Towards 4th generation district heating: Prediction of building thermal load for optimal management

Towards 4th generation district heating: prediction of building thermal load for optimal management

Elisa Guelpa[a], Ludovica Marincioni[b] and Vittorio Verda[c]

[a] Energy department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, elisa.guelpa@polito.it
[b] Energy department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy ludovica.marincioni@polito.it,
[c] Energy department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, vittorio.verda@polito.it

Abstract:
One of the requirements for the transition from conventional district heating (DH) systems to 4th generation DH (4GDH) systems is the knowledge of system dynamics. Forecast of thermal request profiles of buildings is crucial to optimize the operating conditions. In fact, when this is available, the thermal load evolution at the plants can be estimated and proper energy saving actions can be implemented. In this paper, a smart and fast approach for estimating the daily thermal request of buildings in large networks is presented. The methodology uses only data available from the smart meters installed in the building substations (mass flow and temperature data). For the users where smart meter data are not available an alternative approach is proposed. The methodology is shown to be suitable for applications involving a) a very large number of buildings b) necessity of forecast of an area c) measured data, which might be affected by gaps d) low computational time requirements. Experimental data show that, despite the simplicity, the method predicts the thermal request very accurately. Furthermore, the forecasted thermal request are here effectively used with the aim of reducing the peak load in one of the largest DH systems in Europe.

Keywords:
Thermal request forecast, building load prediction, 4 GDH, peak shaving, optimal operation.

1. Introduction

District heating (DH) is a smart technology for house heating [1, 2,3], because it enables the possibility of using heat produced in high efficiency plants, cogeneration plants [4,5], renewable energy sources [6-8] and waste heat [9-11] for house heating. DH is one of the main options [12, 13] for increasing the exploitation of renewable sources and efficient technologies for house heating and domestic hot water production in densely populated areas. Optimal control and management [18-20] is a crucial point for reaching the goals of 4th generation DH [37]: reduced primary energy consumption and pollutant production and increased economic benefit from selling heat and electricity (in the case of cogeneration).

Optimal operation of DH systems usually relies on the forecast of thermal load profiles of the buildings connected to the network. Knowledge of the total request of all the buildings connected to the network is sufficient in various applications. In the literature, many works can be found with this aim [21-23]. Nevertheless, this information is limited because: a) in extended networks the dynamics is not negligible and therefore the sum of the request of all the buildings is different from the load at the thermal plants b) various applications require knowledge of the heat requested in each area of the network or in the single buildings. Rescheduling of the buildings request (virtual storage) and installation of sensible [14, 15] or latent energy [16,17] storage systems are among these applications. Virtual storage in DH systems consists in the variation of building settings, such as the time the heating systems are switched on and off or the set point temperature. This is done to modify the thermal request evolution [24-26]. The innovation of virtual storages stands in enabling for the first time management of the demand side. Therefore the use of thermal storage and virtual storage leads to various advantages:

a) Reducing primary energy consumption, through better exploitation of the high efficiency plants, renewable energy sources or waste heat. In fact, it is possible to increase the heat supplied to buildings when there is availability from these kinds of plants.
b) Reducing pollutant emissions because of the possibility of optimally exploit smart grid potentials and renewable energy sources.
c) Shifting the thermal profile to increase the revenues for the electricity selling.

Analysis on the effect of thermal storage and virtual storage requires a prediction of the request profile within the day for all the buildings (or each area distribution network). Two type of models can be applied [27]: physical models and black box models. The first kind of approach includes advanced models that use physical principles to calculate thermal dynamics and energy behaviour of the building (such as Energy Plus or DOE-2) [28-30]. This approach provides detailed information but requires high computational resources and therefore it is suitable for simulation of a limited number of buildings. Furthermore, a complete description of the structure and several data of the buildings shall be available. Black box (such as neural networks) models [31-33] are simpler because they do not simulate the physic behavior of the building and require low computational costs but the information they provide is less detailed. Data for calibration are necessary.
However, there are some criticism in using conventional forecast methods for thermal storage and virtual storage applications:

a) When extended DH network are considered, very few data are often available. These data are the ones collected by the smart meters installed in the substations (temperature, mass flow rates, and energy consumption), the number of operating hours and the building volumes. For this purpose, a model should use only these data.
b) The evolution of the quantities collected by the smart meters are often affected by problems, such as lack of data for difficulties in data measurement or transmission, or the presence of incorrect or missing values. Robust models, able to predict the thermal demand despite the small amount of data are necessary.
c) The prediction method strongly depends on the usage of the evolution forecast. For instance, the optimal re-scheduling of the users connected to the DH system (virtual storage) requires:
    a. accurate short-term forecasts of the users demand evolution;
    b. accurate prediction of the peak time;

c. prediction of both the thermal request and the water mass flow rate, in order to evaluate network dynamics.

Based on these necessities, a fast and robust approach for the thermal request profile forecast is presented in this work. The approach has been included in a software that automatically predicts the evolution of the thermal request of the various distribution networks.. The new idea of the proposed methodology is to define the main characteristics of the thermal profiles and to predict these quantities. These are: i) the peak height ii) the peak amplitude and iii) the request during steady state conditions. Various factors influencing the thermal profile evolution are considered and a correlation analysis show which are the ones that must be kept as model inputs. The model used for the thermal load prediction is linear and it is based on a black box approach. The methodology also includes a simplified approach for the prediction of thermal profiles of buildings that do not have available smart meter data.

The software has been tested on six distribution networks of the Turin DH system, which is one of the largest in Europe, in order to show its potential for virtual storage usage. The outcomes prove that the software is able to well predict the thermal request also when criticisms occur. The building load evolution is here applied to a virtual storage optimizer with the main aim of rescheduling the building heating systems to shave the total peak load.

The paper is organized as follows:

- Section 2 describes the methodology adopted for the profile forecast (for both the buildings where smart meter data are or are not available), including the description of the main characteristics of the thermal profiles and the pre-processing analysis of the model inputs.
- Section 3 outlines the software that has been created to automatically perform the methodology described in section 2.
- Section 4 reports the DH network where the software has been applied to and describes the model for virtual storage analysis.
- Section 5 includes all the outcomes of the methodology validation.

2. Methodology for thermal profile forecast

2.1 Main characteristics of the thermal load profile

Concerning the prediction model, a new approach has been selected in this paper. The approach is based in the idea of identifying the main points of the curve and predict the position of these points. The main curve characteristics are used to move from the thermal profile to a simplified profile, as depicted in Fig. 1. This approach has been selected for two reasons:

1. It allows reducing the errors due to data gathering and transmission, such as missing values or wrong values.
2. It allows to predict the complete thermal evolution, through prediction of a limited number of quantities. It makes the model reliable and robust.

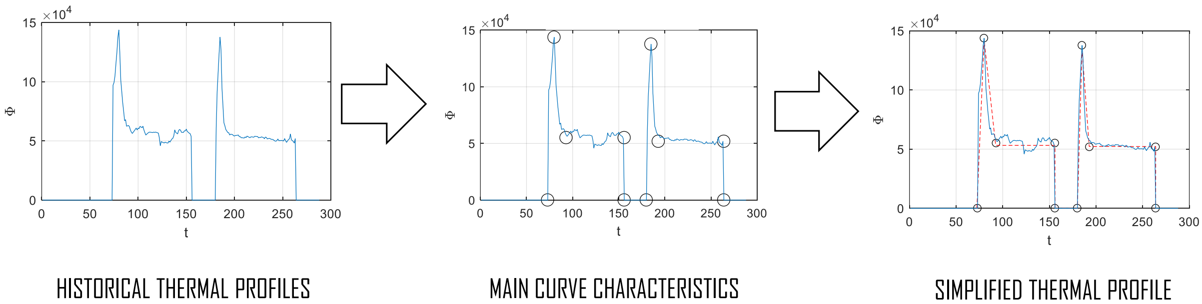


Fig. 1 Simplification of the thermal profile evolution: output definition

The main characteristics of the thermal request are evaluated as the function of the weather forecast. As discussed in the introduction section, the aim of predicting the thermal load is the analysis of virtual storage. Therefore, it is crucial to correctly detect the peak height, the peak amplitude and the steady state request. For detecting correctly these features, the thermal profile is described through the quantities shown in Fig. 2. Evaluation of the four quantities has to be performed in an automatic way due to the large historical dataset (the analysis is performed for building each day for various years):

1. The maximum height of the peak (y_1). It occurs in a building each time the heating system is switched on (item 1 in Fig. 2). This is evaluated as the largest thermal request in each operational period, i.e. between the time the heating system is switched on and off.
2. The steady-state heat request (y_2), which occurs after the end of the peak (item 2 in Fig. 2). This is calculated as the mean value of the heat request between the end point of the peak and the time the heating system is switched off.
3. The time the system needs for reaching the maximum point of the peak, after the heating system is switched on, y_3 (item 3 in Fig. 2). This is evaluated as the time frame between the two events.
4. The time the system needs for reaching the end point of the peak from the maximum point of the peak y_4 (item 4 in Fig. 2). This is evaluated as the time frame between the two events.

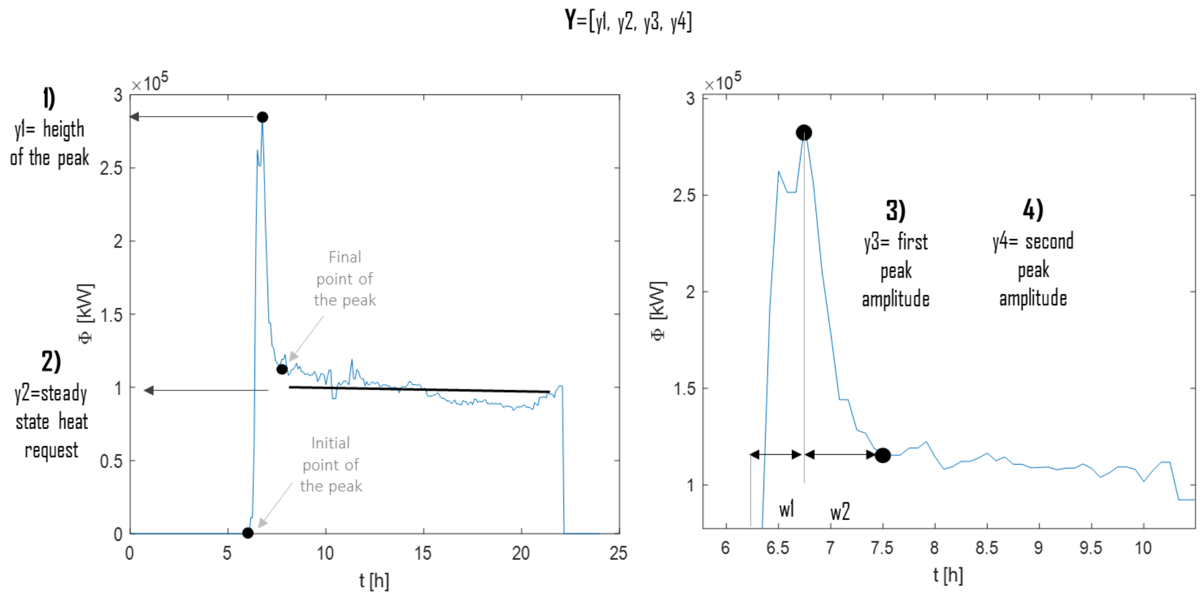


Fig. 2 Quantities detected for the thermal profile description

The evaluation of the final point of the peak requires a careful analysis, because it cannot be obtained by basic mathematical calculations. This is performed through a suitable algorithm. A routine that takes into account all the times after the ones when the maximum peak value occurs is used. The time that is selected as the end of the peak, is the first one complaining with one of the following options:

- The value of the curve in the point is lower than the steady state value.
- The slope of the curve has a value which is, in absolute value, lower than a threshold value.

Some of the heating systems are switched-on and off just once a day. In other cases, there are more than one operating periods (time periods between two starts) each day. In the DH network selected as the test case the operating periods are from one to three per days, as reported in section 4. Each start usually leads to the creation of a peak which is then followed by an almost stationary condition. The software can automatically evaluates the four quantities indicated in Fig. 1 (maximum peak value, the steady state and the two time periods) for each operational periods. Indeed, when a building that presents more than a start per the day, the same approach is used various times.

2.2 Input definition

A two-stage approach is adopted for the identification of the model input. At first, a series of variables affecting the building thermal request are considered. These variables are selected on the basis of a literature review [27] and considering that simulations are performed the day before. Tab. 1 reports the full list of variables at step 1. The temperatures referred to the previous day affect the right behaviour of the building and therefore the thermal demand when the heating system is switched on in the morning.

Description	variable	unit
Mean temperature of the previous day	$T_{m,d-1}$	K
Minimum temperature of the previous day	$T_{min,d.1}$	K
Maximum temperature of the previous day	$T_{max,d-1}$	K
Mean temperature of the day to predict	$T_{m,d}$	K
Solar radiation in the day to predict	I	W/m
Air humidity in the day to predict	φ	-
A wind intensity in the day to predict	v	m/s

Table 1 Main initial input variables

As a second step, a sensitivity analysis is conducted to evaluate how much these quantities influence the main curve characteristics (depicted in Fig. 1). This is conducted using a correlation analysis. Variables that do not significantly influence the main curve characteristics are removed in order to achieve a) a greater model simplicity b) an easier collection of the input data. Table 2 shows the influence of the various variables on the thermal demand at peak and steady state and on the peak amplitude.

	$T_{m,d-1}$	$T_{m,d}$	$T_{min,d-1}$	$T_{max,d-1}$	I	φ	v
Peak height	Medium	Low	Low	Medium	Very low	Very low	Very low
Peak amplitude	**High**	**High**	**High**	**High**	Low	Very low	Very low
Steady state conditions	**High**	**High**	**High**	**High**	Low	Very low	Very low

Table 2. Effects of the input parameters on the curve characteristics

The quantities which mostly affect load evolution are the four temperatures. Solar radiation *I*, humidity *φ* and wind intensity *v* are not taken into account for the model. This assumption is strictly dependent on the site and on the characteristics of the heating systems. As an example, in Turin the average wind velocity is very small and for this reason its effect is small. In other sites, this effect might be much larger. The selected input of the model are:

- the mean temperature in the previous day, $T_{m,d-1}$;
- the minimum temperature in the previous day, $T_{min,d-1}$;
- the maximum temperature in the previous day, $T_{max,d-1}$;
- the mean temperature (forecast) in the current day, $T_{m,d}$.

2.3 Description of the models

The model described in the first part of this section is called MODEL A while MODEL B deals with buildings where smart meter data are not available. MODEL B is described in section 2.3.1. In Fig. 3 both models are summarized together with the approach which has been used to build them.

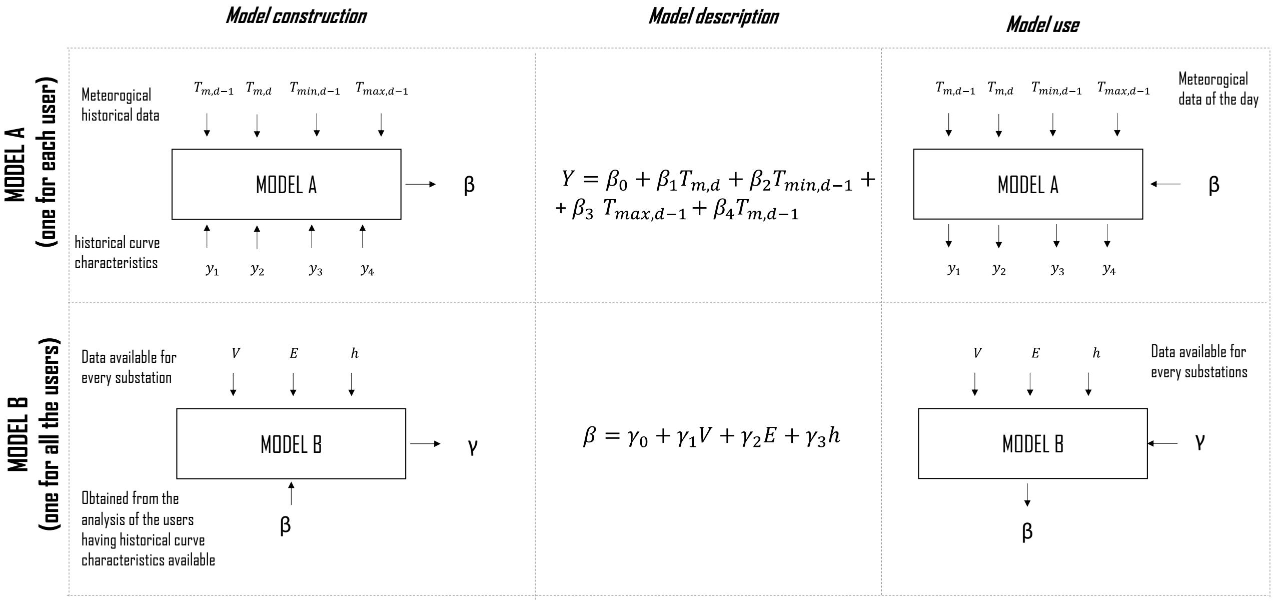


Fig. 3 MODEL A and MODEL b description

2.3.1 MODEL A (smart meter data available)

MODEL A is a linear model described by equation 1:

$$\boldsymbol{Y} = \boldsymbol{\beta} \cdot \mathbf{T} \tag{1}$$

where the unknown vector **Y** for each building includes the main curve characteristics previously described and has the following form:

$$\boldsymbol{Y} = [y_1, y_2, y_3, y_4] \tag{2}$$

and the vector of input parameter has the following form:

$$\boldsymbol{T} = [1, T_{m,d}, T_{min,\,d-1}, T_{max,\,d-1}, T_{m,\,d-1}] \tag{3}$$

The model is used, as shown in top left of Fig. 3, to evaluate vector **Y** once the meteorological conditions are available.

The model A is built for each building, therefore the coefficients **β** are evaluated in our application for about 6500 buildings. The model coefficients **β** are obtained using the historical data of the curve characteristics Y_{hist}, for the temperature conditions $\mathbf{T}_{\boldsymbol{hist}}$ following equation (3), as depicted in top right of Fig.3.

$$\boldsymbol{Y_{hist}} = \boldsymbol{\beta} \cdot \mathbf{T}_{\boldsymbol{hist}} \tag{4}$$

2.3.2 MODEL B (smart meter data unavailable)

In large DH some of the sensors may fail for long time and some of the buildings might not be equipped with a smart meter. A second model is proposed to overcome these issues. In these cases, it is possible to exploit the data available for all the users: the building volume *V*, the annual energy demand for unit volume *E* and the total annual operating hours for unit volume *h*. A correlation between these three variables and the coefficients β is first obtained. The data of the buildings equipped with smart meters are used for this purpose.

MODEL B, is described by the following equation:

$$\boldsymbol{\beta} = \boldsymbol{\gamma} \cdot \mathbf{D} \tag{5}$$

Where the input vector D is defined as follows:

$$\boldsymbol{D} = [1, V, E, h] \tag{6}$$

The buildings where only V, E and h are available are analysed by using eq. (5), as indicated in bottom right of Fig. 3, to obtain the **β** coefficients. Indeed, MODEL B allows evaluating the coefficients β only trough knowledge of V, E, h. The model B is only one for all the users, therefore just a γ is used for all the users. The coefficient matrix **γ** is built by using data of the buildings where smart meter data are available, as indicated in bottom left of Fig. 3.

3. Software infrastructure

The software infrastructure built for the thermal load forecast, including both MODEL A and MODEL B, is shown in Fig. 4. It consists of two parts:

- The first part, represented in the left part of the graph, includes the automatic evaluation of the coefficients for the linear model. It includes:
    - A system for the automatic loading of the historical data. Historical data includes a file for each day and for each quantity analysed. For the network analysed about 50 10^6 historical files are used.
    - An automatic pre-processing of the data to make them suitable for the usage. In this stage all the data are reorganized in various matrices, one for each quantity analysed and one for each distribution network. The matrices have as many column as the number of building located in the distribution network and as many rows as the time considered for the analysis.
    - The evaluation of the number of historical data available for each user;
    - The automatic estimation of the main curve characteristics for all the historical data (the first step in Fig. 1). This step needs various precautions because of the measurement errors that sometime makes data evolution very different to the expected evolution.
    - Evaluation of the coefficient of the linear model $\boldsymbol{\beta}$ (one for each user) as detailed in paragraph 2.2. Coefficients $\boldsymbol{\beta}$ are evaluated only for the users for which smart meter data are available.
    - Evaluation of the coefficient of the linear model $\boldsymbol{\gamma}$ (one for all the users), by using coefficients $\boldsymbol{\beta}$ and V, E and h of the users with smart meter data available. The model is built as detailed in paragraph 2.3 to tackle absence of smart meter data for some users.

  If the buildings request did not change, this part of the software would be run just once, for the creation of the model coefficients for all the buildings. Actually, there are several reasons because the thermal requests change, such as the application of retrofitting measures, the change in heating systems schedule, modification of the occupant habits. This part of the model is thus run periodically (e.g. once a month) in order to take into account all these occurrences.
- The coefficients $\boldsymbol{\beta}$ (MODEL A) and $\boldsymbol{\gamma}$ (MODEL B) are used within the second component of the software for evaluating the main characteristics of the curve $\boldsymbol{Y}$ (described in Fig. 1). MODEL B is run only for the users for which no smart meter data are available. Once these values are estimated, these are used for obtaining the thermal request profiles (second step in Fig. 1)

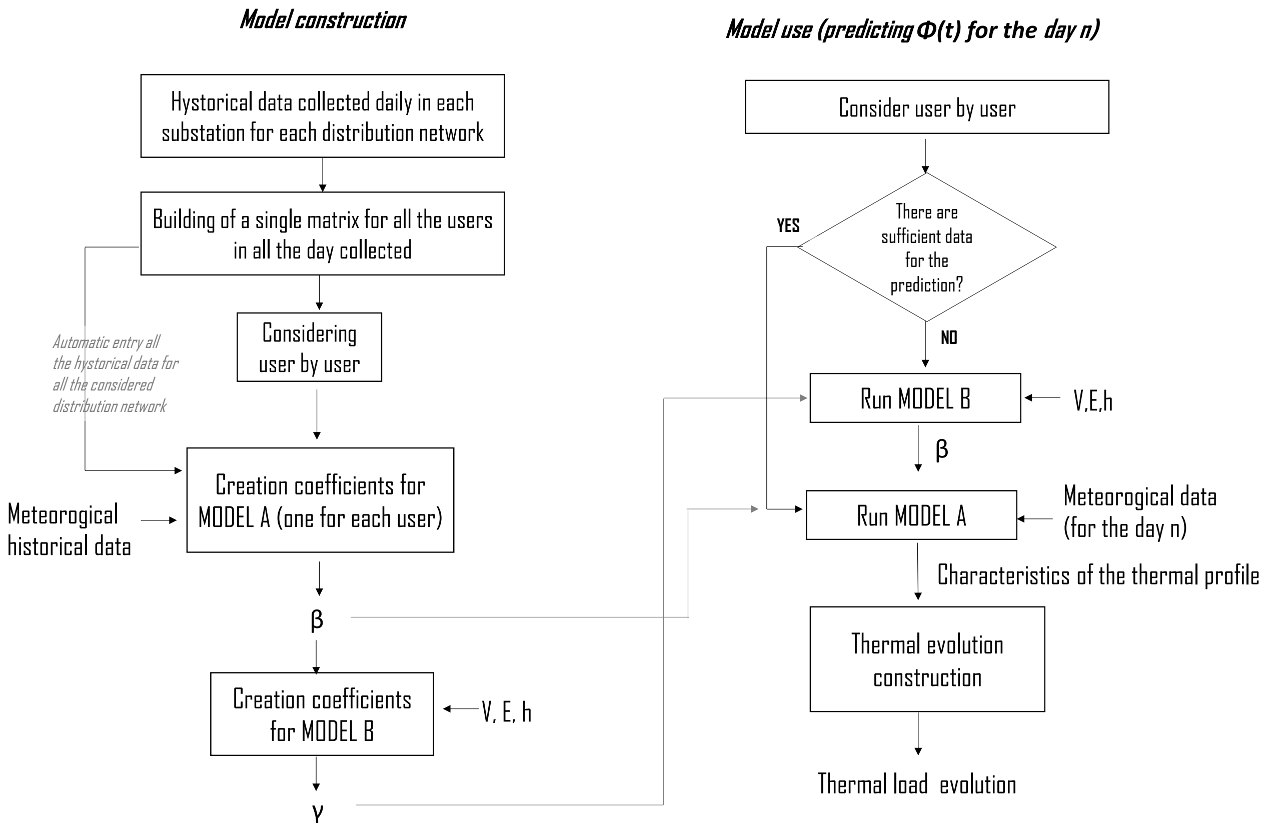


Figure 4. Infrastructure of the tool (Construction of the model and use of the model for the prediction)

4. Application

4.1 System description

The software is applied for the prediction of the thermal profiles of the buildings connected to the Turin DH system. Turin DH network is one of the largest in Europe; indeed it includes more than 6.500 substations, one for each building. The main transport network links the thermal plants to the various distribution networks (182), which connect the transport network to the single buildings. Various pumping groups, located in different areas of the network, pump the water along the network. Five plants are used for feeding the networks, including cogeneration plants, boilers and storage tanks. For further details on the analysed system the reader can refer to [35].

The extension of the network makes it necessary the use of automatic system for the evaluation of the thermal profiles. As mentioned before, in cases where DH networks involves a high number of users, robustness is an important characteristic for the forecast tool. The evaluation of the coefficients **β** for obtaining the thermal profiles is performed using data gathered at the user substations, as detailed in paragraph 2. The DH substations are equipped with a system for controlling and gathering thermodynamic quantities able to transmit measurements to a centralized unit. The detected quantities are reported in Table 3

Quantity	Location	Symbol	Unit
thermal power	substation heat exchanger	ϕ	kW
mass flow rate	primary side	G	kg/s
temperature	inlet section of the primary side	T1	°C
Temperature	outlet section of the primary side	T2	°C
Temperature	outlet section of the secondary side	T3	°C
Temperature	inlet section of the secondary side	T4	°C

Table 3. Main data gathered in the substation

The evolution of some of the data gathered in a distribution network are depicted in Fig.5. From mass flow rates (G) and thermal power (ϕ) evolution it is possible to catch the schedule of each building. Most of the users are switched-off during the night and they are switched on between 5 a.m. and 6 a.m. When the heating systems are switched-on, the mass flow rates (and consequently the thermal powers) present a peak, due to the low temperature of the fluid on the secondary side of the heat exchanger. The number of shutdown of the systems is different in the various buildings (one, two or three times a day). This is a very important point for the thermal load prediction.

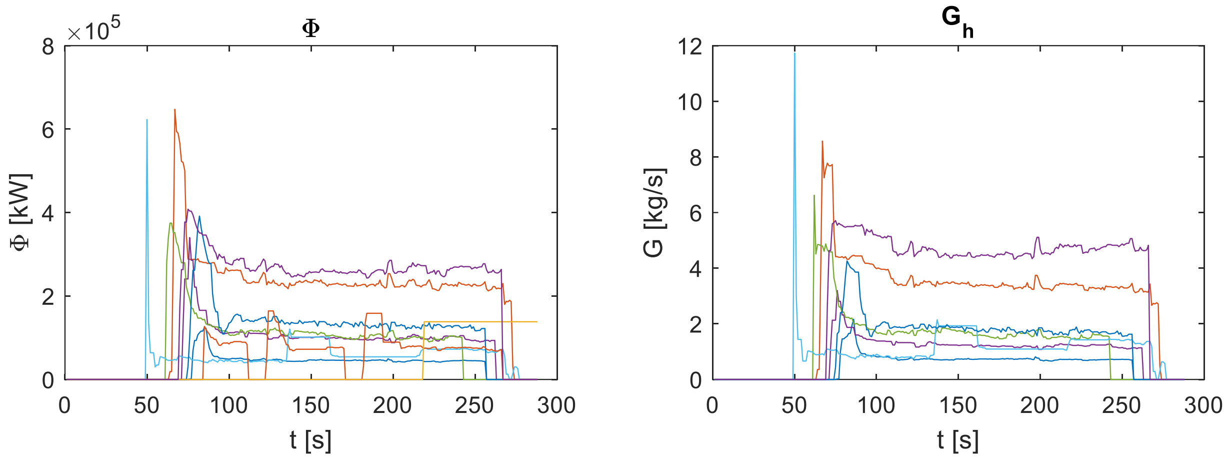


Figure 5. Daily data gathered in the substation

4.2 Forecast application: virtual storage

The thermal profile forecasted by the software are used with the aim of thermal peak shaving by virtual storage. A peak shaving optimizer has been used. The optimizer allows one finding the best set of schedules of substations for minimizing the thermal peak without affecting the comfort conditions.

The virtual storage optimizer includes a fluid-dynamic model of the network in order to take into account for the effects of the long distances involved in the network on temperature distribution. In fact, water exiting the heat exchangers in the substations mixes with the various streams coming from the users located in the other areas. These streams are at different temperatures, due to the different effect of the thermal losses that are affected by the distance of the users and the different effectiveness of the heat exchangers. As a result, the temperature evolution at the plants is significantly different from that at the users. The overall effect of the building thermal profile is evaluated trough the thermal fluid-dynamic model because it significantly differs from the summation of the thermal request profiles of the buildings. The thermal fluid-dynamic model is based on a one-dimensional approach coupled with graph theory [36]. Each pipe of the network is considered as a branch starting from a node, the inlet section, and ending in another node, the outlet section. The incidence matrix **A** is used to describe the network topology by expressing the connections between nodes and branches. The thermal fluid-dynamic model considers the mass conservation equations applied to all the nodes (equation 7), the momentum conservation equations applied to all the branches (equation 9) and the energy conservation equations applied to all the nodes (equation 8). The thermal model is expressed in transient form since thermal perturbations travel the network at the water velocity, which is the order of few meters per second while the fluid dynamic model is expressed in steady state form because the perturbation travel the entire network at the speed of sound.

$$\mathbf{A} \cdot \mathbf{G} + \mathbf{G}_{\mathrm{ext}} = 0 \tag{7}$$

$$\mathbf{G} = \mathbf{Y} \cdot \mathbf{A}^{\mathrm{T}} \cdot \mathbf{P} + \mathbf{Y} \cdot \mathbf{t} \tag{8}$$

$$\boldsymbol{M} \cdot \boldsymbol{T} + \boldsymbol{K} \cdot \boldsymbol{T} = \boldsymbol{g} \tag{9}$$

For further details on the model the interested reader can refer to [37, 38]. The models A and B described in section 2 are used for evaluating both the mass flow rates and the thermal power. The temperature and mass flow rates at the outlet section of the heat exchanger at the primary side (T2 and G) can be thus provided as the input to the thermofluid dynamic model. G is predicted by the forecast model and T2 is obtained trough the relation:

$$\Phi = Gc(T_1 - T_2)$$

where G and Φ are the outcomes of the forecast model while T1 is obtained simulating the supply network considering a supply temperature is known.

5. Results

5.1 Evaluation of the main curve characteristics

After the automatic loading and pre-processing of the historical data, the following step is the evaluation of the main characteristics of the thermal request. This step must be automatic because it includes the analysis of all the buildings connected to the DH network (about 6500) for all the historical data considered (more than 700 days). The outcome of the automatic detector of the main curve characteristics are reported in Fig. 6 for some buildings. The buildings and the days used in the analysis are selected among the available data in order to cover different schedules and weather conditions. The experimental data detected by the smart meter are reported in black, while the points estimated by the automatic detector are depicted with red circles. Fig. 3 clearly shows that the tool is perfectly able to detect all the quantities desired for the construction of the simplified thermal profile.

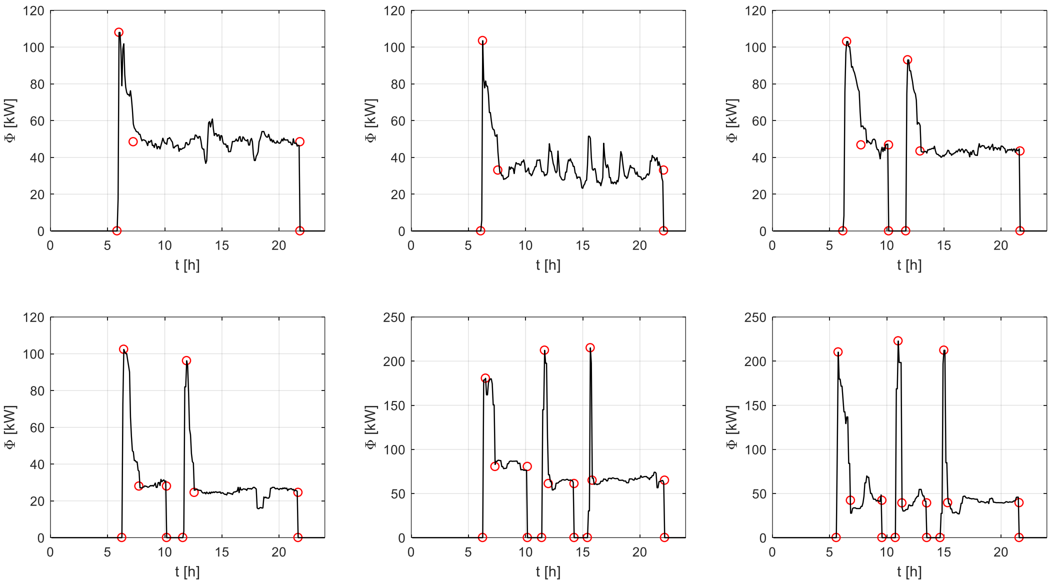


Fig. 6. Results of the tool for the automatic detection of the main curve characteristics

5.2 Thermal profiles forecast

The main curve characteristics of the historical data are used for obtaining the **β** vectors (one for each user) and **γ** (one for all the users). The models are thus used for obtaining the main curve characteristics (and therefore the simplified thermal profiles). The results are compared with experimental data for some days and some buildings and they are shown in Fig. 7. Thermal profiles and mass flow profiles are both reported. The figure depicts these measured evolutions (dashed curves) and the evolutions obtained by the linear model (continuous curves). The criteria adopted to select the buildings and the days are the same considered in the previous section. The thermal request and the mass flow are well detected in all the random cases reported in Fig. 7. For both the quantities, the maximum peak values, the steady state conditions and the peak duration are predicted with a high level of accuracy (about 10%).

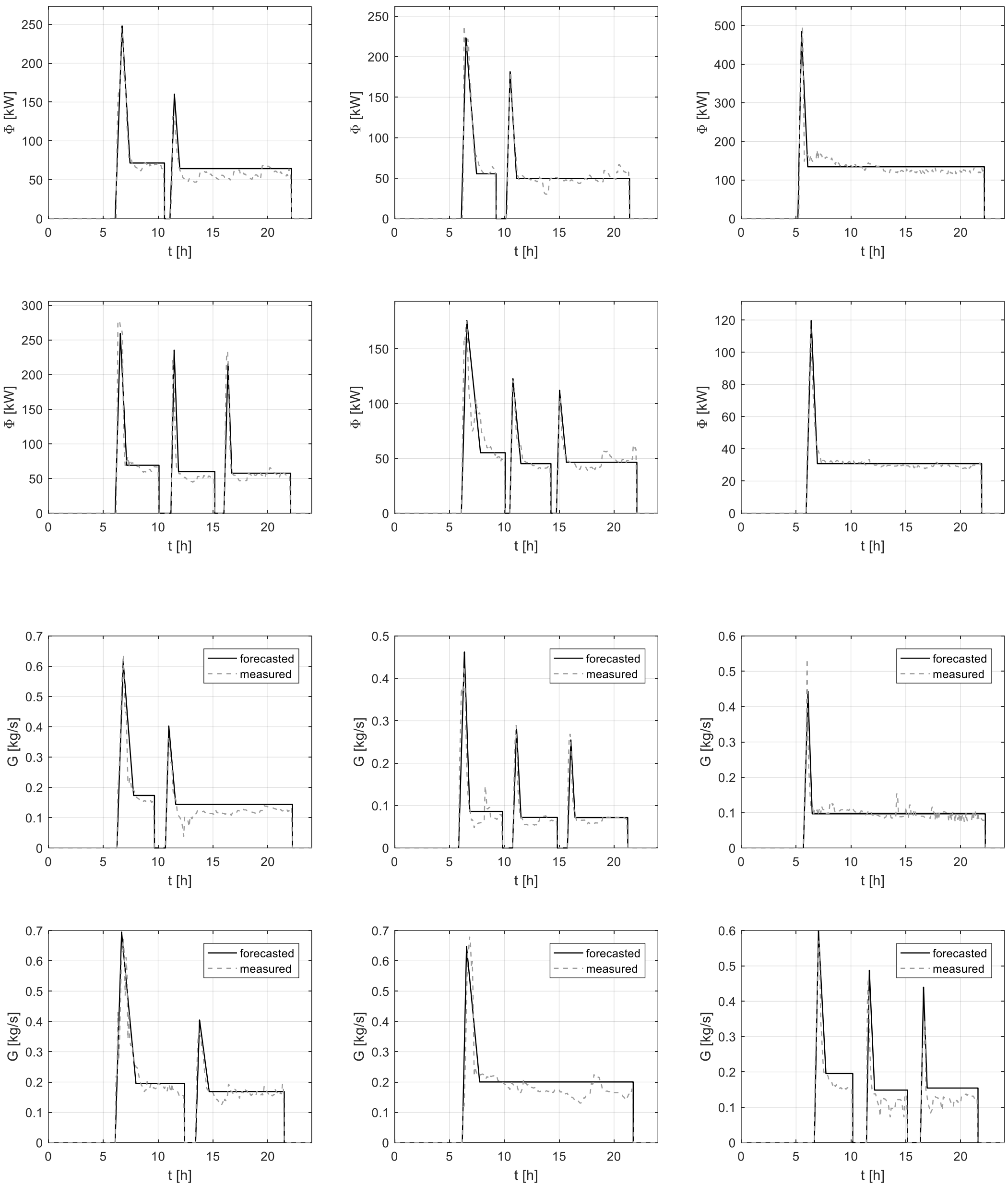


Fig. 7. Predicted thermal request and mass flow profiles predicted (comparison of model results and measured data)

Taking into account all the buildings and all the days considered in the analysis it is possible to perform a more complete error analysis. The mean relative error in the prediction is evaluated as the mean error in the evaluation of the main curve characteristics (comparing predicted and measured values). The percentage relative error on the maximum peak value is evaluated for each considered case; the mean value is about 10 %. As regards the relative error on the steady state prediction the result obtained is higher (although less than 15%) than the ones performed predicting the maximum peak value. This mainly occurs at the beginning and in the end of the season, the steady state values are quite low and even the absolute error is low the relative error is large. These results can be considered satisfying because of:

- the high variability of the thermal request characteristics along the year due to the dependence on several variables which are difficult to include (user behaviour, cloud coverage, etc.);
- the problems related to the detection of data (for instance the eventuality of the lack of a data during the peak request)
- the imprecision of the weather forecast and weather historical data;

- the simplicity of the model used for the profile forecast which is necessary for fastness and robustness reasons.

In order to quantify the effects of the errors on the total demand, the total request of some distribution networks, for a typical winter day, are analysed. The outcome of the analysis are reported in Fig. 8, in terms of thermal power and mass flow rate. It is evident that the model perfectly predicts the thermal evolution and the mass flow rates in most of the cases. The errors in the prediction of the overall network request, they are, on average, lower than the error performed predicting the single building requests. This is because when all the users are considered the various errors performed on the single buildings offset each other, because of their different signs.

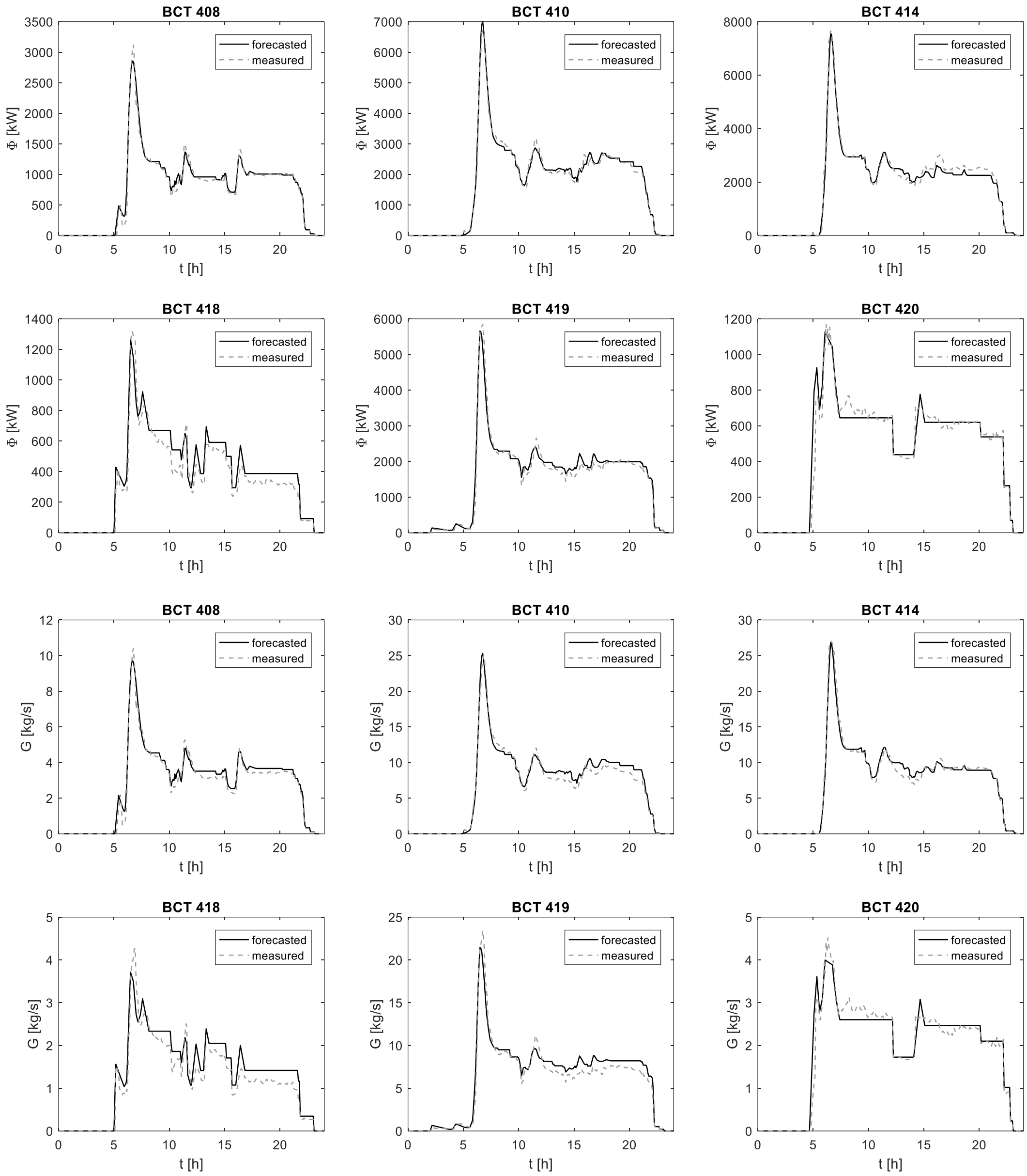
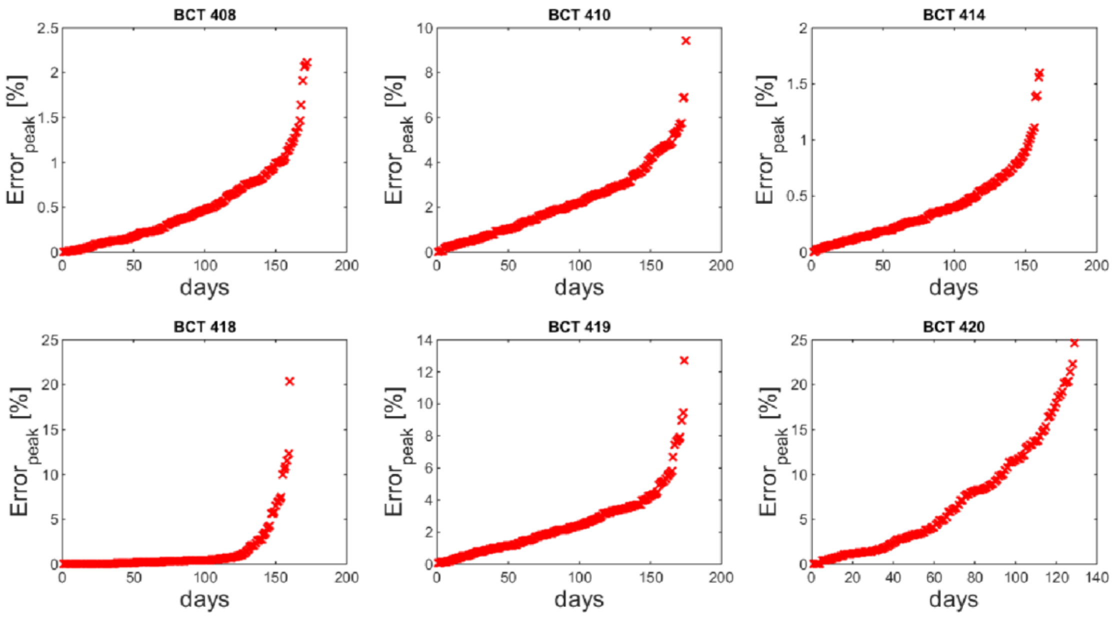


Fig. 8. Thermal request and mass flow profiles predicted for the whole distribution networks (comparison model results and measured data)

Performance of the profile predictor depends on the set of input. In order to evaluate the effects of the input variation in the error performed on the prediction, the model has been tested on a heating season (about 180 days). This is due in order to take into account the high variability of conditions that can occur, such as particularly hot or cold days, or strong temperature variation between two consecutive days or between night and day. Results have been reported in Fig.9. It shows the percentage relative error in both peak value and steady state value predictions. Error is evaluated as the difference between the predicted and the occurred value, divided by the maximum value occurred during the year. This means that errors are weighted respectively on the maximum peak value (in case of peak prediction) and the maximum steady state value (in case of steady state prediction) occurring during the year. This is made to avoid mismatches between the error referred to cold and warm days (characterized by a low thermal request). A frequency plot has been used in order to show how frequently the errors are performed. The figure shows that in all the considered cases the error keeps lower than 5 % for three-quarters of the considered day (except for BCT 420). In two or the distribution network considered (BCT408 and BCT414), the error never reaches 5%. Both peak and steady state request of BCT410 and BCT419 are also well predicted and the errors keep lower than 10% in the whole thermal season. Results obtained for BCT408, BCT410, BCT414 and BCT419 are remarkable. Higher values are obtained in some days for BCT418 and BCT420. This is mainly due to the low thermal demand (as can be seen by the Fig. 8 comparing the request for these networks and the others) and the small number of buildings involved in these distribution networks (respectively 19 and 11). Networks including a low number of buildings usually include one or few large buildings (such as university or hospital) or multiple buildings served by the same substation. For instance, BCT418 serves the Politecnico di Torino, demand is about 95% of the full demand of the network. When an error is made on the large building it automatically affects the whole network. This does not constitute an issue while applying a reschedule of the time heating systems are switched on and off. In fact, when a small number of users are involved or a single building absorbs most of the thermal power of a distribution network, virtual storage is not a good strategy for reducing the peak of the distribution network.



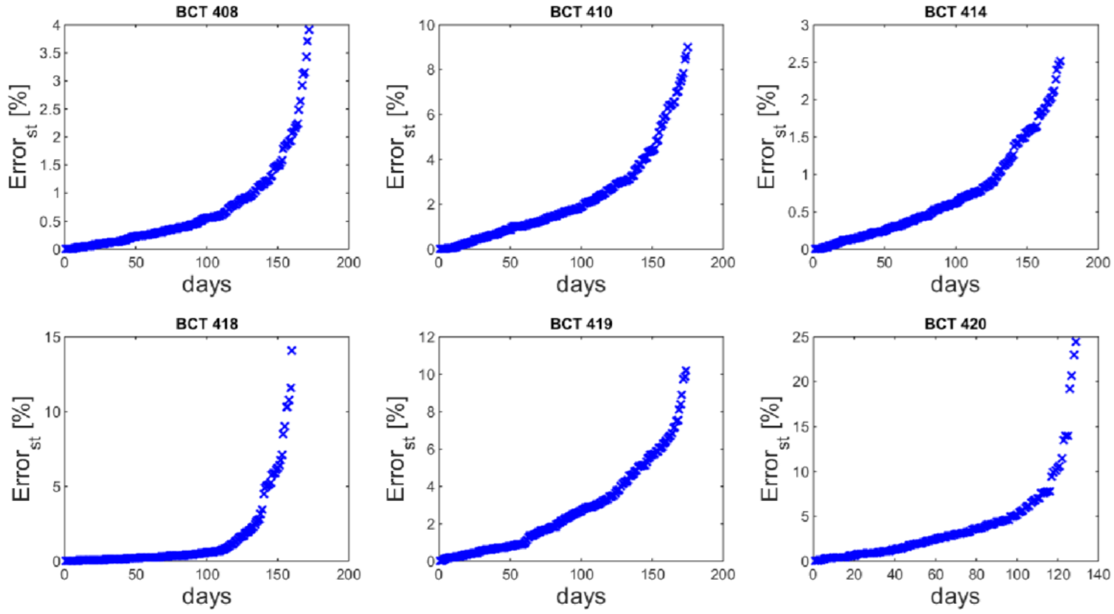


Fig. 9. Error performed on the thermal request for the prediction of the peak condition (in red) and the steady state condition (in blue)

Results prove that the model possess all the characteristics for the particular application:

- it is robust since no meaningful results are obtained;
- it is able to correctly predict the building request both in terms of thermal power and mass flow, with a acceptable error;
- it perfectly detects the overall request of the distribution networks (which is the main aim in this work);
- it takes very small time to provide results, during both the model building and the model use. The model for the power evolution prediction, by using the β coefficients ("model use" in Fig. 4), takes one tenth of a second for a distribution network, including about one hundred buildings. The part of the software performing the evaluation of the β coefficients, which includes the data loading, the data pre-processing, the automatic evaluation of the main curve characteristics, requires few seconds for a distribution network, including about one hundred buildings. These results related to the computational cost is crucial because of the large number of distribution networks included in an extended DH system (for this case study these are 182).

5. 3 Virtual storage application

The thermal profile forecast is used to produce the input for the virtual storage optimization. In this section the effects of the virtual storage are analysed in case of perfect prediction of the thermal request and in case of forecast performed by the software proposed. In particular, the following three cases are compared:

1. Current request.
2. Request obtained with the virtual storage optimizer when the thermal profiles are perfectly predicted.
3. Request obtained with the virtual storage optimizer when the results obtained with the predicted profiles are considered.

Comparison between 1) and 2) provides the ideal peak reductions that can be obtained, while comparison of 1) and 3) provides the expected peak reductions. Results show that both the perfect and the actual prediction produce an evident peak reduction. Six distribution networks have been examined and the outcome of the analysis are reported in Fig. 10.

Nevertheless, for the aim of this work the important point is that, comparing 2) and 3) small differences are observed on the thermal peak shaving. It is possible to see that the variation in the peak reduction in case of perfect and actual prediction are not significant On average, the peak reduction in case of perfect prediction is 1.8 MW, while it is 1.6 MW in case of prediction with the proposed result. This is a significant result, because it means that the incorrect evolution prediction affects the peak shaving of less than 10%.

The proposed application has two main consequences:

1. It makes the connection of additional buildings to the DH network possible, without the installation of storage units or new pipelines. In the six distribution networks, about an additional volume of buildings of 3.5 10^4 m^3 can be connected. Assuming a primary energy cost of about 1.25 for independent heating about 1% of primary energy reduction can be achieved.
2. It produces a better distribution of the daily thermal load at the thermal plants. This allows an increase in the exploitation of cogeneration plants, with a consequent reduction of primary energy consumption.

These improvements would not have been pursued without knowledge of the building request profiles.

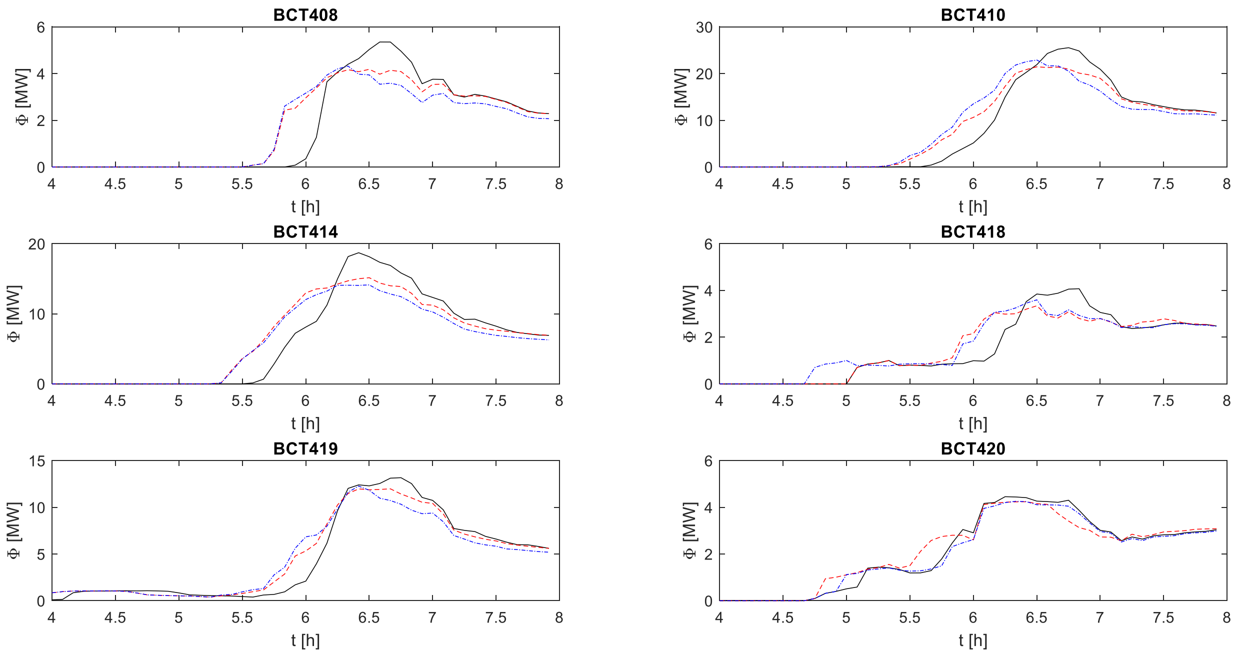


Fig. 10. Application of the virtual storage: current request (continuous line), optimal request with perfect prediction (dashed line), optimal request with forecasted prediction (dashed point line)

6. Conclusions

In this paper, a tool for predicting the thermal request evolution of the buildings connected to a DH network is presented. The methodology developed is goal oriented because it has been conceived for being applied for thermal peak shaving through virtual storage. Virtual storage consists in the variation of the thermal request profiles of the buildings connected to a DH system, acting on the scheduling of the heating systems of each building. The virtual storage model includes a network physical model to take into account the effects of the different distances of the buildings in the total request of the network.

The prediction model is linear, based on a black box approach built through the data gathered through the smart meters installed in the substations. The main idea of the prediction methodology proposed, is to identify of a series of curve characteristics to simplify the thermal profiles. These lead to a simplification of the prediction model because only the peak height, the peak amplitude and the request during steady state conditions are used to identify the request. An analysis for evaluating the inputs that mostly affect the output is performed through a correlation analysis. This allows to minimize the input number.

A simplified model for the buildings where smart meter data are not available, is proposed. Data which are available for all the buildings are used: the building volume, the annual energy request E and the total annual operating hours, h.

A software has been built. It includes: a) the data loading and pre-processing stages b) the automatic evaluator of the curve characteristics c) the model for the forecast d) the simplified model for the buildings without smart

meter data. The software has been applied to some distribution networks of the Turin district heating system, in order to test the capability of predicting the thermal profiles.
The outcomes of the analysis show that:

1. the automatic evaluation of the main curve characteristics allows one to accurately evaluate the desired quantities;
2. the linear model allows an acceptable evaluation of the various daily evolutions;
3. the overall request of the distribution network is predicted with a low error (lower than 5%);
4. the computational costs are very low for both the stage of coefficients evaluation and model use.

The predictor is used together with the virtual storage simulator for the minimization of the peak of some distribution networks in the Turin DH system. The analysis is performed in order to test the capability of the model to be used for the peak shaving applications. Results prove that the model do not strongly impact the thermal peak shaving. On the average the deviation with a perfect prediction is of the order of 10%.

As a conclusion it has been shown that the methodology and the software proposed constitute a very compact, easy to manage and robust tool to forecast the load evolution building of buildings connected to a DH system, when:

- a very large number of buildings are examined;
- the goal is the evaluation of the thermal load evolution of an area (such as applications for thermal peak shaving);
- measured data (affected by data gap) are used for the demand prediction;
- low computational costs are required.

7. References

[1] Lund H., Moller B., Mathiesen B.V., Dyrelund A, The role of district heating in future renewable energy systems. Energy 2010; 35: 1381–1390.

[2] Verda, V., Guelpa, E., Kona, A., & Russo, S. L. (2012). Reduction of primary energy needs in urban areas trough optimal planning of district heating and heat pump installations. Energy, 48(1), 40-46.

[3] Münster, M., Morthorst, P. E., Larsen, H. V., Bregnbæk, L., Werling, J., Lindboe, H. H., & Ravn, H. (2012). The role of district heating in the future Danish energy system. Energy, 48(1), 47-55.

[4] Casisi, M., Pinamonti, P., & Reini, M. (2009). Optimal lay-out and operation of combined heat & power (CHP) distributed generation systems. Energy, 34(12), 2175-2183.

[5] Ziębik, A., & Gładysz, P. (2012). Optimal coefficient of the share of cogeneration in district heating systems. Energy, 45(1), 220-227.

[6] Fahlén, E., & Ahlgren, E. O. (2009). Assessment of integration of different biomass gasification alternatives in a district-heating system. Energy, 34(12), 2184-2195.

[7] Lindenberger, D., Bruckner, T., Groscurth, H. M., & Kümmel, R. (2000). Optimization of solar district heating systems: seasonal storage, heat pumps, and cogeneration. Energy, 25(7), 591-608.

[8] Yildirim, N., Toksoy, M., & Gokcen, G. (2010). Piping network design of geothermal district heating systems: Case study for a university campus. Energy, 35(8), 3256-3262.

[9] Fang, H., Xia, J., Zhu, K., Su, Y., & Jiang, Y. (2013). Industrial waste heat utilization for low temperature district heating. Energy policy, 62, 236-246.

[10] Fang, H., Xia, J., & Jiang, Y. (2015). Key issues and solutions in a district heating system using low-grade industrial waste heat. Energy, 86, 589-602.

[11] Coss S, Guelpa E, Rebillard C, Verda V, Le-Corre O. Industrial waste heat integration for providing energy service to district heating networks. Proceedings of ECOS 2016. Portoroz, Slovenia, June 19-23 June 2016.

[12] Li, Y., Fu, L., Zhang, S., & Zhao, X. (2011). A new type of district heating system based on distributed absorption heat pumps. Energy, 36(7), 4570-4576.

[13] Sciacovelli, A., Guelpa, E., & Verda, V. (2014). Multi-scale modeling of the environmental impact and energy performance of open-loop groundwater heat pumps in urban areas. Applied Thermal Engineering, 71(2), 780-789.

[14] Benonysson, A., Bøhm, B., & Ravn, H. F. (1995). Operational optimization in a district heating system. Energy conversion and management, 36(5), 297-314.

[15] Brundu, F. G., Patti, E., Osello, A., Del Giudice, M., Rapetti, N., Krylovskiy, A., ... & Acquaviva, A. (2017). IoT Software Infrastructure for Energy Management and Simulation in Smart Cities. IEEE Transactions on Industrial Informatics, 13(2), 832-840.

[16] Burer, M., Tanaka, K., Favrat, D., & Yamada, K. (2003). Multi-criteria optimization of a district cogeneration plant integrating a solid oxide fuel cell–gas turbine combined cycle, heat pumps and chillers. Energy, 28(6), 497-518.

[17] Lund, H., Werner, S., Wiltshire, R., Svendsen, S., Thorsen, J. E., Hvelplund, F., & Mathiesen, B. V. (2014). 4th Generation District Heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. Energy, 68, 1-11.

[18] Nielsen, H. A., & Madsen, H. (2006). Modelling the heat consumption in district heating systems using a grey-box approach. Energy and Buildings, 38(1), 63-71..

[19] Idowu, S., Saguna, S., Åhlund, C., & Schelén, O. (2016). Applied machine learning: Forecasting heat load in district heating system. Energy and Buildings, 133, 478-488.

[20] Dotzauer, E. (2002). Simple model for prediction of loads in district-heating systems. Applied Energy, 73(3-4), 277-284.

[21] Nuytten, T., Claessens, B., Paredis, K., Van Bael, J., & Six, D. (2013). Flexibility of a combined heat and power system with thermal energy storage for district heating. Applied Energy, 104, 583-591.

[22] Wang, H., Yin, W., Abdollahi, E., Lahdelma, R., & Jiao, W. (2015). Modelling and optimization of CHP based district heating system with renewable energy production and energy storage. Applied Energy, 159, 401-421.

[23] da Cunha, J. P., & Eames, P. (2016). Thermal energy storage for low and medium temperature applications using phase change materials–a review. Applied energy, 177, 227-238.

[24] Sciacovelli, A., Guelpa, E., & Verda, V. (2014). Second law optimization of a PCM based latent heat thermal energy storage system with tree shaped fins. International Journal of Thermodynamics, 17(3), 145-154.

[25] Guelpa, E., Barbero, G., Sciacovelli, A., & Verda, V. (2017). Peak-shaving in district heating systems through optimal management of the thermal request of buildings. Energy, 137, 706-714.

[26] Verda, V., Guelpa, E., Sciacovelli A., F. G., Acquaviva, A., & Patti. (2016). Thermal peak load shaving through users request variations. International Journal of Thermodynamics, 19(3), 168-176.

[27] Guelpa, E., & Verda, V. (2016). Optimization of the Thermal Load Profile in District Heating Networks through “Virtual Storage” at Building Level. Energy Procedia, 101, 798-805.

[28] Zhao, H. X., & Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586-3592.

[29] Clarke, J. A. (2001). Energy simulation in building design. Routledge.

[30] Al-Homoud MS. Computer-aided building energy analysis techniques. Building and Environment 2001; 36(4):421–33.

[31] Crawley DB, Hand JW, Kummert M, Griffith BT. Contrasting the capabilities of building energy performance simulation programs. Building and Environment 2008; 43(4):661–73.

[32] Kalogirou SA, Bojic M. Artificial neural networks for the prediction of the energy consumption of a passive solar building. Energy 2000; 25(5):479–91.

[33] Olofsson T, Andersson S, Östin R. A method for predicting the annual building heating demand based on limited performance data. Energy and Buildings 1998; 28(1):101–8.

[34] Li Q, Meng QL, Cai JJ, Hiroshi Y, Akashi M. Applying support vector machine to predict hourly cooling load in the building. Applied Energy 2009; 86(10):2249–56.

[35] Sciacovelli, A., Guelpa, E., & Verda, V. (2013, November). Pumping cost minimization in an existing district heating network. In ASME 2013 International Mechanical Engineering Congress and Exposition (pp. V06AT07A066-V06AT07A066). American Society of Mechanical Engineers.

[36] HararyF.,GraphTheory.Narosa Publishing House, New Delhi;1995.

[37] Guelpa, E., Toro, C., Sciacovelli, A., Melli, R., Sciubba, E., & Verda, V. (2016). Optimal operation of large district heating networks through fast fluid-dynamic simulation. Energy, 102, 586-595.

[38] Guelpa E., Sciacovelli A., & Verda V. (2017). Thermo-fluid dynamic model of large district heating networks for the analysis of primary energy savings. Energy.