Concept Drift Detection to Improve Time Series Forecasting of Wind Energy Generation

Tomás Cabello-López(✉), Manuel Cañizares-Juan, Manuel Carranza-García, Jorge Garcia-Gutiérrez, and José C. Riquelme

Department of Computer Languages and Systems, University of Sevilla, 41012 Sevilla, Spain

tclopez@us.es

Abstract. Most of the current data sources generate large amounts of data over time. Renewable energy generation is one example of such data sources. Machine learning is often applied to forecast time series. Since data flows are usually large, trends in data may change and learned patterns might not be optimal in the most recent data. In this paper, we analyse wind energy generation data extracted from the Sistema de Información del Operador del Sistema (ESIOS) of the Spanish power grid. We perform a study to evaluate detecting concept drifts to retrain models and thus improve the quality of forecasting. To this end, we compare the performance of a linear regression model when it is retrained randomly and when a concept drift is detected, respectively. Our experiments show that a concept drift approach improves forecasting between a 7.88% and a 33.97% depending on the concept drift technique applied.

Keywords: Machine learning · Concept drift detection · Data streaming · Time series · Wind energy forecasting

1 Introduction

Today, a growing concern about climate change exists. Climate change has become a major problem for humankind [1]. It is produced by greenhouse gas emissions from human activities such as burning fossil fuels for energy [2]. Renewable energies represent the best alternative to the use of fossil fuels in power generation since they are an inexhaustible, cheap and exploitable source of energy everywhere in the world [3]. However, its increased use in the electricity mix involves an inherent risk of system instability due to renewable energies intermittency problems [4].

Energy production forecasting in advance might be a possible solution to renewable energy issues. In this context, regression models could provide future

Supported by the Spanish Ministry of Science and Innovation (PID2020-117954RB-C22) and the Andalusian Regional Government (US-1263341, P18-RT-2778).

generation and thus improve the estimation of other non-renewable energy sources needed to balance the power generation mix.

Although regression techniques have shown their potential for such forecasting problems, they are sensitive to abrupt changes in data distributions. Especially in situations where data flows are continuous. In this context, it is necessary to retrain the models to improve their results [5].

Concept drift techniques are based on recognising the moments in which a data stream shows significant changes in trend. Thus, when we detect them, we can retrain our model, performing a better adaptive and incremental training. Concept drift detection techniques are often applied as preprocessing for regression techniques. Jean Paul Barddal et al. [6] study an ensemble-based dynamic model using the ADWIN Concept Drifts detection technique. Jan Zenisek et al. [7], evaluate the benefits of using concept drifts detection techniques to help regression models perform preventive maintenance of industrial machinery. Elena Ikonomovska et al. [8], present a regression tree model for data stream processing using concept drifts detection techniques. Lucas Baier et al. [9] elaborate a strategy that uses simple machine learning models when concept drifts are detected and a general complex model otherwise.

Although concept drift is not a novel preprocessing of time series, our work is the first intent (to the best of our knowledge) to evaluate concept drift detection in a real context such as renewable energy forecasting.

In this paper, we have chosen Linear Regression (LR) model (a very well-known technique in the literature [10]) to address wind energy generation forecasting in Spain and evaluate its performance when it is used with and without concept drift detection techniques.

The content of this paper will be organised in the following sections as follows: all relevant information regarding the data, the models used and their comparison method in Sect. 2. The results obtained by the models and the limitations of the study are shown in Sect. 3. Finally, the main conclusions reached based on the results are in Sect. 4.

2 Materials and Method

2.1 Dataset

To train our models, we have used data about historical wind power generation from the API ESIOS [11]. Data was collected from 1st February 2022 to 1st April 2022 in 10-minute intervals. The total number of records is 7640.

Before introducing the data into a LR model, we must normalise it using the min-max technique. The models process the data using a sliding window strategy with size 4. Our training set thus consists of subsets of 4 records (in ten-minute intervals) previous to the wind power generation to be predicted.

2.2 Concept Drifts Detection Techniques

The techniques used for the detection of concept drifts have been Adaptive Windowing (ADWIN) [12], Kolmogorov-Smirnov Windowing Method (KSWIN)

[13] [14] and Page Hinkley [15]. The parameterization used for each technique has been determined by a grid search as can be seen in Table 1 which highlights the parameters finally selected.

Table 1. Grid search carried out in concept drift detection techniques. In bold, the best parameters selected.

Technique	Parameters	Values
ADWIN	delta	0.005, 0.001, 0.05, **0.2**, 0.7, 1, 0.15
KSWIN	alpha	**0.001**, 0.005, 0.05, 0.07, 0.5, 0.2, 0.7
	window_size	100, 150, 120, 110, 80, **72**, 160
	stat_size	30, **40**, 36, 35, 32, 42
PageHinkley	delta	0.005, 0.001, 0.002, 0.007, 0.01, **0.1**, 0.15
	lambda	50, **25**, 10, 5, 30, 70, 100
	alpha	0.99999, 0.999, 0.85, **0.7**, 0.9995, 0.9993

A Python library called scikit-multiflow has been used for the implementation of the Concept Drift detection techniques [16].

To predict the wind power generation and test the performance of the compared concept drift detection techniques we have used LR. The implementation of the LR model has been carried out through the sklearn library [17].

2.3 Comparison Procedure

In this paper, we compare LR performance with and without using concept drift detection techniques. To evaluate each option, we use Blocked Cross-Validation (BCV) method [18,19] which avoids test data leakage issues. BCV also allows to measure errors with lower variability and provide a more accurate estimate of the generalisation error than a hold-out validation procedure does.

In our experimentation, BCV splits the dataset in a number of folds N. N depends on the number of concept drifts detected. Thus, the whole dataset is divided in different blocks which corresponds with the folds used in BCV. Finally, every fold is used to validate a LR model using 80% of the data as training and 20% as test.

To evaluate the goodness of a a concept drift detection which provides N concept drift, we carried out a BCV with a LR model and a random distribution of N blocks within the dataset and compare LR performance with that obtained from the BCV obtained by the same LR but trained with the N concept-drift-derived block distribution. The measures used to estimate the accuracy of the LR models in the predictions are the mean absolute error (MAE) and the WAPE [20,21].

To reduce the random nature of the comparison, BCV on LR models have been repeated fifty times with random blocks. Then, the mean errors in the fifty executions where used to make the comparison with the results obtained from the concept-drift-derived blocks.

3 Results

This results has been obtained on a local computer with 16 gb of DDR4 8 GB $\times$ 2 (2666 MHz) RAM, an Nvidia GeForce RTX2060-6GB GDDR6 graphics card and an Intel Core i7-9750H processor. The concept drifts detected for each detection technique and the mean time in seconds that LR took to process each of the blocks can be seen in Table 2.

Table 2. Number of concept drifts detected for each detection technique.

Technique	Concept drifts	Mean time LR (seconds)
ADWIN	2	0.0016
KSWIN	253	0.0003
PageHinkley	185	0.0017

The results of a LR model with and without concept drifts detection according to several detection techniques can be seen in Table 3.

Table 3. MAE, WAPE results from training the model using concept drifts detectors and random folders. In bold, the best results by each technique.

Concept drift detector	Using detector		Not using detector	
	MAE	WAPE	MAE	WAPE
ADWIN-LR	**890.266**	**0.0077**	1348.324	0.012
KSWIN-LR	**128.386**	**0.0012**	139.366	0.0013
PageHinkley-LR	**126.975**	**0.0011**	140.86	0.0013

As we can see in Table 3, the use of a concept drift detector improves the mean performance of the LR model when it is applied in our case of study in all cases.

In the following figures, we can observe the mean MAE and WAPE obtained by our models along with the fifty random executions. The red line stands for the results obtained by LR when concept-drift-derived blocks were used to retrain. From now on, we will use CD (concept drift) model to refer to the model which uses the concept-drift-derived block and RB (random blocks) model to refer the model which uses the random blocks. We can see in Fig. 1 and Fig. 2 that the models which use ADWIN have very similar results ten out of the fifty times. Furthermore, three out of the fifty times the RB model obtains better results

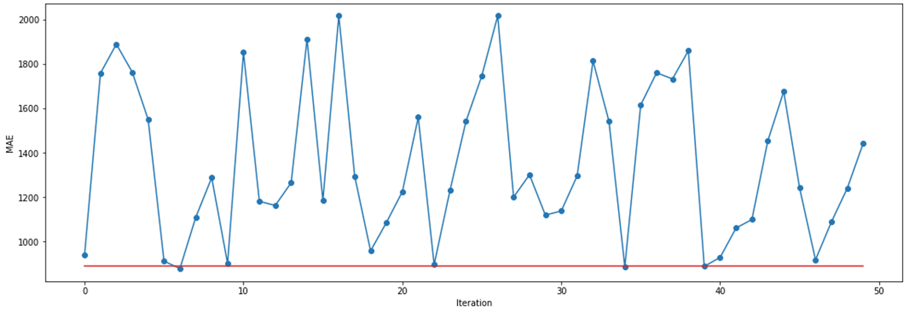

Fig. 1. MAE results using ADWIN. Red line stands for CD best value. (Color figure figure)

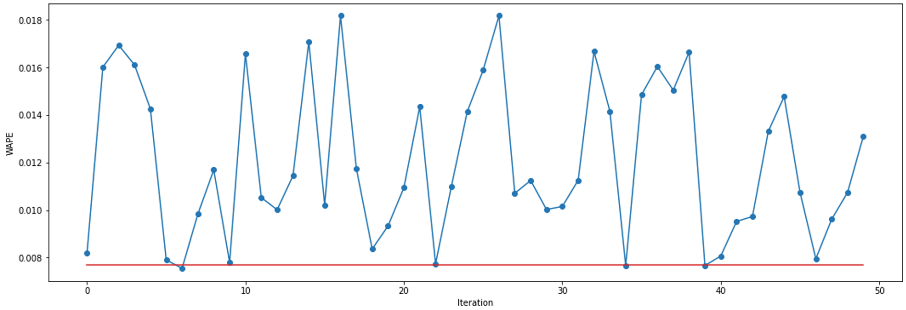

Fig. 2. WAPE results using ADWIN. Red line stands for CD best value. (Color figure figure)

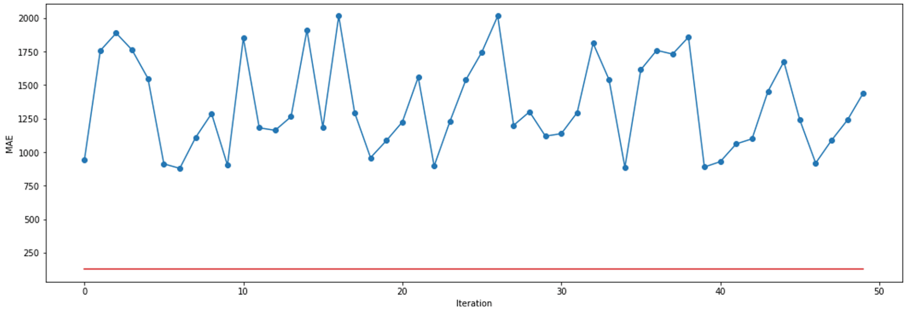

Fig. 3. MAE results using KSWIN. Red line stands for CD best value. (Color figure figure)

than the CD model. Even with better results in a few of iterations, the RB model provides a worse performance throughout the iterations being outperformed by CD model forty-seven times.

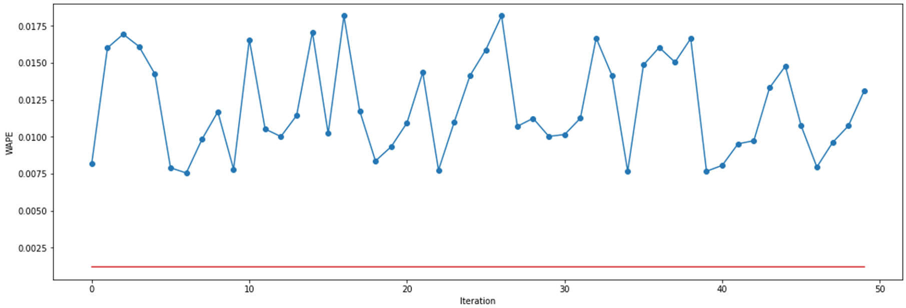

Fig. 4. WAPE results using KSWIN. Red line stands for CD best value. (Color figure figure)

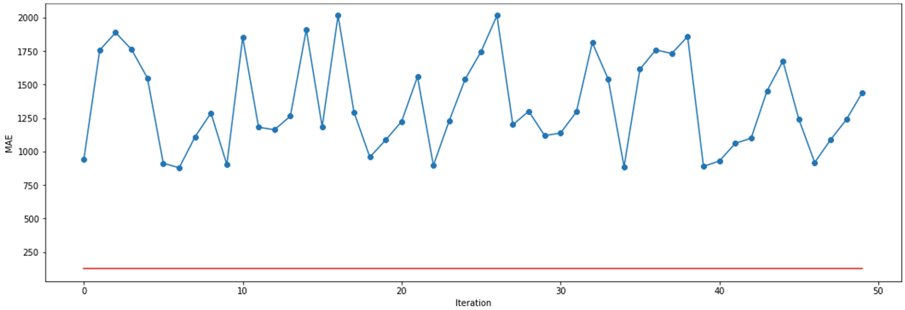

Fig. 5. MAE results using PageHinkley. Red line stands for CD best value. (Color figure figure)

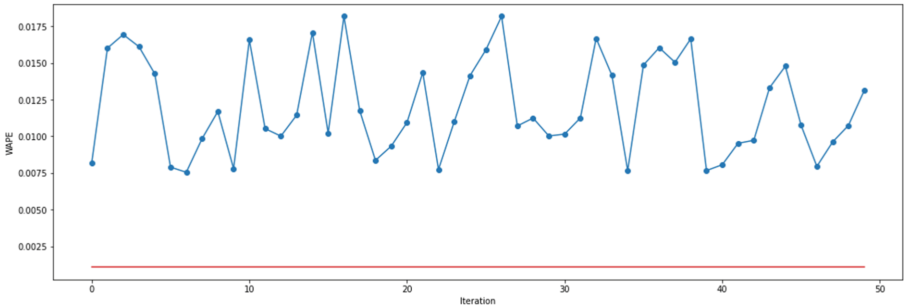

Fig. 6. WAPE results using PageHinkley. Red line stands for CD best value. (Color figure figure)

In the Fig. 3 and Fig. 4 the MAE and WAPE results for the KSWIN models are represented, in this case we can affirm that the CD model outperforms the RB model, being able to achieve much better results in all the iterations. Finally, the Fig. 5 and Fig. 6 show similar results for the PageHikley models. In these cases, the CD model manages to obtain better results than the RB model at every iteration.

Our results seem to confirm that LR model performs better predictions when the concept drifts are previously analysed. They encourage the use of concept drift detection as part of the data pre-processing stage when applying regression models for forecasting.

However, this paper has some limitations, it aims to be a practical exploration of the application of this set of techniques, so we don't attempt to define the best regression technique to deal with power generation forecasting issue but to confirm that concept drift detection can be helpful for regression problems. There are more techniques for time series forecasting with which we could obtain a more complete study. We have fixed the hyperparameters of the concept drifts detection methods through grid search, which does not prove us that there are not better configurations. Finally, we have only taken into account the last two months recorded by ESIOS, it will be interesting to see how the results evolve over time out of our interval of study.

4 Conclusions

In this paper, we analysed the wind energy generation data extracted from the Sistema de Información del Operador del Sistema (ESIOS) of the Spanish power grid. The study we performed indicated that the models improved their results between a 7.88% and a 33.97% by the prior detection of concept drifts in the data depending on the concept drift technique applied. The method of detecting concept drifts with which obtained the best result was the PageHinkley method.

Future work could compare the performance of the proposed model with other models, specially from deep learning area of expertise. This study could also be extended by trying to evaluate the performance of this combination of techniques for a longer time and with meteorological variables useful for the prediction of wind power generation.

References

1. Díaz Cordero, G.: El cambio climático. Ciencia y sociedad (2012)
2. Lizano, B.: Calentamiento Global: "la máxima expresión de la civilización petrofósil". Revista del CESLA. https://www.redalyc.org/articulo.oa?id=243329724003
3. Rafique, M.M., Bahaidarah, H.M., Anwar, M.K.: Enabling private sector investment in off-grid electrification for cleaner production: optimum designing and achievable rate of unit electricity. J. Clean. Prod. **206**, 508–523 (2019). https://doi.org/10.1016/j.jclepro.2018.09.123
4. Mills, A.D., Levin, T., Wiser, R., Seel, J., Botterud, A.: Impacts of variable renewable energy on wholesale markets and generating assets in the united states: a review of expectations and evidence. Renew. Sustain. Energy Rev. **120**, 109670 (2020). https://doi.org/10.1016/j.rser.2019.109670

5. Jaworski, M.: Regression function and noise variance tracking methods for data streams with concept drift. Int. J. Appl. Math. Comput. Sci. **28**(3), 559–567 (2018)
6. Barddal, J.P., Gomes, H.M., Enembreck, F.: Advances on concept drift detection in regression tasks using social networks theory. Int. J. Nat. Comput. Res. (IJNCR) **5**(1), 26–41 (2015)
7. Zenisek, J., Holzinger, F., Affenzeller, M.: Machine learning based concept drift detection for predictive maintenance. Comput. Ind. Eng. **137**, 106031 (2019). https://doi.org/10.1016/j.cie.2019.106031
8. Ikonomovska, E., Gama, J., Sebastião, R., Gjorgjevik, D.: Regression trees from data streams with drift detection. In: Gama, J., Costa, V.S., Jorge, A.M., Brazdil, P.B. (eds.) DS 2009. LNCS (LNAI), vol. 5808, pp. 121–135. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04747-3_12
9. Baier, L., Kühl, N., Satzger, G., Hofmann, M., Mohr, M.: Handling concept drifts in regression problems – the error intersection approach. In: WI2020 Zentrale Tracks, pp. 210–224. GITO Verlag (2020). https://doi.org/10.30844/wi_2020_c1-baier
10. Ray, S.: A quick review of machine learning algorithms. In: 2019 International conference on machine learning, big data, pp. 35–39. cloud and parallel computing (COMITCon), IEEE (2019)
11. Api esios documentation. https://api.esios.ree.es/
12. Bifet, A., Gavalda, R.: Learning from time-changing data with adaptive windowing. In: Proceedings of the 2007 SIAM International Conference on Data Mining, pp. 443–448. SIAM (2007)
13. Raab, C., Heusinger, M., Schleif, F.-M.: Reactive soft prototype computing for concept drift streams. Neurocomputing **416**, 340–351 (2020)
14. Lima, M., Filho, T.S., de A. Fagundes, R.A.: A comparative study on concept drift detectors for regression. In: Britto, A., Valdivia Delgado, K. (eds.) BRACIS 2021. LNCS (LNAI), vol. 13073, pp. 390–405. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-91702-9_26
15. Page, E.S.: Continuous inspection schemes. Biometrika **41**(1/2), 100–115 (1954)
16. Montiel, J., Read, J., Bifet, A., Abdessalem, T.: Scikit-multiflow: a multi-output streaming framework. J. Mach. Learn. Res. **19**(72), 1–5 (2018)
17. Buitinck, L. et al.: API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pp. 108–122 (2013)
18. Bergmeir, C., Benítez, J.M.: On the use of cross-validation for time series predictor evaluation. Inf. Sci. **191**, 192–213 (2012). https://doi.org/10.1016/j.ins.2011.12.028. Data Mining for Software Trustworthiness
19. Bergmeir, C., Costantini, M., Benítez, J.M.: On the usefulness of cross-validation for directional forecast evaluation, computational statistics & data analysis. cFEnetwork: Ann. Comput. Financ. Econometr. **76**, 132–143 (2014). https://doi.org/10.1016/j.csda.2014.02.001
20. Shcherbakov, M.V., et al.: A survey of forecast error measures. World Appl. Sci. J. **24**(24), 171–176 (2013)
21. Hewamalage, H., Montero-Manso, P., Bergmeir, C., Hyndman, R.J.: A look at the evaluation setup of the m5 forecasting competition. arXiv preprint arXiv:2108.03588 (2021)