A Comparative Study of Statistical and Deep Learning Models Used in Energy Load Prediction

Eralda Gjika[1, *], Lule Basha[2]

[1]Department of Applied Mathematics, Faculty of Natural Science, University of Tirana, Albania
[2] Department of Applied Mathematics, Faculty of Natural Science, University of Tirana, Albania
*Eralda Gjika

Abstract.

The objective of this study is to analyse and compare classical time series and deep learning models for energy load prediction. Energy predictions are important for management and sustainable systems. After analysing the climacteric factors impact on energy load (a case study in Albania) we considered classical and deep learning models to perform forecasts. We have used hourly and daily time series for a period of three years. In total respectively 26,280 hours and 1095 days. Average temperature is considered as external variable in both statistical and deep learning models. The dynamic evolution of hourly (daily) load is correlated with hourly (daily) average temperature. The performance of the proposed models is analysed and evaluated based on accuracy measurements (MSE, RMSE, MAPE, AIC, BIC etc.) and graphics results of statistical tests. In-sample and out-of-sample accuracy is evaluated. The models show competitive performance to some recent works in the field of short-and medium-term energy load forecasts. This study addresses the critical challenges of predicting energy load, providing valuable insights for sustainable energy management in the Balkan region. Our findings highlight the effectiveness of statistical and machine learning models, with STL decomposition and ARIMA error emerging as a top-performing model for out-of-sample testing. These results offer practical guidance for short-term and medium-term energy load predictions, with the potential for further improvement by incorporating additional exogenous variables.

Keywords: time series; forecasting; energy consumption; deep learning.

Introduction

The Mediterranean basin is one of the key points of energy efficiency production and use. Every country's energy consumption is specially affected by its economic and industry development, climatic conditions and its energy production sources. Numerous and diverse sources of energy have undergone significant evolution in the last 30 years. There has been a decline in coal use and a significant increase in natural gas use. Although climate change has affected this region over the last decades again the Mediterranean basin is an area which benefits from a mild climate with mild and warm winters and sunny summers. This climate offers the region a great potential for energy production from renewable energy (Soukissian et al., 2017). Albania is a country in which the energy produced by hydropower plants occupies almost 90% of the energy produced in the country. Given that this energy produced relies on the availability of water in large reservoirs of cascades located mainly in the northern part of the country, or the intensity of the flow of rivers that supply these cascades, precipitation and snowmelt. High summer temperatures and droughts are limiting production by hydropower plants.

What is noticed in recent years in the Mediterranean region and in Albania is also the fact that based on the above factors the utilization rate of hydropower plants has decreased. This decline has been followed by an increase in interest in solar energy which is mainly influenced by surface solar radiation whose variations depend mainly on the atmospheric composition (aerosols, water vapour) and clouds (Gjika et al., 2021A; 2021B; 2021C).

An increase in solar radiation has been observed in Europe (Sánchez-Lorenzo et al., 2017) and especially in the Mediterranean basin these solar sources are seen with special interest as one of the areas with medium to high solar radiation on the continent (Hadjipanayi et al., 2016). Exactly at the beginning of winter in 2021, the region was involved in an energy crisis and not only. Experts emphasize the importance of safe and sufficient energy, especially when the energy sources are not numerous and diverse. In this context, they suggest the addition of new and clean energy sources, in the same time they highlight to focus on the importance of optimal management of existing resources. Climate change associated with drought can reduce power generation and result in less electricity produced by the hydropower plants. Significant changes in production and consumption have been observed which have influenced government policies to provide optimal and long-term solutions. Many European countries are part of the energy crisis and have already had wide-ranging impacts on their economy and environment.

There is a lot of work done regarding prediction in different areas. In their work (Allen & Morzuch, 2006), have presented most of the challenges the prediction field has faced with during 25 years of forecasting. More than one decade ago they pointed out the necessity of computational ability for the high complexity amount of data to become the power of prediction in many areas. The relationship between energy consumption and economic growth was analysed in a considerable number of countries in Europe (Topolewski, 2021). They indicate that attention is required to the relation between the efficiency use of resources and climate change in consequence the global warming. Researchers are provided with a systematic literature review of a considerable number of articles on energy demand modelling. Reference (Verwiebe et al., 2021) reviewed and offered a classification of different techniques used in energy demand. There is also a lot of work done especially in machine learning (ML) techniques which rely on historical data and are extensively used to short-term forecasting (Kalimoldayev et al., 2020). Classical statistical techniques on energy prediction are often used as a benchmark for many techniques but in many cases depending on the nature of the data used and exogenous variables these techniques perform comparable with engineering-based models or ML models (Gjika et al., 2021A;2021B).

Given the high ability of deep learning techniques to deal with the change in power generation and load there are many neural network structures that have been utilized to obtain short-term load predictions (Machado et al., 2021). The study undertaken by (Eskandari et al., 2021) presents a forecasting model for hourly load consumption considering external variables unitizing convolutional neural networks (CNNs) to extract the features of variables used. They also show in their study their ability to deal with short-term and long-term memories. ANNs (Artificial Neural Networks) are used (Arvanitidis et al., 2021) for Greek interconnected power system. They point out that the accuracy of the ANNs' prediction depends on the quality and availability of the training data. An analysis of the accuracy of ML and statistical techniques for the Albanian energy sector is done by (Gjika et al., 2021B). In their work, they consider the energy production by hydropower which is the main source of production in the country. They came to the conclusion that neural networks can handle seasonal patterns of monthly energy produced by HPP and provide accurate forecasts in the short term. Reference (Hou et al., 2021) proposed a short-term load forecasting method for hourly data using long short-term memory (LSTM) algorithm as an algorithm that has been shown to deal with the regularity of historical data. They use encoded external factors to predict the load in the next half an hour and showed accurate results.

In their work (Taylor & Buizza, 2003) present forecasting methodology for daily electricity demand using weather ensemble predictions. They show that weather ensemble predictions can improve the accuracy of electricity demand forecasts. When forecasting energy demand, it's often a good practice to use temperature as an exogenous variable.

Reference (Gjika et al., 2021, B) present a methodology of ensemble models to predict energy production by hydropower relying on exogenous variables such as temperature, precipitation, and water inflow. They show accurate results by combining statistical and machine learning models for monthly data.

Depending on many factors when dealing with energy data studies have shown that in special circumstances such as geographical position, climate conditions, variables taken into consideration, seasonality patterns, and frequency of data there is no consensus on the "best" model used in the energy situations. Going through the results of (Ouyang et al., 2021), they show the efficiency of STL decomposition (Seasonal-Trend decomposition using LOESS) when used as a pre-processing step in statistical models. Another study that shows the efficiency of statistical time series models is the one proposed by (Xie & Ding, 2020) which is a simple procedure combining time series models dealing with multi-seasonality. In reference (De Livera et al., 2011) the authors offer a new approach for forecasting time series with complex seasonal.

Although there is plenty of material, research, and competition about recommended models for forecasting in different fields (Hyndman, 2020) there is still discussion of the conditions under which different methods perform best.

1.1 Data

In our study we use hourly time series of energy load for a period of three years (2016-2018) in total there are 26,280 hours and 1095 days offered by (Sevdari et al., 2022). Together with the energy load we have also considered the average temperature (hourly and daily). In this material we have used the terms described below:

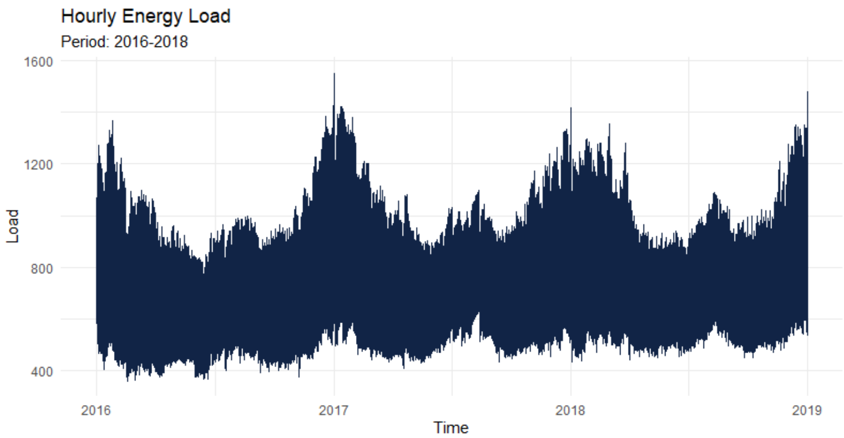
Hour: The time of day for which the variables are expressed. The time is expressed as an integer with values ranging from 0 to 23.

Load: The aggregated energy load (consumption) observed each hour (measured in Mwh).

Temperature: The hourly average value (in Celsius) of the temperature of the day.

When dealing with hourly time series data there are a few models that one can try. Since hourly time series contain multiple seasonal patterns (daily, weekly, and yearly); in your case it contains all these seasonality's because it contains 3 years of hourly data. Many time series exhibit complex seasonal patterns.

Figure 1: Hourly energy load (unit MWh)



Source: (Authors)

Figure 1 shows the hourly energy load for a period of three years. What is clearly observed is the fact that the time series has multi-seasonality patterns. After an accurate investigation of each year, we observe a high load at the beginning of the year which corresponds to the winter season and is accompanied by a noticeable decrease during the spring season. Further an upward trend for the period of summer which in the Mediterranean climate is accompanied by high temperatures, and with a marked decline during the autumn. Patterns are distinct from year to year due to the extreme temperatures and weather situations during the winter months mainly in the entire Mediterranean region and especially in Albania for that year. This behaviour can be observed in Figure 2, Figure 3 and Figure 4.

Figure 2: Hourly energy load (2016)

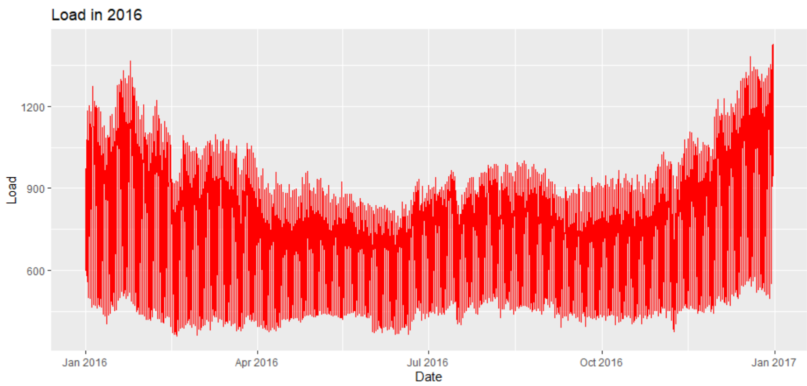


Figure 3: Hourly energy load (2017)

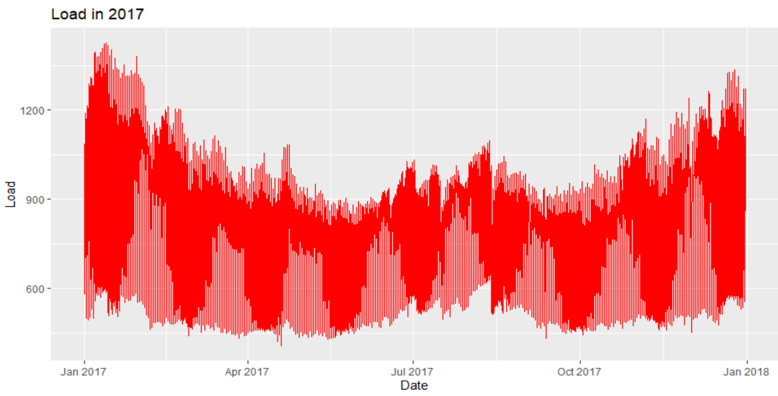
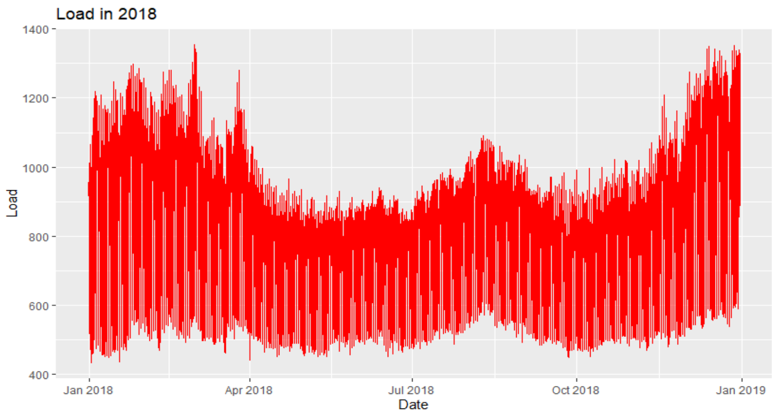


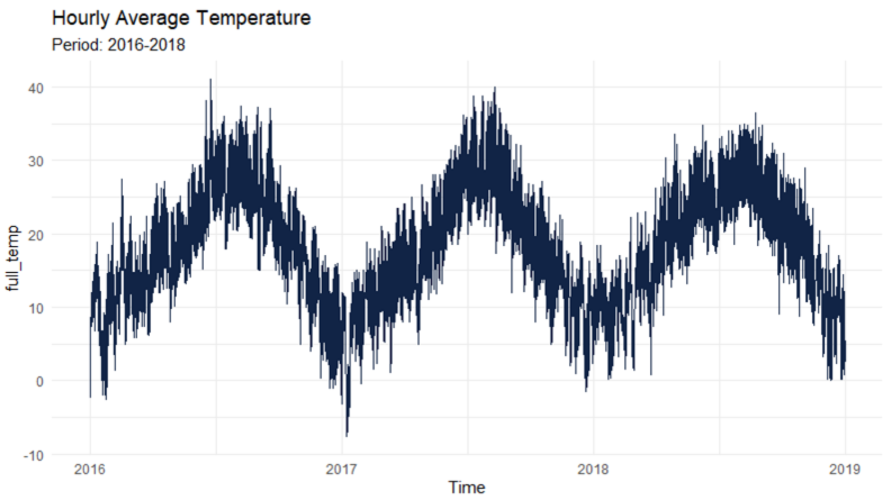
Figure 4: Hourly energy load (2018)



Source: (Authors)

The hourly average temperature (measured in Celsius degree) for the period under consideration is shown in Figure 5. The Mediterranean climate would certainly not be absent in the seasonal behaviour of the average temperature. High temperatures are observed during the summer months (up to 42 degrees Celsius) and low temperatures during the winter months (up to -7 degrees Celsius).

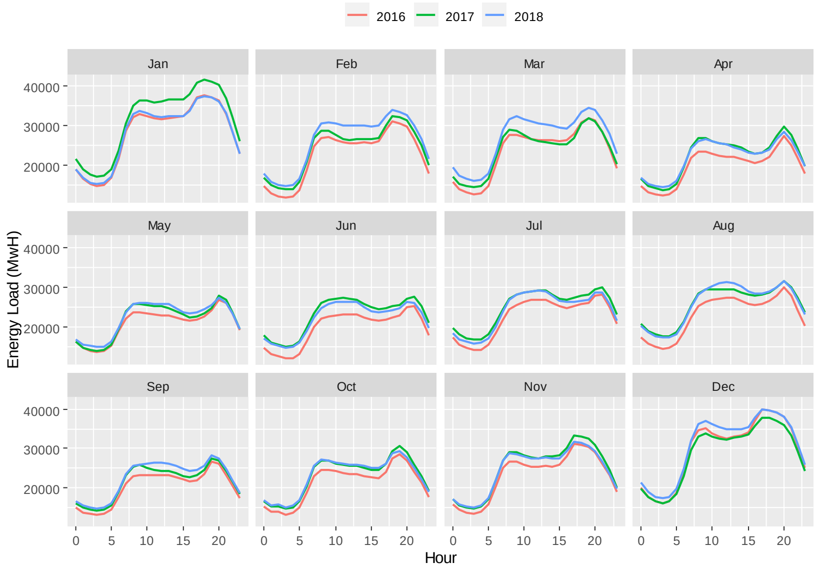
Figure 5: Hourly average temperature (Celsius degree)



Source: (Authors)

These extreme values are especially noticeable during 2016, and there is a decrease in values for both high and low temperatures for 2017 and further for 2018 which is also confirmed by world indicators related to climate change and global warming. Figure 6 shows the different levels of daytime peak energy demand by month. The situation is displayed for three years in separate, and we may notice a flattened pattern from May to September and a clear three peaks for other months. The lower peak in energy demand is observed from midnight to 5am and then a rapid increase of the demand from 5am to 8am, then a steady situation which culminates with the evening hour 8pm and then a decrease again to the lowest levels of the day. Especially for January and December the morning “jump” load is more distinctive and very fast in levels from 20000 Mwh to 35000 Mwh. It is also observed a slightly increase of energy load levels from 2016 to 2018.

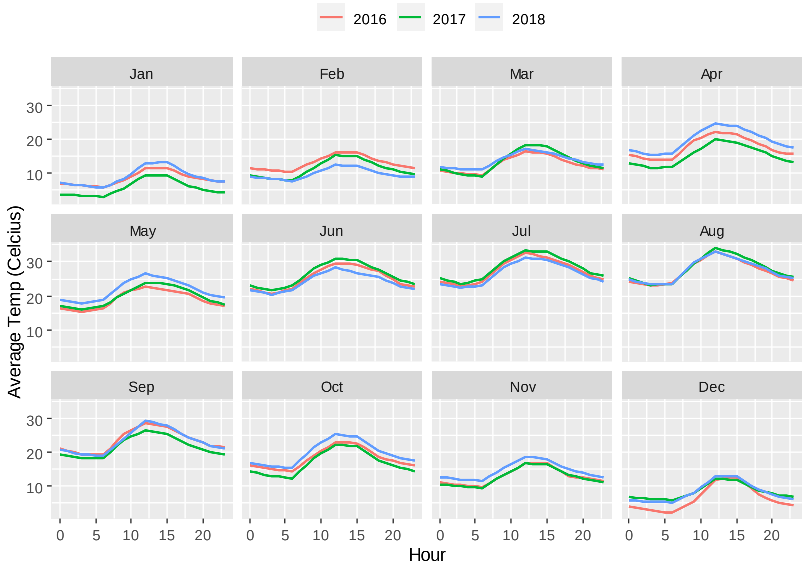
Figure 6: Twenty-four-hour load by month (MWh)



Source: (Authors)

The twenty-four-hour load helps us to investigate the levels of daytime and peak loads which depends also on the solar penetration conditions and variations.

Figure 7: Twenty-four-hour load by month (MWh)



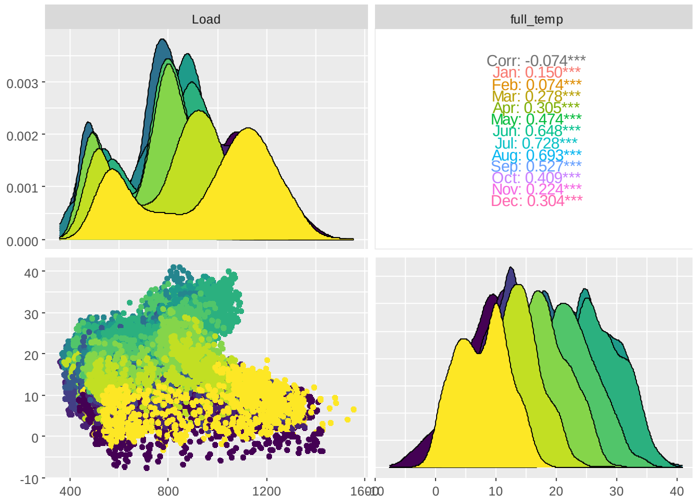
Source: (Authors)

The Mediterranean climate of Albania shows a correlation between energy demand and average temperature. This can be easily observed also by 24 hour evolution of these variables faced by month as shown in Figure 6 and Figure 7. In both variables, we observe the presence of seasonality and one high peak for the average temperature which obviously is reached in midday and high average temperature levels for the summer and low levels for winter. No noticeable differences in levels from 2016 to 2018.

A correlation analysis of energy load and climacteric factors such as temperature is important. Pearson correlation coefficient is a measure of linear correlation between two sets of data. It is defined as: and takes values from -1 to 1. A value close to -1 or 1 indicates a significant (negative or positive) relationship between X and Y.

This correlation analysis between load and temperature is illustrated graphically in Figure 8. The interesting part observed is the fact that aggregated energy load displays a significant correlation with the average temperature (by hour) when observed by month. We notice a significant positive correlation between these two variables especially for the summer season (it varies from 0.64 up to 0.72). Another pattern we clearly observe is the density plot which corresponds to daily and night hours of the day.

Figure 8: Density plot and correlation of energy load (consumption) and average temperature by month

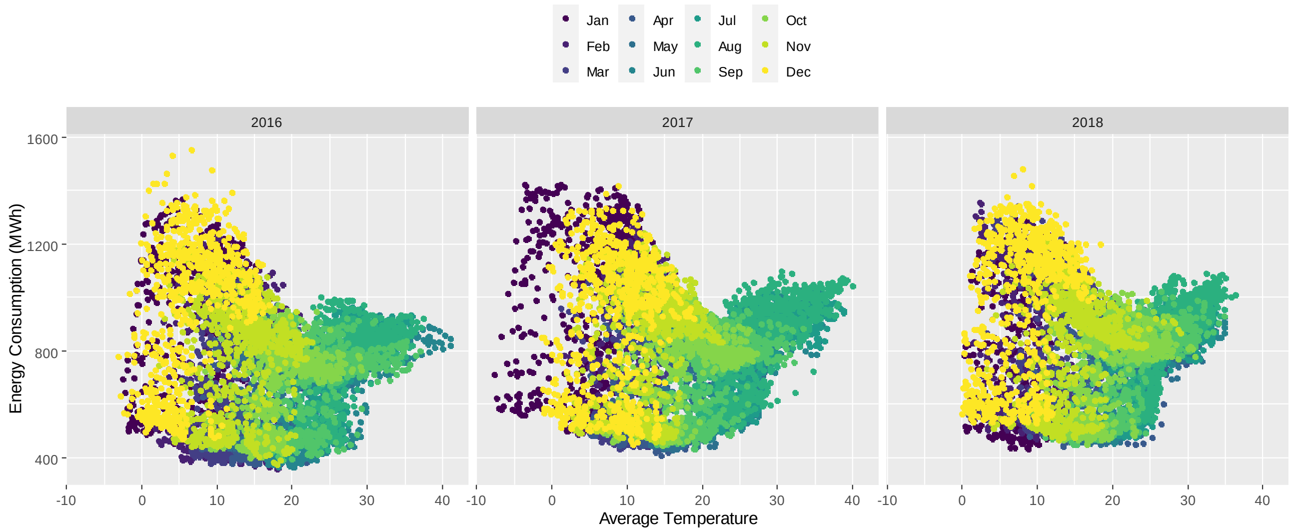


Source: (Authors)

These findings may be used to focus on the research of seasonal power load prediction in order to satisfy and optimize power supply and demand. In this study we have not taken into consideration the seasonal modeling by hour or month which can be further studied.

Figure 9 shows the correlation of energy load (consumption) and average temperature faced by year and also colored by month. The same behavior is observed for each year taken into observation. What makes different the spread are the registered values of the average hourly temperature which clearly display a compression in the amplitude for the last year 2018. The scatterplot of hourly energy load and average temperature shows two clusters that correspond to daily and night hours. The behavior is almost the same apart from the shift of the daily cluster above the night cluster. This shift corresponds also to the higher differences in temperature which suggests the need for electricity due to heating or cooling with respect also to the month or season.

igure 9: Correlation of energy consumption and average temperature faced by year (hourly observations)



Source: (Authors)

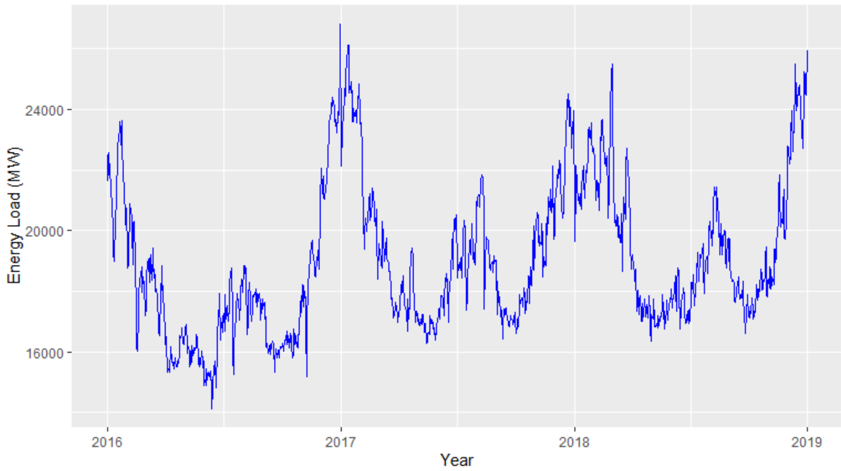
2 Methodology

1.1 Data preprocessing

The data was organized in training and testing datasets respectively (80% and 20%). The testing dataset was used for validation which corresponds to the forecast horizon. Some machine learning methods face difficulties when dealing with missing observations, but this was not the case in our data. The daily energy load was the aggregated load of 24 hours, and the average temperature was the average calculated for 24 hours of that corresponding day. Given the complexity of the data and the lack of other external variables that can handle and better explain multi-seasonal patterns of energy load, we switched to daily aggregated time series of energy load and daily average temperature.

sed on the above analysis of the data and the literature review we have proposed some models which can deal with the multi-seasonality pattern of the energy load, and that can handle exogenous variables. In our case, we have tried the average daily temperature as an external variable in some of our models.

Figure 10: Daily energy load ((aggregated energy load MW/day))



Source: (Authors)

1.2 Statistical Techniques

First, we focused our attention on statistical forecasting models such as: Naïve, moving average, ARIMA, Exponential Smoothing, and Seasonal Naïve. In many situations is shown that STL is a successful try at decomposing the time series into their seasonal, trend, and remainder components. After that STL can be used for modeling purposes. In this study we have used the *Hyndman-Khaldakar* algorithm for performing STL and ARIMA models (Hyndman, 2008A; 2008B; 2021). In reference to (Ouyang et al., 2021) the STL decomposition shows good performance when used in statistical methods and time series with monthly seasonality but was not performing well in machine learning methods. The choice of the best algorithm depends on the nature of the data and the frequency as well. In the presence of seasonality patterns and trends ARIMA and exponential smoothing methods are the ones that can perform well in prediction. In their work (Hyndman et al., 2008B) present a complete modeling framework for time series exponential smoothing models.

The autoregressive integrated moving average (ARIMA(p,d,q)) processes are a combination of autoregressive (AR(p), where p is the degree of the autoregressive model) and moving average (MA(q), where q- degree of moving average model) processes and d is the degree of differences (Newbold, 1983). The comprehensive model can be expressed as follows:

$$y'_t = c + \phi_1 y'_{t-1} + \phi_2 y'_{t-2} + \cdots + \phi_p y'_{t-p} + \theta_1 \varpi_{t-1} + \cdots + \theta_q \varpi_{t-q} + \varpi_t$$

Where p represents the order of the autoregressive component; d signifies the degree of initial differencing applied; q denotes the order of the moving average component. ϕ_i are the autoregressive coefficients; θ_j are the moving average coefficients; ϖ_k represents white noise, which is inherent randomness in the data.

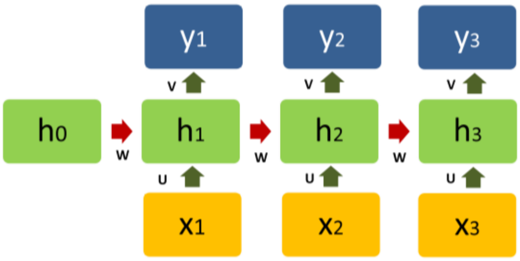
For the implementation of ARIMA models in R, we have used *forecast* package in R which combines unit root tests, minimization of the AICc and MLE to obtain the ARIMA parameters, and coefficient estimates (Hyndman & Athanasopoulos, 2021). There are many models that use STL (Seasonal and Trend decomposition using Loess- a method for estimating nonlinear relationships.) to understand seasonal data and fit appropriate models (Cleveland et al., 1990).

1.3 Deep Learning Techniques

Machine learning techniques and deep learning are attracting more and more attention from researchers in many fields. Especially in the forecasting field these methods have passed through many competitions such as the M Competitions (Makridakis & Hibon, 2000) (Makridakis et al., 2018A; 2018B). Artificial neural networks are forecasting methods that are based on simple mathematical models of the brain. They allow complex nonlinear relationships between the response variable and its predictors. There are many studies of using deep learning methods in energy prediction and reviewed by (Hewamalage et al, 2021). In their study (Mandal et al., 2006) present a neural network approach for short-term energy load prediction paying attention to seasons and using temperature as an external variable. They achieved reliable results for hour-ahead load prediction. In reference to the work presented by (Smyl, 2020) they agreed on the weakness of NNs when dealing with seasonality. Many researchers suggest removing seasonality before modeling, to achieve better predictions. Testing was made by (Hewamalage et al., 2021) on this topic and they showed that for clear seasonality patterns, RNNs are adequate but when this is not the case then a de-seasonal technique should be used.

In this study, we have considered Recurrent Neural Nets (RNNs). The scheme of how this network performs is shown in Figure 11.

Figure 11: RNN Architecture



Source: (Authors)

Here, *x's* in yellow are predictor variables, *h's* in green are hidden layers and *y's* in blue are predicted values.

Recurrent Neural Nets are essentially a bunch of neural nets stacked on top of each other. The output of the model at h_1 feeds into the next model at h_2 as shown. The goal of the learning process is to find the best weight matrices *U*, *V*, and *W* that give the best prediction of *y^(t)*, starting from the input *x(t)*, of the real value *y(t)*.

Neural Network AutoRegression (NNAR) models are developed on the principle of using lagged observations as inputs to a neural network. They are feed-forward networks with one hidden layer. These models perform good with seasonal data, where it add as input the last observed values from the same season. In general, the model *NNAR(p, k)* uses *p-lagged* inputs and k nodes in the single hidden layer. Seasonal *NNAR(p, P, k):* with *k*-neurons in the hidden layer and input $(y_{t-1,}y_{t-2},\ldots,y_{t-p},y_{t-m},y_{t-2m},y_{t-Pm})$. *NNAR(p, P, 0)m* model is equivalent to an *ARIMA(p, 0, 0)(P,0,0)m* model but without stationarity restrictions. More generally, an *NNAR(p,P,k)m* model has inputs :

$(y_{t-1,}y_{t-2},\ldots,y_{t-p},y_{t-m},y_{t-2m},y_{t-Pm})$ and *k* neurons in the hidden layer. If the values of *p* and *P* are not specified, they are selected automatically (Hyndman & Athanasopoulos, 2021). For seasonal time series, the default value of *P* is 1, and *p* is chosen from the optimal linear model fitted to the data after seasonally adjusted. If k is not specified, then it is set to *k=(p+P+1)/2* (which is rounded to the nearest integer).

1.4 Evaluation metrics

The performance of the models presented in this study was evaluated in terms of several metrics. The selection of the most accurate model is made by analyzing and comparing error measurements and information criteria for in-sample and out-of-sample data. As well as extending personal judgment to the advantages offered by each model based on the nature of the data. The metrics used to assess and compare the various methods are:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | \hat{X}_t - X_t |$$ Maximum Absolute Error

$$ME = \frac{1}{n}\sum_{i=1}^{n} (\hat{X}_t - X_t)$$ Mean Error

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n} \frac{| \hat{X}_t - X_t |}{| X_t |}\right) \cdot 100\%$$ Mean Absolute Percentage Error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{X}_t - X_t)^2}$$ Root Mean Square Error

$$MASE = \frac{MAE}{MAE_{in-sample,naive}}$$ Mean Absolute Scaled Error

In many research studies, there are many arguments of using different accuracy measurements of the model. This depends of course on the nature of the data and their complexity. In reference to (Hyndman & Koehler, 2006) MASE offers a straightforward indication of the relative model performance compared with the naïve benchmark. It is a scale-independent measure where a value less than one indicates that the performance of the model is better than the naïve benchmark on average. A value greater than one indicates the opposite. What is important is the fact that this critical value should not conclude the performance of the model but further analysis are suggested.

2 Analysis of the results

This section provides a comprehensive analysis of the results obtained from the modeling process. Results in terms of all error metrics used to evaluate the performance of the models are shown in Table 1. The abbreviations used to denote the “top model” selected from the work done in this study are respectively: ‘Snaive’- seasonal naïve method, ‘STL+ARIMA’- STL decomposition with ARIMA errors, ’Hybrid’- ensemble model with combination of statistical and deep learning models, 'NNAR'- Neural network with autoregression, 'NNAR-Xreg'- neural network with autoregression and average temperature as external variable.

In this study, we present the results of our models for both testing and training data as a fundamental measure to gain valuable insights into the performance and accuracy of each model. This approach allows us to evaluate how well the models generalize to new, unseen data (testing data) compared to the data they were originally trained on (training data). Assessing and reporting model performance on both in-sample and out-of-sample data is a critical aspect of time series analysis. The comparison between in-sample and out-of-sample metrics performance provides insights into whether the model has overfit (performed well on training data but poorly on testing data) or underfit (performed poorly on both training and testing data).

In Table 1, the best model based on the error is indicated in boldface respectively for training and testing dataset. Analyzing the values of error metrics for each model we observe that for the training datasets STL+ARIMA seem to perform better than seasonal naïve. On other hand, NNAR with daily average temperature as a regressor seems to perform better than NNAR without regressors. The difference between NNAR and NNAR-Xreg is not significant. In this situation, we may suggest adding other exogenous variables (such as humidity) to explain the daily energy load. Overall, for the training dataset the neural network with average temperature as an external variable is significantly better compared to other statistical models.

The MASE value for all proposed models is lower than 1 which suggests that all the models perform better than the naïve benchmark on average. The situation changes apparently for the testing data where we have approximately seven months of observations (20% of the three years taken into consideration). Investigating the lowest value of error measurements in this part STL+ARIMA shows significantly better performance compared to the other models. MASE is higher than 1 but close to this value. Comparing the MASE value of STL+ARIMA for the training data and testing data we may gain confidence that this model outperforms the other models. Between NNAR and NNAR with exogenous variable, the first has a slight difference in error values.

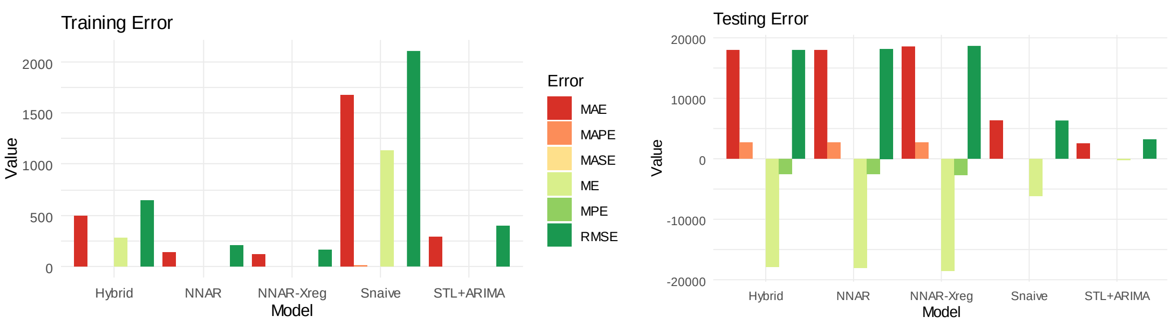
For a better understanding and comparison of the error metrics for training and testing data, we plotted the performance displayed in Figure 12.

Table 1: Model performance error metrics for training and testing dataset

Model	Train						Test					
	ME	**RMSE**	**MAE**	**MPE**	**MAPE**	**MASE**	**ME**	**RMSE**	**MAE**	**MPE**	**MAPE**	**MASE**
Snaive	1134	2105	1671	5.67	8.28	1	-6255	6255	6255	-35.36	35	3.74
STL+ARIMA	6.98	401	298	0	1.57	0.18	**-260**	**3185**	**2405**	**-2.58**	**12**	**1.44**
Hybrid	277	646	501	1.36	2.49	0.3	-17959	18008	17959	-2561	2561	10.7
NNAR	1.16	201	138	-0.02	0.7	0.08	-18048	18132	18048	-2577	2577	10.8
NNAR-Xreg	**0.93**	**160**	**115**	**-0.01**	**0.6**	**0.07**	-18529	18620	18529	-2645	2645	11.1

Source (Authors)

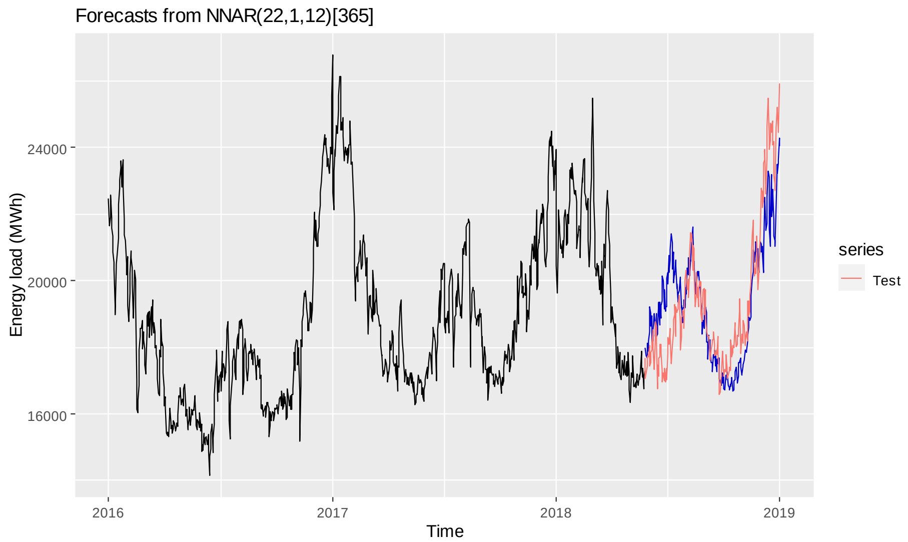
Figure 12: Model performance error metrics for training and testing dataset



Source: (Authors)

On the left side of Figure 12 are displayed the metrics for the training data and on the right are the metrics for testing data. The NNAR with external variables outperformed the traditional univariate methods in the training dataset and it is comparable to the hybrid model. The hybrid model was obtained as a combination with equal weights of four models: nnetar, stlm, tbats and snaive (Shaub, 2020) .

Figure 13 Energy load forecast from NNAR with daily average temperature as regressor (aggregated energy load MW/day)



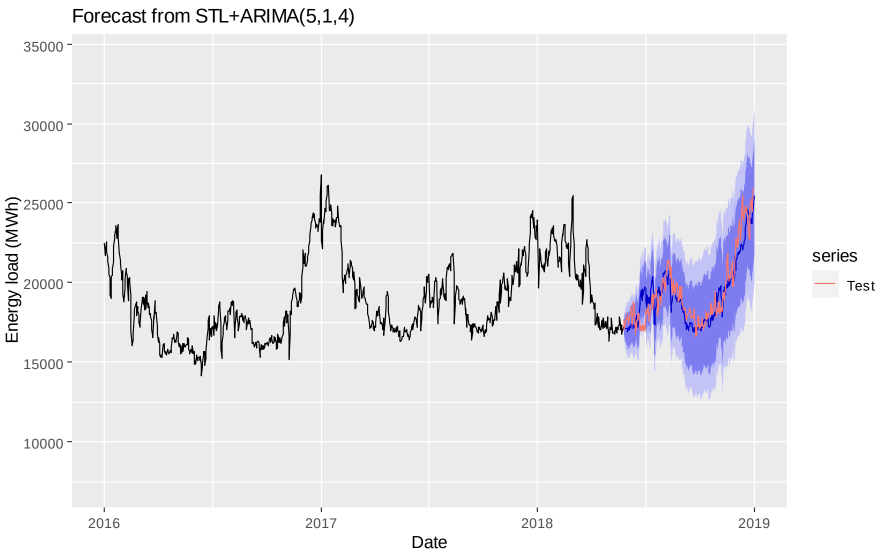
Source: (Authors)

Figure 13 displays the daily energy load prediction obtained from a neural network model where the daily average temperature is used as an external variable.

As we mentioned above these models perform well with seasonal data, where the last observed values from the same season are added as input. For energy load, the model has k=12 neurons in the hidden layer and use $(y_{t-1,}y_{t-2},...,y_{t-p},y_{t-m},y_{t-2m},y_{t-Pm})$ observations as input where *p=22,P=1* and *m=365* daily seasonality.

Figure 14 shows the energy load point forecast and confidence intervals (80% and 95%) from *STL+ARIMA* model.

Figure. 14: Energy load forecast from STL+ARIMA model with no-regressor (aggregated energy load MW/day)



Source: (Authors)

3 Conclusions

Energy supply and demand plays an important role in the economy of a country and the region. Predictions are important for energy management and sustainable systems. The motivation for this study was to address some of the key issues related to the ability to predict energy load using statistical and deep learning models. The work presented here can be used as a reference for researchers and practitioners working in the energy field, especially in the Balkan region.

We showed that due to the high complexity of the hourly data and multiple seasonality, it was easy at the start of our analysis to work with daily data. We performed many statistical and machine learning models that are capable of handling seasonality in time series. In our work, we took into consideration a decomposition method that would give better performance in the modeling process. The models were evaluated on error metrics and a comparative view of in-sample and out-of-sample dataset. NNAR architecture was able to outperform the statistical techniques for in-sample data in terms of all error metrics used at the performance evaluation phase but STL decomposition with ARIMA error was the best model when evaluated to the testing data. The proposed models can be used as a short-term or medium-term prediction model for energy load. Other exogenous variables can have a better effect on the models.

Acknowledgment

The authors want to thank Faculty of Natural Science, University of Tirana, Albania which has financially supported the presentation of this work at the conference.

References

Arvanitidis A. I., Bargiotas D., Daskalopulu A., Laitsos V. M., and Tsoukalas L. H., (2021). "Enhanced Short-Term Load Forecasting Using Artificial Neural Networks." Energies, vol. 14, 7788, https://doi.org/10.3390/en14227788

Cleveland R. B., Cleveland W. S., McRae J. E., and Terpenning, (1990). "STL: A seasonal-trend decomposition procedure based on loess.", Journal of Official Statistics, vol. 6(1), pp. 3–33, http://bit.ly/stl1990

De Livera A. M., Hyndman R. J., and Snyder R.D., (2011). "Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing.", Journal of the American Statistical Association, vol. 106(496), pp. 1513–1527. http://www.jstor.org/stable/23239555

Eskandari H., Imani M., and Moghaddam M. P., (2021). "Convolutional and recurrent neural network based model for short-term load forecasting,", Electric Power Systems Research, vol. 195, 107173, ISSN 0378-7796, https://doi.org/10.1016/j.epsr.2021.107173

Geoffrey Allen P., and Morzuch B. J., (2006) "Twenty-five years of progress, problems, and conflicting evidence in econometric forecasting. What about the next 25 years?", International Journal of Forecasting, vol. 22, Issue 3, pp. 475-492, ISSN 0169-2070, https://doi.org/10.1016/j.ijforecast.2006.03.003 .

Gjika E., Basha L., Ferrja A., and Kamberi A., (2021,A). "Analyzing Seasonality in HPP Energy Production and External Variables", ITISE2021-International Conference on Time Series. Granada, 19th-21th July, 2021. Gran Canaria (SPAIN) , https://itise.ugr.es/index.php (Proceedings) https://www.mdpi.com/2673-4591/5/1/15

Gjika E., Lamberti E. D., and Basha L., (2021,B). "Predicting Energy Production by HPP Using Machine Learning Algorithms with Priority Weights", 41st International Symposium on Forecasting, Online, June,. https://forecasters.org/events/symposium-on-forecasting/

Gjika, E. Basha L., (2021,C) Energy production and consumption relying on climacteric variables (albania case study), "Finance and Accounting towards Sustainable development Goals“ International Conference, Faculty of Economy, University of Tirana 26 November 2021. https://feut.edu.al/lajmerime/1167-international-conference-finance-and-accounting-towards-sustainable-development-goals-november-26-2021-tirana-albania-2

Hadjipanayi M., Koumparou I., Philippou N., Paraskeva V., Phinikarides A. et al., (2016). “Prospects of photovoltaics in southern European, Mediterranean and Middle East regions.”, Renew. Energy 92, pp. 58–74, ,https://doi.org/10.1016/j.renene.2016.01.096

Hewamalage H., Bergmeir Ch., and Bandara K., (2021). “Recurrent Neural Networks for Time Series Forecasting: Current status and future directions”, International Journal of Forecasting, vol. 37, Issue 1, pp.388-427, ISSN 0169-2070, https://doi.org/10.1016/j.ijforecast.2020.06.008

Hou T., Fang R., Tang J., Ge G., Yang D., Liu J., and Zhang W., (2021). “A Novel Short-Term Residential Electric Load Forecasting Method Based on Adaptive Load Aggregation and Deep Learning Algorithms.”, Energies, vol. 14, 7820, https://doi.org/10.3390/en14227820

Hyndman R. J., (2020). “A brief history of forecasting competitions”, International Journal of Forecasting, vol. 36, Issue 1, pp. 7-14, ISSN 0169-2070, https://doi.org/10.1016/j.ijforecast.2019.03.015

Hyndman R. J., and Athanasopoulos G., (2021). Forecasting: principles and practice, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3,

Hyndman R. J., and Khandakar Y., (2008). “Automatic Time Series Forecasting: The forecast Package for R.”, Journal of Statistical Software, vol. 27(3), pp.1–22, https://doi.org/10.18637/jss.v027.i03 .

Hyndman R. J., and Koehler A. B., (2006) .“Another look at measures of forecast accuracy”, International Journal of Forecasting, vol. 22, Issue 4, pp. 679-688, ISSN 0169-070, https://doi.org/10.1016/j.ijforecast.2006.03.001 .

Hyndman R.J., Koehler A. B., Ord, and R. D. Snyder, (2008). Forecasting with Exponential Smoothing, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-71918-2 .

Kalimoldayev M., Drozdenko A., Koplyk I., Marinich T., Abdildayeva A., and Zhukabayeva T., (2020) “Analysis of modern approaches for the prediction of electric energy consumption.”, Open Engineering, vol. 10(1), pp. 350-361. https://doi.org/10.1515/eng-2020-0028

Machado E., Pinto T., Guedes V., and Morais H., (2021) . "Electrical Load Demand Forecasting Using Feed-Forward Neural Networks.", Energies, vol. 14, 7644, https://doi.org/10.3390/en14227644

Makridakis S., and Hibon M., (2000) "The M3-Competition: results, conclusions and implications", International Journal of Forecasting, vol 16, Issue 4, pp. 451-476, ISSN 0169-2070, https://doi.org/10.1016/S0169-2070(00)00057-1 .

Makridakis S., Spiliotis E., and Assimakopoulos V., (2018a). "The M4 Competition: Results, findings, conclusion and way forward", International Journal of Forecasting, vol 34, Issue 4, pp. 802-808, ISSN 0169-2070, https://doi.org/10.1016/j.ijforecast.2018.06.001

Makridakis S., Spiliotis E., and Assimakopoulos V., (2018b) . "Statistical and machine learning forecasting methods: Concerns and ways forward." PLOS ONE, vol. 13, Article e0194889. http://dx.doi.org/10.1371/journal. pone.0194889.

Mandal P., Senjyu T., Urasaki N., and Funabashi T., (2006). "A neural network based several-hour-ahead electric load forecasting using similar days approach", International Journal of Electrical Power & Energy Systems, vol. 28, Issue 6, pp. 367-373, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2005.12.007 .

Newbold P., (1983). "ARIMA model building and the time series analysis approach to forecasting.", J. Forecast. vol. 2, pp. 23–35, https://doi.org/10.1002/for.3980020104

Ouyang Z., Ravier P., and Jabloun M., (2021). "STL Decomposition of Time Series Can Benefit Forecasting Done by Statistical Methods but Not by Machine Learning Ones." Eng. Proc, vol. 5, 42, https://doi.org/10.3390/engproc2021005042

Sánchez-Lorenzo A., Enriquez-Alonso A., Wild M., Trentmann J., Vicente-Serrano S. M. et al., (2017). "Trends in downward surface solar radiation from satellites and ground observations over Europe during 1983–2010.", Remote Sens. Environ. 189, pp. 108–117, https://doi.org/10.1016/J.RSE.2016.11.018

K. Sevdari, A. Islami, E. Haxhiraj and E. Voshtina, "A Data-Driven Assessment of the Electricity Demand-the Case of Albania," 2022 International Conference on Renewable Energies and Smart Technologies (REST), Tirana, Albania, 2022, pp. 1-5, https://doi.org/10.1109/REST54687.2022.10022956 .

Shaub D., (2020). "forecastHybrid: Convenient Functions for Ensemble Time Series Forecasts, Version 5.0.19", Published 28-8-2020, https://cran.r-project.org/web/packages/forecastHybrid/index.html

Smyl S., (2020) “A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting.”, International Journal of Forecasting, vol. 36, pp. 75–8, http://dx.doi.org/10.1016/j.ijforecast.2019.03.017 , M4 Competition

Soukissian Th., Denaxa D., Karathanasi F. , Prospathopoulos A., Sarantakos K. et al., (2017). “Marine renewable energy in the Mediterranean Sea: Status and perspectives.”, Energies vol. 10, 1512, https://www.mdpi.com/1996-1073/10/10/1512

Taylor J. W., and Buizza R., (2003). “Using weather ensemble predictions in electricity demand forecasting”, International Journal of Forecasting, vol. 19, Issue 1, pp. 57-70, ISSN 0169-2070, https://doi.org/10.1016/S0169-2070(01)00123-6

Topolewski L., (2021). “Relationship between Energy Consumption and Economic Growth in European Countries: Evidence from Dynamic Panel Data Analysis.”, Energies, vol. 14, 3565, https://doi.org/10.3390/en14123565

Verwiebe P. A., Seim S., Burges S., Schulz L., and MüllerKirchenbauer J., (2021). “Modeling Energy Demand—A Systematic Literature Review.”, Energies, vol. 14, 7859, https://doi.org/10.3390/en14237859

Xie T., and Ding J., (2020). “Forecasting with Multiple Seasonality”, IEEE International Conference on Big Data (Big Data), pp. 240-245.