THE ENERGY CONSUMPTION FORECASTING IN MONGOLIA BASED ON BOX-JENKINS METHOD (ARIMA MODEL)

Gansukh Zolboo, Bor Adiya, Enkhbayar Bilguun

Abstract

The primary products of the power industry are electric energy and thermal energy. Thus, forecasting electric energy consumption is significant for short and long term energy planning. ARIMA model has adopted to forecast energy consumption because of its precise prediction for energy consumption. Our result has shown that annual average electric energy consumption will be 10,628 million kWh per year during 2019-2030 which approximately 3.3 percent growth per annum. At the moment, there is not a practice solution for the storage of electricity in Mongolia. Therefore, energy supply and demand have to be balanced in real-time for operational stability. Without an accurate forecast, the end-users may experience brownouts or even blackouts or the industry could be faced with sudden accidents due to the energy demand. For this reason, energy consumption forecasting is essential to power system stability and reliability.

Keywords: energy forecasting; energy consumption; ARIMA model; Box-Jenkins method.

Introduction

The primary product of the power system is electric energy. Electric energy is the main factor for production, manufacturing and socio-economic activities. Growing technological and social developments are the main influencing factor for high energy consumption. For this reason, the article involves a time series analysis of the annual energy consumption of Mongolia by using ARIMA modeling. The methodology adopted Box-Jenkins method which is an application of autoregressive integrated moving average (ARIMA) modeling for describing discrete data where continuous observations are correlated(Oppenheim 1978). From mathematical point of view, employing the appropriate forecast could improve the decision to be made(Morales et al. 2014). The Box-Jenkins method ARIMA technique has become considerable recognition in recent years in energy-related forecasting(Albayrak 2010; Hor, Watson, and Majithia 2006; Lai et al. 2014; Li, Han, and Yan 2018; Nichiforov et al. 2017). Electric energy consumption forecasting is not only used for defining short or long-term energy supply but it helps for energy system planning and power system expansion. At the same time, energy demand and supply have to be balanced. At the moment, there is not any scientific research on Mongolian energy consumption forecasting. Hence, there are many researchers hypothesis assumed by the previous year's energy consumption.

In this article, we applied a univariate model for forecasting the future annual electricity consumption as a function of prior annual electricity consumption. ARIMA technique is well known for predicting economic variables and widely used for predicting future values of the power industry, health sector and economic activity and many others. Such as, some researchers(Etuk, Eleki, and Sibeate 2016) have established the ARIMA model by extending seasonal autoregressive integrated moving average (SARIMA) for the monthly natural gas production, the result showed some social constraints, as well as infrastructural inadequacies and economic depression, make limitation for production nevertheless its abundant reserve. The same method has used for forecasting Turkish energy demand between 2005-2020(Ediger and Akar 2007), where researchers concluded that the result of ARIMA is more reliable for future energy prediction. Few researchers(Jiang, Yang, and Li 2018) adopted ARIMA model by integrating several similar forecasting models as well as metabolic grey model (MGM) and back-propagation neural network (BP) for more precise prediction of energy usage. Some scholars(Mohamed and Bodger 2004) have proposed six different forecasting models for electricity consumption. According to the result, ARIMA model ranked as the best forecasting technique in short term prediction. Forecasting energy consumption has a vital role in market stability(Jiang et al. 2018) and reliability.

Methodology

Data source

The historical data regarding electric energy consumption during 1990-2018 are collected from the National Statistical Office of Mongolia(National Statistical Office). Considering the current circumstance of energy consumption of Mongolia over 93 percent(World Bank) generated by fossil fuels. Figure-1 shows the trend of energy consumption based on the collected data which shows that energy consumption has an increasing trend in the last 29 years except for early social transition years (1991-1996) of Mongolia.

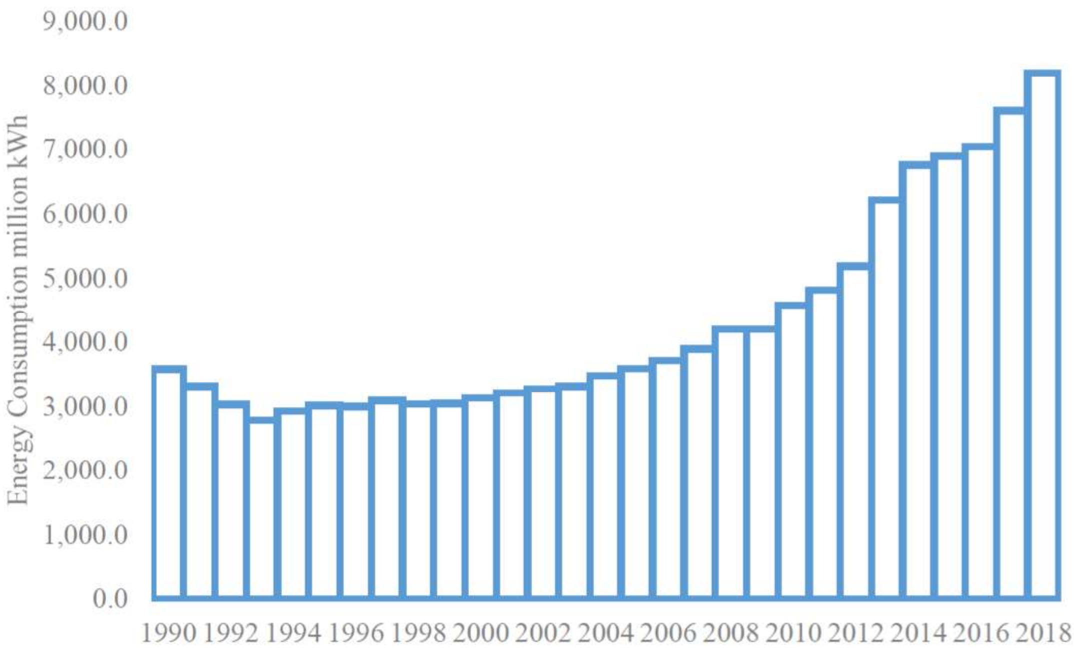

Fig. 1 Annual energy consumption variation

Model structure

ARIMA short for the "Auto-Regressive Integrated Moving Average" model was developed by Box and Jenkins in 1970 which is a time series analysis method based on the theory of random. The forecasting time series divided into two forms. The first is the *univariate time series forecasting* as we have adopted in analyzing which generally uses previous values of the time series to predict its future. The last form called *multivariate time series forecasting*. The ARIMA model has three basic types: Auto-Regressive model (hereinafter AR (p)), the moving average model (hereinafter MA (q)) and auto-regressive moving average model (ARMA (p,q).

Forecasting on the ARIMA model mainly includes four steps: First, the stationary test for an original sequence which means it uses its own lags as predictors because the ARIMA is based on a linear regression model. In most cases, the predictors are not correlated and are independent of each other, as a result need to make it a series stationary by differencing it. Accordingly, need to be a difference by subtracting the previous value from the current value. The value of d, therefore, is the minimum number of differencing needs to make the series stationary where d=0. Second, after time-series stationarity d=0, the parameters p and q will be determined. The parameter p is the order of the AR term. It refers to the number of lags of Y to be used as predictor. The actual mathematical formula for AR model is:

$$Y_t=\alpha+\beta_1 Y_{t-1}+\beta_2 Y_{t-2}+\ldots+\beta_p Y_{t-p}+\varepsilon_t \quad (1)$$

where, Y_t is a function of the lags of Y_t, which means Y_{t-1} is the lag 1 of the series, β_1 is the coefficient of the lag 1 that the model estimates and α is the intercept term.

Next, the parameter q is the order of the MA term. It refers to the number of lagged forecast errors that should go into the ARIMA model. Here the Y_t is depended only on lagged forecast errors. The mathematical formula is:

$$Y_t=\alpha+\varepsilon_t+\phi_1\varepsilon_{t-1}+\phi_2\varepsilon_{t-2}+\ldots+\phi_q\varepsilon_{t-q} \quad (2)$$

where the error terms are the errors of the autoregressive models of the respective lags. The errors ε_t and ε_{t-1} are the errors from the following equation:

$$Y_t=\beta_1 Y_{t-1}+\beta_2 Y_{t-2}+\ldots+\beta_0 Y_0+\varepsilon_t \quad (3)$$

$$Y_{t-1}=\beta_1 Y_{t-2}+\beta_2 Y_{t-3}+\ldots+\beta_0 Y_0+\varepsilon_{t-1} \quad (4)$$

That is our AR and MA model respectively. Third, the estimation of the unknown parameters in the model and examination of the rationality of the model. An ARIMA model where the time series was differenced at least once to make it stationary and by combining the AR and the MA terms. The general equation form becomes:

$$Y_t=\alpha+\beta_1 Y_{t-1}+\beta_2 Y_{t-2}+\ldots+\beta_p Y_{t-p}+\phi_1\varepsilon_{t-1}+\phi_2\varepsilon_{t-2}+\ldots+\phi_q\varepsilon_{t-q} \quad (5)$$

In our ARIMA model predicted Y_t =constant + linear combination of lags of Y + linear combination of lagged forecast errors. The final step is diagnostic analysis to confirm that the obtained model is consistent with the observed data characteristics. Box and Jenkins (Box, et al.) were stated that these do not involve independent variables, but rather make use of the information in the series itself to generate forecasts. Therefore, ARIMA models depend on autocorrelation patterns in the series. The predominant contribution of Box and Jenkins were to grant a common strategy in which three levels of the model building (which are model identification, estimation, and diagnostic checks) had been given prominence(Hipel, McLeod, and Lennox 1977).

Empirical result

The research objective of this paper is the energy consumption forecasting for improving the energy planning of Mongolia. The energy consumption forecast depends on historical data where the dataset belongs to univariate prediction. The energy consumption in the next 11 years forecasting range relies on 2019-2030. Historical data at annual intervals over 28 years period was adopted the Box-Jenkins method iterative procedure to obtain the appropriate model. In the ARIMA modeling, each of the data set is considered independently for stationarity, identification, estimation, and diagnostic checking residuals(Mohamed and Bodger 2004) as we have mentioned earlier. In the first step we need to find the order of differencing (d) in ARIMA model. The aim of differencing is to make the time series stationary. We applied Augmented Dickey-Fuller (hereinafter ADF) test to inspect the existence of unit root in energy consumption. ADF test is the common method of unit root.

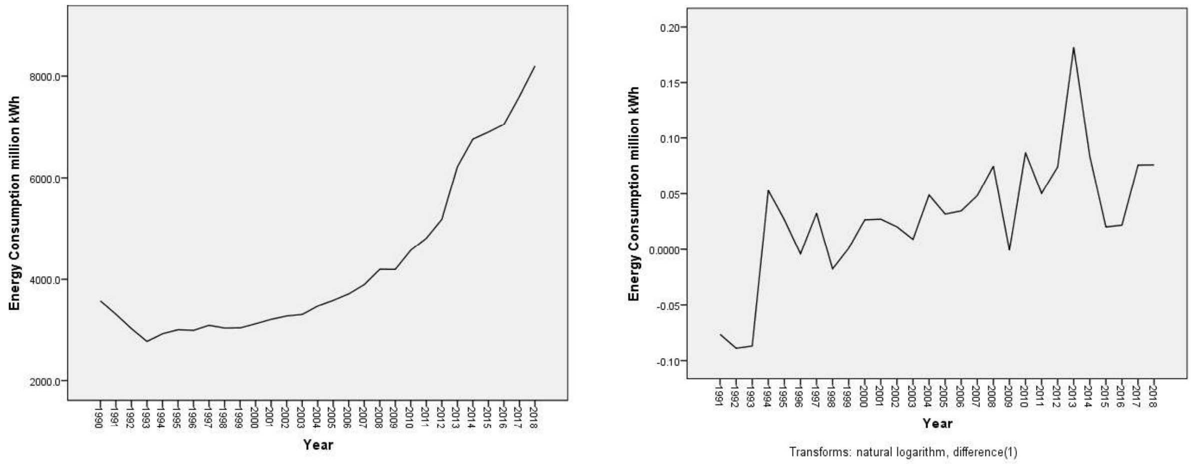

Fig. 2 Unit root level versus Unit root in 1st difference

According to the correlogram in Figure 2, stationary resulted in the first-order difference which is stable. Even though the series is not perfectly stationary but we fix the order of differencing as 1. Afterward, we made model identification which leads to the time series of energy consumption. In this step, we tested correlation coefficients for a stationary sequence. Figure 3 shows autocorrelation and partial-autocorrelation based on stationary sequence difference.

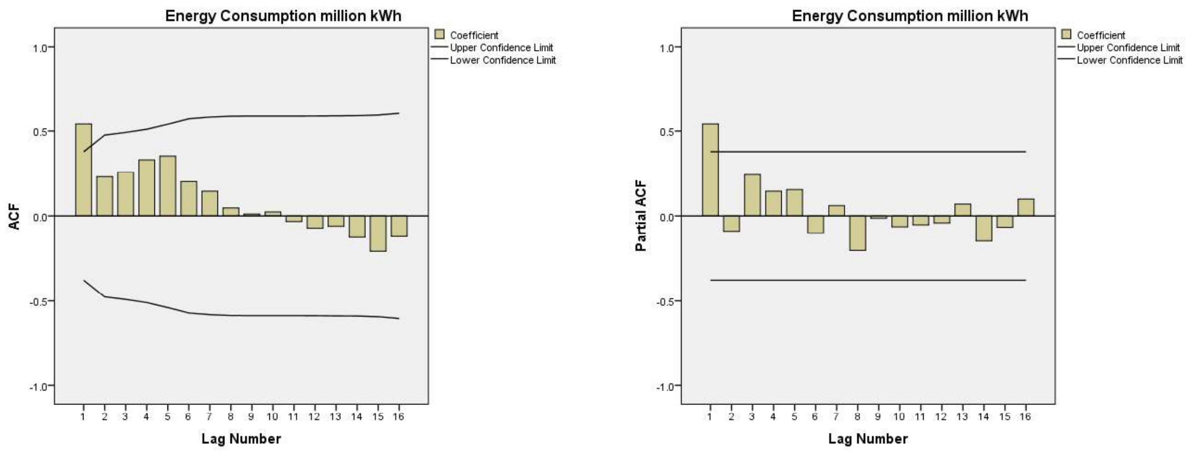

Fig. 3 ACF and PACF coefficients

Partial autocorrelation could be imagined as the correlation between the series and its lag, after excluding the contributions from the intermediate lags. Partial autocorrelation of lag (k) of a series is the coefficient of that lag in the autoregression equation of Y. The mathematical formula of PACF is:

$$Y_t=\alpha_0+\alpha_1 Y_{t-1}+\alpha_2 Y_{t-2}+\alpha_3 Y_{t-3} \tag{6}$$

In our case, suppose, if Y_t is the current series and Y_{t-1} is the lag 1 of Y, which means lag 1 Y_{t-1} is the coefficient in the above equation. According to Figure 3, we can observe that PACF lag 1 is quite significant, thus, we take p as 1. As in figure 3, we can see the ACF plot for the number of MA terms which is technical, the error of the lagged forecast. The ACF shows how many MA terms are required to remove any autocorrelation in the stationarized series. According to the coefficient result criteria of the ARIMA model, ARIMA (1,1,1) is preferable to be used for the forecasting where the process takes formulation as:

$$(1-\beta_1 Y)(1-Y)^d Y_t = (1-\phi_1 Y)\, \varepsilon_t \quad (7)$$

Right after, we have to determine the fit for the model. Figure 4 and Table 1 show that our determination of R^2 is 0.980, which means the fitting effect is good.

Table 1 Parameters of the fit for the ARIMA (1,1,1) model

Model Description

			Model Type
Model ID	EnergyConsumption million kWh	Model_1	ARIMA(1,1,1)

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile 5	10	25	50	75	90	95
Stationary R-squared	.326	.	.326	.326	.326	.326	.326	.326	.326	.326	.326
R-squared	.980	.	.980	.980	.980	.980	.980	.980	.980	.980	.980
RMSE	237.779	.	237.779	237.779	237.779	237.779	237.779	237.779	237.779	237.779	237.779
MAPE	3.815	.	3.815	3.815	3.815	3.815	3.815	3.815	3.815	3.815	3.815
MaxAPE	12.916	.	12.916	12.916	12.916	12.916	12.916	12.916	12.916	12.916	12.916
MAE	162.999	.	162.999	162.999	162.999	162.999	162.999	162.999	162.999	162.999	162.999
MaxAE	727.115	.	727.115	727.115	727.115	727.115	727.115	727.115	727.115	727.115	727.115
Normalized BIC	11.300	.	11.300	11.300	11.300	11.300	11.300	11.300	11.300	11.300	11.300

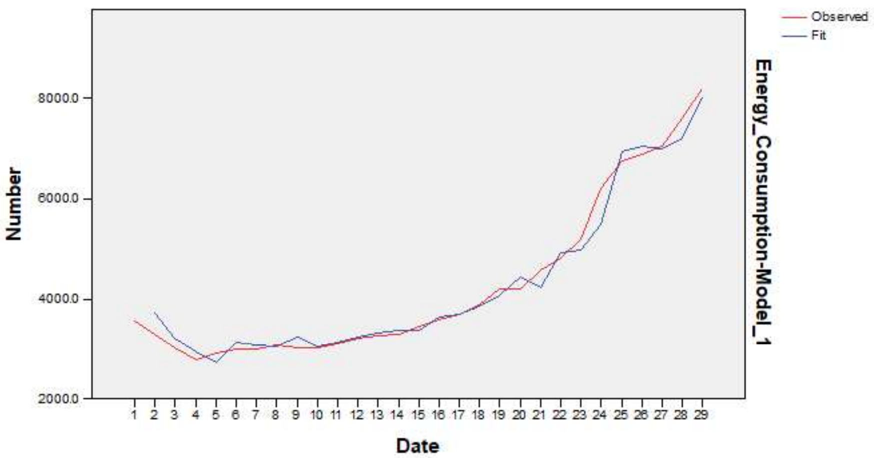

Fig. 4 Model calibration using historical data

The last step of the ARIMA(1,1,1) model is forecasting energy consumption from 2019-2030.

Table 2 ARIMA (1,1,1) model for energy consumption forecasting

ARIMA Model Parameters

					Estimate	SE	t	Sig.
EnergyConsumption million kWh-Model_1	EnergyConsumption million kWh	No Transformation	AR	Lag 1	.958	.089	10.799	.000
			Difference		1			
			MA	Lag 1	.587	.228	2.574	.016

Forecast

Model		2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
EnergyConsumption million kWh-Model_1	Forecast	8635.0	9050.7	9448.8	9830.2	10195.5	10545.4	10880.5	11201.5	11509.0	11803.5	12085.6	12355.8
	UCL	9101.9	9843.0	10579.1	11316.4	12054.9	12793.2	13530.2	14264.6	14995.2	15721.3	16442.1	17156.9
	LCL	8168.1	8258.4	8318.5	8344.0	8336.1	8297.5	8230.8	8138.5	8022.8	7885.7	7729.2	7554.8

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.

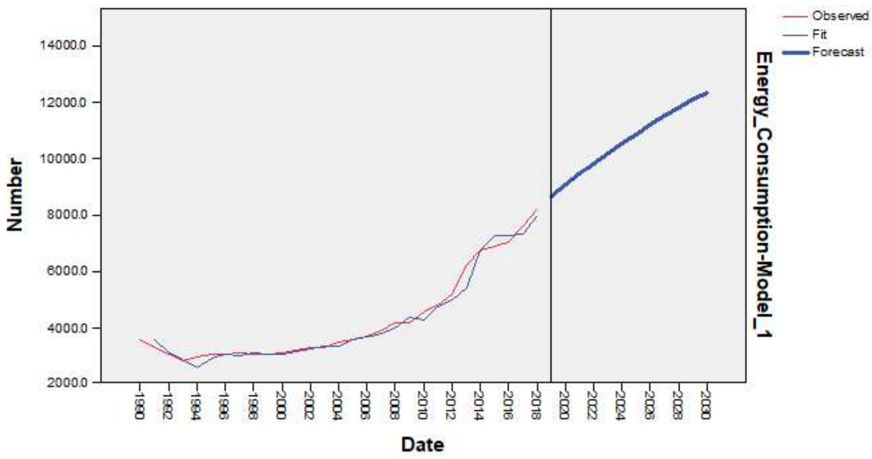

Fig. 5 Historical and simulated energy consumption

From Figure 4-5, the ARIMA(1,1,1) model seems to give a directionally correct forecast. And the actual observed values lie within the 95. The model of total energy consumption is a better overall fit that could be used to predict Mongolia's energy consumption in the next eleven years.

Conclusion

According to the stationarity analysis ARIMA (1,1,1) model has established in this paper for forecasting the energy consumption of Mongolia from 2019-2030. Since the ARIMA (1,1,1) model of total energy consumption is better overall fit. As looking back to figure 3, ARIMA (1,1,1) has high fitting precision and certain stability. As a result, the model adopted in this paper has a higher fitting degree. The result indicates that at the end of 2030 energy consumption predicted more than 12000 million kWh. The annual average growth rate of energy usage 3.3 percent. Moreover, the growth rate of Mongolia's energy consumption during 1991-1996 were declined due to the social structure transition. Currently, 20 percent of electric energy has been importing from Russian Federation and China(Sovacool, D'Agostino, and Bambawale 2011). Increasing energy consumption will result in shortage in the future as well and the imported energy could be increased. Thus, in our point of view Mongolia's energy strategy should integrate allocating more resources such as wind, solar and hydro generated energies for the development of energy technologies and stability of energy supply. And need to establish competitive energy market conditions through liberalization. Forecasting energy consumption has a vital role in energy planning and design to the national energy agency and policy-makers. Within the purpose of balancing energy independence and effective energy policies which promote industrial structure optimization and development the new industry with high technology. The forecasting technique gives more information for maintaining market stability and safety. Although coal reserves are high, this will be exhausting, so we need to find an optimal energy system. Year by the year huge coal reserves has been exhausting, as mitigating environmental impacts of fossil fuel, Mongolia need integrate renewable resources into the energy system. According to the statistics only 5.6% of renewable resources have been supplied by the renewables. Energy consumption forecasting could be done the basis of energy system planning and expansion. Based on this research, we will study system dynamic analysis for promoting renewable energy resource expansion and its integration into existing energy system.

Bibliography

1. Albayrak, Ali Sait. 2010. "ARIMA Forecasting of Primary Energy Production and Consumption in Turkey: 1923–2006." *Enerji, Piyasa ve Düzenleme* 1:24–50.
2. Box, George E. P., M. Ljung Greta, Gwilym M. Jenkins, and Gregory C. Reinsel,. n.d. *Time Series Analysis: Forecasting and Control, 5th Edition*.
3. Ediger, Volkan Ş. and Sertaç Akar. 2007. "ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey." *Energy Policy* 35(3):1701–8.
4. Etuk, Ette Harrison, Alapuye Gbolu Eleki, and Pius Sibeate. 2016. "A Box-Jenkins Model for Monthly Natural Gas Production in Nigeria." *Journal of Multidisciplinary Engineering Science Studies* 2(11):6.
5. Hipel, Keith William, Angus Ian McLeod, and William C. Lennox. 1977. "Advances in Box-Jenkins Modeling: 1. Model Construction." *Water Resources Research* 13(3):567–75.
6. Hor, C. L., S. J. Watson, and Shanti Majithia. 2006. "Daily Load Forecasting and Maximum Demand Estimation Using ARIMA and GARCH." *2006 International Conference on Probabilistic Methods Applied to Power Systems* 1–6.
7. Jiang, Feng, Xue Yang, and Shuyu Li. 2018. "Comparison of Forecasting India's Energy Demand Using an MGM, ARIMA Model, MGM-ARIMA Model, and BP Neural Network Model." *Sustainability* 10(7):2225.
8. Lai, Sue Ling, Ming Liu, Kuo Cheng Kuo, and Ray Chang. 2014. "Energy Consumption Forecasting in Hong Kong Using ARIMA and Artificial Neural Networks Models." *Applied Mechanics and Materials*.
9. Li, Yiyan, Dong Han, and Zheng Yan. 2018. "Long-Term System Load Forecasting Based on Data-Driven Linear Clustering Method." *Journal of Modern Power Systems and Clean Energy* 6(2):306–16.
10. Mohamed, Z. and P. S. Bodger. 2004. "Forecasting Electricity Consumption: A Comparison of Models for New Zealand." in *Engineering: Conference Contributions*.
11. Morales, Juan M., Antonio J. Conejo, Henrik Madsen, Pierre Pinson, and Marco Zugno. 2014. *Integrating Renewables in Electricity Markets: Operational Problems*. Springer US.
12. National Statistical Office. n.d. "Mongolian Statistical Information Service." *Mongolian Statistical Information Service*. Retrieved July 8, 2019 (http://1212.mn/).
13. Nichiforov, C., I. Stamatescu, I. Făgărăşan, and G. Stamatescu. 2017. "Energy Consumption Forecasting Using ARIMA and Neural Network Models." Pp. 1–4 in *2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE)*.
14. Oppenheim, Rosa. 1978. "Forecasting via the Box-Jenkins Method." *Journal of Academy of Management Science* 6(No3):206–21.
15. Sovacool, Benjamin K., Anthony L. D'Agostino, and Malavika Jain Bambawale. 2011. "Gers Gone Wired: Lessons from the Renewable Energy and Rural Electricity Access Project (REAP) in Mongolia." *Energy for Sustainable Development* 15(1):32–40.
16. World Bank. n.d. "World Bank Open Data | Data." *World Bank Open Data*. Retrieved July 8, 2019 (https://data.worldbank.org/).

Gansukh Zolboo - ORCID: 0000-0003-2787-7462

Gansukh Zolboo – currently 4th year PhD student at School of Economics and Management, Yanshan University, PR China, Mongolian citizen. Main research focusing on renewable energy transmission into the Northeast Asian super grid.

Bor Adiya - ORCID: 0000-0002-6301-6902

Bor Adiya – PhD in Engineering at East Siberia State University of Technology and Management, Senior expert of Mongolia, Mongolian.

Enkhbayar Bilguun - ORCID: 0000-0003-1387-7648

Enkhbayar Bilguun – foreign Student in PRC at School of Information Science and Engineering at Yanshan University. 3rd year PhD student, research area generally focused optics and electric engineering.