21, rue d'Artois, F-75008 PARIS
http://www.cigre.org

PS1-04

D2 Colloquium 2019

Combining the strengths of different load modeling methods in short–term load forecasting of a distribution grid area with active demand

P. KOPONEN[1], H. NISKA[2], A. MUTANEN[3]
VTT Technical Research Centre of Finland[1], University of Eastern Finland[2], Tampere University[3]
Finland[1,2,3]

SUMMARY

Power flows are becoming increasingly volatile in power distribution grids. Distributed power generation, electricity storage, electrical vehicles and active demand cause these variations. Dynamic management of constraints, power quality and balancing will be needed. The accurate forecasting of the power flows is a necessary enabler for it. It is important to accurately forecast the whole power balance including distributed generation, storage and loads.

Much measurement data are now available which allows machine learning methods to become popular. A major challenge is that such black box methods are poor when significantly outside the situations included in the learning data. Especially in the presence of dynamic active demand, the black box models completely fail. Physically based model structures forecast those situations reasonably accurately, but require much domain expertise and development work. Modern load curve approaches fit well with grid state estimation and simulation, but like black box machine learning, the models require much more identification data than the models with physically based structures. Hybrid approaches aim at combining the strengths of different forecasting approaches and mitigating their individual weaknesses.

This paper studies short–term load forecasting in a distribution area with about 9000 active consumers subject to both emergency and Time of Use load control. We integrate several modelling approaches to hybrid models to combine the strengths of the component models and avoid the weaknesses of the individual approaches. The component models include 1) models with physically based structures, 2) different machine learning methods, and 3) a similar day forecaster. We developed the methods and analysed their performance using field tests with load control actions and measurements from smart meters, distribution grid and weather.

There are many model hybridisation methods. Here we mainly use sequential modelling of residuals. We begin with the physically based modelling and then model the residual using the other methods. The resulting forecast is the sum of the component methods. We also use physically based models to constraint the other forecasts to remain in a reasonable range and apply simple ensemble forecasting.

The hybrid models were more accurate than the individual component models. The superiority is especially clear in exceptional situations and during dynamic load control actions. A further advantage is that the forecasting task for the other methods becomes easier when the models with physically based structures remove some fast and complex phenomena from the remaining forecasting task.

KEYWORDS

Short–term forecasting -Hybrid forecasting -Machine learning -Big Data

pekka.koponen@vtt.fi

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].

It is increasingly popular to improve forecasting accuracy by combining different forecasting methods to hybrid methods [3]. There are many approaches how to integrate methods into hybrids, including but not limited to 1) ensemble forecasting, for example [4], 2) decomposition to separately estimated load components [5], 3) using one model to limit the forecast of the other model, 4) using one method to forecast the parameters of another method and 5) forecasting the residual of another model. ARIMA models and machine learning were combined using the residual approach in [6]. In our case, the outdoor temperature has large variations and the ARIMA models we found were very inaccurate and unreliable due to nonlinear and nonstationary behaviour of the loads and limited number of dynamic control tests. Thus, we here use models with physically based structures to forecast the control responses, and machine learning models and similar day methods to forecast the residual. The load forecast is the sum of these component forecasts. With this approach, we forecast the hourly interval powers for spot price based control of aggregated loads of full storage heating houses [2]. Also a separate partly physically based forecast for the range of the combined forecast was added to improve the accuracy and reliability further.

The present contribution applies the hybrid approach of [2] to forecasting the next day power of the whole distribution area with 3–minute interval, when two types of load control of partial storage heating houses are applied. We modelled the control responses from aggregated 3–minute interval measurements from primary substations and the hourly interval measurements from the smart billing meters. The results regarding forecasting the hourly interval powers of four controlled groups are in [7]. Now we use the same emergency control response models to forecast the 3–minute interval powers of all controlled five groups, thus forecasting the response for the whole power distribution area. We also add separate partly physically based models for the Time of Use (ToU) load control. The remaining residual is large. We forecast it by machine learning and/or a similar day forecaster. We compare two machine learning methods: support vector machine (SVM) and multilayer perceptron (MLP). It seems reasonable based on a more extensive performance comparison of data driven models in forecasting residential electricity consumption [5]. According to the literature, such as [8], SVM has many benefits, such as good accuracy and insensitivity to outliers. However, SVM is known to be computationally inefficient [5] especially in large problems. We also study adding a similar day forecasts to the hybrids.

A genetic algorithm is used to optimise the structure of the machine learning models. Manual selection tends to require too much effort and is too unreliable to produce comparable results. We identify the parameters of the partly physically based models using nonlinear optimization with physically based constraints.

We simulated and tested the approach with the real field test data of about 7500 active customers and grid measurements. The results achieved and lessons learned regarding the hybrid modelling are explained. The results support the hypothesis that the hybrid models forecast more accurately and reliably than the machine learning models alone, especially, when dynamic load control is applied.

2 CHALLENGES WITH THE MACHINE LEARNING

Much new energy demand related data have become available from various sources such as smart metering, distribution grid automation, building automation, and new public and private databases. Purely data driven methods tend to fail in transforming the data to useful new information. The specific challenges include the following. 1) The system behaviour changes and purely data driven models tend to need so much learning data that the learned model is immediately outdated. 2) Purely data driven models often fail in situations that are not included in the learning data. Good forecasting accuracy is especially important in these situations. 3) Crucial information may be lost in the pre-processing of the identification data and not detected before the forecasting fails. 4) Existing load forecasting models typically model demand as passive. Thus, they fail when forecasting in the presence of substantial amounts of active demand. 5) The MLP identification often fails to converge properly thus requiring repeated runs with different parameters. 6) SVM scales poorly to large problems. 7) Modelling the time dynamics has been rather exploratory and the experience on different approaches is still rather limited.

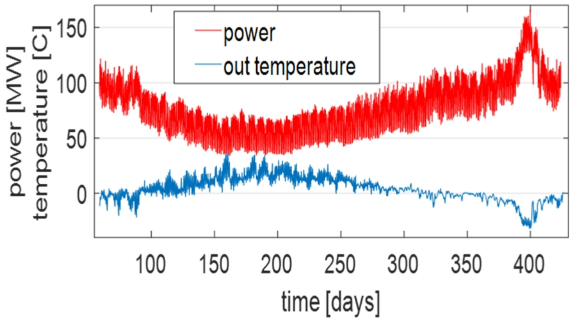
Machine learning has also relative strengths that largely out weight the above challenges. Physically based models are superior in forecasting active demand responses and many new or changing situations, but in forecasting the total load they require very much domain expertise and development work and still most of the time have inferior accuracy compared with machine learning. New load profiling approaches also have their strengths and weaknesses. We compared these approaches in [9].

3 THE SHORT–TERM FORECASTING PROBLEM

The problem studied is to forecast the aggregated power of a power distribution area that includes two types of active demand (AD): 1) emergency load control and 2) ToU control. The focus is on short–term forecasting: each day at 9 a.m. the power during the next day is forecast using 3–minute time resolution. The forecasting of hourly interval powers for electricity market purposes is as in [7].

In this case, the AD comprises over 9000 hourly interval metered electricity customers. The distribution network operator can send a signal that temporarily switches off daytime heating loads and possibly also cooling loads. In addition, all the AD houses have remotely controlled nigh time electrical heating.

The power of the power distribution area is measured at the primary substations with 3-minute time intervals. Hourly interval consumption measurements from the previous day are available from each customer. We used the outdoor temperature and its forecast for Kajaani, the central city of the power distribution area. The impacts from solar radiation, snow cover, wind speed and humidity are much smaller and are ignored here for clarity. The identification period was 12 months long (March 2011 to February 2012), Fig. 1, and the verification period was the two first months of the year 2014, Fig. 2.



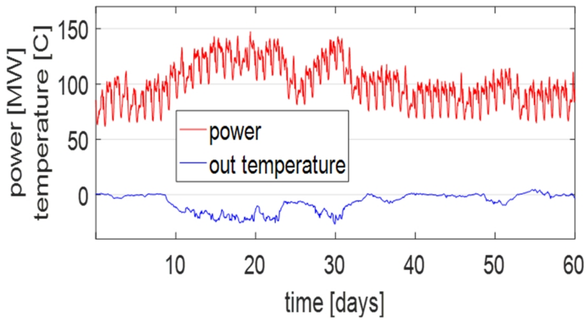


Figure 1: The identification period, starts 1.1.2011 **Figure 2 The verification period, starts 1.1.2014**

4 METHODS

4.1 The main hybrid approach, sequentially forecasting residuals

Fig. 3 shows the main structure of the sequential residual based hybrid forecasting model. Instead of SVM also MLP and a similar day forecaster were used. The input variables include time t, outdoor temperature T_{out}, and, for every controlled group, the AD control signals u, past hourly interval power P_h and the number of sites n. Partly physical models forecast the AD responses for each controlled group and machine learning is taught to forecast the residual. In addition other partly physically based models forecast daily energy demand of each controlled group and the feasible ranges of the power thus improving the forecast slightly during exceptional weather situations, for example. The range forecast based limiter improves accuracy [7], but now we omit it for clarity. The result is the forecast grid area power P_f. Fig. 3 includes a block "physical model of AD". This block comprises many partly physical models in parallel, because each controlled group has its own AD control signals and consequently its own partly physical AD model. Each controlled group also has a separate physically based model for emergency load control and another for ToU load control.

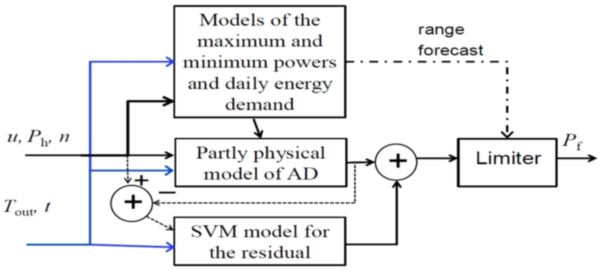


Figure 3: Machine learning forecasts the residual of the partly physically based response model.

4.2 Partly physically based control response models

The structure of the control response models is based on the thermal dynamics of a typical small house. A typical thermodynamic model for forecasting the loads of individual buildings is applied in [5]. Such models have so many parameters that parameter identification is difficult and the accuracy poor in forecasting the aggregated energy consumption of many houses. Thus, we use relatively simple thermodynamic models. The emergency load control and Time-of-Use control have separate models.

4.2.1 Emergency control model

The model for the emergency control responses was developed in [10]. Outdoor temperature T_{out} and the group control signals u_g are the input variables. The state variables comprise four internal temperatures and one temperature controller state. Each internal temperature is associated with a thermal storage capacity. In the real houses, the temperature controllers 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 behaviour unless a large number of models with stochastic disturbances is run in parallel. Thus we use a continuous controller in the aggregated house model that accurately forecasts the aggregated responses also when the heating in the individual houses is controlled on-off. The model output power is scaled by applying a slow first order filtered feedback from P_{hg}, the measured hourly interval consumption of the group, and n_g, the number of houses in the group.

The parameters of the aggregated house model were fitted to the measured aggregated responses using nonlinear constrained optimization. Constraints for parameter ranges were estimated from the building requirements of the climate zone. The emergency load control tests in the identification period did not include so cold temperatures that the identification of heating power saturation would have been possible. In three neighbour power distribution networks, similar aggregated house models had been similarly identified in 1996–1997 from the primary substation power measurements. Then the tests included cold enough temperatures, 13 substations and 6 separately controlled load groups [10]. The measurements covered one year. The model of a large group had very good agreement with the responses identified in the new identification data. Thus we here use that very old aggregated house response model as such as the partly physical emergency load control model in the analysis and results.

The output of the partly physical emergency control model is multiplied by a linear site size dependence function identified from the new identification data (2011 and 2012). The dependence function needs an average site size estimate as input and gets it using slow feedback from the measurements (time constant several weeks) and an even slower forgetting factor.

4.2.2 Time of Use control model

For modelling the Time of Use (ToU) control responses, we first identified how the energy consumed during the day depends on the outdoor temperature and how the temperature dependence differs according to the state of the ToU control signal. Also the temperature dependence of the load increase and decrease at the change of the control signal state were identified. Then we developed a simple heat storage model with first order dynamics and state constraints. It can only fit to the first hours of the control response, when the heating turns on, because in the identification data the load typically increases at the end of heating. Adding a load component that exponentially increases with time improves the fit during the last few hours of the heating period. The model parameters were identified from the identification data. We prefer using minimalistic models, because forecasting the step changes is the main purpose of this ToU response model and the machine learning models applied to the residual model the other aspects of the response well enough.

4.3 The machine learning methods

Two machine learning methods are compared: 1) support vector machine (SVM) and 2) multilayer perceptron (MLP). We used direct prediction scheme for both the machine learning models by using delayed power and temperature values as regressors. Many alternative approaches for modelling the time dynamics of the nonlinear system, such as long short–term memory [11], Elman neural network [12], and [13], use feedback from delayed output values as inputs to the neural network.

SVM is a machine learning technique for data classification and nonlinear regression. The main technical details of SVM are explained in [14]. Epsilon(ε)–SVM with the radial basis kernel function based on the LIBSVM package was used to execute the model runs.

The MLP was trained using Levenberg–Marquart modification of the back propagation. The algorithm iteratively adjusts the weights of the squared errors between the forecast and measurement outputs. One hidden layer was used. More detailed explanation is in [15].

Table I describes inputs variables for the machine learning models. A multi-objective genetic algorithm was applied with sensitivity analysis to select an optimal subset of the inputs variables [15]. A tedious and poorly reproducible trial-and-error effort was thus avoided. We transformed discontinuous timing variables into continuous form using trigonometric transformations. We smoothed the hour of the day to minute level indices using sliding average with one hour window.

It is necessary to control the risk for overfitting when applying machine learning models. In case of the MLP network, we adopted the standard method called early stopping (with 5% sampling of identification data). Contrary to the MLP network, SVM contains the control parameters (ε, C), which define the margin within which the error is neglected (noise) and the smoothness of the approximation, respectively. Values C=100 and ε=0.01 were used based on experimenting.

Table I: Machine learning model inputs and their physical interpretation

Inputs to be optimally selected	**Physical sub load**
Day of the year (1– 365)	Domestic appliance seasonal rhythm
Day of the week (1– 7)	Domestic appliance weekly rhythm
Hour of the day (1– 24)	Domestic appliance daily rhythm
Day length (hours)	Lighting, radiation affected thermal load
Outdoor temperature (ºC) with time-lags of 1– 48 hours	Thermal load (heating and cooling)

4.4 The similar day forecaster

We developed a similar day forecaster for the load in the studied distribution area. In this forecaster, the load on each 3-minute interval is forecast based on earlier intervals with similar characteristics. Typically, there are several similar intervals and average load on these intervals is used. Table II gives the considered characteristics and their averaging windows.

Unlike in some similar day forecasters presented in literature, such as [16] and [17], we modelled the load's dependency on the outdoor temperature separately. The temperature dependency (W/°C) for each forecasted day is determined with simple linear regression. The effects of intra–week fluctuations in electricity demand are eliminated by choosing the dependent and independent variables as follows: Dependent variable is the difference between the daily energy consumption and the average daily energy consumption on similar days of the week. Independent variable is the difference between the daily average of effective hourly temperatures and the average of effective hourly temperatures on similar days of the week. The effective hourly temperature is defined as an average over the previous 24 hourly temperatures. In regression analysis, the effects of seasonal fluctuations are eliminated by using data from only similar days of the year. The identified temperature dependency is then used to correct the average load of similar intervals to correspond the load in the forecasted temperature. Finally, the systematic forecasting errors, possibly caused by rising or falling trends in electricity consumption, are corrected based on (uncorrected) forecasting errors on preceding 30 days.

Table II: Characteristics defining the similar interval

Characteristic (index range)	Size of the moving averaging window
Day of the year (1– 365)	±15 days from the identification data and previous 15 days from the verification data
Day of the week (1– 7)	0 (Must be exactly the same)
Public holiday or other special day (0– 17, 0=normal day)	0 (If index>0, day of the year and day of the week are ignored)
3-minute interval of the day (1– 480)	±1 interval
ToU control signals (0 or 1, two control signals)	0

In this study, similar weights were given to all similar days and intra-day fluctuations in the temperature dependency were not modelled. All 3–minute intervals had at least one similar interval in the measurement history, so we had no need to consider situations with no similar intervals. If ToU control is dynamic, the lack of similar intervals becomes a problem, unless the hybrids mitigate it.

5 RESULTS

5.1 Forecasting the hourly interval powers of controlled customer groups

We studied the forecasting of the hourly interval responses of AD customer groups in [7]. The hybrid model comprises the emergency control models and a machine learning model. The partly physically based ToU response model was not included. Table III compares methods in a 14-month verification. It clarifies and amends the analysis of [7]. In Table III, the performance metric is NRMSE, the Root Mean Square Error normalised to measured mean power over the verification period. Because there are no large errors in the pre-processed measurement data, the sensitivity of RMSE to outliers is not a problem. The number of control events is so small that their impact on these results is very small.

Table III: Forecasting the hourly interval powers in verification, adding model gain scaling feedback

NRMSE %	SVM		MLP		SVM hybrid		MLP hybrid	
Group 1	7.73	7.81	6.81	6.59	6.16	6.49	5.88	6.05
Group 2	9.14	7.96	8.59	7.10	6.71	6.45	9.34	8.12
Group 3	63.27	20.97	82.24	16.94	35.10	15.77	43.95	15.09
Group 4	63.86	20.28	89.49	15.24	35.04	15.17	42.99	16.76

Machine learning forecasts the residual of the partly physically based models, except for the two leftmost methods, where no control response model is applied. According to the results of the groups 1 and 2, the performance of the hybrid approach is slightly better than using the machine learning alone, even when load control is not applied. The poor results of groups 3 and 4 stem from a large change in the average site power of the group compared with identification and the loss of information about this dependency in the grouping applied in the identification. Especially the machine learning methods fail. The right column of each method gives results when slow feedback using the identified group size dependency is applied. The results were similar when we split the groups in identification according to their size [7]. Some other results in [7] focus on 48 hour long time periods that include control actions; according to them the performance of machine learning alone is not satisfactory in the presence of active demand. The hybrids maintain good performance also during those periods. In the group 1, the NRMSE is 9.18 % and 7.26 % for SVM and MLP alone and 5.44 % and 4.97 % for the SVM and MLP hybrids, and in the group 2, they were 11.14 %, 8.48 %, 6.29 % and 6.35 % respectively in a typical run. In the groups 3 and 4, the other forecasting errors largely hide the benefit from the hybrids.

5.2 Forecasting the 3-minute power of the power distribution area

5.2.1 Hybrids of physically based response models and machine learning

We use the same emergency control response models for forecasting the control responses in hourly interval responses of the controlled groups and in the 3-minute interval power of the distribution area. The time step of the response model is 18 s. We forecast the responses for five separately controlled groups comprising over 7500 hourly metered customers, and sum up the results. The verification results in [7] represented only 5188 selected customers in four of the groups, because then coarse identification and elimination of the anomalies in the data was sufficient for the forecasting task studied.

Table IV compares the forecasting performance of the methods studied. These four methods comprise the two machine learning methods both as such and with the proposed hybrid approach. Root Mean Square Error (RMSE) in MW is the performance criterion. The average measured power in the identification period was 77.7 MW and in the verification it was 103.9 MW.

Table IV: Forecasting of 3 minute interval power of the distribution area, machine learning hybrids

RMSE in MW and (NRMSE in %)	Identification	Verification
MLP	2.5936 (3.34)	4.7783 (4.60)
SVM (5% sampling)	3.4444 (4.43)	5.2658 (5.07)
Partly physical emergency control response model and MLP	2.5603 (3.30)	4.2772 (4.11)
Partly physical emergency control response model and SVM (5%)	3.4159 (4.40)	5.2414 (5.04)
Partly physical emergency control and ToU response models and MLP (MLP hybrid)	2.4454 (3.15)	3.8831 (3.74)
Partly physical emergency control and ToU response models and SVM (5%)	2.7849 (3.58)	4.2161 (4.03)

The verification period comprised the two first months of the year 2014. The identification period was one year starting the 1st of March 2011. The MLP was significantly more accurate than the SVM. The amount of 3-minute interval data on the identification period covered 175200 time points, which is too large for our SVM tool and we used only 5% of it using arbitrary sampling.

Our other attempts to compress the identification data have so far led to even more unsatisfactory results. We tried two schemes of self-organizing map (SOM) for the compression: (i) using the prototype vectors of SOM as model input data and (ii) sampling the input data based on the SOM prototypes. Adding the partly physically based model enables a potential solution. It models the fast phenomena and thus allows us to model the residual using less accurate time resolution, such as 15 minutes. Increasing the sampling interval of SVM model to 15 minutes in the hybrids led to inaccurate modelling of 3–minute interval residuals and thus also to inaccurate 3–minute interval forecasts.

With MLP the task was also computationally demanding but not to the same extent as with the SVM. With MLP we could use all the identification data and the forecasting performance was rather good.

The machine learning methods, especially the SVM, were alone poor in forecasting abrupt intra-hour power behaviour. We first tried to enhance the intra-hour performance of SVM by modelling hourly and intra-hourly behaviour separately. These experiments did not bring a substantial improvement.

The models with physically based structure cover only a small part of the total load. In this case, we did not build a physically based model for the total load, because according to our earlier experience [9], it requires much work compared with the achieved forecasting accuracy.

As the Table IV shows, the hybrids that include partly physically based sub-models for both of the different active control types performed best. The one–year long identification period was not long enough for the pure machine learning methods to learn the dynamic control responses and the load behaviour in exceptional situations.

5.2.2 Hybrids with the similar day forecaster

We integrated partly physically based active demand models to the similar day forecaster. Similar day forecasting modelled the residual of the active demand forecasts. Thus, the hybridisation approach was similar to the one applied above for machine learning. Table V summarises the forecasting accuracy in terms of RMSE in MW. In parentheses, the RMSE is normalised to the distribution area average power of the identification or verification period respectively.

Table V: Forecasting of 3-minute interval power of the distribution area, similar day forecaster hybrids

RMSE in MW and (NRMSE in %)	Identification	Verification
Similar day forecaster (SD)	2.3158 (2.98)	3.7466 (3.71)
SD and emergency control response model	2.2556 (2.90)	3.6427 (3.61)
SD, emergency control response and ToU response models (SD hybrid)	2.2355 (2.87)	3.5342 (3.50)
SD, emergency control response and ToU response models + MLP	1.5739 (2.02)	3.4189 (3.29)
SD, emergency control response and ToU response models + SVM	1.6899 (2.17)	3.5341 (3.42)
Simple ensemble (weighted mean of the MLP hybrid and the SD hybrid)	1.9838 (2.55)	3.0913 (3.05)

The hybrid methods are more accurate than the similar day forecaster. An advantage of the hybrid methods is missing from the Table V, because during the measurements, ToU timing was every day the same. The main advantage of the hybrid model is that it forecast accurately also, when the timing varies dynamically, for example based on the electricity market price. Similar day forecaster cannot do that alone. This benefit was clear in separate dynamic ToU control field tests in Helsinki explained in [7], for example.

The tuning of the methods is somewhat preliminary, because the focus in the work has been in the benefits of the hybrid approach. For example, the precautions to prevent overlearning and the optimisation of the SVM structure and control parameters need further development.

According to the results in Tables IV and V, the inclusion of the similar day forecaster in the hybrids improves forecasting performance. The hybrids that include all the developed sub–models, especially a simple ensemble combining the residual based MLP and SD hybrids, preformed best.

6 CONLUSION

A hybrid approach was proposed and studied for forecasting the power of the distribution grid area in the presence of active demand. The approach comprises a response model with a physically based structure and forecasting the residual with machine learning methods and a similar day forecaster. The response model models how the controllable thermal loads respond to the outside temperature and dynamic control signals. It includes the related time dynamics and saturation. For each controlled group, the emergency load control responses and the ToU control responses are forecast separately. The similar day forecaster forecasts the residual of the control response models. The machine learning methods forecast the remaining residual and its time dynamics.

The proposed hybrids were more accurate than their component methods alone. Both in the training data and in the foreseeable real application the identification data do not include so much emergency control events nor exceptional extreme weather situations that purely data driven methods alone learn the related essentially non-linear time dynamics adequately. The machine learning methods and the similar day forecaster have this deficiency. Hybrid modelling methods combined the strengths of different modelling and forecasting methods. The proposed hybrid approaches gave a significant performance improvement. They are also relatively easy to implement and maintain. Different hybrid approaches still need development, analysis, integration and comparison.

7 ACKNOWLEDGEMENT

The authors gratefully acknowledge those who provided funding and field test data. The research was funded by the Academy of Finland via the research project Response. The load control field tests where implemented by the electric power companies Helen and Loiste. The research was also based on models developed in earlier field tests by five other electric power companies. We received weather data both from the electric power companies and the Finnish Meteorological Institute.

BIBLIOGRAPHY

[1] A. Garulli, S. Paoletti, A. Vicino, ”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] P. Koponen, H. Niska, “Hybrid Model for Short-Term Forecasting of Loads and Load Control responses”, *IEEE PES ISGT Europe 2016*, 6 p.

[3] C. Deb, F. Zhang, J. Yang, S. E. Lee, K. W. Shah, “A review on time series forecasting techniques for building energy consumption”, *Renewable and Sustainable Energy Reviews* 74 (2017), pp. 902-924.

[4] B. Cs. Csáji, A. Kovács, J. Váncza,”Online Learning for Aggregating Forecasts in Renewable Energy Systems”, *ERCIM NEWS 107*, October 2016, pp. 40-41.

[5] B. Dong, Z. Li, S. M. M. Rahman, R. Vega, “A hybrid model approach for forecasting future residential electricity consumption”, *Energy and Buildings* 117 (2016) 341-351.

[6] O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L.J. Herrera, A. Guillen, L. Marquez and M. Pasadas, “Hybridization of intelligent techniques and ARIMA models for time series prediction”, *Fuzzy Sets and Systems*, vol. 159, Elsevier 2008, pp. 821-845.

[7] P. Koponen, H. Niska, R. Huusko, “Improving the performance of machine learning models by integrating partly physical control response models in short-term forecasting of aggregated power system loads”. Plenary paper, *ITISE 2017*, 12p.

[8] A. Selakov, S. Ilic, S. Vukmirovic, F. Kulic and A. Erdeljan, “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.

[9] P. Koponen, H. Niska, A. Mutanen, ”Assessment of some methods for short term load forecasting”, IEEE PES Innovative Smart Grid Technologies, Europe, ISGT Europe 2014, 6p.

[10] P. Koponen, *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

[11] S. Hochreiter, J. Schmidhuber, “Long short–term memory”, *Neural Comput.* 1997, 9, 1735–1780.

[12] Feng Zhao, Hongsheng Su, “Short-Term Load Forecasting Using Kalman filter end Elman Neural Network”, *2nd IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 1043-1047.

[13] Gang Lv, Zhiming Liu, Yu Fan, Guo-Guo Li, “Modeling a permanent-magnet linear synchronous motor using hybrid nonlinear autoregressive neural network”, *2008 9th International Conference on Signal Processing*, 2008, pp. 1685-1689.

[14] V. N. Vapnik, *The nature of statistical learning theory.* New York: Springer, 1995, 188 p.

[15] H. Niska, P. Koponen, A. Mutanen, “Evolving Smart Meter Data Driven Model for Short-Term Forecasting of Electric Loads”. *IEEE ISNNIP 2015*, Singapore, 7-9 April, 2015, 6p.

[16] Qingqing Mu, Yonggang Wu, Xiaoqiang Pan, Liangyi Huang and Xian Li, "Short-term load forecasting using improved similar days method," *2010 Asia-Pacific Power and Energy Engineering Conference*, Chengdu, 2010, pp. 1-4.

[17] M. Karimi, H. Karimi, M. Gholami, H. Khatibzadehazad and N. Moslemi “Priority index considering temperature and date proximity for selection of similar days in knowledge-based short term load forecasting method”, *Energy*, Vol. 144, Feb. 2018, pp. 928–940.