VTT Technical Research Centre of Finland

Improving the performance of machine learning models by integrating partly physical control response models in short-term forecasting of aggregated power system loads

Koponen, Pekka; Niska, Harri; Huusko, Reino

Published in:
Proceedings ITISE 2017. Granada, 18-20, September, 2017.

Published: 01/01/2017

Document Version
Peer reviewed version

Please cite the original version:
Koponen, P., Niska, H., & Huusko, R. (2017). Improving the performance of machine learning models by integrating partly physical control response models in short-term forecasting of aggregated power system loads. In O. Valenzuela, & I. Rojas (Eds.), *Proceedings ITISE 2017. Granada, 18-20, September, 2017.* (Vol. 2, pp. 795-806). University of Granada.

Improving the performance of machine learning models by integrating partly physical control response models in short-term forecasting of aggregated power system loads

Pekka Koponen[1], Harri Niska[2], Reino Huusko[3]

[1] Energy Systems, VTT Technical Research Centre of Finland, Espoo Finland, pekka.koponen@vtt.fi
[2] Environmental Informatics University of Eastern Finland, Kuopio, Finland, harri.niska@uef.fi
[3] Loiste Electricity Network Kajaani, Finland, reino.huusko@loiste.fi

Abstract. Combining the strengths of different modelling approaches and various information sources is studied in short-term forecasting of aggregated electrical loads that are controllable and include e.g. thermal storage capacity. Measurement data driven models tend to fail in forecasting power during rare situations such as dynamic control actions and extreme weather conditions. The thermal dynamics of the loads, large outdoor temperature variations, and changes in the technologies contribute to this challenge. Here we study a model integration approach using field trial data covering about 7000 houses and 27 months. Control responses and load saturation are forecast using a physically based structure. The residual is forecast with a machine learning model designed and tuned to learn also system dynamics. The load forecast is the sum of these component forecasts. The forecasting accuracy of this hybrid method is compared with using the machine learning alone. The results show improvement in the accuracy.

Keywords: forecasting, machine learning, physically based models, smart grid.

1 Introduction

Accurate forecasts of the power flows in the distribution system are a critical enabler for high penetrations of distributed power generation and demand response. Ignoring the explicit presence of active demand in the model of the load leads to unsatisfactory forecasts according to [1] and [2].

This contribution belongs to a project Response funded by the Academy of Finland, which studies the following research hypotheses. 1) Hybrid models can combine the benefits of different load modelling approaches, thus providing models that (a) forecast relatively accurately in different situations including also those that have not been experienced before, (b) adapt to expected and unexpected changes in the load behavior, and (c) are easy and fast to maintain. 2) Models that combine all relevant available information forecast dynamically controlled aggregated load more accurately than black box models (purely data driven models) or purely physically based models.

There are several ways to improve forecasting accuracy by combining forecasting methods. An approach is to run several forecasting algorithms in parallel and use a weighted average of the forecasts while adjusting the weights according to the situation as learned in the identification [3]. A hybrid ARIMA-ANN model for time series prediction is proposed and studied by [4]; there a multilayer perceptron forecasts the residual of the ARIMA system. We found forecasting the control responses using ARIMA unreliable and inaccurate. The obvious reasons include nonlinearities, nonstationary behavior and limited amount of test responses. We use a model with a physically based structure to forecast the control responses and the saturation of the load, and the machine learning models forecast the residual. Then the load forecast is the sum of these two component forecasts. We successfully applied this approach for electricity spot price based direct control of the aggregate loads of full storage heating houses [2].

In the present contribution, we explain the methods of [2] and give a new summary of the results. Then we apply and modify the approach of [2] to very seldom activated emergency load control of partial storage heating houses located in a climate with large temperature variations. A further difference is that the control responses are modelled from aggregated 3 minute interval measurements from the primary substations in addition to the hourly interval measurements from the smart billing meters. That enables forecasting the emergency control responses with 3 minute time resolution, which is necessary. We also apply and compare two machine learning methods: support vector machine (SVM) and multilayer perceptron (MLP). According to the literature, such as [5], SVM has many methodological benefits and produces smaller forecasting errors.

2 The forecasting problem

The problem studied is to forecast aggregated powers of customer groups that include active demand (AD). The focus is on short-term forecasting: each day at 9 a.m. the power during the next day is forecast with one hour or 3-minute time resolution. Hourly interval consumption measurements from the previous day are available from each customer. The behavior of individual customers is very stochastic but their aggregated behavior is rather well predictable. The outdoor temperature in the region has large variations and the AD responses and loads have highly nonlinear behavior due to saturation of cooling and especially heating. Accurate forecasts during high load situations, such as very cold temperatures, are very important, because then the balancing errors are exceptionally costly and the operational margins in the distribution grid are small.

Two AD forecasting cases are studied using load control field test data. These are

1) forecasting about 700 full storage electrically heated houses subject to electricity spot price based direct load control in Helsinki, and
2) forecasting partial storage electrically heated houses subject to both emergency load control and Time-of-Use (ToU) load control. (The reported results represent 5188 customers. Slightly over 7500 customers were controlled in the verification tests, but we removed from this study all those sites that had gaps in the data or clearly different load behavior.)

In the first case the identification period was 12 months and the verification period was 7 months. In the latter case, the identification data covered 13 months and the verification data 14 months. For forecasting the emergency load control responses, time resolution of the forecasts must be better than 3 minutes. With the physically based response model structures, this is easy to achieve.

Hourly interval consumption history of each customer is available thanks to ubiquitous smart metering. In addition to them, we used outdoor temperature measurements and forecasts, and power measurements from the primary substations to identify and verify the emergency response models.

3 Background research for the emergency load control case

Paper [6] developed and studied physical model based short term daily energy forecasting using the identification data of the emergency load control case. Fig. 1 shows how the daily mean power per house and outdoor temperature varied in the identification period. The developed forecast is shown denoted as simulation. The figure shows aggregated sliding 24 mean powers and sliding 24 h mean outdoor temperature. Here they demonstrate how large the temperature variation is, how much the loads depend on it and how short the extreme temperatures are. It was found out that most of the temperature dependence comes from heating loads, but there is significant cooling load in summer time, and it has somewhat different dynamics than the heating.

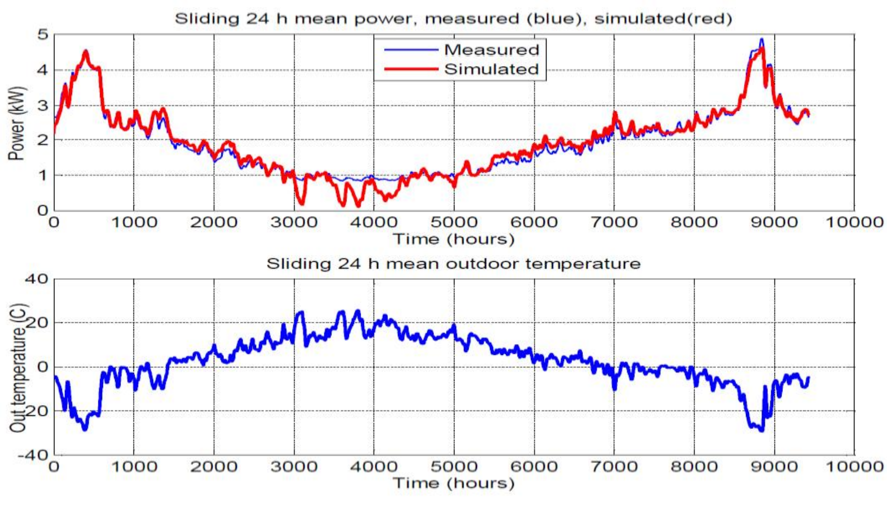

Fig. 1. Temperature dependence of the load in 2011 [6].

4 The hybrid forecasting approach

Machine learning methods alone tend to have challenges in forecasting the dynamics of power during temperature dependent active demand responses and during the load saturation. Here we study a potential solution. We forecast the control responses and load saturation using model structures based on the thermal dynamics of the houses. We identified the parameter values from the identification field tests also taking into account the feasible parameter ranges estimated from the building codes. Then the machine learning methods were taught to forecast the residual of the physically based model. The residual is also a dynamic process so the machine learning models applied need to include capabilities to model the dynamics.

Fig. 2 shows the resulting main structure of the forecasting model.

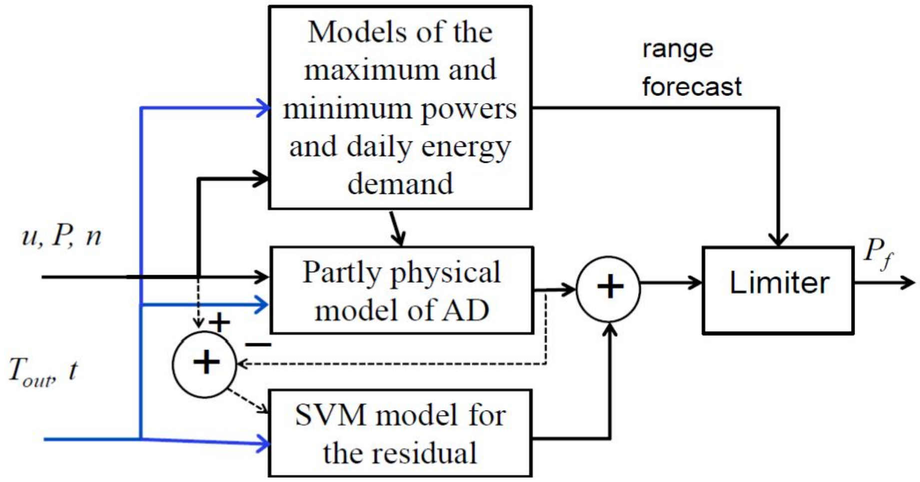

Fig. 2. The machine learning model forecasts the residual of the partly physically based response model.

P is the measured power, P_f the power forecast, u is the control signal and n is the number of houses. T_{out} is a combination of the measured and forecast temperatures as available at the time of forecasting the power. Time t is also an input signal. Each controlled group has its own control signal u and model. The residual model may be common to all groups or each group may have its own residual model. Which one is better depends on the performance and complexity in the particular case.

5 The machine learning methods

In this study, two standard machine learning methods: multi-layer perceptron (MLP) and support vector machine (SVM) are evaluated. Both the methods are largely adopted in the domain of electric load forecasting, e.g. [7].

5.1 Support vector machine (SVM)

Basically SVM is a machine learning technique for data classification and non-linear regression. For more technical details the reader is referred e.g. to [8].

In this study, epsilon(ε)-SVM (or SVR) with the radial basis kernel function based on the LIBSVM package was used to execute the model runs. We have adopted the direct prediction scheme for both the machine learning models by using delayed power and temperature values as regressors. Alternatively, this could be considered as a recursive forecast problem by performing one step ahead prediction. The hyper parameters of SVM were defined as follows. The gamma of kernel function was defined using the LIBSVM default, i.e. 1/number of features whereas C (value 20) was defined based on experimental testing. The value of epsilon ε (0.1) was defined based on the defaults.

The input variables of SVM model were selected based on the previous study [9], which showed that timing variables: day of year, day of week, hour of day, and day length and few delayed outdoor temperature values (-0 hour, -9 hour, and some longer delays, here we used -19 hour) are required to produce sufficient prediction accuracy within the direct prediction scheme. Additionally, we have used delayed power value (here -48 hour) as the SVM model input.

The input data were normalized between -1 and 1. The variance scaling was also tested to prevent influence of potential outliers, but it was not observed to enhance the accuracy.

5.2 Multilayer Perceptron (MLP)

Following, the basic outlines given in the SVM model definition, the standard MLP model was trained to forecast hourly mean powers using timing variables, outdoor temperature and power measurements. The MLP network with one hidden layer (25 nodes) was trained using Levenberg-Marquadt algorithm. In total 3000 training epochs were utilized. A subset of the identification (training) data (5%) were used to control potential over-fitting and to ensure external prediction power. Discontinuous input variables (such as hour of day, day of week) were divided into continuous form by using sine and cosine transformations. This transformation was adopted in case of SVM model, as well.

5.3 Modelling the system dynamics with the machine learning methods

A set of delays is introduced to the forecasting model and during the identification those delays are selected that best improve the fit.

6 Partly physically based control response models

The responses of active demand (AD) to control signals are modelled using models of the thermal dynamics for the buildings and their heat storages. In the houses, the temperature controls are often on-off type. The heating is either on full power or zero power. Such a model is very inaccurate in forecasting the aggregated behavior if a large number of models with stochastic disturbances is not run in parallel. Thus we use a continuous controller in the house model. We fit it to the observed aggregated responses. It turned out that it forecasts accurately the aggregated responses also when the heating in the individual houses is controlled on-off.

For many model parameters a feasible range was estimated from the building codes that set the minimum requirements for the dimensioning and operation of heating, ventilation and insulation. Then the model parameters were estimated by fitting them to the measured test responses in the identification data. In the partial storage heating case the parameters were identified using nonlinear constrained optimization (such as sequential quadratic programming, SQP, or nonlinear conjugate gradient methods). Several initial guesses were used, because multiple local optima were sometimes detected.

The tests of the emergency load control in the identification data set did not include load control actions in cold enough temperatures. Thus it is not possible to model the saturation of heating powers from them. The number of tests was also very small and the information on the rough geographical location of each controlled customers was unknown. Similar emergency load control field tests using power measurements from 13 substations in different cold temperatures had been implemented in1996-1997 and summarized in [10] in chapter 6 and Appendix B. Some of the response models identified from them were applied to the new identification test data. Good forecasting performance was observed. Thus for the emergency load control we use the dynamics and saturation from the old models as such. Only static gain of the model is identified online from the past measurements. A figure of the structure of the emergency load control response model is given in the Appendix B of [10]. It comprises four internal temperatures, the corresponding heat storage capacities, the connecting heat conductivities and ventilation heat loss. The internal state of the temperature controller is in the model, too. The input variables are the following three temperatures: outside air, ground and the set point of the inside temperature.

The model for the responses of full storage heating (for space heating and hot domestic water) includes only the heat storage and its heating element. The thermal dynamics of the building are taken into account only via the forecasting of the heat demand. This response model is given in [2].

The modelling of partial storage heating response still needs some research. For it the model of full storage heating tends to be too simple alone. The very simple response model may nevertheless be adequate in combination with a suitable machine learning model that possibly compensates the shortcomings.

7 Results

Both in the spot price based control case and in the emergency control case the applied machine learning models did not alone forecast the load control responses accurately enough while the hybrid model accurately forecast also the responses. A further initial observation is that also when loads were not dynamically controlled the hybrid models consistently had a slightly better forecasting accuracy than the machine learning models alone. The forecasting performance in exceptional weather situations, and near summer time winter time clock changes, improved.

7.1 Results in the full storage heating case

We studied the dynamically controlled full storage heating case in [2]. There the identification data include 365 days and the verification period was 208 days long. The test included about 700 houses divided in two separately controlled groups.

Table 1 summarizes the results added with a new row. The performance criterion is root mean square error (RMSE) of the forecast normalized to the annual mean power. The first two rows represent models that have a physically based structure. Only the latter one includes a model of the control responses. The method on the last row models the responses using SVM based machine learning. All the other rows are different versions of the approach shown in the Fig. 2, where partly physical models forecast the control responses and saturation, and SVM forecasts the residual. The residual model is common to both groups, because it gave a slightly better performance than forecasting the group residuals separately.

Table 1. Forecasting the residual using SVM improves the forecasting performance

RMSE (normalized)	Identification	Verification
partly physical without response model	0.99105	1.14260
partly physical with response model	0.33606	0.52645
response model and SVM	0.22893	0.36391
response model, SVM and minimum	0.22841	0.34487
response model, SVM and range limit	0.22827	0.34400
SVM	0.17224	0.75300

The SVM alone forecast very well the identification data but not the verification data. This suggests that the 365 dynamically controlled days in the identification data set were not enough for the SVM to generalize the control responses correctly. The hybrid method clearly outperformed its component methods. Using a physically based model for range limitation gives a small further improvement in forecasting performance. Fig. 3 shows a sample of the best forecast of Table 1 compared to the measured power in verification.

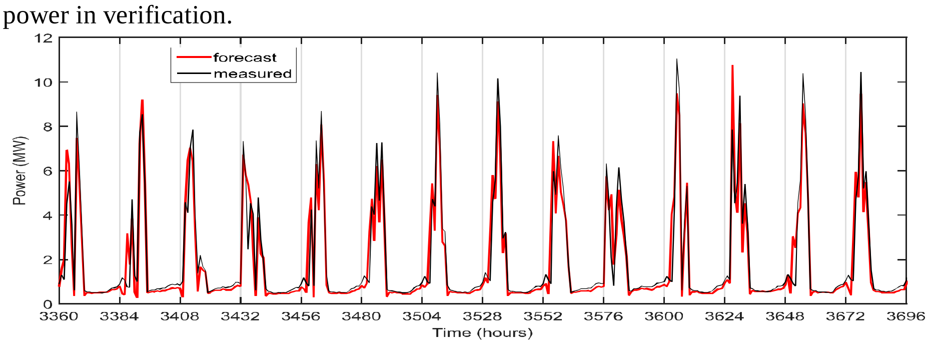

Fig. 3. Forecast and measured aggregated power in verification of the full storage heating case, an example.

The control signals were dynamically scheduled according to the electricity market energy prices and forecast heating needs. When the control signal allows the heating and the heat storage is not full, a thermostat turns heating on and high power peaks occur. The physically based response model models the aggregated behavior of such heat storage system.

7.2 Results in the emergency load control case

The control response model.

The identification period was 13 months long and included some emergency load control tests in early 2013. Then about 8600 electricity customers were subject to the tests and we selected 7062 of them for the response modelling. The test comprised two main groups controlled at different times thus enabling the response identification by reference group comparison. The main groups were split further to subgroups. An identified response of hourly interval powers in outdoor temperature -5 º C is in Fig. 4.

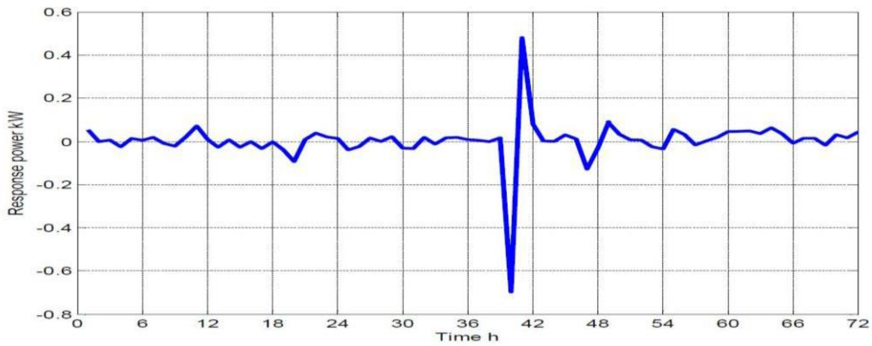

Fig. 4. Emergency load control response identified by reference group comparison, when outdoor temperature was about -5 º C; the control command was applied on hour 40.

The tests in the identification data set were not alone adequate for modelling the emergency load control responses in the temperature range of interest. It was only possible to model the control response in one temperature. Similar emergency load control response identification tests had been implemented in three adjacent power distribution areas in winter 1996-1997 using power measurements from 13 substations [6]. Then the hourly interval measurements of each customer were not available, but the temperature range covered by the tests was -7…-29 º C, which was wide enough to see also the load saturation in some tests. The dimensioning temperature in the area was -32 º C and based on the test results the actual dimensioning of heating and insulation was somewhat better. In the 1996-1997 tests six separately controlled groups of houses were applied based on the thermal dynamics and usage of the buildings. A simple thermal dynamics model was developed, its feasible parameter ranges defined based on building codes and the parameters identified using a nonlinear constrained optimization method, Sequential Quadratic Programming. Now we chose one of those six models for comparison with the 2013 control tests and the results are shown in Tables 2 and 3 below. In the Table 2 the measured responses are uncertain, because the power measurement includes all the loads in the whole distribution area. Especially the reduction of the load in the test 2 has large uncertainty due to apparent simultaneous changes in the other loads. In Table 3 the measured responses are much more accurate.

Table 2. Comparison of the average house responses of power measurements in 2013 with the model identified in 1997 (3 minute time resolution).

Test nr.	**Source**	**Temperature during test ºC**	**Previous mean 24 h temperature ºC**	**Number of houses controlled**	**Load step down kW**	**Load step up kW**
1	measurement	-5.5	-6.9	4757	1.1	2.5
1	old model	-5.5	-6.9	4757	0.95	2.2
2	measurement	-4.5	-9.0	2305	1.2-1.6	2.3
2	old model	-4.5	-9.0	2305	0.98	2.2

Table 3. Comparison of the average house responses of hourly interval powers in 2013 with the model identified in 1997.

Test nr.	**Source**	**Houses controlled**	**reduction in load kWh/h**	**next hour payback kWh/h**
1+2	measurement	7062	0.7	0.5
1+2	old model	7062	0.93	0.41

The old model forecast reasonably well the responses in the emergency load control in the identification data. The time constants are slightly too short and can be adjusted accordingly. For clarity in the following, we use the dynamics of the old model as such. Only the scaling of the response model is identified on-line from the latest power measurements available during the forecasting.

Integration with machine learning models.

Tables 4 and 5 compare the forecasting performance of the hybrid approach with the machine learning methods. Hourly interval powers are forecast. In the verification, the controlled houses were in six groups and the four biggest groups are shown here.

Table 4. Comparison of machine learning with the hybrid methods over the verification period.

	RMSE (normalized)			
Method	**Group 1**	**Group 2**	**Group 3**	**Group 4**
SVM alone	0.1457	0.1756	0.6850	0.6866
MLP alone	0.1283	0.1651	0.8904	0.9622
SVM with response model	0.1161	0.1290	0.3801	0.3767
MLP with response model	0.1108	0.1361	0.4758	0.4622

Table 5. Comparison of machine learning with the hybrid approach when emergency load control was applied in verification; RMSE is evaluated over two 48 h periods, one for each of the two control actions.

	RMSE (normalized)			
Method	**Group 1**	**Group 2**	**Group 3**	**Group 4**
SVM alone	0.2180	0.2784	1.1122	1.1380
MLP alone	0.2037	0.2880	1.3519	1.4536
SVM with response model	0.1162	0.1350	0.5681	0.5820
MLP with response model	0.1126	0.1750	0.6186	0.6493

SVM and MLP produced roughly equal accuracy and they could not predict emergency control load situations. By combining the methods with the physically based response model, also the dynamic control situations were predicted with good accuracy. We prefer the use of SVM models, because MLP has many well-known challenges, such as a risk of over-fitting.

In the verification, the groups were different from the identification. In the identification, the average customer size was the same in all the four main groups. In the verification, the average customer size in the groups 3 and 4 was clearly smaller. The change in the group size from identification to verification resulted in large errors in the forecast. We compared two solutions: 1) each identification test group was split to two subgroups based on the average annual power thus enabling the machine learning to learn the dependence on the average group, and 2) the hybrid forecast was scaled using feedback from the measurement history available when making the short-term forecasts. The same feedback from the measured average power of customer scaled the partly physically based control response model in all alternatives.

Often the on-line feedback scaling of the response model turned out to be the most accurate although the feedback scaling took about two first weeks of data history to converge to a suitable feedback gain. Table 6 shows the results of the comparison.

Table 6. Modelling the dependence on average site power of the group.

Normalized RMSE	**Without scaling**		**Identification from data split based on customer size**		**Feedback scaling to group mean size**	
	Machine learning with response model					
	MLP	SVM	MLP	SVM	MLP	SVM
Group1	0.1108	0.1161	0.1712	0.1587	0.1047	0.1122
Group2	0.1795	0.1290	0.1803	0.1717	0.1431	0.1143
Group3	0.4758	0.3801	0.1365	0.1333	0.1640	0.1718
Group4	0.4622	0.3767	0.1657	0.1330	0.1820	0.1639

Figures 5 and 6 show the machine learning forecasts, hybrid forecasts and the measured responses. The hybrid forecast is the sum of the physically based response forecast and the machine learning forecast of the residual. Alone the machine learning models are not able to forecast the emergency load control responses, see also Table 5.

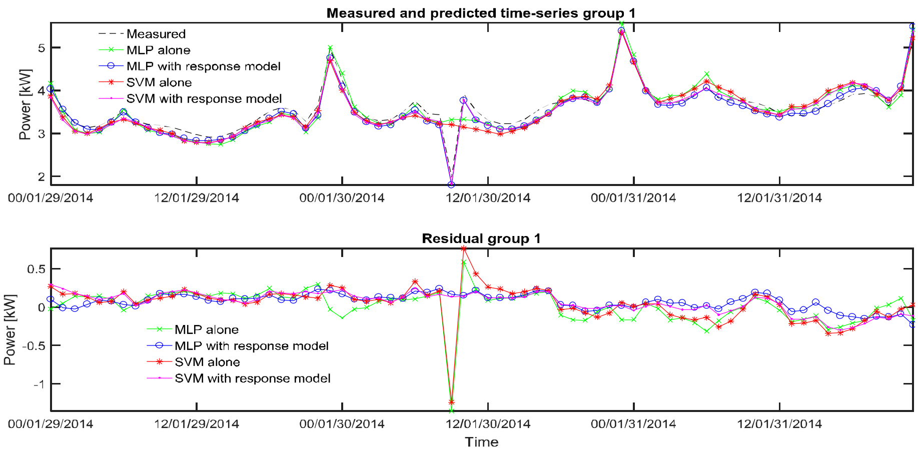

Fig. 5. The responses of the forecasting methods during emergency load control 30 January 2014 in the verification.

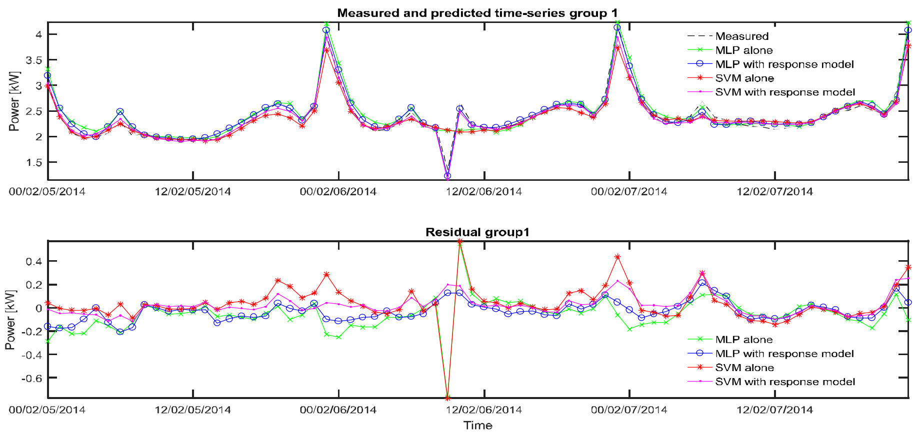

Fig. 6. The responses of the forecasting methods during emergency load control 6 February 2014 in the verification.

8 Discussion

Planned future studies include:

- forecasting the total power of the distribution area with 3 minute time resolution when dynamic load control is applied,
- other hybrid methods in AD forecasting,
- on-line implementation and field testing of the response forecasting,
- tests in cold temperatures,
- analysis and development of criteria for the performance of load forecasting, and
- estimating confidence intervals for the forecasts.

Commonly applied performance criteria reflect poorly or very poorly the costs of forecasting errors. Selection and development of performance criteria should be considered. Splitting the analysis to four groups enables getting some information on the confidence intervals of the forecasts, but further studies are needed.

Common claims are that 1) real time measurements of individual AD customers are necessary and 2) determination of a fair base case for reference and thus the actual response is ambiguous. Individual customer real time measurements improve the performance of forecasting aggregated loads so little that they may be difficult to justify. Our models always forecast the base case or reference case in addition to the responses to the planned control actions. Further studies could clarify these issues.

9 Conclusion

The results show that the hybrid model developed forecasts more accurately than the machine learning models as such. In the hybrid model, the control responses and load saturation are forecast using a physically based structure and the residual is forecast with a machine learning model designed and tuned to learn also system dynamics. The hybrid load forecast is the sum of these component forecasts.

References

1. Garulli, A., Paoletti, S., Vicino, A.: Models and techniques for electric load forecasting in the presence of demand response, *IEEE Transactions on Control Systems Technology*, Vol. 23, No. 3, May 2015, pp. 1087-1097.
2. Koponen, P., Niska, H.: Hybrid model for short-term forecasting of loads and load control responses, *IEEE PES ISGT Europe 2016*, 9-12 October 2016, 6 p.
3. Csáji, B. Cs., Kovács, A., Váncza, J.: Online learning for aggregating forecasts in renewable energy systems, *ERCIM NEWS 107*, October 2016, pp. 40-41.
4. Valenzuela, O., Rojas, I., Rojas, F., Pomares, H., Herrera, L.J., Guillen, A., Marquez, L., Pasadas, M.: Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems*, Vol. 159, Elsevier 2008, pp. 821-845.
5. Selakov, A., Ilic, S., Vukmirovic, S., Kulic, F., Erdeljan, A.: A comparative analysis of SVM and ANN based hybrid model for short term load forecasting, *Transmission and Distribution Conference and Exposition, 2012 IEEE PES*, 7-10 May 2012, Orlando, FL. 5 p.
6. Koponen, P.: Shor-term load forecasting model based on smart metering data, daily energy prediction using physically based component model structure, *IEEE SG-TEP 2012,* Nuremberg, 3-4 December 2012, 4 p.
7. Hahn, H., Meyer-Nieberg, S., Pickl, S.: Electric load forecasting methods: tools for decision making. *European Journal of Operational Research,* Vol. 199, Elsevier 2009, pp. 902-907.
8. Vapnik, V.N.: *The nature of statistical learning theory*. New York: Springer, 1995, 188 p.
9. Niska, H., Koponen, P., Mutanen, A.: Evolving smart meter data driven model for short-term forecasting of electric loads. *IEEE ISNNIP 2015*, Singapore, 7-9 April 2015, 6p.
10. Koponen, P.: *Optimisation of load control, Final report*, VTT Energy. Espoo, 20 November 1997, 26 p. + app. 14 p. Research report ENE6/12/97, http://www.vtt.fi/inf/julkaisut/muut/1997/R-ENE6-12-97.pdf