

Forecasting of Smart Meter Time Series Based on Neural Networks

Thierry Zufferey^(✉), Andreas Ulbig, Stephan Koch, and Gabriela Hug

Power Systems Laboratory, ETH Zurich, Zurich, Switzerland
{thierryz,ulbig,koch,ghug}@eeh.ee.ethz.ch

Abstract. In traditional power networks, Distribution System Operators (DSOs) used to monitor energy flows on a medium- or high-voltage level for an ensemble of consumers and the low-voltage grid was regarded as a black box. However, electric utilities nowadays obtain ever more precise information from single consumers connected to the low- and medium-voltage grid thanks to smart meters (SMs). This allows a previously unattainable degree of detail in state estimation and other grid analysis functionalities such as predictions. This paper focuses on the use of Artificial Neural Networks (ANNs) for accurate short-term load and Photovoltaic (PV) predictions of SM profiles and investigates different spatial aggregation levels. A concluding power flow analysis confirms the benefits of time series prediction to support grid operation. This study is based on the SM data available from more than 40,000 consumers as well as PV systems in the City of Basel, Switzerland.

Keywords: Smart meter · Short-term forecasting · Artificial neural network · Data preparation · Power flow · Big data analytics

1 Introduction

Conventional electricity meters are usually read only once per billing period and give no information as to when energy is consumed at each metered site during that time span. Nevertheless, the current roll-out of new sensor elements called smart meters enables accurate high-resolution measurements on both the spatial scale (on a household level) and the temporal scale (every hour or 15 min) for parts of the distribution grid for which previously only spatially aggregated measurements on the substation and transformer level have been available. At first glance, the main motivations of DSOs to install smart meters are the efficient integration of billing data into the existing billing systems by avoiding manual data gathering, and to facilitate the tracking in case of customers moving to a different property or changing their electricity supplier. Additionally, this evolution can be seen as an excellent opportunity for better operation and planning of active distribution grids, e.g. generation scheduling, load management and system security assessment. This inevitably relies on accurate predictions on the low-voltage level for both distributed generation and end-use consumption, which can be obtained using high-resolution SM data.

Therefore, this paper presents a comprehensive approach for SM forecasting of several types of loads and PV systems, using big data technologies and parallel cloud computing. Predictions are carried out by means of ANNs for several spatial aggregation sizes, going from individual time series to the sum of all available load or production profiles. Measured time series and forecasting results are finally compared by running a power flow analysis for both cases.

The remainder of this paper is organized as follows: Sect. 2 presents the dataset and indispensable preprocessing tasks before performing the forecasting analysis as described by Sect. 3. This is followed by Sect. 4 which shows a power flow simulation based on predicted time series. Eventually, the main contributions of this paper are summarized and future works are given in Sect. 5.

2 Smart Meter Data Preparation

The data used in this study has been collected by IWB [1], the public utility of the City of Basel, between April 2014 and September 2015 and comes from approximately 40,000 small consumers, 1,000 large consumers (commercial and industrial loads) and 400 PV systems that are well distributed across the city. It is made up of energy consumption values with a sampling period of 15 min. Time series whose data is missing during at least one full day in the case of small consumers or one full month for large consumers and PV systems are discarded. In addition, meteorological data measured each 10 min by MeteoSwiss [2] at the weather station of Binningen is also utilized but has to be adapted to comply with the SM sampling rate.

After the above mentioned removals, it appears that 0.94% of energy values coming from small consumers are missing, which is due to both sporadic connection failures for a few SMs and significant data gaps for a majority of devices during several hours. However, this does not impact the billing process as a separate data register exists for the total yearly energy consumption of a given customer. In order not to introduce a substantial bias into the forecasting process, missing data has to be carefully substituted. Since some SM time series are likely to display similar patterns, weighted K-Nearest Neighbors (KNN) is an appropriate imputation method [3]. For the sake of saving time, a reduced training set of 3,000 normalized time series is first created, from which the 5 closest training examples in terms of Euclidean distance are selected for each incomplete load profile. Missing values are then substituted by the weighted average of the corresponding attribute from the 5 nearest neighbors, i.e. each neighbor contributes proportionally to its proximity degree. The KNN implementation is adapted from the open-source software “Knowledge Extraction based on Evolutionary Learning” (KEEL) [4] and supported by the cloud computing engine Apache Spark [5] deployed on a 16-core Azure Virtual Machine (VM).

An anomaly detection is also carried out. On the one hand, it identifies loads with an unusually low energy consumption, i.e. with an average consumption lower than 100 Wh per day for the dataset with small consumers and lower than 100 kWh per month for the one with large consumers. Since the forecasting

algorithm performs very poorly on these load profiles whose energy share among all customers is in fact negligible, they can be excluded from the study. On the other hand, large consumers with a share of zero values higher than 20% as well as PV systems exhibiting a nighttime production are considered as unrealistic and are therefore also removed from the original dataset.

3 Smart Meter Based Forecasting

A wide variety of methods are suggested in the scientific literature concerning time series based prediction. ANNs are nevertheless considered among the most successful machine learning algorithms for this purpose [6, 7]. A feed-forward Multilayer Perceptron (MLP) available in the cloud computing software H₂O [8] is used in this study and deployed in local mode on the Azure VM. Concerning the network architecture, one hidden layer consisting of 200 neurons appears to be a good trade-off between accurate predictions and reasonable computational time. The rectifier $\max(0, x)$ serves as an activation function, notably showing a higher performance than the Sigmoid function for individual SM profiles and low levels of aggregation. Furthermore, a random 50% of incoming weights are zeroed out to prevent overfitting and stochastic gradient descent with backpropagation is used to train the model with a prediction horizon of 24 h, i.e. 96 time steps, starting at midnight. It is assumed that all SM data until midnight is available for the model training and validation. Regarding the meteorological data, values recorded at the same time as the energy value to be predicted are used. This presupposes, though, a perfect weather forecast, which is certainly impossible in reality. The potential impact on the forecasting accuracy is discussed in more detail below, where the presented ANN is assessed on the three different types of datasets.

3.1 Small Consumers

Feature Selection. In this dataset, 27,284 profiles of residential loads, shops, small offices and a few electric storage heaters remain after the preprocessing tasks described in the previous section. Data from April 2014 to March 2015, i.e. one entire year, builds the training set while the month of April 2015 serves as the validation set. Furthermore, four types of data are gathered and used as input features for the neural network. The SM time series itself is the first source of information, from which 16 features are extracted as suggested, in part, by Valtonen et al. in [9]:

- Mean consumption of previous day,
- Last 3 values of previous day,
- Consumption on previous day at the same time, and 3 preceding time steps,
- Average of 3 previous days at the same time, and 3 preceding time steps,
- Average of 3 previous weeks on the same weekday at the same time, and 3 preceding time steps.

Note that multiple consecutive time steps are presented to the ANN simultaneously in order to make use of the temporal structure provided by the time series data. This allows the model not to rely only on a single value that can vary considerably from one time step to the next but to detect a consumption tendency at the considered time period. Instead of a standard feed-forward network, a variant of Time Delay Neural Network (TDNN) is employed, which is known to outperform the simple version [10].

Additionally, three different types of exogenous information are used to train the neural network, which can greatly increase the forecasting accuracy. Concerning weather data, only air temperature is considered in this paper. Since this feature appears to have a limited influence on the prediction performance, errors in the temperature forecast would not significantly impact the result. Another category consists of calendar features such as the hour of the day, the weekday and the month. The energy consumption depends finally to a large extent on social activities. For instance, at a household scale, the start-up of a single device like an electric oven or a washing machine is clearly visible in the load profile. Although it is inconceivable to accurately measure the human activity, one can still account for public holidays. To summarize, 21 features are fed to the ANN, from which a large share directly comes from historical energy values.

Spatial Aggregation. Besides training a model on single time series, different aggregation sizes are investigated, i.e. the aggregation of 10, 100, 1,000, 10,000 and all load profiles. Two options can be implemented which are presented in Fig. 1 with an example of 10 SMs:

- (a) The forecast is carried out for each single profile before building groups of 10 randomly chosen predicted time series and adding them up,
- (b) Original time series are first added up according to the previous group formation and the forecasting algorithm is only applied to aggregate load profiles.

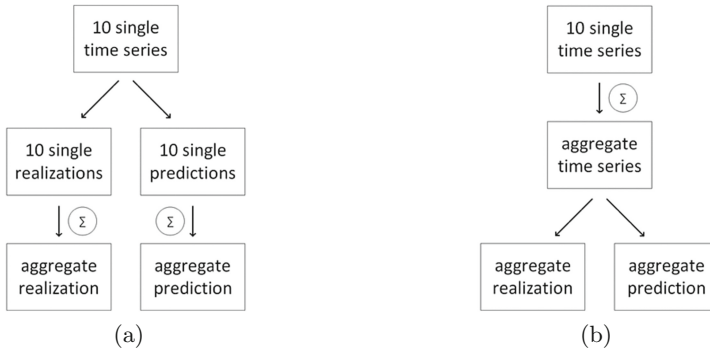


Fig. 1. Two versions to perform prediction with an aggregation of 10 load profiles.

Load Profiles. As previously mentioned, the model is trained during a whole year and evaluated in April 2015, which requires 2 to 4 seconds per SM. Figure 2 illustrates the case of an individual household during the fourth week of April. The predicted profile in red roughly follows the main tendency but fails to fit in with all spikes exhibited by the household, which is in fact hard to predict due to the intermittent nature of home appliances. Since a majority of SMs in this dataset record residential customers with a low and mostly unforeseeable energy consumption, the prediction algorithm inevitably performs poorly on average as described quantitatively in the following subsection. However, one can expect that the performance improves when time series of multiple households are added up because the aggregation exhibits smoother and more seasonal patterns. Figure 3 displays both alternative options (shown in Fig. 1) to combine forecasting and aggregation of 100 load profiles which have been randomly selected from the SM dataset. According to the load shape, this sample consists of a large portion of late evening consumers, typical of households. Here, the ANN overestimates the real consumption for some business days during daytime and notably neglects the brief high spike on Saturday but generally succeeds in estimating the mean load profile hour after hour. Although shops and small offices are a minority among the subsets from 1,000 SMs on, their comparatively larger energy requirements during business hours usually increase the aggregate electricity demand up to the evening peak load level. In this case, the prediction algorithm does fairly well during working days but underestimates the amount of energy consumed at the weekend. The ANN probably places too much faith in features relating to previous days instead of giving weight to weekday based features. Finally, similar smooth profiles and successful outcomes appear for aggregations of 10,000 and of all time series as depicted in Fig. 4. Note that the variant where time series are added up in a second phase typically tends to yield a prediction that shrinks the real load profile due to the smoothing effect of the forecasting algorithm, which is notably visible at high aggregation levels.

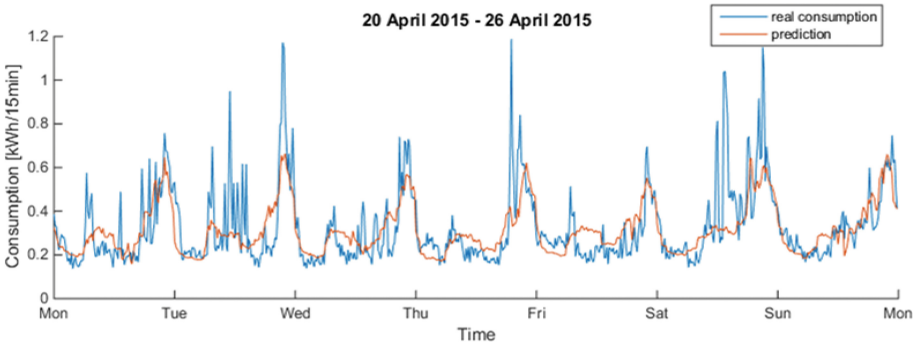


Fig. 2. Forecasting outcome of a single residential load.

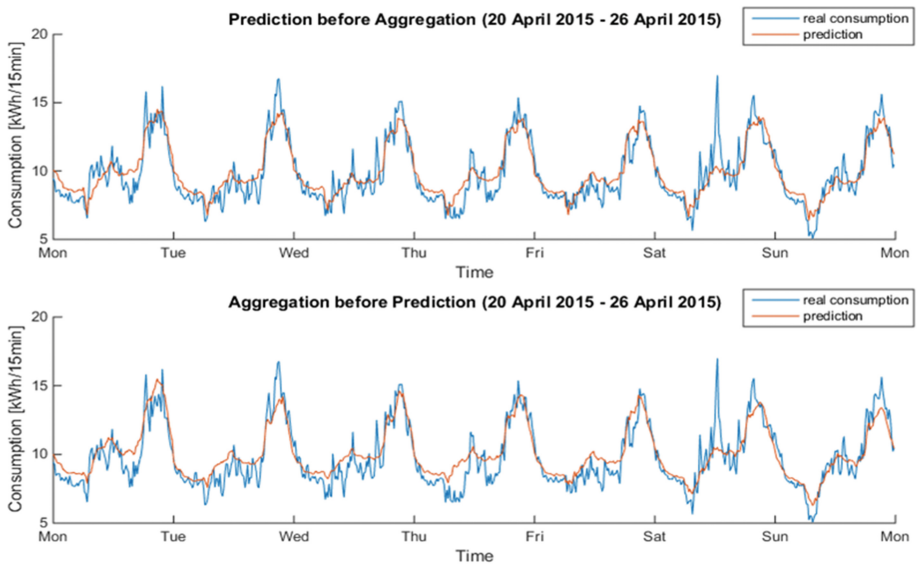


Fig. 3. Forecasting outcome of a random aggregation of 100 load profiles.

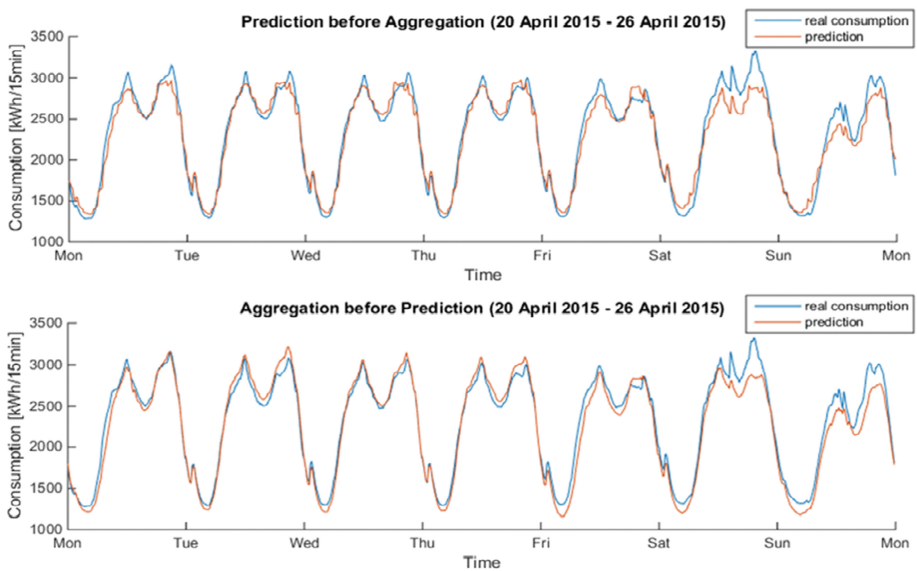


Fig. 4. Forecasting outcome of the sum of all available load profiles.

Performance Evaluation. The Normalized Root Mean Square Error (NRMSE) and the Mean Absolute Percentage Error (MAPE) are standard assessment measures to mathematically evaluate the performance of a forecasting algorithm. Figure 5 shows in detail the procedure used in this paper to obtain the NRMSE averaged over all aggregate groups with an example of 10 SMs per group. The principle is similar for the MAPE and for larger groups.

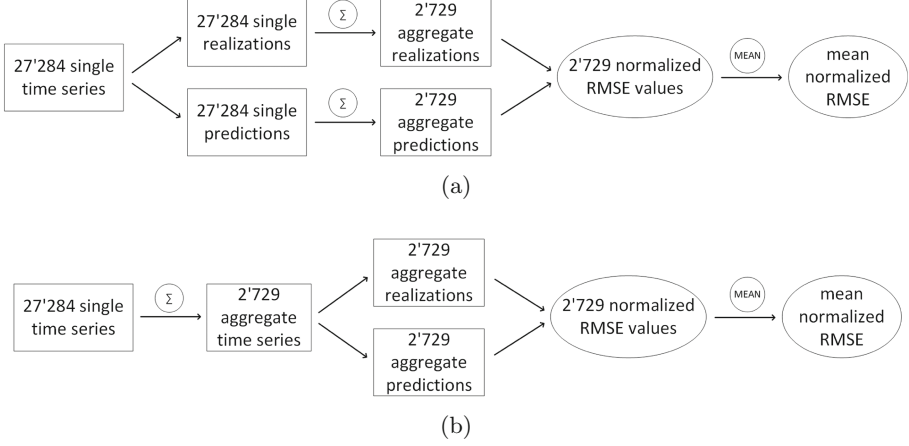


Fig. 5. Computation of the mean NRMSE (a) when groups of 10 individually predicted time series are built and (b) when original time series are aggregated before the prediction is carried out.

However, since a great majority of individual SMs exhibit some zero values, the standard MAPE becomes infinite. A modified metric is therefore used, where instead of normalizing the error at each time step, the Mean Absolute Error (MAE) is first calculated before dividing it by the mean energy consumption in 15 min:

$$\text{MAPE}^* = \frac{\frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} |e_i|}{\frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} y^{(i)}} = \frac{\sum_{i=1}^{m_{\text{test}}} |e_i|}{\sum_{i=1}^{m_{\text{test}}} y^{(i)}} \quad (1)$$

Table 1 presents averages and standard deviations of NRMSE and MAPE* values and the former are plotted in Fig. 6a with respect to the aggregation size. Since the dataset mainly consists of residential loads with a relatively low average electricity consumption but large localized peaks that are hard to forecast, the mean percentage error is skewed although the absolute error is still acceptable. Note, though, that a few loads can be extremely well predicted and yield an accuracy exceeding 99%. Moreover, the high standard deviation indicates the large variety of single consumers. Nevertheless, better predictions are made with larger groups, which can also be seen graphically in Figs. 2, 3 and 4. From an aggregation size of 1,000 SMs and larger, the ANN performance remains more

or less constant. Since many studies related to short-term load forecasting by means of SM data are based on hourly measurements, it would not be judicious to compare them with results acquired in this paper. In addition, the efficiency of a prediction algorithm highly depends on the quality and the type of input data. Nevertheless, an analogous study using 15-minute energy values with aggregation sizes between 20 and 400 consumers, published by SAP Research, achieves a similar level of performance based on a Seasonal Naïve (SN) algorithm [11].

Table 1. Averages and standard deviations of NRMSE and MAPE* for different spatial aggregation levels for both aggregation-prediction variants.

Group size	mean (NRMSE)	std (NRMSE)	mean (MAPE*)	std (MAPE*)
Single SM	264.8	8263	166	6726
10 SMs (a)	47.6	25.6	33.6	20.8
100 SMs (a)	15.0	2.5	11.1	1.8
1,000 SMs (a)	7.9	0.5	5.9	0.3
10,000 SMs (a)	6.8	0.2	4.9	0.1
27,284 SMs (a)	6.7	-	4.9	-
10 SMs (b)	42.7	12.4	29.4	8.2
100 SMs (b)	15.4	2.3	11.4	1.7
1,000 SMs (b)	8.1	0.6	5.8	0.4
10,000 SMs (b)	6.6	0.06	4.4	0.03
27,284 SMs (b)	6.9	-	5	-

Examining the results of Table 1, a surprising fact is the almost identical forecasting accuracy of versions (a) and (b) for any aggregation size. However, the performance of the two versions is uneven when considering the difference between the measured and the forecasted demand averaged over all evaluation days and all consumers. While the real mean consumption is underestimated for any aggregation level as shown by negative values in Fig. 6b, the predicted amount of energy per day is closer to the reality in variant (b), where the energy difference also varies with the aggregation level, in contrast to the other version. Notably, by looking at the load profile examples of Fig. 4, variant (a) frequently underestimates the demand with the exception of a few night hours whereas variant (b) tends to offset shortfalls at night by surpluses during peak load time.

Instead of grouping randomly chosen SMs, an alternative option is to aggregate loads connected to the same Data Concentrator (DC), i.e. located in the same neighborhood. The City of Basel is actually equipped with more than 400 DCs that each collects energy information from a couple to several hundred SMs before forwarding the data to the central server of IWB. With the aim of increasing the forecasting accuracy, all consumers are first clustered into 5 groups by means of the K-Means algorithm provided by H₂O [8]. The Euclidean

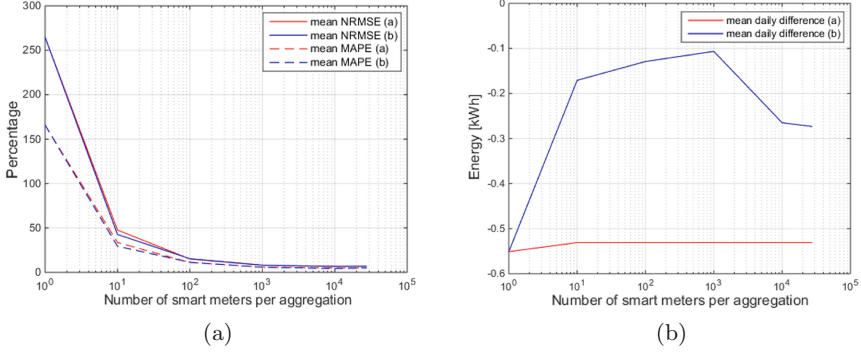


Fig. 6. Performance evaluation of ANN models with respect to the level of aggregation for both prediction-aggregation combinations: (a) average of NRMSE and MAPE*, (b) mean daily energy difference averaged over all consumers.

distance between clustering features such as the typical weekly profile as well as the mean consumption determines the cluster formation. The number of clusters K results from the Elbow method as illustrated by Fig. 7a, which suggests to choose K as the cluster number where the K-Means objective function exhibits an elbow. Consequently, load profiles related to the same DC and that belong to the same cluster are first aggregated together, then the forecast is carried out individually for these aggregate time series which are finally added up to build one profile per DC. In this case, data is trained from April 2014 to March 2015 and tested between April and August 2015. Figure 7b shows the performance at each DC and highlights the considerable dependency of the prediction performance on the energy requirements. Submitting an aggregation of similar time series to the neural network allows to achieve a median MAPE* of 13.06% or a median NRMSE of 18% per DC, and a performance of 3.36% (MAPE*), respectively 5.26% (NRMSE), when considering the sum of all available load profiles.

3.2 Commercial and Industrial Loads

Large consumers are characterized by a higher demand and more periodic patterns, which is reflected by a generally better forecasting accuracy. The dataset considered in this paper consists of 832 commercial and industrial load profiles that have successfully passed through the preparation process. The same features as in the case of small consumers are extracted from individual time series to train the ANN between May and August 2015 and evaluate the model on the data from September 2015. In this case, a median MAPE* of 12.6% and a median NRMSE of 17.6% per large consumer are obtained. Note that the forecast of the sum of all profiles achieves an excellent MAPE* performance of 1.96%.

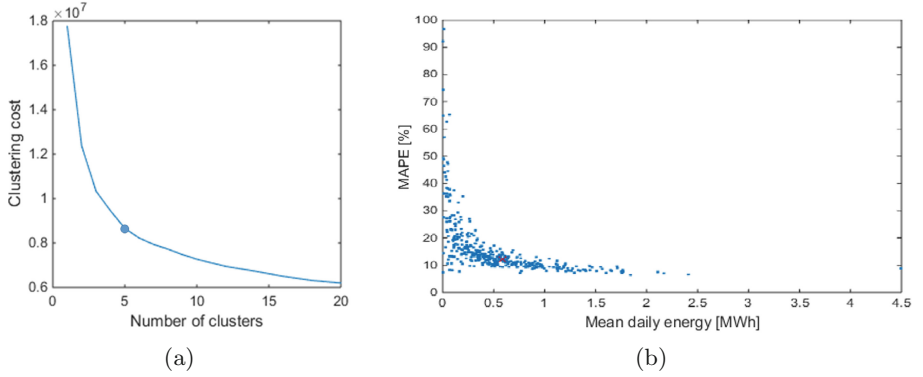


Fig. 7. (a) Elbow method leading to the choice of 5 clusters, (b) MAPE* associated to each DC with respect to the mean daily energy consumption.

3.3 Photovoltaic System Power Production

The same methodology can be used to forecast the PV production even if different input features are required. In particular, it appears that the accuracy gets better without historical SM data. Meteorological data such as global solar radiation, air temperature, atmospheric pressure and relative air humidity are on the other hand of great importance to train the ANN. Hence, weather forecast errors, especially concerning solar radiation, can considerably deteriorate the prediction performance such that presented outcomes must be considered with caution. Furthermore, the hour of the day and the month complete the set of exogenous features. The ANN is then trained from May to July 2015 and evaluated in August 2015. For the sake of consistency, negative predicted values are replaced with zero. Figure 8 illustrates the result during the first 7 days of August for a PV system exhibiting a 9.1% MAPE*. Most PV systems show very similar production patterns such that the performance mainly varies with the corresponding nominal power. In this dataset, a median MAPE* of 13.2% and a median NRMSE of 25.7% are obtained. The forecast based on the sum of all time series finally leads to a result similar to Fig. 8 and a MAPE* of 8.4%.

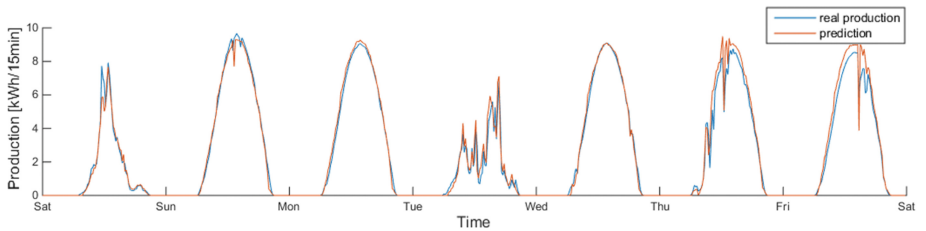


Fig. 8. Forecasting outcome of a PV system.

4 Power Flow Simulation

SM data is usually not available in real-time but in hindsight with a certain delay due to the polling of data only once per day. As shown in this paper, ANNs can nevertheless offer a good approximation of load and PV production profiles. By means of DPG.sim, a simulation environment for active distribution grids developed by the ETH spin-off Adaptricity [12], the impact of day-ahead predictions on power flows and voltages is investigated in an urban area of Basel during 18 days of June 2015. The test grid consists of 12 low-voltage buses and 14 lines in a partially meshed topology with 246 small consumers, 1 industrial load and 1 PV system. Although non-SM loads are not included in the study and data available from loads that do have SMs are arbitrarily assigned to buses since the exact address is unknown for privacy reasons, the case study can still be considered as realistic. Furthermore, all loads are assigned with a fixed inductive power factor of 0.97. The prediction is carried out first for each individual time series, which gives a median MAPE of 16.33% for the active load per bus. As illustrated by Fig. 9a for one of the lines, simulated active power flows based on measured and on forecasted SM profiles are similar and the median MAPE of 17.88% is comparable to prediction accuracy of the active load per bus. In addition, Fig. 9b reveals that bus voltages are barely modified (median MAPE = 0.032%). Note, though, that the voltage is not really sensitive to active power injections. These promising results still need to be validated with other grid topologies and different SM profiles but potentially show that forecasts based on SMs can provide DSOs with additional valuable information, notably for real-time grid operation.

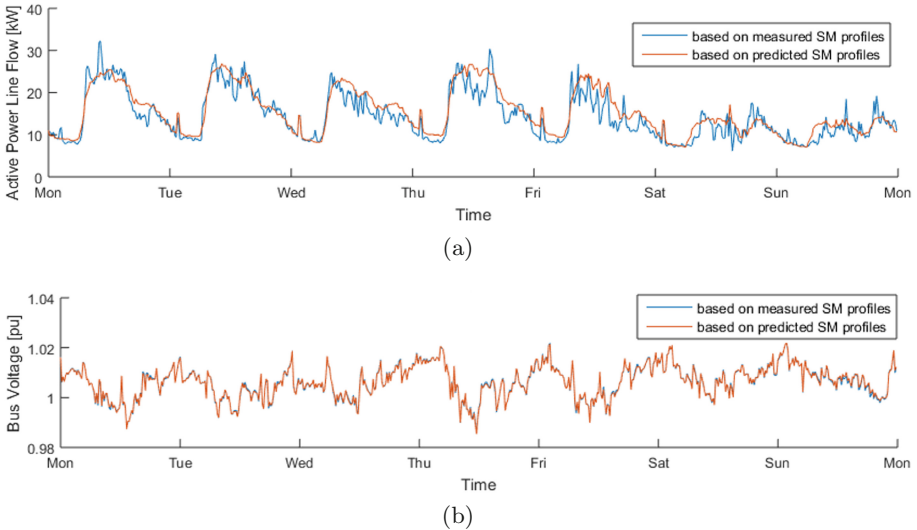


Fig. 9. Simulation of (a) active power line flow (MAPE = 13.67%) and (b) bus voltage (MAPE = 0.029%) based on measured and predicted SM profiles.

5 Conclusion

Based on a large database, this paper proposes an exhaustive approach for forecasting various types of SM profiles, from an appropriate data preparation to the use of predicted time series in a power flow study. While an individual household is difficult to forecast, a considerably improved accuracy is achieved for commercial and industrial loads, PV systems and aggregate load profiles. Furthermore, training an ANN on spatially aggregated time series instead of adding up individual predicted profiles allows to reduce the computational cost without decreasing the forecasting accuracy. An even better efficiency can still be obtained by aggregating profiles of similar shapes. It would be nevertheless worthwhile to consider longer training and validation periods and investigate other prediction algorithms on this type of data, e.g. Support Vector Machine (SVM) or more sophisticated ANNs like Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM). In addition, the impact of weather forecast errors should be closely assessed. Eventually, based on satisfactory SM predictions, DSOs would be able to gain insight into the state of their low-voltage grid in real-time even though real-time measurements are not directly available.

Acknowledgements. This study is part of the project “Optimized Distribution Grid Operation by Utilization of SmartMetering Data” funded by the Swiss Federal Office of Energy and carried out at the ETH Zurich in collaboration with the ETH spin-off company Adaptricity [12] and the public utility of the City of Basel “Industrielle Werke Basel” (IWB).

References

1. Industrielle Werke Basel. <http://www.iwb.ch/>
2. MeteoSwiss. <http://www.meteoswiss.admin.ch>
3. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**(6), 520–525 (2001). Oxford University Press
4. Alcalá-Fdez, J., Sánchez, L., García, S., Del Jesus, M., Ventura, S., Garrell, J., Otero, J., Romero, C., Bacardit, J., Rivas, V., Fernández, J., Herrera, F.: Knowledge Extraction Based on Evolutionary Learning. <http://sci2s.ugr.es/keel/algorithms.php>
5. Apache Spark. <https://spark.apache.org>
6. Koponen, P., Mutanen, A., Niska, H.: Assessment of some methods for short-term load forecasting. In: IEEE PES ISGT Europe (2014)
7. He, W.: Deep neural network based load forecast. *Comput. Model. New Technol.* **18**(3), 258–262 (2014)
8. OxData H2O. <http://www.h2o.ai>
9. Valtonen, P., Honkapuro, S., Partanen, J.: Improving short-term load forecast accuracy by utilizing smart metering. In: CIRED Workshop, Lyon, France (2010)
10. Busseti, E., Osband, I., Wong, S.: Deep learning for time series modeling. In: CS 229 Final Project Report. Stanford University (2012)
11. Ilic, D., Karnouskos, S., Da Silva, P.G.: Improving load forecast in prosumer clusters by varying energy storage size. In: IEEE Grenoble PowerTech (2013)
12. Adaptricity AG. <https://www.adaptricity.com>

Data Analytics for Renewable Energy Integration
4th ECML PKDD Workshop, DARE 2016, Riva del Garda,
Italy, September 23, 2016, Revised Selected Papers
Woon, W.L.; Aung, Z.; Kramer, O.; Madnick, S. (Eds.)
2017, VII, 137 p. 58 illus., Softcover
ISBN: 978-3-319-50946-4