Load ControlResponse Forecast using Machine learning Model

Debarpita Roul
College of Engineering Bhubaneswar, Odisha, India

Abstract: The predictability of active demand's aggregated control reactions is a critical factor in determining its value in the energy and ancillary services markets. The authors of this article investigate compact electrically heated homes equipped with heat storage tanks and remote control via smart metering. They combine the advantages of the individual approaches by integrating a machine learning forecasting technique with a basic physically based model. Now, a support vector regression, an earlier machine learning model, is briefly contrasted with a stacked boosters network, a new deep learning technique. The house's outdoor-dependent heat demand and the thermal dynamics of the heat storage tank are both modeled using the basic physically based model component.

1 Introduction

Often the value of dynamic demand response is small or even negative if the responses are unpredictable and unreliable. Traditional load forecasting methods completely fail in the presence of large dynamic load control actions. Thus, it is necessary to develop and use methods that forecast also the control responses accurately.

As a solution, we study hybrid forecasting models that aim to combine the strengths of different modelling approaches and to forecast more accurately than their component models applied separately [1]. The hybrids include a machine learning method that forecasts the residual of the other model component that is a simple physically based model forecasting the dynamic load control responses.

This paper studies short-term load forecasting in presence of dynamic load control. The aim is to forecast at 9 a.m. for each hour of the next day the aggregated power of loads that include active demand. The methods produce also forecasts for the same day and for the day after tomorrow, but the performance indicators are easier to understand if the forecasts do not overlap each other.

The studied case comprises 727 small houses that are heated by electricity via a large heat storage tank. The heating of the heat storage is remotely controlled via a smart metering system. In this study, the houses were controlled in two separate groups. Group1 had 350 houses and group 2 had 377 houses. The aggregated power of the groups must be forecast accurately and the load control is dynamic. Hourly interval metered power consumption of each house is available from the previous day and before. Dynamic market-based load control was applied both in the model identification period (31 May 2012–31 May 2013) and the verification period (1 January 2015–31 December 2019). A more detailed explanation of the load control system and one of the dynamic load control variants are in [2]. The aggregated mean power in the verification was about 1.235 MW in group 1 and 1.234 MW in group 2. The respective aggregated peak powers were 9.325 and 9.581 MW.

We first present a simple physically based model of the aggregated heat dynamics of the houses. It forecasts the heating needs, the expected load curves and the aggregated load variations in response to dynamic control signals. Next, we show how we improve the forecasting accuracy by adding a machine learning model to forecast the residual of the physically based model. Finally, the achieved results are summarised.

2 Methodology

Physically based response model

We applied the simple physically based model as shown in Fig. 1. The main component is a simple first-order model of the heat storage tank. The state variable of the model corresponds to the state of charge (SOC) of the heat storage tank. The model inputs comprise a control signal and the outdoor temperature. We forecast the daily heat demand of the building from the measured and forecast outdoor temperature according to an empirically identified relation. The forecast heat demand, the SOC, and the control signal define together when the heating is on in the model. Heat losses and saturation are also included in the model. The saturation effects depend on the dimensioning of the heating system and we identified them from the identification data.

An earlier version of the simple physical model did not explicitly model the SOC of the heat storage tank and was explained in [3].

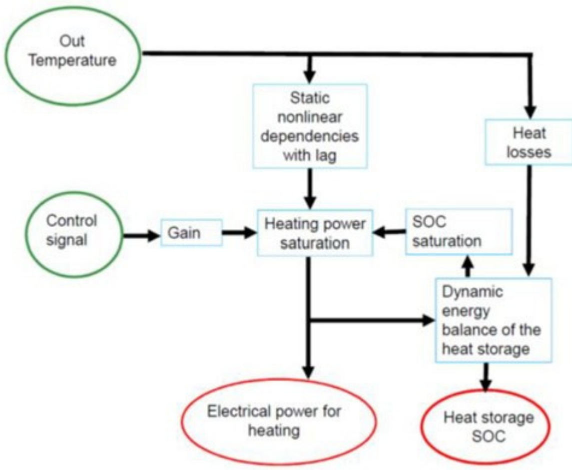

Fig. 1 *Physical heat balance model*

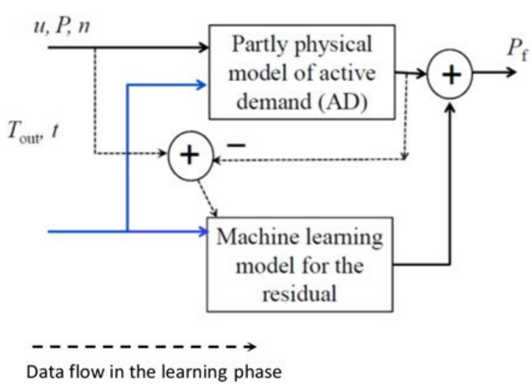

Fig. 2 *Residual hybrid forecaster*

Residual hybrid forecaster

The main structure of the hybrid forecasting model is shown in Fig. 2. The input variables include time t, outdoor temperature T_{out}, and, for every controlled group i, the control signal u_i, past hourly interval power P_i and the number of sites n_i. Partly physical models forecast the responses for each controlled group and machine learning is taught to forecast the residual. The result is the forecast power of the controlled customer group P_f. It is the sum of the partly physical forecast and the forecast residual.

Stacked boosters network (SBN)

SBN [4] is a novel deep learning architecture designed for short-term load forecasting. The network consists of a simple base forecaster and multiple boosting forecasters. Each boosting forecaster operates on the one-time scale: week, day or hour, and corrects systematic errors occurring in that time scale. The applied version of the network is the same as in the original paper [4] with the only exception that the simple base forecaster takes the control signal as an input in addition to the outside temperature.

Support vector regression (SVR)

SVR is a machine learning technique for data classification and non-linear regression. The main technical details of SVR are explained in [5]. Nu(n)-SVR with the radial basis kernel function implemented in the scikit-learn package was used to execute the model runs [6, 7].

With the SVR we modelled the time dynamics using input delays and defined the structure using a genetic algorithm and sensitivity functions [8]. The method selected delays from a set where the longest delay was 48 h and time resolution 1 h. SVR contains the control parameters (g, n, C), which define the kernel width (g), an upper bound on the fraction of training errors and a lower bound on the fraction of support vectors (n), and a regularisation parameter affecting the margin size (C). Based on experimenting, n = 0.5, C = 1.0 and $g = 1/M$, where M is the number of inputs, were selected for both SVR and SVR-hybrid models.

The SVR models as described above may provide biased estimates when the load behaviour changes, because as such they may not include enough feedback from the past prediction error history. That is why we made also versions SVR-b and SVR-hybrid-b by adding a slow first-order low-pass filter with a time constant of 3 weeks to remove the bias in the forecasting error. We leave to further research the more advanced solutions to improve the adaptation of the above SVR methods to typical load behaviour changes.

3 Results

Comparison of the forecasting accuracy using NRMSE

Tables 1–3 compare the forecasting accuracy of the different studied methods. The criterion is NRMSE, the root mean square error normalised to the mean measured power over the period studied.

In the identification, using the physical control response model improved the fit. For group 1 the fit of the SBN was worse than for the other methods. Further analysis is needed in order to find the reasons. In spite of this, SBN had a good performance in the verification also with group 1.

Table 1 Comparison of forecasting accuracy using NRMSE% in the identification

Method	Group 1	Group 2
physical	28.5	28.7
SBN	37.8	28.1
SVR	23.4	18.5
SVR-b	23.4	18.6
SBN-hybrid	22.9	21.6
SVR-hybrid	23.3	21.6
SVR-hybrid-b	23.3	21.7

Table 2 Comparison of forecasting accuracy using NRMSE% in the verification for group 1

Method	2015	2016	2017	2018	2019	All
physical	48.1	38.0	29.9	29.0	32.4	36.0
SBN	43.9	39.7	34.4	30.1	32.0	36.1
SVR	51.7	47.5	45.4	39.0	33.2	44.0
SVR-b	50.2	44.7	40.7	33.2	30.1	40.5
SBN-hybrid	38.6	32.1	23.8	20.1	22.0	27.8
SVR-hybrid	47.9	36.8	27.9	25.3	30.8	34.4
SVR-hybrid-b	47.6	36.4	27.3	24.5	29.9	33.9

Table 3 Comparison of forecasting accuracy using NRMSE% in the verification for group 2

Method	2015	2016	2017	2018	2019	All
physical	47.1	41.8	42.5	39.9	41.1	42.6
SBN	47.9	45.5	38.3	32.0	37.8	40.6
SVR	45.2	47.1	41.8	33.1	30.0	40.4
SVR-b	45.1	46.6	40.8	31.9	29.6	39.8
SBN-hybrid	40.2	37.3	32.7	26.8	31.0	33.9
SVR-hybrid	44.1	41.0	40.5	35.6	35.8	39.6
SVR-hybrid-b	43.9	40.8	40.4	35.5	35.7	39.5

The hybrid methods, especially the SBN hybrid, had the best forecasting performance in the verification. In 2015 and 2016, the forecasting performance was poor, because the load control was subject to communication performance tests and data communication failures made the responses unpredictable. In group 2, the physically based model identified for group 1 was used, because in the identification the fit (NRMSE 28.7%) was almost as good as with the model identified from group 2 (NRMSE 26.8%). The above verification results do not support that choice.

Comparison of the forecasting error time behaviour

Figs. 3–5 show the envelopes of the forecasting errors (residuals) of the forecasting methods over the whole 5-year long verification period. Because the large forecasting errors are of interest, we omit zooming these figures.

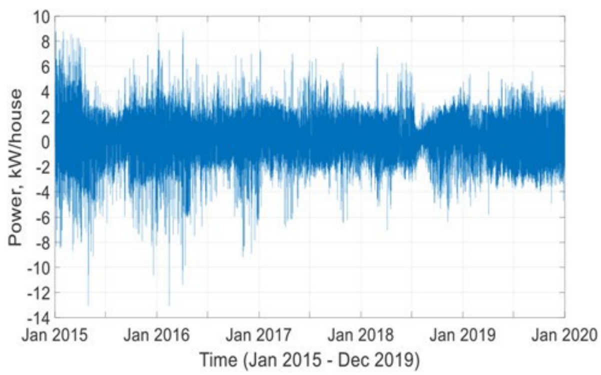

Fig. 3 *Residual of the physically based response model of group 1 in verification*

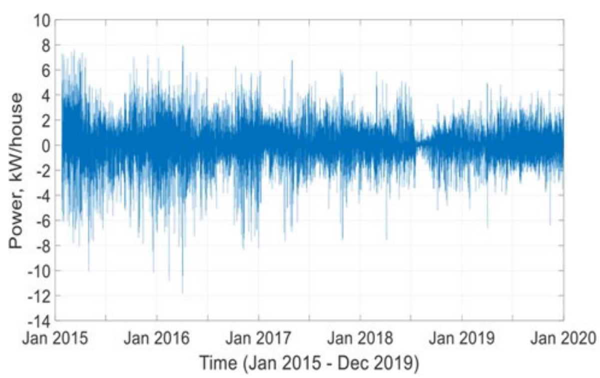

Fig. 4 *Residual of the SBN-hybrid of group 1 in verification*

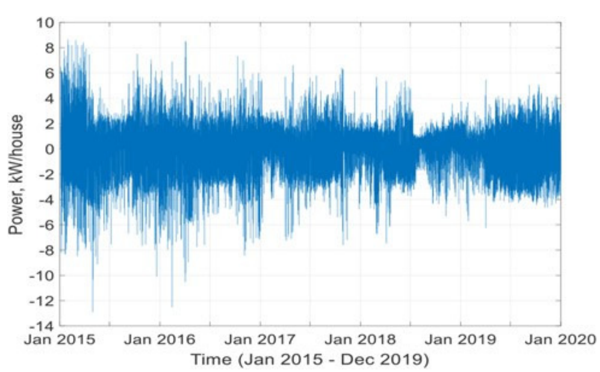

Fig. 5 *Residual of the SVR-hybrid of group 1 in verification*

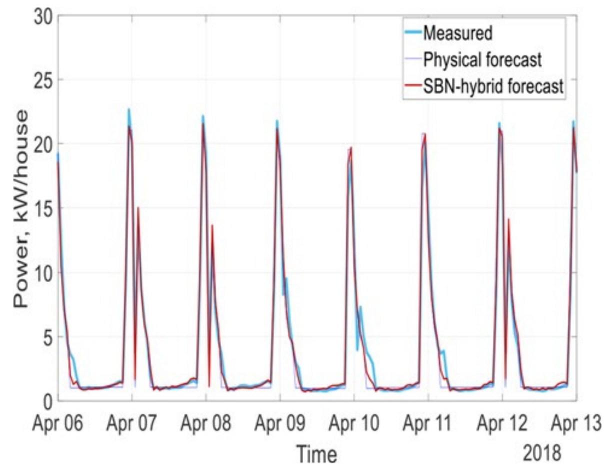

Fig. 6 *Sample of aggregated measurements and forecasts*

Forecasts compared with the measurements

Fig. 6 shows the measured load, the physical forecast and the SBN forecast during a week in the verification period. When the dynamic control signal allows the heating of the heat storage, the load shows high peaks. The control signal turns on heating in those night hours when the electricity price in the day-ahead spot market is the lowest thus minimising the heating electricity costs. In Fig. 6 in some night hours, the control signal has turned heating off as can be seen from the gaps in the heating period. The physically based forecast and the hybrid methods forecast this correctly.

Discussion

The comparisons show relatively high NRMSE values due to the following reasons. (i) The size of the groups is small and thus the stochastic behaviour of individual houses does not completely cancel out in the aggregation. (ii) The loads are dynamically controlled which makes the forecasting task challenging. (iii) The mean load is very small compared to the high load peaks and relatively small forecasting errors of the peak heights are rather large compared to the mean load. In [1] we used a roughly similar hybrid forecasting approach as the SVR hybrid method here for forecasting the power of a distribution area with about 60,000 customers 8000 of them being dynamically controlled, the NRMSE was well below 4%.

The simple physically based forecasting model was used in a simulation study for the estimation of the benefits and costs from the participation of this flexible demand to an ancillary service market. The project EU-SysFlex will report the results. The hybrid models can improve the accuracy of the simulations in possible future studies, if needed.

4 Conclusion

The integration of the machine learning methods (SBN and SVR) with the physically based response model improved the forecasting accuracy in the studied short-term load forecasting problem that has much dynamic load control. The SBN-hybrid had the best prediction performance in the comparison. Accurate forecasting of the control responses improves the value of dynamic demand responses. An additional benefit is that the same physically based response model is also a suitable simulation model for analysing the benefits, costs and CO_2 emission impacts of potential uses of dynamic demand response.

5 References

1 Koponen, P., Niska, H., Mutanen, A.: 'Mitigating the weaknesses of machine learning in short–term forecasting of aggregated power system active loads'. IEEE INDIN19, Helsinki-Espoo, Finland, 22–25 July 2019, p. 8

2 Koponen, P., Seppälä, J.: 'Market price based control of electrical heating loads'. Proc. of the 21st Int. Conf. on Electricity Distribution, CIRED 2011, Frankfurt, 6–9 June 2011. Paper 0796, p. 4

3 Koponen, P., Takki, P.: 'Forecasting the responses of market based control of residential electrical heating loads'. CIRED 2014 Workshop, Rome Italy, 11–12 June 2014, Paper 0178. p. 5

4 Salmi, T., Kiljander, J., Pakkala, D.: 'Stacked boosters network architecture for short term load forecasting in building'. Arxiv Prepr. 2001.08406, 2020

5 Vapnik, V.N.: 'The nature of statistical learning theory' (Springer, New York, 1995), p. 188

6 Schölkopf, B., Smola, A., Williamson, R.C., *et al*.: 'New support vector algorithms', *Neural Comput.*, 2000, 12, pp. 1207–1245

7 Pedregosa, F., Varoquaux, G., Gramfort, A., *et al*.: 'Scikit-learn: machine learning in python', *JMLR*, 2011, 12, pp. 2825–2830

8 Niska, H., Koponen, P., Mutanen, A.: 'Evolving smart meter data driven model for short–term forecasting of electric loads'. Tenth Int. Conf. on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Singapore, 7–9 April 2015, p. 6