# Assessing the Accuracy of Electricity Price Forecasting Models, Before and After, the Impact of Energy Crisis Using Univariate and Multivariate Methods

**Balázs Herczeg**

Doctoral School of Earth Sciences, Faculty of Sciences, University of Pécs
Ifjúság útja 6, H-7624 Pécs, Hungary, hebuabk.pte@tr.pte.hu

**Ágnes Csiszárik-Kocsir**

Keleti Károly Faculty of Business and Management, Óbuda University, Bécsi út 96/b, H-1034 Budapest, Hungary, kocsir.agnes@uni-obuda.hu

**Éva Pintér***

Department of Innovation and Business Incubation, Institute of Entrepreneurship and Innovation, Corvinus University of Budapest, *corresponding author
Fővám tér 8, H-1093 Budapest, Hungary, eva.pinter@uni-corvinus.hu

*Abstract: Forecasting wholesale electricity prices (EPs) is a highly challenging process, especially in an unstable environment. The electricity market is sensitive to crisis events, which can cause significant fluctuations in EPs. Meanwhile, the energy transition and the increasing interconnectedness of the EU's electricity markets add another layer of complexity and further complicate the modeling. This article aims to compare and evaluate several EP forecasting models and methods based on different time horizons, which have unique characteristics. The difference between the periods, reflects the impact of the energy crisis. Therefore, pre- (June 2019 – May 2021) and energy crisis (June 2021 – May 2023) periods were estimated based on best-fit univariate (exponential smoothing and ARIMA) and multivariate (ARIMAX and multiple linear regression) models, built on out-of-sample datasets and the results were assessed primarily with evaluation metrics, such as MAE, MAPE and RMSE. Our empirical results reveal that multivariate methods performed better in estimating monthly average EPs in the EU during pre- and energy crises periods, although the exact models varied between the datasets. Furthermore, regardless of the models utilized, the estimation for the pre-energy crisis period generally resulted in lower error values. Overall, we concluded that different conditions lead to diverse models being more effective.*

*The unprecedented surge in Eps, during 2021-2023, underscores the importance of re-evaluating model performances over time and under different market conditions.*

*Keywords: Electricity price forecast; Energy crisis; Model evaluation; Time series modeling; Statistical models*

# 1 Introduction

Electricity markets globally, and within the European Union (EU) in particular, are complex and dynamic systems influenced by many factors, which range from policy and economic conditions through technological advancements to environmental and geopolitical concerns. Central to these markets are wholesale electricity prices, which play an essential role in determining the electricity cost for residential and non-residential consumers. The ability to accurately forecast these prices is critical, not only for market participants but also for regulators who strive to ensure affordability and market stability.

The wholesale electricity market is basically where electricity is bought between producers and retailers. Prices in this market reflect the balance between supply and demand, and based on the merit order and the underlying marginal pricing system, the wholesale electricity prices are influenced by the fuel costs, the composition of the energy mix used to generate power, power plant outages, electricity demand, weather conditions, and industrial activities among others [1]. Furthermore, the EU's wholesale electricity market is unique due to its high level of interconnectivity and the strong influence of regional and EU-wide policies aimed at strengthening market integration and promoting clean energy [2], which also improves energy security [3] [4]. Although operating nuclear power plants, which can also generate clean energy, concern many EU members due to their disadvantages and risks and therefore it is not considered as a preferred option for moving away from fossil fuels [5], but this is not the case for renewable power plants. Renewables, like wind and solar power, reduce dependence on imported fossil fuels, thereby reducing the risk of depletion of non-renewable natural resources and contribute to reducing emissions considerably, which is essential for combating climate change and complying with international agreements like the Kyoto Protocol (1974), the Paris Climate Accord (2015), or the Renewable Energy Directive(s) (2009, 2018) [6-8]. Accordingly, the energy transition fundamentally reshapes the EU's energy market, requiring improvements in infrastructure, regulation, and market operations and to have a flexible power system for smoother transition [9] [10]. It aims not only at environmental sustainability but also at enhancing energy security, economic resilience, and technological innovation across the EU and in any given region [11], contributing to the broader sustainability of humanity's living space [12].

Wholesale electricity prices are closely related to the energy equity dimension in the concept of energy trilemma, as they have a significant impact on both the accessibility and affordability of electricity [13] [14]. Wholesale electricity prices directly influence the electricity costs for consumers, although, in some regulated markets, this direct impact could not be measured in residential consumer bills due to government price regulations. Nevertheless, higher wholesale prices can increase retail prices, making electricity less affordable for households and businesses. At the same time, it can hinder the expansion of electricity access in undeveloped or rural areas. However, as the factors influencing the dimensions are highly interrelated, the electricity prices can also be considered in the context of energy security and sustainability. Volatile electricity prices can indicate issues with the supply and price of energy used for power generation. Furthermore, the mix of energy sources, hence the share of renewable energy in electricity production, impacts the wholesale electricity prices based on the merit order mechanism [15].

Forecasting wholesale electricity prices is vital for the efficient functioning of energy markets, proper policy-making, operational efficiency in grid management, protecting consumers, and adapting to the evolving energy landscape, among others. Yet, forecasting accurately in the energy market is challenging due to its characteristics and all the influencing factors [16]. Not to mention the impact of a time of distress period, such as the 2021 energy crisis, that brought new challenges and learnings for electricity price forecasting. The interconnectedness of the EU's electricity markets meant the crisis had far-reaching implications across member states, affecting demands and supplies, import dependencies and government policies [17] [18]. The crisis prompted discussions and initiatives to diversify energy sources further, enhance energy storage capacities, and improve market mechanisms to handle such volatile scenarios better. The unprecedented nature of the crisis meant that there was limited historical data on similar events, making it challenging for models to predict the magnitude and duration of the price spikes accurately [19].

In this paper, we attempt to compare and evaluate several forecasting models and methods based on different time series with unique characteristics. Notably, we aim to forecast monthly wholesale electricity prices in the EU for two equally long periods with very different properties. The timeframes of June 2019 – May 2021 and June 2021 – May 2023 were estimated based on best-fit univariate (exponential smoothing and ARIMA) and multivariate (ARIMAX and multiple linear regression) models built on out-of-sample datasets, and the results were assessed with primarily evaluation metrics such as MAE, MAPE, RMSE. The chosen cut-off date relates to the first month impacted by the energy crisis. Consequently, with this approach, we might find evidence of whether assessing electricity prices under different conditions (i.e., pre-energy crisis vs. energy crisis time series) results in other models and methods being more effective.

Concerning this subject, our quantitative analysis shall find answers to following research aims:

1) To find statistical evidence that the forecasting capability of each utilized model before the impact of the energy crisis was better compared to the energy crisis-influenced series.
2) To determine which method (univariate or multivariate) and utilized model can produce a more accurate estimation in a relatively stable or volatile environment.

Our empirical study makes a significant contribution to the existing literature. The novelty of our approach lies in the unique time horizon assessment, the choice of utilizing the coverage of EU average data and conducting both univariate and multivariate approaches in electricity price modeling. Furthermore, we are interested in medium- and long-term trends and a broader view of market behavior and changes over months or years; thus, average monthly electricity price data is utilized, further contributing to the existing literature. Monthly series also smooth out daily volatility. Hence, it is better to observe how the energy crisis impacts prices over the longer term and predict medium- and long-term trends and patterns.

The rest of the research is structured as follows: First, in Section 2, we provide an extensive overview of the current literature on the established electricity price forecasting approaches. Next, Section 3 describes the data used in our empirical study. Finally, Section 4 discusses the results of our analysis and Section 5 concludes the research.

# 2 State of the Art

In the context of the electricity market, many different researchers reviewed and analyzed the trends and volatility of electricity prices in various circumstances. Hence, modeling and forecasting electricity prices have a growing literature in the past decades [20-22]. Among the model types, popular statistical models include multiple linear regression (MLR), autoregressive models (AR, ARMA, ARIMA, SARIMA, autoregressive conditional heteroskedasticity models (ARCH, GARCH), as well as these models augmented by exogenous variables (e.g., ARX, ARMAX, ARIMAX, ARCHX, and GARCHX) and other hybrid solutions. Although, the chosen price forecasting techniques vary based on the properties of the available dataset and the research aims [20] [23].

Weron and Misiorek [24] evaluated a dozen parametric and semiparametric time series models used for day-ahead electricity price forecasting in different markets. They concluded that semiparametric models could lead to better forecasting accuracy in general. In another case [25], various time-series regression modeling approaches were reviewed, and the authors found that only 12% of research papers out of 26 examined papers used multivariate analyses, while the rest utilized univariate analysis, particularly univariate ARCH and GARCH processes. Ziel and

Weron [21] investigated if univariate or multivariate electricity price forecasting models are more accurate. They found that multivariate models do not consistently outperform univariate models, and more accurate forecasts can be achieved with hybrid models. On the other hand, Raviv et al. [26] studied the hourly prices on the Nord Pool electricity market and their multivariate models (VAR, VAR-PCA, and reduced rank regression (RRR) models) based on the full panel of 24 individual hourly prices surpassed the univariate models (dynamic and heterogeneous AR models) of the daily average prices in accuracy. Nevertheless, they also claimed that combining forecasting methods shall grant further improvements. Crespo Cuaresma et al. [27] compared the performance of various AR univariate models in electricity price prediction based on the European Energy Exchange (EEX) hourly time series dataset and primarily found that modeling for each hour of the day, is more accurate than forecasting the whole time-series period.

Besides the AR and ARCH models, researchers also used other models for electricity price forecasting and modeling purposes. Ferreira et al. [28] utilized a multiple linear regression analysis with various explanatory variables to forecast electricity prices in the Iberian electricity market. Ulgen and Poyrazoglu [29] also applied an MLR model on the day-ahead electricity market in Turkey. They found evidence that lagged electricity prices and lagged moving average prices play a key role in the prediction, while they have also used other regressions such as natural gas, oil, and coal prices in the estimation. Saini et al. [30] used a hybrid approach to predict electricity prices, combining MLR and SVM with the final adjustment of the PSO technique, and they found that this hybrid approach shows better accuracy in comparison to other methods. McMenamin and Monforte [31] applied and evaluated several modeling approaches (MLR, exponential smoothing, ARIMA, and ANN models) for forecasting electricity prices, and the results showed that moving from casual to more advanced methods can reduce the forecasting error magnitude.

Many of the studies mentioned above used energy commodity prices as exogenous variables for multivariate modeling purposes. However, the level of renewable energy share in electricity generation also plays a vital role in modeling. The quantile regression model for the German electricity market assessed in Hagfors et al. [32] focused on the impact of wind and solar power on EEX spot prices. This study also found empirical evidence that renewable energy sources have a price-dampening effect on electricity prices. Cevik and Ninomiya [33] used monthly observations for a panel of selected EU members over the period 2014-2021 in conducting a panel quantile regression to determine how the selected variables, such as power generation from renewable energy, electricity load, temperature, and crude oil import price, impact the level and volatility of wholesale electricity prices. According to their results, renewable energy is associated with an average of 0.6% reduction in wholesale electricity prices for each 1%-point increase in renewable share. However, they also found evidence that this association has a non-linear effect, as the higher the proportion of renewables, the higher the impact on electricity prices.

Rathmann [34] analyzed the additional deployment of renewable energy share in the context of the CO2 Emission Trading Scheme (ETS), as the increasing proportion of renewable energy share can not only reduce electricity prices but $CO_2$ emissions as well. Hence, carbon permit prices are also reduced. This study estimated that during the first phase of the EU ETS (2005-2007), the retail electricity prices in Germany lowered by 2.6 EUR/MWh with the addition of renewable units. Gianfreda et al. [35] found empirical evidence based on the Northern Italian zone, which has high renewable penetration, that changes in renewable energy share influence the electricity and fuel prices, pushing gas power plants out of the merit order in the spot market. Paraschiv et al. [36] also found evidence that the increasing penetration of renewable energies reduces EEX electricity spot prices and shifts the merit-order curve. Furthermore, in [37], statistical and machine learning methods were compared across six years and five markets with the integration of RES-E effect on EP. The datasets were divided into training and testing periods for evaluation purposes using MAE, MAPE, and RMSE metrics. Although it was confirmed in Woo et al. [38] with a linear regression model that the increase in wind generation would reduce the day-ahead electricity prices in Texas's electrical grid, the study also highlighted that the price variance might increase as well in parallel. This was also confirmed for Germany, as wind power generation decreases the electricity price level but increases its volatility based on the GARCH model using the January 2006 – January 2012 EEX spot prices dataset [39]. There are other relevant studies in the renewable energy – electricity market context [40-42], which contribute to the need to take exogenous variables into account when creating forecasting models.

Finally, we must recognize the impact of the crisis events that happened in recent years, which fundamentally changed the markets, indicators, and predictions and affected the general research focus regarding our topic. Several authors addressed the pre-COVID-19 and post-COVID-19 differences, regarding energy prices [43-46], but the effect of the energy crisis and warfare also have a growing literature on this matter [47-49].

# 3 Data Used in the Quantitative Analyses

The dependent variable of our study is electricity price (EP). EP datasets were obtained from the EMBER database, which relies on wholesale day-ahead electricity prices for European countries reported by the European Association for the Cooperation of Transmission System Operators (ENTSO-E). Based on the literature reviewed and our understanding of the electricity market, particularly the importance of the merit order mechanism, our selected independent variables for multivariate modeling purposes are the share of renewable energy in gross electricity production (RES-E), net electricity production (NEP), energy commodity prices and the carbon permit price. Electricity data were collected from the International Energy Agency (IEA) database for EU-25[1] countries, while the other regressors, represented by EU benchmark indexes, are commodity (natural gas - TTF, oil - BRENT, and coal - NEWC) and EU ETS (EUA) futures prices, gathered from either Investing.com or Refinitiv Eikon database. Most of the data were available on a daily basis, which we have converted to monthly averages, to have a joint base across the variables and serve our research goals.

As per the time horizons, in public perception the 2021-2023 years were impacted by the global energy crisis we focus on. Figures 1.a and 1.b illustrate this distinction in the development of energy and environmental commodity prices in the EU. The prices surged in late 2021 and accelerated by the onset of the Russia-Ukraine war, which started in February 2022. However, if we take a closer look at the EP trends, monthly average EU EPs significantly stepped out from their average range (previously discussed 30-50 EUR/MWh) for the first time in June 2021. Hence, we define the energy crisis impact on EPs by its starting point in June 2021 and ending in May 2023; as in this month, the average EPs lowered for the first time to the level previously that could be observed prior to June 2021. Hereinafter, this timeframe is called as "energy crisis" period.

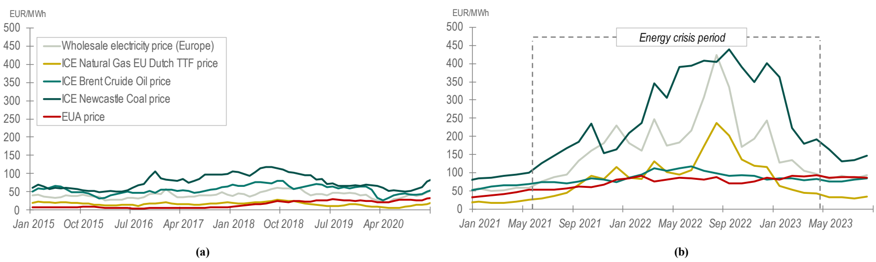



Figure 1

Monthly average energy commodity and EU ETS carbon prices between (**a**) January 2015 – December 2020 and (**b**) January 2021 – August 2023. Source: Created by the authors based on [50-53]

Furthermore, for comparison purposes, we also define a "pre-energy crisis" period with the same length as the energy crisis period to comparably analyze EPs' behavior in a relatively stable period without the influence of the energy crisis. Hence, the pre–energy crisis timeframe covers the June 2019 – May 2021.

A common mistake is that forecasting accuracy is measured based on how well a model fits the historical data [54]. However, it should be assessed on data not used

[1] Instead of relying on the current EU-27, Malta and Cyprus were not included in the utilized datasets as electricity data was limited for bidding zones that have yet to introduce a power exchange, which is the case for these two countries. However, due to the marginal impact of the elimination, we refer to EU-25 as the whole EU.

during model building but available to be compared. For this purpose, the observed time series shall be divided into two sets in approximately 80-20% proportion. As shown in Figure 2, the larger parts are the training periods, while the evaluations are based on the testing periods, which are basically the aforementioned pre- and energy crisis periods. In our case, for the first model, the period from January 2015 to May 2019 served as a training period, while for the second model, this training period was extended to May 2021. The 24-month test periods follow each of these periods.

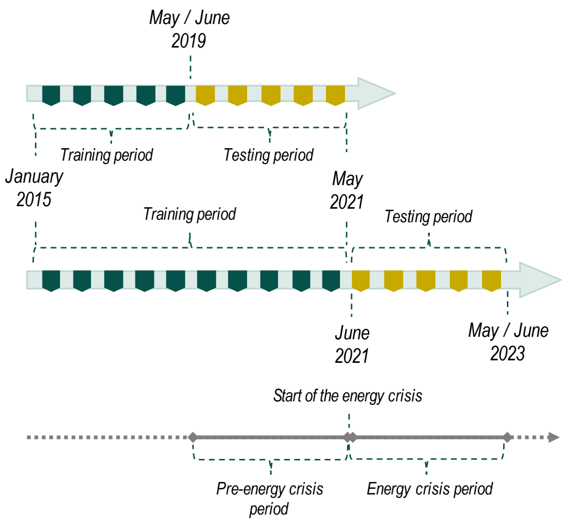



Figure 2

The concept and the exact time frames in our models. Source: Created by the authors

Additionally, Figure 3 presents the monthly development of energy transition in the EU during the same time horizon. Monthly gross electricity production (GEP) indicates a seasonal pattern with lower production during the middle of each year, peaking around year-end. This is attributable to several factors, such as seasonal demand variations (e.g., heating needs and shorter daylight during winter) or hydroelectric power generation (e.g., droughts in summertime and snowmelts during spring).

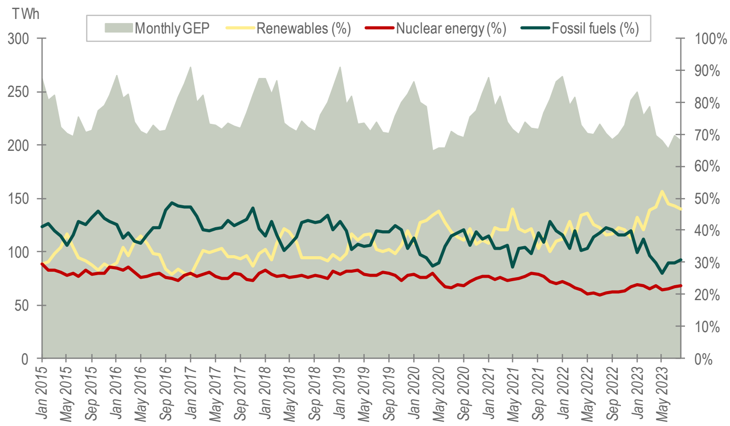



Figure 3

Monthly gross electricity production (GEP) by type of source in EU-25 (Jan 2015 - Aug 2023).

Source: Created by the authors based on [55, 56]

Concerning this, the demand for fossil fuels in electricity production also varies accordingly. More demand requires more production on average, which could not be achieved solely by renewable sources at the moment, especially not in winter, when solar and hydroelectric power generation are generally lower. As a result, expanding and narrowing gaps can be seen between RES-E and the share of fossil fuels in Figure 3, with an inflection point during summer. However, since the time of distress starting with the COVID-19 pandemic, renewable energies have become the leading source of electricity production in the average EU. In the last year, the gap increased more than ever during the middle of the year in favor of RES-E.

Figures 1 and 3 support our selection of variables in forecasting EPs. The association of energy price variables can be seen, although the extent varies based on the observed time series, in other words, the economic circumstances. Furthermore, the importance of RES-E in the merit order and, hence, in setting the clearing price for EPs also contributes to the model setup of forecasting EPs.

# 4 Results and Discussion

The IBM SPSS version 27 program, with its R extension and XLSTAT data analysis tool, was used to conduct the analyses. In this section, after the preliminary data analysis, the univariate models are first assessed, and then we move on to the multivariate methods. The comparison and evaluations are presented in the discussion following the model building.

The summary of the descriptive statistics in Table 1 contains information on the training and test period and their combined dataset, as defined in Figure 2. Our focus is primarily on the training periods, as the model-building process is based on these and on the test periods, which are going to be evaluated and compared. The impact of the energy crisis on EPs can easily be read by looking at the mean, maximum and standard deviation figures between the two test periods. The training period used for modeling the pre-energy crisis period has similar characteristics to the test period, although the moderate right-skew and the normal distribution moved to a left-skew and platykurtic distribution in the latter one. As per the other model, the training period – the aggregated time series of the two just discussed periods – significantly differs from the energy crisis period. This indicates a heavily skewed distribution to the right, while a kurtosis of 1.06 suggests a leptokurtic distribution, implying a higher likelihood of extreme values in EPs. Furthermore, the Jarque-Bera test (JB test) gives further information regarding normality; these indicate that all examined series are likely normally distributed except for the energy crisis period, which significantly deviates from a normal distribution on a 5% significance level.

Table 1

Statistical characteristics of EPs under different time horizons. Source: Created by the authors

| | Modeling the pre-energy crisis period | | | Modeling the energy crisis period | | |
|---|---|---|---|---|---|---|
| **Dataset** | **Whole period** | **Training period** | **Test period** | **Whole period** | **Training period** | **Test period** |
| Observations | 77 | 53 | 24 | 101 | 77 | 24 |
| Mean | 41.893 | 41.818 | 42.058 | 75.081 | 41.893 | 181.559 |
| Minimum | 21.551 | 27.137 | 21.551 | 21.551 | 21.551 | 77.051 |
| Maximum | 61.439 | 61.439 | 59.311 | 425.198 | 61.439 | 425.198 |
| Std. dev. | 9.197 | 8.706 | 10.395 | 72.933 | 9.197 | 85.612 |
| Skewness (Pearson) | 0.269 | 0.709 | -0.343 | 2.464 | 0.269 | 1.114 |
| Kurtosis (Pearson) | -0.148 | 0.052 | -0.481 | 6.352 | -0.148 | 1.060 |
| Jarque-Bera test | 0.998 | 4.451 | 0.700 | 271.988** | 0.998 | 6.085* |
| Ljung-Box test (df=6) | 123.905** | 99.521** | 37.484** | 364.061** | 123.905* | 16.632* |
| ADF-test at level | -2.882 | -3.499* | | -2.096 | -2.882 | |
| ADF-test at I(1) | -4.245** | -3.532* | | -3.241 | -4.245** | |
| PP-test at level | 0.000 | -0.296 | | -1.247 | 0.000 | |
| PP-test at I(1) | -7.782** | -7.071** | | -9.600** | -7.782** | |
| KPSS-test at level | 0.144 | 0.157* | | 0.378** | 0.144 | |
| KPSS-test at I(1) | 0.051 | 0.049 | | 0.092 | 0.051 | |

*** Significant at 1%; * significant at 5%.*

The white noise tests of Ljung-Box (LB tests), which check for autocorrelation in a time series at different lags, jointly imply that there are significant autocorrelations in at least one of the first six flags in each model, resulting in the rejection of the null hypothesis of the randomness of the data. To understand the specifics of the autocorrelation in our data, ACF and PACF plots are visualized on the original series as well as on the log-transformed and log-differentiated datasets. Transformations were necessary as many statistical modeling techniques and analyses assume stationarity, i.e., the time series has a constant mean, variance, and autocorrelation structure; otherwise, the series must be transformed to achieve stationarity. Our original ACF plots of the training sets showed a slow decay to zero, suggesting non-stationarity. As a result, unit root (ADF and PP) and stationarity (KPSS) tests were also performed as Molnár and Csiszárik-Kocsir [57] highlight that ADF test solely could not be powerful enough to determine stationarity. As per the training set of the pre-energy crisis model, the mixed result indicates that the time series may be stationary in terms of not having a unit root (ADF test) in its level, but it may have a deterministic trend or other form of non-stationarity that the KPSS test is picking up. The PP test's non-significance could be due to its robustness to autocorrelation and heteroskedasticity, affecting its sensitivity. However, the results after log-differentiation are more consistent and suggest that the transformations have successfully stabilized the mean and variance over time, leading to a stationary time series. As per the energy crisis model, the non-significant result from KPSS at the level does not necessarily contradict the ADF and PP tests because KPSS tests for stationarity around a trend, and a non-significant result here suggests that the series does not have a deterministic trend. Nevertheless, log-differencing also made the series stationary. We note that, from our chosen methods, ES and MLR require stationary data, while the advanced forms of ES and ARIMA can handle non-stationary data by modeling the underlying components (level, trend, and seasonality) explicitly. It is also possible that the natural log transformation alone makes the series sufficiently stationary for modeling. The statistical properties of the data post-log transformation may not exhibit strong trends or unit roots, leading, e.g., the ARIMA model selection process to conclude that differencing is not needed.

## 4.1 Applying the Selected Methods on the Observed Periods

The first model estimated is an ES model. After our experimentation with various forms, a simple seasonal ES model based on the natural log-transformed series is determined as the best-fit model based on the normalized BIC, R-squared, and accuracy criteria. The model statistics are summarized below (Table 2).

Table 2
Model statistics of the best-fit exponential smoothing models. Source: Created by the authors

| | **Pre-energy crisis model** | | **Energy crisis model** | |
|---|---|---|---|---|
| **Dataset** | **Training period** | **Test period** | **Training period** | **Test period** |
| Observations | 53 | 24 | 77 | 24 |
| Alpha (Level) | 0.700 | | 0.999 | |
| Beta (Season) | 6.24E-07 | | 4.68E-04 | |
| MAE | 2.974 | 11.767 | 3.154 | 109.201 |
| MAPE | 7.068 | 33.873 | 7.761 | 52.346 |
| RMSE | 3.783 | 13.475 | 4.035 | 136.729 |

| | | | | |
|---|---|---|---|---|
| R-squared | 0.808 | | 0.805 | |
| Normalized BIC | 2.849 | | 2.929 | |
| Ljung-Box Q (18) | 44.163** | | 47.016** | |

*** Significant at 1%; * significant at 5%.*

*Best-fit models: Simple seasonal ES based on natural log-transformed datasets*

As per the pre-energy crisis model, the training period results showed higher accuracy metrics compared to what can be seen based on the test period. However, considering the pre-energy crisis training and test period together as a basis for the energy crisis model, the subsequent test period results were much worse, as expected. MAE and RMSE are approximately ten times higher than in the first model. As a result, the model estimating the pre-energy crisis period was more efficient than the one predicting the energy crisis series.

The following method utilized from our toolkit is ARIMA, where the final model is (0,0,3)(0,0,0) for the pre-energy crisis period and (0,0,3)(1,1,0) for the energy crisis period, both based on their log-transformed training sets. The coefficients and the summary statistics are presented in Table 3.

Table 3

Model statistics of the best-fit ARIMA models. Source: Created by the authors

| | **Pre-energy crisis model** | | **Energy crisis model** | |
|---|---|---|---|---|
| **Dataset** | **Training period** | **Test period** | **Training period** | **Test period** |
| Observations | 53 | 24 | 65 | 24 |
| Coefficient estimates: | | | | |
| Constant | 3.714** | | - | |
| MA(1) | 0.1871** | | -1.059** | |
| MA(2) | 0.494** | | -0.716** | |
| MA(3) | 0.411** | | -0.641** | |
| SAR(1) | - | | -0.663** | |
| SDIFF | - | | 1 | |
| MAE | 3.891 | 8.150 | 4.438 | 136.938 |
| MAPE | 9.215 | 22.885 | 10.862 | 68.973 |
| RMSE | 4.915 | 10.142 | 5.574 | 162.113 |
| R-squared | 0.675 | | 0.673 | |
| Normalized BIC | 3.563 | | 3.757 | |
| Ljung-Box Q (18) | 20.937 | | 42.627** | |

*** Significant at 1%; * significant at 5%.*

*Best-fit models: ARIMA(0,0,3)(0,0,0) and ARIMA(0,0,3)(1,1,0) based on natural log-transformed datasets*

The first model is a non-seasonal ARIMA with three MA terms, using the error terms from the previous three periods to estimate the actual values of the series. The model suggests that short-term, lagged shock effects are influential in predicting log-transformed energy prices without the need to account for trends or seasonal patterns. Applying this setup to the test period resulted in a slightly inferior outcome. However, the change in the ARIMA model from pre-energy crisis to energy crisis indicates a shift in the underlying patterns of the log-transformed energy prices. Specifically, it shows that while the short-term effects captured by the MA terms remain relevant, the energy crisis period requires accounting for seasonality and its associated trends or patterns. This shall reflect changes in the energy market dynamics due to the crisis, amplifying seasonal effects that need to be modeled. The seasonal component (1,1,0) indicates that there is one seasonal AR term (P=1) and one order of seasonal differencing (D=1). The latter implies that the log-transformed data exhibits seasonal non-stationarity, which is addressed by taking the difference of the data at the seasonal period. Nonetheless, its forecasting ability on the test period shows high inaccuracies with respect to month-ahead accuracy, with a MAPE of 69% and MAE of 137 EUR/MWh, which is even higher than what was experienced with the simple seasonal ES (52% and 109 EUR/MWh, respectively).

Moving toward the multivariate approach, the ARIMAX models are first estimated (Table 4). Based on the coefficients of model parameters, only RES-E, TTF, and EUA are found to significantly contribute to the pre-energy crisis training period's model goodness and prediction ability. We note that the "I" component in ARIMAX does not automatically apply to the exogenous variables, but these explanatory variables must also be stationary. Hence, after applying and evaluating log-transformation and log-differentiation on all the external variables (as the original data were non-stationary), we found that including log-transformed variables is sufficient for the ARIMAX model to achieve the highest model fit. Moreover, RES-E and EUA are included with a lag of zero periods, which is zero and five for TTF.

Table 4

Model statistics of the best-fit ARIMAX models. Source: Created by the authors

| | **Pre-energy crisis model** | | **Energy crisis model** | |
|---|---|---|---|---|
| **Dataset** | **Training period** | **Test period** | **Training period** | **Test period** |
| Observations | 48 | 24 | 63 | 24 |
| Coefficient estimates: | | | | |
| Constant | 2.156** | | 1.517** | |
| MA(1) | -0.954** | | - | |
| SAR(1) | - | | -0.672** | |
| SDIFF | - | | 1 | |
| MAE | 1.999 | 11.143 | 1.333 | 33.375 |
| MAPE | 4.984 | 28.394 | 3.269 | 14.983 |
| RMSE | 2.399 | 12.893 | 1.760 | 46.338 |
| R-squared | 0.927 | | 0.967 | |
| Normalized BIC | 2.368 | | 1.877 | |
| Ljung-Box Q (18) | 31.514* | | 63.040** | |

*** Significant at 1%; * significant at 5%.*

*Best-fit models: ARIMAX(0,0,1)(0,0,0) based on natural log-transformed datasets, and ARIMAX(0,0,0)(1,1,0).*

As per the energy crisis model parameters, the circle extended with NEP and NEWC; hence, only BRENT did not contribute significantly to either model. In this latter series, all external variables are included with seasonal differencing of one, as well as TTF and EUA with lags of zero, RES-E and NEP with lags of zero and one, and lastly, NEWC with lags of zero and a delay of two periods, implying its immediate past value (adjusted for seasonality) and the values from two periods ago are considered in the energy crisis model.

Even with the inclusion of certain exogenous variables, the base model configurations retained the seasonal component in both cases. However, the MA terms almost completely diminished, and log-transformed EPs are used only in modeling the pre-energy crisis period. Nevertheless, it can be concluded that model extensions resulted in better accuracy and model fit regarding the training set. Also, the forecast for the energy crisis period is less inaccurate – in fact, the MAPE (%) is even better (15%) – than it was experienced by previous methods.

The last method in our framework is the utilization of an MLR model. The output summary can be seen in Table 5.

Table 5
Model statistics of the best-fit CO estimated MLR models. Source: Created by the authors

| | **Pre-energy crisis model** | | **Energy crisis model** | |
|---|---|---|---|---|
| **Dataset** | **Training period** | **Test period** | **Training period** | **Test period** |
| Observations | 52 | 24 | 76 | 24 |
| Coefficient estimates: | | | | |
| Constant | 0.002 | | -0.001 | |
| RES-E | -0.929** | | -0.913** | |
| TTF | 0.334** | | 0.254** | |
| EUA | 0.250** | | 0.43** | |
| NEP | - | | 0.254* | |
| MAE | 1.886 | 2.683 | 1.809 | 73.411 |
| MAPE | 4.588 | 7.394 | 4.476 | 33.784 |
| RMSE | 2.548 | 3.296 | 2.464 | 98.038 |
| Adj. R-squared | 0.732 | | 0.741 | |
| F-stat | 49.500** | | 56.75** | |
| $VIF_{max}$ | 1.258 | | 1.127 | |
| DW stat | 2.151 | | 2.034 | |

*** Significant at 1%; * significant at 5%.*

*Best-fit models are based on natural log-differenced datasets.*

We know that various factors influence EPs, such as demand and supply dynamics, fuel costs, weather conditions, and more. If most of these factors are captured by the variables in the MLR model, and if the time series nature of the examined data is not strongly autoregressive, then lagged variables might not need to be included in the optimal model. The best-fit model selection confirmed this theory, as only exogenous variables are included.

Both of the modeled training sets are based on natural log-differenced datasets due to stationarity concerns, and Cochrane–Orcutt estimation (CO) is used to correct for autocorrelation in the residuals of the regression models, implied by the initial DW statistics (2.585 and 2.207, respectively). The MLR assumptions were also verified by either analytical or graphical way.

As per the pre-energy crisis model, the best-fit model utilized RES-E, TTF, and EUA, while the energy crisis model extended the list with the NEP regressor, out of which RES-E provides the highest contribution to the model. According to the first model, a 1% increase in RES-E from one period to the next is ceteris paribus, associated with an average of 0.91% decrease in EP for the same period. This supports the importance of renewable energy sources in the merit order and their role in balancing the energy trilemma. Regarding the first model's forecasting accuracy, the training and test sets evaluations are substantially better than the previous models. This MLR specification can accurately model EPs. However, it cannot effectively estimate the energy crisis period series, as all the evaluation criteria are significantly higher than the former model.

## 4.2 Comparisons

Table 6 contains the MAE, MAPE, and RMSE metrics of each best-fit method reviewed previously. After evaluating all the applied methods, we can draw several conclusions based on our methodology, the examined timeframes, and results from the evaluation criteria.

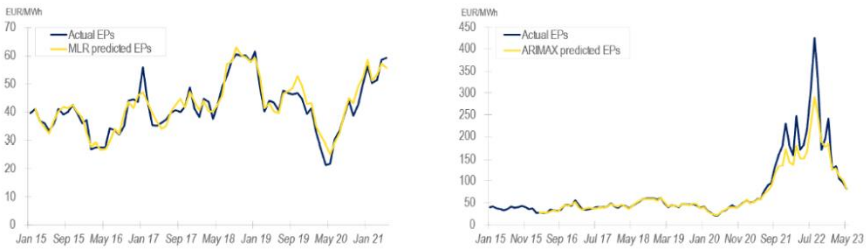
Table 6
Comparison of ES, ARIMA, ARIMAX, and MLR in forecast performance.
Source: Created by the authors

| Test periods | Method | Model | MAE | MAPE | RMSE |
|---|---|---|---|---|---|
| Pre-energy crisis model | Univariate | ES | 11.767 | 33.873 | 13.475 |
| | Univariate | ARIMA | 8.150 | 22.885 | 10.142 |
| | Multivariate | ARIMAX | 11.143 | 28.394 | 12.893 |
| | **Multivariate** | **MLR** | **2.683** | **7.394** | **3.296** |
| Energy crisis model | Univariate | ES | 109.201 | 52.346 | 136.729 |
| | Univariate | ARIMA | 136.938 | 68.973 | 162.113 |
| | **Multivariate** | **ARIMAX** | **33.375** | **14.983** | **46.338** |
| | Multivariate | MLR | 73.411 | 33.784 | 98.038 |

In relation to the pre-energy crisis models, the multivariate MLR model has the best performance across all three-error metrics compared to the univariate models (ES and ARIMA) and the multivariate ARIMAX. The significantly lower values suggest that the use of external regressors has led to a more accurate model for this period. The second in the ranking is the ARIMA, implying that for the univariate models, the ARIMA model is better suited for capturing the time series properties of monthly EPs during the given time horizon. On the other hand, exponential smoothing has the highest error metrics among the pre-energy crisis models, indicating it is the least accurate for this period.

Considering the energy crisis model, the multivariate ARIMAX model shows a remarkable improvement over the other models, as the MAE, MAPE, and RMSE are the significantly lowest of all. This suggests that including external factors alongside ARIMA modeling is particularly effective in accounting for the structural changes in electricity prices during the energy crisis period. In terms of forecasting accuracy, ARIMAX is followed by MLR. It's performance drops compared to the pre-energy crisis period, but it still performs better than using univariate approaches for modeling this time horizon. Both ES and ARIMA models have substantially higher error metrics during the second period, suggesting that they struggle to capture the complexities of monthly EPs in the changed environment.

Overall, regarding the first period and the most accurate MLR model, the actual EPs were EUR 2.683 per megawatt-hour different from the estimated prices or were off by 7.4% on average. These have increased to EUR 33.375 per MWh and 15.0% on average, respectively, while assessing the second period with the most precise ARIMAX model. The model fits are illustrated in Figure 4a and 4b.




(a) (b)

Figure 4

Visualizing the best-fit models' accuracy for (**a**) the pre-energy crisis period and (**b**) the energy crisis period. Source: Created by the authors

## Conclusions

The significance and novelty of our approach is in the time-horizon assessment, which compares "pre-energy crisis" and "energy crisis" periods, based on models built on the immediately preceding time series. In this way, the best-fit models optimized on the training periods could be used adequately to compare and assess 24-month periods with and without the influence of the energy crisis. Furthermore, the choice of utilizing monthly average EPs instead of daily or hourly prices, the coverage of EU average data, and conducting both univariate and multivariate approaches in EP modeling further contribute to the existing literature.

Our results suggest that multivariate models generally led to better forecasting accuracy for electricity prices, particularly in the volatile conditions following the energy crisis. This demonstrates the value of including external regressors in forecasting models to capture more complex dynamics in the data, highlights the influence of external economic and energy variables on the EPs and the interplay of multiple influencing factors. This may also reflect the increased importance of various market factors following the energy crisis. We also highlight the positive role of RES-E in the evolution of EPs, as it significantly contributes to the decrease of EPs in general, regardless of the circumstances. Furthermore, irrespective of the utilized models, the estimation for the pre-energy crisis period resulted in generally lower error values. The energy crisis has introduced more complexity into the electricity price series, making the forecasting task more challenging, as indicated by the overall increase in error metrics for the energy crisis period.

Although our analyses had certain limitations, deriving from the model and parameter selection, further improvements might be achieved by including other terms and specifications into the current framework. Capturing seasonality in the time series with adding seasonal adjustments or using more granular (daily, hourly) time series might lead to more unbiased results.

Furthermore, by improving the chosen series models with e.g., combining ARIMA and GARCH, one can model both the mean (price levels) and variance (volatility) of a series effectively as GARCH focuses on the changing uncertainty or risk around EPs. Besides, electricity markets can be influenced by many factors which may not have a simple linear relationship with EPs or such regressors are treated as nominal values. Non-linear models as well as using dummy regressors can address sudden structural breaks in the data revealing hidden patterns, thus they can capture more complex associations that linear models, such as MLR or ARIMA, might miss.

As a result, sophisticated methods like machine learning and neural networks, which accommodate non-linearity, effectively detect patterns and filter out noise, handle high dimensional datasets, and have high customizability, could be particularly effective for EP forecasting. These options serve as future research topics for us, through the expansion of our methodological toolbox.

## Acknowledgement

This work was supported by the University of Pécs, Doctoral School of Earth Sciences.

## References

[1] European Commission: *Energy prices and costs in Europe*. Brussels, 2016

[2] International Energy Agency: *European Union 2020 Energy Policy Review*. Paris, 2020

[3] Leal-Arcas, R. et al.: *Energy Security, Trade and the EU*. Edward Elgar Publishing, 2016

[4] Zachmann, G. et al.: *The design of the European electricity market. Current proposals and ways ahead*. Luxembourg, 2023

[5] Deutsch, N. et al.: Nuclear power - Additions to wholesale electricity prices and margin of safety. *WSEAS Transactions on Business and Economics*, 15, 2018, pp. 197–212

[6] Moreno, B. et al.: The electricity prices in the European Union. The role of renewable energies and regulatory electric market reforms. *Energy*, 48 (1), 2012, pp. 307–313

[7] Brosset, E., Maljean-Dubois, S.: *The Paris Agreement, EU Climate Law and the Energy Union*. In: Research Handbook on EU Environmental Law. Cheltenham: Edward Elgar Publishing, 2020, pp. 412–427

[8] Erdei-Gally, S., Valisoylu, V.: Natural Resources: Can the Oil and Gas Industry Continue to Support Further Economic Development? *Acta Polytechnica Hungarica*, 20 (10), 2023, pp. 33–51

[9] Carlini, E.M. et al.: EU transition in power sector: How RES affects the design and operations of transmission power systems. *Electric Power Systems Research*, 169 (April 2018), 2019, pp. 74–91

[10] Koltsaklis, N., Knápek, J.: The Role of Flexibility Resources in the Energy Transition. *Acta Polytechnica Hungarica*, 20 (11), 2023, pp. 137–158

[11] IRENA: *World Energy Transitions Outlook 2022: 1.5°C Pathway*. Abu Dhabi: International Renewable Energy Agency, 2022

[12] Csiszárik-Kocsir, Á. et al.: *A fenntarthatóság értelmezésének újragondolása*. In: Vállalkozásfejlesztés a XXI. században 2022/2 kötet - Globális reakciók, lokális akciók a gazdaság rezilienciájának erősítéséhez. Budapest: Óbudai Egyetem, Keleti Károly Gazdasági Kar, 2022, pp. 336–349

[13] World Energy Council: *World Energy Trilemma Index 2022*. London, 2022

[14] Tol, R.: *Navigating the Energy Trilemma during Geopolitical and Environmental Crises*. 1372. Tokyo, 2023

[15] Hafner, M., Luciani, G.: *The Palgrave Handbook of International Energy Economics*. Cham: Palgrave Macmillan, 2022

[16] Hu, L. et al.: A review of short-term electricity price forecasting techniques in deregulated electricity markets. *Proceedings of the Universities Power Engineering Conference*, 2009

[17] European Union Institute for Security Studies, Popkostova, Y.: *Europe's energy crisis conundrum – Origins, impacts and way forward*. 2022

[18] Bhagat, P.R. et al.: Trend Assessment and Comparison of Primary Renewable Energy Consumption in the European Union and India: Possibilities after Covid-19 and War. *Acta Polytechnica Hungarica*, 20 (10), 2023, pp. 253–268

[19] Zhong, H. et al.: Implications of COVID-19 for the electricity industry: A comprehensive review. *CSEE Journal of Power and Energy Systems*, 6 (3), 2020, pp. 489–495

[20] Weron, Rafał.: *Modeling and forecasting electricity loads and prices: a statistical approach*. Chichester: John Wiley & Sons, 2006

[21] Ziel, F., Weron, R.: Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, 70, 2018, pp. 396–420

[22] Weron, R.: Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30 (4), 2014, pp. 1030–1081

[23] Zareipour, H.: *Short-Term Electricity Market Prices: A Review of Characteristics and Forecasting Methods*. In: Handbook of Networks in Power Systems I. Energy Systems (Editors: A. Sorokin et al.). Berlin: Springer Berlin Heidelberg, 2012, pp. 89–121

[24] Weron, R., Misiorek, A.: Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24 (4), 2008, pp. 744–763

[25] Higgs, H., Worthington, A.C.: Modelling Spot Prices in Deregulated Wholesale Electricity Markets: A Selected Empirical Review. *SSRN Electronic Journal*, 17 (1–2), 2008, pp. 1–25

[26] Raviv, E. et al.: Forecasting day-ahead electricity prices: Utilizing hourly prices. *Energy Economics*, 50, 2015, pp. 227–239

[27] Crespo Cuaresma, J. et al.: Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, 77 (1), 2004, pp. 87–106

[28] Ferreira, Â.P. et al.: A linear regression pattern for electricity price forecasting in the Iberian electricity market. *Revista Facultad de Ingeniería Universidad de Antioquia*, (93), 2019, pp. 117–127

[29] Ulgen, T., Poyrazoglu, G.: *Predictor Analysis for Electricity Price Forecasting by Multiple Linear Regression*. In: 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM). IEEE, 2020, pp. 618–622

[30] Saini, D. et al.: *Electricity price forecasting by linear regression and SVM*. In: 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE). IEEE, 2016, pp. 1–7

[31] McMenamin, J.S., Monforte, F.A.: *Statistical Approaches to Electricity Price Forecasting*. In: Faruqui, A., Eakin, K. (eds) Pricing in Competitive Electricity Markets. Boston, MA: Springer US, 2000, pp. 249–263

[32] Hagfors, L.I. et al.: Using quantile regression to analyze the effect of renewables on EEX price formation. *Renewable Energy and Environmental Sustainability*, 1, 2016, pp. 32

[33] Cevik, S., Ninomiya, K.: Chasing the Sun and Catching the Wind: Energy Transition and Electricity Prices in Europe. *IMF Working Papers*, 2022 (220), 2022

[34] Rathmann, M.: Do support systems for RES-E reduce EU-ETS-driven electricity prices? *Energy Policy*, 35 (1), 2007, pp. 342–349

[35] Gianfreda, A. et al.: The RES-Induced Switching Effect Across Fossil Fuels: An Analysis of Day-Ahead and Balancing Prices. *The Energy Journal*, 40, 2019, pp. 1–22

[36] Paraschiv, F. et al.: The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy*, 73, 2014, pp. 196–210

[37] Lago, J. et al.: Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, 293(116983), 2021

[38] Woo, C.K. et al.: The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, 39 (7), 2011, pp. 3939–3944

[39] Ketterer, J.C.: The impact of wind power generation on the electricity price in Germany. *Energy Economics*, 44, 2014, pp. 270–280

[40] Nicolosi, M., Fürsch, M.: The Impact of an increasing share of RES-E on the Conventional Power Market — The Example of Germany. *Zeitschrift für Energiewirtschaft*, 33 (3), 2009, pp. 246–254

[41] Sáenz de Miera, G. et al.: Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. *Energy Policy*, 36 (9), 2008, pp. 3345–3359

[42] Sensfuß, F. et al.: The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. *Energy Policy*, 36 (8), 2008, pp. 3086–3094

[43] Nyga-Łukaszewska, H., Aruga, K.: Energy Prices and COVID-Immunity: The Case of Crude Oil and Natural Gas Prices in the US and Japan. *Energies*, 13 (23), 2020

[44] Amri Amamou, S., Aguir Bargaoui, S.: Energy markets respond to Covid-19 pandemic. *Resources Policy*, 76(102551), 2022

[45] Alam, M.S. et al.: Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools. *Resources Policy*, 81(103342), 2023

[46] Khan, K. et al.: Examining the behaviour of energy prices to COVID-19 uncertainty: A quantile on quantile approach. *Energy*, 239(122430), 2022

[47] Poggi, A. et al.: Electricity Price Forecasting via Statistical and Deep Learning Approaches: The German Case. *AppliedMath*, 3 (2), 2023, pp. 316–342

[48] Bâra, A. et al.: Forecasting the Spot Market Electricity Price with a Long Short-Term Memory Model Architecture in a Disruptive Economic and Geopolitical Context. *International Journal of Computational Intelligence Systems*, 16 (1), 2023, pp. 1–22

[49] Zakeri, B. et al.: The role of natural gas in setting electricity prices in Europe. *Energy Reports*, 10, 2023, pp. 2778–2792

[50] International Energy Agency: *Monthly Electricity Statistics*. Paris. [2023.10.18]

[51] EMBER: *European wholesale electricity price data*. London. [2023.10.18]

[52] Investing.com: *Investing.com*. Madrid. [2023.10.18]

[53] Refinitiv Eikon: *Eikon Energy Commodities*. London. [2023.10.18]

[54] Hyndman, R.J., Athanasopoulos, G.: *Forecasting. Principles and Practice*. Melbourne, Australia: OTexts, 2018

[55] Eurostat: *Gross production of electricity and derived heat from non-combustible fuels by type of plant and operator*. Luxembourg. [2023.10.24]

[56] Eurostat: *Gross production of electricity and derived heat from combustible fuels by type of plant and operator*. Luxembourg. [2023.10.24]

[57] Molnár, A., Csiszárik-Kocsir, Á.: Forecasting Economic Growth with V4 Countries’ Composite Stock Market Indexes-a Granger Causality Test. *Acta Polytechnica Hungarica*, 20 (3), 2023, pp. 135–154