Indonesian Journal of Contemporary Multidisciplinary Research (MODERN)
Vol.2, No.6, 2023: 1309-1320

Modelling Primary Energy by Long Memory Time Series

Gumgum Darmawan[1]*, Budhi Handoko[2]
Universitas Padjadjaran
Corresponding Author: Gumgum Darmawan gumgum@unpad.ac.id

ARTICLEINFO

Keywords: Long Memory, Primary Energy, ARFIMA, MAPE

Received : 4 October
Revised : 15 October
Accepted: 30 November

ABSTRACT

This research employs long memory modeling techniques to analyze and forecast global energy data spanning from 1965 to 2022. Focusing on the ARFIMA (Autoregressive Fractionally Integrated Moving Average) model, the study demonstrates its efficacy in predicting energy consumption trends. The evaluation of forecasting results for the subsequent four years reveals a remarkable Mean Absolute Percentage Error (MAPE) below 5%. This outcome underscores the effectiveness of incorporating long memory components in energy modeling, offering a robust approach for accurate and reliable predictions. The findings contribute to the advancement of energy forecasting methodologies, providing valuable insights for policymakers, energy analysts, and researchers in the pursuit of sustainable and informed energy planning

DOI: https://doi.org/10.55927/modern.v2i6.6970
E-ISSN: 2985-6108
https://journal.formosapublisher.org/index.php/modern

INTRODUCTION

Accurate energy demand modelling is crucial for informing and framing energy policy decisions, ensuring energy availability and quality, and improving energy access across countries, thereby alleviating energy poverty, particularly in developing nations (Berbesi & Pritchard, 2023). The use of fundamental energy resources, such as fossil fuels, renewable sources, and nuclear energy, has a substantial influence on worldwide energy security, environmental well-being, and economic steadiness (Cao et al., 2019). In confronting the issues of climate change and working towards a shift to a more sustainable energy future, precise prediction and modeling of primary energy consumption grow progressively vital(Zaharia et al., 2019). This research paper focuses on employing long memory time series modeling techniques to enhance the precision and reliability of forecasting primary energy consumption trends. Long memory time series analysis is particularly relevant in capturing the persistent and autocorrelated nature of energy consumption data, offering insights that traditional short memory models may overlook (Li et al., 2020; Monge & Gil-Alana, 2019).

Long memory time series models, such as fractional integration and autoregressive fractionally integrated moving average (ARFIMA), are designed to capture the memory effects present in economic and energy-related time series data (Adekoya, 2020; Balagula, 2020) . By incorporating these long memory components, the proposed modeling approach aims to provide a more accurate representation of the primary energy consumption (Shalalfeh et al., 2021) . This research seeks to contribute to the existing literature by exploring the applicability and effectiveness of long memory time series modeling in forecasting primary energy consumption.

The primary objectives of this research paper are to investigate the presence of long memory in primary energy consumption time series data, develop and validate long memory time series models for forecasting primary energy consumption. Through a comprehensive analysis of historical primary energy consumption data, this study aims to identify patterns, trends, and dependencies that extend over longer time horizons. The research will employ rigorous statistical methods to evaluate the accuracy and reliability of long memory time series models, offering insights into their potential as tools for improved energy forecasting.

This research contributes to the broader field of energy modeling by bridging the gap between traditional short memory time series analysis and the long memory characteristics inherent in primary energy consumption data. By advancing the methodology for modeling primary energy consumption with long memory time series techniques, this research aims to provide a valuable framework for developing more accurate and robust energy forecasts, ultimately supporting sustainable energy planning and management.

This research paper is organized as follows. Section 2 provides a comprehensive review of existing literature on primary energy consumption modeling, emphasizing the shortcomings of short memory models and the potential advantages of long memory time series techniques. Section 3 outlines the data sources and methodology employed in this study, detailing the steps taken to preprocess the data and implement long memory time series models. Section 4 presents the results of the analysis, including the identification of long memory components in primary energy consumption data and the performance evaluation of the proposed models. Finally, Section 5 concludes the paper with a summary of key findings, implications for future research, and the potential impact of incorporating long memory time series modeling in the field of primary energy consumption forecasting.

LITERATURE REVIEW

Long Memory Time Series Model

Long memory time series modeling is a statistical technique used to model time series data that exhibits long-range dependence (Li et al., 2021). This means that the autocorrelation function (ACF) of the time series decays slowly, or even not at all. Long memory time series models are often used to model phenomena such as financial data, climate data, and internet traffic.

There are a number of different long memory time series models, but the most common is the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model. The ARFIMA model is a generalization of the ARMA model, which is a more traditional time series model. The ARFIMA model includes a fractional differencing parameter, which controls the decay rate of the ACF.

The general form of the ARFIMA(p, d, q) model is given by (Wei, 2006);

$$\phi(B)(1-B)^d(Z_t-\mu)=\theta(B)a_t \tag{1}$$

where :
t = Index of observations,
d = Differencing parameter (number of fractions),
μ = average of observations,
$a_t \sim$ IIDN($0,\sigma_a^2$),
$\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p\, B^p$ is a polynomial AR(p),
$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p\, B^p$ is a polynomial MA(q),
$(1-B)^d = \sum_{k=0}^{\infty}\binom{d}{k}(-1)^k B^k$ Fractional differentiation operator,

$$\sum_{k=0}^{\infty}\binom{d}{k}(-1)^k B^k = F(-d,1;1;B),$$

$$= \sum_{k=0}^{\infty}\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)}B^k.$$

METHODOLOGY

Identification

Basically, the identification of long memory patterns in time series can be a critical element in the development of the ARFIMA (Autoregressive Fractionally Integrated Moving Average) model. The ARFIMA model is a type of time series model that takes into account the presence of long-term dependencies or long memory in data (Tokhmpash et al., 2020). Identification of these long memory patterns can be done through spectral analysis and parameter estimation methods (Arteche & Robinson, 1998) .

First, in the identification of long memory patterns, spectral analysis can be used to evaluate frequency behavior in time series. If the spectrum shows a frequency that decreases slowly as the frequency increases, this may indicate a long memory. Spectral analysis can identify typical frequency characteristics of time series, helping us understand whether the data shows long-term dependencies.

Furthermore, parameter estimation methods are also important in the identification of long memory patterns. Parameter estimation in the ARFIMA model can involve determining the degree of fractional integration which measures the degree to which stochastic processes retain information from past observations.

Estimation Parameter

The parameters of the ARFIMA model can be estimated using a variety of methods, such as maximum likelihood estimation (MLE) and least squares estimation. Parameter estimation in the ARFIMA model is carried out in two stages, namely differentiating parameter (*d*) and parameter estimation $\emptyset$ and θ.

The assessment of ARFIMA model parameters through the spectral regression method for differentiating parameters *(d)* was proposed by Geweke and Porter-Hudak (Geweke & Porter-Hudak', 1983) , first forming a spectral density function into a linear regression equation and estimating the d parameter through the *Ordinary Least Square* (*OLS*) method. Meanwhile, the assessment of autoregression and moving average parameters is carried out after the fractional differentiation process using the maximum likelihood method (Palma, 2007).

A good ARFIMA model is one whose parameters are significant, or whose parameter values differ by zero. In general, if θ is a parameter of the ARFIMA model and $\hat{\theta}$ is the estimated value of that parameter, serta $SE(\hat{\theta})$ is *Standard error* of estimated value $\hat{\theta}$, Then the parameter significant test can be carried out with the following stages :

1. Hipotesis

   H_0 : $\theta = 0$ or the parameters of the ARFIMA model are equal to zero.

   H_1 : $\theta \neq 0$ or the parameters of the ARFIMA model are not equal to zero.

2. Test Statistics

$$t = \frac{\hat{\theta}}{SE(\hat{\theta})}.$$

3. Rejection Criteria H_0

Subtract H0 if $|t| > t_{\propto/2,df}$, with *df*= T-M, *T* = number of observations and *M* = number of parameters in the ARFIMA model, or by using *p-value*, i.e. reject H0 if p-value< ∝. In this study using ∝= 0.05.

Diagnostics Check

Diagnostic *checking* can be divided into two parts, namely parameter significance test and model conformity test (consisting of white noise assumption test and Normal distribution of residue). In ARFIMA modeling, the diagnostic check procedure involves assessing the goodness of fit and the adequacy of the model to the observed time series data. After estimating the parameters of the ARFIMA model, diagnostic checks typically include examining the residuals for stationarity, independence, and normality. Stationarity is crucial to ensure that the model captures the underlying patterns in the data, while independence of residuals is essential for the absence of systematic patterns or autocorrelation. Normality checks assess whether the residuals follow a Gaussian distribution. Additional diagnostics may involve inspecting the autocorrelation and partial autocorrelation functions of the residuals to identify any remaining patterns. Overall, a comprehensive diagnostic check ensures that the ARFIMA model adequately captures the temporal structure of the time series and that the assumptions underlying the model are met.

Forecast

The ARFIMA model can be used to forecast future values of a time series. The forecasts are generated by iteratively applying the ARFIMA model to the past values of the time series. Forecasting in the ARFIMA model is basically the same as the ARIMA model, in equation (1) it can be formed into the following equation;

$$\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p\right)(1-B)^d Z_t=\left(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q\right)a_t$$

$$Z_t-\phi_1 Z_{t-1}-\phi_2 Z_{t-2}-\ldots-\phi_p Z_{t-p}=\frac{\left(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q\right)a_t}{(1-B)^d}$$

From the Fractional differentiation operator, formed

$$(1-B)^d=\sum_{k=0}^{\infty}\frac{\Gamma(-d+k)}{\Gamma(-d)k!}B^k=\left(\sum_{k=0}^{\infty}\lambda_k(d)\right)B^k$$

By multiplying each quarter of the above equation by $\frac{a_t}{a_t}$ then the equation becomes,

$$Z_t=\phi_1 Z_{t-1}+\phi_2 Z_{t-2}+\ldots+\phi_p Z_{t-p}+\frac{a_t^2}{fd(t)}-\frac{\theta_1 a_{t-1}^2}{fd(t-1)}-\ldots-\frac{\theta_q a_{t-q}^2}{fd(t-q)}$$

With;

$$fd(t) = \left(\sum_{k=0}^{\infty} \lambda_k (d) \right) B^k a_t,$$

$$fd(t-1) = \left(\sum_{k=0}^{\infty} \lambda_k (d) \right) B^k a_{t-1},$$

.

.

.

$$fd(t-q) = \left(\sum_{k=0}^{\infty} \lambda_k (d) \right) B^k a_{t-q}.$$

The estimated *h* step forward is obtained by changing the t index to *T*+*h*

$$\hat{Z}_{T+h} = \hat{\phi}_1 Z_{T+h-1} + \ldots + \hat{\phi}_p Z_{T+h-p} + \frac{a_{T+h}^2}{fd(T+h)} - \ldots - \frac{\hat{\theta}_q a_{T+h-q}^2}{fd(T+h-q)}$$

Forecasting Evaluation

In this study, evaluate the model using the *Mean Absolute Percentage Error*

$$MAPE = \left(\frac{1}{H} \sum_{h=1}^{H} \left| \frac{e_h}{z_{T+h}} \right| \right) \times 100\%$$

with :

$$e_h = Z_{T+h} - \hat{Z}_T(h),$$

$h = 1, 2, 3, \ldots, H,$

RESEARCH RESULT

In this study, data on world primary energy consumption from 1965 to 2022 was used from https://ourworldindata.org/energy.

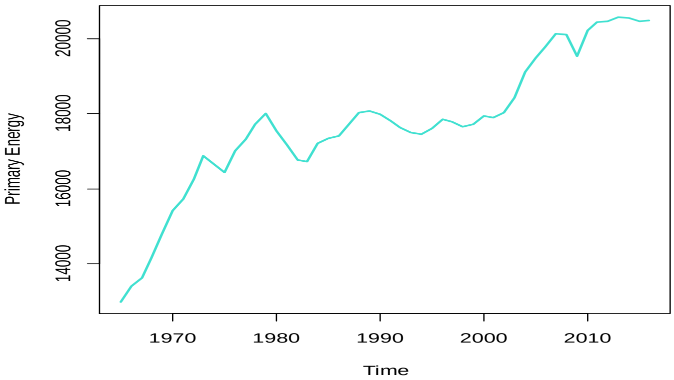

Figure 1. Primary Energy Consumption in The World

The primary energy of the world has grown steadily over the past 50 years, from about 7,200 TWh in 1965 to about 19,600 TWh in 2022. This growth has been driven by population growth and economic development. The figure shows that fossil fuels (oil, coal, and natural gas) continue to be the dominant source of primary energy, accounting for about 82% of global consumption in 2022. However, the share of renewables, such as hydropower, solar, and wind power, is growing rapidly. In 2022, renewables accounted for about 13% of global primary energy consumption.

The growth of primary energy consumption has had a significant impact on the environment. The burning of fossil fuels releases greenhouse gases into the atmosphere, which contribute to climate change. In addition, the extraction and processing of fossil fuels can have negative environmental impacts, such as water pollution and air pollution.

The transition to a clean energy future is essential to address the challenges of climate change and other environmental problems. Renewable energy sources offer a sustainable and low-carbon alternative to fossil fuels. However, the transition to a clean energy future will require significant investment in renewable energy technologies and infrastructure.

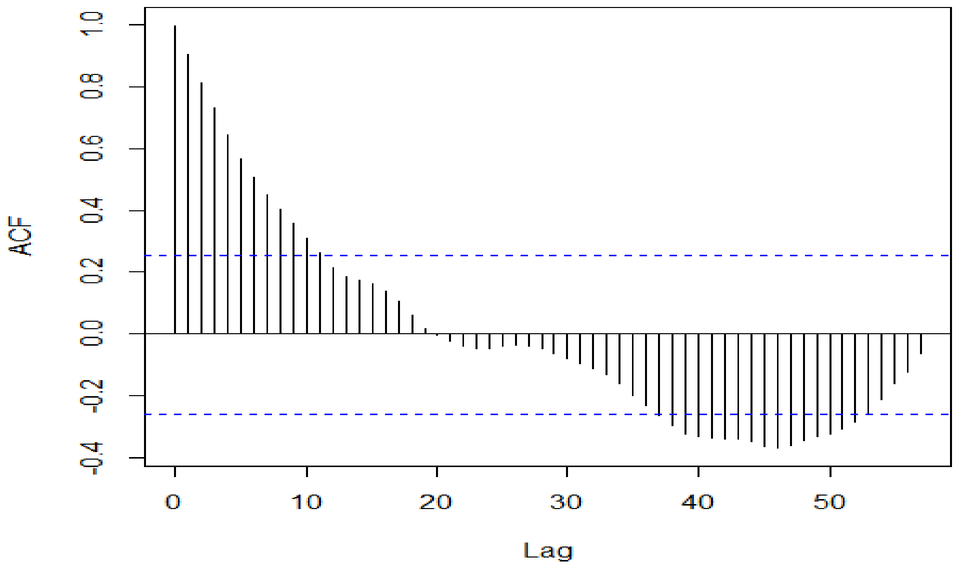

Figure 2. Autocorrelation Function plot of data

The ACF of the annual primary energy consumption in the world from 1965 to 2022 shows a strong positive correlation at lags of 1 to 10 years. This means that the consumption of energy in a given year is highly correlated with the consumption of energy in the previous 1 to 10 years. This suggests that there is a strong trend in energy consumption, with consumption increasing steadily over time.

The ACF also shows a weaker positive correlation at lags of 11 to 30 years. This suggests that there is also some seasonal variation in energy consumption, with consumption being higher in some years than others. However, the seasonal variation is not as strong as the trend.

Finally, the ACF shows a negative correlation at lags of 31 to 50 years. This suggests that there is a long-term cycle in energy consumption, with consumption increasing for a period of time followed by a period of decline. However, the long-term cycle is not very strong.

Overall, the ACF of the annual primary energy consumption in the world from 1965 to 2022 shows a strong trend in energy consumption, with consumption increasing steadily over time. There is also some seasonal variation in energy consumption, but it is not as strong as the trend. Finally, there is a weak negative correlation at lags of 31 to 50 years, which suggests that there may be a long-term cycle in energy consumption.

Table 1. MAPE with h=4 and h=6

Model	MAPE	Model	MAPE
ARFIMA(1,d,0)	1.58%	ARFIMA(1,d,0)	3.39%
ARFIMA(0,d,1)	4.29%	ARFIMA(0,d,1)	5.97%
ARFIMA(2.d,0)	1.55%	ARFIMA(2.d,0)	2.37%
ARFIMA(0,d,2)	3.71%	ARFIMA(0,d,2)	5.72%
ARFIMA(1,d,2)	3.71%	ARFIMA(1,d,2)	2.22%
ARFIMA(2,d,1)	1.38%	ARFIMA(2,d,1)	2.24%
ARFIMA(1,d,1)	1.52%	ARFIMA(1,d,1)	2.28%
ARFIMA(2,d,2)	1.48%	ARFIMA(2,d,2)	2.24%

The table presents the Mean Absolute Percentage Error (MAPE) values for different models, the models considered are ARFIMA models with different parameters. The table has two columns for each model, one for h=4 and one for h=6, with h is length of forecasting.

ARFIMA(2, d, 1) model has the lowest MAPE value of 1.38% for h=4 and 2.24% for h=6. The ARFIMA (1, d, 2) model also has low MAPE values of 1.48% for h=4 and 2.24% for h=6.The ARFIMA (0, d, 1) model has relatively higher MAPE values of 4.29% for h=4 and 5.97% for h=6. The ARFIMA (0, d, 2) model has even higher MAPE values of 3.71% for h=4 and 5.72% for h=6. The ARFIMA (1, d, 0) model has the highest MAPE values of 1.58% for h=4 and 3.39% for h=6.The MAPE values suggest that the ARFIMA (2, d, 1) and ARFIMA (1, d, 2) models are the most accurate in predicting energy consumption trends, for h =4 and h=6 respectively.

Statistical model of ARFIMA(2,d,1) is

$$(1-\emptyset_1 B-\emptyset_2 BB^2)(1-B)^d Z_t=(1-\theta_1 B)a_t$$

$$Z_t=\frac{1}{\Gamma(d+1)}(\emptyset_1 Z_{t-1}+\emptyset_2 Z_{t-2}) + \sum_{k=1}^{\infty}\left(\frac{(-1)^k\binom{d}{k}+\emptyset_1 B(-1)^k\binom{d}{k}+\emptyset_2 B^2(-1)^k\binom{d}{k}}{\Gamma(d-k+1)}\right)Z_{t-k} + (1+\theta_1 B)a_t$$

With

d = 0.7368, $\emptyset_1$= 0.085, $\emptyset_2$= 0.88, θ_1= -0.99

DISCUSSION

Long memory time series models, such as ARFIMA, capture the persistent and autocorrelated nature of energy consumption data, providing insights that traditional short memory models may overlook. By incorporating long memory components, these models aim to provide a more accurate representation of primary energy consumption.

Accurate energy demand modeling is crucial for informing energy policy decisions, improving energy access, and addressing climate change. Long memory time series modeling techniques can enhance the precision and reliability of forecasting primary energy consumption trends

The study found that the sectoral energy consumption in the world long memory behaviour. The energy consumption series showed strong non seasonal patterns and highly significant autoregressive and moving average components. Overall, the results of the study have important implications for the formulation and implementation of effective energy.

The findings of this research contribute to the advancement of energy forecasting methodologies, offering valuable insights for policymakers, energy analysts, and researchers in the field of primary energy consumption..The use of long memory time series modeling can support sustainable and informed energy planning, aiding in the development of effective energy policies and strategies.

CONCLUSIONS AND RECOMMENDATIONS

Data of primary energy consumption showed Presence of long memory pattern. Data from far away still have a reasonably strong association with each other, and data from the present has a relationship with data from a long time ago. The Proposed model (ARFIMA) to forecast primary energy consumption showed good performance (According to MAPE Values). By advancing the methodology for modelling primary energy consumption with long memory time series techniques, this research provides a valuable framework for developing more accurate energy forecasts.

The paper suggests that incorporating long memory time series modelling techniques can improve the accuracy and reliability of primary energy consumption forecasts. It recommends further research and exploration of long memory time series models to better understand the underlying dynamics and dependencies in primary energy consumption data. Theis research also emphasizes the need for continuous monitoring and analysis of primary energy consumption data to capture any evolving long memory patterns and adjust forecasting models accordingly.

ADVANCED RESEARCH

Further research was tried for seasonal data with different patterns such as additive or multiplicative patterns. Further research can also explore the application of long memory time series modelling techniques to other sectors of energy consumption, beyond primary energy, to gain a comprehensive understanding of energy dynamics and improve forecasting accuracy.

Investigate the potential of combining long memory models with other forecasting methods, such as machine learning algorithms, to enhance the predictive capabilities of energy consumption models. Explore the use of alternative performance evaluation metrics to assess the accuracy and reliability of long memory time series models in energy forecasting, providing a more comprehensive analysis of model performance.

Conduct comparative studies between long memory models and other traditional time series models, such as ARIMA, to determine the superiority of long memory techniques in capturing long-term dependencies and improving forecasting accuracy. Investigate the impact of incorporating exogenous variables, such as economic indicators or climate data, into long memory time series models to enhance the predictive capabilities and robustness of energy consumption forecasts

REFERENCES

Adekoya, O. B. (2020). Modeling of Persistence and Seasonality in Sectoral Energy Consumption in the USA Using Fractionally Integrated Processes: Implications for Economic Policy. Natural Resources Research, 29(4), 2787–2800. https://doi.org/10.1007/s11053-019-09599-x

Andrysiak, T., & Saganowski, Ł. (2015). Network Anomaly Detection Based on ARFIMA Model. Advances in Intelligent Systems and Computing, 313 AISC, 255–261. https://doi.org/10.1007/978-3-319-10662-5_31

Arteche, J., & Robinson, P. M. (1998). SEMIPARAMETRIC INFERENCE IN SEASONAL AND CYCLICAL LONG MEMORY PROCESSES Ã.

Balagula, Y. (2020). Forecasting daily spot prices in the Russian electricity market with the ARFIMA model. Applied Econometrics, 57, 89–101. https://doi.org/10.22394/1993-7601-2020-57-89-101

Berbesi, L., & Pritchard, G. (2023). Modelling Energy Data in a Generalized Additive Model—A Case Study of Colombia. Energies, 16(4). https://doi.org/10.3390/en16041929

Cao, Z., O'Sullivan, C., Tan, J., Kalvig, P., Ciacci, L., Chen, W., Kim, J., & Liu, G. (2019). Resourcing the Fairytale Country with Wind Power: A Dynamic Material Flow Analysis. Environmental Science and Technology. https://doi.org/10.1021/acs.est.9b03765

Geweke, J., & Porter-Hudak', S. (1983). THE ESTIMATION AND APPLICATION OF LONG MEMORY TIME SERIES MODELS.

Li, D., Robinson, P. M., & Shang, H. L. (2020). Long-Range Dependent Curve Time Series. Journal of the American Statistical Association, 115(530), 957–971. https://doi.org/10.1080/01621459.2019.1604362

Li, D., Robinson, P. M., & Shang, H. L. (2021). Local Whittle estimation of long-range dependence for functional time series. Journal of Time Series Analysis, 42(5–6), 685–695. https://doi.org/10.1111/jtsa.12577

Monge, M., & Gil-Alana, L. A. (2019). Automobile components: Lithium and cobalt. Evidence of persistence. Energy, 169, 489–495. https://doi.org/10.1016/j.energy.2018.12.068

Palma, W. (2007). LONG-MEMORY TIME SERIES Theory and Methods.

Shalalfeh, L., Bogdan, P., & Jonckheere, E. A. (2021). Fractional Dynamics of PMU Data. IEEE Transactions on Smart Grid, 12(3), 2578–2588. https://doi.org/10.1109/TSG.2020.3044903

Silver, S. D., & Raseta, M. (2021). An ARFIMA multi-level model of dual-component expectations in repeated cross-sectional survey data. Empirical Economics, 60(2), 683–699. https://doi.org/10.1007/s00181-019-01757-7

Tokhmpash, A., Hadipour, S., & Shafai, B. (2020). On Analysis of Fractional Order System Identification. https://doi.org/10.0/Linux-x86_64

Wei, W. W. S. (2006). Time series analysis : univariate and multivariate methods. Pearson Addison Wesley.

Zaharia, A., Diaconeasa, M. C., Brad, L., Lădaru, G. R., & Ioanăs, C. (2019). Factors influencing energy consumption in the context of sustainable development. Sustainability (Switzerland), 11(15). https://doi.org/10.3390/su11154147