International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-10, Issue-5, September 2022
https://doi.org/10.55524/ijircst.2022.10.5.12
Article ID IRP1312, Pages 87-94
www.ijircst.org

Exploratory Data Analysis and Comparison of Total Energy Consumption in Major World Countries Using Artificial Intelligence

Mohammad Sameer Parray [1], Dharminder Kumar[2] and Satish Saini[3]

[1]M. Tech Scholar, Department of Electrical Engineering, RIMT University, Mandi Gobingarh, Punjab, India
[2]Assistant Professor, Department of Electrical Engineering, RIMT University, Mandi Gobingarh, Punjab, India
[3]Professor, Department of Electrical Engineering, RIMT University, Mandi Gobindgarh, Punjab, India

Correspondence should be addressed to Mohammad Sameer Parray; sameerparray904@gmail.com

ABSTRACT- Accurate power forecasting for electrical consumption prediction is crucial for national energy planning since it is a method for understanding and anticipating market energy demand. In a deregulated market, their power output can be modified accordingly. In order to explicitly deal with seasonality as a class of time-series forecasting models, Persistence Models (Naive Models), Seasonal Autoregressive Integrated Moving Averages with Exogenous regressors, and Univariate Long-Short Term Memory Neural Network (LSTM) are utilized. The main goal of this project is to do an exploratory data analysis of the major nations' power systems, followed by the application of several forecasting models to determine the daily peak demand and the next 24 hours' worth of energy need once every day. Energy is a vital resource for society's development. The development of the economy depends heavily on an accurate assessment of its consumption. A database of historical energy consumption statistics from all the main nations of the globe was created. The complete information was modelled and simulated using machine learning techniques to anticipate the energy consumption of each country because national patterns may be transferrable from one country to another. Understanding the commonalities between the characteristics of other nations might boost prediction accuracy and enhance the corresponding public policies.

KEYWORDS- LSTM, Autoregression, World Countries, Energy Consumption

I. INTRODUCTION

Since the Industrial Revolution, the world's population has been growing, and in recent years, as seen in Fig. 1, this expansion has accelerated. From 4.84 billion in 1985 to 7.59 billion in 2018, the global population increased (multiplied by a factor of 1.57). Furthermore, as technology has advanced, so have individual energy demands. For instance, the amount of electronic gadgets that each person has, the demand for transportation, and the need for air conditioning to enhance living conditions and workplace comfort have all grown [3] Indeed, as mentioned in the 2030 United Nations Sustainable Progress Goals, energy availability is a crucial component of economic development and human wellbeing [5] As a result, over the same time period, the world's gross energy generation increased dramatically. From 9880 TWh in 1985 to 26,615 TWh in 2018, it rose. As can be seen in Fig. 1, it was subsequently increased by a factor of 2.69.

Today, each nation strives to use domestic fossil fuel sources, if available, uses sustainable and renewable energy generating programs, or imports energy from neighbouring nations, in case of supply shortages, to address the ever-increasing demand for energy and decide policies [2] Blackouts have a detrimental impact on industries reliant on energy when the supply is insufficient to meet the demand. In contrast, power plant surplus energy capacity may suggest over capacity.

. Consequently, it is crucial to model historical energy usage and forecast future demand within a nation in order to plan the building of new power plants, which are required for sustained economic growth [7].

Identifying the factors that affect energy consumption and future policymaking is the first crucial step in creating an accurate energy consumption forecasting model (Yuan et al., 2017). The population of the nation is one factor that is directly related to consumption.

The number of people consuming electricity and the number of homes connected to the grid both rise as the population [9]. Along with the demographic component, a country's economic development level reflects the customs of its residents [11] . More living standards often correspond to higher consumption. The link between per-person gross domestic product (GDP) and energy usage allows for the observation of this relationship [13] In contrast, some elements like inflation and a high unemployment rate may have a negative impact on consumption. For instance, during a financial crisis, inflation and unemployment rates rise. People try to save costs as a result, which causes a drop in energy use[10] .

Numerous papers discuss various techniques for predicting future energy usage. Sadly, the concepts and findings presented in these books are unique to the nations under investigation. However, it's feasible that national patterns of energy use may be carried over from one nation to another, which would improve the accuracy of predictions. There may be parallels in the energy consumption characteristics of different member nations given that they all aim to meet the major country criteria. As a result, it is feasible to compile historical data on Energy consumption from all of the main nations into a large database, spot broad patterns and similarities, and use machine learning to estimate the future for each nation using a global model.

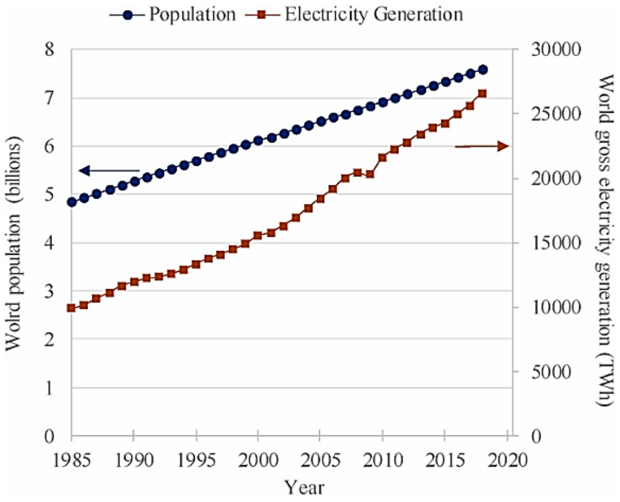

Figure 1: Changes in the world population (World Bank, 2019) and gross Energy generation (BP, 2019) over time

In order to build a worldwide machine learning model that can forecast each country's energy consumption, this study compiles socioeconomic data from all of the main nations and historical yearly energy consumption into a database. This strategy is predicated on the notion that certain nations may mimic the energy consumption patterns of other nations. In other words, the future of one nation may be influenced by the experiences of another.

Thus, using machine learning techniques on this information may assist to identify patterns in the energy profiles of different nations. Additionally, these parallels can boost forecast precision and boost the dependability of the linked policies. We created this document to address the demand for a worldwide forecasting model based on the results of our literature research. To our knowledge, no research has ever used a worldwide model to anticipate energy use across a range of nations and develop future policy

II. LITERATURE REVIEW

[2] proposed a novel hybrid model including eps-SVR and nu-SVR models to develop a performance prediction model for forecasting the Energy load of buildings. They employed a differential evolution (DE) algorithm to optimize the performance of the Support Vector Regression (SVR) model, finding the best model parameters and corresponding weights for both eps-SVR and nu-SVR models to forecast both half-hourly as well as daily Energy consumption for an institutional building in Singapore. The results of their proposed model showed a lower mean absolute percentage error (MAPE) for both the daily and half-hourly energy consumption data.

[4] compared the ML model's accuracy in predicting the hourly HVAC energy consumption of a hotel in Madrid, Spain, utilizing two machine learning-based methods, namely artificial neural networks and random forests (RF). It was found that ANN is capable of performing marginally better than RF with a root-mean-square error (RMSE) of 4.97 and 6.10 respectively. However, it was seen that both of the models can be feasible and effective in building energy prediction

III. METHODOLOGY

Data from 1994 to 2019 for 36 countries were collected from various sources, and a database with 936 entries was constructed. The population, gross domestic product, inflation rate, and unemployment rate were used as the descriptor variables (input variables) influencing Energy consumption. Information related to the population was collected from the "World Population Prospects" section of the United Nations website [3] Then, gross domestic product (based on purchasing power parity in US dollars) data for each country was extracted from the "GDP and spending" section of the MAJOR website [14] Data related to inflation rates (based upon the consumer price index) and the unemployment rate were taken from the World Bank Finally, the gross Energy consumption (including losses related to Energy transmission and distribution) of each country was calculated by adding the annual gross Energy production (IEA, 2020; OECD, 2020a) to the net Energy imports (COMSTATS, 2020; IEA, 2020). As a result, a database composed of four descriptors and one target variable was formed.

A. Computational Details

As indicated in the literature, most studies that have proposed energy-related forecasting techniques applied artificial neural networks and support vector machines. In addition to these methods, decision trees and random forest regression have been used extensively in regression-based learning approaches. Thus, random forest regression, artificial neural networks, and support vector machines were applied in this study. All of the computational codes related to the artificial neural networks (one hidden layer multi-layer perception type) were written in MATLAB (version number 9.4; R2018a). First, the "map minmax" function was used to normalize the input variable range between − 1 and +1. The networks were then trained using the backpropagation algorithm with the "*trainbr*" network training function. This function uses Bayesian regularization to obtain high generalization accuracy when predicting results from unexplored data regions [6]

Levenberg-Marquardt optimization was used to update weight and bias values. Several artificial neural network topologies with different numbers of hidden neurons were tested using the hyperbolic tangent

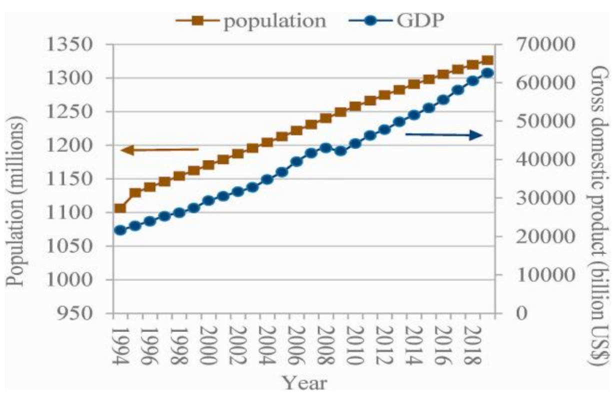

Figure 2: Change in the combined gross domestic product (based on purchasing power parity) and population of the Major countries

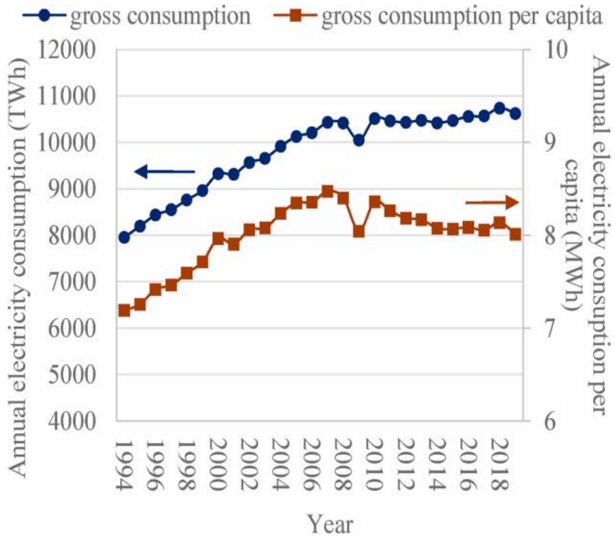

Figure 3: Changes in the annual gross Energy consumption of all major countries, combined

Support vector regression models were developed using ‘sklearn’ in the Python (version 3.7.4) environment. The training process was performed using symmetrical loss functions, and positive and negative misestimates were penalized equally. Therefore, the absolute values of the errors were minimized by forming a flexible tube via the radius minimization approach [10]

Moreover, in cases where the available data was linearly inseparable, nonlinearity tools were adopted via a kernel trick. Based on our preliminary studies, radial basis function (RBF) was adopted as a nonlinearity tool for support vector regression. The selected support vector regression kernel enabled the utilization of three main prediction parameters: cost (C), gamma (γ), and epsilon (ε). The parameter C is forming a trade-off between the accuracy and allowance of crossing from the margin with error. margin between the two groups was expected to decrease as the value of the parameter C increased for all training points. However, the margin was expected to increase as the parameter decreased, yielding a higher training error and weak predictive capabilities. The parameter γ, on the other hand, represented the farness (low values) and closeness (high values) of the influence of an individual training sample. The third parameter, ε, was used to define the tolerance for exceeding the margin when no penalty is applied in response to errors. A larger epsilon indicated that larger errors were allowed in the solution. In contrast, every error was penalized if the parameter converged to [3] Since there are many trade-offs among these three parameters, a grid search algorithm was adopted to tune them

Table 1: The rule of thumb for interpretation of the Pearson correlation coefficients

Pearson Correlation Coefficient Interval	Interpretation
[0.90, 1.00]	very high positive correlation
[0.70, 0.90)	high positive correlation
[0.50, 0.70)	moderate positive correlation
[0.30, 0.50)	low positive correlation
(-0.30, 0.30)	negligible correlation
(-0.50, − 0.30]	low negative correlation
(-0.70, − 0.50]	moderate negative correlation
(-0.90, − 0.70]	high negative correlation
[-1.00, − 0.90]	very high negative correlation

Likewise, random forest regression models were developed using *‘sklearn’* from the Python environment. The random forest regression technique uses multiple decision trees and is known as bootstrap aggregation [7] The method trains each decision tree using a separate data sample, and sampling is performed with replacement. Bootstrap aggregation is commonly referred to as ‘bagging.’ Since the decision trees are specific to the particular data on which they are trained, the resulting tree can yield inaccurate and computationally expensive results via high overfitting. Concerns regarding decision tree algorithm inaccuracy occur mainly due to their inability to return to the previous phase after a split. This inability can cause the algorithm to reach a local optimum rather than the optimum global state. Therefore, a random forest regression algorithm is used instead of individual decision tree

B. Overview of the Countries

In this section, changes in the total GDP, total population, and total annual Energy consumption among Major countries are evaluated. First, Fig. 2 shows the change in the total population (primary axis) and gross domestic product (secondary axis) for all 36 countries combined. According to the figure, the GDP increased from US $21.6 trillion in 1994 to US $62.5 trillion in 2019, while the total population increased from 1.11 billion to 1.33 billion during the same period.

Fig. 3 shows changes in the total annual gross Energy consumed and Energy consumed per capita for the 36 MAJOR countries combined. As indicated in the figure, the total Energy consumption increased from 7954 TWh in 1994 to 10,625 TWh in 2019. Consumption did not change significantly after 2007 (except in 2009, which corresponds to the global financial crisis), probably due to a combination of the widespread use of energy-efficient equipment and devices and technology development. The figure also shows that the average annual Energy consumption per capita had a peak of 8.47 MWh in 2007 and exhibits a sharp decrease in 2009 followed by a slight decrease after 2010, possibly due to the reasons mentioned above.

IV. SYSTEM IMPLEMENTATION

The main motivation for this work is to combine past annual Energy consumption and socioeconomic data for all the major countries and create a database that can be used to produce a global model of future Energy consumption. This approach was designed based on the expectation that a particular major world country may follow the Energy consumption profile of some other country. In other words, the past experience of a particular country may be reflected in the future of another country.

The conceptual approach is shown in Fig. 4, which summarizes all of the steps applied in this work. In the first step, a global correlation analysis is performed using all of the descriptor variables and Energy consumption. For this purpose, the population, GDP per capita, inflation rate, and unemployment rate are used as the descriptor variables and correlated with the Energy consumption during the previous year (lag term). Hence, the descriptor variables that correlate with Energy consumption throughout the entire database are determined by checking the Pearson correlation coefficients (R values), and any descriptor variables without significant correlation with the target ($|R|<0.30$) are eliminated. This finding is different from the previous correlation examination, based on data from each country. Thus, the optimal set of input variables is determined.

In the second step, the significant descriptor variables are used to build several global models, and their performance characteristics were analysed deeply. For this purpose, the forecasting ability of a global model was checked by dividing the dataset into three subsets. The first subset included data from all countries for 1994–2012 and was used as the training set when building the model, whereas the 2013–2015 period was used as the validation set. The global model parameters were tuned to minimize the combined training and validation error. The data from 2016 to 2019 were used as the test set and was not seen by the model until the model was complete. Three global models were built using artificial neural networks, support vector regression, and random forest regression. The model with the smallest test error was chosen as the one with the highest forecasting accuracy, and future predictions were performed using the best model for each country.

After the best global model with the highest predictive power was identified, it was used to make predictions using the same (1994–2012) training set. However, at the final stage, the validation period was expanded to include the 2013–2019 period for each country, and

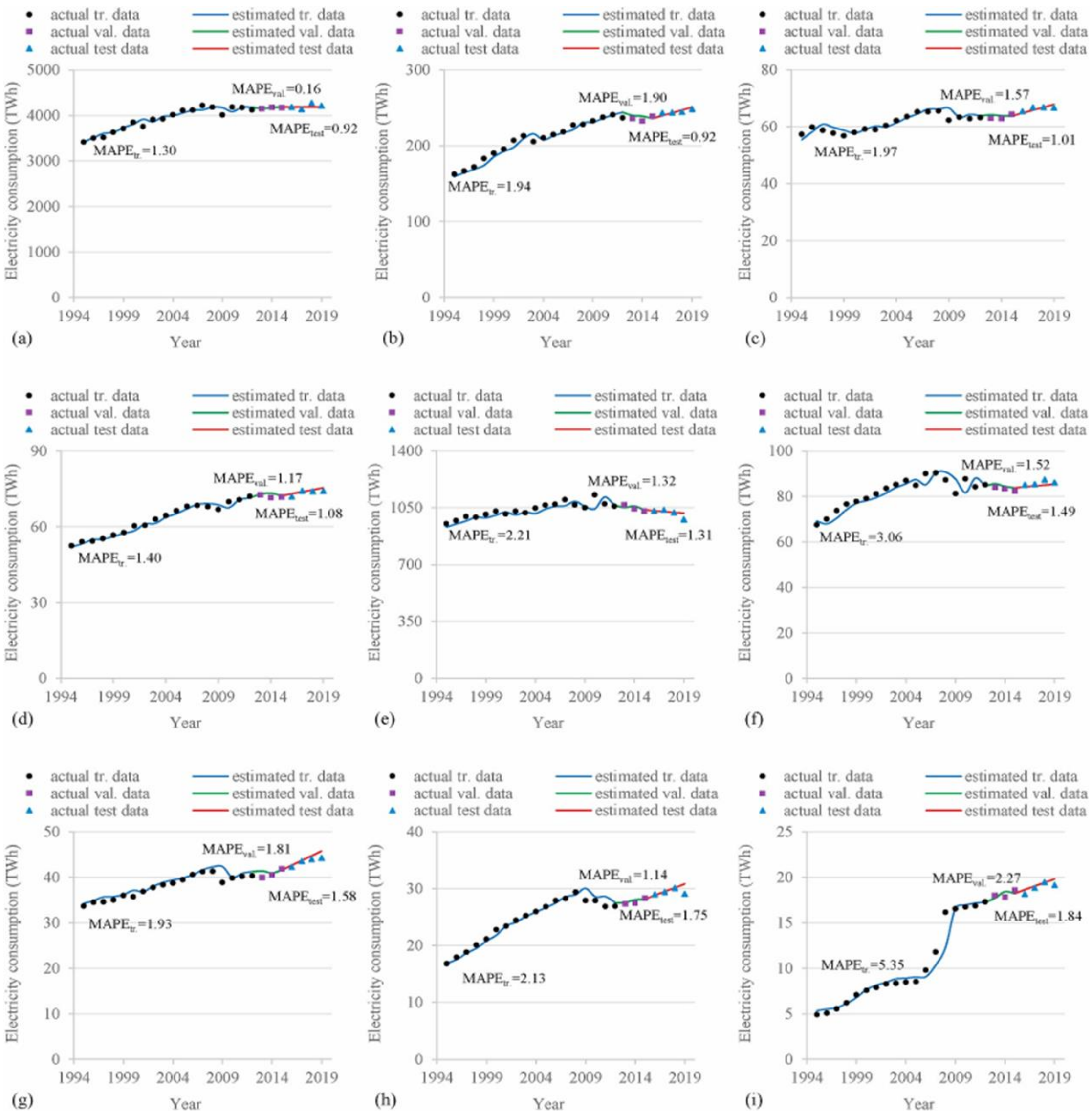

Figure 4: Tuning the global models and testing their predictive capabilities

A. Determining the Optimal Set of Descriptor Variables

During the first step, a global correlation analysis between descriptor variables as well as between descriptor variables and the target variable was performed using the training data set (1994–2012) (Table 1). The most significant correlations between descriptor variables are between the GDP per capita and the unemployment rate (R = – 0.40), and between the GDP per capita and the inflation rate (R = – 0.33), both of which correspond to "low negative correlation." In order to check for multicollinearity between the descriptor variables, the variance inflationary factor (VIF) of each descriptor variable was calculated. The values are 1.02, 1.35, 1.13, and 1.21 for the population, GDP per capita, inflation rate, and unemployment rate, respectively. As a rule of thumb, the VIF factor should be less than ten to avoid multicollinearity (Kennedy, 2008; O'Brien, 2007). In the current case, all of the variables follow this rule. On the other hand, the correlation analysis between the descriptor variables and the target variable indicates that only the population term exhibits a significant correlation (very high positive correlation). The GDP per capita, inflation %, and unemployment % indicate no significant correlation based on the overall data.

Although the GDP per capita, inflation rate, and unemployment rate correlate with the Energy consumption information for some individual countries (Table 1) no general trend is observed between these variables and the Energy consumption when the entire database (including data for all MAJOR countries) is considered. An ARIMA analysis was used to determine the number of lag terms. It was found that 30 out of 36 countries have one or zero lag terms, whereas four countries have two lag terms and only two countries have three lag terms. Hence, it is preferred to use one lag term to eliminate redundant lag terms throughout the database and make the model more generalizable. In the following sections, only the population is used as the only descriptor variable together with the lag term to design the global models. The generic artificial neural network, support vector regression, and random forest regression global model structures.

Several artificial neural network, support vector regression, and random forest regression models were built, and their performances were compared. Because the objective was to predict the target variable several years (steps) ahead, the model parameters were tuned to minimize the validation error corresponding to a point a couple of years later than the training data (holdout method for cross-validation).

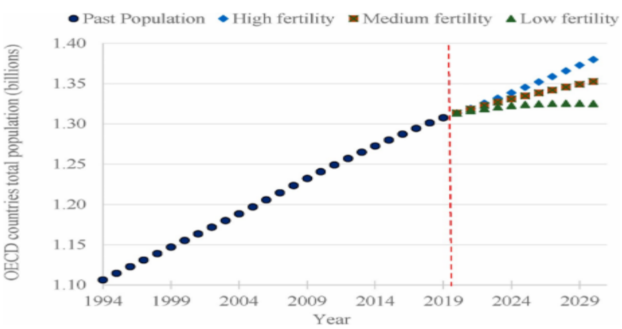

Figure 5: Combined population of all Major countries based on various future population projection assumptions (United Nations, 2020)

The period from 1994 to 2012 was selected as the training set, while the data from 2013 to 2015 was chosen as the validation set. The model that could estimate the target variable in the validation set with the smallest error was selected as the optimal model and was used to predict the target variable based on a separate test set (never before seen by the model) corresponding to 2016 to 2019. It should be noted that lag terms corresponding to the validation and test periods (2013–2019) were generated step-by-step from the predicted values of the targets (i.e., the predicted Energy consumption for 2018 was used as the lag term to predict the consumption in 2019). Thus, the global models were forced to predict future target values without any information related to the future, which avoided information leakage (see step #2 in Fig. 2).

For the artificial neural networks, one hidden layer containing various numbers of hidden neurons was tested. Because artificial neural networks may produce different results during different runs due to random initialization of weights, each neural network topology was run 1000 times. The model with the lowest validation error was recorded.

Fig. 6 shows the training and validation errors versus the number of neurons used. As indicated by the figure, the training errors tend to decrease as the network grows, whereas the validation error first decreases and tends to increase as the network grows. Hence, the selected topology with the lowest validation error (highest accuracy when predicting unknown future data) has ten neurons in the hidden layer. This network produces a training root mean square error (RMSE) value of 10.1 and a validation RMSE value of 10.7.

B. Predicting future Energy Consumption Using the Global Model

Finally, a new global support vector regression model was created (trained using the 1994–2012 data and validated using the 2013–2019 data), as shown in the 3rd step in Fig. 4. The goal is to forecast the energy consumption of each MAJOR country for 2020–2030 using the population and the lag term as the descriptor variables. As explained in Section 5.2, lag terms for the future years are generated step-by-step from the predicted target values (i.e., the predicted 2029 Energy consumption was used as the lag term for prediction of 2030 consumption). Future population data for each country were collected from "World Population Prospects" under the United Nations website (UnitedNations, 2020) based on low, medium, and high fertility assumptions. For instance, according to the medium-fertility assumption, it is expected that worldwide fertility will decrease from 2.47 births per woman in 2015–2020 to 2.38 births per woman in 2025–2030. How-ever, in Europe, overall fertility is expected to rise from 1.61 births per woman in 2015–2020 to 1.64 births per woman in 2025–2030. In North America, fertility is expected to rise from 1.75 births per woman to 1.76 births per woman during the same periods. Based on medium-fertility projections, the world population is expected to reach 8.55 billion in 2030. However, compared to the medium variant, global fertility for the high variant was assumed to be 0.25 births higher for the 2020–2025 period and 0.4 births higher for the 2025–2030 period. Low fertility produces 0.25 fewer births per woman in 2020–2025 and 0.41 fewer births in 2025–2030, relative to medium fertility.

The projected combined population of the Major countries is shown in Fig. 10. Based on high fertility assumptions, the combined population of the MAJOR countries is projected to rise linearly, reaching 1.38 billion in 2030. Under medium-fertility assumptions, the increase in population is expected to slow to 1.35 billion in the same year. Finally, under low fertility assumptions, the population remains nearly constant at approximately 1.33 billion throughout the projected period. In this paper, the medium-fertility assumption is chosen to predict the future population of each MAJOR country, and the global support vector regression model is built accordingly.

The global support vector regression model was used to forecast Energy consumption among individual MAJOR countries. Medium- fertility predictions corresponding to the same nine countries (with the smallest mean percent error for forecasts of 2016–2019) are shown in Figure 8 and 9 compares past and forecasted Energy consumption for all MAJOR countries. The reliabilities of these forecasts are based on the ability of the global model to predict target variables that correspond to a period ahead of the training data. In other words, the global model that provides the smallest errors when forecasting 2016–2019 can be expected to provide small errors for future predictions (i.e., consumption estimates for the USA can be assumed to be the most reliable, whereas estimates for Latvia can be assumed to be the least reliable). Among the most reliably forecasted nine countries, the Energy consumption of the USA, Japan, and Finland is predicted not to change significantly during the next ten years. In contrast, Ireland, Hungary, Iceland, Australia, Czechia, and Austria are expected to exhibit increasing Energy consumption during the same period.

It should be noted that the actual Energy consumption can be lower than forecasted numbers, especially if governments conduct energy-saving programs. For instance, a successful demand-side management program saves energy and thus helps to reduce Energy consumption in the future (Ardakani and Ardehali, 2014). In the United States, such a program began in the 1970s and became more active after the 1990s. The efficient implementation of this program resulted in energy savings of 87.8 TWh in 2010 (roughly 2.3 percent of total Energy consumption) and 139.2 TWh in 2012. (about3.7 percent of total Energy consumption) (Ardakani and Ardehali, 2014; EIA, 2018). The application of this type of program to other countries can reduce future Energy consumption significantly. Moreover, smart energy monitors, which allow homeowners to track the Energy usage of various appliances in real-time, are used widely by utilities in various countries. A machine learning-based study recently reported that such devices reduced household Energy consumption by an average of 7.1% in Hong Kong (Wang et al., 2020). Finally, many governments have passed laws that phase out incandescent light bulbs in favour of more energy- efficient lighting replacements. For instance, the replacement of conventional light bulbs with light-emitting diodes in Japan has reportedly reduced household Energy usage by 1.96% (Onuma et al., 2020). Likewise, it was recently estimated that Switzerland would reduce household lighting energy consumption by 62% by 2035 by replacing old bulbs (Heidari et al., 2018).

V. SIMULATION AND RESULTS

Table 1: Types of energy and its consumption throughout the major countries

	Entity	Code	Year	Fossil	Renewables	Nuclear	continent
0	Austria	AUT	1965	76.060	23.940	0.0	Europe
1	Austria	AUT	1966	75.340	24.660	0.0	Europe
2	Austria	AUT	1967	75.069	24.931	0.0	Europe
3	Austria	AUT	1968	76.058	23.942	0.0	Europe
4	Austria	AUT	1969	78.788	21.212	0.0	Europe

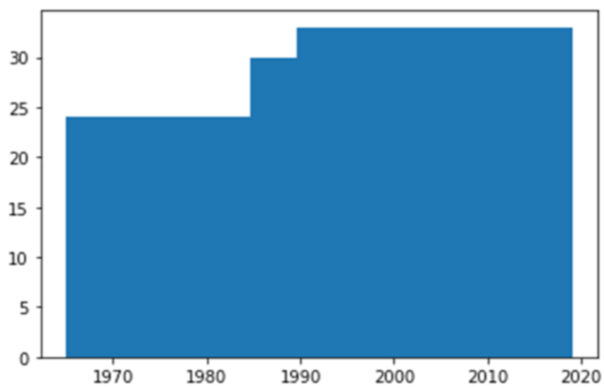

Figure 6: Energy consumption through year 1970 to 2022

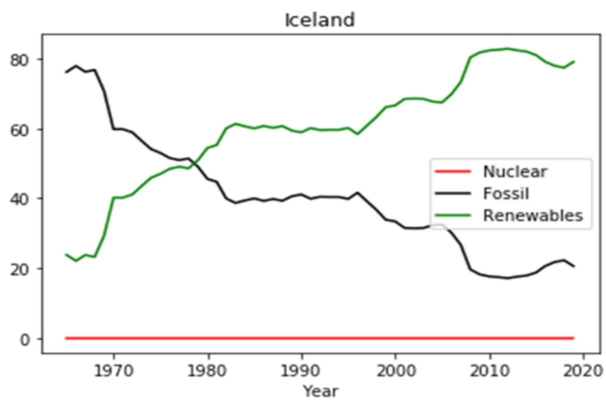

Figure 7: Nuclear fossil and renewable energy throughout the years in Iceland

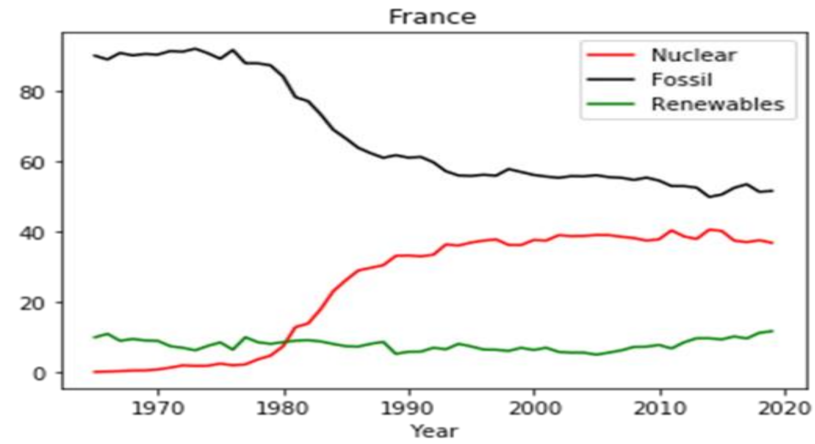

Figure 8: Nuclear fossil and renewable energy throughout the years in Iceland

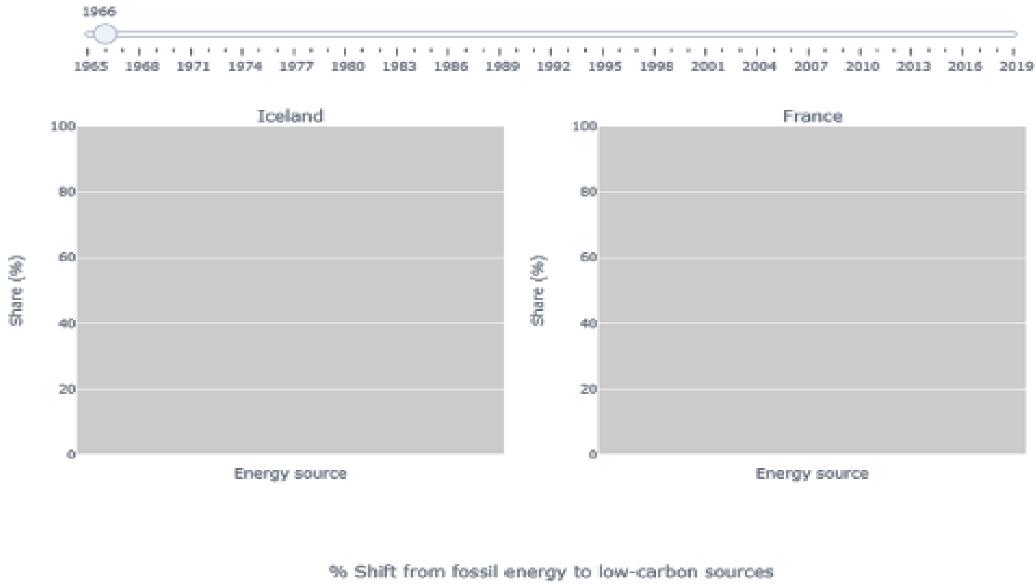

Figure 9: Comparison of energy consumption in Iceland and France

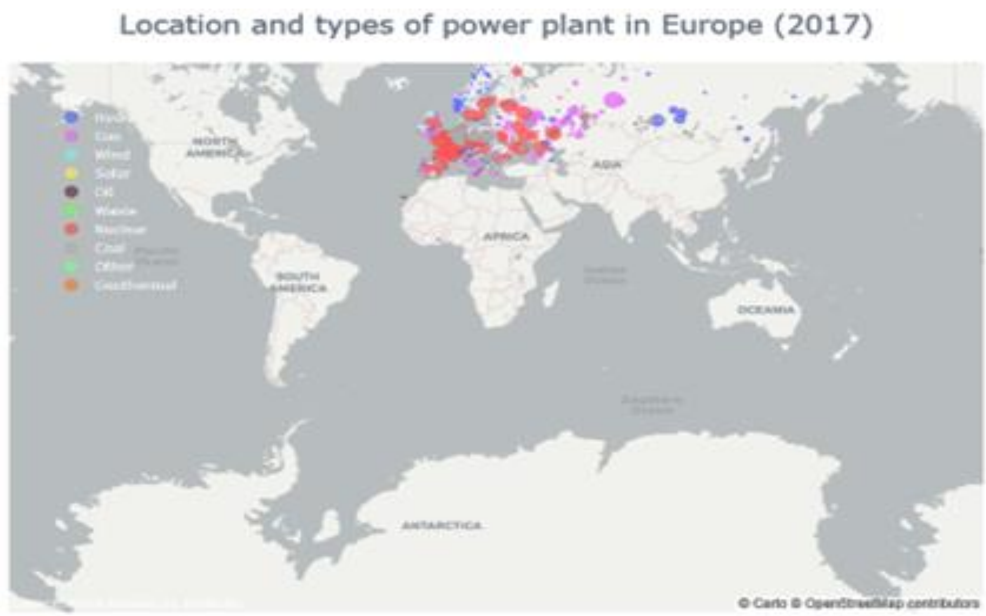

Figure 10: Types of power plants in Europe

VI. CONCLUSION

In this study, global models based on machine learning that were trained using a database that includes data from all major countries were created to model and predict the yearly patterns of energy consumption in various nations. The hypothesis that historical trends in one nation may accurately anticipate future trends in other countries served as the study's inspiration.

A database with 936 data points was created using historical socioeconomic (population, gross domestic product, inflation rate, and unemployment rate) and energy consumption statistics for 36 nations between 1994 and 2019 from a variety of sources. Artificial neural networks, support vector regression, and random forest regression were three different machine learning global models that were built. The support vector regression model's forecasting precision was higher than that of the artificial neural network model and marginally better than that of random forest regression. To predict each major country's yearly energy consumption from 2020 to 2030, a new support vector regression model was developed. Countries with modest validation errors were anticipated to produce better accuracies in the predicted horizon since the global model was utilized to create country-based forecasts.

In conclusion, creating a database using historical data from a number of nations and using machine learning techniques on the database can aid in predicting the expected energy profiles of other nations. These parallels can boost forecast precision and dependability, supporting the creation of pertinent laws. It is difficult and essential to estimate energy usage accurately since it is important for economic and social as well as environmental reasons... It is anticipated that predictions made using the global modelling technique would assist governments and policymakers in finding solutions to pressing problems. The methodology used in this paper can also be extended to other geographic or developmental groups of nations, such as the European Union, countries in South America, or countries in Africa (e.g., developed countries or developing countries)

REFERENCES

[1] Al-Musaylh, M.S., Deo, R.C., Adamowski, J.F., Li, Y., 2018. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Adv. Eng. Inf. 35, 1–16.

[2] Angelopoulos, D., Siskos, Y., Psarras, J., 2019. Disaggregating time series on multiple criteria for robust forecasting: the case of long-term Energy demand in Greece. Eur. J. Oper. Res. 275, 252–265.

[3] Awad, M., Khanna, R., 2015. Support Vector Regression, Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Apress, Berkeley, CA, pp. 67–80.

[4] Bannor, B.E., Acheampong, A.O., 2019. Deploying artificial neural networks for modeling energy demand: international evidence. Int. J. Energy Sect. Manag. 14, 285–315.

[5] Bessec, M., Fouquau, J., 2018. Short-run Energyload forecasting with combinations of stationary wavelet transforms. Eur. J. Oper. Res. 264, 149–164.

[6] Bianco, V., Manca, O., Nardini, S., 2009. energy consumption forecasting in Italy using linear regression models. Energy 34, 1413–1421.

[7] Bianco, V., Manca, O., Nardini, S., 2013. Linear regression models to forecast electricity consumption in Italy. Energy Sources B Energy Econ. Plann. 8, 86–93.

[8] Bianco, V., Scarpa, F., Tagliafico, L.A., 2014. Analysis and future outlook of natural gas consumption in the Italian residential sector. Energy Convers. Manag. 87, 754–764.

[9] Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. Association for Computing Machinery, Pittsburgh, Pennsylvania, USA, pp. 144–152.

[10] BP, 2019. Statistical Review of World Energy. https://www.bp.com/en/global/corpora te/energy-economics/statistical-review-of-world-energy.html. (Accessed 21 April 2021).

[11] Cabral, J.d.A., Freitas Cabral, M.V.d., Pereira Júnior, A.O., 2020. Elasticity estimation and forecasting: an analysis of residential Energy demand in Brazil. Util. Pol. 66.

[12] COMSTATS, 2020. Energy Consumption, Export & Import 1980-2017. http://comsta t.comesa.int/sqhysj/electricity-consumption-export-import-1980-2017. (Accessed 21 April 2021).

[13] Cutler, A., Cutler, D.R., Stevens, J.R., 2012. Random Forests, Ensemble Machine Learning, 1 ed. Springer-Verlag, New York, pp. 157–175.

[14] Dangeti, P., 2017. Statistics for Machine Learning: Techniques for Exploring Supervised, Unsupervised, and Reinforcement Learning Models with Python and R. Pact Publishing, Birmingham. de Oliveira, E.M., Cyrino Oliveira, F.L., 2018. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. Energy 144, 776–788.