Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 116 (2017) 208–216

Energy Procedia

www.elsevier.com/locate/procedia

The 15th International Symposium on District Heating and Cooling

# Operational Demand Forecasting In District Heating Systems Using Ensembles Of Online Machine Learning Algorithms

Christian Johansson[a]*, Markus Bergkvist[a], Davy Geysen[b,c], Oscar De Somer[b,c], Niklas Lavesson[d], Dirk Vanhoudt[b,c]

[a]*NODA, Biblioteksgatan 4, 374 35 Karlshamn, Sweden*
[b]*EnergyVille, Thor Park 8310, 3600 Genk, Belgium*
[c]*VITO, Boeretang 200, 2400 Mol, Belgium*
[d]*Bleking Institute of Technology (BTH), 371 79 Karlskrona, Sweden*

**Abstract**

Heat demand forecasting is in one form or another an integrated part of most optimisation solutions for district heating and cooling (DHC). Since DHC systems are demand driven, the ability to forecast this behaviour becomes an important part of most overall energy efficiency efforts.
This paper presents the current status and results from extensive work in the development, implementation and operational service of online machine learning algorithms for demand forecasting. Recent results and experiences are compared to results predicted by previous work done by the authors. The prior work, based mainly on certain decision tree based regression algorithms, is expanded to include other forms of decision tree solutions as well as neural network based approaches. These algorithms are analysed both individually and combined in an ensemble solution. Furthermore, the paper also describes the practical implementation and commissioning of the system in two different operational settings where the data streams are analysed online in real-time.
It is shown that the results are in line with expectations based on prior work, and that the demand predictions have a robust behaviour within acceptable error margins. Applications of such predictions in relation to intelligent network controllers for district heating are explored and the initial results of such systems are discussed.


Peer-review under responsibility of the Scientific Committee of The 15th International Symposium on District Heating and Cooling.

* Corresponding author. Tel.: +46 735 30 95 02.
*E-mail address:* cj@noda.se

1876-6102 
Peer-review under responsibility of the Scientific Committee of The 15th International Symposium on District Heating and Cooling.
10.1016/j.egypro.2017.05.068

*Keywords:* district heating and cooling networks; heat load forecast; algorithms; machine learning

## 1. Introduction

Operational data analytics where traditional engineering and modern data science solutions are merged is a driving force behind the development of innovative 4th generation district heating networks as well as the upgrading of current 3rd generation systems [1]. A key aspect of most such energy efficiency schemes is the ability to predict future behaviour within the network. Since district heating, by basic design, is demand driven, it follows that demand predictions are vital to the success of such endeavours. Heat demand is also an important support tool for traditional optimisation schemes due to the substantial time delays in heat delivery throughout a district heating system. By using heat demand forecasting in such situations it is possible to increase the efficiency of heat generation in relation to the actual heat demand within the dispersed topology of the heating grid. Furthermore, the ability to control the demand is at the core of many operational optimisation solutions relating to smart grid technology such as demand side management and active load control, which further increases the value of heat demand forecasting [2].

### *1.1. Demand forecasting in district heating*

The heat demand in a district heating system generally originates from space heating and tap water heating. While space heating is primarily weather dependent, the tap water usage is related to social behaviour [3]. This combination leads to a nonlinear, stochastic and non-stationary characteristic of the system that increases the complexity of any sufficiently good solution [4]. Furthermore, the development of modern real-time supervision systems increase the availability of real-time data, which although providing a valuable resource for extensive data analytics also increases the complexity of the situation.

Demand forecasting in district heating systems is not a new subject. Throughout the years a number of forecasting approaches have been proposed, including statistical models as well as machine learning solutions such as neural networks [5]. The statistical approach is normally focused on trying to separate the weather dependent heating demand from the tap water usage based on social behaviour. Such solutions can range from rather simple solutions, featuring linear functions, to more complex solutions combining physical knowledge of the system with statistical modelling [6, 7]. Using the physical knowledge of the network as a basis together with statistical models for identifying system parameters is also underpinning other similar approaches [8], in which the Box-Jenkins methodology is applied to an autoregressive moving average model (ARMA). The use of statistical models for demand forecasting also includes the application of seasonal autoregressive integrated moving average models (SARIMA) in which the forecasting values are derived using Kalman filtering [9].

The other major branch of demand forecasting is based on more machine learning related approaches such as neural networks or support vector machines [10, 11, 12]. This should in theory increase the ability of the solution to handle nonlinear and non-stationary behaviour in the data. These ideas have since been further explored, e.g. by introducing recurrent neural networks to improve the ability to handle non-stationary heat demands [13].

There are many influencing factors making up the total heat demand in a district heating system and it is in practice impossible to make an exact model of this behaviour. This is a contributing factor to the use of both statistical models and machine learning approaches in which exact physical models are not required. Furthermore, in addition to being dependent on the quality of historical data, the forecasting is also conditioned on the quality of the influencing factors during operational usage. For example, many demand forecasting models make use of an outdoor temperature forecast, which in itself can be of varying quality.

### *1.2. Machine Learning*

Machine Learning is a methodology for finding and describing structural patterns in data [14]. Machine learning is a subfield of computer science and is usually regarded as a part of artificial intelligence research. The basic idea of machine learning is to construct algorithms, which in turn can generate models that can then make data-driven predictions or make decisions regarding classification of the data.

Normally a machine learning algorithm will use a set of training data to create a model. This model can then be used for subsequent predictions or classification tasks. A machine learning algorithm can consider static datasets as well as streaming data.

### *1.3. Ensemble learning*

Ensemble learning is basically about taking the advice of several shareholders instead of only one, thereby arriving at a better conclusion. In the context of this paper, a forecasting model is a shareholder in the endeavour of generating sufficiently correct heat demand forecasts. Instead of only using one algorithm with one training set generating one single model to predict the future heat demand, several training sets and/or algorithms can be used to generate a set of models. This set can be combined in an intelligent way to obtain a better forecast. First results of applying this method are available, however in future research we will do an in-depth analysis of the different methods existing to define an optimal ensemble.

## 2. PROJECT OUTLINE

### *2.1. Previous work*

Heat demand forecasting is an integrated part of our overall on-going endeavour to develop innovative smart grid technologies for district heating and cooling systems. The first part of this specific forecasting project was published at the 14th International Symposium on District Heating and Cooling in Stockholm, Sweden in 2014 [15]. In that paper we presented an online machine learning algorithm for heat demand forecasting. That the algorithm is online means that it automatically updates its model as new data becomes available. Due to the non-stationary nature of the heat demand such solutions represent an important step forward in domain specific data analytics.

The solution is based on a combination of decision tree machine learning algorithms and online functionality. The decision tree generation is based on ensemble bagging using the Fast Incremental Model Trees with Drift Detection (FIMT-DD) algorithm [16]. As the name implies FIMT-DD has the ability to detect concept drift in the data stream that helps it adapt to the non-stationary behaviour of an operational district heating system. The algorithm will grow sub-trees to replace those parts of the decision tree that becomes obsolete during the drift. Techniques for handling missing data and outliers were also added to the algorithm. All in all, this produces a solution that is efficient at processing streamed data while providing a robust forecasting ability.

As part of this previous work the performance of the algorithm was evaluated using operational data from the Karlshamn district heating system in the south of Sweden. Heat demand and outdoors temperature data was collected throughout a full heating season from residential buildings, commercial buildings and schools. This data was then analysed using the open source WEKA and MOA platforms [17, 18]. Two different approaches were studied in relation to the data. The first approach created one model for each building and then aggregated the predictions, while the second approach was to aggregate the building data and then create one single overall model.

The performance of the different approaches was evaluated using Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) [19]. The results showed that the second approach (aggregate data before building a model) was slightly superior to the first approach (create individual models and then aggregate results). The mean absolute percentage error showed an error of 5,1 % (first approach) and 4.8 % (second approach) over the studied period.

It was concluded that the algorithm possessed a strong predictive ability during the experiment. The algorithm is memory efficient since it does not require the storage of large sets of data and each instance of data can be discarded after the algorithm updates the model. Furthermore, the algorithm is able to learn incrementally while processing large amounts of data in real time.

### *2.2. Current work for this paper*

The primary purpose of this current study is to implement online machine learning forecasting algorithms in an industrial, fully operational real-time environment using actual weather forecasts as input for the system. Basically to run the system as it would have been used in an operational environment.

A selection of three algorithms was used to perform the tests, one of which is similar in construction to the algorithm presented in the previous work. Based on the results from the previous study only aggregated data was used to construct models (second approach). This also means that data such as meter data from production facilities can be used as input for the training. The first algorithm used is the Extra-Trees Regressor (ETR), which uses randomized decision trees in relation to sub-sets of the dataset in combination with averaging to increase predictive accuracy and to minimize over-fitting. The ETR algorithm is the one similar to algorithms used previous work. The second algorithm is Extreme Learning Machines (ELM) which is a feed-forward neural network used for regression analysis. An ELM uses a single layer of hidden nodes with randomized weights assigned to the input to the hidden layer. The third algorithm is an expansion of the ELM algorithm, in which a regularisation factor was added to prevent over-fitting of the training data. These algorithms were implemented according to the process described in the following section of the paper.

The main outcome of the work is to verify the accuracy of the forecasting schemes in an operational setting and to evaluate the influence of the added uncertainties due to the usage of actual weather forecasts.

## 3. EXPERIMENTAL SET-UP

### *3.1. The STORM project*

STORM is a European Union Horizon 2020 project aimed at developing an innovative district heating and cooling network controller for enhanced district energy efficiency. The project started in 2015 will continue until 2018. The previous work was done outside of STORM, but since then the work has been merged with the overall effort of STORM.

The theoretical work in STORM is based on self-learning algorithms for efficient control of components within a thermal system, which is in line with the machine learning based efforts presented in this current work. The STORM controller is based on three generic modules relating to algorithms for forecasting, planning and tracking the operational behaviour with the thermal grid. The work presented in the paper is obviously related to the first of these three modules.

STORM will be implemented in two demonstration sites in Sweden and The Netherlands. The Swedish site is located in the city of Rottne in the south of Sweden. This district heating system is operated by Växjö Energi and is a traditional 3rd generation system with two bio-fuel boilers, complemented with a peak load oil boiler. The IT-platform used in STORM is already operational in Rottne. This made it convenient to implement the forecasting algorithms described in this paper.

More information on the STORM project can be found in [1], a paper that will also be presented in this symposium.

### *3.2. Data management*

Since the district heating system in Rottne is equipped with the STORM IT-platform, it is possible to access the operational data in real-time. This was used during the experiment for this study, in which historical heat demand and weather forecast data was used to train the forecasting models. Weather forecasts were used to train the models since such forecasts would later be used to evaluate the system. In addition to this, weather forecast data was continuously accessed on an hourly basis to make the actual predications for the coming days.

The forecasts were made every day based on weather forecast data available at 2:00pm. This time of day is used since it relates to the setting time of most spot price markets for electricity, which is relevant in relation to combined heat and power generation. This then generated hourly head demand forecasts for the coming day. The reason to use this set-up was that we wanted to evaluate the system in relation to combined heat and power generation, in which the day-ahead spot price market is relevant to optimize against. Since this market sets during the early parts of the afternoon the day before, it was relevant for us to follow those context boundaries. Also, in general it provides a more accurate prediction scenario since weather forecasts tend to become better and better the closer in time they get to the forecasted value. So using a deadline at around a day ahead for the weather forecast provides a better estimation of the actual operational ability of the system.

### *3.3. System implementation*

The forecasting module is implemented on the STORM IT-platform, which in turn is based on the NODA Smart Heat Grid framework. This includes all required hardware and software, including communication infrastructure. The implementation of the operational module was a joint effort between VITO and NODA. Since they are located in different countries this required the sharing of data in a robust and secure manner.

Part of the NODA Smart Heat Grid framework is Linckii, which is a graphical user interface for managing data in graphs and other graphical schemes. For the forecasting module charts were created in Linkii by NODA. These charts were then accessed by VITO through HTTP GET requests on the chart data using the Python request package [20]. This data can be retrieved in either JavaScript Object Notation (JSON) or Comma-Separated Values (CSV) format. Both these formats are open standards for human-readable data exchange. This made it possible to exchange data between the NODA and VITO databases without affecting the overall function of the operational system. The forecasting module has been implemented with support for adding different database schemas through pluggable adapters set during configuration. Currently support for the NODA Linckii and VITO schemas are implemented. The database schemas supported are only limited by the dialects supported by SQLAlchemy, which at the time of writing includes Firebird, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, SQLite and Sybase.

The development code for the project was separated on two different code repositories to retain proprietary information and data. Throughout the project GitLab was used as repository management and issue tracker [21]. Each new feature was tracked as an issue on GitLab, and all discussions regarding that feature was done in that specific issue. All development was done on feature branches. Then when a feature was considered complete a merge request was sent to another system developer for peer review. After a complete peer review of the code the feature branch was merged with the master development branch. The two different repositories were synchronised manually. This work process facilitated smooth integration and deployment throughout the prototype development effort.

### *3.4. Method*

The forecasting module was evaluated during January, February and March of 2016 in the Rottne district heating system. The operational behaviour of each of the three algorithms was evaluated for each month separately as well as for the whole period. This provided us with a diverse set of evaluation scenarios since the weather changed from winter to spring during that period of time. For the purpose of evaluation the algorithms were analysed using the common metrics of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). MAE and MAPE were also used during the previous work in the forecasting project, which makes it convenient to compare results.

In the context of this work the MAE shows the absolute difference between actual and predicted heat demand expressed in kW. The MAPE, however, is a relative metric, which means that it relates to the scale of the values being analysed. For example, consider two values of 50 and 100, both of which have an error of 10. Using MAE they would have the same error (i.e. 10), while having very different errors using MAPE (i.e 10 is a larger percentage of 50 than of 100).

The forecasting module was part of the operational system, which basically means that is was run once a day at 2pm for the coming day using the weather forecast available at that point in time. This was repeated throughout the experimentation period.

## 4. RESULTS

Figure 1 shows the correlation between the outdoor temperature and the heat demand using forecasted as well as measured outdoor temperatures. During lower temperatures the weather forecast tends to consistently overestimate the outdoor temperature, which is why there are more low temperatures within the measured dataset.

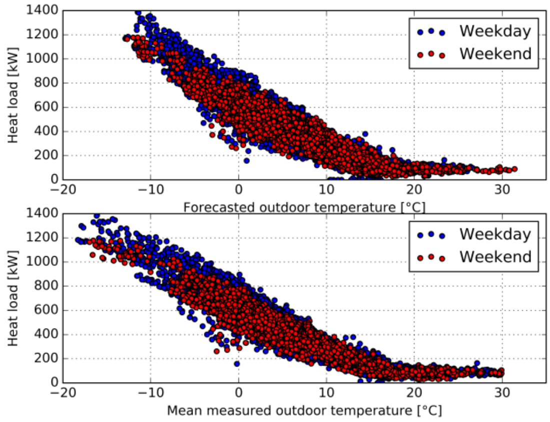


Figure 1: Correlation between outdoor temperature and heat demand

Results are shown in relation to the evaluation metrics described in the previous section. Table 1 shows the evaluation metrics for the Extra-Trees Regressor (ETR) and the Extreme Learning Machines (ELM), for the whole experiment period as well as for each month individually. The results of the extended ELM are almost identical to the normal ELM, which is why they are not shown in Table 1.

Table 1: Evaluation metrics

| | ETR MAE | ETR MAPE | ELM MAE | ELM MAPE |
|---|---|---|---|---|
| Jan-Mar | 73.24 | 11.7 | 91.72 | 14.31 |
| Jan | 89.84 | 10.2 | 145.25 | 17.62 |
| Feb | 48.05 | 7.59 | 43.13 | 6.84 |
| Mar | 76.28 | 16.56 | 75.26 | 16.86 |

Figure 2 shows a graph of the operational behaviour of the ELM algorithm during the period from January to March. There is a short period during late February with data missing due to communication error.

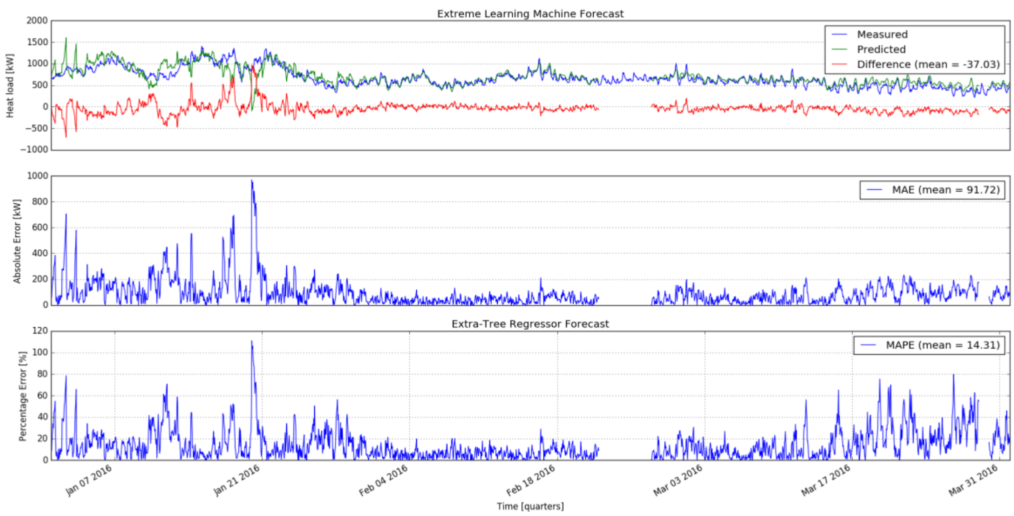


Figure 2: Operational behaviour of the ELM algorithm

## 5. DISCUSSION

As shown in Figure 1 a strong correlation (-0.934) exists between the measured outdoor temperature and the heat load. A slightly less strong correlation (-0.909) holds between the forecasted outside temperature and the heat load. It follows that the temperature forecast is the most important feature for a heat load forecast. Due to the lower energy demand during the weekend it is also useful to add a feature representing the day of the week. A last feature representing the quarter of the day is also added because of the heat load being time dependent.

Table 1 shows that the accuracy of both forecasters is best in February. There are several reasons to explain this behaviour. In January there was a period with outdoor temperatures outside of the temperature span used to train the models, which leads to less accurate results. This is especially true for ETR, which is not able to extrapolate predictions outside the span of the training data. Then in February the accuracy of the predictions is already significantly higher since the models are trained using the online data as this becomes available throughout the experiment period. There is also a decrease in accuracy during March due to systematic overestimation of the heat load in the second part of March. Reasons for this systematic overestimation cannot be found in the data as all algorithms lead to the same results in this time window. In all probability this is due to a behaviour change in the system, which is not represented in the training set. Due to the origin of the errors being lacking training data, the models would be expected to improve their accuracy once this data is added into the training set.

Considering the above discussion February is the most appropriate metric to use since it most accurately resembles a long-term operational setting. Thus it can be concluded that ELM is superior to ETR.

It should be noted that the MAPE metric in March is high due to the lower total heat demand. This is due to the percentage centric approach by this evaluation metric. Considering the MAE metric it is clear that the accuracy is still acceptable given the context. This can be seen in Figure 2, as the forecast is following the measured data even though the head demand decreases.

There are multiple possibilities available to enhance the performance of the forecasters discussed in this paper. Preliminary analysis shows that adding a feature representing historical heat load information, more specific of the previous day, is important to improve the accuracy. Next to this a feature representing the day of the year can be added to capture seasonal behaviour.

Furthermore, the ensemble concept can be improved by an optimal combination of the forecasting algorithms, for example by using intelligent weighting in more advanced machine learning methods.

## 6. CONCLUSION

It has shown that the forecasting algorithms work well during operational scenarios when relevant training data was available. The forecasting ability deteriorated when the models were confronted with scenarios not covered by their training data, which is especially apparent for ETR. However, it was also shown that through continuous re-training as new data becomes available, they were able to adapt which can be seen in the increased of accuracy in the transition between January and February. This shows that the algorithms become better over time and that they have the ability to improve using new datasets.

Several potential improvements have been identified during the practical experimentation phase, which can be implemented to increase the predictive ability of the system. The most significant of these is to add short term (24h) historical head demand data to the model. Also adding the day of year for seasonable purposes improves the accuracy. Adding the control signal corresponding to the historical heat demand only introduces slight improvements.

Artificial neural networks such as ELM provide the best forecasting ability of the studied algorithms, and they are able to handle data outside the training dataset.

## Acknowledgements

In this project EnergyVille/VITO, NODA and Växjö Energi are working within the context of the STORM project which is funded by the European Union's Horizon 2020 Programme under Grant Agreement no. 649743. For Blekinge Institute of Technology this work is part of the research project "Scalable resource-efficient systems for big data analytics" funded by the Knowledge Foundation (grant: 20140032) in Sweden.

## References

[1] D. Vanhoudt, B. Claessens, J. Desmedt and C. Johansson, "Status of the H2020 STORM", in proceedings of the 15th Symposium on District Heating and Cooling, Seoul, Korea, September 4-7 2016

[2] C. Johansson, On Intelligent District Heating, Doctoral Thesis, Blekinge Institute of Technology, 2014

[3] S. Fredriksen and S. Werner, District Heating and Cooling, Studentlitteratur AB, Lund, 2013

[4] P. Bacher, H. Madsen, H. Aalborg and B. Perers, "Short-term Heat Load Forecasting for Single Family Houses", in Energy and Buildings, 2012, vol. 65, pp. 101-112

[5]P. Giridhar (ed), Energy Management Systems, InTech, 2011

[6] E. Dotzauer, "Simple Model for prediction of Loads in District Heating Systems", in Applied Energy, 2002, vol 73, pp. 277-284

[7]H.A. Nielsen and H. Madsen, "Modelling the heat consumption in district heating systems using a grey-box approach" In Energy and Buildings, 2006, vol. 38, pp. 63-71

[8]B. Chramcov, "Heat Demand Forecasting for Concrete Dsitrict Heating Systems", in International Journal of Mathematical Models and Models in Applied Sciences, 2010, vol. 4, pp. 231-239

[9] S. Grosswindhager, A. Voigt and M. Kozek, "Online Short-Term Forecast of System Heat Load in District Heating Networks". In proceedings of the 31st International Symposium on Forecasting, Prague, Czech Republic, 2011

[10] M. Grzenda and B. Macukow, "Demand Prediction with Multi-Stage Neural Processing" in Advances in Natural Computation and Data Mining, 2006, pp. 131-141

[11] K. Kato, M. Sakawa, K. Ishimaru, S. Ushiro and T. Shibano, "Heat Load Prediction Through Recurrent Neural Network in District Heating and Cooling Systems" in Systems, Man and Cybernetics, 2008, pp. 1401-1406

[12] C.P. Tae, S.K. Ui, K. Lae-Hyun, W.J. Byung and K.Y. Yeong, "Heat Consumption Forecasting Using Partial Least Squares, Artificial Neural Netowrk and Support Vector Regression Techniques in District Heating Systems" in Korean Journal of Chemical Engineering, 2010, vol 27, pp. 1063-1071

[13]M. Sakawa, H. Katagiri, T. Matsui, K. Ishimaru and S. Ushiro, "Heat Load Prediction in District heating and Cooling Systems Through a Recurrent Neural Network with Data Characteristics", in Scientiae Mathematicae Japonicae Online, 2010, pp. 449-464

[14] I.H. Witten, E. Frank and M.A. Hall, Data Mining – Practical Machine Learning Tools and Techniques 3rd ed. Elsevier Inc. 2010
[15] S. Provatas, N. Lavesson and C. Johansson, "An Online Machine Learning Algorithm for Heat Load Forecasting in District Heating Systems", in proceedings of The 14th International Symposium on District Heating and Cooling, Stockholm, Sweden, 2014
[16] E. Ikonomovska, J. Gama and S. Dzeroski, "Learning Model Trees from Evolving Data Streams" in Data Mining and Knowledge Discovery, 2011, vol. 23, pp. 128-168
[17] R.B. Remco, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald and D. Scuse, "WEKA Manual for Version 3-6-10", 2013
[18] A. Bifet and R. Kirkby, "Data Stream Mining – A Practical Approach", 2009
[19] U. Graham and I. Cook, A Dictionary of Statistics, 3rd ed. Oxford University Press, 2014
[20] J. Danjou, The Hacker's Guide to Python – 3rd ed. Julien Danjou (Standard Copyright License), 2016
[21] J.M. Hethey, GitLab Repository Management, Packt Publishing Limited, 2013