

Load Forecasting in District Heating Systems Using Stacked Ensembles of Machine Learning Algorithms

Till Faber
 Kempten University of Applied Sciences
 Kempten, Germany
 till.faber@hs-kempten.de

Matthias Finkenrath
 Kempten University of Applied Sciences
 Kempten, Germany
 matthias.finkenrath@hs-kempten.de

Abstract— For district heating, heat demand forecasting is playing a key role for an optimised power plant dispatch. Machine Learning can help to significantly improve forecasts of thermal loads. The prediction quality of neural networks is higher than that of decision trees in most cases. However, compared to decision trees neural networks have weaknesses when extrapolating outside known ranges. This work presents a novel method called "Deep DHC" (Deep Learning for District Heating and Cooling), which combines these two approaches in order to benefit from strengths of both methods. On the one hand, the novel approach uses conventional decision tree based regression algorithms such as the AdaBoost and Random Forest, as well as artificial neural networks. In addition to common feed forward neural networks (FNN), a deep learning network structure, which consists of long short-term memory (LSTM) cells, is used for the first time. The LSTM method has already proven to be very powerful in modern speech recognition. In order to achieve best possible heat demand forecasts, the aforementioned methods for load forecasting are combined and weighted by an additional machine learning method. Results show that it is possible to achieve a further improvement in forecasting quality for district heating loads by purposefully combining individual forecasting methods. Hence, mean and absolute deviations are significantly reduced in comparison to the individual methods.

Keywords—district heating and cooling, load forecast, machine learning, stacking

I. INTRODUCTION

With total lengths of often several hundreds of kilometres, the efficient supply of district heating networks is a considerable challenge. Typically, heating network operators use multiple power plants with different technical and economic characteristics. The increasing use of thermal energy storage within the heating networks offers further optimisation options for plant operation, for example in order to avoid the frequent use of fossil-fuelled peak-load boilers. However, the diversity of plants increases the complexity of plant dispatch planning. A key boundary condition for an optimised operation is to know as precisely as possible the time and the amount of the expected heat load in the network.

In recent years, a large number of different forecasting methods have been investigated. The focus of these investigations is on energy forecasting for buildings [1, 2, 3]. Common load forecasts for district heating networks are based on decision trees or simple neural networks [4, 5]. A decision tree is a hierarchical data structure, as shown in Fig. 1 using a

simple example of district heating load forecasting. In the example shown, the first step is to differentiate between weekdays and weekends. In the second step, the temperature range is narrowed down, which results in the expected load. In the present study, classical decision trees of the type "Ada-Boost" and "Random Forest" were used for load prediction [6, 7]. "Ada-Boost" or Adaptive Boosting refers to a popular training method, in which incorrectly classified data points are given more weight during the next training step, i.e. the repeated comparison between measurement and forecast results. In contrast, the "Random Forest" method randomly generates several uncorrelated decision trees, from which the final configuration is determined during the learning process.

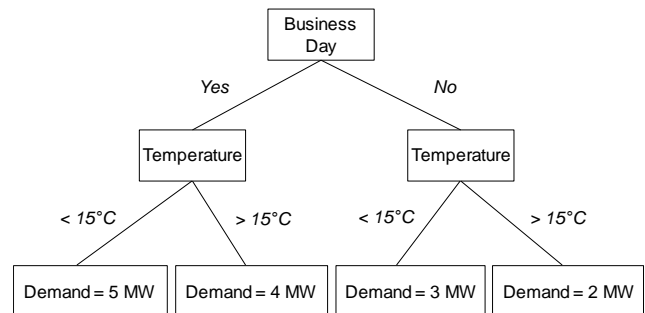


Fig. 1. Example of a heat demand forecast using a decision tree

In addition to decision trees, artificial neural networks have increasingly been used for load forecasting since the early nineties.

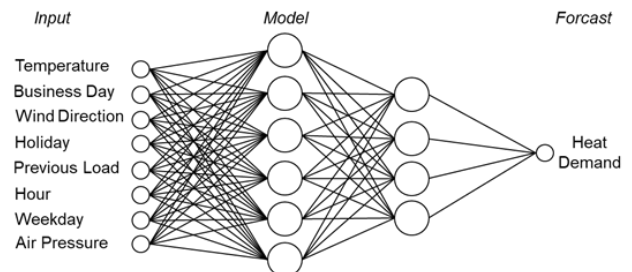


Fig. 2. Example of a heat demand forecast using an artificial neural network

Fig. 2 shows the schematic structure of a neural network for load forecasting of a district heating network. For this purpose, a neural network is trained with the help of measurement data, so that the trained model can determine the

dependent output variable (heat demand) within a certain accuracy using the input variables (e.g. temperature, working day, etc.). Neural networks can have various structures, which usually differ in how many artificial neurons are located on how many layers and how they are connected to each other. In the present work, feed-forward neural networks (FNN) are considered from the group of classical neural networks [6]. In this topology, a layer is always only connected to the next higher layer in the direction of the output variable.

II. MODERN LEARNING METHODS AND FURTHER DEVELOPMENT "DEEP DHC"

Due to rapid developments in information technology and significantly faster computers, very powerful machine learning methods have become available. Algorithms from the field of so-called "deep learning" - and there of the type of recurrent neural networks such as Long-Short-Term-Memory (LSTM) - already dominate commercial speech and pattern recognition [8, 9]. In contrast to classical neural networks, these procedures can store information states for a longer period and they can decide independently when they are to be retrieved again. Thus, events that are separated by a longer period can be linked together more easily. In its structure, this type of neural network is characterized not only by its long-term storage capacity, but also by the fact that connections from neurons of one layer to neurons of the same or a preceding layer are possible. This characteristic is very promising for load forecasting in heating networks, since the consideration of current as well as longer-term information - e.g. in the forecast over several days or over weekends - on heat loads or weather conditions is of great importance with respect to forecast quality.

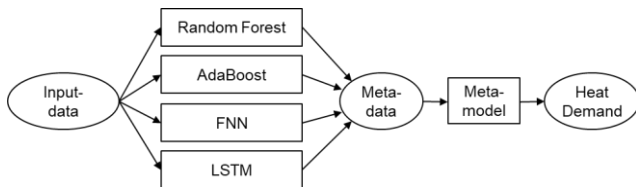


Fig. 3. Heat demand forecast using "Deep DHC" method

In most cases, the prediction quality of neural networks is higher than that of decision trees. In contrast to decision trees, however, neural networks show weaknesses in extrapolation outside of known value ranges. Therefore, the research project has developed a novel method that can combine the strengths of both approaches. This method named "Deep DHC" (Deep Learning for District Heating and Cooling) combines the two approaches presented in order to make use of the strengths of both methods [10]. The "Deep DHC" method combines conventional machine learning methods of the decision tree type with artificial neural networks. In addition to simple feed forward neural networks (FNN), a deep learning network structure of long-short term memory (LSTM) cells is used. The core of the "Deep DHC" procedure is a combination and weighting of the above-mentioned initial procedures for load prediction by means of a further machine learning procedure, in order to achieve the best possible load predictions. This procedure is illustrated in Fig. 3. The machine learning procedure, which determines the best possible combination and weighting of the previously executed load predictions of

the individual procedures through training, is called a "meta model".

III. CONSIDERED DISTRICT HEATING NETWORK AND DATABASE

The above-mentioned heat load forecasting methods were analysed based on a subnetwork of Fernwärme Ulm GmbH with a total length of 40 km, which covers the annual heat demand of 75 GWh or 1,100 households, on average. For this purpose, measurement data of the last 15 years were used in an hourly resolution, and hence a total of 131,400 data series. Each hourly data set consisted of 20 parameters, i.e. measurement data from the district heating network or weather information such as air temperature, wind direction or wind speed. The correlation between individual parameters and the associated district heating load was previously checked by means of a correlation analysis [11].

IV. TECHNICAL VALIDATION

The results of the load forecasts of the various methods over 72 hours in advance were compared with corresponding measurement data from the operation of the subnetwork of Fernwärme Ulm GmbH described in the beginning. The input data records from plant operation used for this purpose had not previously been used for training, or for the validation of the individual processes. Archived weather forecasts were used for the weather data, so that the load forecast boundary conditions during the test were identical to those of a real power plant operation. For each of the various load forecasts, the mean absolute percentage error (MAPE) was determined, which indicates the mean (absolute) relative distance of the true measured values from the respective load forecast. Together with the maximum forecast error, this is compared in Table 1 for an exemplary test period of 72 h for the various load forecast procedures.

TABLE I. Comparison of forecast and measured loads

	Ada-Boost	Random Forest	FNN	LSTM	Deep DHC
Maximum Error [%]	20.53	17.69	19.7	13.84	8.53
MAPE [%]	4.87	5.72	5.73	4.32	3.56

All procedures show a comparatively high accuracy in the selected test period, with errors of less than 6 percent. However, the "Deep DHC" method described above can increase the average accuracy by almost one percentage point (3.56%) compared to the best single run LSTM (4.32%). The maximum deviation of the load forecast, which fluctuates between 14 % (LSTM) and 20 % (Ada-Boost) for the individual processes, can also be reduced to below 10 % with the new process, which is particularly important for power plant dispatch optimisation. These data on maximum and percentage deviations for all compared methods as well as a direct comparison of the load forecast of the "Deep DHC" method with the actually measured load is shown in Fig. 4 for a period of three days as an example. The load peaks as well as the basic course can be predicted very well with the new method.

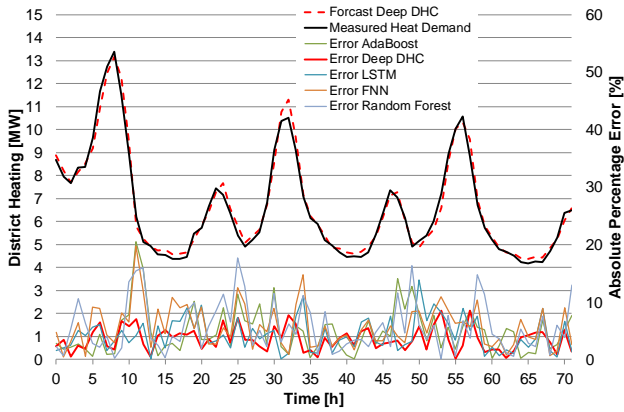


Fig. 4. Comparison of forecast and measured heat demand

V. ECONOMIC ANALYSIS

For operators of district heating networks it is crucial to know the expected district heating load as precisely as possible, as this enables a particularly efficient use of the available power plants, thermal storage facilities or even "power-to-heat" plants. Although it is intuitively obvious that the forecasting quality for the district heating load plays a key role for plant dispatch, an exact quantification of the economic benefit of a reduced load forecast error is a challenge, since the real plant dispatch of a certain period of time cannot simply be repeated based on different load forecasts. However, the economic impact of the load forecast quality can be assessed based on a numerical plant design optimisation tool, together with real district heating load data of past years of operation. A detailed plant dispatch planning tool based on mixed-integer linear numerical optimisation was developed in 2015 for the entire district heating network and all plants of Fernwärme Ulm GmbH. The tool was further improved within the framework of the KWK-Flex research project.

In order to assess the economic impact of the load forecast quality, the ideal plant dispatch of all generation or storage units is first determined for a reference year using the numerical plant dispatch optimisation tool. This reference case is determined under the boundary condition of a complete and exact knowledge of the expected district heating load three days in advance for the entire reference year, which reflects the best possible operating case. In order to estimate the influence of an increasingly inaccurate district heating load forecast, forecast errors are then randomly applied to the actually measured annual load profile at different levels, whereby the total amount of heat produced over the year under consideration is kept constant. Fig. 5 shows, in comparison to the measured load, correspondingly generated load profiles with forecast errors of 5 % and 15 % for a period of three days. Higher forecast errors cause higher expected load ramps and peaks in this evaluation. A more inaccurate forecast of the district heating load for the following three days hence results in a higher need for corrections in plant operation, which in practise is usually realized by an increased use of peak-load boilers. This changed plant dispatch due to a deteriorated forecast quality can be determined in each case using numerical plant dispatch optimisation.

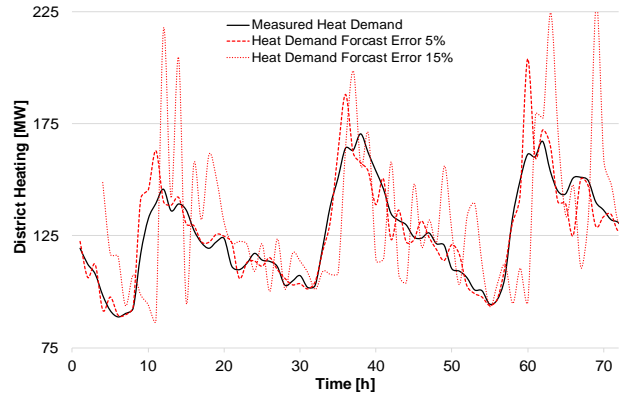


Fig. 5. Comparison of measured district heating load and heat load with forecast errors

Over a whole year, the results of the numerical plant dispatch optimisation show that with a forecast error of 5 %, heat generation by peak-load boilers deviates by only 1 % from the reference case with exact knowledge of the district heating load. The change in the peak load boiler input relative to the reference case is therefore small. This is shown in Fig. 6 for the period shown in Fig. 5.

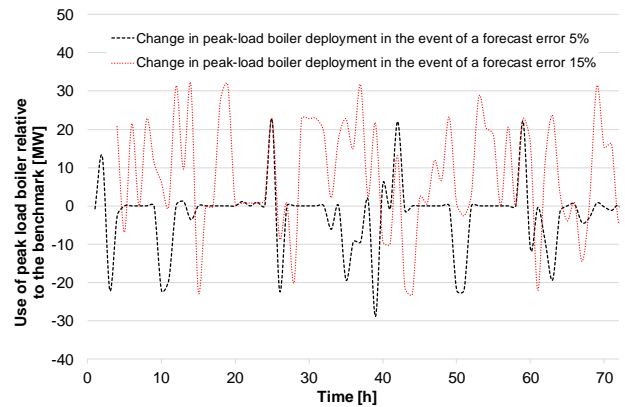


Fig. 6. Change in the peak load boiler insert compared to the reference case

A forecast error of 15 % over a whole year leads to a 22 % increase in heat generation in the peak-load boiler compared to measured operation. This can also be seen in Fig. 7 from the curve "Change in peak-load boiler input in the event of a 15% forecast error", which is exemplary for a period of three days. In principle, as shown in Fig. 7, numerical plant dispatch optimisation results in an increasing number of start-ups and full load operating hours of the peak-load boiler with increasing forecast error, and consequently an increasing heat generation in the peak-load boiler as a whole.

Increased use of generally fossil-fired peak-load boilers is associated with increased consumption of fossil fuels and higher wear and tear. As this increases emissions and operating costs, the analysis illustrates the considerable benefits of improved forecasting accuracy for district heating loads. A comparable benefit should also be expected for the operation of district cooling networks.

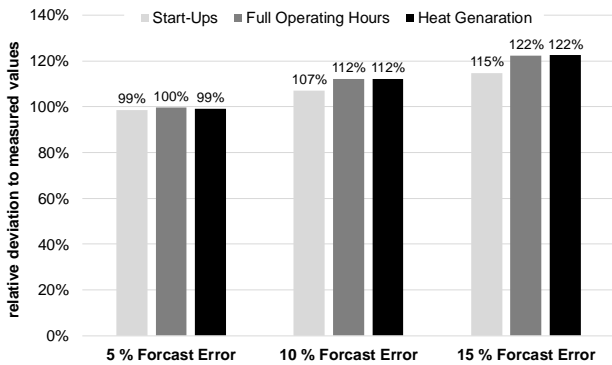


Fig. 7. Changes in peak-load boiler operation compared to the reference case

VI. CONCLUSION AND OUTLOOK

Conventional and novel methods for load prediction in district heating or cooling networks, which are based on decision trees and neural networks, were presented. These can be beneficially used for plant dispatch optimisation. The strengths of the various individual load forecasting methods can be further boosted by combining and weighting the individual forecasts. This is done using the newly developed load forecasting method "Deep DHC", which allows a significant improvement in the forecast quality to a few percent compared to the measured district heating load. Simulation calculations show that this results in a considerable optimisation potential for the planning and operation of power plants, thermal energy storage or power-to-heat plants that are used for district heating or cooling. In particular, the use of fossil-fired peak-load boilers could be significantly reduced by means of improved district heating load forecasting. It is planned to further develop the methods from the KWK-Flex research project in further research work.

ACKNOWLEDGMENT

This work was funded by the Federal Ministry of Economics and Energy of Germany under the funding code 0324111. The responsibility for the content of this publication lies with the authors.

REFERENCES

- [1] H. Xiang Zhao, F. Magoulés: A review on the prediction of building energy consumption, *Renew. Sustain. Energy Rev.*, pp. 3586-3592, 2012
- [2] N. Fumo: A review on the basics of building energy estimation, *Renew. Sustain. Energy Rev.*, pp. 53-60, 2014
- [3] Dagdougui H., Bagheri F., Le H., Dessaint L.: Neural network model for short-term and very short-term load forecasting in district buildings, *Energy and Buildings* Volume 203, 15 November 2018
- [4] Meisenbach C.; Gnüchtel S.: Lastprognose konventionell oder mit neuronalen Netzen - ein Vergleich. VDI Report Nr. 1508. 1999.
- [5] Kato, K.; Sakawa, M.; Ishimaru, K.; Ushiro, S.; Shibano, T.: Heat load prediction through recurrent neural network in district heating and cooling systems. *IEEE International Conference on Systems, Man and Cybernetics*. Singapore. October 2008.
- [6] Geron, A.: *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media. Peking, Boston, Farnham, Sebastopol, Tokyo. 1. Edition. 2017.

- [7] Antonopoulos I, Robu V., Couraud. B Kirli D., Norbu S., Kiprakis A., Flynn D., Elizondo-Gonzales S., Wattam S.: Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review, *Renew. Sustain. Energy Rev.*, Volume 130, September 2020
- [8] Sak, H.; Senior, A.; Beaufays, D.: *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modelling*. 15th Annual Conference of the International Speech Communication Association. Singapore. 2014.
- [9] Pironkov G., Wood S., Dupont S.: Hybrid-task learning for robust automatic speech recognition, *Computer Speech & Language*, Volume 64, November 2020
- [10] Faber T.; Groß J.; Finkenrath M.: Innovative Lastprognosen mit Deep Learning-Methoden, *EuroHeat & Power*. 47. Ed. Issue 1-2. 2018.
- [11] Faber T.; Brauer J.; Finkenrath M.; Mayer W.; Schott M: Optimization of a cross-linked system using neural networks. European patent application EP3.432.234.A1. 2017.