Energy 118 (2017) 231–245

Contents lists available at ScienceDirect

Energy

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

# Day-ahead natural gas demand forecasting based on the combination of wavelet transform and ANFIS/genetic algorithm/neural network model

Ioannis P. Panapakidis [a, b, *], Athanasios S. Dagoumas [b]

[a] Department of Electrical Engineering, Technological Educational Institute of Thessaly, 41110 Larisa, Greece
[b] Energy & Environmental Policy Lab., School of Economics, Business & International Studies, University of Piraeus, 18532 Piraeus, Greece

ARTICLE INFO

*Article history:*
Received 28 July 2016
Received in revised form
5 December 2016
Accepted 8 December 2016
Available online 19 December 2016

*Keywords:*
Artificial neural networks
Computational intelligence
Fuzzy inference
Genetic algorithms
Natural gas demand forecasting

ABSTRACT

Accurate forecasts of natural gas demand can be essential for utilities, energy traders, regulatory authorities, decision makers and others. The aim of this paper is to test the robustness of a novel hybrid computational intelligence model in day-ahead natural gas demand predictions. The proposed model combines the Wavelet Transform (WT), Genetic Algorithm (GA), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Feed-Forward Neural Network (FFNN). The WT is used to decompose the original signal in a set of subseries and then a GA optimized ANFIS is employed to provide the forecast for each subseries. ANFIS output is fed into a FFNN to refine the initial forecast and upgrade the overall forecasting accuracy. The model is applied to all distribution points that compose the natural gas grid of a country, in contradiction to the majority of the literature that focuses on a limited number of distribution points. This approach enables the comparison of the model performance on different consumption patterns, providing also insights on the characteristics of large urban centers, small towns, industrial areas, power generation units, public transport filling stations and others.



## 1. Introduction

Robust forecasting is the fundamental process that the energy system planning is based on [1]. Predictions of key variables like energy demand, market prices, fuel reservoirs are crucial in the design, implementation and evaluation phases of different energy projects. Especially, electricity demand forecasting has held an important role in the power systems community in terms of published researches, pilot programs and practical applications [2]. The literature on forecasting of other variables like natural gas demand is more limited [3]. This is mainly due to the fact that the utilization of natural gas in power generation was rather limited during the last decades. However, the deregulation of energy markets and the low capital cost of natural gas units, compared to coal plants, have created opportunities for market players, leading to numerous investments on natural gas projects [4]. The increasing environmental awareness as well as the importance of energy security have facilitated this trend together with the penetration of renewable energy resources [5]. Other common usages of natural gas are public transportation, manufacturing and residential heating. Hence, natural gas demand is an important carrier in the energy mix and its accurate forecasting is of fundamental importance in many applications.

The respective literature on natural gas demand forecasting can be categorized based on the various criteria such as forecasting horizon, type of tools, covered area and others [3]. Actually the forecasting area defines the types of inputs involved and model used. Our study belongs to the sub-categories of daily forecasting and per distribution level. In Ref. [6] a comparison between various Artificial Neural Networks (ANNs) takes place in order to derive the most effective topology for the natural gas demand prognosis of a city in Poland. Two different types of ANN are used for day-ahead forecasts in a providence of Turkey [7]. The inputs do not consider historical natural gas values but only external variables. Authors of [8] propose a combined model consisted of 2 ANN. To test the models robustness, they provide a comparison between various approaches, including ANN, in order to determine the appropriate model. The target set corresponds to the data provided by 8 natural gas distribution providers. In Ref. [9] the area of

* Corresponding author. Department of Electrical Engineering, Technological Educational Institute of Thessaly, 41110 Larisa, Greece.
*E-mail address:* ipanap@ee.auth.gr (I.P. Panapakidis).

http://dx.doi.org/10.1016/j.energy.2016.12.033
0360-5442/

application is Ankara, Turkey. The inputs that the ANN uses are gas prices, population, selling prices, degree-day and exchange rate. A recurrent neural network is used in Ref. [10] as a gating model of an existing one that is used by a gas utility. Authors of [11,12] are concerned with daily and weekly forecasting. A statistical analysis is conducted to determine which factors are appropriate to serve as input variables. For each forecasting horizon, they compare 7 different ANN topologies that differ in terms of the training function used. In Ref. [13] the proposed model uses 3 ANN working in parallel to provide separate forecasts. Then these forecasts are combined to obtain the final one. Various parameters are used as inputs like calendar information, temperature and historical gas consumption values. The data are obtained from 4 gas distribution companies. Authors of [14] deal with 2 regions in Poland and examine 3 horizons, namely day-ahead, week-ahead and four-week ahead. A comparison takes place between a FFNN and a fuzzy neural network. Also, they test some regression models. In all forecasting horizons, the FFNN outperforms all the rest. In Ref. [15] the GA is incorporated in order to define the optimal number of layers, the weights and bias of an ANN. The proposed model predicts daily gas consumption as a function of degree-day, relative humidity, rainfall, and wind speed and is applied on a city in Iran. In Ref. [16] the GA is used to produce the optimal weights and threshold values. The proposed model is tested with several others on the data corresponding to a city in China. Day-ahead forecasting of the consumers demand of a local distribution company in Croatia is the subject of [17]. There are several linear and non-linear models that are used for the aforementioned purpose. Among them, the authors include a recursive neural network and a FFNN. The comparison denotes the recursive linear autoregressive model with exogenous inputs as the optimal tool for the data under study. In Ref. [18] the Multilayer Perceptron (MLP) is compared with ANFIS. The latter leads to better predictions in the 2 examined cases. These differ in terms of the number and types of inputs used for both models. The authors also present, according to a literature survey, the advantages of ANFIS over traditional models. In Ref. [19] the authors deal with national wide forecasting of 6 countries of South America. The inputs include the population and the gross domestic product of the country and the output refer to the total annual gas demand of the country.

Based on the above brief literature survey, it is obvious that the natural gas demand forecasting problem has been tackled by many approaches involving sole and hybrid forecasting systems and algorithms for advancements in neural network training. Contrary to the majority of the literature that focuses on a limited number of distribution points, the paper aims as developing a generic methodology that is applied to all distribution points that compose the natural gas grid of a country. This approach enables the comparison of its performance on different consumption patterns, providing also insights on the characteristics of large urban centers, small towns, industrial areas, power generation units, public transport filling stations and others.

The consideration of the whole distribution points of a national system, lead to a large variety of demand patterns both in shapes and magnitude, in order to capture geographical differences and different types of consumers. This inherent attribution of the natural gas demand pattern of the country leads to difficulties in modeling and prediction of future demand trends. Thus, a potential forecasting system should be able to capture and simulate the non-linear behavior of the natural gas demand with regard to the parameters that influence it. A promising category of tools belong to the general technical field of computational intelligence. ANN based models have found numerous applications in a variety of engineering tasks and are a favorable tool in forecasting problems [20]. The paper develops a robust natural gas demand forecasting model, combining the Wavelet Transform (WT), Genetic Algorithm (GA), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Feed-Forward Neural Network (FFNN) approaches. The WT is used to decompose the original signal in a set of subseries and then a GA optimized ANFIS is employed to provide the forecast for each subseries. ANFIS output is fed into a FFNN to refine the initial forecast and upgrade the overall forecasting accuracy. The model is used for the day ahead natural gas demand forecasting of Greece.

The objectives of the present paper can be summarized in the following:

a) The paper aims at providing forecasts for all distribution points that consist the national natural gas grid of a whole country. To the best of the authors, knowledge, this is the first study that examines all the distribution points corresponding to the largest amount of data in the respective literature.
b) A novel hybrid forecasting model is proposed. We combine an ANFIS with a FFNN to construct a cascaded model that leads to high accuracy in most distribution points. The proposed model is characterized by high flexibility, comprehensive operation and low execution time requirements.

Our paper serves as an initial study focusing on the national gas distribution system of a country by examining the consumptions pattern per distribution point. The findings of this work can be used by distribution operators or other utilities to upgrade their operational long-term planning and validate their decision making procedure when dealing with emergencies, unit constructions, gas network expansion and other issues.

## 2. Natural gas demand forecasting framework

The paper investigates the consumption patterns of all distribution points that compose the national natural gas system of Greece [21]. Totally there are 44 distribution points, from which 4 are not considered due to zero or extremely low consumption, reducing the number of points under investigation to 40. The distribution points are spread over the Greek territory, covering different types of loads. In our study, a novel computational intelligence model is constructed and tested separately on the distribution points of the Greek system. The available data set corresponds to aggregated daily consumption and concerns the period 01/01/2014-30/06/2016. The data set is split to training and test sets. The training set corresponds to the period 01/01/2014-31/12/2015 and is employed so that to define the optimal model's parameters. The test set covers the period 01/01/2016-30/06/2016 and is used for the overall assessment of the trained model. We are concerned with day-ahead prediction, i.e. the purpose is to estimate next day's demand of each distribution point using historical natural gas consumption and other influential variables.

### 2.1. Data description

Fig. 1 displays the National Natural Gas (NNGTS) Transmission System of Greece. The NNGTS transports gas from the Greek-Bulgarian Border and Greek-Turkish border in continental Greece. The main transmission pipeline with total length equals to 512 km and design pressure 70 barg extends from the Greek-Bulgarian border at Promachonas to Attica. Transmission branches with total length 947 km extend from the main transmission pipeline and supply natural gas to the regions of Eastern Macedonia, Thrace, Thessaloniki, Platy, Volos, Trikala, Oinofyta, Antikyra, Aliveri, Korinthos, Megalopoli, Thisvi and Attica. The data used in this study were provided by The National Natural Gas System Operator (DESFA) S.A. The purpose of the company is the operation,

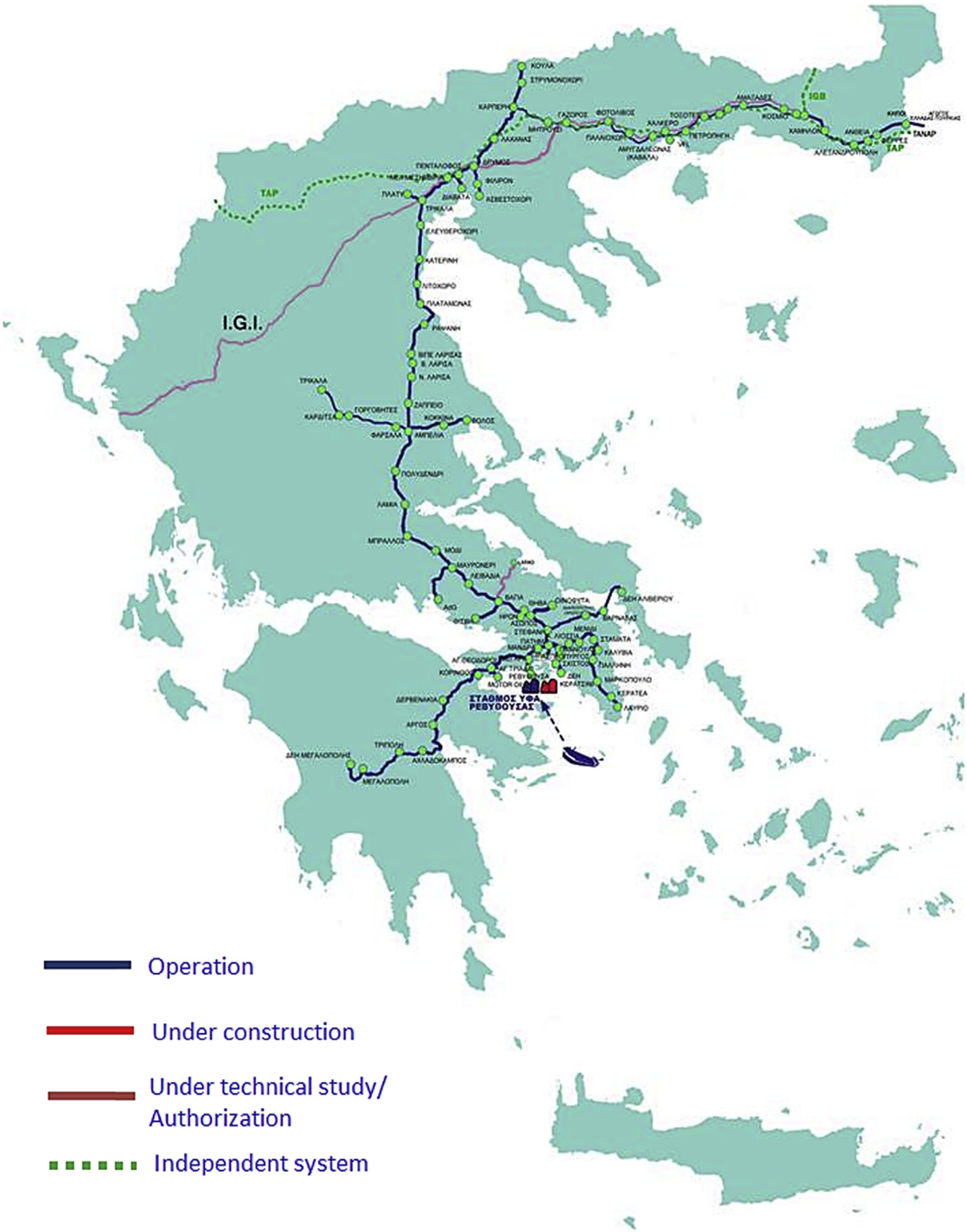


**Fig. 1.** National natural gas transmission system [21].

management, exploitation and development of the NNGTS in order to cover the needs of the consumers in a safe, reliable and economically efficient way. The available data refer to large urban areas, towns, industrial areas and other special consumers [21]. The number of data points is 40 and can be categorized in the following types of consumers:

- Large cities: Athens, Alexandroupolis, Volos, Thessaloniki, Katerini, Komotini, Lamia, Larissa, Xanthi, Serres, Trikala.
- Towns: Agia Triada, Sidirokastro, Kipi, Agioi Theodoroi, Drama, Thriasio, Karditsa, Kilkis, Kokkina, Spata, Oinofyta, Platy.
- Industrial areas: ELPE, VIPE Larissa, VFL, Motor Oil, Motor Oil II.

- Power generation units: Alouminion, Alouminion II, Alouminion III Aliveri (PPC), Komotini (PPC), Lavrio (PPC), Energeiaki Thess. (ELPE), Heron II, Heronas, Thisvi.
- Public transport supplementation stations: SALFA Anthousa, SALFA Ano Liossia.

As an initial exploration of the data, a correlation analysis takes place. The distribution point of Athens serves as the basis of the analysis. Fig. 2 shows the correlation between the time series of Athens with respect to the rest. The time series refer to the training period. With the correlation analysis we check the degree of similarity between the time series shapes between the distribution points. The higher correlations to Athens have the points of Alexandroupolis, Volos, Thessaloniki, Karditsa, Larisa, Serres and Trikala. It can be concluded that most city center present similar consumption patterns. Aloumimion and Heronas display the lowest similarities. This points refer to privately owned natural gas fired generation units.

### 2.2. Wavelet transform

In many forecasting problems, prior to entering the data into the model for training and test, a pre-processing phase takes place. The presence of atypical data such as outliers, extremely low values and others influence the prediction accuracy. However, in real-time applications, for instance in electricity grid operators the forecasting refers to out-of-sample data. This means that the test data are not available and no pre-processing is applicable. In our study no pre-processing took place since there were no mission values. In many points the consumption of some day was zero. This fact, as it will be denoted in a following Section, influences the forecasting accuracy. We conducted a preliminary analysis of the data and found that in many cases showed volatilities and trends. To overcome this fact, the WT is applied to split up the original gas consumption series into one low-frequency and some high-frequency subseries in the wavelet domain [22]. These subseries present a better behavior compared to the original signal and thus their forecast will lead to lower error. The WT provides a filter to the original series.

Generally, WTs are distinguished in Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). Let $f(x)$ and $\Phi(x)$ be the original series and a mother wavelet, respectively. The CWT $W(a,b)$ of $f(x)$ is expressed as:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x)\Phi\left(\frac{x-b}{a}\right)dx \tag{1}$$

where the scale parameter $a$ controls the spread of the wavelet and the translation factor $b$ determines its central position. The wavelet representation of $f(x)$ with respect to the mother wavelet $\Phi(x)$ refers to the set of all wavelet coefficients $W(a,b)$. The CWT is accomplished by continuously scaling and translating the mother wavelet. But this concept may lead to increased redundant information. An alternate to this, is to consider certain scale, an approach known as DWT. In the DWT, each coefficient $W(m,n)$ is expressed as:

$$W(m,n) = 2^{-\left(\frac{m}{2}\right)} \sum_{t=0}^{T-1} f(t)\ \Phi\left(\frac{t-n2^m}{2^m}\right) \tag{2}$$

where $T$ is the length of the signal $f(t)$ and $t$ is the discrete time index. A fast DWT has been proposed in Ref. [23]. It consists of 4 filters: decomposition low-pass, decomposition high-pass, reconstruction low-pass, and reconstruction high-pass filters. This approach leads to approximation (which is the low-frequency representation) and details (the difference between the high-frequency representations) of the original series. The original

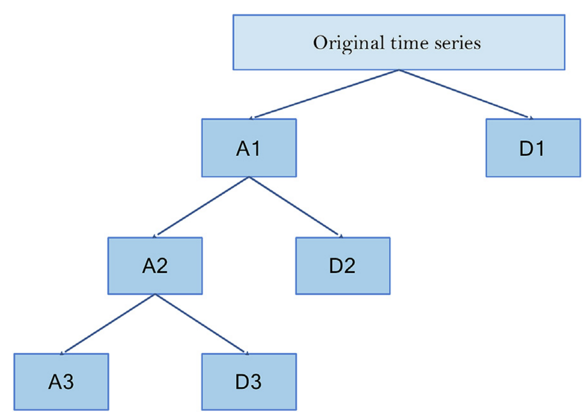


**Fig. 3.** Multi-level decomposition process.

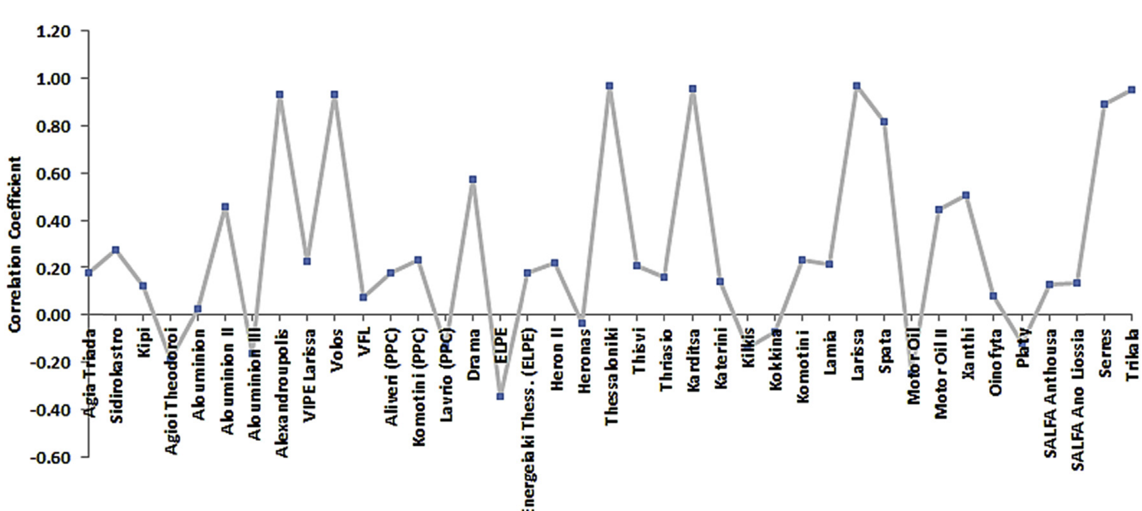


**Fig. 2.** Correlation coefficient values between Athens the other distribution points under study.

series is split by successive decompositions into lower resolution components. The process is depicted in Fig. 3. The original series is the sum of the low-frequency component A3 and the high-frequency components D1, D2, D3, i.e. $f = A3 + D1 + D2 + D3$. In our study, the wavelet function of type Daubencies of order 4 serves as the mother wavelet $\Phi(t)$.

As illustrative examples of the WT, Figs. 4 and 5 show the original series and the low- and high-frequency components of Athens and ELPE, respectively. The wavelet components have the same length with the original series. The component A3 is an approximation of the original signal.

### 2.3. GA/ANFIS/FFNN approach

A crucial study on the overall success of a forecasting task is the proper selection of the number and types of inputs. The historical values of the parameter under study are input candidates. In order to explore the natural gas consumption series periodicity, the Pearson correlation coefficient is used to measure the degree of dependence between current values and values up to 30 days before [24]. The correlation analysis is held for each wavelet component separately. The results for the series of Athens are presented in Fig. 6. In order to form a reduced set of inputs for the purpose of faster training, only the first two most correlated values are selected: for component D1 the selected days are 3 and 4, for component D2 the selected days are 1 and 7 and both for component D3 and A3, the selected days are 1 and 2. All the other inputs have been chosen via trial and error. A series of different models were examined in order to reach safe conclusions about the inputs sets that need to be considered for the current data set. According to the experimental outputs, the mean daily temperature is also selected. The temperature data are obtained by the nearest to the distribution gas point station. Finally, calendar indicators are considered. Specifically, we use 1 input for months and days coding, respectively. Let $k = 1,2,\ldots,12$ be the number of months. We considered the following matching: January→1, February→2,…,December→12. Let $l = 1,2,..,7$ be the number of days. The day type matching is the following: Monday→1, Tuesday→2,…,Sunday→7.

A schematic representation of the proposed hybrid model is depicted in Fig. 7. The original gas consumption series are decomposed into the wavelet components. From the correlation analysis, the first two most correlated values are selected. Also, the external variables (mean daily temperature and calendar indicators) are used to form the input pattern of the ANFIS. The output refers to the total daily demand of the specific distribution point. The GA is employed to optimize the parameters of ANFIS. The output of ANFIS with the same external variables are led to the input layer of the FFNN in order to refine the initial forecast. The 4 ANFIS/GA/FFNN models are run in parallel. In the last layer of the model, the inverse WT takes place and in order to obtain the predicted gas series. Fig. 8 shows the general structures of FFNN and ANFIS that are combined to formulate the forecasting engine. The FFNN is trained by the Levenberg-Marquardt algorithm [25]. FFNNs are built with information processing units, i.e. the neurons. The neuron is composed by a set of synapses and a summation junction fired by an activation function. The neuron is fed with a set of discrete signals that are modified by the weights of the synapsed. The role of the junction is to sum all the modified signals. It generates an output based on the form of the activation function of the neurons. For full mathematical description, the reader is referred to [26].

ANFIS is composed by 5 layers and each layer contains several nodes [27]. The nodes are described by a node function. Let $O_i^j$ be the output of the $i$th node in layer $j$. In the 1st layer, every node I is an adaptive node with node function:

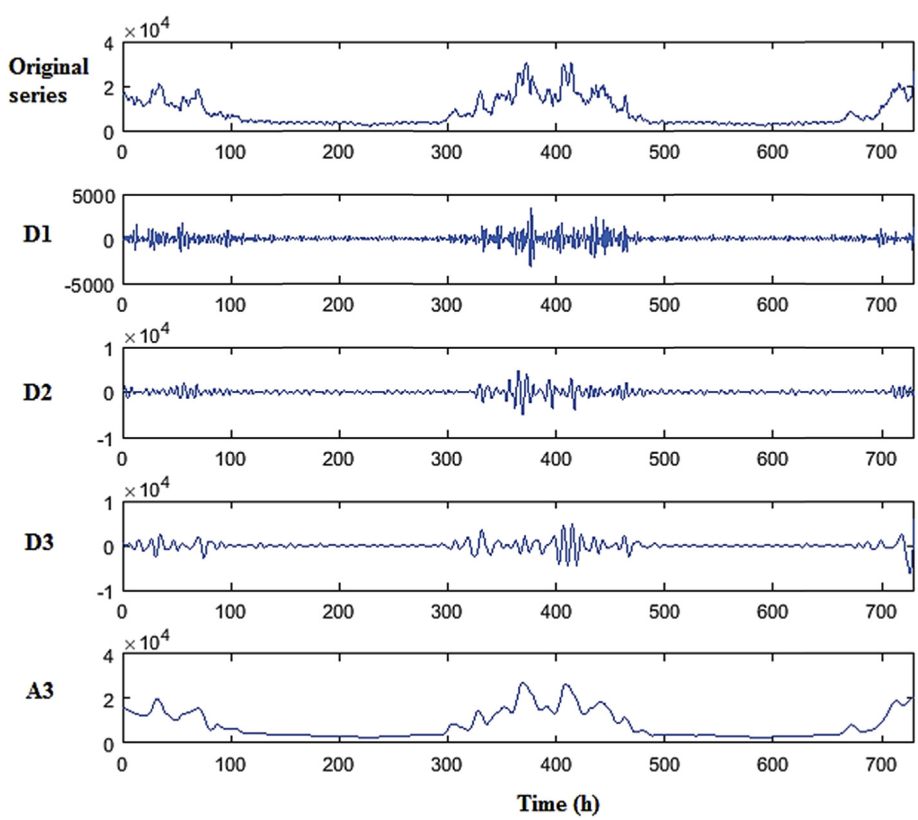


**Fig. 4.** Original series and wavelet components of the training period of Athens.

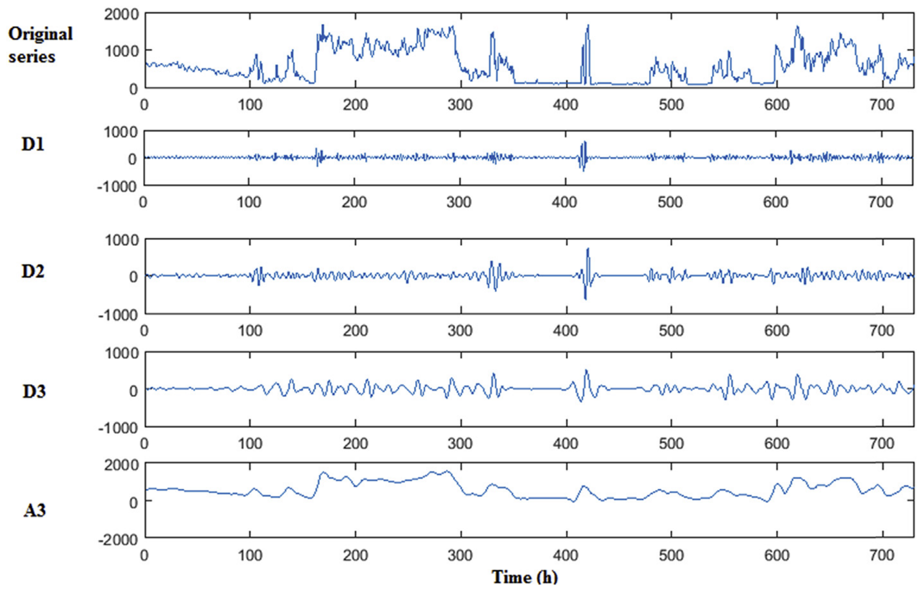


**Fig. 5.** Original series and wavelet components of the training period of ELPE.

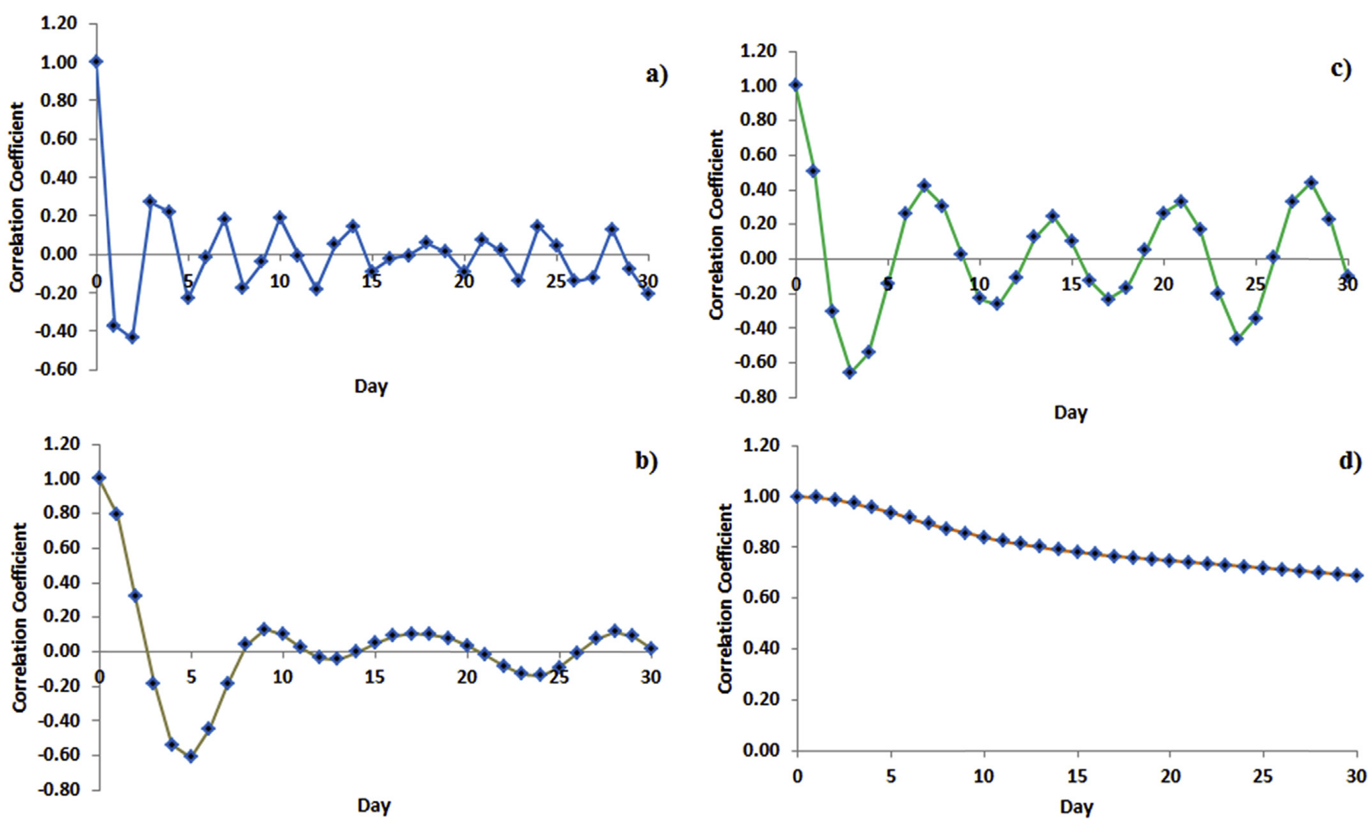


**Fig. 6.** Correlation coefficient between the current value up to the value of 30 days before for a) D1, b) D2, c) D3 and d) A3 wavelet components.

$$O_i^1 = \mu A_i(x), \quad i = 1, 2 \tag{3}$$

or

$$O_i^1 = \mu B_{i-2}(y), \quad i = 3, 4 \tag{4}$$

where $x$ or $y$ is the input of the $i$th node and $A_i$ or $B_{i-2}$ is a linguistic label associated with the node. Hence, $O_i^j$ is the membership grade of a fuzzy set $A_1, A_2, B_1$ or $B_2$ and it specifies the degree to which the input $x$ or $y$ satisfies the quantifier $A$ or $B$. Any continuous and piecewise differential function can be used for node functions in the 1st layer. In the 2nd layer, each node $\Pi$ multiplies the inputs and sends the product in output:

$$O_i^2 = w_i = \mu A_i(x)\mu B_i(y), \quad i = 1, 2 \tag{5}$$

In the 3rd layer, each node N computes the following ratio:

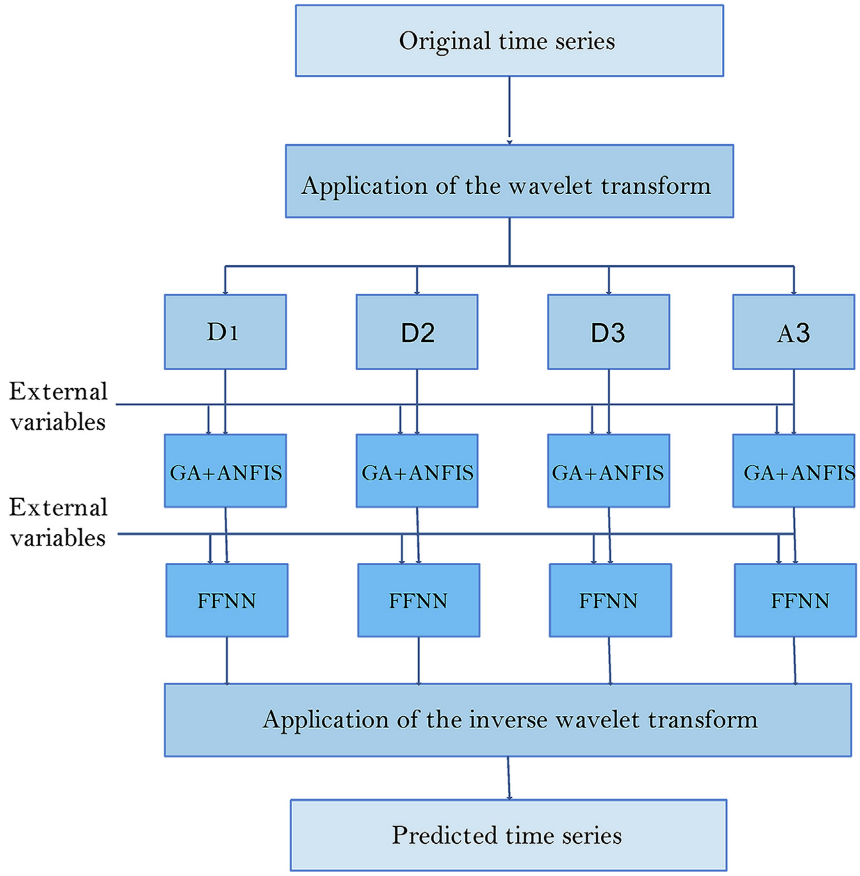


**Fig. 7.** Schematic representation of the proposed model.

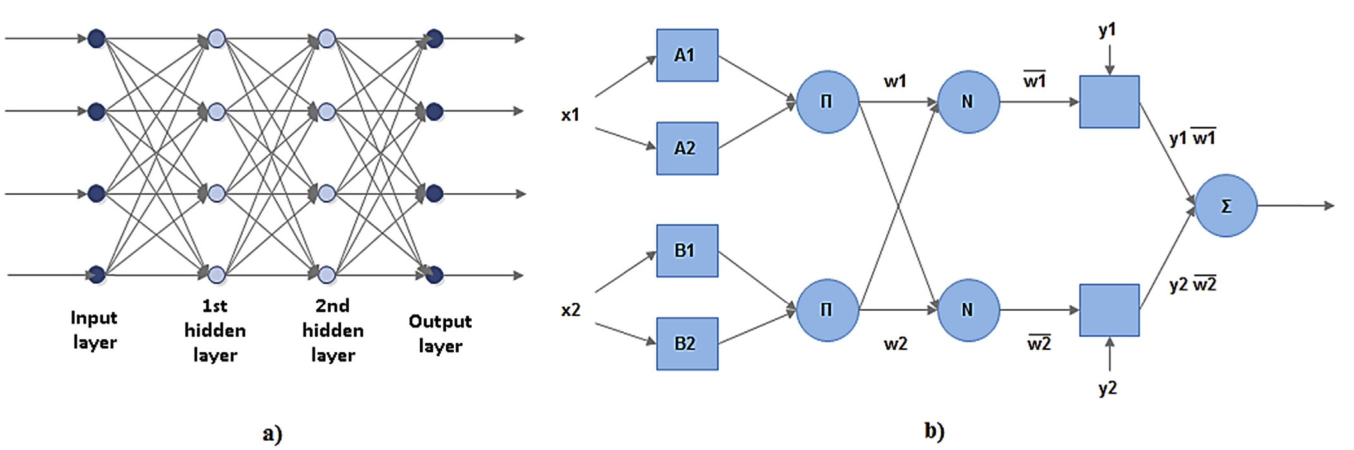


**Fig. 8.** Illustrations of the structures of a) FFNN and b) ANFIS.

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{6}$$

In the 4th layer, each node computes the contribution of the $i$th rule to the overall output:

$$O_i^4 = \overline{w}_i z_i = \overline{w}_i (a_i x + b_i y + c), \quad i = 1, 2 \tag{7}$$

where $\overline{w}_i$ is the output of the third layer and $a_i, b_i, c$ are a set of parameters.

Finally, in the 5th layer, the node Σ computes the final output as the summation of all inputs:

$$O_i^5 = \sum_i \overline{w}_i z_i \tag{8}$$

GA is well-known evolutionary optimization technique with numerous applications. In this study is used to define ANFIS parameters. More specifically, the operational steps of the proposed model can be summarized in the following:

1) Determination of the ANFIS structure. In this step, the inference mechanism is defined, i.e. whether we will consider Sugeno-type or Mamdani-type fuzzy inference system. Also, the type of membership function is decided, i.e. whether we will use Triangular-shaped, Trapezoidal-shaped, Gaussian or others. Other parameters are number of training epochs, error function threshold between subsequent epochs and others.
2) Determination of the GA parameters such as maximum iteration, population size, mutation probability, crossover probability, mutation rate and others.
3) Formulation of the GA's fitness function. The optimization problem has the following decision variables: number of membership functions and values of $a_i, b_i, c$.
4) Prediction of the WT component via the optimized ANFIS structure.
5) Determination of the FFNN parameters such as number of training epochs, number of hidden layers, number of neurons in the hidden layer, type of activation function and others.
6) Prediction of the WT component and inverse WT to obtain the final series.

Fig. 9 shows the flow-chart of the function of the hybrid model. The optimal FFNN parameters are determined via trial and error set of simulation. If the forecasting error is not acceptable, the GA parameters can be altered and run the optimization procedure again. Also, the FFNN parameters can be changed accordingly.

## 3. Simulation results

### 3.1. Evaluation framework

The evaluation framework consists of a set of mathematical criteria that measure the forecasting errors. A relatively large of validity indicators supports the justification of a models usage. Let $P_m^a$ and $P_m^f$ are the real and predicted natural gas demand of the $m$-th day of the test set, $m = 1,2,\ldots,M$ and $M = 182$, respectively. The

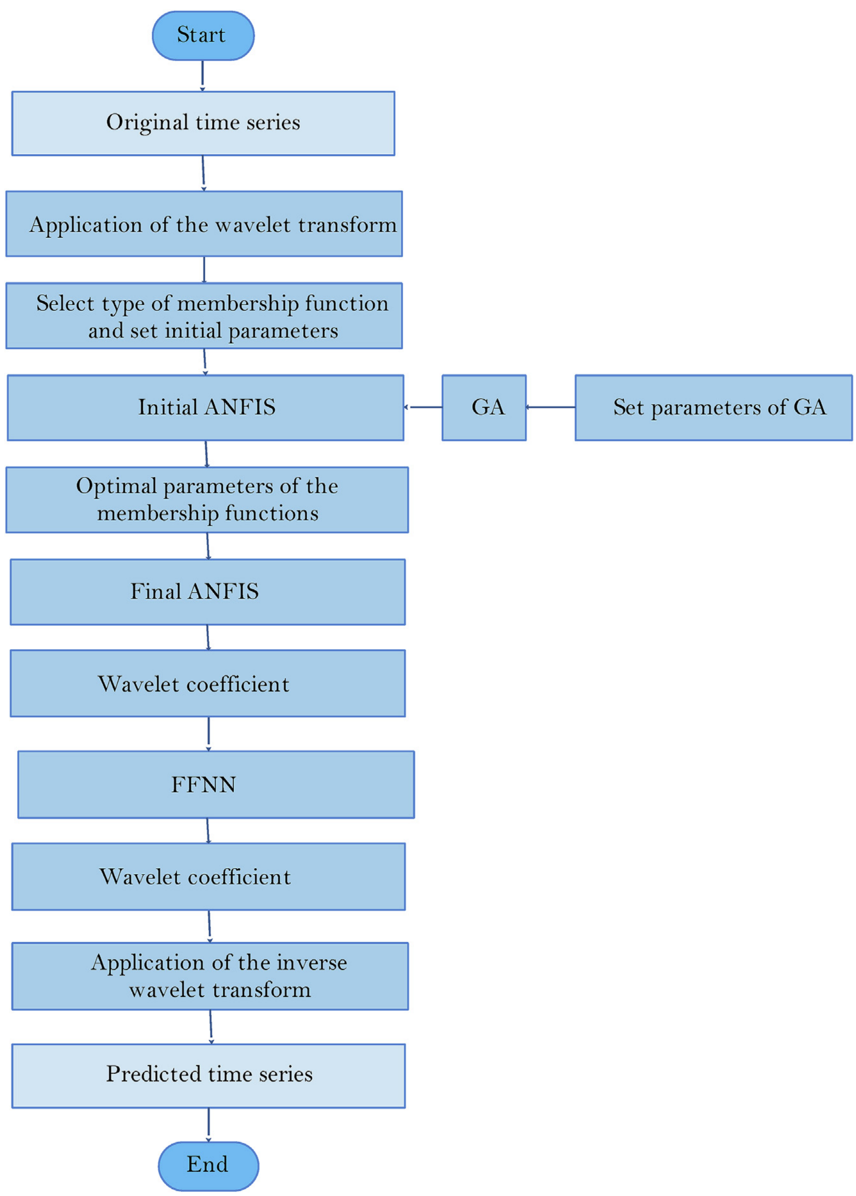


**Fig. 9.** Flow-chart of the operation of the proposed model.

Absolute Error (AE) is defined as:

$$AE = \sum_{m=1}^{M} \left| P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \right| \tag{9}$$

The Mean Absolute Error (MAE) corresponds to the sum all AEs:

$$MAE = \frac{1}{M} \sum_{m=1}^{M} \left| P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \right| \tag{10}$$

The Mean Absolute Percentage Error (MAPE) is a common indicator in forecasting problems and it is given by:

$$MAPE = \frac{1}{M} \sum_{m=1}^{M} \frac{\left| P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \right|}{P_m^{\mathrm{a}}} \times 100 \tag{11}$$

Another common indicator is the Root Mean Squared Error (RMSE), which is expressed as:

$$RMSE = \frac{1}{M} \sqrt{\sum_{m=1}^{M} \left( P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \right)^2} \tag{12}$$

The Mean Absolute Range Normalized Error (MARNE) is the absolute difference between the real and forecast natural gas demand, normalized to the maximum gas demand [28]:

$$MARNE = \frac{1}{M} \sum_{m=1}^{M} \frac{\left| P_m^{\mathrm{a}} - P_m^{\mathrm{f}} \right|}{\max\left(P_m^{\mathrm{a}}\right)} \times 100 \tag{13}$$

In order to validate the proposed model, a comparison will take place with a model without the FFNN part, i.e. hybrid model composed by ANFIS/GA. All the above indicators have been calculated using the test data.

**Table 1**
Error measures per distribution point generated by the ANFIS/GA model.

| | MAE | MAPE | RMSE | MARNE | | MAE | MAPE | RMSE | MARNE |
|---|---|---|---|---|---|---|---|---|---|
| Agia Triada | 1366.92 | >100 | 1802.52 | 7.35 | Thessaloniki | 2147.44 | 16.20 | 21962.28 | 8.67 |
| Sidirokastro | 78639.56 | >100 | 150227 | 64.90 | Thisvi | 9841.79 | >100 | 45526.76 | 56.90 |
| Kipi | 2313.51 | 14.81 | 14549.22 | 8.97 | Thriasio | 28.18 | 8.64 | 37.79 | 5.12 |
| Agioi Theodoroi | 45.12 | >100 | 145.51 | 20.23 | Karditsa | 200.55 | 31.99 | 2347.31 | 20.05 |
| Alouminion | 479.04 | 5.58 | 1830.22 | 4.40 | Katerini | 10.77 | 4.16 | 16.90 | 3.33 |
| Alouminion II | 2189.79 | 36.46 | 7713.85 | 13.08 | Kilkis | 65.21 | 13.39 | 81.21 | 4.75 |
| Alouminion III | 73.33 | 3.64 | 115.70 | 2.35 | Kokkina | 10.69 | >100 | 38.20 | 8.32 |
| Athens | 750.32 | 7.59 | 1989.08 | 2.56 | Komotini | 10410.11 | >100 | 42959.69 | 47.42 |
| Alexandroupolis | 10.25 | 8.47 | 61.52 | 5.18 | Lamia | 155.56 | >100 | 676.43 | 64.41 |
| VIPE Larissa | 15.65 | 14.70 | 20.19 | 7.48 | Larissa | 142.33 | 8.69 | 327.15 | 2.50 |
| Volos | 218.52 | 11.54 | 494.50 | 4.15 | Spata | 89.84 | 27.22 | 612.00 | 18.35 |
| VFL | 367.97 | >100 | 1078.93 | 7.48 | Motor Oil | 315.78 | >100 | 596.92 | 4.25 |
| Aliveri (PPC) | 2287.06 | >100 | 8354.39 | 12.30 | Motor Oil II | 11965.78 | >100 | 40328.37 | 64.00 |
| Komotini (PPC) | 2204.44 | >100 | 5122.09 | 13.88 | Xanthi | 16.82 | 17.69 | 22.85 | 4.90 |
| Lavrio (PPC) | 1523.51 | >100 | 2147.52 | 5.92 | Oinofyta | 199.49 | 8.43 | 427.04 | 5.46 |
| Drama | 26.85 | 3.41 | 61.77 | 2.17 | Platy | 20.07 | 7.52 | 43.83 | 4.57 |
| ELPE | 135.30 | 82.37 | 411.36 | 6.23 | SALFA Anthousa | 54.45 | 28.68 | 208.67 | 17.67 |
| Energeiaki Thess. (ELPE) | 2585.53 | >100 | 12049.59 | 14.80 | SALFA Ano Liossia | 25.14 | 31.27 | 45.54 | 6.58 |
| Heron II | 1831.61 | >100 | 4061.27 | 13.38 | Serres | 39.61 | 8.23 | 214.95 | 3.46 |
| Heronas | 80.81 | >100 | 341.20 | 7.97 | Trikala | 9841.79 | >100 | 45536.76 | 56.90 |

**Table 2**
Error measures per distribution point generated by the ANFIS/GA/FFNN model.

| | MAE | MAPE | RMSE | MARNE | | MAE | MAPE | RMSE | MARNE |
|---|---|---|---|---|---|---|---|---|---|
| Agia Triada | 1285.07 | >100 | 1763.83 | 6.91 | Thessaloniki | 585.49 | 7.67 | 1534.56 | 2.36 |
| Sidirokastro | 7173.08 | 12.97 | 13558.93 | 5.92 | Thisvi | 1550.32 | >100 | 2195.03 | 8.96 |
| Kipi | 1147.02 | 9.69 | 2134.97 | 4.45 | Thriasio | 28.67 | 8.67 | 37.63 | 5.21 |
| Agioi Theodoroi | 21.17 | >100 | 33.59 | 9.49 | Karditsa | 21.29 | 11.47 | 97.04 | 3.13 |
| Alouminion | 318.05 | 3.63 | 599.69 | 2.92 | Katerini | 9.89 | 3.64 | 15.00 | 3.05 |
| Alouminion II | 1449.16 | 28.18 | 2106.12 | 8.65 | Kilkis | 60.18 | 11.63 | 76.83 | 4.38 |
| Alouminion III | 72.19 | 3.58 | 114.76 | 2.32 | Kokkina | 10.29 | >100 | 24.76 | 8.01 |
| Athens | 667.51 | 6.60 | 1802.53 | 2.28 | Komotini | 1618.33 | >100 | 2943.51 | 7.37 |
| Alexandroupolis | 5.11 | 4.86 | 12.48 | 2.58 | Lamia | 16.55 | 22.33 | 21.00 | 6.85 |
| VIPE Larissa | 14.55 | 14.87 | 18.59 | 6.96 | Larissa | 132.77 | 8.17 | 242.68 | 2.33 |
| Volos | 161.47 | 9.42 | 287.30 | 3.06 | Spata | 16.91 | 7.60 | 24.42 | 3.45 |
| VFL | 310.10 | >100 | 768.50 | 6.30 | Motor Oil | 314.17 | >100 | 632.22 | 4.23 |
| Aliveri (PPC) | 1435.48 | >100 | 2420.67 | 7.72 | Motor Oil II | 2265.99 | >100 | 3132.46 | 12.12 |
| Komotini (PPC) | 1909.63 | >100 | 3457.55 | 8.69 | Xanthi | 16.49 | 16.60 | 21.47 | 4.81 |
| Lavrio (PPC) | 1595.21 | >100 | 2261.37 | 6.20 | Oinofyta | 166.34 | 8.05 | 214.15 | 4.55 |
| Drama | 23.07 | 3.01 | 37.96 | 1.87 | Platy | 22.58 | 9.32 | 38.69 | 5.14 |
| ELPE | 92.41 | 53.96 | 179.70 | 4.26 | SALFA Anthousa | 14.63 | 12.43 | 20.99 | 4.74 |
| Energeiaki Thess. (ELPE) | 1801.63 | >100 | 2513.27 | 10.31 | SALFA Ano Liossia | 18.73 | 23.17 | 27.10 | 4.90 |
| Heron II | 1642.25 | >100 | 2328.45 | 12.00 | Serres | 23.89 | 5.89 | 77.21 | 2.08 |
| Heronas | 104.36 | >100 | 429.61 | 10.30 | Trikala | 33.79 | 13.67 | 79.10 | 3.92 |

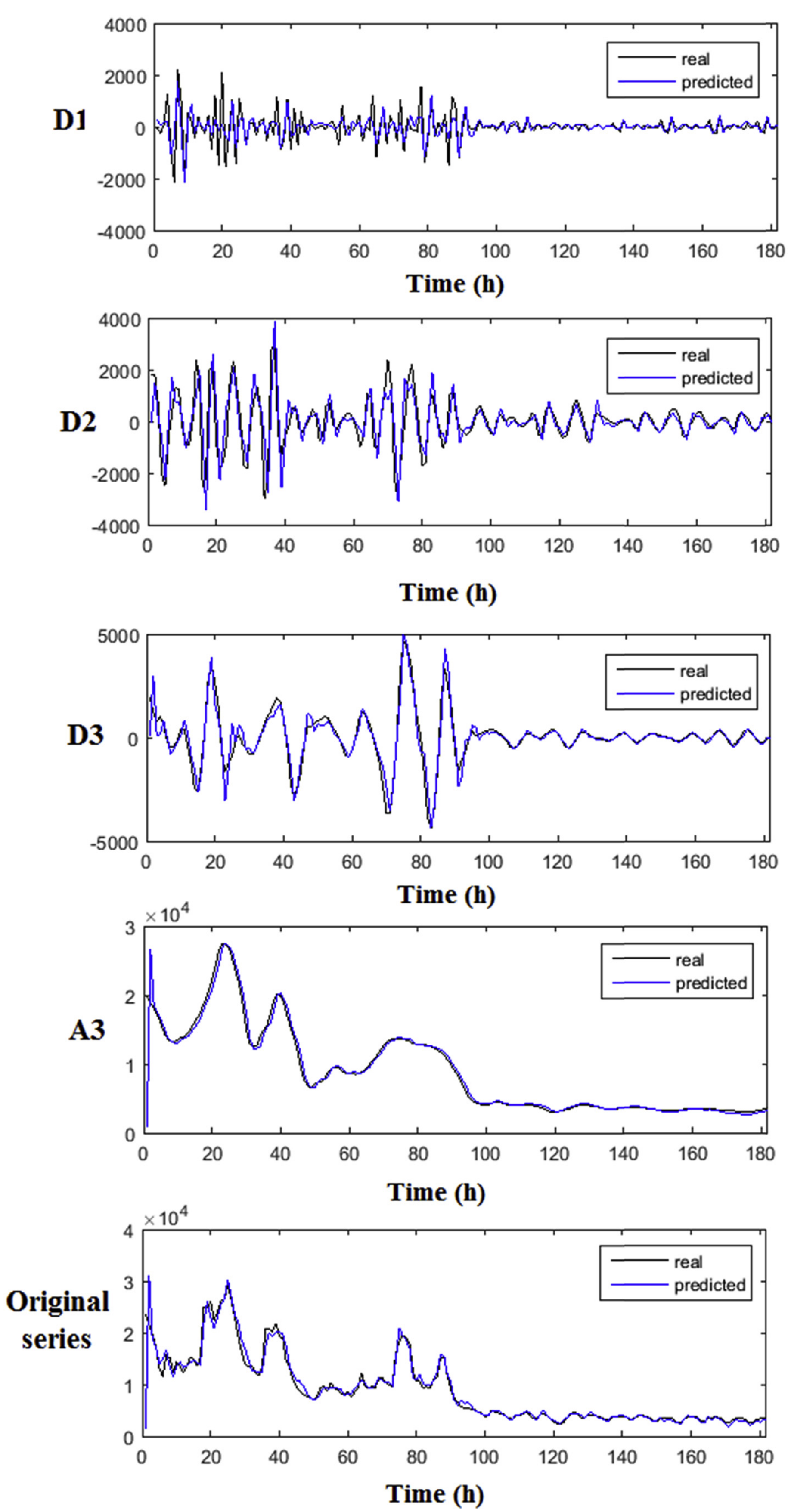


**Fig. 10.** Real and predicted series (original and WT components) of Athens.

### 3.2. Comparison of the models

For both models, the basic GA parameters were set to: maximum iteration→800, population size→25, mutation probability→70%, crossover probability→40% and mutation rate→15%. The basic FFNN parameters were set to: number of hidden layers→1, activation function in the hidden layer→ tangent sigmoid, activation function in the output layer→ tangent sigmoid and maximum number of epochs→500. The FFNN was executed for variable number of neurons in the hidden layer and more specifically, from to 2 to 30 with an increasing step equals to 2. In most cases, the lowest error was met in different number of neurons for every distribution point. Also other configurations were tested (i.e. using 2 hidden layers and other activation functions) but they were rejected due to their poor forecasting efficiency.

When dealing with large data sets, a working practice should be followed to systemize the whole forecasting procedure. This practice is built on three phases. The first one is the "pro-forecasting" phase and refers to an initial statistical analysis of the data to study descriptive statistics, trends and others. The atypical data should be tracked and if necessary, removed from the data set. The analyst should consult the data provider for potential special attributes of the data and/or elaborate data mining processes. The second phase is the "main-forecasting" phase. The analyst should choose to construct a special model according to the needs or reproduce a model proposed in the literature. Most codes are available from commercial software packages, which have user friendly environment and are easily executed in common computer systems. The last phase is the "meta-forecasting" phase. Here the analyst should seek the expert's (i.e. utility, grid operator, etc.) feedback. If expertise knowledge is available, it should be used accordingly. The experimental outcomes should be presented to the expert for further insights and recommendations.

Tables 1 and 2 register the error metrics per distribution point considering the ANFIS/GA and the ANFIS/GA/FFNN models, respectively. It can be concluded that there is a large diversity of errors among the distribution points. In all cases the proposed model lead to lower errors, a fact that denotes its robustness and exploitability considering different data sets. The models performances are close to each other in the cases of Agia Triada, Alouminion III, Athens, VIPE Larissa, Thriasio, Katerini, Kilkis, Kokkina, Larisa, Motor Oil, and Xanthi. These points cover all the types that are mentioned in Section 2.1. The proposed model has more evident superior performance compared to ANFIS/GA in small towns like Agioi Theodoroi, Thirsvi and others. Generally, city centers correspond to the lowest errors followed by the towns. Among the generation units, the privately owned units Alouminion, Alouminion II and Alouminion III have more stable demand patterns. All the Public Power Corporation S.A. (PPC S.A.) stations have less predictable patterns. The natural gas demand at the power generation is linked to the dynamics of the wholesale electricity market. Therefore, a more robust forecast of natural gas demand in power generation should elaborate more specific approaches, such as the unit commitment and the power systems expansion planning problems [29,30]. According to Table 2, the prediction of the demand of transport supplementation points is relatively satisfactory. The industrial area of VIPE Larisa results in medium sized errors contrary to the industrial consumer VFL.

There is a variety of consumption levels among the points and this is reflected by the large divergences of the indicators especially in MAE and RMSE. These indicators measure directly the sum of differences of the predicted and real demand value per day in the

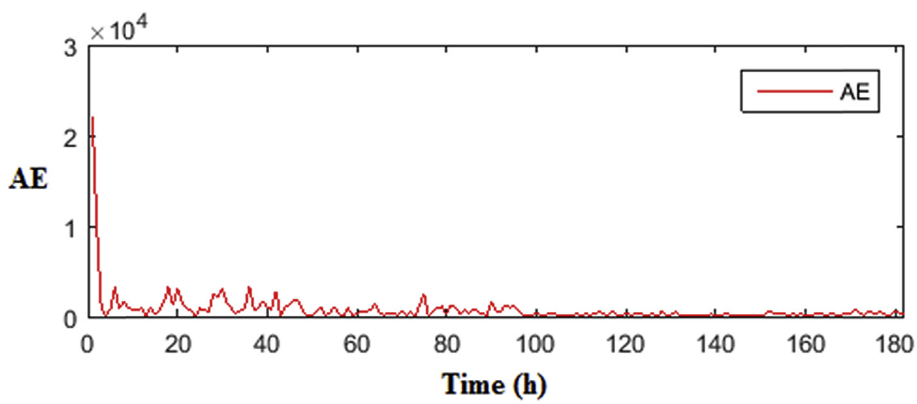


**Fig. 11.** Absolute Errors per day referring to Athens.

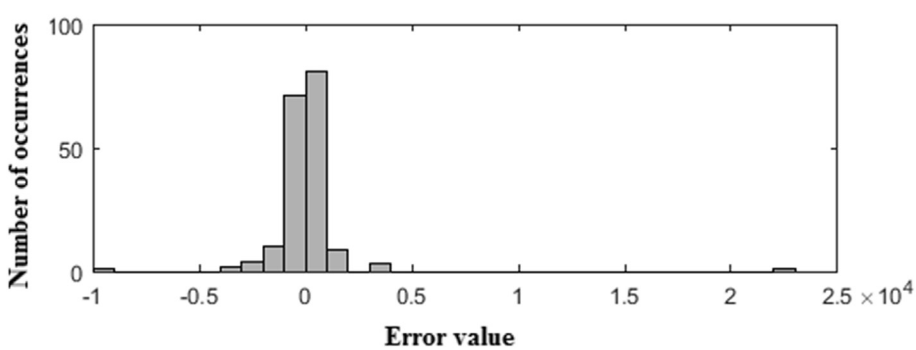


**Fig. 12.** Histogram of errors referring to Athens.

test set. Also, they indicate indirectly the demand level. The values of the MAE and RMSE with those of MAPE are not always proportional. For example, according to Table 2 Agioi Theodoroi have low MAE and RMSE but high MAPE. The opposite is the case

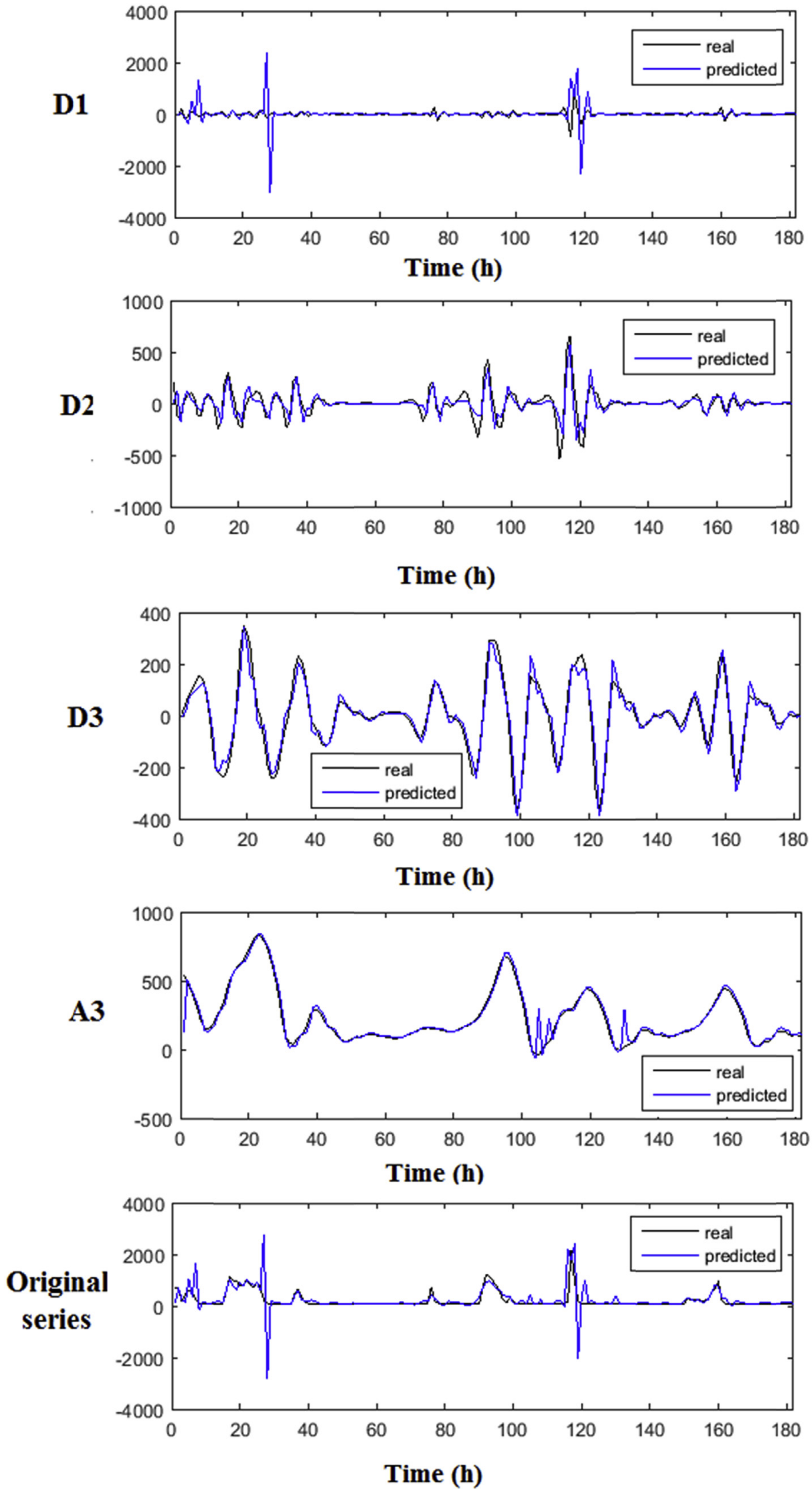


**Fig. 13.** Real and predicted series (original and WT components) of ELPE.

for Athens. As mentioned earlier, no data pre-processing have employed. The presence of very low values and zeros affect negatively the models operation, if the assessment is held with the MAPE indicator. When the denominator of MAPE (i.e. the real value) is an extremely low value, the average percentage error of the specific instance (i.e. day) is extremely high. Thus, MAPE receives large values. In most generation units, there are instances that the consumption is zero. These instances differ in terms of number and position within the gas demand series, a fact that provide limitations in the forecasting process. The operation of gas fired generation units is affected mostly by the energy market actions (i.e. unit commitment, economic dispatch and others). Therefore, they do not follow the patterns of other points, a fact that is preliminary shown by Fig. 2. Note that large MAPE values are referred also to other studies [31]. When we need to measure percentage errors, the MARNE indicator is more suitable. Contrary to MAE and RMSE, it is expressed in per unit values and this fact allows us to have a clear view of the comparison between the models.

For a graphical examination of the forecasting performance, Fig. 10 presents the real and predicted series of Athens after the utilization of the proposed model. All vertical axis are expressed in MWh. It is shown that in each case the predicted curve follows in a large portion the real one. According to the original gas series, the demand of Athens is high in the winter months of 2016 and progressively decreases until the summer months where it reaches the lowest levels. From this pattern, at least theoretically, it is shown that natural gas is utilized mainly for heating loads. Fig. 11 displays the absolute errors per day that refer to the original series prediction shown in Fig. 10. The lower errors are met at the end of the series. The largest error is found in the first instance. If more research effort is placed in eliminating this error, MAPE of the point of Athens will decrease in a large portion. This large error is met only once according to Fig. 12 that shows the histogram of errors. This histogram is another indication of the robustness of the proposed model.

Another example is found in Fig. 13, where the real and predicted for original series and each WT component of ELPE point are compared. ELPE refers to a gas fired generation unit located in Northern Greece and owned by the largest petroleum processing company of Greece. The D1, D2 and D3 components are almost symmetrical around zero. The original series is not volatile and present values close to zero. According to the A3 component and original series curves, there are higher predictions errors. This is also proven by Fig. 14 which displays the absolute errors per day of the test. Combing the fact that is many instances where the consumption is zero, the MAPE indicator receives large values. The histogram of errors is presented in Fig. 15. While the larger gathering of error values are met close to zero, there are some instances that correspond to large aberrations between the prediction and real series.

Fig. 16 presents the comparison between the real and the predicted series for Alouminion, Komotini PPC, VIPE Larisa and Drama distribution points. Again it is shown that the model succeeds by a large portion to capture the demand patterns of the various distribution points. Alouminion and Komotini PPC correspond to generation units. Here the errors are higher compared to the other two points. The pattern VIPE Larisa is almost periodic; it implies a specific pattern of industrial activity followed in all months of the test set. Drama is a town located in north Greece. The natural gas consumption is high during the winter months and exhibits a continuous decrement towards the summer. Just like the city center of Athens, it is implied that natural gas is used mainly to cover heating loads.

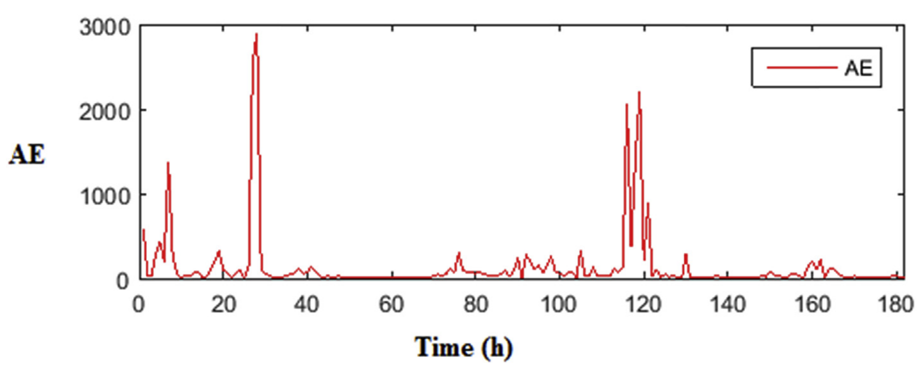


**Fig. 14.** Absolute Errors per day referring to ELPE.

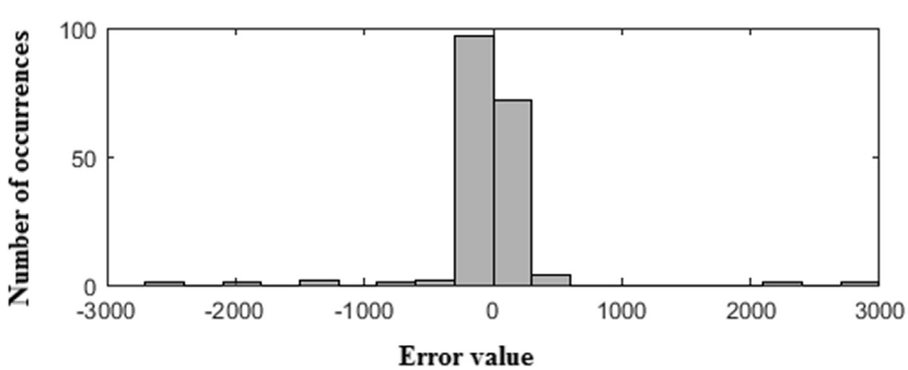


**Fig. 15.** Histogram of errors referring to ELPE.

## 4. Conclusions

Accurate forecast of energy demand is essential for utilities, energy traders, regulatory authorities, decision makers and others. The significance of robust forecasts of crucial parameters like electricity load, fuel demand, energy prices and others is evident for the decision making strategy in short-term horizon. Natural gas demand forecasting is also essential in long-term planning, towards supporting the decisions regarding gas imports, tariff design, maintenance, pipeline system expansion and others. In recent years, the electricity load forecasting literature has witnessed an enormous growth of research papers. On the contrary, natural gas demand forecasting is relatively more limited. While contemporary energy policies seek ways to limit the utilization of coal and petroleum, natural gas and renewables energy resources appear as strong candidates to gradually replace these aforementioned energy carriers.

The aim of this paper is to test the robustness of a novel hybrid

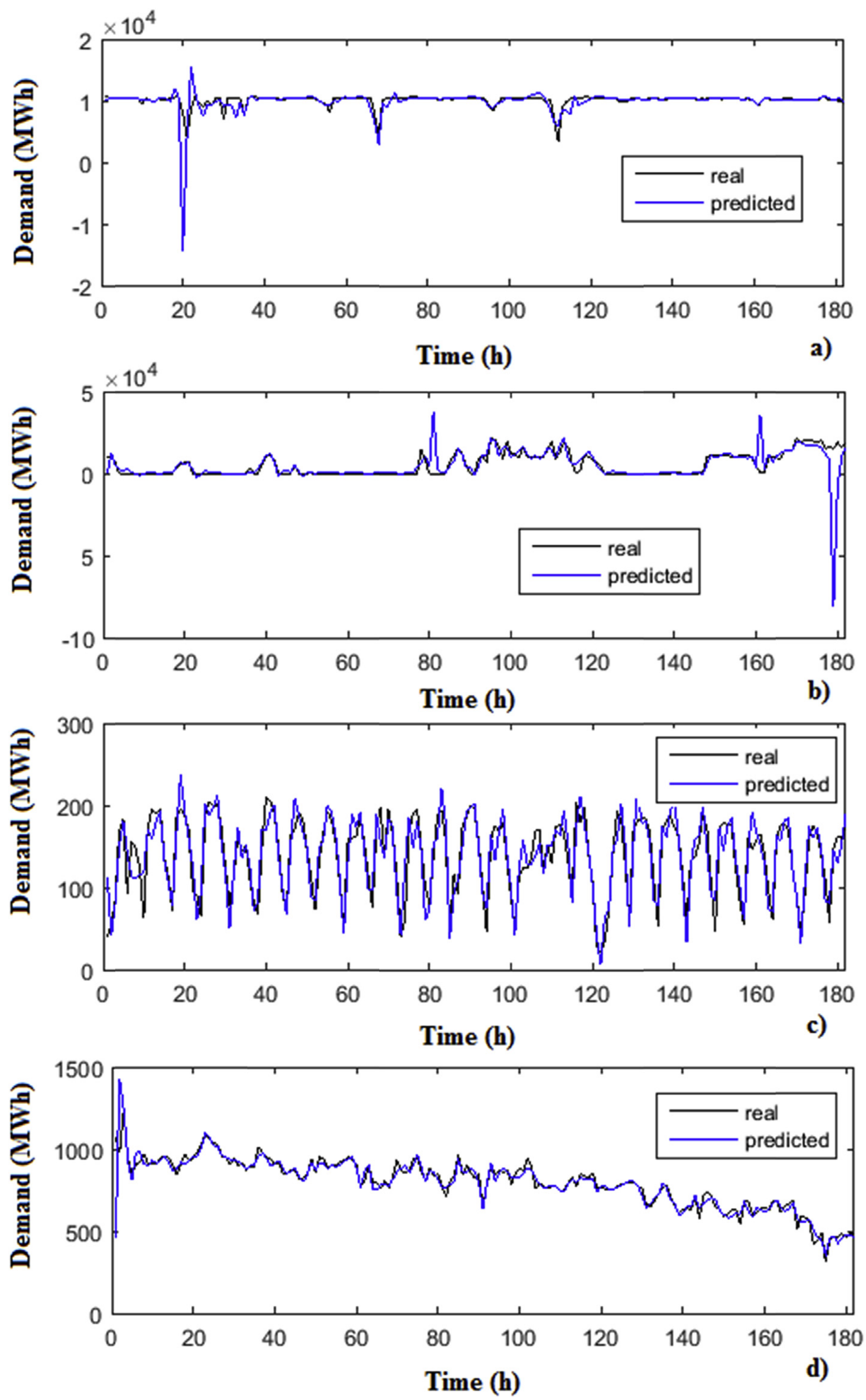


**Fig. 16.** Real and predicted series of a) Alouminion, b) Komotini PPC, c) VIPE Larisa and d) Drama.

computational intelligence model in day-ahead natural gas demand predictions. The proposed model combines the Wavelet Transform (WT), Genetic Algorithm (GA), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Feed-Forward Neural Network (FFNN). The WT is used to decompose the original signal in a set of subseries and then a GA optimized ANFIS is employed to provide the forecast for each subseries. ANFIS output is fed into a FFNN to refine the initial forecast and upgrade the overall forecasting accuracy.

The model is applied to all distribution points that compose the natural gas grid of a country, in contradiction to the majority of the literature that focuses on a limited number of distribution points. The reported errors in load forecasting tasks corresponding to aggregated system level are below 3%. However, this is not the case for natural gas series. This is mainly due to the high volatilities that the gas series exhibit in many distribution points. This approach enables the comparison of the model performance on different consumption patterns, providing also insights on the characteristics of large urban centers, small towns, industrial areas, power generation units, public transport filling stations and others.

According to the experimental outcomes, the lowest errors are met in large city centers. The gas fired generation units lead to higher errors in most cases, but this due to the dynamics of the wholesale electricity market, demanding more specific modeling approaches such as the unit commitment and the power systems expansion planning problems, but as well to the presence of zero values that affect the calculation of the MAPE indicator. For this reason, more forecasting validity indicators are needed to assess the performance of the forecasting tool. In the case of presence of zero or extremely low values, MARNE indicator appears more suitable.

The proposed model is characterized by high flexibility, comprehensive operation and low requirements for computation resources. Thus, it can be used by modern utilities, grid operators and market participants.

The natural gas forecasting problem is a very challenging engineering task. Gas demand series are volatile and can be influenced by a diverse group of variables. Therefore, more effort should be placed towards the goal of increasing the prediction accuracy. The future challenges in the natural gas demand forecasting can be summarized in the following: a) To explore data-preprocessing techniques in terms of lowering both the forecasting error and the execution time, b) to construct new models within the concept of "combined forecasts", c) to involve the clustering tool in order to extract natural gas profiles and use them in the forecasting process.

## References


[1] Seifi H, Sepasian MS. Electric power systems planning: issues, algorithms and solutions. Berlin: Springer-Verlag; 2011.
[2] Suganthi L, Samuel AA. Energy models for demand forecasting - a review. Renew Sustain Energy Rev 2012;16:1224–40.
[3] Soldo B. Forecasting natural gas consumption. Appl Energy 2012;92:26–37.
[4] Rothwell G, Gómez T. Electricity economics: regulation and deregulation. Canada: Wiley-IEEE Press; 2003.
[5] Deetman S, Hof AF, Pfluger B, van Vuuren DP, Girod B, van Ruijven BJ. Deep greenhouse gas emission reductions in Europe: exploring different options. Energy Policy 2013;55:152–64.
[6] Szoplik J. Forecasting of natural gas consumption with artificial neural networks. Energy 2015;85:208–20.
[7] Taspinar F, Çelebi N, Tutkun N. Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods. Energy Build 2013;56:23–31.
[8] Khotanzad A, Elragal H, Lu T-L. Combination of artificial neural-network forecasters for prediction of natural gas consumption. IEEE Trans Neural Netw 2000;11:464–73.
[9] Gorucu FB, Geris PU, Gumrah F. Artificial neural networks modeling for forecasting gas consumption. Energy Sources 2004;26:299–307.
[10] Musilek P, Pelikan E, Brabec T, Simunek M. Recurrent neural network based gating for natural gas load prediction system. In: The 2006 IEEE international joint conference on neural network proceedings, vancouver, Canada, 3736–3741.
[11] Kizilaslan R, Karlik B. Comparison neural networks models for short term forecasting of natural gas consumption in Istanbul. In: International Conference on the Applications of digital information and web technologies; 2008. p. 448–53. Ostrava, Czech Republic.
[12] Kizilaslan R, Karlik B. Combination of neural networks forecasters for monthly natural gas consumption prediction. Neural Netw World 2009;19:191–9.
[13] Khotanzad A, Elragal H. Natural gas load forecasting with combination of adaptive neural networks. In: Proceedings of the international joint conference on neural networks; 1999. p. 4069–72. Washington, U.S.A.
[14] Viet NH, Mandziuk J. Neural and fuzzy neural networks in prediction of natural gas consumption. Neural Parallel Sci Comput 2005;13(3–4):265–86.
[15] Karimi H, Dastranj J. Artificial neural network-based genetic algorithm to predict natural gas consumption. Energy Syst 2014;5:571–81.
[16] Yu F, Xu X. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. Appl Energy 2014;134:102–13.
[17] Potocnik P, Soldo B, Šimunovic G, Šaric T, Jeromen A, Govekar E. Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia. Appl Energy 2014;129:94–103.
[18] Azadeh A, Asadzadeh SM, Ghanbari A. An adaptive network-based fuzzy inference system for short-term natural gas demand estimation: uncertain and complex environments. Energy Policy 2010;38:1529–36.
[19] Behrouznia A, Saberi M, Azadeh A, Asadzadeh SM, Pazhoheshfar P. An adaptive network based fuzzy inference system-fuzzy data envelopment analysis for gas consumption forecasting and analysis: the case of South America. In: International conference on intelligent and advanced systems; 2010. p. 1–6. Kuala Lumpur, Malaysia.
[20] Hippert HS, Pedreira CE, Souza RC. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans Power Syst 2001;16:44–55.
[21] Hellenic Gas Transmission System Operator S.A. (DESFA): http://www.desfa.gr/?lang=en.
[22] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009;34:46–57.
[23] Mallat S. A theory for multiresolution signal decomposition-the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11:674–93.
[24] Xu R, Wunsch D. Clustering. New Jersey: John Wiley & Sons. Inc.; 2006.
[25] Levenberg K. A method for the solution of certain non-linear problems in least squares. Quar J Appl Math 1944;2:164–8.
[26] Graupe D. Principles of artificial neural networks. Singapore: World Scientific Publishing Company; 2007.
[27] Jang J-SR. ANFIS: adaptive-network-based fuzzy inference system. Syst Man Cyber 1993;23:665–85.
[28] Soldo B, Potocnik P, Simunovi G, Sari T, Govekar E. Improving the residential natural gas consumption forecasting models by using solar radiation. Energy Build 2014;69:498–506.
[29] Koltsaklis NE, Dagoumas AS, Kopanos GM, Pistikopoulos EN, Georgiadis MC. A spatial multi-period long-term energy planning model: a case study of the Greek power system. Appl Energy 2014;2014(115):456–82.
[30] Koltsaklis NE, Dagoumas AS, Georgiadis MC, Papaioannou G, Dikaiakos C. A mid-term, market-based power systems planning model. Appl Energy 2016;179:17–35.
[31] Sánchez-Úbeda EF, Berzosa A. Modeling and forecasting industrial end-use natural gas consumption. Energy Econ 2007;29:710–42.