Control and Cybernetics

vol. **41** (2012) No. 1

Consumer-oriented heat consumption prediction*

by

Maciej Grzenda

Warsaw University of Technology,
Faculty of Mathematics and Information Science,
00-661 Warszawa, Pl. Politechniki 1, Poland
e-mail: M.Grzenda@mini.pw.edu.pl[1]

Abstract: The advent of modern low-cost monitoring and wireless transmission systems results in unprecedented availability of measurement data potentially available in near real-time mode. In particular, some of the remote meter reading systems can be used to collect data on an hourly or even sub-hourly basis. This allows the utility companies to model and predict consumer behaviour more precisely than before. In this study, the way the monitoring data can be used to model heat consumption at individual premises supplied with heat by a district heating system, is proposed.

The proposed algorithm is based on customer partitioning used to devise a number of group models serving the needs of consumers sharing similar consumption profiles. Self-organising maps are used to group averaged long-term time series, while the short-term time series provide a basis for group prediction models. Particular attention has been paid to a wider hydraulic modelling perspective, as the application of the proposed method to provide assumed demand for hydraulic model of a district heating system is envisaged. The approach has been validated using a real data set. Results show that in spite of a limited number of monitored consumers, group prediction models, constructed using the algorithm proposed in this study, can significantly reduce demand prediction error.

Keywords: district heating systems, demand prediction, neural networks.

1. Introduction

1.1. Control strategies for utility networks

The operators of utility networks, such as water supply, power grid networks or district heating systems (DHS) have to provide sufficient quality of services on

*Submitted: November 2010; Accepted: January 2012

[1]and Megabit sp. z o.o., 04-501 Warszawa, ul. Plowiecka 1/3, Poland

regular basis. In case of pipe-based networks, hydraulic modelling (Bhave and Gupta, 2006; Walski et al., 2004) has been used to predict system behaviour under known operational settings of sources, e.g., pressures applied in water purification plants and assumed consumer behaviour. In the case of water supply systems, being the most widely used pipe-based utility networks, consumption profiles have been usually developed based on individual studies or the investigation of daily or weekly water consumption profiles for flow control zones. For planning and design purposes, even simpler approach is considered i.e. the variation of water consumption is accounted for by a peaking factor (Bhave and Gupta, 2006). While this approach is sufficient for steady state simulations, extended period simulations (EPS) (Bhave and Gupta, 2006; Walski et al., 2004) are used to consider changing demand throughout the day. Such simulations are required to achieve flow and pressure time series comparable to real time series collected in the system and predict the hydraulic behaviour of a system.

Limited consumption data available for such studies have frequently resulted in only rough estimation of consumer needs in modelling projects. Unfortunately, billing records, provide enough information to determine a baseline demand, but not enough to determine fluctuations in demand on a finer time scale required for EPS (Walski et al., 2004). These temporal variations can be described for 24-hour cycle by *diurnal* demand pattern based on the investigation of the data acquired from water meters. In general, the problem remains an open issue as temporal demand variations are known to change not only on daily basis, but also weekly and annually and depend on weather conditions, too (Bhave and Gupta, 2006; Walski et al., 2004).

At the same time, the rapid development of low-cost remote metering systems has started to provide a vast volume of data. Both data collected in consumers' premises and key network points have become available. Such data can be used to model consumer behaviour and optimise the configuration of an entire utility system including the changes in production strategies aiming to minimise the cost of operation while ensuring the quality of service.

The development of an appropriate control strategy is crucial also for a DHS (Balate et al., 2007). As transported heat is used for space heating and hot tap water, the required volume of heat constantly changes. At a consumer level, it depends on a number of factors, weather conditions, life style and time of day being the most important of them. Moreover, due to its strong relation to at least two leading factors i.e. time of the day and week and outdoor temperature, heat consumption profile can not be developed as a function of one argument only. What makes the control strategy for DHS systems particularly difficult to develop is also the fact that it may take even several hours to transfer extra heat from a heat source to a heat consumer. In other words, particularly large inertia of DHS systems compared to other systems, e.g., water supply networks has to be emphasized (Youen, 2009). Not surprisingly, errors in load prediction result in significant problems in managing a DHS (Park et al., 2009; Sandou et al., 2005).

At the same time, district heating system operators should minimise the cost of operation by eliminating unnecessary heat production. Hence, it is extremely important to build an appropriate control strategy for a DHS. A control system for DHS requires a number of problems to be solved (Balate et al., 2007). One of them is the implementation of a demand prediction module. Still, even recent studies (Park et al., 2009) notice that most optimisation solutions reported so far do not include demand prediction.

The development of heat meters (Móczar, Csubák and Várady, 2002; Ye, Zhang and Diao, 2005) allows to capture heat consumption of every consumer on hourly basis, which provides crucial data for demand modelling. However, for this data to be used in a control system in near real time manner, on-line transmission is needed. Complex systems including both wired and wireless communication from a heat meter to the central control system have been proposed (Móczar, Csubák and Várady, 2002; Ye, Zhang and Diao, 2005) and could be potentially used to achieve this objective. Still, even though it is technically possible to monitor all the heat meters of all the consumers in an on-line manner (Ye, Zhang and Diao, 2005), this solution is not feasible due to the cost of transmission devices and transmission itself. Thus, an important aspect of the study is to verify whether the data from a limited number of constantly monitored consumers can be used to develop prediction models.

1.2. The objectives of the study

Unlike the previous works on the topic, discussed in detail in Section 2.2, the paper concentrates on load prediction on the client side - not at a heat source. In other words, instead of predicting total load in a heat source, being the sum of heat consumption and heat losses, individual needs of different consumer groups are modelled. The research is based on newly available heat consumption time series data from a group of monitored consumers. Moreover, the hypothesis that averaged long-term time series can be used to partition heat consumers and develop group prediction models based on short-term time series data collected from monitored consumers is investigated. An important assumption is that only the data easily available in utility companies are used. In particular, long-term consumption time series have been developed using the billing data. Groups of consumers sharing similar average long-term behaviour have been identified using self-organising maps (SOM), see Haykin (1999). Both models predicting an average consumption of all the consumers and models answering the need for consumption prediction at a group level have been constructed. These models are referred to as *global* and *group* models, respectively. The study summarises previous works on the topic (Grzenda and Macukow, 2006, 2009; Grzenda, 2008). In the previous work, Grzenda and Macukow (2009), two different techniques, namely modified evolutionary construction of multilayer perceptrons (MECoMLP), Grzenda (2008), and multilayer perceptrons (MLP), Haykin (1999), trained with gradient methods, were used as prediction models. The first technique was based on autonomous selection of MLP architecture, connection weights and a number of standard prediction algorithms via evolutionary programming (Fogel, 1999) to construct a prediction model. The MECoMLP follows the results of the previous work devoted to load prediction in power systems (Grzenda and Macukow, 2002). In the current study, particular attention is paid to multilayer perceptrons and the analysis of error rates in case MLP networks are used to construct both group and global models.

The work concentrates on heat consumption modelling in district heating systems. Nevertheless, the proposed techniques can be adapted to the needs of other utility networks. More precisely, the main objective of the paper is to propose and validate the method of constructing consumption prediction models that satisfies real-life limitations and can be potentially used by different utility network operators. Equally importantly, the question whether aggregate consumption data provide basis for partitioning consumers relevant for short-term prediction is answered. Moreover, in order to develop a method for DHS utility companies, the following system constraints had in particular to be considered:

- a limited number of monitored consumers,
- significant inaccuracy of heat meters influencing the calculation of prediction errors,
- periodically noisy and close to zero consumption time series, caused by the nature of heat consumption and the local scale of DHS, not found in typical country or region-wide electrical power systems,
- inertia of space heating systems making the temporal inadequacy of heat supply partly acceptable.

All these aspects were addressed, when developing the proposed method, too.

The proposed method can be also adapted to the needs of water and gas supply systems for water and gas consumption modelling, respectively. Moreover, the same techniques could be applied in electrical power consumption modelling. Nevertheless, the unique features of district heating systems, which make precise demand modelling at a consumer group level both especially desired and more attainable, compared to other systems, are:

- The long latency of the system. Should the heat production be too low, the time needed to transfer extra heat from the sources to the consumers may reach several hours. This phenomenon is a particular feature of district heating systems due to the unique characteristics of heat transfer. Hence, the need for high precision demand modelling, as potential errors in heat demand estimation can not be immediately corrected by increasing heat production.
- Heat meters require the integration of flow and temperature in consumer's heat exchanger. Hence, as heat meters are partly electronic devices, they can be easily extended to transfer the hourly data to the central server. Thus, the hourly data become available for district heating systems. The hourly data are usually not available to this extent in many other systems such as gas and water supply systems, as the latter systems still use many

non-electronic meters for measuring the consumption.

As a consequence, the proposed technique can be applied to district heating systems, but also to other utility systems. The development of remote meter reading systems, observed within last few years, should provide a basis for group-based prediction in other utility systems, too.

The remainder of the work is organised as follows:

- The modelling background and the differences between source-oriented and consumer-oriented prediction are discussed in Sectjion 2.
- The way the partitioning of consumers can be performed is described in Section 3.
- Section 4 outlines the construction of group and global models.
- Prediction errors for both categories of models are compared in Section 5. The section includes the discussion of the impact of measurement accuracy on model evaluation.
- Finally, conclusions and the main directions of future work are outlined in Section 6.

2. The role of demand prediction in hydraulic models

2.1. Hydraulic and thermodynamic models of DHS

Modern DHS (Kato et al., 2008; Park et al., 2009; Sandou et al., 2005; Youen, 2009) can benefit from hydraulic and thermodynamic models. A model represented by a graph of edges, being an abstraction of pipes, and points, representing junctions and consumers, can be used to simulate the behaviour of a system. The hydraulic and thermodynamic model can provide the data on temperatures, flows and pressures in any point of the system. An illustration of a model, depicted on Fig. 1, shows a simple model with one heat source, several supply pipes marked with solid lines and receive pipes marked with dotted lines. The heat transported through the piping system is used to supply heat via heat exchangers (Davidsson and Vernstedt, 2004; Youen, 2009) to internal piping systems in customers' buildings. Receive pipes are used to transfer heat carrier, typically hot water, back to the source, where its temperature is increased again. The model graph is usually developed basing on the content of a geodatabase (Arctur and Zeiler, 2004). An increased interest in hydraulic modelling can be also considered as a way of optimising network operation and exploiting the benefits of geographical information systems (GIS), Maguire, Kougoumijan and Smith (2008). An in-depth discussion of the generation of a model graph basing on geodata can be found in Walski et al. (2004).

What is important, is that even in a middle-size DHS thousands of customers are supplied with heat through a number of heat exchangers. In a typical layout one heat exchanger provides heat to one building or several residential buildings. This means hundreds of consumption nodes exist in a model graph. Assumptions regarding heat consumption in these nodes have to be made to perform hydraulic and thermodynamic simulations and to optimise heat production and network settings, in turn.

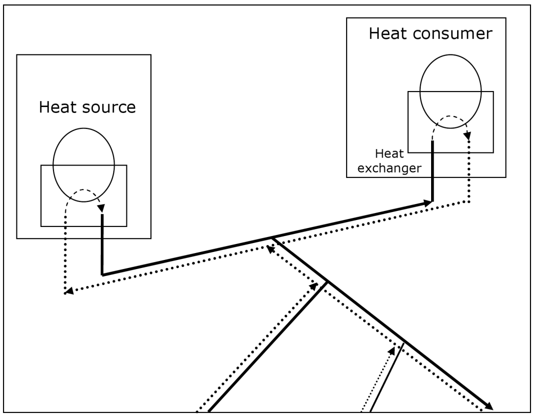

Figure 1. Network model of a DHS system

To simulate flows, pressures and temperatures, the volume of heat transferred by a heat exchanger to a consumer and consumed by a consumer has to be assumed. Once an assumption regarding heat consumption is made, different control strategies regarding heat production can be applied to the model. By evaluating their impact on the hydraulic and thermodynamic behaviour of a system, the best control strategy can be selected.

2.2. Total load vs. individual consumption

2.2.1. The decomposition of heat production

Demand for heat at a heat source is significantly different from the sum of consumption at consumers' premises, as heat loss occurs in a piping system (Park et al., 2009). In general

$$\sum_{s=1}^{S} Q_s = Q_{\mathrm{L}} + \sum_{c=1}^{C} Q_c \tag{1}$$

where Q_s stands for the volume of heat produced in source s, Q_{L} for heat loss and Q_c denotes heat consumption at consumer c, $S \ll C$. The proportion of Q_{L} and total consumption $\sum_{c=1}^{C} Q_c$ constantly changes and depends not only on weather conditions and time of day, but also pressure applied at individual sources. Moreover, the hydraulic and thermodynamical properties of the system affect heat distribution. What is yet more important, is that the heat produced at time t reaches the consumer after a delay of even up to several hours. This delay depends not only on the distance, but also flow speed and other hydraulic properties of the system.

In addition, any Q_s time series reflects the total consumption of many individual consumers. This results in a much smoother profile than an individual consumption Q_c time series, which is largely affected by stochastic heat consumption for hot tap water needs.

2.2.2. Source-side production modelling techniques

Until recently, the only detailed time series data spanning long periods of time were available for heat sources. Hence, the research on heat demand modelling has been referring to heat source needs. Numerous techniques have been applied to predict the load at heat sources. Research has been concentrated on the direct prediction of the total volume of heat Q_s required in every source in a DHS during each hour. A number of methods have been used for this purpose. One of the first approaches was to use a neural network model with a two-layer associative memory using Learning Vector Quantisation to predict the load of the system (Kashiwagi and Tobi, 1993). A related problem of cooling load prediction in a district heating and cooling system has been discussed in Sakawa et al. (1999). The authors have compared the results of load prediction using a nonlinear auto regressive moving average (NARMA) and different filtering methods. Another approach is to apply superposition of Box-Jenkins models based on the correlation analysis of time series (Balate et al., 2007). Still, the problem of load prediction in DHS is reported to be in general an open issue (Park et al., 2009; Sandou et al., 2005), as most studies assume that consumers' demands are given and perfectly known. The works described above are largely inspired by power load prediction systems. In the latter case, some of the state-of-the-art solutions were developed for the Polish power system (Siwek and Osowski, 2009; Siwek et al., 2010). At the same time, the authors notice that most papers concerning power systems are devoted to large power systems, usually spanning entire countries (Siwek et al., 2010). Moreover, the prediction of load in local power systems, which is the scale of all DHS, is reported to be more challenging, due to high variability of load, which makes the prediction more difficult (Siwek et al., 2010).

To sum up, in the case of DHS, traditionally, the total volume of heat needed by consumers was assumed known and was usually experimentally estimated based on weather conditions and time of day. Moreover, the entire load of the system observed at a heat source was predicted. This approach refrains from determining to what extent individual needs of different consumers are satisfied. As a consequence, the predicted values can not be used for hydraulic calculations of the distribution system, as the latter calculations require heat consumption at individual heat exchangers to be provided. The latter requirement is mandatory, as flows and temperatures in the system largely depend on the spatial distribution of the overall load in the network of a DHS.

At the same time, extensive studies have shown that hot tap water demand profiles vary greatly and can differ even in the same region (Lane and Beute, 1996). The differences in normalised peak consumption in different towns can reach 25% (Lane and Beute, 1996). This phenomenon has been analysed using real data acquired in municipalities close to Cape Town (Lane and Beute, 1996) and clearly illustrates that different groups of consumers are likely to have diverse heat demand profiles. Finally, it is important to note that the heat consumption for hot tap water needs occurs in the form of sudden relatively short peaks, during which this form of heat demand largely exceeds heat consumption for space heating.

2.2.3. The limitations of source-side prediction

The source-side demand models can be based on:

- The prediction of heat demand based on the investigation of past heat production time series.
- The calculation of minimal production needed. Such calculations can be based on input data being the demands of individual consumers and heat losses calculated through hydraulic simulation.

It is important to note the following limitations of the former approach, i.e. the source-side demand prediction based on past heat production data:

- The total source production observed in the past, when used to devise prediction models, affects the prediction models. More precisely, the data mining models trained with past production will predict insufficient production, in case they are trained with too low production observed under some conditions in the past. Similarly, should the past production be higher than needed, the prediction models trained with the past excessive production will predict higher source-side demand than actually occurring.
- Moreover, any attempts to change the pressure applied in a heat source to a new value, not used in the past under the same weather conditions, would make the source-side predictions incorrect. For instance, by increasing flow speed in the system, the engineer controlling the system may reduce the delays in heat transportation while increasing the cost of pumping. Thus, the peak production could be started later than with the previous flow speed. Obviously the scale of this change, would not be known to a data mining model not provided with such historical data. In other words, the source-side prediction model based on historical heat production is improper for the investigation of new control strategies. Similarly, it could easily recreate the deficiencies of a former control strategy, e.g., insufficient or excessive heat production. Consequently, it could not be used to support new control strategy. Therefore, little or no economic gains from a better control strategy can be expected as long as source-side prediction models trained with historical source data are used.
- Once the source-side prediction models are trained with source-side data,

the impact of new large consumers on the systems remains unknown. Their impact on source-side demand would need to be experimentally observed over a longer period of time to retrain the models with new source-side heat production data.

- Consumer-side consumption profiles are mandatory for the modelling of individual sections of an entire DHS. In particular, by running hydraulic and thermodynamic simulations, it can be determined that pipe diameters are insufficient to transport required volume of heat in some sections of the system - currently or in the future, once new consumers are connected. Obviously, for such calculations heat demand on a consumer side needs to be known.

Therefore, the only reliable technique to estimate production needed in a heat source and validate the control strategy is to perform hydraulic simulations taking into account the demand of individual consumers, their distance from heat source, and temperatures and pressures set by the system engineer in every heat source. This corresponds to a standard approach to water supply system modelling. In the latter case, a hydraulic model is configured with water source production settings and daily or weekly consumer demand patterns. Hydraulic simulations are executed next to verify if water pressures set in the pumping stations are sufficient to deal with predicted demand of individual water consumers (Walski et al., 2004). Hence, production settings can be tuned to minimise the cost of pumping while ensuring sufficient level of service, i.e. water pressure in client premises (Bhave and Gupta, 2006; Walski et al., 2004).

In the case of water supply systems, diurnal curves showing daily or weekly variation of water consumption are created for every consumer group. Such curves can be easily created for an entire season in water supply systems, as they are a function of time only. They are used as an input for hydraulic modelling. As a result, simulated pressures and flows become available for every part of a water supply system (Bhave and Gupta, 2006; Walski et al., 2004). Unfortunately, as stated before, heat consumption is related to time of the day and outdoor temperature. Hence the need for prediction models at consumer side rather than simple diurnal curves i.e. functions of time of the day and week only.

To sum up, the required heat production in a heat source should result from hydraulic and thermodynamic calculations of the DHS taking into account current control strategy and predicted consumer needs. As long, as source-side heat production is estimated based on past production, the potential for system optimisation is largely diminished. The consumer-side prediction provides necessary input to the calculation of source settings-temperature and pressure set by the system engineer and resulting heat production. So, consumer-side prediction is not a replacement, but a vital step towards objective calculation of source-side heat demand and production.

2.2.4. The requirements for consumer-side prediction

Before client-side prediction models are constructed, key requirements for developing them should be analysed. Thus, while individual prediction models for every heat source can be constructed, it is impossible to model every consumer separately. One of the main reasons is the lack of real measurement data for most consumers. Hence the need for methods predicting the load for consumers and taking into account existing constraints. Such predictions could be used as input data for hydraulic calculations of the network. Flow, pressure and temperature in the DHS network can not be calculated without assumptions on the volume of heat consumed by the clients via heat exchangers. Consumer-side prediction could be used instead of rough a priori assumptions based on averaging heat consumption in the entire DHS or a part of it. The remainder of this paper investigates the way newly available consumer-side data can be used to build such prediction models and the scale of advantages arising from separate prediction models for different consumer groups.

3. The identification of consumer groups

Before prediction models serving different groups of consumers can be built, consumer groups have to be established. Detailed attributes of individual consumers could be used for this process. These could include thermal properties of buildings, the data on the number of inhabitants and other attributes. However, even this data would not provide a basis for the entire load profile. One of the factors that are hardly quantifiable is the individual notion of heat comfort. Therefore, the actual heat consumption profiles have been used instead. Such profiles have been obtained from the billing database and normalised. Unfortunately, the billing database contains aggregate consumption only, measured on a monthly basis. Detailed consumption time series, acquired by a monitoring system on hourly basis, are available for a limited number of consumers only.

The data sets used in this study come from one of the Polish district heating systems, operating a medium size system supplying heat to over 1000 heat exchangers. $N = 1109$ sales profiles have been obtained from the company billing data. Every sales profile of a consumer $c_i, i = 1, \ldots, N$ is represented by a vector $S_i = (s_{i,1}, \ldots, s_{i,12})$, where $s_{i,m}$ denotes the average heat sale to a consumer i during month m. Only 72 consumers are or were monitored in the past. Hence, the heat consumption data $\tilde{d}^i_h$, collected for individual hours h is available for $i = 1, \ldots, 72$ consumers. As stated before, the cost of maintenance of a larger consumer monitoring system would be too high for a DHS. Therefore, the question of whether the available hourly data is sufficient for modelling consumer groups, should be answered.

Based on the $S_i, i = 1, \ldots, N$ vectors, normalised profiles $\tilde{S}_i, i = 1, \ldots, N$ have been obtained. These have been used to build a SOM network. Winner Takes All with Conscience (CWTA), Haykin (1999), and Neural Gas (NGAS), Martinetz, Berkovich and Schulten (1993), algorithms have been used to tune neuron weights. The two latter algorithms were applied to overcome deficiencies of the standard WTA algorithm. Among these deficiencies, the overrepresentation of regions with low input density is not of least importance, Haykin (1999).

As a consequence, a limited number of significant consumer groups was established through SOM network investigation. In fact, some 40% of neurons of the 10x10 lattice represented at most two consumers (Grzenda and Macukow, 2009). In other words, even though the algorithm promoting neuron diversity was used, it was revealed that the majority of consumers were associated with 8 neurons only (Grzenda and Macukow, 2006), which shows relative concentration of similar consumer demand vectors. In general, the average demand profiles of these large consumer groups followed inversed outdoor temperature profile. Nevertheless, the peak consumption is reached by different groups in different months - January, February, August or December. Moreover, the share of summer heat consumption in the overall yearly consumption is another factor differentiating the groups. The details of consumer grouping and identified consumer groups are discussed in Grzenda and Macukow (2006).

Based on a limited number of large consumer groups identified by 10x10 lattice, the SOM lattice was reduced to a lattice of 5 rows per 5 columns i.e. 25 neurons. The self-organising process was based on the Neural Gas (Martinetz, Berkovich and Schulten, 1993) algorithm. The algorithm defines the neighbourhood function of neuron i and input vector x as $G(i,x) = exp(-\frac{m(i)}{\lambda})$, where $m(i)$ is the index of a neuron on the sorted list of distances between neurons and input vector x. The algorithm parameter λ, playing a similar role to neighbourhood radius, and the learning coefficient η were linearly reduced during the process, as suggested in Osowski (2000). More precisely, λ was set to linearly decrease from 5 to 0 during the process. Similarly, the learning coefficient of a SOM neuron η (Haykin, 1999; Osowski, 2000) was linearly decreasing from initial $\eta_0 = 0.1$ (Haykin, 1999). Finally, 100 000 iterations of the ordering and tuning process were performed altogether.

Next, for each neuron j in the lattice, at the end of the network tuning phase, a winning count Y_j has been calculated as follows:

$$Y_j = \sum_{i=1}^{N} eval(i,j) \tag{2}$$

while

$$eval(i,j) = \begin{cases} 1 & \text{if } \parallel \tilde{S}_i - w_j \parallel = min_k \parallel \tilde{S}_i - w_k \parallel \\ 0 & \text{otherwise} \end{cases} . \tag{3}$$

In this way, neurons representing at least 40 consumers, i.e. $Y_j \geq 40$, have been identified. $Y_j : j = 1, ..., J$ denotes here the number of demand patterns $\tilde{S}_i$ that are the most similar to the weight vector w_j in terms of Euclidean distance. As a consequence, five neurons matching the condition $Y_j \geq 40$ were determined. Moreover, over 80% of consumers were partitioned into $C_i, i = 1, \ldots, 5$ consumer groups, defined by the corresponding weight vectors w_i.

It is worth noticing that there are significant differences in long term demand profiles of consumer groups identified in this way. In particular, for some of the consumer groups identified, peak consumption is observed in January, while in case of other groups the maximum consumption is observed in February or even August. In the latter case, this clearly suggests that these consumers do not use heat purchased from DHS for space heating needs.

The numbers of monitored consumers $card(M_i)$ in groups $C_i, i = 1, \ldots, 5$ are 34, 31, 4, 1 and 0, respectively, $M_i \subset C_i$. The monitored consumers consist of two categories: consumers with transmission devices and consumers with data loggers, which can be used to collect data for several days, but can not be remotely read.

Because of the insufficient number of monitored consumers in groups C_4 and C_5, prediction models for these groups can not be constructed. A global prediction model has to be used for any consumer not in $\bigcup_{i=1}^{m} C_i$, $m = 3$. Selection of consumers that should be constantly monitored to provide input data for demand prediction models was discussed in Grzenda (2009). It has been shown that from the demand modelling perspective, the existing layout of monitored consumers does not guarantee that all the important consumer groups are sufficiently monitored.

4. The construction of prediction models

4.1. Model development

In developing a prediction model providing basis for DHS modelling and its potential application in near-real time, the following constraints must be considered:

- Due to high volatility of heat consumption, prediction should be made on an hourly basis. Longer time intervals could be potentially useful for long-term sales prediction, but do not provide valuable input for the hydraulic modelling of DHS.
- Prediction should be made a few hours ahead of the actual consumption to take into account the time period between the change to production settings is made and the time the heat carrier with a modified temperature reaches the most distant consumers.

Two approaches to modelling heat consumption at consumer premises can be considered:

- a prediction model ξ serving the needs of all the consumers i.e. predicting an average normalised consumption of all consumers. This model is referred to as a *global model*,

- group prediction models μ_i predicting the averaged normalised consumption of consumers belonging to a group $C_i, i = 1, \ldots, m$. Every *group model* μ_i is trained with the data from $M_i, i = 1, \ldots, m$ group of monitored consumers. Moreover, it can take advantage of a normalised consumption of all consumers, which is also used as a part of its input data.

Finally, once the group and global models are developed, the prediction error rates can be calculated. The entire process is summarised in Fig. 2.

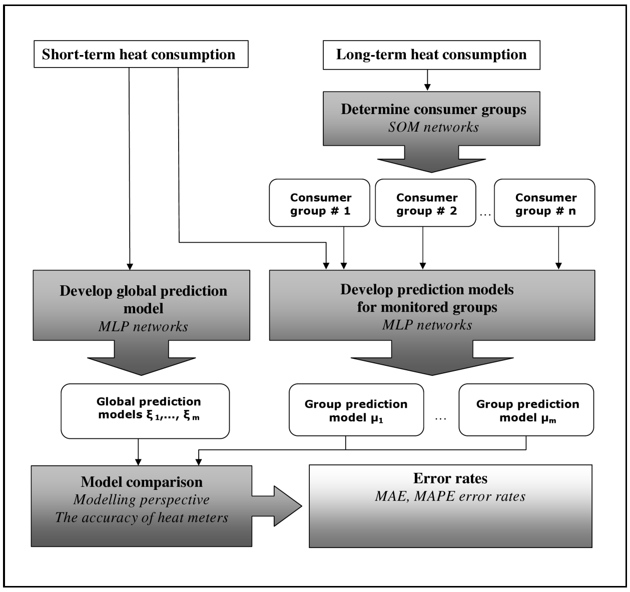

Figure 2. The construction of group and global models - an overview of the process

4.2. Data preprocessing

To provide data for model development, for each consumer group $C_i, i = 1, \ldots, m$ the following data sets have been constructed:

- training and validation *group data set* composed of ca. 75% of available data,
- testing group data set composed of the remaining data not used in the training group data set,
- training and testing *global data sets* prepared in the same way, but using the data from all monitored consumers $\bigcup_{i=1}^{J} M_i$.

In the case of the DHS being analysed, the heat consumption varies greatly over the year, depending on weather conditions, the time of the day, day of week and also the season, as many consumers are hotels or houses offering rooms for rent.

Thus, in order to ensure unbiased model evaluation and comparison, the data from different hours and weather conditions (coming from every fourth calendar day) has been placed in the testing data set.

For every consumer group $C_i, i = 1, \ldots, m$ two test cases have been considered. In each case $r = 50$ algorithm runs have been performed. The primary objective of the first test case was to build and evaluate the prediction models μ_i forecasting average heat consumption for consumers of groups C_i, using *group data set* i.e. both average consumption acquired from monitored consumers M_i and average consumption from all monitored consumers $\bigcup_{i=1}^{J} M_i$.

Let $A_h = avg_{i \in \bigcup_{j=1}^{J} M_i}(\tilde{d}_h^i)$ and $G_h^j = avg_{i \in M_j}(\tilde{d}_h^i)$ stand for average heat consumption of all consumers and consumers of M_j group during hour h, while h stands for both date and hour index. An assumption has been made that it may take up to 5 hours for the water to reach a heat consumer. Thus, in both test cases, G_{h-6}^j and A_{h-6} are the latest values that can be used for prediction. Prediction made using more recent data could not be used to obtain required heat supply at time h.

Thus, each data pattern p_h in a group data set contains the following values: average consumption among group members G_{h-6}^j, G_{h-7}^j,G_{h-8}^j,G_{h-24}^j, average consumption by all consumers A_{h-6}, A_{h-7},A_{h-8},A_{h-24}, hour h, outdoor temperature τ_h, day of week index, time of day status for annotating peak hot tap water hours, and the value to be predicted G_h^j.

Global models ξ_i were trained using *global data sets*. Each data pattern p_h in a global data set contains the following values: average consumption by all consumers A_{h-6}, A_{h-7}, A_{h-8}, A_{h-24}, hour h, outdoor temperature τ_h, day of week index, time of day status for annotating peak hot tap water hours, and the value to be predicted A_h.

Global data sets have been used for the second test case aiming to verify the efficiency of the global prediction model based on averaged data from all consumers. On the one hand, such a model does not address any specific group of consumers, on the other hand, by averaging the data from a larger number of consumers the global model is less likely to be affected by the chaotic behaviour of individual consumers of a group.

It is worth emphasising that other attributes, e.g., G_{h-9}^j can be also included. However, the number of attributes should be limited in order not to diminish generalisation capabilities of the models.

The real heat consumption data at monitored consumers from two different heating seasons, namely the years 2005 and 2007, was used. It was applied after normalisation as raw time series $\tilde{d}_h^i$. The data are from January, February, March and April, i.e. the time when heat was used both for hot tap water needs and for space heating. Both data coming from data loggers and acquired through on-line transmission of meter readings is present in the data sets. Due to the gaps in measurements causing incompleteness of the data, minor differences in the number of available records p_h are observed. The number of available measurements $\tilde{d}_h^i$ and resulting p_h records placed in individual group data sets is summarised in Table 1. The number of available measurements is the highest for group 1 and equal to $card(\{\tilde{d}_h^i : i \in M_1\}) = 20137$, which is caused by the largest number of monitored consumers in the group $card(M_1) = 34$. Finally, different gaps in the measurement data of every group result in a group of global models $\xi_i, i = 1, \ldots, m$ being developed, as shown in Fig. 2.

It is worth emphasising that inevitable incompleteness of data is a common issue when dealing with measurement data. Two factors contribute to the problem: sudden failures of measurement devices and data transmission problems.

Table 1. The number of data patterns p_h in group data sets

Group i	$card(M_i)$	$card(\{\tilde{d}_h^i : i \in M_i\})$	No. of p_h records for group i
1	34	20137	1663
2	31	18217	1683
3	4	1918	1536

5. Results

5.1. Comparison of absolute errors for group and global models

In all the cases, group and global prediction models in the form of MLP networks have been trained using Levenberg-Marquardt (LVM) algorithm. The number of neurons in hidden layers has been manually set. The group models have been trained on group data sets, while the global models on the global data sets, as defined in Section 4.2. The resulting global and group MLP models are denoted by ξ_i and μ_i, respectively. $E()$ stands for mean absolute error (MAE) calculated based on $r = 50$ separate LVM runs. The original $E()$ rate was largely affected by individual algorithm runs trapped in local optima with extremely high error rates. Thus, the reported $E()$ rate has been calculated without considering the outliers, in order to reliably compare average error rates. Average error rates $E()$ calculated on testing data sets have been summarised in Table 2. All the group models were tested on the testing group data sets. The global models were provided with the testing global data sets constructed for exactly the same sequence of days as these contained in testing group data sets. As a consequence, the input of global models in the testing phase was the averaged data from all monitored consumers. The output of both group and global models was compared with the averaged group consumption observed at the testing time points. Hence, the error rates shown in Tables 2 and 3 are calculated by comparing the output of individual prediction models with the actual average behaviour of a group. In particular, the output of global

Table 2. Results summary

			MAE errors	
Group	card(C_i)	card(M_i)	$E(\xi_i)$	$E(\mu_i)$
1	363	34	0.01732	0.00961
2	339	31	0.01872	0.01206
3	136	4	0.02855	0.02426
Avg.err.	n.a.	n.a.	0.02153	0.01531

models trained with the averaged data from all consumers is compared with the actual heat consumption in every group. Hence, the benefits arising from training the global models with much more smooth averaged data collected from all consumers are compared with the advantages of training a number of group models with much more noisy data, but coming from the respective group of consumers only.

The average error rate $\overline{E(\mu_i)} = 0.01531$ of the group models on the testing data set is on average 28.9% lower than the error rate $\overline{E(\xi_i)} = 0.02153$ of the global models tested on the same set of time points. This clearly shows the potential of group prediction models. At the same time, the error rate of a group prediction model depends greatly on the number of monitored consumers. The higher the number of monitored consumers, the lower the error rate of a model built using their data. This confirms previous results obtained with MECoMLP method (Grzenda and Macukow, 2009) and MLP networks. In particular, the error rate of the group models trained with evolutionary programming was lower than the corresponding error of the global models trained using the same MECoMLP technique (Grzenda and Macukow, 2009).

For the two largest groups, the improvement, measured as a difference between $\overline{E(\mu_i)}$ and $\overline{E(\xi_i)}$, is even larger and equal to 39.9% of the latter MAE error rate. At the same time, in Grzenda and Macukow (2009) it was observed that the average error rate obtained for the training data set $\overline{E(\alpha_i)}$ was lower by 22% than $\overline{E(\mu_i)}$, α_i and μ_i standing for global and group hybrid prediction model trained with MECoMLP technique, respectively. Similarly, the error rate of the MLP-based models trained with a gradient method, measured on the training data sets, $\overline{E(\xi_i)}$, was lower by 29% than $\overline{E(\mu_i)}$ (Grzenda and Macukow, 2009). In other words, in general it is easier to train a global prediction model using the data from all consumers than a group prediction model based on the data coming from a lower number of consumers. Still, group prediction models produce lower error rates for consumers from their groups.

To sum up, the results of simulations show that averaged long-term time series can be successfully used to identify consumer groups suitable for prediction purposes. Prediction models constructed for such consumer groups produced significantly lower error rates on the testing data sets than the global prediction models. This improvement is even larger when the two biggest consumer groups $C_i, i = 1, 2$ are considered. Nevertheless, an analysis of percentage errors is needed to evaluate the practical merits of the models.

5.2. Investigation of percentage errors

In addition to mean absolute errors reported in Section 5.1 an investigation of percentage errors was made. The following factors were considered:

- The accuracy of the measurement data used for this study. Heat consumption is calculated in heat meters as an integrated multiplication of the flow rate and the temperature difference (Kusui and Nagai, 1990). Hence, the normative error rate resulting from integrating three underlying measurements is significant and equal to $\pm 7.5\%$ of reported value. In some cases it is believed to exceed even this rate due to long periods of near-zero flow causing sedimentation problems. This rate can not be neglected, as it definitely exceeds typical error rates e.g. for water flow measurements, which can be as low as $\pm 2\%$.
- In hydraulic modelling of utility networks, a common acceptance criterion for the model calibration process is defined as an acceptable error rate in a specified percentage of cases. For instance, in the case of water distribution systems, one of the criteria for accepting a hydraulic model is that 95% of field test measurements should be within $\pm 0.75m$ or $\pm 7.5\%$ of the maximum head loss across the system, whichever is greater (Walski et al., 2004). The error rates for 100% and 85% of field test measurements are expected not to exceed $\pm 15\%$ and $\pm 5\%$, respectively (Walski et al., 2004). These criteria accept the inevitable random behaviour of system users that may affect the system hydraulics. In the case of DHS and heat consumption, as stated in Seciton 2.2, a largely stochastic heat consumption for hot tap water has to be considered.
- Taking into account a limited number of monitored consumers in a group $card(M_j)$, averaged heat consumption $G_h^j = avg_{i \in M_j}(\tilde{d_h^i})$ may temporarily be equal to zero.
- Heat meters have been polled regularly but at slightly different times in a round robin manner. Hence, the hourly consumption time series, which provides the basis for this study, is affected by time differences of the underlying raw heat consumption data.

Therefore, first a formal algorithm for calculating percentage error rates in view of measurement errors E_{MAPE} is described in Algorithm. 1. The algorithm is an extension of standard Mean Absolute Percentage Error (MAPE), defined as: $MAPE = \sum_{i=1}^{n} \frac{|y_i - d_i|}{d_i}$. In particular, $MAPE = E_{MAPE}(\epsilon_a, \epsilon_p, d, y); \epsilon_a = 0, \epsilon_p = 0$. The main difference is that in the case of $\epsilon_a > 0$ absolute measurement errors are taken into account, which makes it possible to calculate error rates for $d_i = 0$ at the same time. Moreover, $\epsilon_p > 0$ can be used to address the relative accuracy of the measurements. As stated before, heat consumption measurements are affected by a large relative measurement error, too large to be neglected when dealing with prediction error rates.

Then, investigation of percentage error rate, as a function of the percentage of time steps considered in the evaluation, is made. The latter algorithm is formally defined as Algorithm 2. It follows the idea of model calibration acceptance criteria, cited above. In other words, instead of calculating a single MAPE rate calculated for an entire testing set, a number of MAPE error rates are calculated to show the MAPE error rates as a function of the percentage of time steps considered.

Hence, by investigating $\alpha < 100$ one can determine whether a prediction model generates significant errors most of the time or in a limited number of cases. Moreover, the latency of a DHS system and central heating systems in consumers' premises makes short-term problems more acceptable than in the case of electrical power or water supply systems. Simply, due to thermal inertia of the buildings temporarily insufficient heat supply may not negatively affect the heat comfort of the residents. Finally, as stated before, the proposed method follows a common hydraulic modelling practice, i.e. setting the boundary limits that should be met by 90% or 95% of simulated time steps rather than by 100% of cases.

Input: $d = [d_1, \ldots, d_n]$ - a vector of expected demand, $y = [y_1, \ldots, y_n]$ - demand prediction, ϵ_p - allowed percentage error, ϵ_a - allowed absolute error
Result: $E_{MAPE}(\epsilon_a, \epsilon_p, d, y)$
begin
    $E_{MAPE}(\epsilon_a, \epsilon_p, d, y)$=0;
    for $i = 1, \ldots, n$ **do**
        $d_{max} = max(d_i \times (1 + 0.01 \times \epsilon_p), d_i + \epsilon_a)$;
        $d_{min} = min(d_i \times (1 - 0.01 \times \epsilon_p), d_i - \epsilon_a)$;
        if $d_{max} \leq y_i$ **then**
            $E_{MAPE}(\epsilon_a, \epsilon_p, d, y) = E_{MAPE}(\epsilon_a, \epsilon_p, d, y) + \frac{y_i - d_{max}}{d_{max}}$
        end
        if $y_i \leq d_{min}$ **then**
            $E_{MAPE}(\epsilon_a, \epsilon_p, d, y) = E_{MAPE}(\epsilon_a, \epsilon_p, d, y) + \frac{d_{min} - y_i}{d_{min}}$;
        end
    end
    return $\frac{E_{MAPE}(\epsilon_a, \epsilon_p, d, y)}{n}$;
end

Algorithm 1: The calculation of percentage error rate considering limited accuracy of measurement data

Input: $d = [d_1, \ldots, d_n]$ - a vector of expected demand, $y = [y_1, \ldots, y_n]$ - demand prediction, ϵ_p - allowed percentage error, ϵ_a - allowed absolute error, α - the percentage of cases considered in calculating error rates
Result: $E^{\alpha}_{MAPE}(\epsilon_a, \epsilon_p, d, y)$
begin
$d^s, y^s = sortvectors(d, y, MAE(d, y))$;
$n_\alpha = \lceil n \times \alpha \times 0.01 \rceil$;
$d_\alpha = d^s_1, \ldots, d^s_{n_\alpha}$
$y_\alpha = y^s_1, \ldots, y^s_{n_\alpha}$
return $E_{MAPE}(\epsilon_a, \epsilon_p, d_\alpha, y_\alpha)$;
end

Algorithm 2: The calculation of percentage error rate for α% of time steps

Algorithm 2 describes the way the calculations used for the latter problem were made. The averaged MAPE error rates $\overline{E^{\alpha}_{MAPE}}(\epsilon_a, \epsilon_p, d, y)$ are reported for every consumer group in Table 3.

In addition, two figures illustrate the averaged $\overline{E^{\alpha}_{MAPE}}(\epsilon_a, \epsilon_p, d, y)$ error rate as a function of α. In all cases $\epsilon_p = 7.5$, $\epsilon_a = 0.00005$ were applied and $r = 50$ algorithm runs were performed to provide averaged error rates. $\overline{E_{MAPE}()}$ will stand for $\overline{E^{100}_{MAPE}}(\epsilon_a, \epsilon_p, d, y)$.

Fig. 3 shows an averaged error rate as a function of α for group and global models developed in 50 algorithm runs for group 2. The averaged error rate $\overline{E^{\alpha}_{MAPE}}(\epsilon_a, \epsilon_p, d, y)$ is equal to 7.69% for $\alpha = 100$, i.e. when all time steps are considered. For 95% of time steps this error is equal to 5.54%.

A similar investigation was performed for group 3. Only four consumers are monitored in this group. Thus, even averaged heat consumption is more likely to go down to zero or near-zero values, which is actually observed in the data. As a consequence, large percentage errors are observed. In the latter case, by treating the error rate as a function of α, a significant insight into the error rates can be provided. While for group models $\overline{E^{100}_{MAPE}}(\epsilon_a, \epsilon_p, d, y) = 5931.34\%$,

Table 3. $\overline{E^{\alpha}_{MAPE}}(\epsilon_a, \epsilon_p, d, y)$ error rates - group and global models

	Group models			Global models		
Group index	$\alpha = 100$	$\alpha = 95$	$\alpha = 90$	$\alpha = 100$	$\alpha = 95$	$\alpha = 90$
1	9.63	5.74	4.47	22.66	14.48	11.20
2	7.69	5.54	4.33	13.20	11.09	9.61
3	5931.34	71.30	14.11	7542.31	209.10	16.52

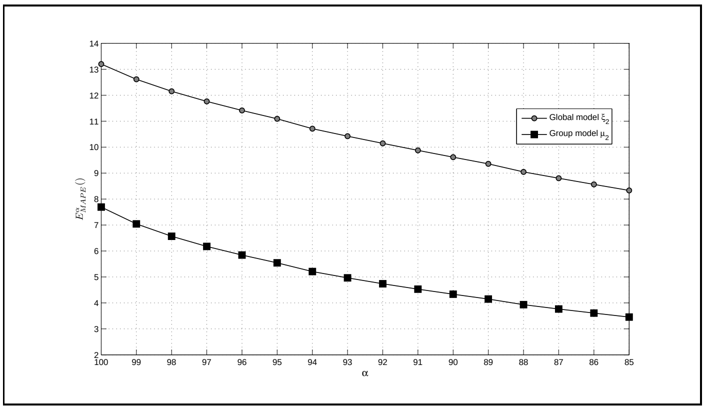

Figure 3. MAPE error rates for α% of time steps - consumer group 2

$\overline{E^{90}_{MAPE}}(\epsilon_a, \epsilon_p, d, y) = 14.11\%$. In fact, the error rate curves visualised in Figs. 3 and 4 show the percentage of time α, for which the error rate of $k\%$ can be expected. In the case of heat prediction models, this turns out to be more informative than a single value, reported for $\alpha = 100\%$. The reason is the fact that heat consumption, unlike power consumption, is more likely to take zero or near-zero values. Moreover, DHS typically spans a few hundred or thousands consumers. While a single heat consumer group may contain just a few hundred consumers, power load prediction is usually developed for large systems used by hundreds of thousands consumers, which results in much more smooth non-zero averaged consumption profiles. The prediction at a smaller scale of individual cities, which is the scale typical for DHS, has been reported more challenging also in the case of power systems (Siwek et al., 2010).

Still, no matter whether group or global models are applied, the error rate for the third consumer group is significant. The limited number of monitored consumers in the group contributes to the problem. On the one hand, this group has its own long-term consumption profile. On the other hand, insufficient number of short-term time series for the group is available. As a consequence, the prediction errors can not be accepted for this group. Moreover, group model provides similar prediction accuracy as a global model. Hence, in the case of consumer groups not monitored by a sufficient number of devices, a global model can be used instead. Still, the only way to obtain more accurate demand predictions for significant groups of consumers is to install additional transmission devices capturing their unique short-term consumption time series. The impact of the number of monitored consumers on prediction error has been investigated in a greater detail in Grzenda (2009).

Finally, graphical comparison of real heat consumption and prediction for the consumer groups C_1 and C_2 was made. High prediction accuracy can be observed most of the time. Relatively significant prediction errors occur only

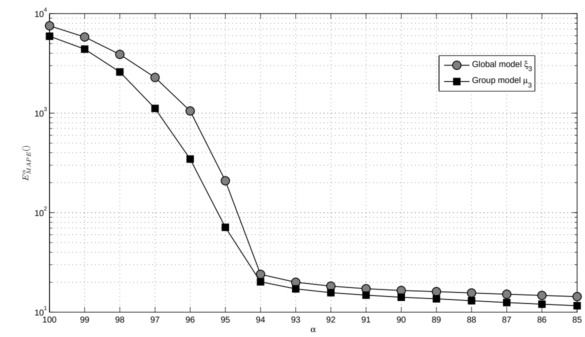

Figure 4. MAPE error rates for $\alpha\%$ of time steps - consumer group 3

during some hours of high volatility periods present in the last part of the testing set. For clarity reasons, the results for two sequences of testing days are shown. First of all, in Figs. 5 and 6 the results for the period of limited demand volatility are shown. It can be observed that prediction made by group models μ_1 and μ_2 matches with relatively high accuracy the actual consumption. Moreover, group models predict the actual consumption of their groups more precisely than the global models. For instance, this can be observed during the second and third testing days shown in Fig. 5.

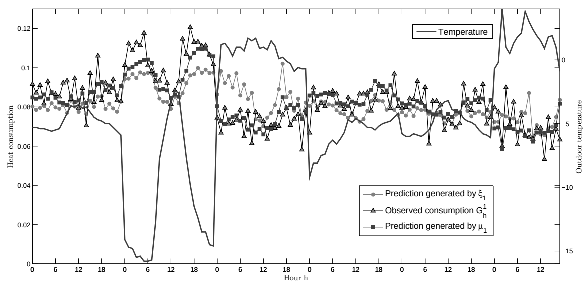

Figure 5. Prediction - consumer group C_1 - the first subset of testing days

To provide a better illustration of the problem, a sequence of testing days was also selected to visualise some days with relatively high volatility of heat consumption, which results in lower accuracy of prediction. Fig. 7 shows the actual heat consumption and the averaged prediction produced by group and global models in the case of C_1 group for this subset of testing days. Fig. 8 shows the same comparison made for the same set of time points and C_2 group.

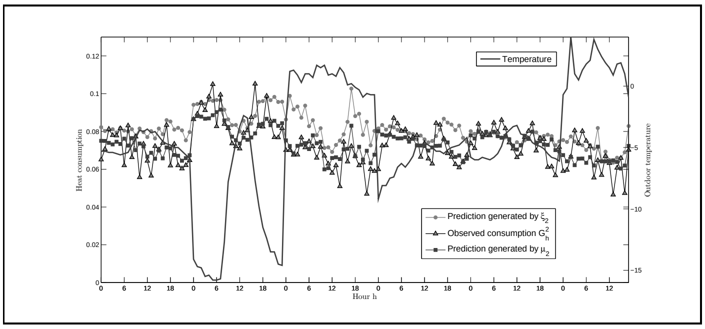

Figure 6. Prediction - consumer group C_2 - the first subset of testing days

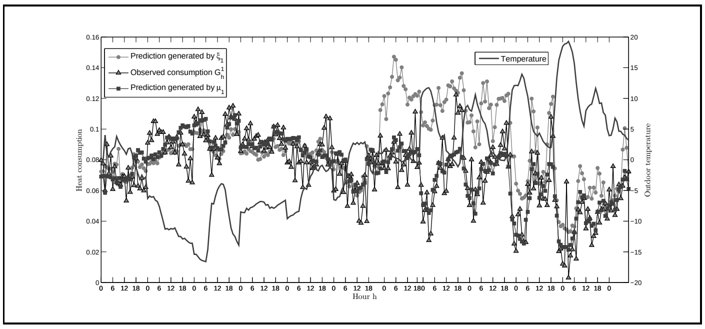

Figure 7. Prediction - consumer group C_1 - the second subset of testing days

It should be noted that while the first group decreased heat consumption in the second half of the analysed period, as visualised in Fig. 7, the opposite tendency can be observed in Fig. 8, i.e., in the case of C_2 group. Whereas global models predict an averaged behaviour of all consumers, the group models captured the unique tendencies in their groups. In Fig. 7 it can be observed that the global prediction ξ_1 is reaching 0.15 and is largely higher during the second half of the testing period than the actual consumption of the group at these days, which is at this time approximately in the range [0.08,0.12]. While Fig. 7 shows the global prediction exceeding the actual group needs, the results visualised in Fig. 8 illustrate the opposite tendency occurring for group 2. In the latter case the prediction delivered by global model is in the range [0.1,0.15] and is significantly lower than the actual consumption, which at the same time remains in the range [0.1,0.25]. In both cases, the group prediction models μ_1 and μ_2 forecast the heat consumption of the corresponding groups more accurately.

However, at some periods a global model may be more accurate than its group counterpart. This may be expected in the periods of limited diversity between consumer groups. Nevertheless, the summary results shown in Table 3 show the overall superiority of group models over global models. More precisely, for the first group $\overline{E_{MAPE}()}$ of group prediction μ_1 is lower by 57% than the corresponding rate for global model ξ_1. In the case of the second group, the percentage error $\overline{E_{MAPE}()}$ of group prediction μ_2 is lower by 41% than the error made by global models ξ_2.

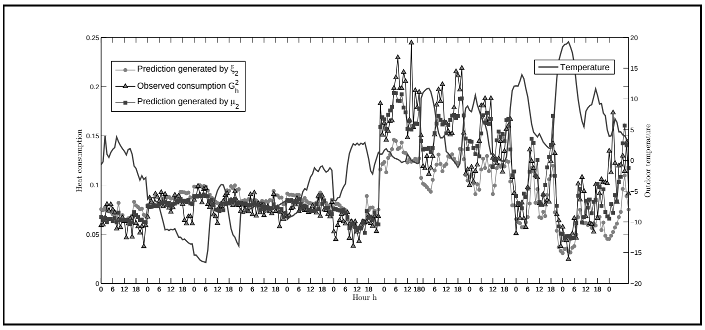

Figure 8. Prediction - consumer group C_2 - the second subset of testing days

5.3. Comparison of percentage error measures

Finally, the comparison of error rates calculated with the proposed $E^{\alpha}_{MAPE}()$ function against other norm-based error rates was made. The comparison was performed using the same neural networks and data sets as in the preceding part. Two error functions following the same pattern i.e. $E_{MAPE}() = \frac{\|y-d\|}{\|d\|}$ were used. As before, $d = [d_1, \ldots, d_n]$ denotes a vector of expected demand and $y = [y_1, \ldots, y_n]$ - demand prediction. The first error function was defined as $E_{MAPE_1}() = \frac{\|y-d\|_1}{\|d\|_1}$, i.e. it was based on the L_1 norm. In addition, $E_{MAPE_2}() = \frac{\|y-d\|_2}{\|d\|_2}$, based on the Euclidean norm, was included. In the case of the error function $E^{\alpha}_{MAPE}(\epsilon_a, \epsilon_p, d, y)$, defined in Algorithm 2, ϵ_p, ϵ_a were set to the same values as in Section 5.2, i.e. $\epsilon_p = 7.5$ and $\epsilon_a = 0.00005$, respectively.

The same predicted time series y as these illustrated in Figs. 5, 6, 7, and 8 were used. More precisely, as before, for every consumer group, and model category (group or global), the prediction y was generated by averaging the prediction made by individual MLP-based models, created in LVM-based training sessions. Each of the error functions $E^{\alpha}_{MAPE}(), E_{MAPE_1}()$ and $E_{MAPE_2}()$ was calculated using prediction y obtained on the testing data sets and the actual consumption d. The results of these calculations are summarised in Table 4.

Moreover, $\alpha = 100$ was used, i.e. E^{100}_{MAPE} is reported. Hence, all the available testing time points were considered in the calculations of all the error functions. As before, it can be observed that the error rates attained with

Table 4. Comparison of error measures - group and global models

	Group models			Global models		
Group	$E^{100}_{MAPE}()$	E_{MAPE_1}	E_{MAPE_2}	$E^{100}_{MAPE}()$	E_{MAPE_1}	E_{MAPE_2}
1	8.48	0.1130	0.1417	22.033	0.2138	0.2795
2	6.38	0.1195	0.1617	12.61	0.2017	0.2718
3	5791.06	0.2510	0.3245	7535.41	0.3063	0.3829

$E^{\alpha}_{MAPE}()$ function are lower for the group models than for the global models. This tendency is confirmed by the remaining error functions. Both in the case of $E_{MAPE_1}()$ and $E_{MAPE_2}()$, the error rates are lower for the group models than for the global models. For instance, in the case of the first group and its group-based prediction $E_{MAPE_1}() = 0.1130$, which is less than the corresponding value for the global prediction equal to 0.2138. This superiority of group models is observed for all the groups. However, the highest difference is observed for the two largest groups, in terms of the number of monitored consumers. In the latter case, the reduction of both $E_{MAPE_1}()$ and $E_{MAPE_2}()$ exceeds 40%, which corresponds to the results for MAE rates, reported in Section 5.1. This reduction is significantly lower for group 3, where the reduction of $E_{MAPE_1}()$ error is equal to $1 - \frac{0.2510}{0.3063} = 0.19$ of the global error rate. Also this fact confirms previous discussion and shows the positive impact of the number of monitored consumers on group prediction accuracy.

What should be emphasised is that the proposed error function $E^{\alpha}_{MAPE}()$ is based on the sum of individual error proportions, which makes it much more susceptible to individual, but very high percentage error rates. It can be observed that this impact is much lower in the case of E_{MAPE_1} and E_{MAPE_2} rates. On the other hand, by setting $\alpha < 100$ in $E^{\alpha}_{MAPE}()$ one can control the number of considered time points and observe whether high overall error rate is caused by a limited number of problematic cases or a limited prediction accuracy for the majority of cases.

To sum up, the results attained with other error functions based on Manhattan and Euclidean norms confirm the superiority of the group models created with the algorithm proposed in this study. The impact of errors made by the prediction models at individual time points on the error rates differs. The decisions on the possible use of each of the error functions can be made by an engineer in charge of heat production.

5.4. Global models with group input

Finally, one more series of experiments was performed. The objective was to verify the potential benefits of training global models with the entire avail-

Table 5. Summary of results - the hybrid approach

			MAE errors	
Group	card(C_i)	card(M_i)	$E(\xi_i)$	$E(\mu_i)$
1	363	34	0.01186	0.00961
2	359	31	0.01498	0.01206
3	136	4	0.02737	0.02426
Avg.err.	n.a.	n.a.	0.01807	0.01531

able data set, i.e. with the global data set, but performing prediction with a group data. The global models were trained as before, but when prediction was performed, the MLP-based global models were supplied with an average consumption among group members G^j_{h-6}, G^j_{h-7},G^j_{h-8},G^j_{h-24}, instead of an average consumption by all consumers A_{h-6}, A_{h-7},A_{h-8},A_{h-24}. In this way, a global model could potentially benefit from the most recent data describing group behaviour. This approach will be referred to as a hybrid approach in the remainder of the work. The results of this series of experiments are summarised in Table 5.

The results obtained in the hybrid approach are better than the results of using a global model with global data A_{h-6}, A_{h-7},A_{h-8},A_{h-24} as input. A reduction of 16% of MAE error rate is observed. Nevertheless, group models provide even lower error rates. A hybrid approach might, though, be a promising alternative to group models in the periods of relatively similar behaviour of different consumer groups.

6. Summary

Load prediction techniques for DHS have traditionally concentrated on heat sources. With the advent of low-cost monitoring systems, detailed data describing short-term consumer behaviour become available. This helps to predict heat consumption on the consumer side, which significantly differs from source-side prediction. The work proposes the way such prediction models can be constructed and used to model heat consumption by different consumers.

It has been shown that aggregated long-term time series, when used for SOM network construction, can be successfully applied to identify groups of consumers relevant for short term prediction. The groups identified in this way share similar short term demand profiles. Therefore, it has been possible to obtain more accurate predictions than those produced by global models.

Group prediction models were successfully used to answer the need for prediction at a consumer rather than at a heat source level. At the same time, the method of building consumer groups and their own prediction models has been proposed for a real data set from one of the DHS. For both group and global models absolute and percentage errors have been investigated. Particular attention has been paid to the latter category of errors in view of common hydraulic calibration processes. Unlike most studies, we analysed, no single error rate, but MAPE rates as functions of the percentage of time steps. This provides an insight into the distribution of percentage error rates calculated for individual time steps and can be used to evaluate the quality of a model.

An improved accuracy of group models when compared to global models has been obtained in spite of using the data from a lower number of consumers for every group model than in the case of a global model. What is important, little or no extra investment is required to obtain the data needed for the proposed method. Finally, consumer-side prediction provides data necessary for hydraulic calculations of DHS.

In the future, further studies and simulations aiming to investigate the results of applying the group prediction models to hydraulic modelling are planned. One of the objectives is the development of a control strategy based on dynamic modelling with different scenarios of heat source settings i.e. temperature and pressure settings. Another question to be answered is whether a global model or a hybrid approach can be selected in a self-adaptive manner to deliver in some periods even more accurate prediction than the group-based one.

References

Arctur, D., Zeiler M. (2004) *Designing Geodatabases. Case Studies in GIS Data Modelling.* ESRI Press.

Balate, J. et al. (2007) Strategy evolution of control of extensive district heating systems. *International Conference on Power Engineering, Energy and Electrical Drives, POWERENG 2007.* IEEE, New York, 678–683,

Bhave, P.R., Gupta, R. (2006) *Analysis of Water Distribution Networks.* Alpha Science, Oxford.

Davidsson, P., Wernstedt, F. (2004) *Embedded Agents for District Heating Management.* In: *Proceedings of Third International Joint Conference on Autonomous Agents and Multi-Agent Systems.* HACM, New York, 1148–1155.

Fogel, D.B. (1999) An Overview of Evolutionary Programming. In: *The IMA Volumes in Mathematics and its Applications*, **111**. Springer, New York, 89–109,

Grzenda, M., Macukow, B. (2002) Evolutionary Neural Network-Based Optimisation for Short-Term Load Forecasting. *Control and Cybernetics*, **31**, 371–382.

Grzenda, M., Macukow, B. (2006) Demand Prediction with Multi-Stage Neural Processing. In: L. Jiao et al., *Advances in Natural Computation and Data Mining.* Xidian University Press, Xi'an, China, 131–141.

Grzenda, M. (2008) Load Prediction Using Combination of Neural Networks and Simple Strategies. In: *Frontiers in Artificial Intelligence and Applications*, **173**, IOS Press, Amsterdam, 106–113.

Grzenda, M. (2009) SOM-Based Selection of Monitored Consumers for Demand Prediction. In: E. Corchado and H. Yin, eds., *IDEAL 2009*, **LNCS 5788**, Springer-Verlag, Berlin-Heidelberg, 807–814.

Grzenda, M., Macukow, B. (2009) Heat Consumption Prediction with Multiple Hybrid Models. In: S. Omatu et al., eds., *IWANN 2009, Part II.* **LNCS 5518**, Springer-Verlag, Berlin-Heidelberg, 1213-1221.

Haykin, S. (1999) *Neural Networks: a Comprehensive Foundation.* Prentice-Hall Inc.

Kashiwagi, N., Tobi, N. (1993) Heating and cooling load prediction using a neural network system. In: *Proceedings of 1993 International Joint Conference on Neural Networks, IJCNN '93-Nagoya*, **1**, IEEE, New York, 939–942.

Kato, K. et al. (2008) Heat load prediction through recurrent neural network in district heating and cooling systems. *IEEE International Conference on Systems, Man and Cybernetics.* IEEE, New York, 1401–1406.

Kusui, S., Nagai, T. (1990) An electronic integrating heat meter. *IEEE Transactions on Instrumentation and Measurement*, **39**(5), 785–789

Lane, I., Beute, N. (1996) A Model of the Domestic Hot Water Load. *IEEE Transactions on Power Systems*, **11**,4, 1850–1855.

Maguire D., Kouyoumjian V., Smith R. (2008) *The Business Benefits of GIS. An ROI Approach.* ESRI Press.

Martinetz, M., Berkovich, S., Schulten, K. (1993) "Neural-gas" Network for Vector Quantization and Its Application to Time Series Prediction. *IEEE Transactions on Neural Networks*, **4**, 558–569.

Móczar, G., Csubák, T., Várady, P. (2002) Distributed Measurement System for Heat Metering and Control. *IEEE Transactions on Instrumentation and Measurement*, **51**,4, 691–694.

Osowski, S. (2000) *Neural Networks for Information Processing* (in Polish). Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.

Park, T., C. et al. (2009) Optimization of district heating systems based on the demand forecast in the capital region. *Korean Journal of Chemical Engineering*, **26**(6), 1484–1496.

Sandou, G. et al. (2005) Predictive Control of a Complex District Heating Network. In: *44th IEEE Conference on Decision and Control, 2005 European Control Conference. CDC-ECC '05*, IEEE, New York, 7372–7377.

Sakawa, M. et al. (1999) Cooling load prediction in a district heating and cooling system through simplified robust filter and multi-layered neural network. *IEEE International Conference on Systems, Man, and Cybernetics, IEEE SMC '99 Conference Proceedings*, **3**, 995–1000.

Siwek, K., Osowski, S. (2009) Two-Stage Neural Network Approach to Precise 24-Hour Load Pattern Prediction. In: E. Corchado et al., eds., *HAIS 2009*, **LNAI 5572**, 327–335.

Siwek K. et al. (2010) Prediction of Power Consumption for Small Power Region Using Indexing Approach and Neural Network. In: K. Diamantaras, W. Duch, L.S. Iliadis, eds., *ICANN 2010*. **LNCS 6352**, Springer, 54–59.

Ye, X., Zhang, X., Diao, W. (2005) A Networked Heat Meter System for Measuring the Domestic Heat Supply. In: *IEEE International Conference on Industrial Technology ICIT 2005*, IEEE, New York, 225–230.

Youen, Z. (2009) A Generalized Adaptive Predictive Controller Design-based Direct Identification for District Heating System. *Chinese Control and Decision Conference*, IEEE, New York, 3426–3431.

Walski T.M. (2004) *Advanced Water Distribution Modeling and Management. Haested Methods*, Bentley.