

HEAT LOAD NUMERICAL PREDICTION FOR DISTRICT HEATING SYSTEM OPERATIONAL CONTROL

D. Rusovs^{1*}, L. Jakovleva¹, V. Zentins¹, K. Baltputnis²

¹Riga Technical University,
Department of Heat and Power Engineering Systems
6B Kipsalas Str., Riga, LV-1048, LATVIA

²Riga Technical University, Institute of Power Engineering
12-1 Azenes Str., Riga, LV-1048, LATVIA

*e-mail: dmitrijs.rusovs@rtu.lv

To develop an advanced control of thermal energy supply for domestic heating, a number of new challenges need to be solved, such as the emerging need to plan operation in accordance with an energy market-based environment. However, to move towards this goal, it is necessary to develop forecasting tools for short- and long-term planning, taking into account data about the operation of existing heating systems. The paper considers the real operational parameters of five different heating networks in Latvia over a period of five years. The application of regression analysis for heating load dependency on ambient temperature results in the formulation of normalized slope for the regression curves of the studied systems. The value of this parameter, the normalized slope, allows describing the performance of particular heating systems. Moreover, a heat load forecasting approach is presented by an application of multiple regression methods. This short-term (day-ahead) forecasting tool is tested on data from a relatively small district heating system with an average load of 20 MW at ambient temperature of 0 °C. The deviations of the actual heat load demand from the one forecasted with various training data set sizes and polynomial orders are evaluated for two testing periods in January of 2018. Forecast accuracy is assessed by two parameters – mean absolute percentage error and normalized mean bias error.

Keywords: *Forecasting, heating demand, heating curve slope, multiple regression.*

1. INTRODUCTION

In line with the recently proposed European Climate Law, energy consumption and greenhouse gas (GHG) emissions should be dramatically reduced in order to achieve the European Union's long-term climate neutrality objective [1]. Energy production for district heating system (DHS) operation makes up to 14 % of the total energy balance in Latvia [2].

Recent energy market liberalization for heat and electrical energy supply creates additional challenges for DHS, especially in case of integrated operation with combined heat and power plants (CHPPs). The requirement for competition between

different energy producers in heat and electrical energy markets calls for the creation and utilization of accurate forecasting tools for various important parameters (e.g., electricity market price, heating demand etc.) necessary for operational planning and the optimization of bidding in the energy markets.

In Latvia, the introduction of Climate Finance became an important support for significant improvements in DHS efficiency. During the last few years, at least three heat energy storage sites integrated with CHPPs were constructed in Latvia, summarised in Table 1.

Table 1. The Largest Heat Storage Sites in Latvia

Heat storage site	Tank volume, m ³	Heat storage capacity, MWh	Max. charge/discharge capacity, MW/h
Riga CHPP-2, Latvenergo [3]	18000	550	150
Salaspils Siltums [4]	8000	418	15/25
Jelgava, Fortum Latvia [5]	5000	170	15

The application of heat storage tanks and utilization of the thermal inertia of consumers (i.e., buildings) can cover imbalance between heat demand in DHS and the request for electrical energy produced by CHPPs from the perspective of the electricity wholesale market [6]–[8]. DHS renovation together with the already ongoing pre-insulated network pipe exchange and building thermal insulation improvements will enhance the overall operational efficiency. For instance, the heat losses from DHS network can be reduced by 40 % if old pipes are replaced by new pre-insulated pipes as estimated for a particular case in Tallinn DHS [9]. In another study [10], ranking of the DHS based on efficiency and relative losses from networks is offered. For the overall technical evaluation of network performance, a factor is suggested, which

depends on the heat transfer coefficient of the pipes. Another way to decrease losses in DHS networks is forward temperature reduction in the supply lines [11]. Moreover, the heat load can be modelled by controlling the forward temperature in the network and the indoor temperature in housing [12]. However, ultimately, all these steps will ensure the best performance only if the smart operation of DHS is also supported by accurate heat load prediction.

Furthermore, the simulation of DHS performance requires the application and modelling of heat demand profiles, forward and return temperature in the supply line, as well as the estimation of pressure losses in the network. Therefore, forecasting of heat load is an important part of the overall DHS planning process [13]. The planning of the electricity generation schedule of CHPPs

for participation in the electricity wholesale market requires day-ahead hourly forecast of the heat load for DHS to be supplied by the particular CHPPs [14]. Overall, there are several approaches to heat demand forecasting in the literature. For instance, the forecasting of heat demand for the case of Riga by an application of artificial neural networks, linear regression and a combined method ensured the daily mean average percentage error in the range of 4.76–5.83 % for the particular data set studied [14].

While currently, in some DHS networks, the prediction of load is carried out for daily volumes, the need to adapt a more market-oriented operational and planning model in line with the highly volatile electrical energy prices in trading platforms (such as Nord Pool) requires shifting the forecast resolution to at least hourly basis. Therefore, our research is devoted to day-ahead forecasting of hourly heating load. The overarching goal thereby is to devise and successfully validate a short-term forecasting approach able to provide satisfactory accuracy on any DHS, at the same time maintaining simplicity of implementation, so that such an approach could be applied by any CHPP operator aiming to improve their heat and power production planning co-optimization.

Consequently, the heating load forecasting algorithm initially described in [14]

and more thoroughly elaborated in [15] is used in this study with the aim to validate its accuracy in DHS load forecasting in a particular system during different periods of the heating season and with varied model parameters. Up until now, this forecasting approach has been tested on only one DHS, which, furthermore, could be characterised as a generally very large system. Thereby, this paper envisions to add to prior research by validating the short-term forecasting approach on a significantly different DHS, e.g., notably smaller in terms of average consumption.

Moreover, the obtained forecast deviations from the actual consumption values are to be analysed by the implementation of the mean average percentage error (MAPE) and normalized mean bias error (NBIAS). When applied together, these metrics allow drawing more complete conclusions on the performance of a forecasting model.

The other major goal of this study is to introduce and apply a novel parameter for DHS performance characterisation – heating load curve normalized slope, which, to the best of the authors’ knowledge, has not previously been considered in the literature and, consequently, its definition offers additional novelty. The extensive datasets collected for this research allow for the calculation of this parameter for a number of different DHSs.

2. METHODOLOGY

2.1. Simple Linear Regression and Normalized Slope

For this study, the data about hourly DHS parameters (hourly heat load in MW and mean hourly ambient temperature in °C) were collected for five distinctly different DHSs from 2015 to 2019 in various parts of Latvia.

The studied DHS networks supply heat energy for two main purposes: space heating and domestic hot water supply. Heat consumption for air heating in ventilation systems is insignificant in the scope of the collected data. The value of heat load depends

on many factors, such as the season of the year, ambient temperature, desired inside temperature, air humidity, solar radiation, wind speed and direction etc. The size of the heating area and building insulation parameters as well as the number of inhabitants and their activities also affect the resulting heat demand. Therefore, heat load modelling is a very complicated process and certain simplifications need to be made. Most importantly, these simplifications are based on the available data.

A novel parameter characterising DHS performance is proposed – normalized slope. To obtain it, it is first necessary to perform a linear regression analysis on the data sets of various DHSs, using heat load as a dependent variable and mean hourly ambient temperature as an independent variable. From the regression analysis, a linear equation is obtained in the form of

$$Q = aT + Q_0, \quad (1)$$

where

Q – the dependent variable (heat load);
 T – the independent variable (ambient temperature);
 a – the slope;
 Q_0 – the intercept term (equal to the average heat load at ambient temperature of 0 °C).

The coefficients a and Q_0 from Eq. (1) are found for each DHS by minimising the sum of least-squares, which is a mathematical approach to calculate the dispersion of experimental data. The goal of the analysis is to get the smallest possible sum of squares and draw a line according to the linear regression equation, which best matches the collected DHS performance data.

Goodness of fit of the obtained load curve can be assessed by evaluating the coefficient of determination (R^2) for each studied system. It shows the deviations of the measured DHS load from the one calcu-

lated by the regression line. The value of R^2 is the sum of the squared deviations of the DHS load from the average value. It is common to consider that the obtained regression models are strong if R^2 is close to 1.

After the regression analysis, the DHS load curve normalized slope can be calculated as the ratio of the temperature multiplier (slope) in the linear regression equation (1) vs the average load (intercept) from the same equation (i.e., a/Q_0). This coefficient shows the DHS load change relative to the average load when the temperature changes by one degree.

To perform this analysis, only the heating season is considered (i.e., when DHS heating load consists of both space heating and hot water supply). Thereby data about the heating load of various DHS networks during December–March period of the year are used. We eliminated from consideration the data for October–November and April, when the heating load is not stable.

For clarity, it should be noted that a simple linear regression model described here, the results of which are summarised in Section 3.2, is not used as a forecasting approach, but rather as a tool to obtain DHS characteristics necessary to calculate the heating load curve normalized slope (i.e., a DHS performance assessment approach). Though, in principle, it can be used as a forecasting method; it would generally be applicable as such more likely to long-term planning (e.g., preparing for the next heating season), as opposed to short-term forecasting and planning. The key difference in short-term planning, which is necessary for effective participation in established and emerging energy markets, is the need to more dynamically adapt to changes in heating consumption patterns occurring during the season. Such a forecasting approach specifically suitable for short-term operational planning is offered in the next section.

2.2. Heat Load Short-Term Forecasting Approach

There are various data-driven approaches for generating linear and multilinear regression models with different influencing parameters for heat load prediction. For instance, in [16], the forward selection method is used to identify the most important influencing parameters for the forecasting of daily heat energy consumption. There it is found that for a particular DHS studied, high coefficient $R^2 = 0.9369$ can be achieved when only daily mean ambient temperature is taken into account. In comparison, this coefficient grows to 0.9663 when input variables also include the heating consumption of the previous day, maximum daily temperature, wind data, humidity and month of the year [16]. Nevertheless, the selection of influencing parameters also heavily depends on data availability.

For this case study, the heating demand short-term (day-ahead) forecasting methodology from [15] has been adapted for the use on one of the DHS described in this paper. The forecasting model is based on multiple linear regression with the forecast of the day-ahead hourly ambient temperature being the main predictor. However, for the purposes of this study, to not include temperature forecast inaccuracy in the heating demand forecasting and only isolate the effects of the model itself, the actual historical temperature data will be used as input. The additional effects temperature forecast inaccuracy can cause in heat demand prediction have previously been discussed in [17].

Overall, the forecasting algorithm used in this study can be summarised as follows:

1. Select forecasting model parameters – the look-back horizon (training data set) in days b and the order of the polynomial to be constructed with multiple linear regression k .

2. Select the number of days the forecast has to be performed for, h .
3. Start with the day number $d = 1$.
4. Read the hourly temperature and heat load data from previous days ($d-b$ to $d-1$).
5. Perform multiple linear regression on the dataset (obtained in step 4) in the form of

$$Q_i = a_0 + \sum_{n=1}^k a_n \cdot T_i^n + \varepsilon_i, \quad (2)$$

where Q_i – the dependent variable at point i ; T_i – independent variable at point i ; n – power of each term; k – power of the last term (i.e., order of the polynomial); ε_i – the error term at point i ; a_0 – the intercept term; a_n – the coefficient for the corresponding function of the independent variable.

Using the least-squares method, this allows identifying the model parameters – a_n coefficients.

6. The error terms for each point i are then averaged over the training dataset based on the hour of the day, thereby obtaining a 24-hour profile of the training model residuals.
7. Read the hourly temperature for the next day $d+1$.
8. Input the day-ahead temperature in the model obtained by the linear regression and output the initial 24-hour day-ahead heating load forecast.
9. Subtract from the initial forecast the 24-hour error profile and output the final heating demand forecast for day $d+1$. Save the result.
10. Increment day number d by one and repeat steps 4 to 10 until the day number d exceeds the preselected number of days for the forecasting experiment h .

Afterwards, the results can be assessed. The forecasting model performance is evaluated using the mean absolute percentage error and normalized mean bias error.

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{Q_i - \hat{Q}_i}{Q_i} \right| \cdot 100\%, \quad (3)$$

$$NBIAS = \frac{1}{m} \sum_{i=1}^m \frac{(Q_i - \hat{Q}_i)}{Q_{max} - Q_{min}} \cdot 100\%, \quad (4)$$

where

Q_i – the actual heat load at point i ; \hat{Q}_i – forecasted heat load at point i ; m – the total number of points in the forecast; Q_{max} –

maximal value of heat load in the actual observation series; Q_{min} – minimal value of heat load in the actual observation series.

These two error measures have distinctly different applications and, consequently, they can be used in parallel, as each of them describes the accuracy of the tested forecasting approach from a different angle. MAPE shows the average error disregarding the bias of it. This allows drawing conclusions on the overall accuracy of the model. On the other hand, NBIAS specifically focusses on the sign of the errors, allowing for the identification of systemic inaccuracies, i.e., a tendency of the model to over- or underestimate.

3. RESULTS

3.1. Example of the Case Study Data

Figure 1 shows an example of the duration curve for hourly ambient temperature and the corresponding DHS heat load for a relatively small DHS with the average heating power of about 8 MW during 2016. Figure 2 demonstrates for the same system the seasonal profile of average temperature and DHS heat load during 2016 and 2017.

The yearly profile can be classified into two parts – the heating season, when the ambient temperature is lower than 8 °C, and the off-season. The duration of the heating season in Latvia is usually about 200 days or 4800 hours per year. In the second part of the yearly profile (the off-season), the DHS heating load is mostly for hot water supply.

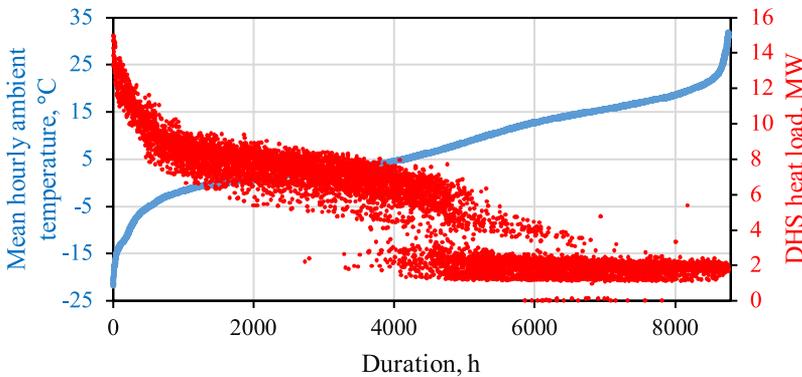


Fig. 1. Ambient temperature duration curve and the corresponding heat load in a small DHS in 2016.

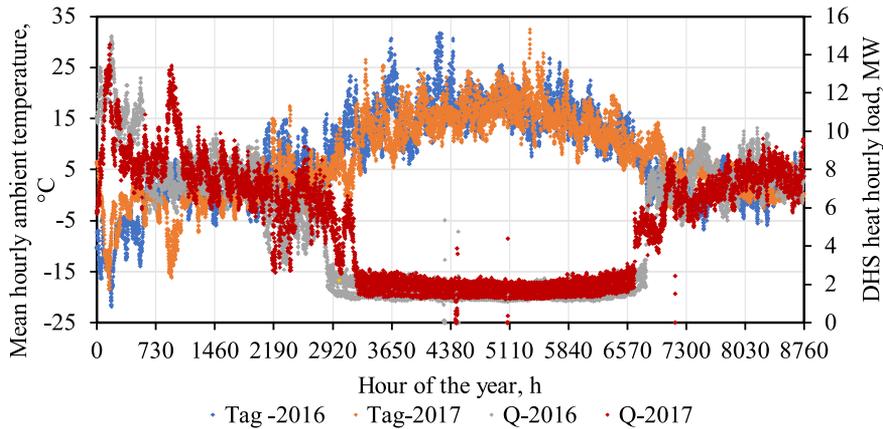


Fig. 2. Ambient temperature and DHS heat load annual profile during 2016 and 2017.

It can already be seen from Figs. 1 and 2 that there is a profound impact of the ambient temperature on the heating load in the DHS during the heating season. Since the average temperature during a heating season is about 0 °C, we introduce the definition of rated power (RP) of DHS as the average heating load (MW) when the ambi-

ent temperature is equal to the average. The RP values of the various DHS networks studied in this paper are summarised in Table 2, whereby data regarding the regression analysis of heating load dependency on the ambient temperature on an hourly basis are also shown, explained in more detail in the next section.

3.2. Simple Linear Regression Models of DHS Heat Load

Table 2 presents the analysis of a considerable volume of statistical data of hourly heat load, which allows performing linear regression for DHS heat load curves. Data were extracted from a five-year period (2015–2019) for five different DHSs with rated power (i.e., average consumption) ranging from 8 to 453 MW. In Table 2 and in the following figures, the RP value

is used to distinguish and identify various DHS networks considered, since that is a parameter which varies greatly among the studied systems.

The regression analyses presented in Figs. 3–7 demonstrate how the DHS load variable changes when the independent variable represented by ambient temperature varies.

Table 2. Data of DHS Case Studies for Different RP and Time Periods

DHS RP/ figure number	Heating season	Average load Q_0 , MWh/h	Minimum/ maximum load, MWh	Linear regression equation ($Q = aT + Q_0$)	R^2	Normalized slope, a/Q_0
RP8/ Fig. 3	2016–2018	8.04	3.8...15	$Q = -0.3229T + 8.0426$	0.8600	-0.040
RP10/ Fig. 4	2018	9.66	3.8...18.2	$Q = -0.4199T + 9.664$	0.8894	-0.044
RP20/ Fig. 5	2017–2018	18.56	14...38	$Q = -0.7972T + 18.563$	0.8770	-0.043
RP40/ Fig. 6	2017	42.41	22...80	$Q = -1.8123T + 42.407$	0.8579	-0.043
RP40	2018	40.94	12...75	$Q = -1.8357T + 40.937$	0.8977	-0.045
RP40	2019	41.95	22...62	$Q = -1.6395T + 41.946$	0.8039	-0.039
RP460/ Fig. 7	2015	457.48	210...1100	$Q = -22.877T + 457.48$	0.9400	-0.050

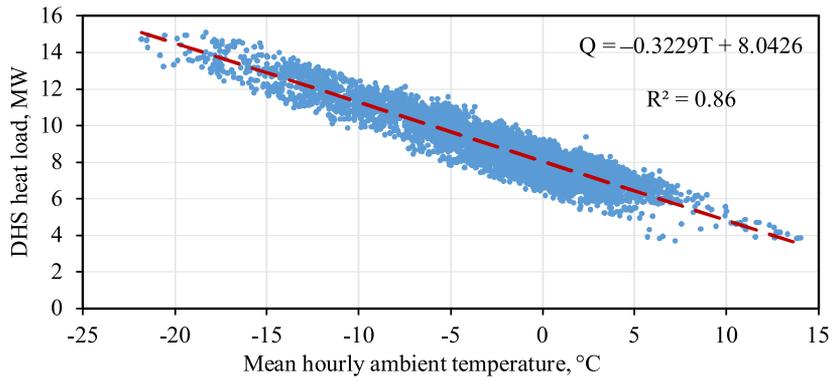


Fig. 3. DHS (RP=8 MW) heat load dependency on ambient temperature (2016–2018).

Fig. 3 demonstrates the performance of a DHS with RP of 8 MW during three heating seasons of 2016–2018. The obtained

linear regression equation represented in Table 2 describes the average performance during these three seasons.

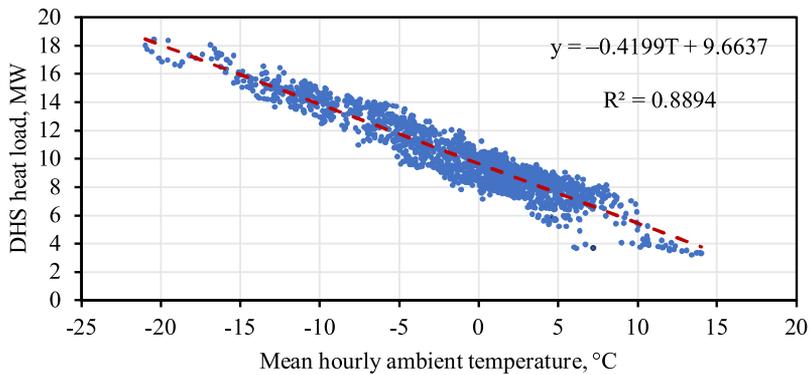


Fig. 4. DHS (RP=10 MW) dependency on ambient temperature in 2016.

In contrast, Fig. 4 represents data for a DHS with RP of 10 MW during only part of

the 2016 season (January–March).

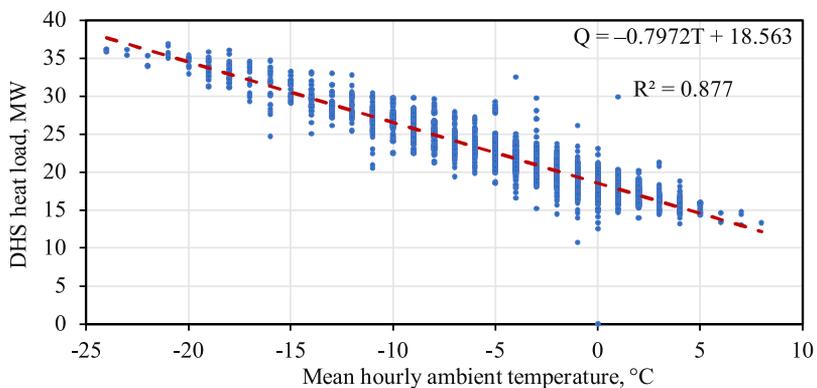


Fig. 5. DHS (RP=20 MW) dependency on ambient temperature (2017–2018).

For DHS with RP of 20 MW presented in Fig. 5, the data extracted during December of 2017 and January–March of 2018 were used. Note that for this DHS, the

ambient temperature data were only available in integer values, hence the peculiarity of Fig. 5 compared to the other figures.

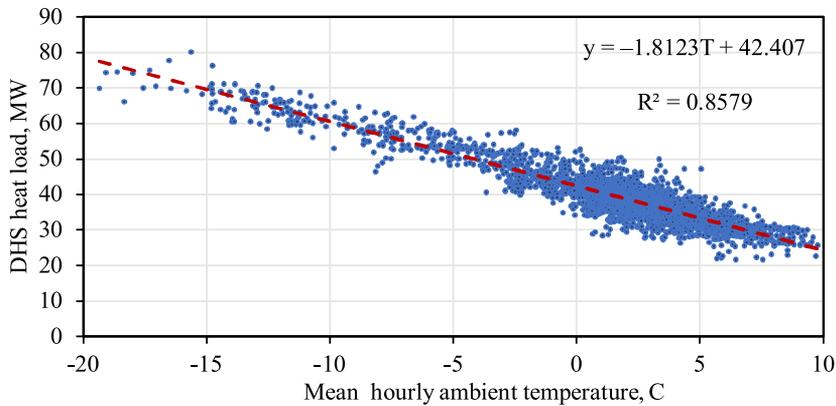


Fig. 6. DHS (RP=40 MW) dependency on ambient temperature in 2017.

Every single heating season had its own set of parameters (like ambient temperature profile) and therefore linear regression equation could be different for the same system. To exemplify this situation, data of

a DHS with RP of 40 MW were compared during three distinct heating seasons (2017, 2018, 2019) and the results were consequently presented in Table 2 and Fig. 6.

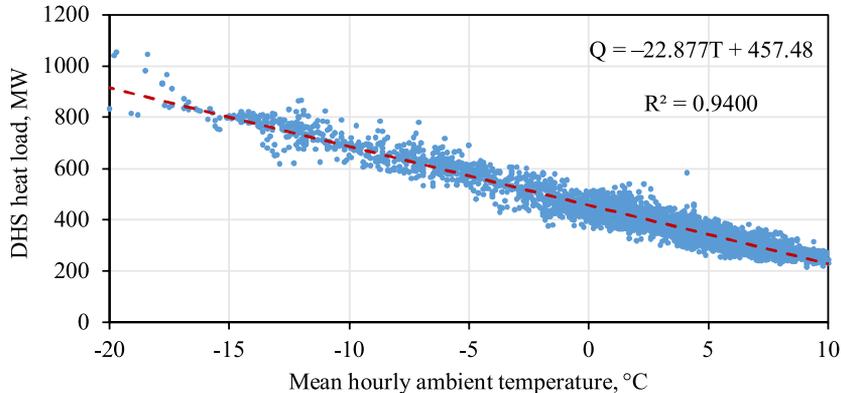


Fig. 7. DHS (RP=457 MW) dependency on ambient temperature during 2015 and 2016.

Finally, the last row in Table 2 describes the performance of a large DHS with RP of 460 MW during the heating season of 2015–2016. These data are visualised in Fig. 7.

As explained before, the simple linear regression equations presented in Table 2 represent the dependency of the recorded heating demand on the ambient temperature. Additionally, Table 2 contains the

coefficients of determination (R^2) for each case study. For example, in the case of RP460 in Fig. 7, the value of R^2 is 0.94, which is sufficiently close to 1. However, in the other case studies the coefficient is noticeably smaller. Nevertheless, overall, the results presented do confirm a general dependency of the heating load on the ambient temperature in the studied systems and,

consequently, this is a promising direction in the further development of regression-based forecasting tools applicable to these particular DHSs.

Interestingly, linear regression of heating demand versus ambient temperature generally shows good correlation if studied from the whole system perspective, as in this paper; however, when performing the analysis on a smaller scale (i.e., residential building level) the correlation is noticeably weaker. For instance, the authors of [18] found that for a particular large residential building in Latvia the R^2 of the daily heat consumption versus the ambient temperature was only 0.5459. This signifies that while for an individual consumer linear regression against solely the ambient temperature is not well suited for heating demand estimation, on a larger scale, the individual consumer deviations seem to some extent cancel out, and the regression results for the whole system are markedly better.

Furthermore, in other studies, the ambient temperature has been proven to be a major impacting variable also on other important DHS characteristics. For instance, the second-degree polynomial equations developed for DHS return temperature in [19] had determination coefficient of 0.9 for function of ambient temperature.

Nevertheless, the fact that for most of the case studies presented in this paper the coefficient of determination for linear regression is below 0.9, and therefore shows a bigger gap between the obtained equation and the collected data, needs to be addressed. The R^2 value for four of the studied DHS networks is from 0.9 to 0.8 or, in other words, 10 to 20 % of the calculated DHS load data cannot be predicted by the selected independent variable – ambient temperature.

The observed deviations of real DHS

heat load from the values obtained by the linear regression equations could be explained by the impact of factors other than ambient temperature. For example, in some of the studied DHS networks, the heating energy for hot water consumption can be equal to about 10–20% of the RP and this consumption, evidently, does not depend on ambient temperature, as can be seen in the off-season part in Fig. 2. If the presented linear equations were used for DHS load operational forecasting (e.g., during the following heating seasons), the obtained results could provide insufficient accuracy because only one parameter (ambient temperature) was taken as argument in the linear model and, more so, because of the fact that various DHS characteristics could change both during and in-between heating seasons. This latter argument is well supported by the results in Table 2, particularly, the significantly changing parameters of the RP40 DHS in the three seasons considered.

In devising and testing a short-term DHS heat load forecasting approach described in Section 2.2, this second shortcoming is alleviated by incorporating an adaptive data set, i.e., using only the most recent data of a certain time period for each new forecasting operation.

However, before addressing the results of the short-term forecasting approach, we should also consider the last column of Table 2, which shows the DHS load curve normalized slope represented by a ratio of the temperature multiplier (slope) in the linear regression equation vs the average load (intercept) from the same equation (i.e., a/Q_0). This dimensionless coefficient shows the DHS load change relative to the average load when the temperature changes by one degree. To the best of the authors' knowledge, this is the first time when the normalized slope parameter has been used to characterise district heating systems.

In our opinion, the value of the normalized slope describes the performance of DHS (including energy losses in the buildings and the distribution networks). The lesser the value, the better system resistance

to ambient temperature drops. In other words, statistical data analysis, also including the calculation of the heating curve normalized slope, can be used to evaluate the total performance of a DHS.

3.3. Testing of DHS Heat Load Forecasting

For validation of the multiple regression model, historical data from a DHS with RP = 20 MW (Table 2) were used. The dataset contains heat load and ambient temperature records from 1 September 2017 to 31 May 2018. The forecasting simulation experiments were run for two time periods from 30 December 2017 to 7 January 2018 (9 days; the first period) and from 23 January 2018 to 26 January 2018 (4 days; the second period). Both periods are in the middle of the heating season. The recorded temperatures were used as predictors as they

were considered to be the most important predictor in heating load forecasting. Multiple regression with polynomials up to the 3rd order and for two training data set sizes (29 and 60 days) was tested.

An example of the actual and forecasted heating demand is provided for both periods with a case of 29-day training set size in Figs. 8 and 9. Note that for better readability of the illustrations and to emphasise the differences between the lines, the y-axis in both figures does not start from zero, i.e., the charts are zoomed in.

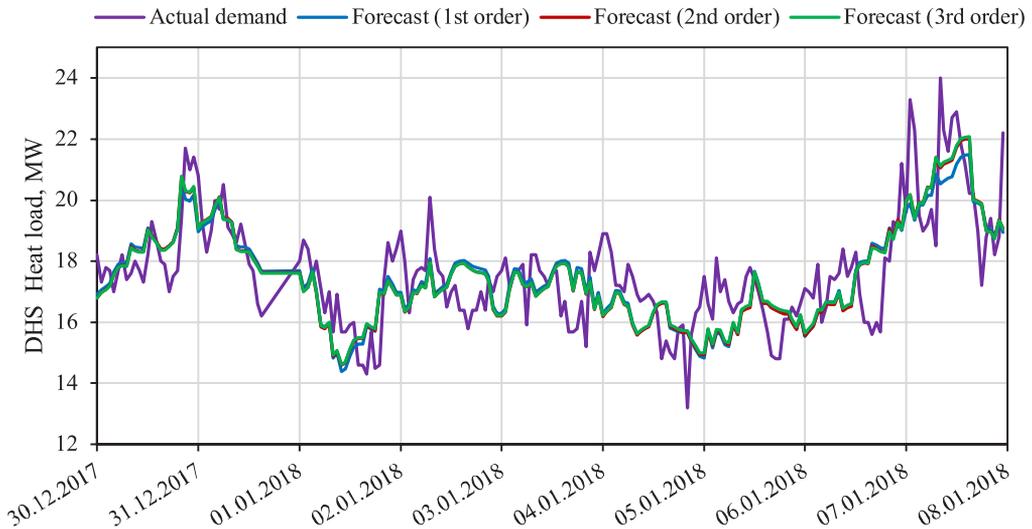


Fig. 8. Actual and forecasted demand in the 1st period with 29-day training data (RP20).

For the forecasting example from the second period, some additional information is provided in Fig. 9. Namely, the dynamically changing forecasting equations have

been extracted from the rolling-horizon experiment as displayed in the figure for each day and for each polynomial order. The term $\bar{\epsilon}_t$ in each of these equations

denotes the mean model residual (from the training dataset) averaged over the hour-of-day, i.e., this term is different for each hour within the day. To obtain the final forecasted values, this term has to be subtracted from the initial forecast evaluated with the polynomial. These equations do not have any intrinsic value on their own; however, displaying them in the figure allows seeing their change over time. In the short horizon (four days shown), this change is moderate, since only one of the 29 days in the training set gets exchanged at each step. Nevertheless, in a longer horizon they can be expected to evolve more noticeably. This

can be particularly well understood when comparing these equations to the simple linear regression equation for RP20 from Table 2, which is considerably different. The linear regression model from Table 2 is based on all the data from the stable part of the heating season, including data which are not known by the rolling-horizon forecasting model in this section, as, from its perspective, part of the data is in the future.

Table 3 summarises the performance of each of the three polynomial models tested depending on the training set size (29 or 60 days).

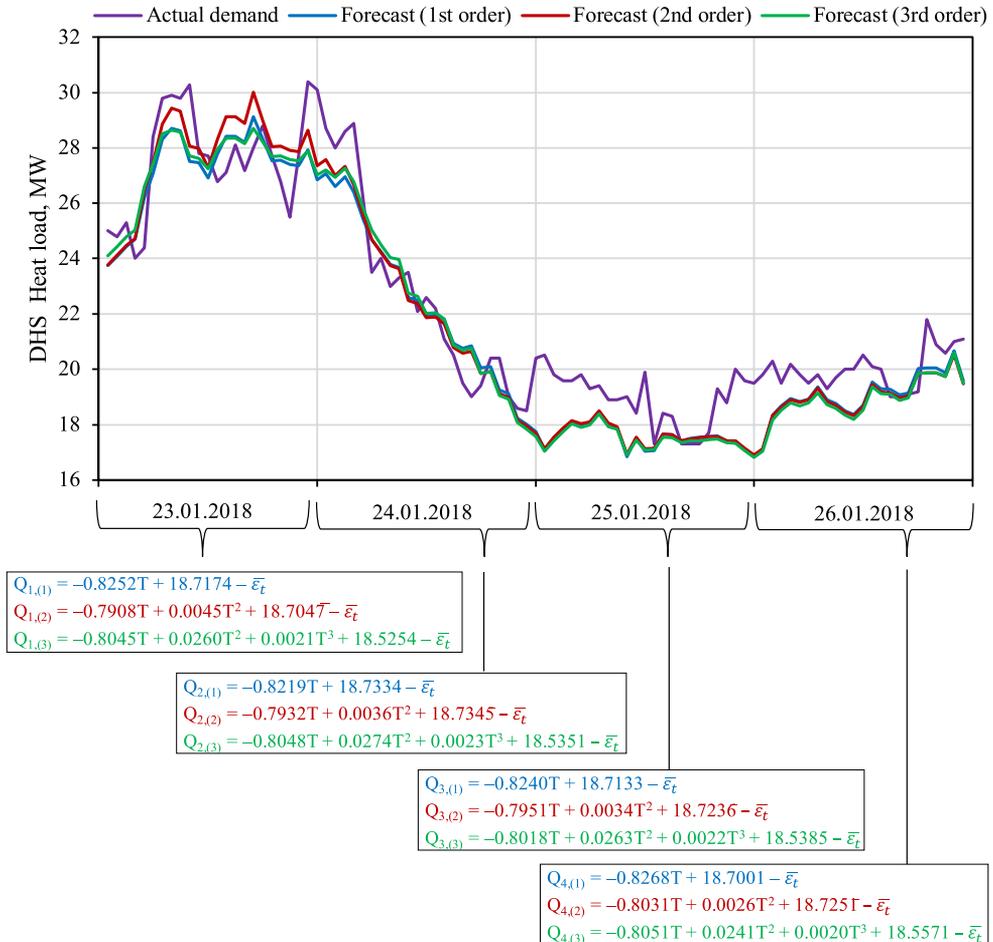


Fig. 9. Actual and forecasted demand in the 2nd period with 29-day training data (RP20).

Table 3. DHS Heat Load Prediction Accuracy

	The first period				The second period			
	Training set 29 days		Training set 60 days		Training set 29 days		Training set 60 days	
	MAPE	NBIAS	MAPE	NBIAS	MAPE	NBIAS	MAPE	NBIAS
1 st order polynomial	5.76 %	-0.40 %	5.77 %	1.30 %	4.31 %	3.84 %	5.10 %	5.45 %
2 nd order polynomial	5.56 %	-1.13 %	5.74 %	1.32 %	4.19 %	3.48 %	5.06 %	4.65 %
3 rd order polynomial	5.60 %	-1.08 %	5.70 %	1.17 %	4.55 %	4.50 %	5.24 %	5.50 %

For the first period, the 2nd order polynomial with training set of 29 days provided the best accuracy with a MAPE of 5.56 %, while the 3rd order polynomial was close behind with 5.60 % and the same training set size. In regard with NBIAS, an overall tendency to have a negative bias with a 29-day training set and a positive bias with a 60-day set can be observed. If calculated for the whole forecast set, NBIAS ranges from -1.13 % to 1.32 %, with 1st order polynomial with a 29-day training set providing the best results (-0.40 %). However, when calculating NBIAS for individual days, there is more variety and the daily errors range from -12.93 % to 8.25 %.

For the second period, similar results were received. The 2nd order polynomial with a training set size of 29 days provided the best (lowest) MAPE value - 4.19 % and is followed by the 1st order polynomial with the MAPE equal to 4.31 % and training set of 29 days. The NBIAS values also show that the accuracy of the forecast is higher for the training set of 29 days and for the 2nd order polynomial. Evidently, for the particular DHS studied, larger training set size introduces additional risk of overfitting the regression model. This effect was

especially pronounced in the second evaluated period. The reason could potentially be discerned from Fig. 9 – in the second period the actual heating demand was subject to sharp changes, presumably induced by similarly sharp deviations in the ambient temperature. A larger training set might have not allowed the forecasting model to adequately react to these changes, if a significant portion of the dataset contained records from more stable periods.

Moreover, in regard with NBIAS for the second period of the case study, there is a strong tendency for a positive bias. If calculated for the whole forecast set, NBIAS ranges from 3.48 % to 5.50 %. However, when calculating this error measure for individual days, it varies from -1.70 % to 36.55 %, but in only one of the days it is negative. Evidently, in certain instances, the forecasting approach tested can accrue bias. This likely is also connected to the notable ambient temperature change in the second period, especially since by the other metric, which disregards bias (MAPE), the second period of the case study actually has better results. Thereby, in future work, the approach should be improved to identify and eliminate bias.

4. CONCLUSIONS

The historical data statistical analysis presented in this paper demonstrates that simple linear regression for the description of different DHS heat load from a single parameter (ambient temperature) has coefficient of determination for DHS load in the range of 0.8 to 0.94 for five distinct and different DHSs in Latvia. Evidently, the heat load dependency on the ambient temperature is peculiar to particular systems. Moreover, as demonstrated in the case of one of these DHSs, this dependency can also vary notably from season to season.

A major contribution of this paper is the introduction of a descriptive parameter for DHS – a dimensionless heating curve normalized slope, calculated as the ratio of the temperature coefficient in linear regression equation (i.e., the slope) to the average DHS heat load or RP (the intercept in the regression equation). This parameter demonstrates how the heat load will change relative to its average value when the ambient temperature changes. It can be used for the evaluation of DHS performance and it can be especially useful when comparing different systems. The heating curve normalized slope can be recommended for application as a system descriptive metric also in other countries for a wide range of DHSs.

The accuracy of the linear regression models for heat load obtained in this study is to some extent insufficient as the models do not consider the ambient temperature-independent component of heating demand – hot water consumption. To overcome this problem, in devising a heat load short-term forecasting approach with satisfying accuracy, we test in this paper the application of a regression-based forecasting method that takes into account the testing model residuals averaged over the hour of the day, which presumably may contain implicit information pertaining to hot water consumption. The aim of the testing was to prove the stability of prediction accuracy for different data set sizes in different periods within the heating season.

For the selected system (RP = 20), the forecasting inaccuracy expressed by MAPE ranged in value from 4.19 % to 5.76 % and by NBIAS, from –1.13 % to 5.45 %, depending on the forecasting model parameters. Evidently, the forecasting approach is generally suitable for heat load day-ahead forecasting also in small DHSs and provides an overall satisfying accuracy; however, further improvements are necessary, especially to improve forecasting performance in periods that contain sharp changes in ambient temperature.

ACKNOWLEDGEMENTS

The research has been funded by the Ministry of Economics of the Republic of Latvia, project “Innovative Smart

Grid Technologies and their Optimization (INGRIDO)”, project No. VPP-EM-INFRA-2018/1-0006.

REFERENCES

1. Proposal for a Regulation of the European Parliament and of the Council Establishing the Framework for Achieving Climate Neutrality and Amending Regulation (EU) 2018/1999 (European Climate Law), COM/2020/80 final. Available at

- <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1588581905912&uri=CELEX:52020PC0080>
2. Central Statistical Bureau. (n.d.). *Statistical Yearbook of Latvia 2020*. Available at <https://www.csb.gov.lv/en/statistics/statistics-by-theme/economy/gdp/search-in-theme/400-statistical-yearbook-latvia-2020>
 3. Latvenetrgo (n.d.). *Ražošana*. Available at <https://latvenetrgo.lv/en/par-mums/razosana/tec2-siltuma-akumulacija>
 4. Epp, B. (2019). *15 MW SDH Plant Inaugurated in Latvia*. Available at <https://www.solarthermalworld.org/news/15-mw-sdh-plant-inaugurated-latvia>
 5. Burkhardt. (n.d.). *Biomass Power Plants: Cogeneration with Wood*. Available at <https://www.fortum.com/about-us/our-company/our-energy-production/our-power-plants/jelgava-chp-plant>
 6. Soročins, A., Rusovs, D., Nagla, J., & Žentiņš, V. (2020). The influence of the thermal storage on the electricity production in a co-generation in peak and off-peak time range. In *2020 IEEE 61st International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON 2020)*, (pp. 1–4), 5–6 November 2020, Riga, Latvia.
 7. Sauhats, A., Kozadajevs, J., Dolgicers, A., Zalitis, I., & Boreiko, D. (2019). Thermal energy storage for CHP in power market conditions. In *2019 16th International Conference on the European Energy Market (EEM)*, 18–20 September 2019, Ljubljana, Slovenia.
 8. Ivanova, P., Sauhats, A., Linkevics, O., & Balodis, M. Combined heat and power plants towards efficient and flexible operation. In *2016 IEEE 16th International Conference on Environmental and Electrical Engineering (EEEIC)*, (pp. 2434–2439), 7–10 June 2016, Italy, Florence.
 9. Latõšov, E., Volkova, A., Hlebnikov, A., & Siirde, A. (2018). *Technical Improvement Potential of Large District Heating Network: Application to Case of Tallinn, Estonia*. *Energy Procedia*, 149, 337–344.
 10. Howard, D.A, Filonenko, K., Busk, F.S., & Veje, Ch. (2020). Methodology for evaluation of district heating network efficiency. In *3rd International Conference on Electrical Engineering and Green Energy (CEEGE 2020)*, (pp. 1–9), 27–29 June 2020, Munich, Germany.
 11. Turlajs, D., Soročins, A., Cers, A., & Žīgurs, Ā. (2009). The Dependence of Quality and Efficiency of District Heating on the Type of Load Regulation. *Mehānika*, 31, 137–146.
 12. Kral, E., Kostalova, A., Capek, P., & Vasek, L. (2018). Algorithm for central heating heat load modelling. In *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, (pp. 1438 – 1439), 12–14 December 2018, Las Vegas, USA.
 13. Soročins, A., Nagla, J., & Žentiņš, V. (2019). District heating simulation model development to solve optimization problems in the market conditions. In *2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON 2019)*, (pp. 1–5), 7–9 October 2019, Riga, Latvia
 14. Petrichenko, R., Baltputnis, K., Sauhats, A., & Sobolevsky, D. (2017). District heating demand short-term forecasting. In *17th International Conference on Environment and Electrical Engineering (IEEE EEEIC 2017)*, 6–9 June 2017, Milan, Italy.
 15. Baltputnis, K., Petrichenko R., & Sobolevsky, D. (2018). Heating demand forecasting with multiple regression: Model setup and case study. In *2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, 8–10 November 2018, Vilnius, Lithuania
 16. Jovanović, R.Ž., Sretenović, A.A., & Živković, B.D. (2015). Ensemble of Various Neural Networks for Prediction of Heating Energy Consumption. *Energy and Buildings*, 94, 189–199.
 17. Baltputnis, K., Petrichenko, R., & Sauhats, A. (2017). ANN-based city heat demand forecast. In *12th IEEE PES PowerTech Conference towards and beyond Sustainable Energy Systems*, (pp. 1–6), 18–22 June 2017, Manchester, United Kingdom.

18. Blumberga, D., Blumberga, A., & Vitolins, V. *Energy Audits in Dwelling Buildings in Latvia. Data Analysis*. In Dynastee 2005 Scientific Conference, 12–14 October 2005, Athens, Greece.
19. Talcis, N., Dzelzītis, E., & Līckrašiņa, A. (2018). Return Temperature in DH as Key Parameter for Energy. *International Journal of Modern Engineering Research*, 8 (7), 88–92.