

Natural Gas Consumption Forecasting with Kolmogorov–Arnold Networks: A Comparison with MLP

Kürşad Arslan^{1*} , Emrah Dönmez¹ 

¹Bandırma Onyedi Eylül University, Department of Software Engineering, Faculty of Engineering and Natural Sciences, Bandırma/Balıkesir, Türkiye, ror.org/02mtr7g38

Corresponding author:

Kürşad Arslan,
Bandırma Onyedi Eylül University
Department of Software Engineering
karslan@bandirma.edu.tr



Article History:

Received: 07.08.2025
Revised: 25.10.2025
Accepted: 03.11.2025
Published Online: 11.12.2025

ABSTRACT

Natural gas remains a vital resource for meeting residential heating energy needs, particularly during the winter months. Accurate demand forecasting is essential for maintaining supply-demand balance, optimizing operational costs, and supporting effective energy management. In this study, the natural gas consumption prediction performance of Kolmogorov-Arnold Networks (KAN), a new neural network architecture, was compared with that of the basic model, Multi-Layer Perceptrons (MLP). Both models were trained and tested on the same dataset using monthly consumption data. While MLPs are diversified through the number of neurons, activation functions, and layer configuration, KAN models are configured by modifying B-spline parameters, grid size, and layer structure. The results show that the KAN model achieved the highest R2 value despite having fewer trained parameters. Although some versions of the MLP model yielded lower Mean Absolute Percentage Error (MAPE) values, they fell short of KAN in terms of overall fit. These findings demonstrate the superior ability of KAN to capture nonlinear patterns in energy demand forecasting, offering computational efficiency.

Keywords: Kolmogorov-Arnold Networks (KANs), Natural gas, Demand forecasting, MLP, Time series

1. Introduction

Time series analysis is widely used to model and forecast variables that evolve over time, such as natural gas consumption. It remains one of the most commonly applied techniques in energy forecasting today. Natural gas is one of the most heavily utilized energy sources across both industrial and residential sectors. While consumption in the industrial sector tends to remain relatively stable, residential consumption exhibits strong seasonality due to its primary use for heating. Notably, the volume of natural gas consumed for heating during winter months constitutes more than half of the total daily usage. Numerous studies have been conducted on forecasting natural gas consumption. Arslan et al. proposed two artificial neural network (ANN)-based models designed to detect abnormal consumption patterns resulting from gas meter malfunctions. Their approach combined ANN models with statistical methods and was fundamentally structured around demand forecasting principles [1]. Wei et al. explored the influence of missing data on the accuracy of daily consumption forecasts [2]. Akpınar et al. evaluated gas consumption forecasting through two distinct modeling scenarios involving ANN architectures trained using backpropagation (BP) and artificial bee colony (ABC) optimization algorithms. Their results indicated that the ABC-trained model outperformed its counterpart, underscoring the importance of the training algorithm [3]. Some studies use multiple linear regression, ARIMA, and exponential smoothing models to predict natural gas consumption [4], [5], [6]. A gray forecast model was used to estimate natural gas consumption in Chinese households [7]. In addition to these methods, Bayesian models and genetic algorithms were used to improve the accuracy of natural gas consumption forecasts [8], [9], [10]. Several studies have also focused on short-term consumption forecasting. Anđelković et al. incorporated weather forecasting into their short-term natural gas demand prediction model, enhancing its predictive performance [11]. Similarly, Qiao et al. applied Gaussian smoothing to consumption data for short-term forecasting and reported promising outcomes [12].

This study explores the application of Kolmogorov-Arnold Networks (KAN), a recently introduced neural network architecture [13], in time series forecasting, with a particular focus on energy demand prediction. Given the novelty of the method, existing research is still limited. Among the few studies available, the work of J. Vaca-Rubio et al. is particularly notable. The authors employed KAN to forecast real-world satellite traffic data and reported that the model achieved strong predictive performance with fewer parameters [14]. KAN has also begun to attract attention in energy-related applications. M. Sulaiman et al. proposed KAN for estimating the state of charge (SoC) in electric vehicle batteries. They compared its

performance against that of artificial neural networks (ANN) and a hybrid Barnacles Mating Optimizer–Deep Learning (BMO-DL) model. Their findings revealed that KAN-Model 1 achieved the most accurate and reliable results, while ANN and BMO-DL exhibited less consistent performance. The study emphasized the potential of KAN to enhance prediction accuracy in SoC estimation tasks [15]. In another study, M. Sulaiman et al. applied KAN to forecast chiller energy consumption using real data from a commercial building. They compared the performance of KAN with that of ANN and the Teaching-Learning-Based Optimization–Deep Learning (TLBO-DL) model. The results showed that KAN achieved an R^2 value of 0.9465 and a root mean square error (RMSE) of 6.1023, outperforming both comparison models in terms of accuracy [16]. Additionally, studies applying the KAN model to various problems are also available in the literature. Peng et al. used the KAN model to predict the pressure and flow velocity of pumps. They compared the performance of the KAN model with MLP and Random Forest (RF) models. The authors reported that the KAN model achieved superior accuracy with lower MSE values in pressure and flow rate predictions. Additionally, the formulas derived from KAN revealed nonlinear relationships between parameters and pump performance [17]. Granata et al. examined the short-term river flow prediction performance of KAN and Transformer models. The study revealed that KAN significantly outperformed Transformer models in short-term predictions, highlighting the potential of this new method to improve water resource management in complex hydrological systems [18].

In this study, the KAN model was used to estimate natural gas consumption. This model, which has not yet been widely used in energy demand forecasting, was experimentally investigated to determine whether it possesses a higher capacity for learning complex relationships compared to traditional methods in the field. The KAN model was evaluated in comparison with the MLP; both models were trained on the same dataset and compared in terms of prediction accuracy and parameter efficiency. To the best of our knowledge, this study represents the first application of the KAN to natural gas consumption forecasting, highlighting its innovative structure and potential advantages over conventional ML/DL approaches. The data set used in the study covers the years 2015–2022. This dataset consists of natural gas consumption data from Istanbul, Türkiye's most populous city, and includes monthly consumption amounts for each district. The data was obtained from the Open Data Portal and is provided by IGDAS [19]. Additionally, only historical consumption data were used in model training, without considering external factors (such as meteorological conditions) that influence consumption. This reveals how well the models adapt to sudden changes during periods when seasonal effects on natural gas consumption are observed.

The methodologies used in Section 2 of the study are described. Section 3 describes the data modeling, while Section 4 presents the study's results. Section 5 compares the results obtained with those of some existing studies in the literature, and the next section evaluates the study's results.

2. Method

2.1. Kolmogorov-Arnold Networks (KAN)

Kolmogorov-Arnold Networks are a neural network architecture developed based on the Kolmogorov-Arnold representation theorem. It was first proposed by Z. Liu et al. [13]. According to this theorem, multivariate continuous functions can be represented as a finite sum of univariate continuous functions. This theorem states that a multivariate function f can be expressed as a combination of specific univariate functions and addition operations. That is, the function $f : [0, 1]^n \rightarrow \mathbb{R}$ can be represented using univariate functions $\phi_{q,p} : [0, 1] \rightarrow \mathbb{R}$ and $\Phi_q : \mathbb{R} \rightarrow \mathbb{R}$. Therefore, it has a different architecture from classical neural networks. In this network, single-variable functions are used instead of weights. These functions are spline-based and learnable. Therefore, unlike traditional neural network architectures, the KAN model does not contain linear weights, which is the most fundamental difference between KAN and traditional neural networks. Spline functions are curves that model non-linear relationships between inputs and outputs. One of the features that significantly changes the learning process is that the activation functions and weights are flexibly adapted for KAN. Activation functions are non-linear but fixed. Weight values in traditional neural networks, on the other hand, are linear and learnable. The KAN model has the learnable and non-linear properties of these two structures. This enables activation functions to be learnable (Figure 1). Learnable activation functions enable the KAN model to solve more complex patterns than other neural networks.

At each layer of the network, activation values are computed and represented through corresponding nodes (Figure 1). For instance, expressions such as $x_{1,1}$ and $x_{1,2}$ denote activation functions located in the first and second layers, respectively. These activations follow a forward flow from the input to the output layer, akin to the mechanism observed in conventional neural networks. Within the KAN architecture, the functions $\phi_{a,b}$ serve as trainable univariate components that replace the fixed weights commonly found in traditional neural models. Rather than relying on static weight matrices, KAN utilizes B-spline functions to model activations along network connections. Figure 1(b) presents an enlarged view of an example spline function [13]. In this framework, the activation function $\phi_{(x)}$ is expressed as a parameterized B-spline, enabling smooth adaptation across both coarse and fine grid configurations. Two distinct configurations are illustrated for comparison. In the first example, where $G_1 = 5$, the structure involves fewer control points and a coarser grid layout. Conversely, in the second case with $G_2 = 10$, the model operates with a denser set of control points and a finer grid, which enhances the ability to capture more subtle patterns in the data. The spline-based representation is constructed by summing weighted contributions defined at control points, thereby accurately reflecting the input data structure.

Equation (1) defines a two-layer network structure in which the functions are positioned along the edges, while a basic

summation operation is carried out at the nodes [13]. By stacking this fundamental structure K times, the resulting deep-KAN architecture can be formally described by Equation (2).

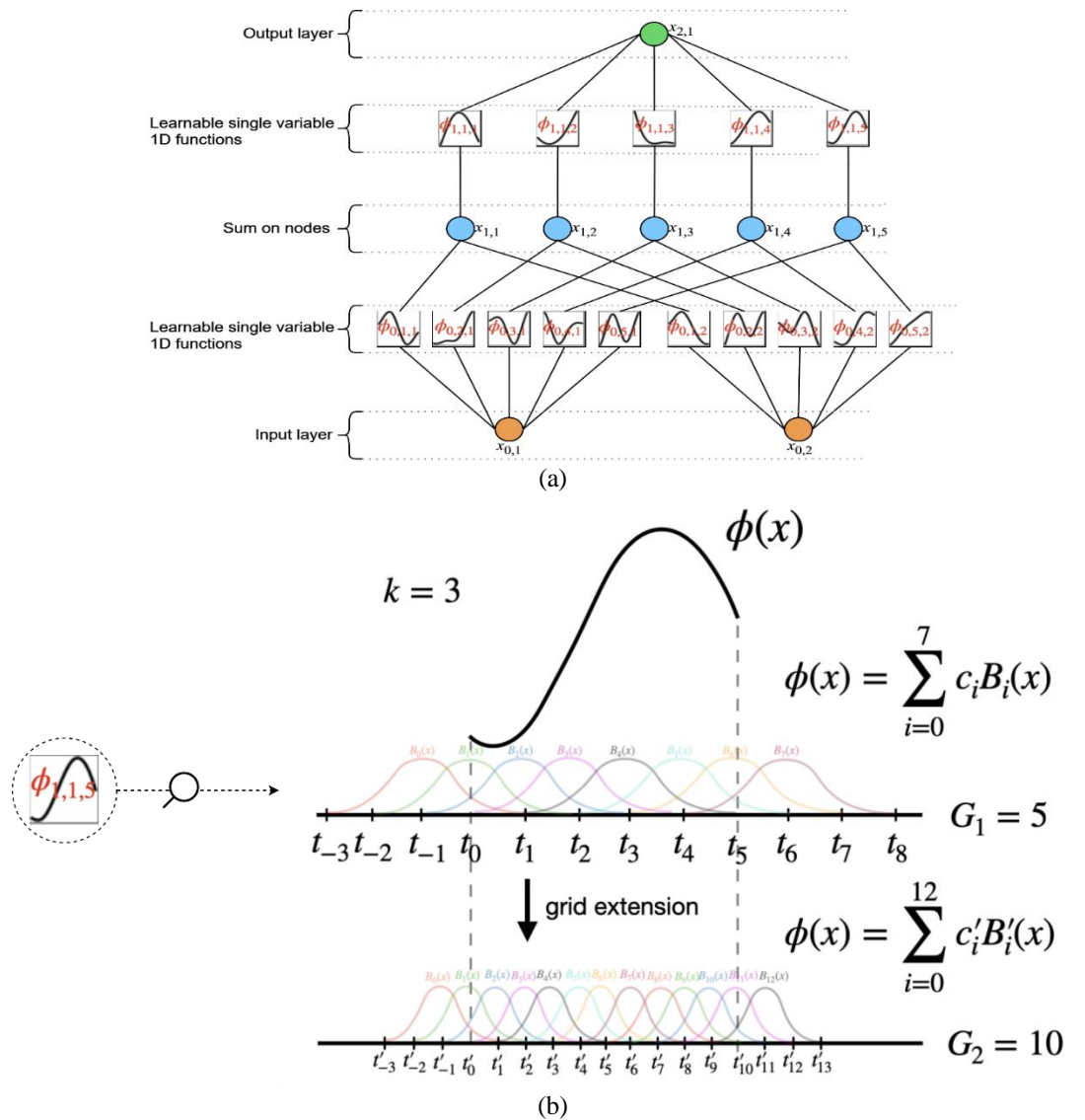


Figure 1. General architecture of the KAN model; (a) architecture consisting of a hidden layer, (b) detailed look at the spline function [13].

$$f(x) = \sum_{a=1}^{2n+1} \Phi_a \left(\sum_{b=1}^n \phi_{a,b}(x_b) \right) \tag{1}$$

$$KAN(x) = (\Phi_K \circ \Phi_{K-1} \circ \dots \circ \Phi_1)(x) \tag{2}$$

$$spline(x) = \sum_a c_a B_a(x) \tag{3}$$

Spline functions are an approach that enables continuous functions to be represented by piecewise polynomials, and their general form is as shown in Equation (3). The term $B_a(x)$ denotes the basis functions that form the spline, while c_a denotes the coefficients corresponding to these functions. Each basis function effectively controls the local behavior of the spline within a specific interval. The learning process in KAN begins with dividing the input data into grid intervals (G), where spline functions are defined. A spline function is applied to each input data point, and the output values calculated by the spline functions are produced. The difference between the network's output and the target output is calculated using an error

function (usually the mean squared error). Based on the values obtained from the error function, the parameters of the spline functions (spline coefficients and knot points) are updated. Updating spline parameters is one of the most critical steps in improving the representational power of the model. Throughout the training process, optimization methods are applied to the spline parameters to minimize the total error.

This process ensures that the model captures changes in the input space more accurately by modifying the coefficients of the spline functions.

2.2. Multi-Layer Perceptron (MLP)

Multi-layer perceptrons, which draw inspiration from the operational mechanisms of the human brain, are constructed by integrating multiple neurons and hierarchical layers [20]. This structural framework serves as the basis not only for deep neural networks but also for the KAN, which constitutes the primary focus of this study. Although the backpropagation (BP) algorithm is the standard method for training, it may also be incorporated into hybrid frameworks alongside various optimization techniques [3]. In feedforward neural networks, the computed outputs are evaluated against target values using specific error metrics, and the weight parameters (i.e., constant coefficients) are iteratively adjusted to reduce the error [1]. The corresponding summation process is mathematically represented in Equation (4).

$$NET_a = \sum_{a=1}^n x_a \cdot w_a + b \quad (4)$$

Following this, the learning mechanism proceeds in a nonlinear fashion via activation functions, as illustrated in Equation (5).

$$OUT_a = \varphi(NET_a) \quad (5)$$

After obtaining the results produced by the models, one of the most important steps in making a fair comparison is the error measurement process. The error measurement metrics used are presented in Table 1. In this study, the error measurement metrics used were the mean square error (MSE) (Equation 6) and the mean absolute percentage error (MAPE) (Equation 7). In addition to these metrics, it is also important to evaluate the similarity between the models' outputs and the target values, as well as the model fit. This evaluation was performed using the R^2 value (Equation 8).

Table 1. Error measurement metrics

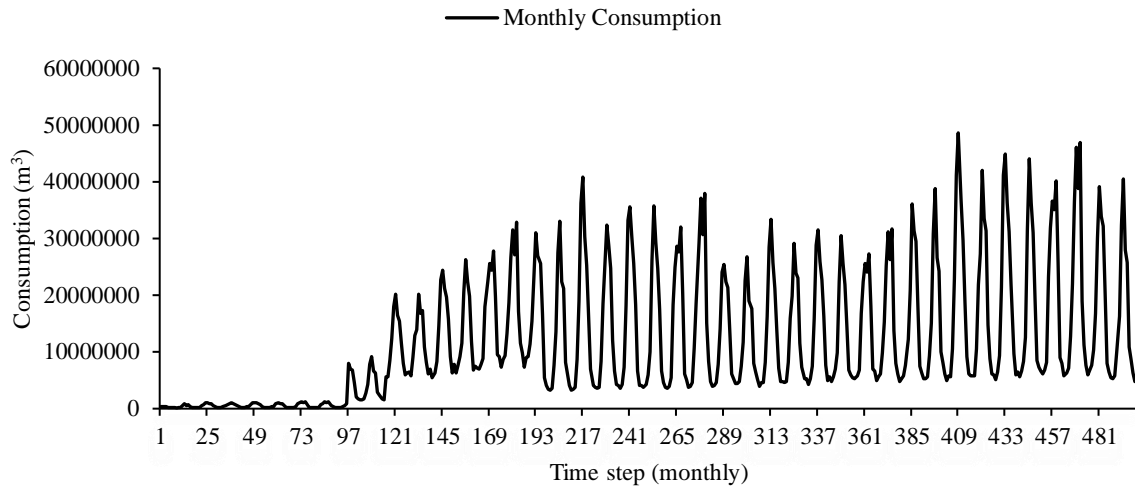
Equation No	Metric	Equation
6	MSE	$MSE = \frac{1}{n} \sum_{a=1}^n (\hat{y}_{Pa} - y_{Ta})^2$
7	MAPE	$MAPE = \frac{1}{n} \sum_{a=1}^n \left \frac{\hat{y}_{Pa} - y_{Ta}}{y_{Ta}} \right $
8	R^2	$R^2 = 1 - \left[\frac{\sum_a (\hat{y}_{Pa} - y_{Ta})^2}{\sum_a (y_{Ta} - \bar{y}_T)^2} \right]$

In this equation, \hat{y}_{Pa} denotes the predicted value at step a, y_{Ta} indicates the target value, and \bar{y}_T represents the mean of the target values. Evaluating the relationship among these three quantities is crucial in thoroughly analyzing the model's prediction performance. Specifically, the R^2 value quantitatively expresses how closely the model outputs align with the target values and is therefore widely used as a performance metric in model comparisons.

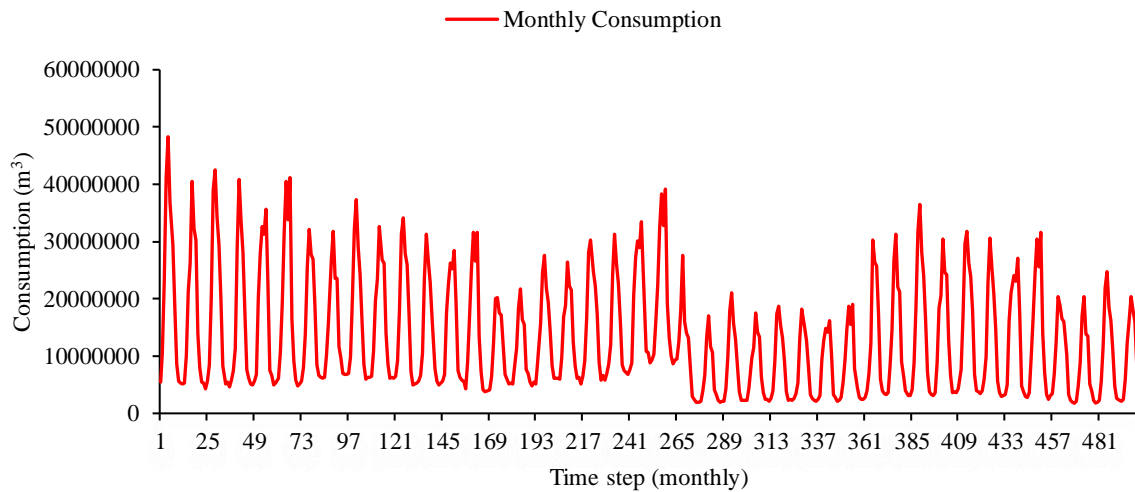
3. Modeling and Consumption Forecasting

The amount of natural gas consumed in a city can be influenced by various external factors, some of which are interdependent, while others are independent of one another. Examples of such influences include the season during which consumption occurs, air temperature, consumer habits, the regional population, the household population, and calendar events during the consumption period. The data obtained in this study represent the monthly natural gas consumption in Istanbul province between 2015 and 2022. Upon examination of the data, it is evident that total natural gas consumption has increased annually. The highest increase occurred in 2017, while the second-highest increase occurred during the global COVID-19 pandemic. As of 2022, the province's total natural gas consumption has increased to 7 billion cubic meters. The province's natural gas consumption graph is shown in Figure 2. In this figure, (a) shows the first 500 consumption data points available in the dataset, while (b) represents the second 500 consumption data points.

In the study, 75% of all the data was allocated for training, while the remaining portion was used to test the models. The same data was used for both models (KANs-MLP). Among the most significant factors affecting consumption were the date and the district, so this information was also digitized and included in the model. Specifically, the data used in the study are collected every month. Each consumption record includes the year, month (1-12), district name, and consumption amount. The district names (1-39) have been numbered and digitized (there are 39 districts in total). Since the year information may vary, it was not used. Thus, the dataset includes the month number (1-12), district number (1-39), and consumption amount (in cubic meters). In this model, the month and district information were used as input variables, while the predicted variable was the amount of natural gas consumption.



(a)



(b)

Figure 2. Monthly natural gas consumption data instances: (a) The first 500 samples of the dataset, (b) The second set of 500 samples

Since monthly data is used in the study and the consumption amounts represent the total consumption of the respective districts, the values in this column are high. This situation may reduce the impact of other information (month number-district number) on the model during its use. To overcome this, the data was subjected to min-max normalization. Thus, using equation (9), the data in each column was modeled within the range of 0 to 1.

$$y_{a(0-1)} = \frac{y_a - y_{min}}{y_{max} - y_{min}} \quad (9)$$

4. Approach and Results

In the study, two different models were developed to predict the city's natural gas consumption. These are the KAN approach, which was not previously used for predicting natural gas consumption, and the MLP model. Both models were trained and tested with the same data. The MLP model was tested using different numbers of neurons, activation functions, and layer

counts. In the KAN model, this process was achieved by adjusting the parameters of the B-spline function, as well as the number of neurons and layers. Our goal is to investigate whether the KAN model outperforms previously used methods for predicting natural gas consumption. In this way, a total of 48 different models were created, consisting of 24 MLP models and 24 KAN models. When calculating the results, the data was transformed back to its original form, and the error values were calculated accordingly. Additionally, the number of parameters in each model and the execution time of the models (in minutes) were recorded.

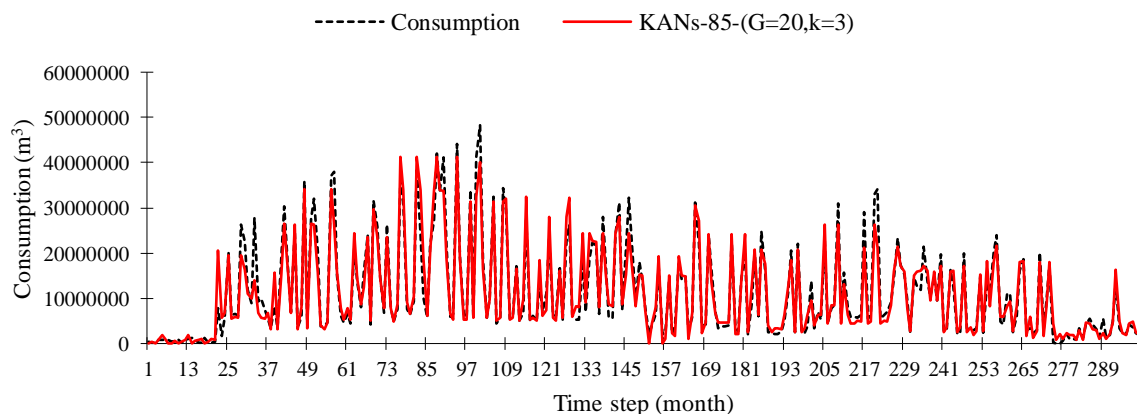
In the study, the lowest MAPE value was 0.897, while the highest was 3.25. These models were, respectively, the MLP model with 3 layers and the ReLU activation function, and the MLP model with a single layer and a linear activation function. Despite having the lowest MAPE value, this model did not yield the best result. The top 10 models, based on R^2 values for both KANs and MLP, are presented in Table 2.

Table 2. Evaluation results of models

No	Model	Neuron*	B-spline	Act. Func.	MAPE	MSE	R^2
1	KAN	85	$G=20, k=3$	learnable	1.44	0.003	0.929
2	KAN	85	$G=20, k=5$	learnable	1.157	0.003	0.927
3	KAN	120+85	$G=20, k=3$	learnable	2.149	0.003	0.927
4	KAN	300+120+85	$G=20, k=3$	learnable	1.304	0.003	0.925
5	KAN	120+85	$G=20, k=5$	learnable	1.811	0.003	0.925
6	KAN	40	$G=20, k=3$	learnable	0.931	0.003	0.924
7	KAN	300+120+85	$G=20, k=5$	learnable	1.138	0.004	0.919
8	KAN	40	$G=20, k=5$	learnable	1.196	0.004	0.902
9	KAN	300+120+85	$G=10, k=3$	learnable	1.032	0.006	0.867
10	KAN	120+85	$G=10, k=3$	learnable	1.798	0.006	0.864
11	MLP	300+120+85	N/A	ReLU	0.897	0.016	0.67
12	MLP	300+85	N/A	ReLU	0.954	0.018	0.633
13	MLP	300	N/A	ReLU	1.09	0.019	0.608
14	MLP	85	N/A	ReLU	1.386	0.02	0.586
15	MLP	40	N/A	ReLU	1.25	0.021	0.576
16	MLP	20	N/A	ReLU	1.92	0.024	0.517
17	MLP	300+120+85	N/A	tanh	2.059	0.024	0.516
18	MLP	40	N/A	tanh	2.08	0.025	0.499
19	MLP	85	N/A	tanh	2.312	0.025	0.491
20	MLP	20	N/A	tanh	2.515	0.027	0.456

*: The (+) symbol means adding a new layer.

The best R^2 value was obtained in the single-layer KAN model. This value corresponds to the KAN model with a grid size of $G = 20$ and a spline order of $k = 3$. The grid size specifies the number of segment sizes into which the data will be divided, with each segment represented by a spline function. This parameter enables the model to solve complex relationships and learn more effectively about situations that change over time due to various reasons. The spline order determines the degree of the spline functions used. Upon examining the results, it was found that the $k=3$ value not only provided higher accuracy but also resulted in faster performance. When examining the number of parameters in the models, the best MLP model contained 44,300 parameters, while the best KAN model contained 6,300 parameters. This demonstrates that the KAN model achieved better results with fewer parameters. Additionally, it was observed that the change in the number of neurons and layers in the KAN models with $G = 20$ did not significantly affect the model's success. The target and predicted values for the best KAN and MLP models are shown in Figure 3.



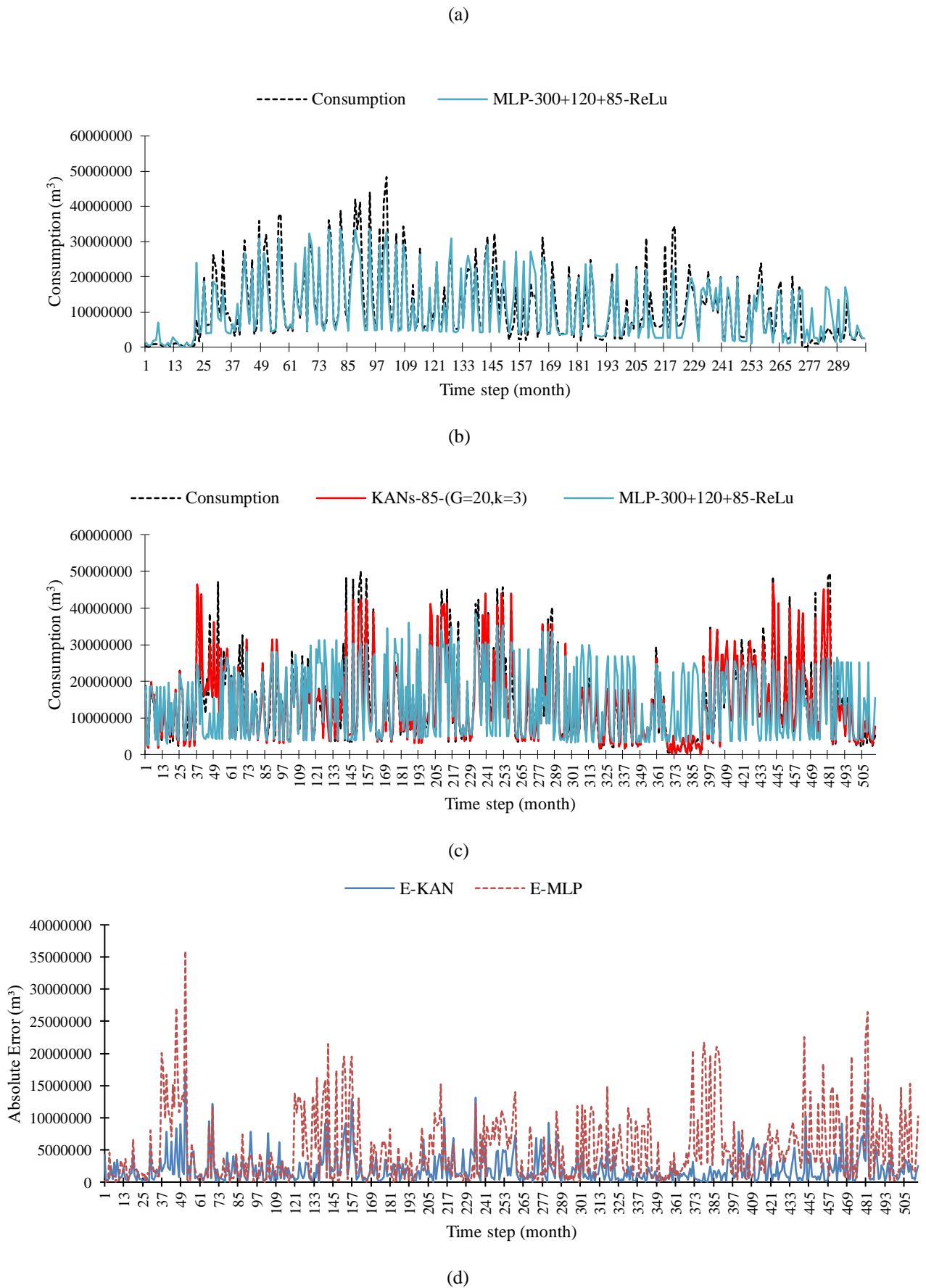
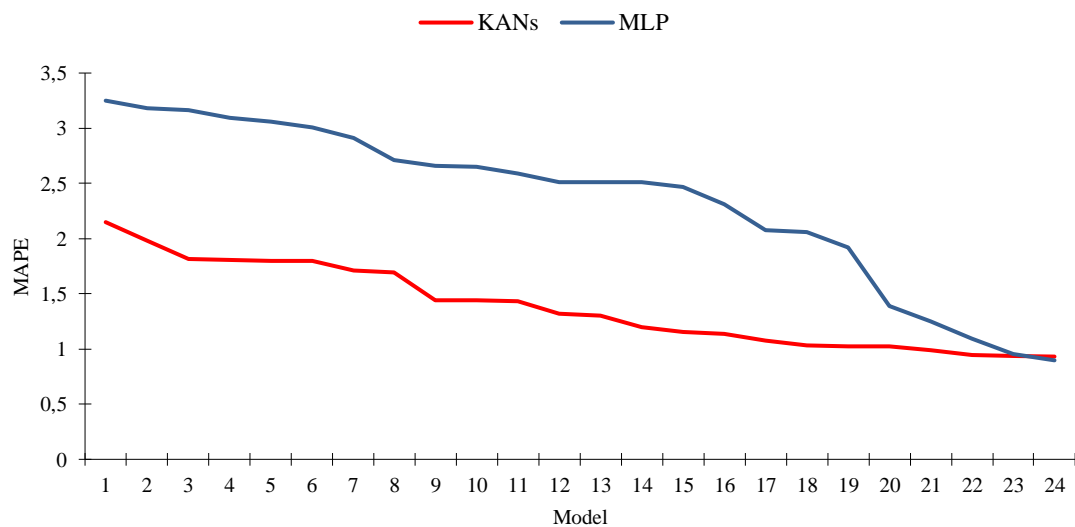


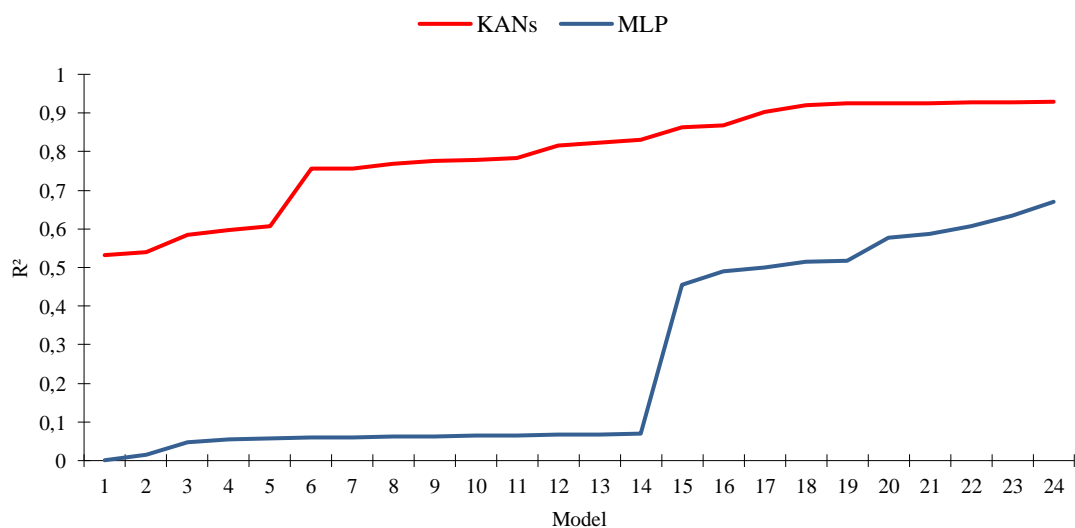
Figure 3. Forecasting results of the best models for KAN and MLP

In the figure, the first 300 data points of the prediction series are presented (Figures 3a and 3 b). The remaining predictions

are shown in a single graph (Figure 3-c). Additionally, Figure 3-d presents the absolute prediction error curves (E-KAN and E-MLP) corresponding to these predictions. It can be clearly seen that while both models generally follow the overall trend of the actual data, the MLP model exhibits more abrupt and higher fluctuations in error values, especially during sudden consumption changes or seasonal transitions. In contrast, the KAN model shows smoother and more stable error behavior over time, indicating better generalization and consistency in prediction performance. Upon examining the graphs, it is observed that the KAN-80-(G=20, k=3) model demonstrates a higher alignment with the actual consumption data patterns and provides more accurate predictions (Figure 3-a). However, although the MLP model yields a MAPE value of 0.543, the R^2 value is not within reasonable limits. This is because MAPE takes the average of the absolute values of the errors, thus evaluating each error size with equal weight. As a result, the impact of regions with high prediction errors is almost the same as that of regions with low prediction errors, and no distinction is made between them. On the other hand, R^2 indicates how well the model fits the data (Figure 4-b). However, it does not provide detailed information about the distribution of the errors. In this scenario, the model may make accurate predictions for high consumption but less accurate predictions for low consumption. The important point here is that even if the MAPE value is at its lowest levels up to a certain consumption point, it suddenly increases during seasonal transitions or at points where subscriber behaviors begin to change. Figure 4 shows the changes in the sequential MAPE and R^2 values of the models.



(a)

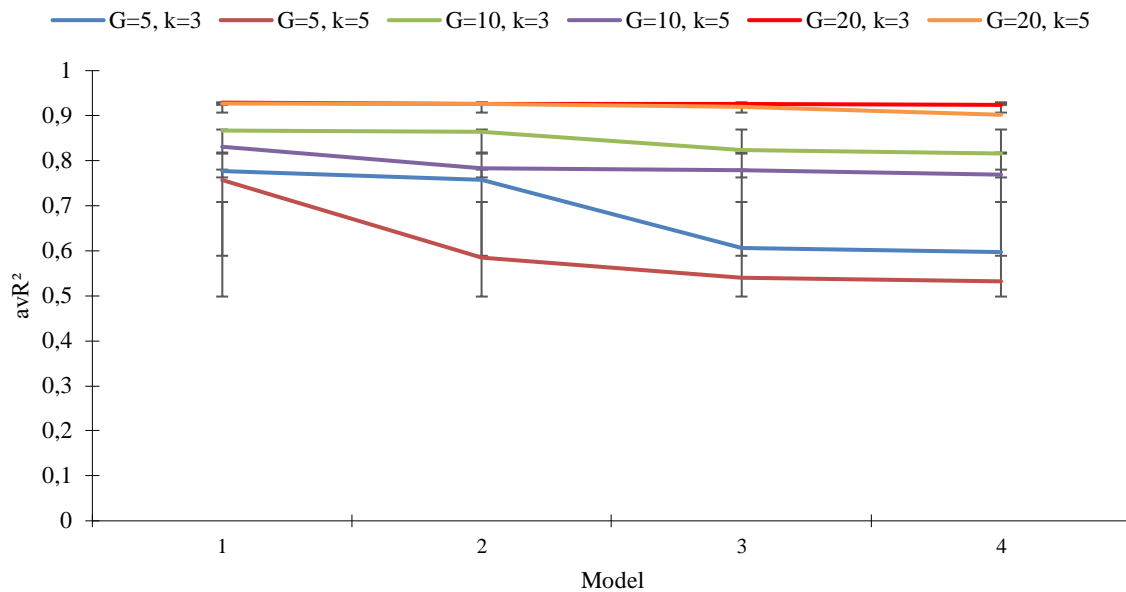


(b)

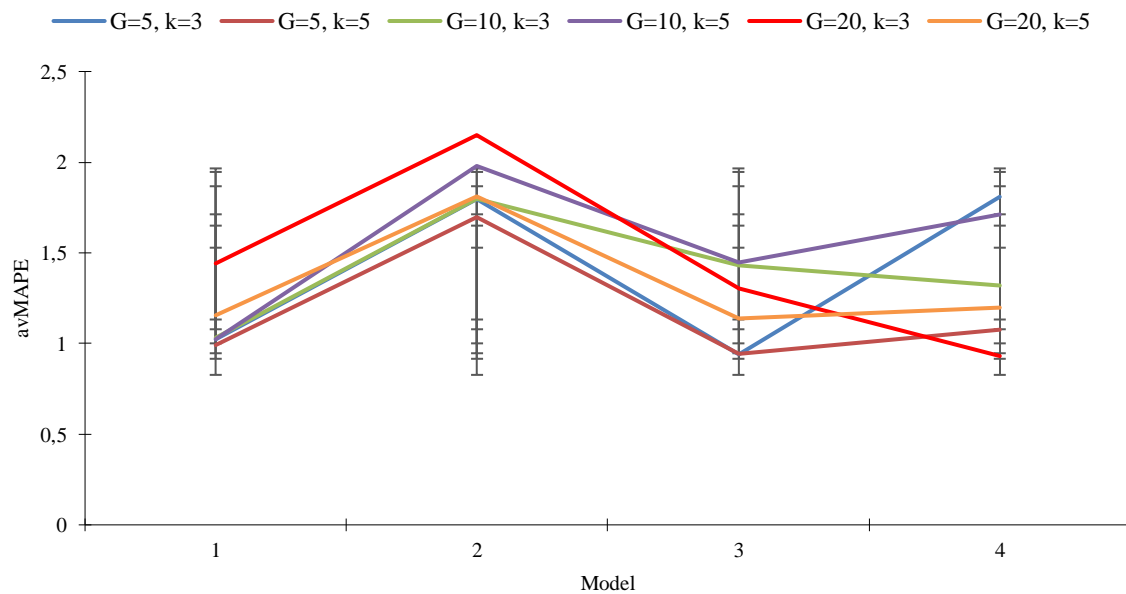
Figure 4. MAPE and R^2 values of KAN and MLP models

Upon examining the results, the prediction performance of the KAN model using different B-spline parameters (G and k) was compared based on MAPE and R^2 values. The graphs visually summarize the effects of changes in the B-spline parameters on the model's performance (Figure 5). However, when preparing the graphs, differences in the number of layers and neurons in the model were ignored, and only the impact of B-spline parameters on prediction accuracy was analyzed.

These findings provide an in-depth understanding of how parametric changes affect model performance, offering valuable insights into the adaptability and flexibility of the KAN model.



(a)



(b)

Figure 5. Change in MAPE and R^2 values of KAN models according to different B-spline parameters

In this graph, the performance of the KAN model was evaluated using B-Spline parameters with different G and k values. In part (a) of the graph, the model's fit was examined through R^2 . Generally, differences in R^2 values were observed based on G and k values, with higher R^2 values achieved for certain parameter combinations. Specifically, with the B-Spline parameters set to $G=20$, the R^2 value reached its highest level, indicating better model fit (Figure 5-a). In part (b) of the graph, prediction accuracy was analyzed through MAPE values. Lower MAPE values indicate that the model has a lower error rate and better prediction performance. It was particularly observed that models with a B-spline degree of 3 ($k=3$) had lower MAPE values (Figure 5-b). The KANs model, with 85 neurons in its hidden layer, $G=20$, and $k=3$, yielded the best results, achieving a MAPE of 1.44 and an R^2 of 0.929.

5. Discussion

In this study, two different models were used to predict natural gas consumption: Kolmogorov-Arnold Networks (KAN) and Multi-Layer Perceptrons (MLP). Both models were trained and tested with the same dataset. Our goal is to evaluate the performance of the KAN model by comparing it with the traditional MLP and examining its success in this field. The dataset used in the modeling processes covers a relatively larger time span compared to other studies in the existing literature.

Table 3. Comparison of the proposed method with existing studies in the literature

Study	Forecasting model	Forecasting type	Data size	Performance
Taspınar et al. [21]	ANN-MLP	Daily	2007-2011	0.814 (%MAPE) 0.441 (RMSE) 0.823 (R ²)
Akpınar and Yumusak [6]	ARIMA	Daily	2009-2012	9.84–18.51 (%MAPE)
Soldo et al. [22]	MLP	Daily	2011-2012	5.06 (%MAPE) 0.635 (R ²)
Szoplik [23]	MLP	Daily	2009-2011	5.5–11.0 (%MAPE) 582.2–2166.6 (RMSE)
Akpınar et al. [24]	ABC-MLP	Daily	2011-2014	14.9 (%MAPE)
Wang and Jiang [25]	NMGM-ARIMA	Monthly	2012-2018	3.16 (%MAPE) 0.79 (%RMSE)
	ARIMA-ANN			2.06 (%MAPE) 0.57(%RMSE)
Fabbiani et al. [26]	Gaussian Process	Daily	2007-2017	2.56 (Av. MAE) 4.25 (Av. RMSE)
Table 3(Continued)				
Hribar et al. [27]	RNN	Hourly	8 winter seasons	9.3 (%MAPE) 1×10^{-3} (MAE)
Qiao et al. [28]	IWOA-RVM	Hourly	360 h	0.02 (%MAPE) 82.51 (MAE) 9064.71 (RMSE)
This study	KAN-MLP	Monthly	2015-2022	0.929 (R ²) 0.897-1.44 (MAPE) 3×10^{-3} (MSE)

In the literature, studies on natural gas consumption typically focus on short-term predictions. For example, in the study by Akpınar and Yumusak [6], the ARIMA method was employed, yielding MAPE values ranging from 9.84% to 18.51%. In another study, Taspınar et al. [21] applied the MLP model to a 4-year dataset, achieving an MAPE of 0.8 and an R² value of 0.8. These percentages serve as significant comparison points for measuring the accuracy of predictions in natural gas consumption. In our study, the lowest MAPE value obtained is 1.44, and the highest R² value is 0.929. This is a significantly successful result compared to other studies.

One of the main advantages of the KAN model is its ability to achieve similar or better results with fewer parameters. For example, the best MLP model has 44.3k parameters, while the best KAN model contains only 6.3k parameters. This demonstrates that the KAN model operates more efficiently and is computationally less costly. Additionally, the dataset used in this study is based solely on historical natural gas consumption data and date information. External factors, such as weather conditions, were not included in the model. This indicates that the model can still yield highly successful results even with just consumption data, highlighting the KAN model's ability to learn the hidden relationships between variables.

In conclusion, the results obtained with the KAN model are more successful in predicting natural gas consumption than those obtained with traditional methods. The model's parameter efficiency significantly reduces computational costs. Therefore, it is expected that the KAN model will find more applications in energy demand forecasting and time series analysis in the future. Furthermore, it is anticipated that future studies in the literature will further advance the success of KAN by testing it with larger datasets and different independent variables.

6. Conclusions

In this study, time series forecasting was performed using seven years of natural gas consumption data from Istanbul. Both the traditional Multi-Layer Perceptrons (MLP) model and a new method, Kolmogorov-Arnold Networks (KAN), were employed. The forecast results compared the performance of both models, with particularly high accuracy rates achieved in many variations of the KAN model. The high performance of the KAN model, despite its low number of parameters, demonstrates that it provides an effective solution for energy demand forecasting and time series analysis. Only historical natural gas consumption data were used during the study. Despite this, the KAN model delivered highly successful results,

particularly in energy demand forecasting. However, to further enhance the success of the KAN model, it would be beneficial to incorporate meteorological data and additional information regarding subscriber behavior. This study shows that KAN presents a powerful model that can serve as an alternative to traditional methods in energy demand forecasting, such as natural gas consumption.

References

- [1] K. Arslan, M. Akpınar, and M. Fatih Adak, "The detection of unaccounted natural gas consumption: A neural networks and subscriber-based solution," *Engineering Science and Technology, an International Journal*, vol. 52, p. 101669, Apr. 2024, doi: 10.1016/j.jestch.2024.101669.
- [2] N. Wei *et al.*, "Data complexity of daily natural gas consumption: Measurement and impact on forecasting performance," *Energy*, vol. 238, 2022, doi: 10.1016/j.energy.2021.122090.
- [3] M. Akpınar, M. F. Adak, and N. Yumusak, "Forecasting natural gas consumption with hybrid neural networks — Artificial bee colony," in *2016 2nd International Conference on Intelligent Energy and Power Systems (IEPS)*, IEEE, Jun. 2016, pp. 1–6. doi: 10.1109/IEPS.2016.7521852.
- [4] M. Akpınar and N. Yumusak, "Year ahead demand forecast of city natural gas using seasonal time series methods," *Energies*, vol. 9, no. 9, 2016, doi: 10.3390/en9090727.
- [5] M. Akpınar and N. Yumusak, "Estimating household natural gas consumption with multiple regression: Effect of cycle," in *2013 International Conference on Electronics, Computer and Computation, ICECCO 2013*, 2013, doi: 10.1109/ICECCO.2013.6718260.
- [6] M. Akpınar and N. Yumusak, "Forecasting household natural gas consumption with ARIMA model: A case study of removing cycle," in *AICT 2013 - 7th International Conference on Application of Information and Communication Technologies, Conference Proceedings*, 2013, doi: 10.1109/ICAICT.2013.6722753.
- [7] Q. Wang, S. Liu, and H. Yan, "The application of trigonometric grey prediction model to average per capita natural gas consumption of households in China," *GS*, vol. 9, no. 1, pp. 19–30, Feb. 2019, doi: 10.1108/GS-08-2018-0033.
- [8] Z. Mi *et al.*, "China's Energy Consumption in the New Normal," *Earth's Future*, vol. 6, no. 7, 2018, doi: 10.1029/2018EF000840.
- [9] G. De and W. Gao, "Forecasting China's natural gas consumption based on adaboost-particle swarm optimization-extreme learning machine integrated learning method," *Energies*, vol. 11, no. 11, 2018, doi: 10.3390/en11112938.
- [10] X. Wang, D. Luo, J. Liu, W. Wang, and G. Jie, "Prediction of natural gas consumption in different regions of China using a hybrid MVO-NNGBM model," *Mathematical Problems in Engineering*, vol. 2017, 2017, doi: 10.1155/2017/6045708.
- [11] A. S. Anđelković and D. Bajatović, "Integration of weather forecast and artificial intelligence for a short-term city-scale natural gas consumption prediction," *Journal of Cleaner Production*, vol. 266, 2020, doi: 10.1016/j.jclepro.2020.122096.
- [12] W. Qiao, Z. Yang, Z. Kang, and Z. Pan, "Short-term natural gas consumption prediction based on Volterra adaptive filter and improved whale optimization algorithm," *Engineering Applications of Artificial Intelligence*, vol. 87, p. 103323, Jan. 2020, doi: 10.1016/j.engappai.2019.103323.
- [13] Z. Liu *et al.*, "KAN: Kolmogorov-Arnold Networks," 2024, *arXiv*. doi: 10.48550/ARXIV.2404.19756.
- [14] C. J. Vaca-Rubio, L. Blanco, R. Pereira, and M. Caus, "Kolmogorov-Arnold Networks (KANs) for Time Series Analysis," 2024, *arXiv*. doi: 10.48550/ARXIV.2405.08790.
- [15] M. H. Sulaiman, Z. Mustaffa, A. I. Mohamed, A. S. Samsudin, and M. I. Mohd Rashid, "Battery state of charge estimation for electric vehicle using Kolmogorov-Arnold networks," *Energy*, vol. 311, p. 133417, Dec. 2024, doi: 10.1016/j.energy.2024.133417.
- [16] M. H. Sulaiman, Z. Mustaffa, M. S. Saecalal, M. M. Saari, and A. Z. Ahmad, "Utilizing the Kolmogorov-Arnold Networks for chiller energy consumption prediction in commercial building," *Journal of Building Engineering*, vol. 96, p. 110475, Nov. 2024, doi: 10.1016/j.job.2024.110475.
- [17] Y. Peng *et al.*, "Predictive modeling of flexible EHD pumps using Kolmogorov–Arnold Networks," *Biomimetic Intelligence and Robotics*, vol. 4, no. 4, p. 100184, Dec. 2024, doi: 10.1016/j.birob.2024.100184.
- [18] F. Granata, S. Zhu, and F. Di Nunno, "Advanced streamflow forecasting for Central European Rivers: The Cutting-Edge Kolmogorov-Arnold networks compared to Transformers," *Journal of Hydrology*, vol. 645, p. 132175, Dec. 2024, doi: 10.1016/j.jhydrol.2024.132175.

- [19] IGDAS, “Monthly Natural Gas Consumption by District.” Accessed: Oct. 09, 2024. [Online]. Available: <https://data.ibb.gov.tr/>
- [20] X. Feng, G. Ma, S.-F. Su, C. Huang, M. K. Boswell, and P. Xue, “A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan,” *Ocean Engineering*, vol. 211, p. 107526, Sep. 2020, doi: 10.1016/j.oceaneng.2020.107526.
- [21] F. Taşpınar, N. Çelebi, and N. Tutkun, “Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods,” *Energy and Buildings*, vol. 56, pp. 23–31, Jan. 2013, doi: 10.1016/j.enbuild.2012.10.023.
- [22] B. Soldo, P. Potočnik, G. Šimunović, T. Šarić, and E. Govekar, “Improving the residential natural gas consumption forecasting models by using solar radiation,” *Energy and Buildings*, vol. 69, pp. 498–506, Feb. 2014, doi: 10.1016/j.enbuild.2013.11.032.
- [23] J. Szoplik, “Forecasting of natural gas consumption with artificial neural networks,” *Energy*, vol. 85, pp. 208–220, Jun. 2015, doi: 10.1016/j.energy.2015.03.084.
- [24] M. Akpınar, M. Adak, and N. Yumusak, “Day-Ahead Natural Gas Demand Forecasting Using Optimized ABC-Based Neural Network with Sliding Window Technique: The Case Study of Regional Basis in Turkey,” *Energies*, vol. 10, no. 6, p. 781, Jun. 2017, doi: 10.3390/en10060781.
- [25] Q. Wang and F. Jiang, “Integrating linear and nonlinear forecasting techniques based on grey theory and artificial intelligence to forecast shale gas monthly production in Pennsylvania and Texas of the United States,” *Energy*, vol. 178, pp. 781–803, Jul. 2019, doi: 10.1016/j.energy.2019.04.115.
- [26] E. Fabbiani, A. Marziali, and G. D. Nicolao, “Ensembling methods for countrywide short term forecasting of gas demand,” *IJOGCT*, vol. 26, no. 2, p. 184, 2021, doi: 10.1504/IJOGCT.2021.10035077.
- [27] R. Hribar, P. Potočnik, J. Šilc, and G. Papa, “A comparison of models for forecasting the residential natural gas demand of an urban area,” *Energy*, vol. 167, pp. 511–522, Jan. 2019, doi: 10.1016/j.energy.2018.10.175.
- [28] W. Qiao, K. Huang, M. Azimi, and S. Han, “A Novel Hybrid Prediction Model for Hourly Gas Consumption in Supply Side Based on Improved Whale Optimization Algorithm and Relevance Vector Machine,” *IEEE Access*, vol. 7, pp. 88218–88230, 2019, doi: 10.1109/ACCESS.2019.2918156.

Article Information Form

Authors Contributions

The first author contributed to data collection, modeling, data analysis, evaluation of results, visualization, and revision. The second author contributed to modeling, data analysis, and discussion of results.

Acknowledgments

We gratefully acknowledge IGDAS (Istanbul Gas and Natural Gas Distribution Co.) and the Open Data Portal for providing the data used in this study.

Conflict of Interest Notice

The author has no conflicts of interest to declare.

Ethical Approval

It is declared that during the preparation process of this study, scientific and ethical principles were adhered to, and all studies cited are listed in the bibliography.

Availability of data and material

Data: <https://data.ibb.gov.tr/>

Artificial Intelligence Statement

No artificial intelligence tools were used while writing this article.

Plagiarism Statement

This article has been scanned by iThenticate™.