Support of machine learning in assessing distribution network consumer flexibility potential

Jaka Rober[1], Miloš Beković[1], Leon Maruša[2]

[1]*University of Maribor, Faculty of electrical engineering and computer science, Koroška cesta 46, 2000 Maribor*
[2]*Elektro Celje d.d., Vrunčeva ulica 2a, 3000 Celje*
E-mail: jaka.rober@student.um.si

Abstract. *This paper presents a step-by-step approach to assessing the flexibility potential of residential consumers to manage congestions. A case study is presented where a selected transformer station exhibits signs of overloading. Based on historical load data, analysis has been made to evaluate the magnitude of overloading and timing of overload occurrence. The four most prominent consumers have been chosen for flexibility assessment based on historical load data. Machine learning algorithms, specifically multiple linear regression and support vector machines were employed for load profile forecasting during overload occurrences. The generated models were evaluated and compared with forecasting based on the average load of the past days. Based on the evaluated profiles, a scenario of flexibility has been made for each consumer that has been selected as having highest potential for flexibility services. The results demonstrate the effectiveness of the machine learning models, which outperform the average-based forecasting method and provide more realistic estimates of flexibility potential. The proposed approach can be applied to other overloaded transformer stations but with a limited number of consumers.*

1 Introduction

The electric power system has undergone drastic changes in the last two decades. With the so-called green transition and the introduction of carbon-neutral electric energy production, the number of distributed energy resources, such as solar and wind power plants, has increased [1]. Part of the green transition also introduces new loads, most noticeable are heat pumps used for heating [2] and electric vehicles used for transportation [3]. Apart from high energy consumption, the usage of described loads is not evenly distributed across the day, which means high power demand that usually coincides with peak load time and therefore causes transformer station overload. One way of solving these issues is to reinforce network, which is expensive and at the same time we risk not utilizing infrastructure to maximum. It is also important to acknowledge that grid reinforcement may not be able to keep up with the increasing demand of rising consumption.

The current distribution network was particularly placed under constrain, because at the time of planning it was not dimensioned for loads with high consumption and two-way energy transmission [4]. Transformers and cables are essential components that form the foundation of the grid. Temporary overloading of these elements is possible, however it can lead to shortened life expectancy, increased losses, and in the worst case outages. The emergence of a new smart grid environment places consumers in an active role, which means that they can adjust their consumption or generation to provide the necessary ancillary services to the distribution network operator. Consumers with this ability are called prosumers and can provide flexible services in scope of demand response programs. In this way, flexibility plays a crucial role in mitigating issues of network congestion and insufficient voltage profile by enabling prosumers to shift or reduce their consumption or generation during peak demand hours [5].

In this paper use of consumer flexibility is proposed to overcome problem of transformer station overloading. Our approach aims to detect most prominent consumers for flexibility services and evaluate their flexibility potential with aim to enhance grid stability and prevent outages by effectively managing energy demand through consumer participation and optimizing consumption patterns. By forecasting consumers load, we seek to develop a novel approach that can provide assessment of consumer flexibility based on wheather the magnitude and pattern of consumption is sufficient to impact load reduction of transformer station. In other related works [6] users were segmented by load profile and appliance activity. In [7] clustering has been used to determine customer's load profile based on hourly data.

2 Methodology

As mentioned, we proposed load forecast of consumers as method for assessing their flexibility potential. We have assumed that the load forecast can represent the minimum flexibility that a consumer can provide. Load forecast can offer information whether the load pattern is consistent or there has been occurrence of one-time event. Below used forecast methods are described along with methods for their evaluations. Proposed forecast methods for residential buildings are described in [8].

2.1 Forecasting methods

Average of last Y days: last Y working days are selected from the last Z calendar days and then samples of the same time are averaged While it's a straightforward approach, it may not capture the complex patterns and dynamics present in the load profile data.

Multiple linear regression (MLR): models' linear relationship between the independent variable such as weather components (temperature) or calendar events (months or day of week) with the dependent variable (load).

Support vector regression (SVR): is effective for handling non-linear relationships and is particularly useful when dealing with complex datasets with high dimensionality. The radial basis function kernel was used with parameters C = 1, ε = 0,3, and γ was set to scale.

Input parameters of machine learning models were correlations of month, day, and hour. Further information on the average load of the past three days and the week before was included. Also, a rolling average with a window of three hours over the past three days was included. As for weather parameters, a linear relationship with temperature was considered.

2.2 Evaluation metrics

The following metrics were used to evaluate the forecasting performance of models: mean absolute average (*MAE*) (1), mean absolute percentage error (*MAPE*) (2), and root mean square error (*RMSE*) (3). P_i presents measured power and $\hat{P}_i$ is predicted power.

$$MAE = \frac{\sum_{i=1}^{n} |P_i - \hat{P}_i|}{n} \quad (1)$$

$$MAPE = \frac{100\ \%}{n} \sum_{i=1}^{n} \left| \frac{P_i - \hat{P}_i}{P_i} \right| \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_i - \hat{P}_i)^2}{n}} \quad (3)$$

3 Dataset description

For this study a dataset has been provided by distribution network operator Elektro Celje for transformer station Križ Bertot, covering the period between January 1, 2020, and June 1, 2023. With the help of advanced metering infrastructure, 15-minute readings of transformer apparent power were provided, along with active power readings of 19 consumers out of 22.

The dataset was reprocessed so that missing values were filled with the values of a week before at the same time. The same approach has been used to replace extreme values, which were caused by disturbances in the meter reading of data. For consumers with solar power plants, generated and consumed power was summarized. After processing data, four customers had invalid or not complete readings for the whole period of observation.

First two years of dataset have been used for training models and third year has been used for validation. On the rest half a year of data, testing of models was made.

4 Transformer station analysis

Figure 1 shows apparent power measurements of the transformer for the whole duration of a provided dataset.

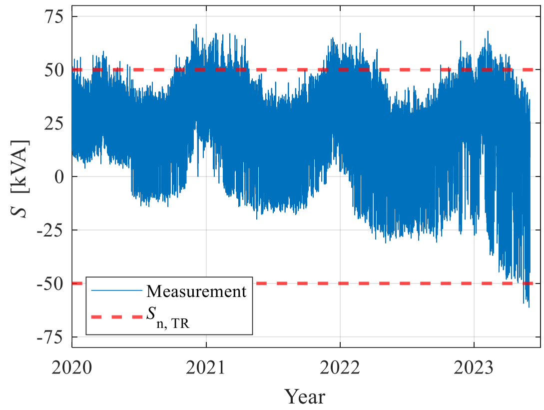

Figure 1: Transformer power measurements.

As it can be deduced from Figure 1, transformer overloading is being occurred in the winter season. From this year on overloading is even being occurred in the summer season because of generated power of solar power plants. In Figure 2 the duration of overload by hour of the day for year 2023 is presented

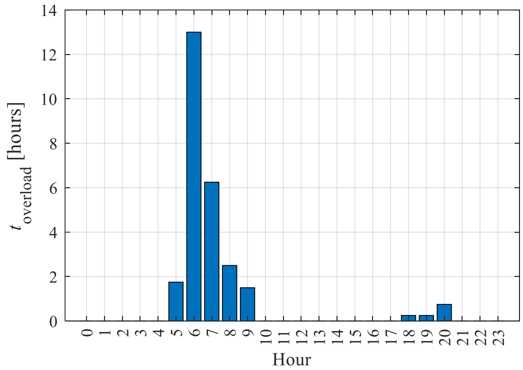

Figure 2: Duration of overload by hour for the year 2023.

From the bar diagram, we can determine which hours of the day are critical for transformer overload. In this case interval between hours 6 in 7 shows the most overloadings.

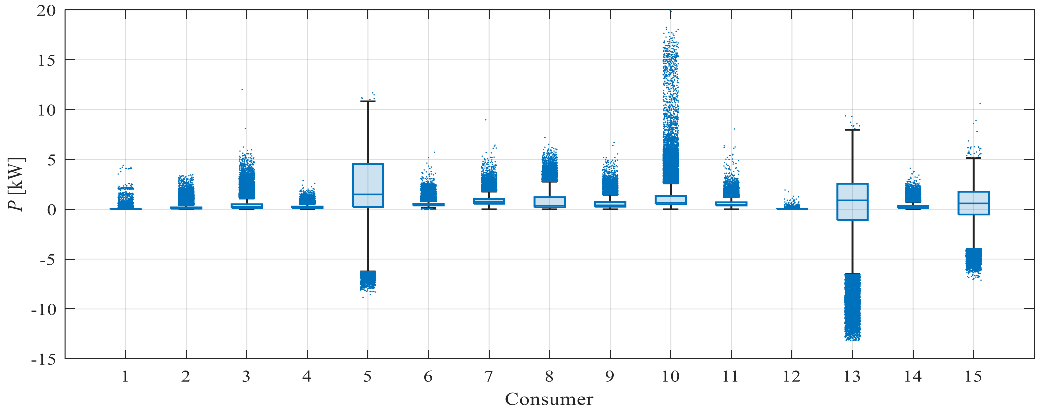

Figure 3: Box plot of consumers load measurement for year 2022.

5 Flexibility assessment of study

To assess the load profiles of consumers, a box plot was created for the load data of 2022 as it is shown in Figure 3. Based on that box plot, the four most prominent consumers were selected by criteria of the high value of median and quartiles that cover as much of high values as is the case for consumers 5, 13, and 15. In addition, consumer 10 was selected because of a high number of outliers with high values, which indicates a unique pattern of consumption. Consumers 3, 7, 8, and 9 were also considered, but it was later shown that at the time of overloading, they did not overload the transformer.

On selected four consumers further analysis has been carried out in such a way that the load forecast has been made on the test dataset for the day with the biggest overload. Highest overload occurred on January 30 between 6:30 and 7:15 hour. Peak value in that time was 64 kVA, which means that transformer was at 128 % of nominal power. In theory that means that minimum power of 14 kVA is needed in worst case scenario. Transformer load profile for that day is shown in Figure 4. After occurrence of overload, there is indication of another peak few hours later, but the power from solar power plant lowers the spike. In the evening, there was nearly another occurrence of overload.

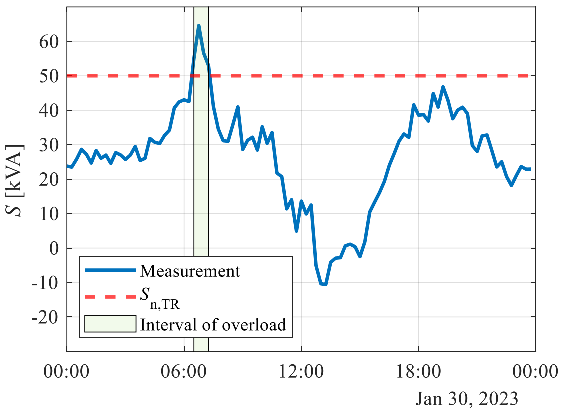

Figure 4: Transformer load profile on the day of overload.

The measured daily load profile of selected consumers for the day of overload is shown in Figure 6 and in Figure 7 load of each consumer at time of overload is shown. Consumer 5 slowly increased power before the overload occurrence and again in the evening. Consumer 10 has a rapid increase of power in times of overload. Consumer 13 has constant power consumption throughout the day if we neglect the generation of solar plants. For consumer 15 load pattern throughout the whole day is hard to predict.

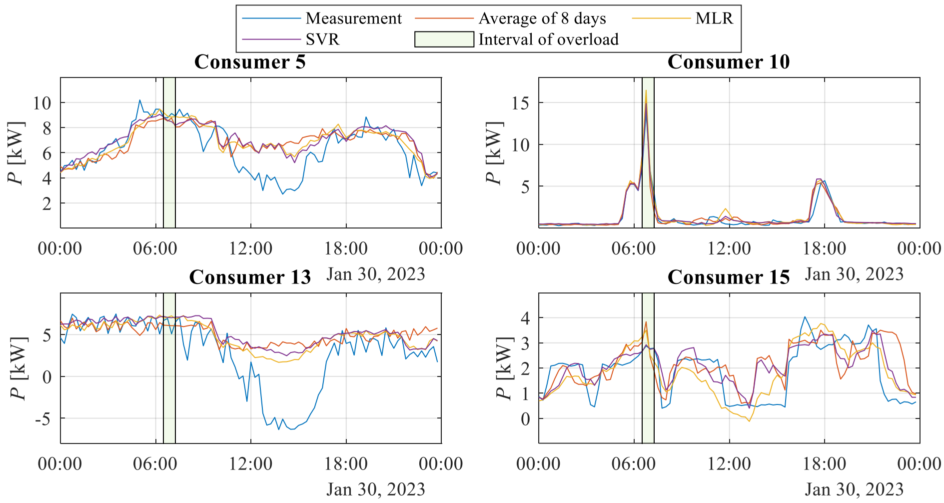

Figure 6: Load profile of consumers 5, 10, 13 and 15 on the day of highest overloading in test dataset.

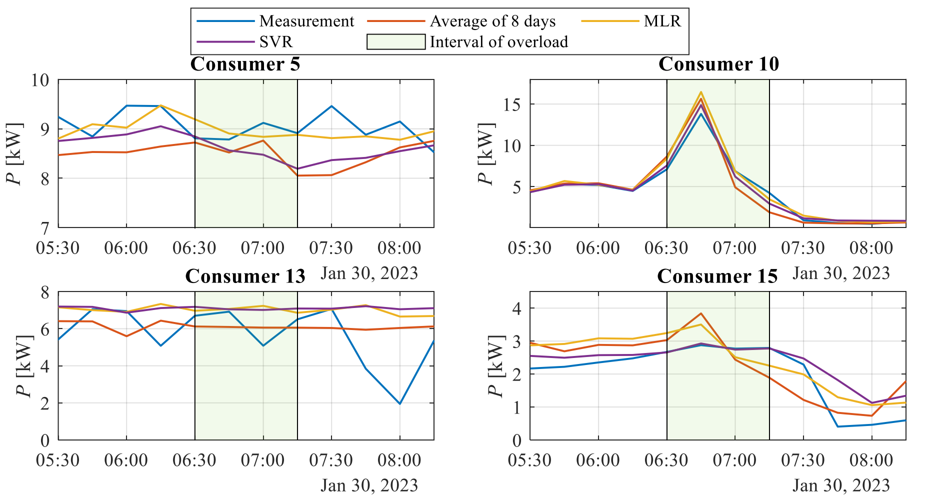

Figure 7: Load profile of consumers 5, 10, 13 and 15 in interval of highest overloading in test dataset.

But in times of overload pattern is more or less predictable but the power is low compared to others.

In Table 1 evaluations of created models are presented for the whole day on which overload has occurred. Both regression models give better results than averaging. The evaluation of the whole test dataset is presented in Table 2. In tabels bold values represent best evaluation for selected method for given consumer.

Table 1: Evaluation of models for a day of overload.

Consumer		5	10	13	15
MAE [kW]	Averaging	1,122	0,322	2,647	0,748
	MLR	**0,982**	0,305	**2,271**	0,581
	SVR	1,029	**0,272**	2,514	**0,578**
MAPE [%]	Averaging	24,39	33,9	187	81,89
	MLR	**21,4**	37,2	**161,6**	**51,0**
	SVR	22,4	**33,4**	202,1	65,7
RMSE [kW]	Averaging	1,54	0,61	4,027	0,99
	MLR	**1,363**	0,561	**3,425**	**0,667**
	SVR	1,411	**0,471**	3,790	0,741

Table 2: Evaluation of models for test dataset.

Consumer		5	10	13	15
MAE [kW]	Averaging	1,491	**0,466**	1,7	0,973
	MLR	1,326	0,494	1,510	0,913
	SVR	**1,307**	0,467	**1,489**	**0,897**
MAPE [%]	Averaging	144	**44**	**187**	204
	MLR	126	52	226	**175**
	SVR	**121**	46	210	178
RMSE [kW]	Averaging	1,929	0,892	2,399	1,298
	MLR	1,760	0,886	**2,083**	**1,179**
	SVR	**1,732**	**0,873**	2,128	1,203

Regression models were only outperformed by a narrow margin in the case of consumer 13. Averaging model is better only in cases where the consumption pattern is constant as is in the case of consumer 10. Regression models outperform averaging models in all other cases. MLR and SVR are in terms of performance equal as they give similar results. It is necessary to consider that the MLR model was built with more input interaction than SVR. The SVR model has been optimized on a limited range, because of the long computation time. Values of *MAPE* are large for consumers with solar power plants, because they increase error due to poor forecast and small values when transitioning in or from generation operation. In assessing flexibility potential, MLR has been used due to its low computing requirements compared to SVR, while yielding similar results.

Flexibility potential can therefore be determined as the highest forecasted value during times of overload for each selected consumer. Consumer 5 can theoretically provide 9,5 kW, consumer 10 15 kW, and consumer 13 7,5 kW of flexibility. The temperature dependency of consumers has been analyzed, but due to limitations of this paper, it cannot be presented. Consumer 5 and 13 show high temperature dependency, while consumer 10 has no temperature dependency. Conusmer with temperature dependency have option to reduce and shift consumption. Scenarios to overcome transformer overloading in this case study are following:

- Load shift of consumer 10.
- Load reduction of consumer 5 and 13, where each lowers load by 80 - 60 %.
- Combinations of consumer 10 load shift and load reduction of consumer 5 and 13.

Consumer 15 has not been included in scenarios due to low consumption at the time of overload. Consumer 5 has more potential for shifting load, compared to consumer 13 which has a constant load through entire day. It is necessary to consider the rebound effect that could cause overload to happen at another time.

6 Conclusion

This paper presents an alternative approach to assessing the flexibility potential of distribution network consumers. A case study was made on transformer station Križ Bertot where regular overloading has been occurring. Along with the simple averaging model two regression models were used to forecast consumption in time of transformer overloading. Out of all consumers connected to a transformer station, four consumers with appropriate historical load data were selected for further analysis. In the end, three consumers were chosen as appropriate candidates for providing flexibility services. Based on consumption pattern scenarios of flexibility were provided.

Models were evaluated using standard evaluation metrics. As expected, regression models outperformed simple averaging models. The proposed approach is limited to a certain number of consumers and transformer station which have consumers with high loads. Created models can also be used in evaluating prosumers' activity during events.

Literature

[1] Perez-Arriaga, I. J. The Transmission of the Future: The Impact of Distributed Energy Resources on the Network. *IEEE Power and Energy Magazine*, 2016,14, pp. 41-53,

[2] Fischer, D., Madani, H. On heat pumps in smart grids: A review. *Renewable and Sustainable Energy Reviews*, 2017, 70, pp. 342-357.

[3] Gemassmer, J., Daam, C., Reibsch, R. Challenges in grid integration of electric vehicle in urban rural area. *World Electric Vehicle Journal.* 2021, 12, 206.

[4] Vahidinasab, V. Overview of Electric Energy Distribution Networks Expansion Planning. *IEEE Access*, 2020, 8, pp. 34750-34769.

[5] Dalhues, S. Research and practice of flexibility in distribution systems: A review. *CSEE Journal of Power and Energy Systems*, 2019, 5, 3, pp. 285-294.

[6] Afzalan, M., Jazizadeh, F. Residential loads flexibility potential for demand response using energy consumption patterns and user segments. *Applied Energy*, 2019, 254.

[7] Kwac, J., Flora, J., Rajagopal, R. Household energy consumption segmentation using hourly data. *IEEE Trans Smart Grid*, 2014, 5(1), pp. 420–30.

[8] Jazaeri, J., Alpcan, T., Gordon, R., Brandao, M., Hoban, T., Seeling, C. Baseline methodologies for small scale residential demand response. *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, Melbourne, 28. November – 1. December 2016. IEEE, str. 747-752.