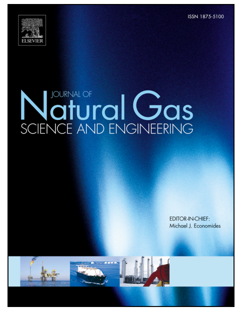
# Accepted Manuscript




Natural gas consumption and correlation with the uses of thermal energy: analysis of the Italian case

Alessandro Franco

PII: S1875-5100(16)30213-X

DOI: 10.1016/j.jngse.2016.03.094

Reference: JNGSE 1406

To appear in: *Journal of Natural Gas Science and Engineering*

Received Date: 28 January 2016

Revised Date: 13 March 2016

Accepted Date: 29 March 2016

Please cite this article as: Franco, A., Natural gas consumption and correlation with the uses of thermal energy: analysis of the Italian case, *Journal of Natural Gas Science & Engineering* (2016), doi: 10.1016/j.jngse.2016.03.094.

# Natural gas consumption and correlation with the uses of thermal energy: analysis of the Italian case

**Alessandro Franco**

Department of Energy, Systems, Territory and Constructions Engineering,
University of Pisa
Largo Lucio Lazzarino 2,
56126 Pisa, Italy

**Abstract**

In this paper, after a brief review of the methodologies proposed in the literature for the forecasting of heat load and thermal energy consumption, a specific analysis of the trends of natural gas consumption applied to the Italian case. Italy is a country in which the use of natural gas can be mainly connected to the uses of thermal energy for heating purposes. The natural gas consumption data are analyzed characterized and clustered referring to the last four years (2012-2015). The dependence of the typical Italian consumption profile in different days of the week, in different seasons and for the different users are analyzed: residential/civil, industrial and use referred to thermoelectric power plants are in particular analyzed. The analysis of the data shows that natural gas consumption profile is mainly related to seasonality pattern and to the weather conditions (in particular temperature below 15 °C) for the civil/residential sector, but it is shown that there is also an important daily pattern both related to industrial and civil energy consumption that, at a lower degree than the previous one, does affect the consumption profile and have to be taken into account for defining an effective short and mid term thermal energy forecasting method. A possible mathematical structure of the natural gas consumption profile is proposed. Due to the strong link between thermal energy use, mainly connected to heating purposes, and natural gas consumption this analysis could be considered the first step for the development of a model for thermal energy forecasting.

**Keywords:** Natural gas consumption; Data analysis; Thermal energy use; Clustering; Forecasting;

* Corresponding author:
Alessandro Franco
Department of Energy, Systems, Territory and Constructions Engineering
Largo Lucio Lazzarino,
56126, Pisa, Italy
Phone: +39 050 2217154    Fax: +39 050 2217150
e-mail: alessandro.franco@ing.unipi.it

## 1. Introduction

The recent years have witnessed significant efforts worldwide at multiple levels, from research to policy initiatives to support the interaction of the various sectors of energy at different levels, from demand to production. If in the past each specific sector (electricity, thermal energy and mobility) have been analyzed in an independent manner; today this approach is no longer possible. Over the years, new technologies that create links between the various energy uses have been developed. For example systems for Combined Heat and Power (CHP) binds electric and thermal uses while the diffusion of Heat Pumps (HP) shifts thermal energy consumption to the electricity sector, with possible contribution of electricity produced with renewable sources in the thermal sector. Moreover the important penetration of Renewable Energy Sources (RES), in particular Photovoltaic (PV) plants and Wind power plants, determines new problem in term of balance of the energy fluxes. Therefore a global perspective in which the single energy system is seen as a component of a complex macro system is required (Franco and Salza, 2011).

The future development of energy systems will be not only in the use of more efficient components and technologies, but mainly in the perspective of optimizing control, distribution and consumption of the various energy vectors in an integrated mode. A fundamental part of this process is the development of control strategies that allow a minimization of inefficiencies and energy waste. In this perspective it will be of primary importance the development of methods and tools for energy planning with the perspective of mid and long term (Huang et al., 2015) but also the development of provisional instruments for the short and mid term (Ardakani et al., 2014). Information about the patterns that govern the energy demand can determine significant energy savings. Accurate models for electric and thermal power load forecasting are essential to the operation and planning of a utility company. For example, the electricity generation can be optimized if the electric load to cover is known: this is well considered in the literature since ten years ago (Mirasgedis et al., 2006). Until few years ago the development of the electrical forecasting was the first to catch on. This because the energy system was still broken down into its individual sub-parts and the electrical component was the most important. But it was also a matter of simplicity. In fact the electrical consumption can be well divided between working days and weekends and has a low temperature dependence. It is a quite stable system and therefore it is easier to describe.

But important advantages connected to the development of forecasting methods will be obtained in the field of thermal energy and natural gas uses too. This because detailed analysis of this sector is more recent and more recently connected with the energy market. Moreover, the thermal energy is not easily measurable as the electric energy, its time and inertias are bigger and it presents strong dependence on external factors. So if methods for electric energy load and production forecast are well developed in the literature, (Sandels et al., 2014) the same cannot be said for the thermal load, even if some attempts are present in the literature (Feinberg and Genethliou, 2005) and (Nielsen and Madsen, 2006). These concerns both simple cases, like the single family houses (Bacher et al., 2013), large-scale district energy system (Powell et al., 2014) or more complex urban systems (Kwak et al., 2013 and Dirks et al., 2015).

The thermal energy consumption is in direct connection with the weather parameters (temperature, humidity, wind speed and solar radiation); but different parameters, like the day of the week as well as economic and behavioural elements, are important too. Considering that it is rather difficult to manage data directly connected with the use of thermal energy, the idea developed in the paper is to analyze the data of consumption of natural gas in order to obtain from those data support information for the definition of a model for thermal energy forecast.

The specific objective of the paper is to analyse the natural gas consumption profiles in Italy, considering the last four years and to infer the characteristic patterns of such profiles in order to define a short and mid term forecasting method of natural gas consumption profiles and to obtain information about the structure of a forecasting model.

Considering the strong link observed in Italian market between the natural gas consumption and the use of thermal energy, the analysis and clustering of natural gas data consumption can be really considered as the first preliminary step for the definition of a model for thermal energy forecasting.

**2. The problem of forecasting heat load and thermal energy use**

Forecasting is defined as a tool used for predicting future demand based on past demand information. As stated in the previous section, forecasting of energy use is really important for general development of energy policy, for designing efficient demand response functionalities to adapt the energy suppliers supply strategy to the customer load as much as possible and to support the development of environmentally friendly urban planning, without redundancy (Yeo et al., 2013). Though if modeling of end-use energy consumption are available in the literature since some years ago, both for industrial and residential sector, (Schellong, 2011) and (Swan and Ugursal, 2009), forecasting of the thermal energy use, mainly in connection with heating purposes and industrial processes, appears to be a quite complex task. Heating load is challenging to forecast as it requires forecasts of weather factors and the knowledge of customer behaviour.

The customer base generally is divided into three categories: residential, commercial and industrial. The demand characteristics of these three categories differ in a significant way. The use of thermal energy depends on a lot of variables, primarily on the weather conditions, but a lot of dependent variables can be identified; those can be organized at two different hierarchical levels, as summarized in Fig. 1 (Yu et al., 2011).

As stated before, an important amount of the thermal energy consumed is caused by the heating of civil sector, but the part relating to industrial uses, such as the production of process steam, that is used for a lot or processes, is important too. While the first component can be directly correlated to climatic variables (in particular temperature), the second component is more directly related to the planning of production (day of the week). The topic of the prediction of energy consumption has been widely debated in the literature as there are interests of system optimization (reduction of waste energy and of the difference between supply and demand) and also economic reasons mainly as a result to the increase number of customers in the free market.

If the approaches for the forecasts of the electricity demand are quite reliable and well-tested, the same cannot be said for thermal energy uses and heat load. So there is an entire sector of the energy system in which the application of appropriate forecasting tools could be of special interest. In general the methods for load forecasts, including thermal load can be divided into three categories: short-term forecasts which are usually from one hour to one week, medium forecasts (from a week to a year) and long-term forecasts. Surely the more interesting are the first two types. Even if long-term forecasting is important in order to define capital expenditure on projects, with supply and demand fluctuating and the changes of weather conditions and energy prices increasing by a factor often or more during peak situations, short and mid-term load forecasting are vitally important for utilities in order to make decisions that can prevent distribution problems. Moreover, in the perspective of diffusion of deregulated market, heat and thermal load forecasting is also important for contract evaluations. In addition to their type, models can also be divided based on spatial levels in which they operate. At a spatial level it is possible to work on data that, referring to different sizes of groupings, require a different aggregation: single elements (buildings, factories, etc.) with a "low-level" aggregation data; specific sectors or geographic zones (e.g. office sector, educational sector, a specific industrial sector, district, municipality, single province) with a "mid-level" aggregation data, or considering a macro-zone (regional, country, state) with a "high-level" aggregation data.

The combination of time and space determines a perspective that can be variable from a systems analysis (with general and strategic perspectives) to an operational perspective (Fig. 2). In order to define a model for the forecasting of thermal energy use a clear definition of the system under analysis and of the objectives have to be carried out.

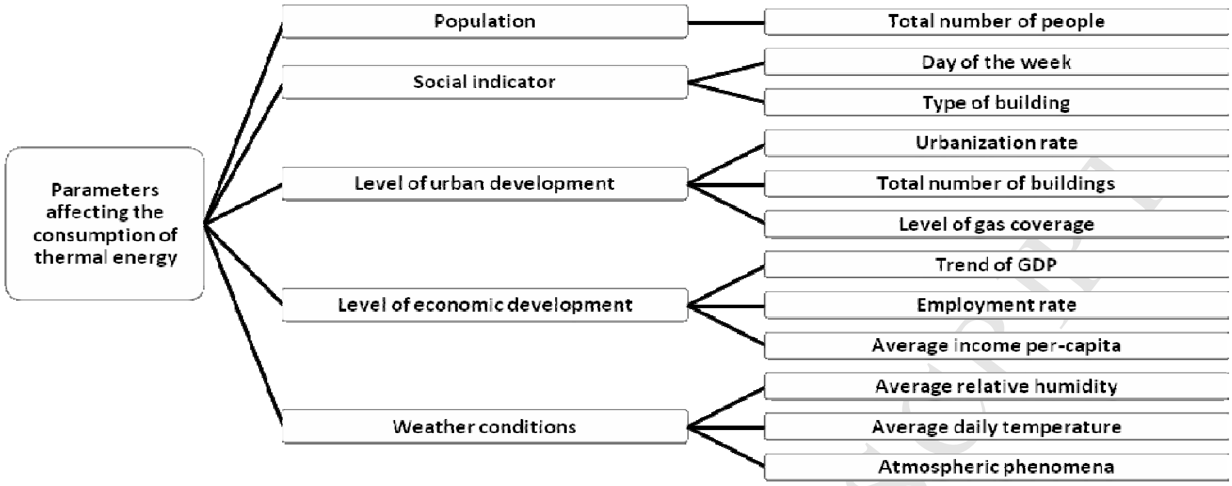



**Fig. 1: Possible system indicators for thermal energy consumption**

However different perspectives can be also assumed. Concerning the use of thermal energy, all models analyzed in the literature can be divided into 3 types: models based on the analysis of data ("*Statistical*"); models based on a more clear definition of the system and on the physical characteristics of it ("*Physical-engineering*") and models that try to combine the two previously mentioned approaches ("*Hybrid*"). The various models available in the literature for forecasting are summarized in Fig. 3.

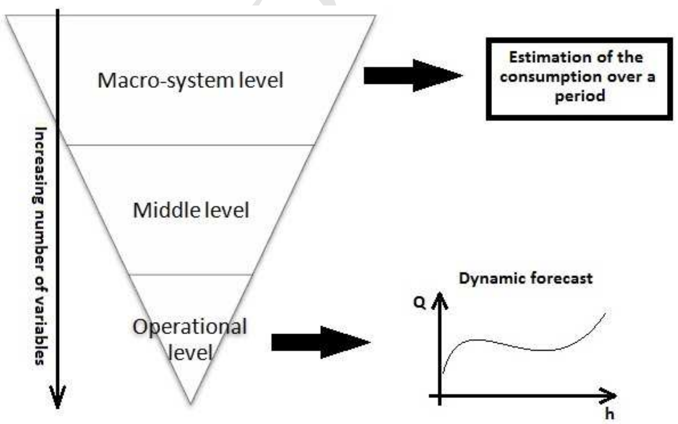



**Fig. 2. Approach levels to the problem of forecasting**

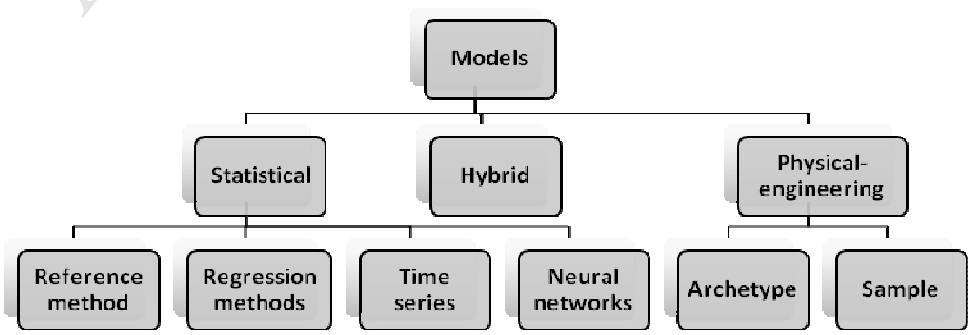



**Fig. 3. Classification of the forecasting models**

*2.1 Model based on the analysis of the data ("Statistical models")*

Statistical models use computational techniques to find relationship between the required data, in this case heat load and thermal energy consumption, and the main variables on which it is considered dependent (Sughanti and Samul, 2012). A connection between thermal energy consumption and some specific variables (the weather, the days of the week, the user profiles, the trend of economy, the development of certain technologies) can be identified. These relationships are then used to estimate the thermal load once known quantities on which it depends.

It is therefore necessary to know the time series of the variables involved in order to be able to calibrate these relationships. The correlation is defined basing on a statistical analysis of the data. A lot of statistical models are available in the literature. The "*reference method*" is rather simple and it does not require a proper mathematical structure. It contains a database of thermal energy consumption and the variables on which it depends. Having information on the variables for the day, it simply searches in the database one day with the variables as similar as possible. The load is set to be the same as the day in the database.

The "*regressive models*" represents a second option: it is supposed that there is a link between the objective function (for example the dependent variable, $Q$) and the selected variables (independent variables, $x, x'$), as temperature, humidity or wind speed, dependent on some multipliers $c_i$. The multipliers are defined by minimizing a certain type of error with respect to a set of data of calibration. For simplicity it is usual to use a link of the linear type and the error is the average quadratic error, so that

$$Q = f\left(x, c_0, c_1, c_2\right) = c_0 + c_1 \cdot x + c_2 \cdot x' + \varepsilon \qquad (1)$$

$$\min_{c_0, c_1, c_2} \sum_{i=1}^{n} \left[ Q_i - \left( c_0 + c_1 \cdot x_i + c_2 \cdot x_i' \right) \right]^2 \qquad (2)$$

A third interesting group of statistical models is those referred to the category of "*time series*". They are based on the assumption that the consumer has an inside regular pattern. In practice it is assumed that the observed values of a specific datum will continue in the future if the same boundary conditions will be observed. There are various types of time series, depending on the characteristics of the magnitude under examination. The basic form consists of the models ARMA (*Auto Regressive Moving Average*), modified and expanded in SARMA (*Seasonal Auto Regressive Moving Average*) in the case of series with seasonal patterns, ARIMA (*Auto Regressive Integrated Moving Average*) in the case of non-stationary trends and ARMAX (*Auto Regressive Moving Average with eXogenous input*) in the case of introduction of dependence on external variables. Another method belonging to the statistical methods is the well known "*Neural networks*" that had a great development in recent years. (Hernandez et al., 2014). A neural network consists of a number of calculation cells, interconnected with each other according to a specific pattern. Each cell communicates with the others: it receives data, transforms them using an internal function and then retransmits them to the next cell. Interesting aspects are the ability to model non-linear systems using non-linear internal functions and the ability to self-adapt to the problem. After calibrating the weights of the connections over an appropriate set of data, the variable upon which the consumption depends are used as input data and as output we get the estimate for consumption. The models for thermal energy forecasting based on Neural Networks have been brought within the field of statistical models but may represent an interesting option. The Neural Networks are very promising in case of systems rather stable, like for example the electricity sector (Hassan et al., 2015) and (Anbazhagan and Kumarappan, 2014). In case of thermal energy forecasting, the ability of the neural network to reproduce an environment behavior could be counterproductive when the particular situations (like particularly cold days) have a fundamental importance.

*2.2 Model based on the physical description of the system: "Physical-engineering" models*

The models based on a more direct physical description of the system and of direct correlation with the output variables directly calculate the thermal load and the thermal energy consumption basing on physical considerations (Sandels et al., 2014). In this case, the structure of the user is well defined and the correlation with data, such as the external temperature, thermal properties of building envelope, the type of generation and distribution systems present is introduced. Those models can be quite simple in case of single building or small district, but could be rather complex or based on important simplifications. Anyway they do not depend on the data time series and they are able to directly calculate the energy consumption. This makes them able to assess the impact of any improvements or replacements in the environment analyzed. The main physical-engineering found are: "archetypes models" and "models samples". The "*archetypes models*" does not really analyze existing buildings. They make a preliminary analysis to identify the average values of the quantities that are interested in the calculation, and then use these averages to construct a dummy building having those properties. The buildings are constructed as a sort of average buildings of the area concerned. The "*samples models*" instead analyze a number of representative buildings of the housing stock. The overall result is obtained by multiplying the results by appropriate weights representative of the actual distribution of that building types in the whole stock. The physical models are surely the best solution in case of small aggregation level (particular buildings and small districts). However in case of complex aggregations (e.g. regional level) they require a complex mathematical modeling and a lot of data.

*2.3 Hybrid models*

Considering the good and bad qualities of the methods previously analyzed, it seems quite clear that a good method for thermal energy forecasting, the strengths of both statistically based and physically based models need to be used. A kind of "hybrid" model, in which a physical description of the system including end users and appliances, in the civil, industrial and thermoelectricity production sectors is joined with statistical elements or artificial intelligence algorithms in order to predict the short, mid and long term energy consumption. Even if a univocal definition is difficult in this case, these models are very promising. An interesting example is the model proposed by (Nielsen and Madsen 2006) in which the authors use a physical approach based on weather variables (temperature, wind and radiation) to infer the parameters to be set out in a time series model. Among the various methodologies analyzed, hybrid models seem the most appropriate to be used in a sector as complex as the thermal energy system.

**3. The link between natural gas consumption and thermal energy use: the Italian case**

As previously discussed, while the development of forecasting method for electricity can be based on available data of consumption and electricity load, the same approach cannot be used for the thermal energy.

For a country like Italy, the maximum amount of energy used consists of electricity (the mix and spatial distribution of generation technologies, transmission, and distribution) and natural gas (production, transmission, distribution, and storage).

Italy, a country with more than 60 million of inhabitants, is one of the main consumers of natural gas within European Union after UK and Germany. Natural gas is widely utilized in Italy for different purposes such as production of energy for heating demand using small, medium and large scale boilers, industrial processes, generation and residential other activities, in particular heating, cooking and sanitary water production.

The customer base generally is divided into four categories: residential, commercial, industrial, and electric power generation. Some important drivers like the use of natural gas for thermal energy production and the quite high number of combined cycle power plants have modified the traditional cyclical demand of natural gas.

Another important factor is the cost and the availability of natural gas, which is itself related to many political and economical factors, many times not fully predictable. The analysis of the drivers of natural gas consumption is currently object of research, as suggested by a recent papers (Dilaver et al. 2014, Szoplik, 2015), but it is clear that in a particular country like Italy, the use of thermal energy use is the most important driver.

Furthermore forecasting of natural gas consumption has been already considered in the literature, for example by (Sabo et al, 2011) and (Soldo, 2012) with a general perspective. An interesting analysis of the mathematical models that can be applied for natural gas consumption forecasting is also available in Vitullo et al. (2009).

Numerous researchers and practitioners have analyzed various issues and focused on developing appropriate models for natrural gas demand forecasting with reference to specific countries of different size and growth trends like China (Li et al., 2011), Turkey (Sarak and Satman, 2003; Erdogdu, 2010 and Taspinar et al., 2013), Iran (Forouzanfar et al., 2010), Bangladesh (Wadud et al., 2011). Prognostic models consider calendar-(weekday, daytime, month, season), weather-related (temperature, humidity, sunshine, wind speed) factors, as well as demographic and economic factors and each country has a specific behaviour.

The residential and non residential consumption of natural gas in Italy has been analyzed in the literature in two papers (Bianco et al. 2014). They proposed a model to forecast residential and non-residential consumption of natural gas; consumption drivers are identified and discussed and a single equation demand model is identified. Anyway the object of the two papers is the obtainment of long term scenario and the effect of economic drivers is considered of primary importance. It is important to remark that the renovation of the natural gas wholesale and retail markets together with a technological review of current smart energy practices is a key step to accomplish an optimal energy consumption pattern, to avoid energy wastes and to implement a sustainable energy system. In this vision smart meters installed at the utility endpoints to monitor usage could be an important component of the smart grid. As discussed in the previous section, high resolutions automated meter reading system for residential use monitoring, which can be used to record gas consumption for each appliance have been already experimentally tested, even if in most cases they are not yet in the commercial development phase (Tewolde et al., 2013). Research on smart meters appears to be interesting from a technological point of view, but the development and application of smart metering is a particularly complex task mainly for aspects related to data treatment and communications.

The true challenge and difficult task in the field of smart meters' development, which also occurs in the use of Automatic Meter Reading (AMR) devices, appears to be the fact that monitoring energy usage data poses privacy risks that might not be easy to solve (Nist, 2010 and Rouf et al., 2012). On the other hand, fine-grained energy consumption data collected by AMR could reveal sensitive information coming from home (number of occupants, vacancies, typical energy usage, etc.). While there is a general consensus that energy monitoring devices can be an opportunity to "*provide consumers with a more comprehensive and nuanced understanding of their usage patterns*", the only real perspective of considering demand side management is the gas flow monitoring (Kerr and Tondro, 2012). In this case even if there has recently been a lot of research in this topic (Tewolde et al., 2013) and the italian distributors were required to install Smart Gas Meters (SGMs) at 100% of all non-domestic consumers' premises by the end of 2012 and at 80% of domestic consumers' premises by the end of 2016 (Di Castelnuovo and Fumagalli, 2013), the availability of data is limited to integral data.

Natural gas has been used since the '70s to satisfy the needs of many commercial and residential users throughout the world through a huge and complex network. Demand for natural gas has traditionally been cyclical, simply following the characteristic seasonal patterns. As a consequence, the cyclical nature of the demand was used in the past to accurately predict the required natural gas: demand was highest during the coldest months in winter and lowest during the warmest months of the year in summer; also, a direct correlation with average medium environmental temperature could be observed. In addition, until a few years ago, the connection between gas and electricity sectors was marginal. Things began to change at the beginning of the '90s, when the commercial development of combined cycle power plants led to a shift towards the use of natural gas for the generation of electricity as well. This had a strong impact on the traditional cyclical gas demand. As a further complication to this scenario, in the last few years there has been a consistent increase in the penetration of intermittent renewable energy production, particularly in countries like Italy, Denmark and Germany. In turn, this has implied that gas is being used as a back-up for electricity production when renewables are not available (e.g., to back-up solar plants at night time), thus contributing to tighten the connections between gas and electricity. The significant penetration of RES in electricity production has imposed specific logic to the consumption of natural gas in thermoelectric plants, mainly natural gas combined cycle power plants. The topic is discussed by one of the author in (Franco and Salza, 2011). Due to the particular development in the 80 and 90s of thermal energy systems based on the use of natural gas, Italy represents a particular case where the use of thermal energy for heating purposes could be correlated with the consumption of natural gas

*3.1 Database definition*

For the Italian case, daily gas flows are available from the database of national network for natural gas distribution (SNAM Rete Gas): for the aim of the present analysis, the data referred to the last four years 2012-2015 have been considered. The data are of public domain and are available on the website.

It is important to remark that in those four years, mainly between 2012 and 2014, Italy has been interested by an important development of renewable energy systems (PV, Wind, Bioenergy). Both the use of heat pumps and of thermal energy systems based on biomass has been largely increased. As additional element, it is possible to observe that in 2013 and 2014 the effect of economic crisis has influenced the consumption of natural gas, so that the year 2012 must be the last year in which it is exactly clear the connection between thermal energy use and natural gas consumption. The weather is represented by an ideal average temperature, obtained from a specific database (Il Meteo) considering the daily average temperatures of six different towns in Italy, that are well representative of the various climatic conditions of the country (two in the north, Torino and Trento, two in the center, Roma and Firenze and two in the south, Bari and Palermo).

The natural gas consumption data are represented in aggregated form and in the three main customer components (gas for thermoelectric plants, gas for industrial sector and gas for civil-residential sector). The effect of season is represented by an average daily value of the temperature. Fig. 4 shows the strong fluctuation of the natural gas demand, expressed in $Mm^3$ ($1\ Mm^3 = 10^6\ m^3$), taking into separate account the two main sectors of use: gas for heating and gas for thermoelectricity production.

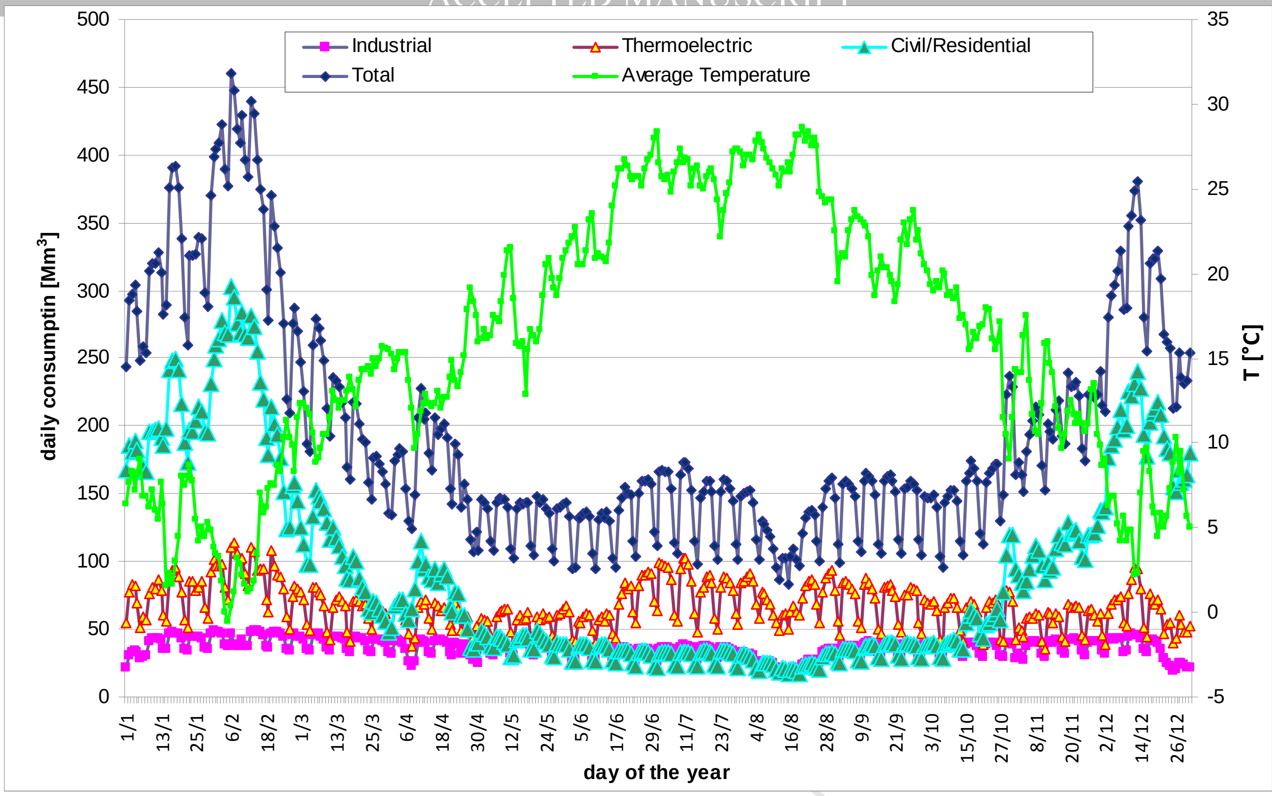



**Fig. 4. Consumption of natural gas in Italy in 2012**

As can be seen in Fig. 4, gas for heating varies in a significant fashion depending on the time of the year (cold vs. warm seasons); on the other hand, gas for thermoelectricity remains of the same order of magnitude over the year, and fluctuates according to other dynamics (e.g., weekly patterns).

However, even if it is not trivial to predict accurately the evolution of natural gas demand, the trend of natural gas consumption analyzed in Fig. 4 for the year 2012 can be evidenced with some quantitative differences in the following three years, 2013, 2014 and 2015, as reported in Figs. 5-7. To appreciate the quantitative difference related to different extreme weather conditions and some other evolutions, te limit data (maximum and minimum values of natural gas consumption) observed in the various sectors are grouped in Table 1.

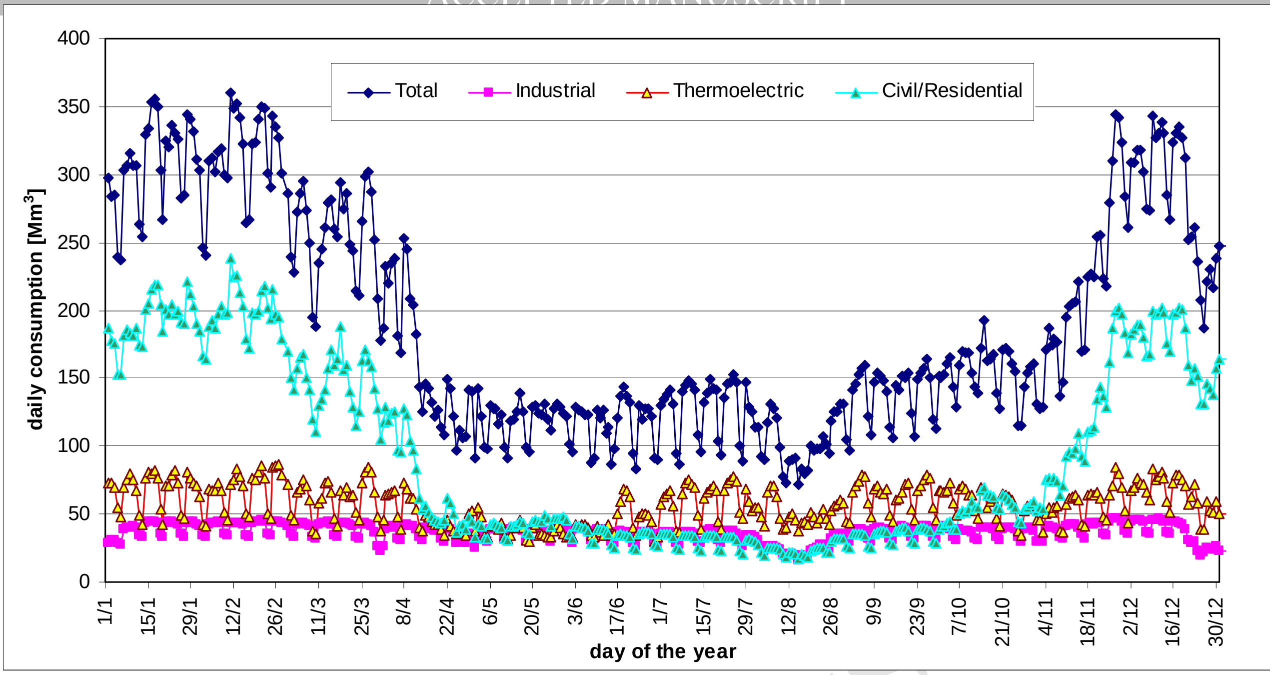



**Fig. 5. Consumption of natural gas in Italy in 2013**

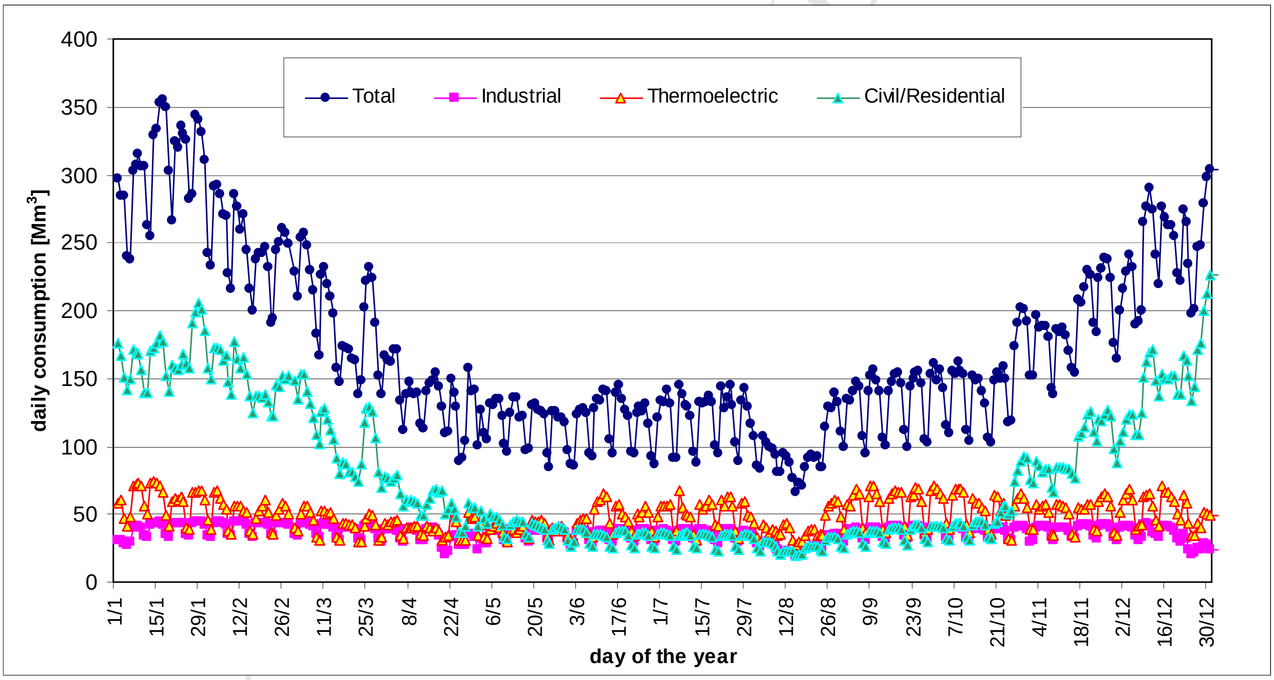



**Fig. 6. Consumption of natural gas in Italy in 2014**

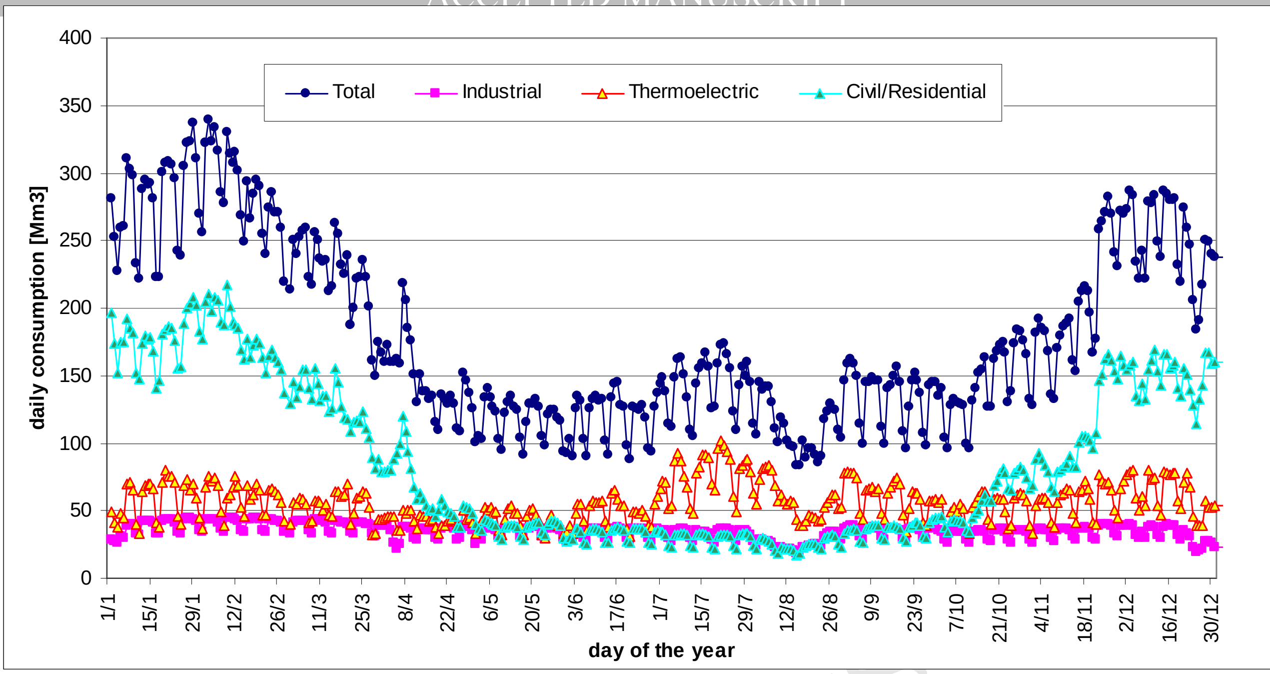



**Fig. 7. Consumption of natural gas in Italy in 2015**

**Table 1. Maximum and minimum values of the natural gas consumption in the period 2012-2015 (data in Mm$^3$)**

| | **2012** | | | **2013** | | | **2014** | | | **2015** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Min.** | | **Max.** | **Min.** | | **Max.** | **Min.** | | **Max.** | **Min.** | | **Max.** |
| **Thermoelectricity** | 35.7 | - | 113.4 | 29.7 | - | 86.3 | 27.2 | - | 74.8 | 30.1 | - | 101.4 |
| Dates | (11/11) | - | (7/02) | (29/06) | - | (27/02) | (15/08) | - | (14/01) | (24/05) | - | (21/07) |
| **Industrial sector** | 17.0 | - | 49.5 | 16.4 | - | 46.4 | 18.9 | - | 44.7 | 18.3 | - | 47.7 |
| Dates | (15/08) | - | (14/02) | (15/08) | - | (26/11) | (15/08) | - | (4/02) | (15/08) | - | (27/01) |
| **Civil sector** | 17.0 | - | 302.9 | 17.1 | - | 238.6 | 19.0 | - | 226.3 | 17.3 | - | 217.4 |
| Dates | (15/08) | - | (6/02) | (15/08) | - | (11/02) | (15/08) | - | (31/12) | (15/08) | - | (9/2) |
| **Total** | 83.6 | - | 459.6 | 71.6 | - | 359.7 | 66.0 | - | 355.3 | 82.9 | - | 339.7 |
| Dates | (15/08) | - | (6/02) | (15/08) | - | (11/2) | (15/08) | - | (17/1) | (15/08) | - | (3/2) |

While natural gas consumption in thermoelectric power plants is linked to the electricity market, because it is directly referred to the electricity generated in gas fuelled power plants and it is not directly dependent on the need for thermal energy use, the other components have major correlation. Residential consumption of natural gas is determined by the demand for heating, sanitary water and cooking facilities in residential buildings. In this case the main consumption drivers are represented by external temperature of the heating season, population and buildings characteristics (i.e. insulation, facilities, etc.). Non-residential natural gas consumption represents the usage of natural gas related to economic activities and services (i.e. offices, shops, etc.). Concerning natural gas consumed in the industrial sector (i.e. manufacturing, food, construction, etc.) it is mainly used in production processes, even though a substantial share difficult to estimate is also utilized for heating purposes (i.e. heating of very large industrial buildings).

As observed in Fig. 4 for the year 2012, it is possible to appreciate the clear link between an average outside temperature and the natural gas consumption. But the influence of other elements, like the day of the week (in particular the difference between working days and holidays) cannot be neglected. The civilian/urbane consumption is the share of gas flowing in the network that is used to heat the civil sector (residential and tertiary). By making an initial clustering of the data between the various sector, analyzing the trend of the civilian consumption during the year (Fig. 8) it is possible to clearly identify the dependence of the structure of the demand not only on the level of the outside temperature.

*3.2. General analysis of the data*

The most significant factor for modelling natural gas consumption is temperature, since most gas is used for space heating. Gas consumption versus average temperature for individual days is plotted in Fig. 8.

By visual analysis of the data it is possible to propose two clusters related to two different meteorological seasons: the "cold" season (starting from the middle of October and ending at the end of April) in which the natural gas consumption can be clearly linked to the average external temperature and the "warm" season (from May to middle of October), in which the consumption is not particularly influenced by the outdoor temperature.

Considering the problem on a quantitative perspective it is possible to state that when temperatures are cold, as temperature increases, gas consumption decreases in a nearly linear way, although once the ambient temperature reaches approximately 15 degrees consumption levels off. Once the average temperature reaches a certain temperature, space heating no longer occurs; consumption levels are near some constant value known as base load.

The data reported in Fig. 8, expressing the natural gas daily consumption in the civil sector for the k-th day, expressed in $Mm^3$, can be approximately described by an exponential function of the temperature T, of the type:

$$y = A \cdot \exp(-0.1 \cdot T_{k,avg}) \tag{3}$$

where $T_{k,avg}$ is the average temperature for the k-th day expressed in °C and A is a constant (in the case analyzed for example A = 320-330). It is also possible to refer the total consumption with reference to the degree days with a function of the type

$$\begin{cases} y_k = B + K(T_{ref} - T_{avg}) & T_{k,avg} \geq T_{ref} \\ y_k = B & T_{k,avg} < T_{ref} \end{cases} \tag{4}$$

where B is a basic consumption value (for example B=30-35 $Mm^3$) and K is a coefficient defining the correlation between the consumption and the difference between a current value of the average temperature and a basic reference temperature (e.g. $T_{ref}$ = 15 °C). In the case under analysis the value of K can be between 17.5 and 18).

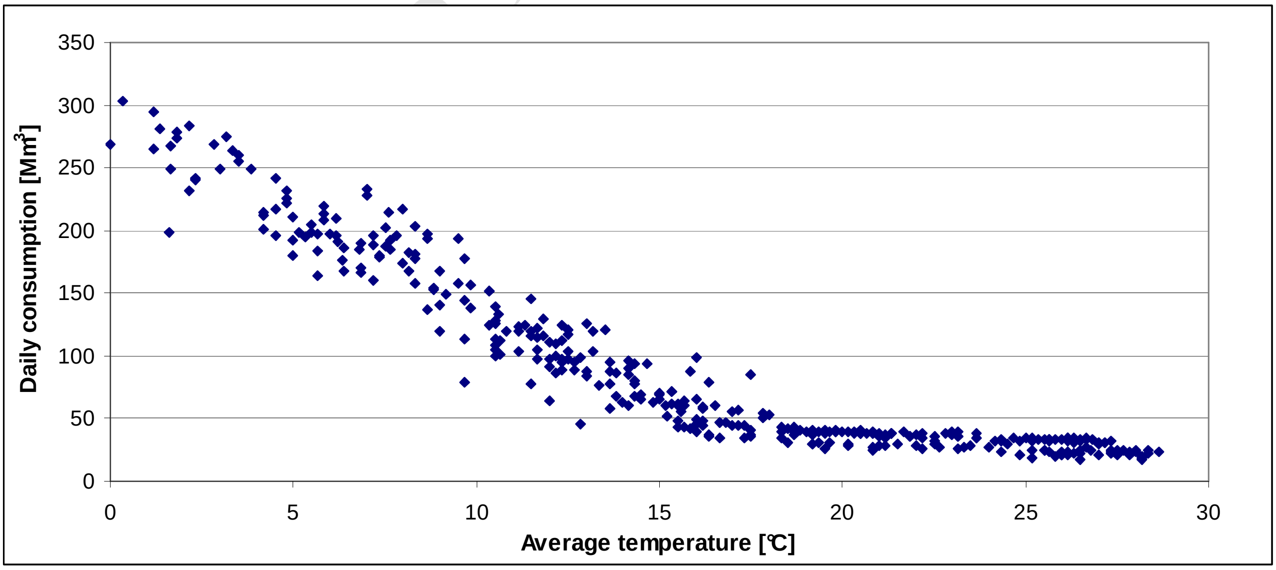



**Fig. 8. Link between total natural gas consumption and average temperature in civil/residential sector**

But analyzing in a more careful way the aggregated data of natural gas consumption it can be also evidenced that the gas consumption data, both in the "cold" season and in the "warm" season rely on the specific type of the day, e.g., working days, vs. holidays, vs. pre-holidays, two weeks of August. In general the gas consumption is smaller on Saturdays (due to the fact that some activities, like offices, schools and commercial activities are closed) and even smaller on Sundays. This suggests that the load should be clustered according to two or three additional clusters.

## 4. Analysis and clustering of natural gas consumption data: research of peculiar patterns and general elements

Classification and clustering of time series date is recognized as an important area of research for energy data both considering demand and production. Clustering refers to the ability to aggregate data with similar characteristics; the basic clustering operation corresponds to take a set of a given number n of objects and/or data and group them into k clusters.

A good clustering method has predictive power; in this case: clusters can be used to predict future trend of the data, for example concerning energy consumption. Moreover clusters allow compressing a great number of information into a single or a reduced number of information. In addition clusters permits to identify the "outliers", i.e., the cases in which clusters fail to accurately represent particular data. The objective of this section is to propose possible clustering of natural gas consumption data. Analyzing natural gas consumption data, the seasonality effects connected with the average temperature can be clearly observed with sensibly higher consumption in winter, lower values in autumn and spring and the lowest values in summer. But analyzing the data it is also interesting to distinguish the various components of natural gas consumption, in particular referred to industrial load, civil/residential load and thermoelectric load and the correlation with the temperature is different in the three cases. Also, three main types of energy consumption users can be identified. The first type (majority of users) is characterised by the domestic consumers.

The second type of consumers is characterised by the tertiary activities, having a quasi-flat consumption profile during the day. Finally, the third type involves industrial use of natural gas, both for thermal energy and other industrial uses.

### *4.1 Analysis of the different sectors*

Concerning the residential use of natural gas, the analysis of the consumption data shows that it can be mainly referred to seasonality pattern: this as can be evidenced by the strong link with the average external temperature, mainly when the temperature is below 15 °C. However different profiles in different days of the week can be also evidenced. Weekly consumption variation is significant and particularly visible as a decrease in weekend consumption for the same weather conditions. It can be easily seen by visual inspection that natural gas consumption data rely on the specific type of the day, e.g., working days, vs. holidays, vs. preholidays. Clearly, the consumption is smaller on Saturdays, due to the fact that some activities (like offices and schools) are closed and even smaller on Sundays (a lot of commercial activities are closed too). This suggests that the load can be splitted into three clusters corresponding to the category of the day of the week (e.g., working days, Saturdays and Sundays). In case of industrial users, by classifying load consumption into two categories (working days and holidays), it is possible to identify two clusters while the natural gas consumption appears to be less connected with the outside temperature. On the other hand, visual inspection of the aggregated natural gas consumption data concerning the civil sector, reported in Fig. 9, suggests that the consumption, for a well defined value of the outside temperature is different depending on whether the day is a working day or not. But the difference between the two types of days is less remarkable than for industrial sector.

More generally, it is possible to recognise whether a given datum belongs to a holiday (i.e., Sundays, or other Italian holidays), to a working day, or to a pre-holiday. With the term preholidays it is defined the set of Saturdays, the days within two holidays when most offices are closed, and the working days during conventional Italian summer and Christmas holidays (i.e., the two weeks in the middle of August, around 15 and the two weeks between Christmas and January 6). Fig. 9 and 10 provides the link between the daily consumption of natural gas and the average outside temperature in the weekdays (Fig. 9) and during all the other days of the year, a total of 112 days (Fig. 10), grouping preholidays, holidays and special days, according to the detailed description reported in Table 2. This kind of analysis can be carried proposed for the other three years under analysis, without modifying the perspective.

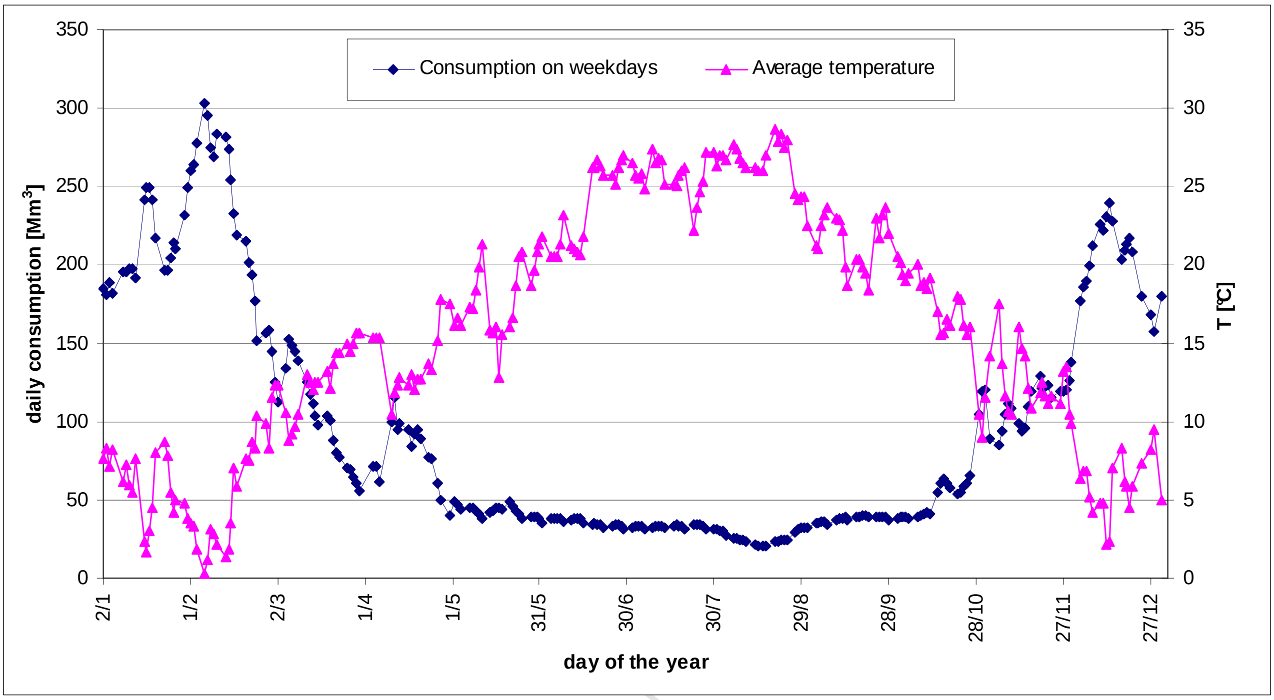



**Fig. 9. Natural gas consumption data during the days of the week**

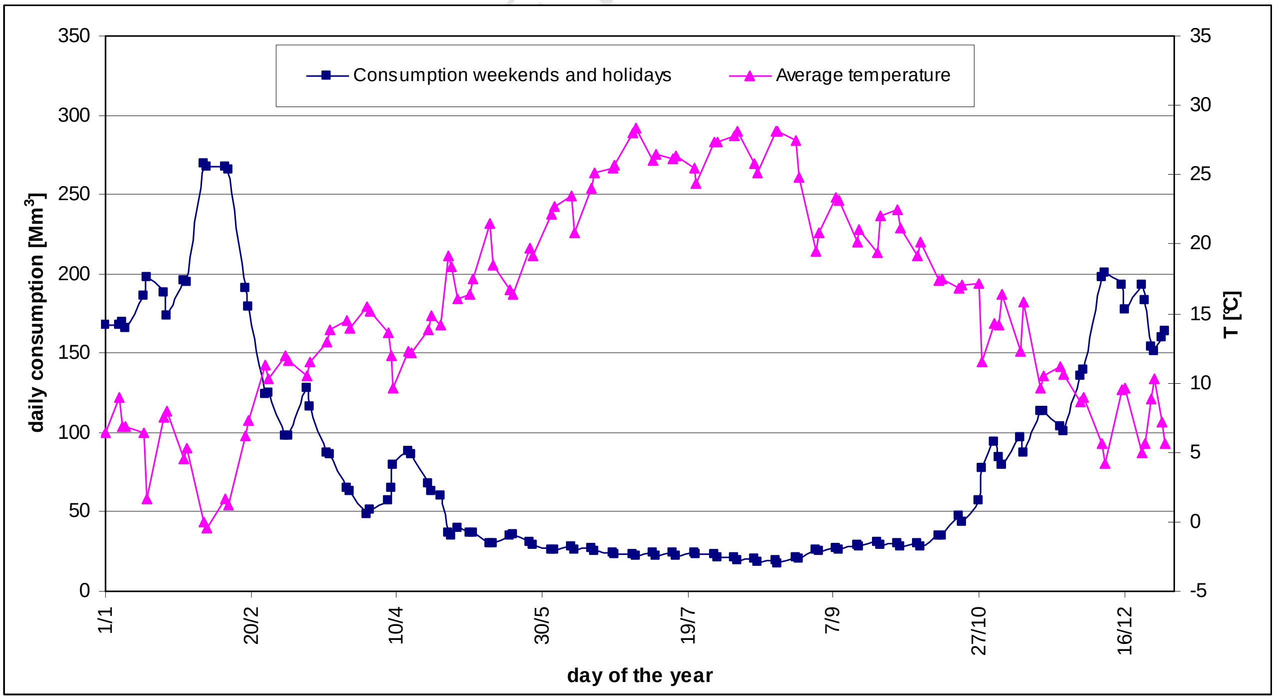



**Fig. 10. Natural gas consumption data during preholidays and holidays**

**Table 2. Working days (W), preholidays (P), holidays (H) during the year 2012 in Italy.**

| Mon | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Jan | *H* | W | W | W | W | *H* | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W |
| Feb | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | | |
| Mar | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* |
| Apr | *H* | W | W | W | W | W | *P* | *H* | *H* | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | *H* | W | W | *P* | *H* | *P* | |
| May | *H* | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W |
| Jun | *P* | *S* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | |
| Jul | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W |
| Aug | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | *H* | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W |
| Sep | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | |
| Oct | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W |
| Nov | *H* | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | |
| Dec | *P* | *H* | W | W | W | W | W | *H* | *H* | W | W | W | W | W | *P* | *H* | W | W | W | W | W | *P* | *H* | W | *H* | *H* | W | W | *P* | *H* | W |

The data on the horizontal axis represents the progressive number of the day in accordance with the description of Fig. 8, including all holidays, pre-holidays, working days and special days in Italy in the 2012 calendar. Considering the aggregated data analyzed, a multiplicative model seems to be the useful in order to predict the natural gas consumption in the civil sector in Italy. The mathematical formulation will be the following:

$$C(t) = F(d(t), f(w(t))) + R(t) \qquad (5)$$

where C(t) is the actual consumption at time t, d(t) is the day of the week, F(d) is the daily component, w(t) is a function of the weather data that include temperature, humidity and wind chill, f(w) is the weather function, and R(t) is a term of correction (random). This last term reflects for example the fact that thermal energy uses depends not only on the current weather conditions but also on the weather during the previous days. In particular special conditions occur when the cold weather continues for several days or when a cold day arrives after some "hot" days.

Analysing the consumption of natural gas referred to industrial sector (Fig. 11), it is possible to evidence the reduced dependence on the temperature and a most important influence of the particular days of the week. In particular a remarkable difference between working days and holidays can be clearly evidenced.

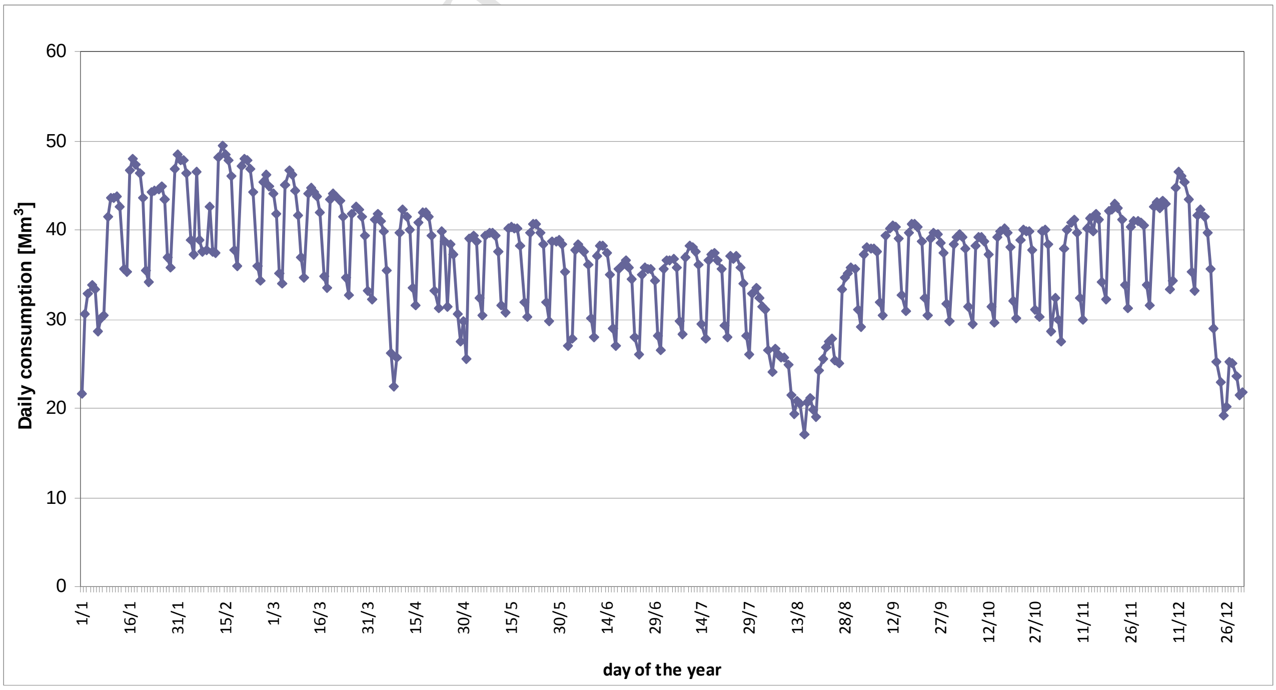



**Fig. 11. Natural gas consumption in the industrial sector**

Elasticity values with respect to the temperature have been estimated, for the natural gas consumption in industrial sector. The next figures permit a quantitative analysis of the results. Fig. 12 reports the link between the daily consumption and the temperature in the weekdays, while Fig. 13 and 14 provide the same data for the three different types of days: pre-holidays and holidays (excluding special holidays).

As it can be evidenced from the various figures, there is a basic consumption level and an additional amount directly linked to the temperature. It is possible to conclude this analysis putting in evidence that if in the case of natural gas consumption in the civil/residential sector, the link with weather conditions and temperature is clearly evident (the difference between the minimum level is remarkable, going from 20-30 $Mm^3$ for each day up to 300 $Mm^3$ for each day in the coldest period of the year), the industrial sector appears to be different. As it can be evidenced from the analysis of the three following figures, a minimum level of 20 $Mm^3$ for each day is consumed in each day of the year, while an additional amount of maximum 30 $Mm^3$ for each day can be related to the temperature and to the day of the week.

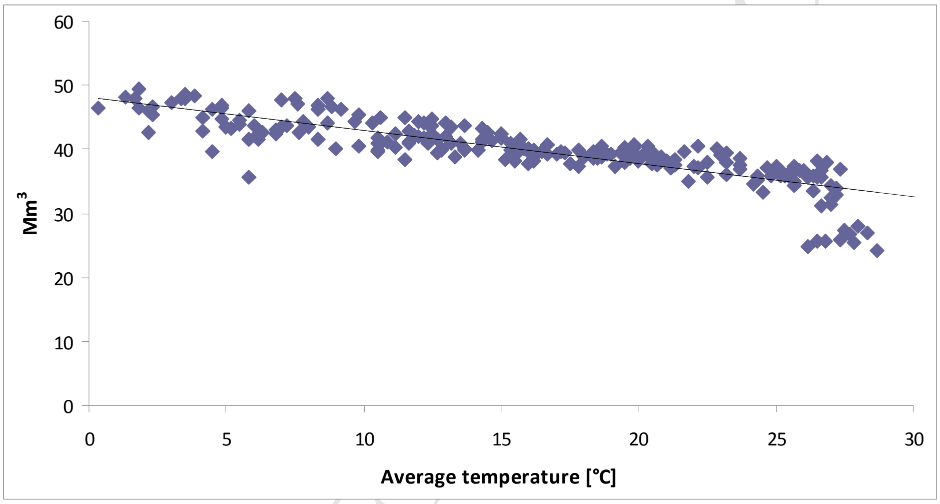



**Fig. 12. Daily natural gas consumption trend in industrial sector during the weekdays**

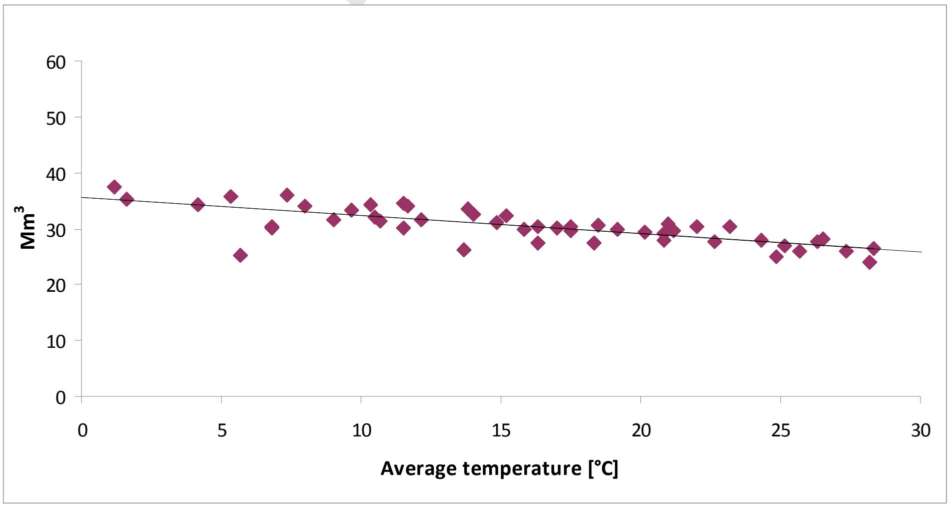



**Fig. 13. Daily natural gas consumption trend in industrial sector during the holidays**

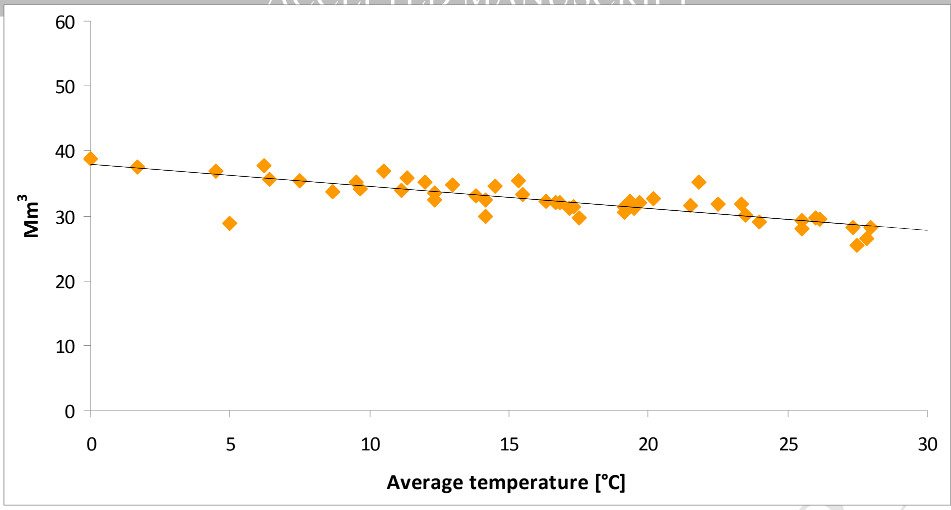



**Fig. 14. Daily natural gas consumption trend in industrial sector in the preholidays**

Analyzing the data of natural gas consumption in the industrial sector, for the modelization an additive model can be proposed for the detailed description: this takes the following form as the function of four components:

$$C = C_n + C_w + C_d + C_R \tag{6}$$

where C is the total consumption, $C_n$ represents the "baseload" component, which is a set of standardized load shapes for each "type" of day that has been identified as occurring throughout the year, $C_w$ represents the weather sensitive part, $C_d$ is a special day component that determines a substantial deviation from the usual load pattern and $C_R$ is a completely random term.

*4.2. Discussion and use of the analysis for thermal energy use forecasting*

The analysis of the aggregated data of natural gas consumption is interesting in itself but it also permits, if it is joined with meteo-data, to obtain information about the use of thermal energy and for defining a model for short (daily) and mid term (weekly) forecasting. In principle, considering the analysis of the particular Italian case, it seems that both for natural gas consumption (strongly linked to the thermal energy end uses) and for thermal energy use in general, a combination of end-use model and econometric model can be proposed as the more appropriate one.

The end-use approach directly estimates energy consumption by using extensive information on end uses and end users. For example end-use models focus on the various uses of thermal energy in the residential, commercial, and industrial sector. Concerning the civil/residential sectors informations about sizes of houses, appliances, customer uses, and so on, are required. Statistical information about customers along with dynamics of change is another additional element for the forecast. In general, weather conditions influence in maximum part both the load and the natural gas consumption. The accuracy of load forecasting depends not only on the techniques, but also on the accuracy of forecasted weather data. Time factors including the period of the year and the day of the week (mainly the day type) and in principle the hour of the day are quite important too. While it is really difficult to elaborate hourly models, as observed analyzing the natural gas consumption data, some remarkable differences can be evidenced between weekdays and weekends, mainly in the industrial sector and less in the civil/residential sector. For this reason, a combination of additive and multiplicative models can be used in order to develop a model for thermal energy use and heat load forecasting. In addition, economic factors such as per capita incomes and employment levels can influence the natural gas consumption and the thermal energy use. For example natural gas prices can be included in econometric models.

## 5. Conclusions

This paper reports a detailed analysis of the natural gas consumption in the Italian market basing on the data of the last four years. Considering the particular link of thermal energy use and natural gas consumption in Italy the work can be considered as a first step of the development of a method for the forecast of thermal energy use of a complex system. The analysis is focused on a specific objective: to analyze and describe the structure of the data and their connection with physical and statistical elements. Among this the season of the year, a given day in the week (weekday, preholidays and holidays) and the various customer categories (civil/residential, industrial and thermoelectric).

The natural gas consumption data referred to one year are analyzed, organized and opportunely clustered. The analysis provides interesting information concerning the trend of natural gas consumption data.

In particular, the natural gas consumption in civil/residential, that represents the main component (up to 65% of the total consumption in the coldest days of the year) appears to be strongly connected with the outside temperature and at low level with the day of the week.

In particular, the link between natural gas consumption data and the weather conditions, mainly represented by an average value of the outside temperature, is analyzed. The particular value that can be considered as a lower boundary limit is 15 °C. The reduction of consumption in the civil/residential sector during in the holidays can be estimated in the amount of 9-12 $Mm^3$ during the coldest period.

The natural gas consumption in the industrial sector appears to be only partially connected with the temperature while a meaningful connection with the day of the week and in particular with the holidays can be evidenced. A variation up to 30% can be evidenced with absolute values of 15 $Mm^3$. Considering that a maximum of daily consumption of less than 48 $Mm^3$ has been observed in 2015, it possible to understand the quantitative meaning of the variation connected to the holidays. The additional amount connected to the thermoelectric production is more difficult to be considered: a variation between about 30 $Mm^3$ and over 100 $Mm^3$ can be evidenced even if a clear connection with external elements is difficult in this case.

The analysis appears to be particularly informative for the Italian market but the methodological approach can be extended to other countries. Some of the indications coming from this work can furnish an important basic element as a first step for the development of a model for thermal energy forecasting, conceived for the definition of thermal energy use consumption in complex and largely extended energy systems.

**References**


Anbazhagan S., Kumarappan N., 2014. Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT. Energy Conversion and Management 78, 711–719.

Ardakani F.J., Ardehali M.M., 2014. Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. Energy 65, 452-461.

Bacher P., Madsen H., Nielsen H., Perers B., 2013. Short-term heat load forecasting for single family houses. Energy and Buildings 65, 101-112.

Bianco V., Scarpa F., Tagliafico L.A., 2014. Analysis and future outlook of natural gas consumption in the Italian residential sector. Energy Conversion and Management 87, 754–764.

Di Castelnuovo M., Fumagalli E., 2013. An assessment of the Italian smart gas metering program. Energy Policy 60, 714-721.

Dirks J.A., Gorrissen W.J., Hathaway J.H., Skorski D.C., Scott M.J., Pulsipher T.C., Huang M., Liu Y., Rice J.S., 2015. Impacts of climate change on energy consumption and peak demand in buildings: a detailed regional approach. Energy 79, 20-32.

Dilaver O, Dilaver Z., Hunt L.C., 2014. What drives natural gas consumption in Europe? Analysis and projections. Journal of Natural Gas Science and Engineering 19, 125-136.

Erdogdu E., 2010. Natural gas demand in Turkey. Applied Energy 87 211–219.

Forouzanfar M., Doustmohammadi A., Menhaj M. B., Hasanzadeh S., 2010. Modeling and estimation of the natural gas consumption for residential and commercial sectors in Iran, Applied Energy 87 268–274.

Feinberg E., Genethliou D., 2005. *Load Forecasting*. State University of New York, Stony Brook. Applied Mathematics for Restructured Electric Power Systems Power Electronics and Power Systems, ISBN 978-0-387-23470-0, Springer, 269-285.

Franco A., Salza P., 2011. Strategies for optimal penetration of intermittent renewables in complex energy systems based on techno-operational objectives. Renewable Energy 36, 743-753.

Hassan S., Khosravi A., Jaafar J., 2015. Examining performance of aggregation algorithms for neural network-based electricity demand forecasting. Electrical Power and Energy Systems 64, 1098-1105.

Hernandez L., Baladron C., Aguiar J.M., Carro B., Sanchez-Esguevillas A., Lloret J., 2014. Artificial neural networks for short-term load forecasting in microgrids environment. Energy 75, 252-264.

Huang Z., Yu H., Peng Z., Zhao M., 2015. Methods and tools for community energy planning: A review. Renewable and Sustainable Energy Reviews 42, 1335-1348.

Kwak Y., Seo D., Jang C., Huh J., 2013. Feasibility study on a novel methodology for short-term real-time energy demand prediction using weather forecasting data. Energy and Buildings 57, 250–260.

Kerr R, Tondro M., 2012. Residential feedback devices and programs: opportunities for natural gas. US Department of Energy, Energy Efficiency & Renewable Energy – Report 2012.

Il Meteo. http://www.ilmeteo.it/portale/archivio-meteo/ [in Italian].

Li J., Xiucheng D., Jianxin S., Mikael H., 2011. Forecasting the growth of China’s natural gas consumption. Energy 36 1380-1385.

Mirasgedis S., Sarafidis Y., Georgopoulou E., Lalas D.P., Moschovits M., Karagiannis F., Papakonstantinou D., 2006. Models for mid-term electricity demand forecasting incorporating weather influences, Energy 31, 208–227.

Nielsen H., Madsen H. 2006. Modelling the heat consumption in district heating systems using a grey-box approach. Energy and Buildings 38, 63-71.

NIST U.S., 2010. Guidelines for smart cyber security: vol. 2, privacy and the Smart Grid, NISTIR 7628.

Powell K.M., Sriprasad A., Cole W.J., Edgar T.F., 2014. Heating, cooling, and electrical load forecasting for a large-scale district energy system. Energy 74, 877-885.

Rouf I., Mustafa H., Xu M., Xu W., Miller R., Gruteser M., 2012. Neighborhood watch: security and privacy analysis of automatic meter reading systems, In: Proceedings of the 2012 ACM conference on computer and communications security. Raleigh, North Carolina, USA: ACM 462–473.

Sabo K, Scitovski R, Vazler I, Zekic´-Sušac M., 2011. Mathematical models of natural gas consumption. Energy Conversion and Management 52, 1721-1727.

Sandels C., Widén J., Nordström L., 2014. Forecasting household consumer electricity load profiles with a combined physical and behavioral approach. Applied Energy 131, 267–278.

Sarak H., Satman A., The degree-day method to estimate the residential heating natural gas consumption in Turkey: a case study. Energy 28 (2003) 929–939.

Schellong W. 2011, *Chapter 5: "Energy Demand Analysis and Forecast*", Energy Management Systems, Dr Giridhar Kini (Ed.), ISBN: 978-953-307-579-2.

Snam Rete Gas. The gas transportation in Italy: day ahead capacity booking, http://www.snamretegas.it/en/services/Thermal_Year_2012_2013/Transportation_capacity/

Soldo B., 2012. Forecasting natural gas consumption. Applied Energy 92, 26–37.

Swan L., Ugursal I.V., 2009. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renewable and Sustainable Energy Reviews 13, 1819-1835.

Suganthi L., Samuel A.A., 2012. Energy models for demand forecasting - A review. Renewable and Sustainable Energy Reviews 16, 1223-1240.

Szoplik J., 2015. Forecasting of natural gas consumption with artificial neural networks. Energy 85, 208-220.

Taspınar F., Celebi N., Tutkun N., 2013. Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods, Energy and Buildings 56, 23–31.

Tewolde M., Longtin J.P., Das S.R., Sharma S., 2013. Determining appliance energy usage with a high-resolution metering system for residential natural gas meters. Applied Energy 108, 363-372.

Vitullo S. R., Brown R. H., Corliss G. F., Marx B. M., 2009. Mathematical Models For Natural Gas Forecasting, Canadian Applied Mathematics Quarterly 17 (4), 807-827.

Yeo I., Yoon S., Yee J.J., 2013. Development of an urban energy demand forecasting system to support environmentally friendly urban planning. Applied Energy 110, 304-317.

Yu W., Li B., Lei Y., Liu M., 2011. Analysis of a Residential Building Energy Consumption Demand Model. Energies, 4, 475-487.

Wadud Z., Dey H.S, Kabir M.A., Khan S., 2011. Modeling and forecasting natural gas demand in Bangladesh, Energy Policy, 39 7372–7380

**Highlights**

1) A preliminary analysis about forecasting thermal energy uses is developed

2) The data of consumption of natural gas in Italy are analyzed and clustered

3) Natural gas consumption data are organized according to different customer profiles

4) The research of some peculiar patterns natural gas consumption data is provided

5) Physical, statistical elements and main features for clustering of the gas data are included