Energy and Buildings 122 (2016) 222–227

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Extreme learning machine for prediction of heat load in district heating systems

Shahin Sajjadi [a], Shahaboddin Shamshirband [b,*], Meysam Alizamir [c], Por Lip Yee [b], Zulkefli Mansor [d], Azizah Abdul Manaf [e], Torki A. Altameem [e], Ali Mostafaeipour [f]

[a] Department of Construction Management, University of Houston, Houston, TX, USA
[b] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Young Researchers and Elite Club, Hamedan Branch, Islamic Azad University, Hamedan, Iran
[d] Research Center for software technology and management (SOFTAM), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Malaysia
[e] Advanced Informatics School, Universiti Teknologi Malaysia, Malaysia
[f] Industrial Engineering Department, Yazd University, Yazd, Iran

ARTICLE INFO

Article history:
Received 25 January 2016
Received in revised form 23 March 2016
Accepted 9 April 2016
Available online 12 April 2016

Keywords:
District heating systems
Heat load
Estimation
Prediction
Extreme Learning Machine (ELM)

ABSTRACT

District heating systems are important utility systems. If these systems are properly managed, they can ensure economic and environmental friendly provision of heat to connected customers. Potentials for further improvement of district heating systems' operation lie in improvement of present control strategies. One of the options is introduction of model predictive control. Multistep ahead predictive models of consumers' heat load are starting point for creating successful model predictive strategy. In this article, short-term, multistep ahead predictive models of heat load of consumer attached to district heating system were created. Models were developed using the novel method based on Extreme Learning Machine (ELM). Nine different ELM predictive models, for time horizon from 1 to 24 h ahead, were developed. Estimation and prediction results of ELM models were compared with genetic programming (GP) and artificial neural networks (ANNs) models. The experimental results show that an improvement in predictive accuracy and capability of generalization can be achieved by the ELM approach in comparison with GP and ANN. Moreover, achieved results indicate that developed ELM models can be used with confidence for further work on formulating novel model predictive strategy in district heating systems. The experimental results show that the new algorithm can produce good generalization performance in most cases and can learn thousands of times faster than conventional popular learning algorithms.

1. Introduction

District heating is utility energy service based on provision of heat to remote customers (connected to system via heating substations) from available heat sources [1]. It enable utilization of waste or low-cost heat, which is the main precondition for district heating systems (DHS) competitiveness, when compared to onsite, individual, boilers. Another precondition is high heat density. Although some previous studies showed that sparse DHS (systems with low heat density) can be economic [2–4] their success is highly correlated with energy taxation which is country specific, even in EU. However, according to [5,6], 73% of EU 27 residents (502 million) lived in urban areas which indicate high heat density and wide prospects for further growth of DHS in Europe.

One possible way for further increase of competitiveness of DHS lies in improvement of present control strategy used. Prevailing control strategy is based on weather compensation control, where the primary supply temperature (temperature of the water from heat source) is determined from so-called "sliding diagram" [7]. This diagram provides the functional dependence between the momentary outdoor temperature and primary supply temperature of water pumped in DH network. Correlation between these two variables is almost linear, with two or three knee points and often corrected by operators after the years of usage.

However, district heating systems are complex, dynamic systems with high inertia and marked heterogeneity of users and this precondition of static correlation between the outdoor tem-

* Corresponding author. Tel.: +60 146266763.
E-mail addresses: shamshirband@um.edu.my (S. Shamshirband), altameem@ksu.edu.sa (T.A. Altameem).

http://dx.doi.org/10.1016/j.enbuild.2016.04.021
0378-7788/

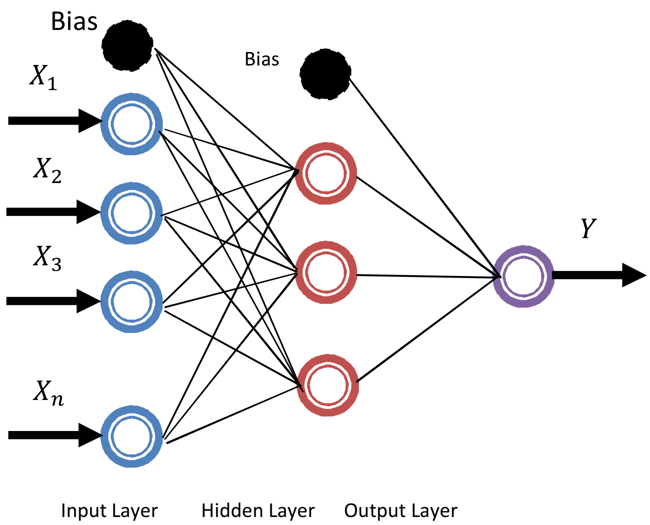

Fig. 1. The topological structure of the extreme learning machine network used in this study.

perature and heat pumped in the network does not hold. As a consequence, primary return temperature is frequently higher than needed, which is especially evident during the periods of moderate cold weather. Therefore, in order to reach the higher efficiency in DHS operation, control strategy should be directed towards lowering primary return temperatures. By lowering primary return temperature in DHS following effects on overall operation could be achieved: increased electrical output from CHP-plants (usually used as heat sources), increased heat recovery from industrial excess heat and geothermal heat, lowered distribution losses and increased coefficient of performance if heat pumps are used in heat generation [8]. Economic benefit of reduced primary return temperature was estimated at 0.05–0.5 €/MWh for 1 °C of reduced temperature [9].

In order to get the lowered primary return temperatures, produced and delivered heat from heat supply units must closely correspond to heat demanded by consumers, increased for appropriate distribution losses. Taking into account that changes in supply temperature in heat source will be "sensible" at the peripheral parts of network after considerable time, due to transport delay (water temperature changes "travel" at speed of water in network) in distribution network, the required heat to be pumped in network should be known in advance (few hours ahead), in order to avoid excessive heat input in network with very important precondition that consumers get the required heat. Therefore, predictive heat load models of all, or at least most influential, consumers in system are indispensable as inputs for advanced model predictive control. Predictive models should provide few hours ahead forecasts of required heat, where the prediction horizon can be defined according to endmost consumer in network.

In this article, we motivate and introduce the heat load prediction models of consumers, for different prediction horizon, using the data acquired from one heating substation in DHS Novi Sad, Serbia. Proposed models are developed using the soft computing approach, namely Extreme Learning Machine (ELM) [10–17]

2. System and data description

Collection of data used for subsequent analysis and building the prediction models was conducted in District heating system Novi Sad, Serbia. District heating system Novi Sad is the second biggest in Serbia with 6 heat supply units (Table 1) and one main manifold

Table 1
Heat supply units in District heating system "Novi Sad".

Name of heat supply unit	Installed power of heat supply [MW]	Fuel type
"West"	256	Natural gas
"East"	104	Natural gas
"North"	46	Natural gas
"South"	185	Natural gas
"Petrovaradin"	11,6	Natural gas
"Dudara"	2,9	Natural gas

which is used for transferring the heat from CHP unit "Novi Sad" towards the three supply units inside the town and backward.

Base load for consumers connected to heat supply units "East", "North" and "South" is covered from CHP unit "Novi Sad" via central manifold and additional heat is provided from heat supply units themselves (peak load).

Consumers are connected to district heating network (214,2 km) via heating substations. There are 3.795 heating substations in district heating system. Distribution network is implemented as 3 pipe system. One pipe is used for distribution of heat for space heating, second is used for distribution of heat for preparing domestic hot water and third is common return.

Data used for development of predictive models were taken during the heating season 2009/2010 in one of the systems' heating substation. Following variables were measured on 15 min interval:

- Outdoor temperature [°C]
- Primary supply temperature [°C]
- Primary return temperature [°C]
- Flow on primary side [m^3/h]

All the data (for above mentioned variables), together with heat load data (which were internally calculated in Ultrasonic heat meter) were regularly downloaded via SCADA system. Data was averaged on 1 h interval. No data cleaning was performed. Initially five time series were created and subsequently used for further analysis and model building. Summary of collected data is provided in Table 2. The heat power provided to the district heating network by the heat plant depends on the supply and return temperature and the flow of the water. These values can be measured, and the supplied power can be calculated as

$$P_{sup}(t) = c_p \dot{m}(t)(T_S(t) - T_R(t)) \quad (1)$$

where

$P_{sup}(t)$– is the power supplied to the district heating system by the heat plant,

c_p– is the specific heat capacity of water,

$\dot{m}(t)$ – is the mass flow of the water,

T_S – is the supply temperature at the power plant,

T_R – is the return temperature at the power plant.

3. Soft computing prediction algorithms

3.1. Extreme learning machine

Huang et al. [18] developed Extreme Learning Machine (ELM) as a novel learning algorithm for single hidden layer feed forward networks (SLFNs).

This approach has some priority compared with conventional neural networks including:

1) ELM is easy to use, and it method increase not only makes learning extremely fast but also produces good generalization performance [18];

Table 2
Statistical summary of collected data.

Variable	Minimum value	Maximum value	Mean	Median
Outdoor temperature [°C]	−7.6	24.4	5.6	4.9
Primary supply temperature [°C]	21.7	102.4	60.1	58.4
Primary return temperature [°C]	21.1	52.4	37.1	36.6
Flow on primary side [ml/s]	0.0	1630.7	1310.7	1374.9
Heat load [kW]	0.0	298.3	125.2	117.9

Table 3
List and description of potential predictors for heat load modeling.

PredictorID	Predictor	Predictor ID	Predictor	Predictor ID	Predictor
1	Outdoor temperature (t)	12	Heat load (t)	23	Primary return temp. (t)
2	Outdoor temperature (t-1)	13	Heat load (t-1)	24	Primary return temp. (t-1)
3	Outdoor temperature (t-2)	14	Heat load (t-2)	25	Primary return temp. (t-2)
4	Outdoor temperature (t-3)	15	Heat load (t-3)	26	Primary return temp. (t-3)
5	Outdoor temperature (t-4)	16	Heat load (t-4)	27	Primary return temp. (t-4)
6	Outdoor temperature (t-5)	17	Heat load (t-5)	28	Primary return temp. (t-5)
7	Outdoor temperature (t-6)	18	Heat load (t-6)	29	Primary return temp. (t-6)
8	Outdoor temperature (t-7)	19	Heat load (t-7)	30	Primary return temp. (t-7)
9	Outdoor temperature (t-8)	20	Heat load (t-8)	31	Primary return temp. (t-8)
10	Outdoor temperature (t-9)	21	Heat load (t-9)	32	Primary return temp. (t-9)
11	Outdoor temperature (t-10)	22	Heat load (t-10)	33	Primary return temp. (t-10)

Table 4
User-defined parameters for the ELM, ANN and GP models.

ELM		ANN		GP	
Number of layers	3	Number of layers	3	–	–
Neurons	Input: 6 Hidden: 3, 6, 10 Output: 1	Neurons	Input: 6 Hidden: 3, 6, 10 Output: 1	Neurons	– – Output: 1
–	–	Number of iteration	1000	Population size	512
–	–	Activation function	Sigmoid Function	Head size	5–9
Learning rule	ELM for SLFNs	Learning rule	Back propagation	Chromosomes	20–30
–	–	–	–	Number of genes	2–3
–	–	–	–	Mutation rate	91.46
–	–	–	–	Crossover rate	30.56
–	–	–	–	Inversion rate	108.53

2) In conventional neural networks all the parameters of the networks such as learning rate, learning epochs and local minima are tuned iteratively by using such learning algorithms;
3) ELM can be easily implemented and can obtain the smallest training error and the smallest norm of weights [18].

In Fig. 1 the schematic topological structure of ELM network is shown. For M arbitrary samples $(\mathbf{x}_i, \mathbf{t}_i)$, in which $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathrm{T}}\ \mathbf{R}^n$ and $\mathbf{t}_i = [t_{i1}, t_{i2}, \ldots, t_{in}]^{\mathrm{T}} \in \mathbf{R}^m$, standard single hidden layer feed forward networks (SLFNs) with N hidden nodes and activation function $g(x)$ are modeled as follows [18]:

$$\sum_{i=1}^{M} \beta_i g_i(\mathbf{x}_j) = \sum_{i=1}^{M} \beta_i g(w_i . x_j + b_i) = o_j,\ j = 1, \ldots, N \quad (2)$$

where $w_i = = [w_{i1}, w_{i2}, \ldots, w_{in}]^{\mathrm{T}}$ is the weight vector between input and hidden nodes, $\beta_i = [\beta_i 1, \beta_i 2, \ldots, \beta_i n]^{\mathrm{T}}$ is the weight vector between output and hidden nodes, and b_i is the threshold of the ith hidden node. That standard single hidden layer feed forward networks with M hidden nodes with activation function $g(x)$ as follows:

$$\sum_{i=1}^{M} \beta_i g(w_i \times \mathbf{x}_j + b_i) = \mathbf{t}_j,\ j = 1, \ldots, N. \quad (3)$$

These equations can be written as follows [18]:

$$\mathbf{H}\beta = \mathbf{T} \quad (4)$$

where

$$\boldsymbol{H} = \begin{bmatrix} g(w_1 \times \mathbf{x}_1 + b_1) & \cdots & g(w_M \times \mathbf{x}_1 + b_M) \\ \vdots & \cdots & \vdots \\ g(w_1 \times \mathbf{x}_N + b_1) & \cdots & g(w_M \times \mathbf{x}_N + b_M) \end{bmatrix}_{N \times M}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{M \times m} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_M^T \end{bmatrix}_{N \times m}.$$

where $\mathbf{H}$ is the hidden layer output matrix on neural network. The output weights can be constructed by finding least square solutions mentioned equation which the result represents the following equation:

$$\beta = \mathbf{H}^{\dagger}\mathbf{T} \quad (5)$$

where $\mathbf{H}^{\dagger}$ is the Moore–Penrose generalized inverse of the hidden layer output matrix $\mathbf{H}$.

3.2. Artificial neural networks

A very common neural network architecture is the multilayer feed forward network with a back propagation learning algorithm, which has been intensely studied and extensively used in numerous areas. Normally, it has 3 layers, including output, input hidden

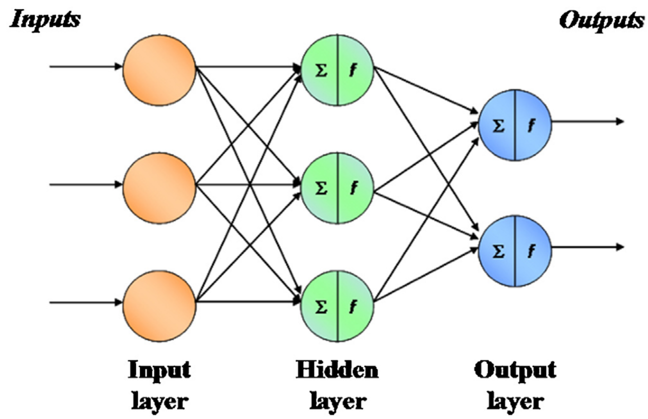

Fig. 2. ANN structure.

layer. $D = (X_1, X_2, \ldots, X_n)^T$ and$D \in R^n$ are the input vectors; in the hidden layer, $Z = (Z_1, Z_2, \ldots, Z_n)^T$ is the output of q neurons; and $Y = (Y_1, Y_2, \ldots, Y_n)^T$ and $Y \in R^m$ are the output layer's outputs. Supposing that and the threshold and the weight between the hidden and input layers are y_j and w_{ij}, respectively, and that the threshold and the weight between the output and hidden layers are y_k and w_{jk}, respectively, the following will be output of every neuron in an output and hidden layers:

$$Z_j = f\left(\sum_{i=1}^{n} w_{ij} X_i - \theta_j\right) \tag{6}$$

$$Y_k = f\left(\sum_{j=1}^{q} w_{kj} Z_j - \theta_k\right) \tag{7}$$

Here, transfer function is represented by f, which is the mapping rule for the total input of neuron to its output, can be a good instrument to add non-linearity to the design of network. The sigmoid function is one of the most popular functions, which ranges from 0 to 1 and increases monotonically. The ANNs details can be accessed from [19]. Fig. 2 shows the ANN structure. MATLAB software was used for the ANN simulations.

3.3. Genetic programming

As an evolutionary algorithm, GP defines how the relationship between output and input variables. The GP assumes a preliminary set of randomly produced programs, resultant from the accidental grouping of input variables, functions and numbers, which comprise of mathematical functions (log, exp, sin and cos) arithmetic operators and comparison/logical functions. Then, this set of probable solutions are exposed to an evolutionary process and the evolved programs' 'fitness' is assessed (the extent to which they solve the problem). Then, the programs with best fitness are nominated from the first population to exchange information and develop better programs through 'mutation' and 'crossover'. When the best program parts are exchanged with each other, it is crossover, and when programs are changed randomly to generate new programs, it is termed as mutation. Those programs with lower fitness are eliminated. This process of evolution is repeated to find representative expressions labeling the data that are possible to interpret scientifically to gain knowledge. The GP has been explained in details in [20].

4. Results and discussion

Predictive models of heat load were developed for: 1, 2, 3, 4, 5, 8, 10, 12 and 24 h ahead. Five models, for prediction horizon from 1 to 5 h ahead, are intended for control purpose, while the models for 8, 10, 12 and 24 h ahead for planning purpose. When developing multi step ahead prediction models, two approaches are possible:

- Iterated prediction approach and
- Direct prediction approach.

In iterated prediction scheme, one step ahead prediction is used for building the subsequent predictions (i.e. for p steps ahead). Contrary to this, in direct prediction approach, separate (direct) prediction models are developed for each prediction horizon.

Merits of each of these approaches are discussed extensively [29–31]. Here direct approach is adopted. Main benefit of direct approach is that procedure is intuitive [32] and that there is no accumulation of forecasting error, which is not the case for iterated scheme. Moreover, direct approach was successfully applied in electricity load forecasting [33–35] the domain quite similar to one discussed in this work.

4.1. Input variables for model building

From available time series the list of most influential variables on heat load, to be used for predictive model building, was created. Previous studies suggest that the most influential variables are: outdoor temperature and primary return temperature [36–38]. Moreover, due to pronounced inertia of DHS, additional predictors (features), made of time lagged heat load, outdoor temperature and primary return temperatures were created and subsequently used for building predictive models. List of all, 33 potential predictors, is shown in Table 3.

4.2. Evaluating accuracy of proposed models

Predictive performances of proposed models were measured for training and testing data set and presented as root means square error (RMSE), Coefficient of determination (R^2) and Pearson coefficient (r). These statistics are defined as follows:

1) root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}, \tag{8}$$

2) Pearson correlation coefficient (r)

$$r = \frac{n\left(\sum_{i=1}^{n} O_i \times P_i\right) - \left(\sum_{i=1}^{n} O_i\right) \times \left(\sum_{i=1}^{n} P_i\right)}{\sqrt{\left(n\sum_{i=1}^{n} O_i^2 - \left(\sum_{i=1}^{n} O_i\right)^2\right) \times \left(n\sum_{i=1}^{n} P_i^{\,2} - \left(\sum_{i=1}^{n} P_i\right)^2\right)}} \tag{9}$$

3) coefficient of determination (R^2)

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right) \times \left(P_i - \overline{P_i}\right)\right]^2}{\sum_{i=1}^{n}\left(O_i - \overline{O_i}\right) \times \sum_{i=1}^{n}\left(P_i - \overline{P_i}\right)} \tag{20}$$

where P_i and O_i are known as the experimental and forecast values of heat load, respectively, and n is the total number of test data.

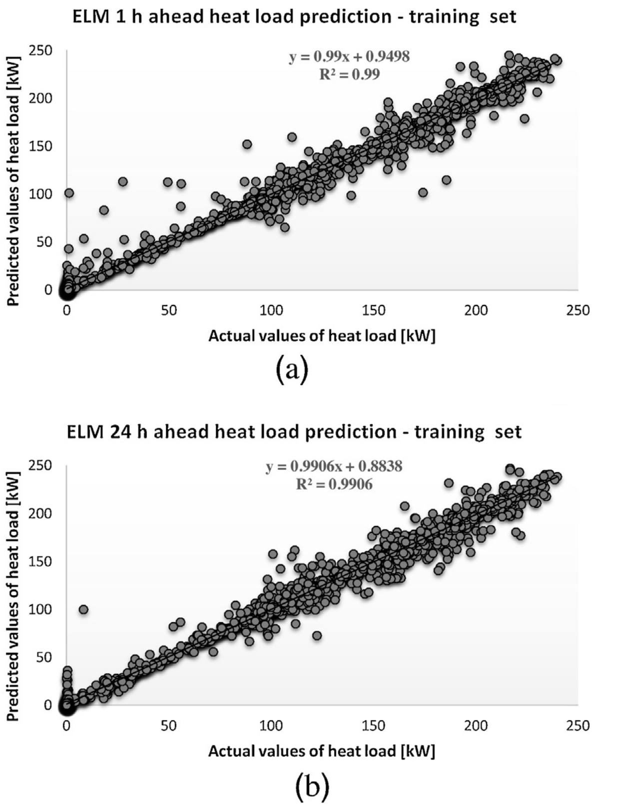

Fig. 3. Actual and predicted values of heat load (training data set) using ELM method for (a) 1 h ahead and (b) 24 h ahead predictions.

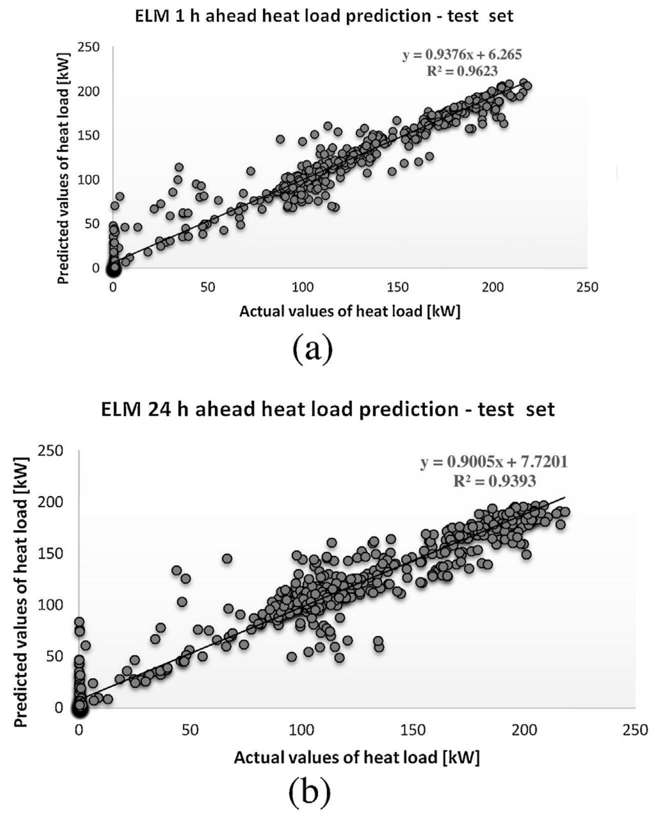

Fig. 4. Actual and predicted values of heat load (test data set) using ELM method for (a) 1 h ahead and (b) 24 h ahead predictions.

4.3. Architecture of soft computing models

The optimal parameters of the ELM, ANN and GP modelling frameworks employed in this study are presented in Table 3. Data are divided in random training and testing groups and average results are reported for all different training and testing groups in order to generalize the models' application. In this case all groups are included in the training of the models and in later in testing as well.

The average computation time for the ELM modelling was around 350 s using a PC with Intel Core Duo CPU E7600 @3.06 GHz and 2-GB RAM. The average computation time for the ANN with back propagation training algorithm and GP modelling using the same PC with the same performances was 428 and 439 s, respectively. For the ELM modelling, MATLAB software was used (Table 4).

4.4. ELM results

In this section performance results of ELM heat load predictive models are reported. Fig. 3 presents the accuracy of developed ELM heat load predictive models for horizon of 1 h ahead Fig. 3(a) and 24 h ahead Fig. 3(b), using the training data set. It can be seen that the most of the points fall along the diagonal line for 1 h ahead prediction model. Consequently, it follows that prediction results are in very good agreement with the measured values for ELM method. This observation can be confirmed with very high value for coefficient of determination, especially for 1 h ahead prediction. As was expected prediction accuracy deteriorates with enlarged prediction horizon.

To further prove the potential of the proposed ELM method for heat load prediction in DHS, in the following the method's suitability in terms of testing prediction is assessed. The scatter plot of the observed head load values against predicted ones for testing phase is illustrated in Fig. 4(a) using the ELM model for 1 h ahead prediction and in Fig. 4(b) using the ELM model for 24 steps ahead prediction. The number of either overestimated or underestimated values produced is limited. Consequently, it is obvious that the predicted values enjoy high level precision.

Table 5
ELM, ANN and GP heat load predictive models for different prediction horizons.

Prediction horizon		ELM		ANN		GP	
		RMSE	R^2	RMSE	R^2	RMSE	R^2
1 h ahead	training	7.6839	0.99	12.0273	0.9754	14.6806	0.9633
	testing	13.8049	0.9623	14.2846	0.9596	16.1586	0.9482
2 h ahead	training	8.7360	0.987	17.3725	0.9487	19.9091	0.9326
	testing	21.7397	0.9064	23.5433	0.8904	27.0437	0.8557
3 h ahead	training	10.2172	0.9822	18.0250	0.9447	22.1102	0.9169
	testing	23.0337	0.8963	24.6608	0.8806	26.4655	0.8617
4 h ahead	training	9.7348	0.9839	15.4833	0.9592	18.3963	0.9425
	testing	24.2002	0.8852	25.4779	0.8722	26.2306	0.864
5 h ahead	training	8.7450	0.987	15.3421	0.96	19.9393	0.9324
	testing	19.9588	0.9223	20.2960	0.9197	21.1994	0.9122
8 h ahead	training	6.5716	0.9927	13.7415	0.9679	18.9755	0.9389
	testing	15.3016	0.9547	15.9618	0.9503	19.9248	0.9211
12 h ahead	training	10.5097	0.9812	16.8430	0.9518	21.2769	0.9231
	testing	29.0023	0.8377	30.3166	0.8208	31.3007	0.8078
24 h ahead	training	7.4101	0.9906	14.8911	0.9622	17.9667	0.945
	testing	17.5764	0.9393	17.9663	0.9366	18.8980	0.9294

In order to confirm the ELM approach on a more tangible basis, ELM models' results was compared with ANN and GP methods. For each of three aforementioned methods, nine different direct predictive models were developed (for different prediction horizons). Table 5 presents the results for the three methods.

For training data set, all three methods provide similar results. However, ELM models for 4, 5, 10 and 12 h ahead outperform GP and ANN models.

For testing data set results are markedly different. ELM model outperforms other methods in all cases. While the prediction results for 1 and 2 h ahead are quite similar for all methods, for higher prediction horizons RMSE for ELM models is fairly lower. Additionally, RMSE values for all values of prediction horizon fluctuate in very narrow range.

5. Conclusion

Predictive models for consumers' heat load in DHS were developed using the ELM method.

Observations collected in heating substation in DHS during the heating season 2009/2010 were used for model building and testing. Nine predictive models, for different prediction horizons (from 1 h to 24 h ahead), were created using the ELM method. A comparison of ELM method with GP and ANN was performed in order to assess the prediction accuracy. Accuracy results, measured in terms of RMSE, r and R^2, indicate that ELM predictions are superior then GP and ANN, which particularly stands for higher prediction horizons. Additionally, results reveal robustness of method. Developed method can be used with confidence in further work, directed in creating model predictive control in DHS.

Acknowledgment

This project was supported by the Fundamental Research Grant Scheme (FRGS) - FP071-2015A from the Ministry of Higher Education, Malaysia. Also, the authors would like to extend their sincere appreciation to the Deanship of Scientific Research at king Saud University for its funding this research group No. (RGP – 1436-035).

References

[1] S. Frederiksen, S. Werner, District Heating and Cooling, Studentlitteratur AB, Lund Sweden, 2013.
[2] C. Reihav, S. Werner, Profitability of sparse district heating, Appl. Energy 85 (2008) 867–877.
[3] S. Amiri, B. Moshfegh, Possibilities and consequences of deregulation of the European electricity market for connection of heat sparse areas to district heating systems, Appl. Energy 87 (2010) 2401–2410.
[4] S.F. Nilsson, C. Reidhav, K. Lygnerud, S. Werner, Sparse district- heating in Sweden, Appl. Energy 85 (7) (2008) 555–564.
[5] Population Division, Department of Economic and Social affairs, United Nations World Urbanization Prospects: The 2009 Revision (Data in digital form POP/DB/WUP/Rev.2009).
[6] Population Division, Department of Economic and Social affairs, United Nations (2010).
[7] B. Jaćimovic, B. Zivković, S. Genić, P. Zekonja, Supply water temperature regulation problems in district heating network with both direct and indirect connection, Energy Build. 28 (3) (1998) 317–322.
[8] H. Gadd, S. Werner, Achieving low return temperatures from district heating substations, Appl. Energ. 136 (2014) 59–67.
[9] S. Frederiksen, S. Werner, District Heating and Cooling, Lund, Studentlitteratur.
[10] Shahaboddin Shamshirbanda, Kasra Mohammadib, Por Lip Yeea, Dalibor Petkovićc, Ali Mostafaeipourd, A comparative evaluation for identifying the suitability of extreme learning machine to predict horizontal global solar radiation, Renewable Sustainable Energy Rev. 52 (December) (2015) 1031–1042.
[11] Shahaboddin Shamshirbanda, Kasra Mohammadib, Hui-Ling Chenc, Ganthan Narayana Samyd, Dalibor Petkoviće, Chao Maf, Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: a case study for Iran, J. Atmos. Sol. Terr. Phys. 134 (November) (2015) 109–117.
[12] Q. Yu, Y. Miche, E. Séverin, A. Lendasse, Bankruptcy prediction using extreme learning machine and financial expertise, Neurocomputing 128 (2014) 296–302.
[13] X. Wang, M. Han, Online sequential extreme learning machine with kernels for nonstationary time series prediction, Neurocomputing 145 (2014) 90–97.
[14] L. Ghouti, T.R. Sheltami, K.S. Alutaibi, Mobility prediction in mobile ad hoc networks using extreme learning machines, Procedia Comput. Sci. 19 (2013) 305–312.
[15] D.D. Wang, R. Wang, H. Yan, Fast prediction of protein–protein interaction sites based on extreme learning machines, Neurocomputing 128 (2014) 258–266.
[16] R. Nian, B. He, B. Zheng, M.V. Heeswijk, Q. Yu, Y. Miche, et al., Extreme learning machine towards dynamic model hypothesis in fish ethology research, Neurocomputing 128 (2014) 273–284.
[17] P.K. Wong, K.I. Wong, C.M. Vong, C.S. Cheung, Modeling and optimization of biodiesel engine performance using kernel-based extreme learning machine and cuckoo search, Renew. Energ. 74 (2015) 640–647.
[18] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.
[19] N.Y. Liang, G.B. Huang, H.J. Rong, P. Saratchandran, N. Sundararajan, A fast and accurate on-line sequential learning algorithm for feedforward networks, IEEE transactions on neural, Networks 17 (2006) 1411–1423.
[20] J.R. Koza, Genetic Programming: On the Programming of Computers by Natural Selection, MIT Press, Cambridge, MA, 1992.