SHORT-TERM LOAD FORECASTING ON MV/LV TRANSFORMER LEVEL

Rik Fonteijn
Eindhoven University of Technology
the Netherlands
r.fonteijn@tue.nl

Thomas Castelijns
Enexis Netbeheer
the Netherlands
thomas.castelijns@enexis.nl

Marinus Grond
Enexis Netbeheer
the Netherlands
marinus.grond@enexis.nl

Phuong Nguyen
Eindhoven University of Technology
the Netherlands
p.nguyen.hong@tue.nl

Johan Morren
Eindhoven University of Technology
& Enexis Netbeheer
the Netherlands
johan.morren@enexis.nl

Han Slootweg
Eindhoven University of Technology
& Enexis Netbeheer
the Netherlands
han.slootweg@enexis.nl

ABSTRACT

Distribution system operators (DSOs) are installing increasing amounts of measurement equipment in the distribution networks. The measurements provide the DSOs insights in the challenges distributed energy resources bring to the network. The next step is to utilise the data generated from the field measurements. This paper will introduce a practical implementation of a short-term load forecast for residential loads on medium-to-low voltage level. As this forecast is applied in a congestion management setting, the error over time is visualised, and the model is optimised for the moment at which the congestions are expected. The practical implementation of the forecasting algorithm is part of the Dutch demonstrator of the Horizon 2020 InterFlex project.

INTRODUCTION

Background

As a result of the ongoing energy transition, increasing amounts of distributed energy resources (DER) are installed. Distribution networks are faced with new challenges, such as network congestions. Traditionally, Dutch distribution system operators (DSOs) do not place permanent real-time measurement equipment in the low-voltage networks. However, in order to gain insights in the new challenges in their distribution networks, DSOs started to install measurement equipment, which provide information on power, current, and voltage on a 15-minute resolution.

Normally, DSOs reinforce their distribution networks when a congestion problem occurs. This is however costly and time intensive. Alternative methods to manage congestion are evaluated. An example of such method is the application of demand-side flexibility. Research has shown that flexibility can be used for congestion management [1]. The next step is unlocking this flexibility through a market mechanism, for example on a flexibility market. The Horizon 2020 project InterFlex is demonstrating a local flexibility market for congestion management in Eindhoven, the Netherlands [2].

In the local flexibility market of the InterFlex demonstrator, the DSO can obtain demand-side flexibility for congestion management, in a day-ahead and intraday setting. A short-term load forecast is used to determine the expected load profiles of the two medium-to-low voltage (MV/LV) transformers on which congestion can occur. The loads behind these transformers consists of four elements: residential loads (inflexible), loads from electric vehicles (flexible), a battery energy storage system (flexible), and a solar photovoltaic installation (flexible). The scheduled behaviour of the flexible loads is assumed to be known through the flexibility market, the inflexible loads need to be forecasted for every program time unit (PTU) of 15-minutes, 48h ahead of time.

This paper will introduce a practical implementation of a short-term rolling-window load forecast on a MV/LV transformer. The forecasting algorithm is specifically developed for residential loads. This paper will furthermore elaborate on an alternative way to visualise the forecasting error, in relation to the forecasting application in the distribution network. All work is done in the context of the Dutch demonstrator of the H2020 InterFlex project.

Approach

Based on the requirements of the Dutch demonstrator of the H2020 InterFlex project, a load forecasting algorithm specialised at forecasting residential loads is developed. The algorithm provides a 48h rolling window forecast, with 15-minute resolution at a MV/LV transformer level. This paper provides insight in the data and features used for the forecast.

The flexibility market interfaces are based on the universal smart energy framework (USEF) [3]. Through this framework, the flexibility market provides the necessary information on the ahead of time behaviour of the flexibility sources, though so-called D-prognoses. The total load of the transformer can be constructed by combining all prognoses with the forecasted residential load.

The load forecast is used as input for a decision-making

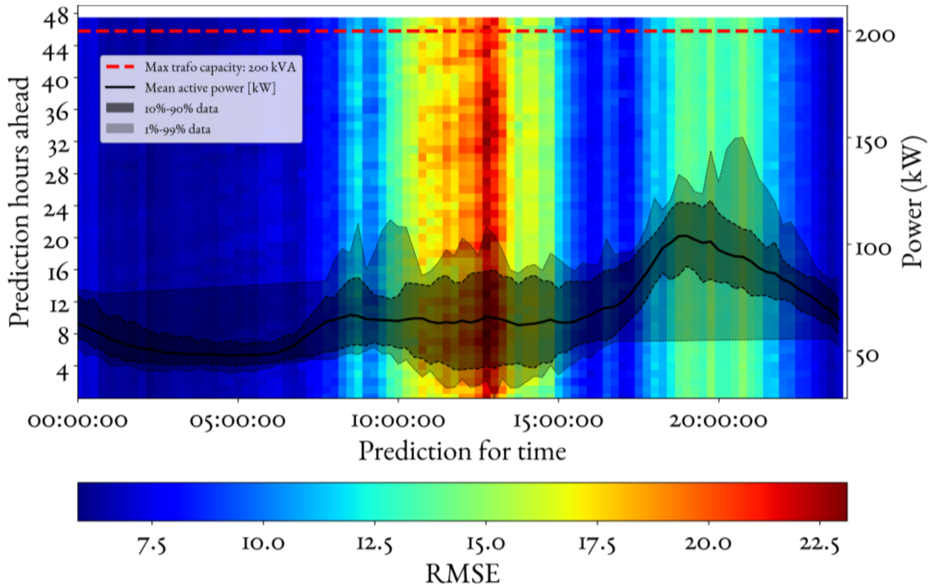

Figure 1: Example of error visualization.

process, in which the DSO decides where, when, and how much flexibility is needed for congestion management, and at what price this flexibility can be obtained. The need for flexibility is primarily expected during the traditional on-peak hours in the evening (17:00-21:00) [4]. Therefore, a methodology to visualise the error is developed to evaluate the performance of the load forecast at certain time windows.

Outline

This paper is organised in the following manner. First, we introduce the methodology to visualise the error. Then, the forecasting algorithm is presented, including insights in the features used, and the model performance. Finally, we present our conclusions and highlight future work.

ERROR METRIC

Error measurement method

Two common metrics to determine the error of a forecast are the mean absolute percentage error (MAPE) and the root mean square error (RMSE)[1]. Whereas the MAPE weighs every error equally heavy by dividing the difference between forecasted and actual value by the actual value, the RMSE is proportional to the size of the error squared. With the RMSE therefore, larger differences between forecasted and actual value are penalised harder.

Since the application of the load forecasting algorithm is a flexibility market to mitigate network congestions in a day-ahead and intraday setting, larger differences between forecast and actual value are more critical than small deviations from the actual values. Therefore, the RMSE is used for the evaluation of the load forecasting model.

Error visualisation

The accuracy (or error) of the forecast lays at the basis of operationalising flexibility through a market mechanism. Therefore, the output of the forecast model direct influences the DSO's perceived flexibility need. Forecasting models are optimised to minimise the overall error. However, congestion problems in the distribution networks often occur at specific times, for example during hours of peak production, or peak loading. Therefore, visualizing the error in relation to time is necessary to evaluate the accuracy of the model over the desired time horizon, and in relation to the moment of the day

Figure 1 provides an example of this proposed error visualisation. On the x-axis, the time of the day is visualised. On the left y-axis, the 192 PTUs (48h) ahead of time are visualised. The colour intensity reflects the RMSE, where dark blue reflects RMSE values around 0, and dark red reflects RMSE values around 40.

The right y-axis shows (average) transformer loadings during the day for recent weeks of data. The transformer load is plotted as a mean, supplemented with the bandwidths in which 10%-90% and 1%-99% of the data lay. Furthermore, the rated power of the transformer is plotted as a dashed red line.

With this, at every time of the day the error for the next 192 PTUs can be evaluated by taking the diagonal, providing insights in the uncertainty and deviations at for example the transformer's (critical) peak loading.

In the example in figure 1, it can be observed that forecast

[1] Ref. [7] for example uses the MAPE and RMSE to compare various forecasting algorithms.

between 24h and 48h ahead result in large errors, in particular for the period between 10:00 and 16:00. Forecasting a few hours ahead of time leads to relatively low errors, in particular during the night.

FORECASTING ALGORITHM

Measurement data

Measurement equipment is installed on two MV/LV transformers and all their outgoing LV feeders. This measurement equipment provides data of active and reactive power, voltage, current, total harmonic distortion, and energy throughput on a 15-minute resolution. Ref. [5] further elaborates on the measurement setup in place for the Dutch InterFlex demonstrator.

For the Dutch InterFlex demonstrator the number of households connected to each feeder, and to which feeders the flexibility sources are connected is known. As the forecast should only reflect the residential loads, the total residential load is computed by adding up of the active power measurements of the feeders connecting households.

Timeseries decomposition

The load profile on a transformer can be considered as a time-series. In order to obtain the best accuracy in the forecasting algorithm, the load profile signal is therefore decomposed (figure 2) into a trend (light-blue line), a daily pattern (green line), a weekly pattern (dark-blue line), and a residual signal (red line). For the decomposition, a decomposition method from the Facebook Prophet open-source library is applied [6]. By eliminating trends and time-dependent patterns, a residual signal remains. This residual signal is forecasted, after which the decomposed signals are added up with the forecasted residual signal.

Feature analysis

Additionally, a number of additional features have been included in the forecasting algorithm. These features can roughly be clustered in two clusters, a time-related cluster and a weather-related cluster.

The time-related features include the quarter (or PTU) of the day, quarter of the week, the quarters ahead, the quarters, the weekday, number of the week, and whether or not a day is a holiday. A number of these features are cyclic (quarter of the day, quarter of the week, number of the week). In order to maintain the cyclic behaviour, those features are modelled as a sine and cosine wave. This way, the modelling maintains a strong link between for example the last quarter of the day and the first quarter of the next day, ensuring the model knows these are successive. The non-cyclic time-related features are modelled with a (binary) dummy variable.

The weather-related features are based on weather forecast data, and include the sun altitude, weather measurements and predictions of the past, current, and next hour, solar irradiance, precipitation, temperature, wind speed, wind direction, weather ahead of time, and weather in the past.

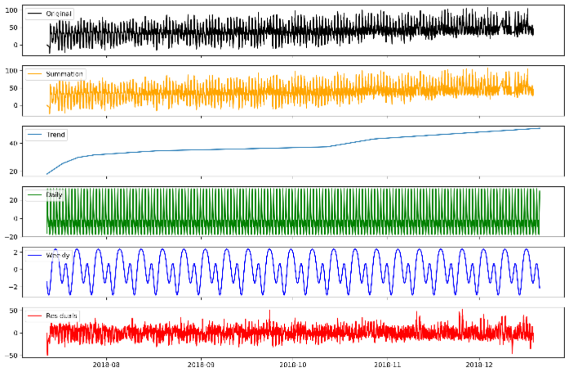

Figure 2: Timeseries decomposition steps.

In order to be able to compare all features with one another, features are scaled to a standardised scale. For this, the standard normal distribution is used, following eq. 1,

$$x'_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

where x'_i is the standardised feature, x_i the original feature, $\underline{\mu}$ the median, and σ the standard deviation. The median of all features ($\underline{\mu'}$) is set to 0, and the standard deviation (σ’) is set to 1. Cyclic (between -1 and 1) and dummy variables (0 or 1) do not need to be scaled further.

An overview of the impact of individual features on the relative RMSE of the forecasting model is provided in figure 3. This is the case-specific impact for one of the two transformers in the Dutch InterFlex demonstrator. From the figure it can be observed that especially between PTU 40 and 80 (the afternoon and early evening), features have a significant influence on the accuracy of the load forecast.

Models

Three different machine learning models have been considered to forecast the time-series of active power: a Linear Regression model, a XGBoost Regression model, and a Random Forest Regression model. To tune the optimal settings of these regression models, also known as hyperparameter tuning, a weighted RMSE is used. In order to ensure maximal accuracy during peak hours, the tuning is made such that the RMSE values of the PTUs around the peak (i.e. 19:00 +/- 8 PTUs) weigh for 50%, whereas the other PTUs combined weigh for the other 50%. Eq. 2 shows this as an objective function,

$$obj = \min\left(\varepsilon \sqrt{\sum_{t \in S_{test}^{peak}} \frac{\left(y_t - y_t\right)^2}{\left|S_{test}^{peak}\right|}} + \left(1-\varepsilon\right) \sqrt{\sum_{t \in S_{test}^{off-peak}} \frac{\left(y_t - y_t\right)^2}{\left|S_{test}^{off-peak}\right|}} \right) \tag{2}$$

where ε is the share of the peak PTUs in the loss function and set to 0.5, S_{test} is the test-set, split up in a peak and an off-peak set, $\hat{y}$ the predicted, and y the measured value.

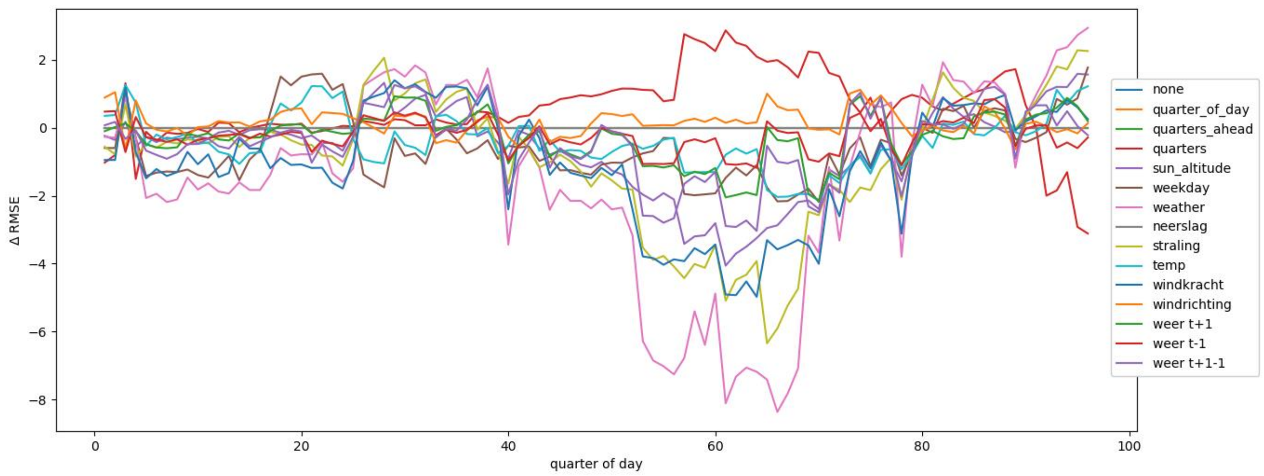

Figure 3: Relative change in RMSE when removing a single feature from the forecasting algorithm on one of the two transformers in the Dutch InterFlex demonstrator. The following features are compared: quarter of the day, quarters ahead, quarters, sun altitude, weekday, weather, precipitation (neerslag), irradiation (straling), temperature, wind speed and direction (windkracht and -richting), and the weather of the past (weer t-1), present (weer t+1-1) and next (weer t+1) hour.

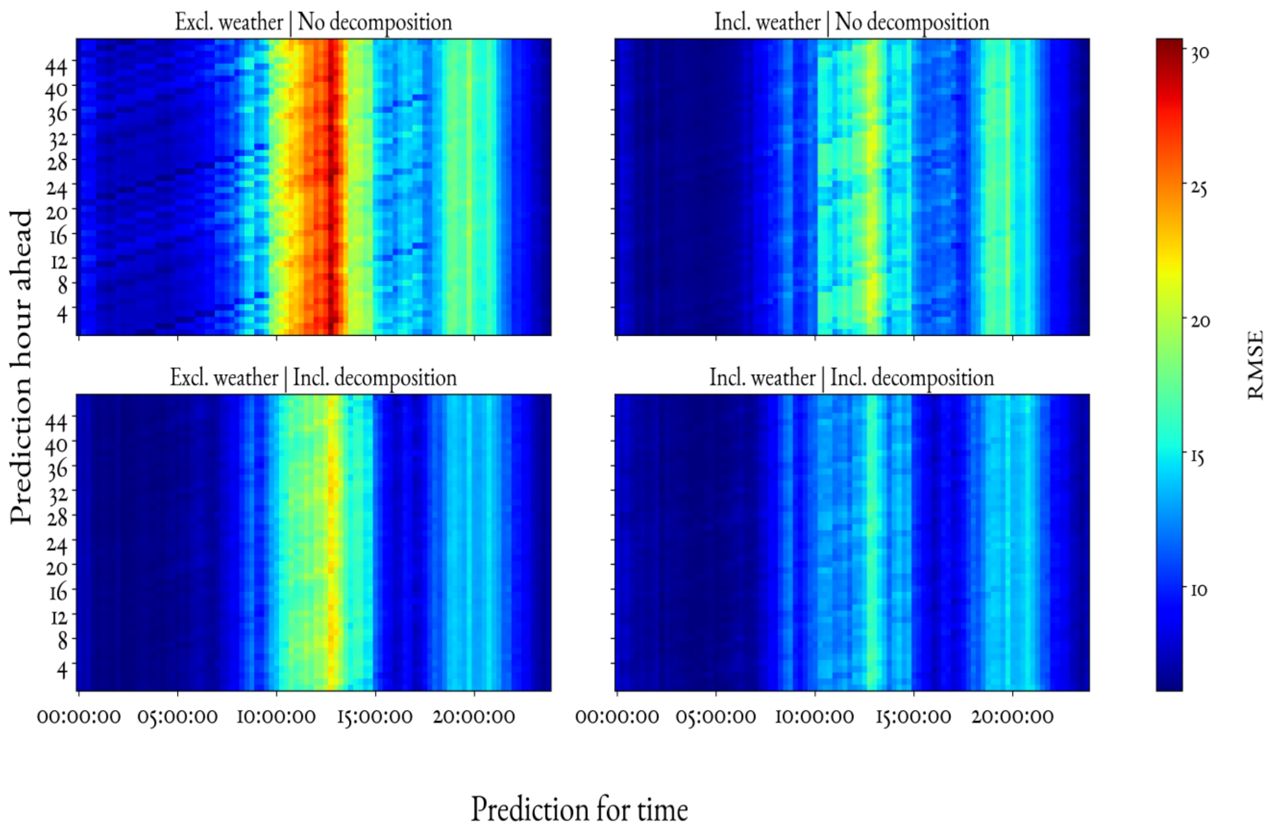

Figure 4: Example of model performance for each moment of the day, without time-series and weather (top-left), without timeseries and including weather (top-right), including timeseries and without weather (bottom-left), and including timeseries and weather (bottom-right).

From the three compared models, the XGBoost regression model has the best performance (weighted RMSE score of 7.75) using the objective function in eq. 1. The linear regression and random forest models scored 8.52 and 9.86 respectively.

Model performance

The performance of the XGBoost regression model is compared for four cases: 1) excluding weather information and timeseries decomposition, 2) including weather information and excluding timeseries decomposition, 3) excluding weather information and including timeseries decomposition, and 4) including weather information and timeseries decomposition.

Figure 4 shows the performance of these four cases, by applying the error visualisations. It can be observed the main uncertainty is found in the late morning and early afternoon. Depending on the connections behind a transformer (e.g. in case of small-scale PV), adding elements of weather information and time-series decomposition results in an increased performance. An example of the output of the model is provided in figure 5. Here, 70% of the historical data is used to train the model, and 30% to test it. Then, a 48h forecast is generated. It can be observed that the forecast follows the pattern of the measurements.

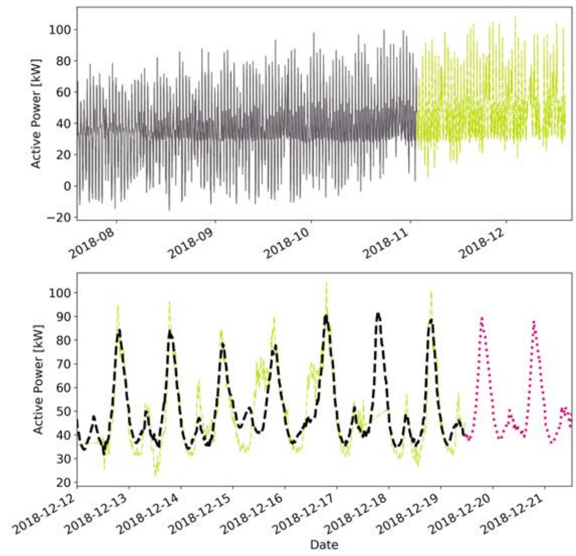

Figure 5: Output of the forecasting data. Top figure – division of test-set (gray) and training-set (green). Bottom figure – field measurements (green), output model (black) and 48h forecast (red).

CONCLUSIONS

This paper presents a practical implementation of a 48h rolling window load forecast on MV/LV level, used in a congestion management environment. To this end, a method to visualise the error in relation to the time-of-day and time-ahead is presented.

Three models have been evaluated, using an objective function tailored to the specific application of the forecast. Of these models, the XGBoost regression model has the best performance. It is shown that this performance can be improved further by applying timeseries decomposition, and introducing weather information and time-related features.

The current model has been developed specifically for residential loads, and for the application in combination with congestion management measures. To investigate further applications of this model, future research is recommended. This should include a broader analysis of the applicability of this forecasting algorithm in the field, both for application in a non- or partly residential area, and for applications outside the congestion management domain.

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement Nº 731289

REFERENCES

[1] E. A. M. Klaassen, "Demand response benefits from a power system perspective," PhD thesis, Eindhoven University of Technology, 2016.

[2] R. Fonteijn, T. Van Cuijk, P. H. Nguyen, J. Morren, and J. G. Slootweg, "Flexibility for congestion management: A demonstration of a multi-mechanism approach," *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2018.

[3] USEF Foundation, "USEF: The Framework Explained," Brussels, 2015.

[4] P. van Oirsouw, *Netten voor distributie van elektriciteit*. Arnhem: Phase to phase, 2011.

[5] M. Roos, R. Fonteijn, P. H. Nguyen, J. Morren, and J. G. Slootweg, "The Strijp-S living lab for embedded microgrid studies," *CIRED Workshop 2018 on microgrids and local energy communities*, 2018

[6] "Facebook Prophet." [Online]. Available: https://facebook.github.io/prophet/.

[7] N. Mohan, K. P. Soman, and S. Sachin Kumar, "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model," *Appl. Energy*, vol. 232, no. June, pp. 229–244, 2018.