Applicability of Machine Learning to Short-Term Prediction of Changes in the Low Voltage Electricity Distribution Network

Piotr Cofta[1] [0000−0002−4269−6590], Tomasz Marciniak[1] [0000−0002−0862−8500], and Krzysztof Pałczyński[2] [0000−0003−3989−3400]

[1] UTP University of Science and Technology, Al. Prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland

[2] UTP University of Science and Technology; Doctorate School, Al. Prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland

piotr.cofta@utp.edu.pl

Abstract. Low voltage electricity distribution network actively maintains the stability of its key parameters, primarily against the predictable regularity of seasonal changes. This makes long-term coarse prediction practical, but it hampers the accuracy of a short-term fine-grained one. Such predictability can further improve the stability of the network. This paper presents the outcome of re-search to determine whether Machine Learning (ML) algorithms can improve the accuracy of the prediction of next-second values of three network parameters: voltage, frequency and harmonic distortions. Four ML models were tested: XGBoost Regressor, Dense neural networks (both one and two layer) and LSTM networks, against static predictors. Real data collected from the actual network were used for both training and testing. The challenging nature of this data is due to the network executing corrective measures, thus making parameter values return to their means. This results in non-normal distribution with strong long-term memory impact, but with no viable correlation to use for short-term prediction. Still, results indicate improvements of up to 20%, even for non-optimized ML algorithms, with some scope for further improvements.

Keywords: Machine Learning · power distribution network · prediction.

1 Introduction

The distribution network is a complex system that actively maintains the values of its key parameters (voltage, frequency, distortions etc.) within ranges required by regulators. In this process, it uses both proactive and reactive measures, the latter based on the predictions of changes. The more accurate the prediction, the more effective and efficient the network is in maintaining required values of its parameters.

Currently, the network uses coarse predictions that benefit from the cyclical and seasonal nature of changes in demand. However, the emergence of smart

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

networks creates opportunities for additional, fine-grained near-instantaneous proactive measures, where the prediction of the next few seconds will be required.

The objective of this research is to conduct the preliminary screening of the class of ML models to determine whether their use stands a chance of increasing the predictability of changes for the time horizon of a second, or few seconds, as compared with simple static predictors. This paper addresses this question through a preliminary study in the applicability of ML algorithms for forecasting the next-second value of three time series composed of voltage, frequency and harmonic distortions. Four ML regressors were tested: XGBoost Regressor, Dense neural networks (both one and two layer) and LSTM networks, using actual time series.

The analysis demonstrated the challenging nature of those time series when it comes to forecasting. Those time series have no normal distribution, and there is a long-term memory dependence with tendencies to switch monotonicity in order to return to their means. Both correlation and autocorrelation analysis found out that there is little to none linear dependencies between the current and past value resulting in signal being very difficult to forecast.

The research demonstrated that despite those challenges, the use of ML algorithms improves the accuracy of prediction by up to 20%, as compared to static predictors, with XGBoost and dense neural networks slightly outperforming the LSTM neural networks. The novelty of this research lies in the following:

- Addressing short-term predictability, as contrasted with more popular long-term predictability;
- Establishing the reference performance with the use of static predictors;
- Analysis of the use of selected ML algorithms for predictability that demonstrated their relative advantage over the reference performance;
- Indicating the most promising ML algorithms.

2 Literature review

Liu et al. in [5] focused on improving the time series prediction accuracy using wavelet filtering and neural network. Their method eliminates effect of measurement noise, thus improving prediction precision, specifically when wavelet filtering is used.

Górriz et al. in [3] suggested a new method of time series forecasting with the use of Independent Component Analysis (ICA) algorithms and Savitzky-Colay filtering as pre-processing to introduce data to an Artificial Neural Network (ANN) based on radial basis functions (RBF). The presence of pre-processing improved the prediction.

Tao et al. in [8] proposed a chaotic time series prediction method, based on the RBF network. This CIFCA-ROLSA method includes Iterative Fuzzy Clustering Algorithm (CIFCA) and Regularized Orthogonal Least Squares Algorithm (ROLSA). The proposed method was verified on the basis of the known Rollser chaotic system. The results shown in the paper were worse for a ROLSA-only network than for a CIFCA-ROLSA network.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

Zhao et al. in [13] compares the results achieved by back propagation (BP) neural network and the RBF neural network in chaotic time series prediction. RBF and BP neural networks are compared based on the Logistic equation which is a known chaotic equation. The prediction performance of the RBF is better than the BP. The use of RBF to predict the Shanghai Composite Index shows that neural networks can be effectively used to short-term stock price forecasts.

Park et al. in [6] proposed a method based on BiLinear Recurrent Neural Network (BLRNN) for time series prediction, enhanced by applying the multi-resolution learning algorithm to BLRNN training to make it more reliable for predicting time series data. The normalized mean square error (NMSE) was used to assess the efficiency of long-term prediction. Both on the Mackey-Glass Series data and on the Sunspot Series data, the predictor proposed by the author achieves better results than the Multiple-Layer Perceptron Neural Network (MLPNN).

Yang et al. [11] suggested genetic algorithms (GAs) for time series prediction. The results showed that there is a correlation between the observed and predicted deformations, and the resulting models have interpretative forms that can be easily used for further analysis and inform decisions. The proposed method has been tested with promising results.

Yu et al. in [12] proposed using hybrid prediction algorithm using wavelet analysis of time series to obtain trend prediction for a spacecraft, where forecasting results can be obtained by adding the predicted value from each layer. In order to verify the results obtained in the model, they were compared with mean absolute error (MAE) and root mean square error (RMSE). The results indicate a high accuracy.

Wu et al. in [10] proposed a method for energy consumption prediction using BP neural networks in combination with exogenous series. For comparison of effectiveness, the tested model (called NARX) was compared with the normal TDNN model, showing better predictions by the NARX model.

Grant et al. in [4] demonstrated benefits of the Cellular Simultaneous Recurrent Neural Network (CSRN) to identify and predict the dynamics of a 12-bus voltage power system, comparing with a standard single SNR. Two types of disturbances were assessed, including perturbations in power system generators and the least stable loads. The method was also assessed in the event of a transmission line failure.

Qiu et al. in [7] proposed a model based on support vector machine (SVM) to predict the voltage breakdown of rod-plane air gaps, using the binary classifier. The predicted results correlate with the 29 experimental values. The influence of atmospheric parameters on the breakdown voltage obtained on the basis of the analysis of the predicted results are almost the same as those obtained experimentally.

Adebayo et al. [2] proposed a model capable of predicting a loss of a bus voltage. The critical bus voltage stability index (CBVSI), is based on the system load bus volt-age deviation, the maximum load capacity, and the total number of steps needed to achieve the minimum allowable load on each receive bus.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

Simulation results against the conventional fast voltage stability index (FVSI) show that the CBVSI-based method can serve as an alternative tool for power system engineers in the operation and planning of the system.

3 Analysis

The time series used for the analysis were selected from the database created from the measurement of parameters of the electricity at a fixed point at the University campus in Bydgoszcz, Poland. The ND20 meter [1] was used to measure three parameters: one phase voltage (V1, V), frequency (F, Hz) and total harmonic distortion of the voltage (THDV1, %). The rated basic error for ND20 is V1: +-0.2%, F: +-0.2%, THDV1: +-5%. Data from two months were used: July 2019 and December 2019, being representatives of two typical seasonal patterns. For each time series, the 'delta' series was created that consisted of second-to-second changes in the specific parameter. Minimal data cleansing was performed prior to the analysis.

3.1 Statistical properties of time series

The analysis of statistical properties shows that time series satisfy regulatory requirements. Shapiro-Wilk test indicate that they are not sampled from normal distribution. Augmented Dickey-Fuller (ADF) test indicated that they are stationary.

Hurst exponent proved to have significant impact on the understanding of the behavior of time series [9]. Its value, which was below 0.5 for all time series, suggests significant long-term memory dependence and tendencies to switch monotonicity in order to preserve oscillations centered in its mean.

Strong long-term memory dependence, suggested by the value of Hurst exponent, turned out to be difficult to extract. The values of extracted seasonality are around four orders of magnitude smaller than the residual (error).

Autocorrelation and partial autocorrelation of each time series indicated some temporal, short-term memory dependence. Values of correlation with lag of 1 hour and 24 hours are significantly above the error threshold, although it did not later translate into any improvements in the forecast.

Pearson's and Spearman's coefficients indicated that acquired data contain certain correlations between themselves, albeit they is not significant. Nonetheless, ma-chine learning algorithms deployed in this research managed to use these connections.

3.2 Normalization and reference metric

Time series were normalized using the Z-score normalization. Normalized data has been divided into non-overlapping feature vectors containing 10 samples of each signal or the differential signal based on the type of experiment conducted.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

In total, 4,853,209 vectors were created, of which 3,397,245 formed a training set.

The reference metric consisted of the outcome of applying two static predictors (persistence and moving average), each in two variants, to the time series. The persistent predictor always returns the last observed value. The moving average always returns the average over the set number of samples. Errors were calculated using mean absolute error metric (MAE).

Machine learning algorithms' performance was calculated as a difference between the MAE of the most accurate static predictor in regards to the forecasted series (V1, F, THDV1 or their derivatives) and the MAE of a given algorithm, relative to the MAE of the static predictor. Higher values of performance are better.

3.3 Selection and parametrization of methods

This research focused on exploring simple versions of the selected few algorithms, rather than elaborating on a single one. Three algorithms were selected, each one theoretically satisfying the characteristics of time series:

- XGBoost Regressor,
- Dense Neural Networks with one and two linear layers (DNN and DNN2, re-spectively),
- LSTM (Long Short-Term Memory) Neural Networks.

Each model was used to predict the values of only one time series (V1, F, THDV1 or their derivatives) despite using input made out of every signals' samples. It is because prediction of three time series' elements at the same time requires using a combination of loss functions for every metric in order to train the models, leading to the loss in precision.

Both XGBoost and Artificial Neural Networks are well-regarded for their robustness and ability to learn and recognize non-linear patterns. XGBoost Regressor was evaluated with 100 estimators, each having maximum depth equal to 6. Regularization L1 was not performed. Weight of L2 regularization was equal to 1. Minimum sum of instance weight needed in child was set to 1.

Artificial Neural Networks were trained using learning rate equal to 1e-5. The loss function used during the training calculated a mean squared error. First network (DNN) was made of one linear layer with 30 inputs (one for every value of three signals) and one output. Second network (DNN2) contained two linear lay-ers. The first layer had 30 inputs and 60 outputs and the second one had 60 in-puts and one output. Layers were separated with a ReLU activation function.

LSTM Neural Network required splitting input data in three channels for every signal (V1, F, THDV1). Every channel was trained with separate LSTM layer with 10 gates and 10 outputs with ReLU activation function. Results of LSTM layers processing were concatenated into one-dimensional, 30-element array which is an input to the linear layer with one output.

Additional analysis was conducted to see whether whether selected ML algorithms would allow for the prediction of subsequent values, i.e., 11th, 12th etc.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

3.4 Results

The results are presented in Table 1, showing relative gain for each combination of a time series and ML algorithms. Higher values indicate greater improvement over the best static predictor for the same timeseries.

The results indicated that the use of ML for predictions outperformed static predictors for all time series. The maximum improvement was 20%, achieved by both XGBoost Regressor and DNN for differential frequency. It indicates that there is indeed an added benefit of using ML for prediction.

Table 1. Results of forecasting using Machine Learning algorithms.

	XGBoost	DNN	DNN2	LSTM
$V1$	7.7%	7.9%	1.6%	6.1%
$\Delta V1$	13.2%	14.0%	14.0%	12.4%
F	17.8%	17.8%	17.8%	17.8%
ΔF	20.0%	20.0%	17.8%	17.8%
$THDV1$	7.9%	7.4%	7.1%	7.4%
$\Delta THDV1$	8.0%	7.5%	8.8%	8.7%

Predictions were more accurate for differential time series, which confirms early indications regarding the nature of the time series. Considering differential time series, the dense neural network with one layer (DNN) was the most accu-rate predictor for the differential voltage and frequency time series while the DNN2 with two layers was the most accurate for differential harmonic distortions.

For longer-term prediction, there is a decrease in performance for all models, for both absolute and differential time series, most visible for the frequency time series. It indicates that while there may be a value in using ML for short-term prediction, this value diminishes for the longer-term ones.

4 Conclusions

The second-by second characteristics of the low-voltage distribution network makes short-term fine-grained prediction difficult. The network is not only governed by the law of physics, but it also actively responds to changes in its key parameters, to maintain its stability. Statistical analysis of data sampled through-out summer and winter of 2019 indicated that voltage (V1), frequency (F) and distortions (THDV1) show signs of such self-correction, with tendencies to switch monotonicity in order to preserve oscillations centered in their respective means.

The research demonstrated that the use of ML algorithms for short-term prediction can improve the accuracy of the next second prediction by up to 20%, comparing to static predictors. However, this advantage diminishes for predictions with a longer time horizon. Non-optimized DNN and DNN2 performed

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22

slightly better comparing to non-optimized LSTM and XBoost Regressor, specifically for differential time series. Considering that dense neural networks are the simplest models in comparison with XGBoost and LSTM neural networks, and that no optimization has been performed, results are encouraging. It can be expected that certain optimization, as well as certain tuning of the model, may improve the accuracy of prediction even further.

References

1. Lumel homepage. https://www.lumelcom.pl (December 2020), [last access: 2020/12/07]
2. Adebayo, I.G., Sun, Y.: Performance evaluation of voltage stability indices for a static voltage collapse prediction. In: 2020 IEEE PES/IAS PowerAfrica. pp. 1–5. IEEE (2020)
3. Górriz, J.M., Puntonet, C.G., Salmerón, M., Lang, E.: Time series prediction using ica algorithms. In: Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2003. Proceedings. pp. 226–230. IEEE (2003)
4. Grant, L.L., Venayagamoorthy, G.K.: Voltage prediction using a cellular network. In: IEEE PES General Meeting. pp. 1–7. IEEE (2010)
5. Liu, Z., Pei, X., Huang, G.: Real time prediction method of sensor output time series. In: 2009 9th International Conference on Electronic Measurement & Instruments. pp. 2–969. IEEE (2009)
6. Park, D.C.: A time series data prediction scheme using bilinear recurrent neural network. In: 2010 International Conference on Information Science and Applications. pp. 1–7. IEEE (2010)
7. Qiu, Z., Ruan, J., Xu, W., Huang, C.: Breakdown voltage prediction of rod-plane gap in rain condition based on support vector machine. In: 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE). pp. 1–4. IEEE (2016)
8. Tao, D., Hongfei, X.: Chaotic time series prediction based on radial basis function network. In: Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD 2007). vol. 1, pp. 595–599. IEEE (2007)
9. Wang, N., Li, Y., Zhang, H.: Hurst exponent estimation based on moving average method. In: Advances in wireless networks and information systems, pp. 137–142. Springer (2010)
10. Wu, B., Cui, Y., Xiao, D., Zhang, C.: Prediction of energy consumption time series using neural networks combined with exogenous series. In: 2015 11th International Conference on Natural Computation (ICNC). pp. 37–41. IEEE (2015)
11. Yang, C.X., Zhu, Y.F.: Using genetic algorithms for time series prediction. In: 2010 Sixth International Conference on Natural Computation. vol. 8, pp. 4405–4409. IEEE (2010)
12. Yu, H., Liu, J., Wang, M., Hu, S.L., Guo, R.: The trend prediction for spacecraft state based on wavelet analysis and time series method. In: 2014 11th International Computer Conference on Wavelet Actiev Media Technology and Information Processing (ICCWAMTIP). pp. 88–91. IEEE (2014)
13. Zhao, H.: A chaotic time series prediction based on neural network: Evidence from the shanghai composite index in china. In: 2009 International Conference on Test and Measurement. vol. 2, pp. 382–385. IEEE (2009)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_22