FORECASTING PROCEDURE AND TECHNOLOGIES FOR A CONGESTION MANAGEMENT IN LOW VOLTAGE GRIDS

Marc-Aurel FRANKENBACH
Netze BW GmbH, Germany
m.frankenbach@netze-bw.de

Carmen EXNER
Netze BW GmbH, Germany
c.exner@netze-bw.de

Rebecca HOFMANN
Netze BW GmbH, Germany
r.hofmann@netze-bw.de

Ariane HOECK
FZI Research Center for Information Technology, Germany
hoeck@fzi.de

Martin SPITZENPFEIL
EnBW Energie Baden-Württemberg AG, Germany
m.spitzenpfeil@enbw.com

Daniel FOS
EnBW Energie Baden-Württemberg AG, Germany
d.fos@enbw.com

ABSTRACT

The increase in decentralised generation and the massive expansion of e-mobility are causing short-term load and generation peaks in the distribution grid. To prevent these power peaks from overloading operating equipment, the flexQgrid project is testing automated congestion management. This congestion management is based on forecasts of the future grid status generated by building energy management systems and an external provider.

INTRODUCTION

The low voltage grid faces new challenges due to further increase in renewable decentralized power generation and flexible consumption units such as private charging infrastructure for electric vehicles in low voltage grids. The simultaneity of charging processes is a widely discussed topic. However, considering current funding programs and the following strong increase of installed charging stations, even a lower simultaneity of charging processes can cause grid congestions. To avoid these situations, charging stations requiring approval can be rejected in the future if the number of charging stations is already too high in the considered grid cluster. In order to enable the energy transition and a quick and secure integration of the requested number of charging stations into the grid of the future, new concepts for the coordination of flexibilities to avoid grid congestions are required.

The approach tested in the federally funded project *flexQgrid* uses quotas at the grid connection point of buildings and plant sites to avoid predicted congestion as well as to optimise the utilisation of the existing grid and the individual preferences of customers. The operation of charging stations for electric vehicles, heat pumps, photovoltaics and battery storages within households is optimised and coordinated by building energy management systems (BEMS). Each BEMS can react differently in situations of a predicted congestion and adapt its planned operation schedules – for example throttling the charging power of the electric vehicle or shifting the charging times.

FORECASTING PROCEDURE AND TECHNOLOGY

To reliably detect future congestion, forecasts of the power at the critical equipment (transformers and line feeders) are required. The load on the grid is calculated in quarter-hourly time increments. Accordingly, the forecasts are also required in a resolution of 15 minutes. [1]

In the previous project *grid-control*, forecasts based on standard load profiles and forecasts based on measured values were tested [2]. As load curves of prosumer households, especially with a high degree of flexibility, are highly individual due to the influence of the BEMS, the forecast models used in *flexQgrid* are based on measured values.

Furthermore, a distinction must be made in the forecasts between flexible and inflexible power. The distinction is needed to be able to decide on the extent to which the flexibilities can be used to avoid a potential congestion. [1]

The load of a household equipped with an intelligent BEMS is expected to be predicted with higher accuracy if the forecast is made by the BEMS instead of an external party for two main reasons: Since the BEMS knows the technical characteristics of all installed flexible consumption units and has direct communication access to these units, it can differentiate between the measurements of inflexible and flexible load. Moreover, it only needs to predict the household's inflexible load, because it controls the operation of all flexible units. The BEMS' forecast models are described in the next section.

In the case of inflexible households (without BEMS), an external system needs to generate the required load forecasts. These forecasts of inflexible power are aggregated forecasts at the critical equipment to economise measurement technology.

The measurements from the flexible households are aggregated and subtracted from the total measurement at the critical equipment. The result is the inflexible power at the equipment, which serves as the basis for forecasting the inflexible power. The forecast of the flexible power at the resource is composed of the aggregation of the individual household forecasts of the flexible households.

Since forecasts of medium and high voltage transformers are already in use, the aggregate forecast uses existing technology. At Netze BW, the forecasts are currently requested by the NETZlive initiative and generated by EnBW Datalab. For the low-voltage forecasts, the existing forecast models are used and have been checked for their transferability. This subset of low voltage forecasts is therefore called flexQgrid objects in the following.
Contrary to that, household forecasts are not yet in use and were developed for the project by the project partner FZI.

Forecasting procedure for flexible prosumer household

The fundamental task of a building energy management system is the optimised operation of controllable generation, consumption and storage units within a building. While minimisation of overall operation cost presents the most common target [3], other goals such as minimum greenhouse gas emissions can be defined. As optimisation is conducted for future periods, any parameter of the optimisation problem that depends on external factors, such as PV generation, the presence and state of charge of the electric vehicle or a household's general consumption (lights, dishwasher, TV, etc.), needs to be predicted.
The BEMS in *flexQgrid* employs four different forecast models to predict the PV generation, inflexible load and availability (connection state) of the electric vehicle (EV). The individual models are described in the following, starting with the inflexible load forecast.

Load of inflexible consumption

The BEMS' optimisation model requires a forecast of the active power consumption in the household that cannot be controlled. It is termed *inflexible load* in the following. The load of electric vehicle charging, for instance, is not included because the charging process is controlled by the BEMS.
To develop the load forecast model, the data set provided by the HTW Berlin comprising (synthetic) load data of 74 single-family homes [7] was used for training and evaluation. Load data from the buildings participating in the field test has not been available when developing the forecast models.
The inflexible load within each household is predicted using a reference-based model. A forecast may span up to 24 hours with 1 minute resolution and the reference data consists of inflexible load measurements from the past 30 days. The reference-based model works as follows: From the reference data, three different daily load curves are generated. They are based on the average values of all 30 days, of the previous 7 days and of the days with the same day of week as the predicted days, respectively. The final load curve – the forecast – then results from simply taking the mean of these three load curves.
Besides the reference-based approach, a neural network (NN) model has been developed. Its features consist of the inflexible load measurements of the previous day, the day of week and the hour of the day. As load data from the households participating in the field test has not been available for training, the model, when deployed, is faced with load data from a previously unseen household. Accordingly, the testing data set did not include any households from the training data set. A more detailed description of the model and its variants is beyond the scope of this paper. Overall, the performance of the reference-based model was slightly better than the one of the NN models in terms of MRE (3.17% vs. 3.44%, mean of all tested household data sets). For this reason and the requirement for a model that can work with previously unseen data, the reference-based model was chosen for application in the BEMS for the field test.

Electric vehicle connection state

In *flexQgrid*, the BEMS has total control of the charging processes at the households' charging station, i.e., it starts and stops the process and sets the charging current. In order to know when the electric vehicle will be available and in need for charging, the BEMS predicts at which time periods the electric vehicle is connected to the charging station. The battery's state of charge (SOC) cannot be predicted due to a lack of corresponding data, because charging stations operating with alternating current (AC) are not able to communicate SOC values. Since all charging stations used in the project's field test use AC, forecasting the vehicle's SOC is out of scope for the project.
In general, the prediction is made using the same reference-based model as for the inflexible load forecast. Instead of active power measurements, the reference data consists of the connection state which is permanently requested from the charging station and stored in the BEMS' database. The connection state is a binary value, with connected=1 being defined as "the charging cable is plugged into the vehicle and the charging station, and the charging process is authorized". Any other state translates to connected=0 ("unconnected" or "absent").
However, there are two cases where the reference-based model is not exclusively applied. If the vehicle gets connected to the charging station, the BEMS notices the change of the connection state and triggers a new forecast as prerequisite for the optimization of all device schedules, including an EV charging schedule. In this case, first, the average continuous duration of the state of connection and of absence is calculated, respectively, based on all available data from the last six months. Secondly, the vehicle is predicted to stay connected for the resulting average number of hours, then to be unconnected for the corresponding average duration. Only the remaining periods of the forecast horizon, if any, are predicted by the mentioned reference-based model. Otherwise, the model may predict "unconnected" for the upcoming time period based on the historical data and the EV would not be charged. The second case concerns the disconnection of the vehicle from the charging station during an active charging process. In this case, the vehicle is assumed to

stay unconnected for the corresponding average duration and the remainder of the forecast horizon is again predicted by the reference-based model.
In the project, field test participants are provided with a simple BEMS user interface and are asked to enter their planned departure time when connecting their vehicle to the charging station. If a user enters their planned departure time, this information is used instead of the average connection duration.

PV active power generation

For the prediction of the PV active power generation two models have been developed: a short-term model and a 24-hour model. The short-term model is used every hour to predict the active power for the following hour with a lead time of only five minutes, whereas the forecast of the 24-hour model spans 24 hours and further has temporal characteristics that are specific to the requirements of the quota system realized in the project *flexQgrid* The main motivation for the additional short-term model is the ability to update the PV generation forecast frequently, with focus on the next few time periods, and consider the latest PV measurements as well as weather forecast data.
Both models are feed-forward neural networks with two hidden layers. The feature vector comprises the following parts: PV active power measurements normalized by nominal power and the hourly forecast of the sunshine duration from the three weather stations with minimum distance to the field test location (Freiamt), over 1 hour (short-term model) and 24 hours, respectively. The short-term model additionally takes the hour of the day as input. The PV measurements used for training originate from data recorded in 2017/2018 during the predecessor project *grid-control* in four participating prosumer households. The sunshine duration forecast is provided by the German Weather Service (Deutscher Wetterdienst (DWD)) [4]. The sun duration was chosen as a feature instead of sun radiation or irradiance because the latter parameters are not available for all the three weather stations close to Freiamt, especially not for the one with least proximity.
The features described above result from training and testing models with different variations regarding the number of hours of PV measurements to include (1-3 hours (short-term model) / 1-4 days (24h model)) and how to input the sunshine duration forecast (only from closest station, unweighted average of the three closest stations or weighted by distance to Freiamt, direct input of the three forecasts (chosen)). The resulting best models yield an MRE of 7.29% (24h model) and 7.52% (short-term model) on the test set. Accordingly, these models perform better than their respective baseline models, a persistence forecast, with an MRE of 16.26% (24h model) and 8.64% (short-term model).

The performances of all the forecast models described in this section remain to be evaluated during the field test.

Aggregated forecasting procedure inflexible power

The aggregated forecasts of inflexible power at the critical equipment as well as additional PV-generators are calculated by custom forecasting libraries running on the *EnBW DataLab* platform.
Day-ahead and intraday forecasts are generated on demand for an arbitrary amount of measured grid objects (currently around 1000), including transformers and a variety of renewable energy generators. However, the power flows in *flexQgrid* objects are orders of magnitude smaller scale than their equivalents in *NETZlive*: By a factor of around 125 for transformers/substations and a factor of roughly 12 for photovoltaic power plants. Since all applied machine learning algorithms are invariant to scale this does not pose a fundamental technical problem. Yet the lower level of aggregation of individual power consumers/producers can be expected to make these objects harder to predict which in turn could influence the best choice of model hyperparameters.
For both substation and photovoltaic generators recent measurements are included as explaining variables in the model. The photovoltaic generator forecast additionally uses forecasts for local weather conditions such as direct and diffuse radiation (i.e., sunshine) and the local temperature provided by the DWD [3] as well as information about the time of the day.
The substation power flows are an aggregate of heterogenous generators and consumers and allow a larger variety of patterns. Therefore, they are provided with more input variables, including a superset of weather variables including precipitation, and windspeeds, and a detailed encoding of calendar information such as seasons, weekdays, and local holidays. Furthermore, standard and temperature dependent load profiles are calculated based on their definitions by the BDEW [5] and local weather forecasts and included in the set of explanatory variables. Based on experience and empirical performance in backtesting trials, further variables have been derived from the mentioned inputs (‘feature engineering’) for both the substation and the generator model.
For a large set of heterogenous transformers one forecasting procedure is used. Therefore, the relevant subset of input information is not known for individual grid objects. Also, there is no general reliable prior knowledge on the functional relation between input and target variable. For photovoltaic stations the most relevant input variables are known, so that production can always be expected to increase with global radiation. However, even in that case, the mapping from input to output is partially unknown due to inaccurate metadata such as orientation and roof inclination and other local conditions. Last, relying on the availability of recent measurements at prediction time is not always possible. This may be due to delays in the delivery of measurement values that can be caused by failure of IT components. A different reason for missing recent data is classification as implausible or an

outlier by the forecasting library.

Gradient boosting algorithms have the advantages of being robust with respect to irrelevant input variables and completely flexible with respect to the functional form. Recent implementations are highly efficient both in terms of algorithmic complexity and memory usage. [6]
As availability of 100% reliable measurements with zero latency is not given, the general definition of the forecast horizon is altered for direct multi-step modeling. The implemented forecast horizon is defined as the difference between the most recent measurement and the timestamp to be predicted. Forecasting the next 24 hours requires models covering the horizons [t +1, t + 2, …, t + 24 * 4], where each step corresponds to the applied measurement frequency of 15 min. However, measurements can arrive with an arbitrary time delay. Therefore, the forecasting horizons covered by the machine learning models hast to be extended to [t +1, t + 2, …, t + 24 * 4, …, t + 24 * 4 + delay].
Including a trivial model for long horizons that only depends on calendar data such as the time of the day or the weekday ensures the reliable return of a basic plausible forecast, irrespective of the health of data ingestion pipelines.
The models for different horizons can be trained independently in parallel. An alternative recursive approach was clearly outperformed by the direct modeling strategy, such that the increased forecast accuracy outweighed the increased computational demand to train one model for each horizon instead of a single model [8].

The machine learning algorithms come along with a large set of hyperparameters specifying the algorithmic details of how the relation between input and output is learned. A random-search algorithm was applied over a large set of hyperparameter-candidates on all transformers and generators of the *NETZlive* project and selected those that maximize average prediction accuracy in backtesting runs. To prevent overfitting, an early stopping criterion is applied: Model training will stop, once further model refinements do not increase accuracy on an isolated subset of the data (the 'validation set'). The machine learning models are kept up to date by a fully automated process that triggers a retraining on a weekly basis. Forecasts and measurements, as well as forecast accuracy metrics and detailed information from the model training process are easily accessible in a custom dashboard, allowing for quick response times in case of unexpected behavior or poor forecast quality.
Figure 1 shows the distribution of absolute normalized forecast errors grouped by forecast horizon and object type. Normalization is achieved by dividing the absolute error by the object's maximum capacity. Since the forecast horizon is closely correlated with the time of day, the errors were normalized via mean shift such that each time of day has the same average absolute forecast error. This allows to decouple the effect of forecast horizon from the effect of time of day. Two generators and two substations are forecasted, as indicated by the color coding. The bottom row shows the number of observations for each horizon and object. A bimodal distribution is observed for all stations: Most of the forecasts have a horizon of 3 to 6 hours, but there is another peak for the range of 24 to 48 hours which is possibly caused by technical issues of data delivery in the initial project phase. The boxplots show the distribution of the normalized absolute error which is defined as the absolute value of the difference between forecast and measurement divided by the capacity of the object. There is no strong trend in the quality of photovoltaic generator predictions, except for the outliers in the sparsely populated categories of >= 12 hours and >= 336 hours. Note that the forecast horizon is defined by the availability of recent measurements. Earlier analyses have shown that the relevance of recent measurements for photovoltaic forecasts decreases quickly and tends to zero after a horizon of roughly 60 minutes. For longer horizons few additional information was found that is not already contained in the weather data. For substations the forecast error clearly increases in the forecast horizon.

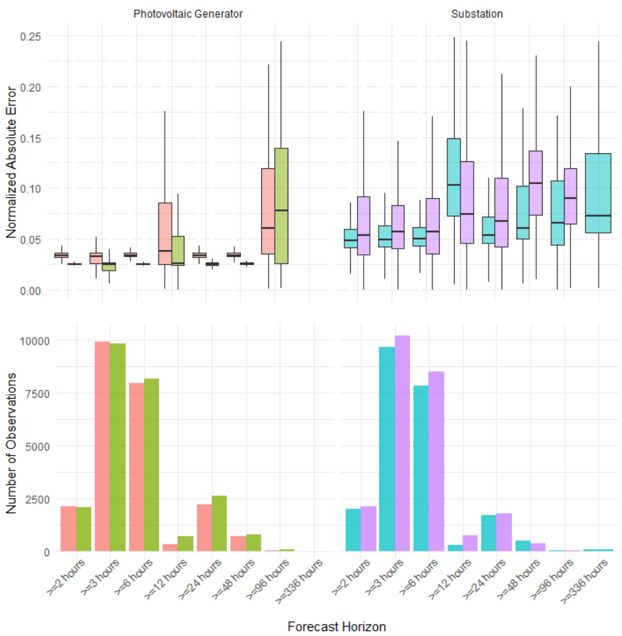

Figure 1 Forecast accuracy by forecast horizon

Plans for future developments include a refined algorithm for the detection of structural breaks, that occur when the load attached to a subnet changes significantly. Besides, the potential of allowing for different hyperparameters for different forecast horizons or for distinct groups of transformers will be investigated.

FIELD TEST OF PROJECT FLEXQGRID

To ensure a practical implementation, the forecasting procedure as well as the congestion management is tested in a field test in the southern Baden municipality of Freiamt.

The proactive avoidance of grid congestions is tested in three low voltage grids with a total of 23 participating households with flexible units. Depending on the household, these can include all or multiple flexible units such as a photovoltaic system, battery storage, heat pump or charging station for the electric vehicle. These units are forecasted, optimised, and controlled by BEMS.

In total 22 electric vehicles participate in the field test divided between three low voltage grids and separate low-voltage feeders. The share of electric vehicles to grid connection points in these feeders are between 7% and 33%, latter share being more realistic in the future. There is even one large prosumer with several charging stations and a separate feeder.

Measurements in the secondary substations and from smart meters in the households are used for the forecasting process. These measurements have a resolution of one minute. To be able to serve as a basis for the quarter-hourly forecasts, the quarter-hourly mean value of the measurements is formed in each case.

As expected, the overall quality of the aggregated prognosis decreases when solar radiance occurs which can be seen in Figure 2a. It also shows that the quality of the forecast loses precision with increasing solar radiance. At the solar peak the error accounts for up to 40% of the power actually fed into the grid. The figure shows the MSE and RMSE by quarter-hourly values over a time span of 14 days in January in one of the three tested low voltage grids. Most days the shape of the solar curve is predicted well, but the actual amount of power fed into the grid is mispredicted. Events like sudden shadowing of a large PV-area for only a short time are difficult to predict which impacts the quality of the prognosis significantly. In figure 2b this effect is visible shortly after 2 p.m. for a single day out of the considered time window from figure 2a.

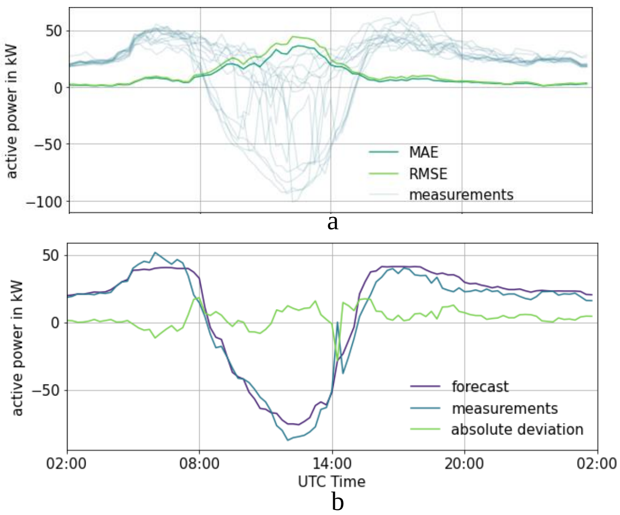

Figure 2: exemplified quality of aggregated forecast (a) quarter hourly quality, (b) shadowing effect

CONCLUSION

Proven forecasting methods for critical equipment from medium and high voltage could be transferred to the application in low voltage. The first evaluations show that the power in substations can be predicted with sufficient accuracy. Together with the forecasts from the BEMS, this forms the basis for automated congestion management. Because the BEMS can predict the use of the respective electric vehicle better than an external system, electric vehicles can also be integrated into this solution. Thus, the introduced forecasts form the basis to intelligently integrate a large number of electric vehicles into the distribution grid. Congestion management and the associated forecasting and control of electric vehicles will be practically investigated in the field test.

ACKNOWLEDGMENTS

The project “flexQgrid” is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi; funding no. 03EI4002A). The authors would like to acknowledge the support from the project consortium as well as from BMWi.

REFERENCES

[1] M. A. Frankenbach, C. Exner, S. Koch, F. Eckl and B. Wasowicz, 2021, "Avoiding congestion and optimizing grid utilization – an advanced multilevel quota realized in a comprehensive smart grid," CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution, 2223-2227.

[2] „Gemeinsamer Abschlussbericht zum Verbundvorhaben grid-control“, https://assets.ctfassets.net/xytfb1vrn7of/5jahPuSzQpCV0TpdoO2ZNr/35a543237b5288ba353e8b6fc21f5ca5/Abschlussbericht_grid-control_final.pdf [2022-01-21]

[3] L. Fiorini, M. Aiello, 2019, “Energy management for user’s thermal and power needs: A survey”, Energy Reports, 5, 1048–1076.

[4] https://opendata.dwd.de [2022-01-21]

[5] https://www.bdew.de/energie/standardlastprofile-strom/ [2022-01-21]

[6] J. H. Friedman, 2001, "Greedy function approximation: A gradient boosting machine." Ann. Statist. 29 (5) 1189 - 1232.

[7] S. B. Taieb, R. J. Hyndman, 2012, ”Recursive and direct multi-step forecasting: The best of both worlds.“ Monash University, Department of Econometrics and Business Statistics Working Paper Series 19/12. https://robjhyndman.com/papers/rectify.pdf [2022-01-21]

[8] T. Tjaden, J. Bergner, J. Weniger, V. Quaschning, „Repräsentative elektrische Lastprofile für Einfamilienhäuser in Deutschland auf 1-sekündiger Datenbasis“, data set, Hochschule für Technik und Wirtschaft (HTW) Berlin, License: CC-BY-NC-4.0, downloaded on 08.12.2020