopoulos, N. Hatziargyriou, S. Papathanassiou, J. Garopoulos, E. Karagounis, J. Halliday, G. Dutton, J. Pedas-Lopes, A. Androutsos, and P. Pligoropoulos: Development and Implementation of an Advanced Control System for Medium Size Wind-Diesel Systems. Proceedings of the EWEC '94 in Thessaloniki, 10.-14. Okt, pp. 599-604

[11] Tantareanu, C.: Wind Prediction in Short Term: A first step for a better wind turbine control. Nordvestjysk Folkecenter for Vedvarende Energi, October 1992, ISBN 87-7778-005-1

[12] Dutton, A.G., G. Kariniotakis, J.A. Halliday, and E. Nogaret: Load and Wind Power Forecasting Methods for the Optimal Management of Isolated Power Systems with High Wind Penetration. Wind Engineering 23(2), pp. 69-87 (1999) .

[13] Kariniotakis, G., E. Nogaret, and G. Stavrakis: Advanced Short-Term Forecasting of Wind Power Production. Proceedings of the European Wind Energy Conference held in Dublin, Ireland, October 1997, pp. 751-754, ISBN 0 9533922 0 1 .

[14] Kariniotakis, G.N., E. Nogaret, A.G. Dutton, J.A. Halliday, and A. Androutsos: Evaluation of Advanced Wind Power and Load Forecasting Meghods for the Optimal Management of IsolatedPower Systems. Proceedings of the European Wind Energy Conference, Nice, France, 1-5 March 1999, pp. 1082-1085, ISBN 1 902916 00 X .

[15] Fukuda, H., S. Tamaki, M. Nakamura, H. Nagai, F. Shijo, S. Asato, K. Onaga: The Development of a Wind Velocity Prediction Method Based on a Data-Mining Type Auto-Regressive Model. Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 2-6 June 2001, pp. 741-744, ISBN 3-936338-09-4 .

[16] Hunt, K., and G.P. Nason: Wind speed modelling and short-term prediction using wavelets. Wind Engineering 25 (1), pp. 55-61 (2001).

[17] Gong Li and Jing Shi. On comparing three artificial neural networks for wind speed forecasting. Applied Energy Volume 87, Issue 7, July 2010, pp. 2313-2320.

 [18] J. F. G. de Freitas, M. Niranjan, A. Gee, and A. Doucet, "Sequential Monte Carlo methods to train neural network models," Neural Comput.,vol. 12, no. 4, pp. 955–993, Apr. 2000.
[19] N. de Freitas, C. Andrieu, P. Højen-Sørensen, M. Niranjan, and A. Gee,"Sequential Monte Carlo methods for neural networks," in SequentialMonte Carlo Methods in Practice, A. Doucet, N. de Freitas, andN. Gordon, Eds. New York, USA: Springer-Verlag, 2001, pp. 359–380.

[20] T. Kurihara, Y. Nakada, K. Yosui, and T. Matsumoto, "Bayesian online learning: A sequential Monte Carlo with importance resampling," inProc. IEEE Signal Process. Soc. Workshop Neural Netw. Signal Process.,Sep. 2001, pp. 163–172.

[21] Yohei Nakada and Takashi Matsumoto, Online Bayesian Learning WithNatural Sequential Prior Distribution, IEEE transactions on neural networks and learning systems, vol. 25, no. 1, january 2014.