Commun.Fac.Sci.Univ.Ank.Series A2-A3
Volume 61, Number 1, Pages 68-75 (2019)
DOI: 10.33769/aupse.525368
ISSN 1303-6009 E-ISSN 2618-6462
http://communications.science.ankara.edu.tr/index.php?series=A2A3

COMMUNICATIONS
SERIES A2-A3

ESTIMATING CO2 EMISSIONS BY USING ENERGY INTENSITY DATA OF OECD COUNTRIES

Recep ERYİĞİT and Semra GÜNDÜÇ

ABSTRACT. It is discussed that economic development has an essential effect on the country's CO2 emission which plays an important role in global warming. In this research well-known machine learning algorithm Extreme Learning Machine, ELM, is used to investigate the relationship between CO2 emission and energy intensity for countries in OECD. The results indicate a strong correlation and the method perform well for estimation.

1. INTRODUCTION

The rising amount of Greenhouse gases in the atmosphere contribute one of the most critical problems that we have today, global warming. Even though there are many sources of greenhouse gases, CO_2 is heading with a contribution of more than 60% in total [1]. For this reason, to concern global warming it is essential to reduce CO_2 emission. Since pre-industrial revolution period, it is clear that CO_2 emission has increased intensively [2], and correspondingly result in an increase in the global temperature approximately 1.3 degrees [3].

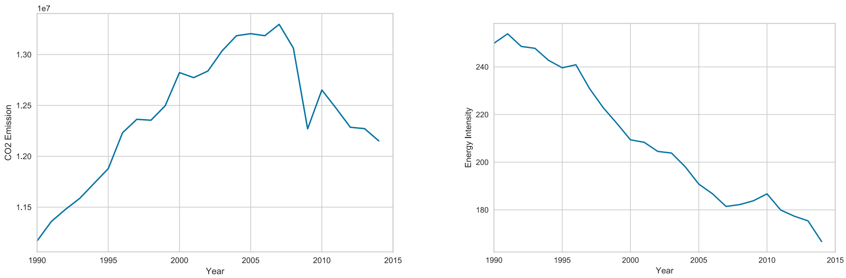
The population of the world is also increasing. As a result, energy demand (mainly produced from fossil sources), pollution, usage of natural sources and deforestation (essential for the absorption of CO_2), is increasing. It is clearly discussed in [4] that human activities have many effects on the environment.

Also, the development in the economy for many countries is rising. Besides, Gross Domestic Product, (GDP), mostly used as a metric to measure this development, is increasing. For this reason the studies which aim the relation of many factors, mostly economic, with CO_2 emission has become popular in the last decades. In [6] it is supposed that economic output is the main driving factor for CO_2 emission. The

Received by the editors: February 11, 2018; Accepted: May 24, 2019.
Key word and phrases: Extreme Learning Machine (ELM), Energy Intensity, CO2 emission.

work [7] says that about half of the increase in CO_2 emission is due to economic growth.

The total amount of CO_2 emission of the countries in Organisation for Economic Co-operation and Development (OECD) in years from 1990 to 2015 is introduced in Figure 1a. The data is taken from the World Data Bank [8]. The effect of strategies which are taken by policymakers to decrease the CO2 emission is seen in the figure since 2005.



(a) (b)

Figure 1. (a)Total CO2 emission values (b) Total Energy Intensity values, of OECD Countries in the years between 1990 and 2015.

Both increasing energy usage/demand and economic development are considered as main factors affecting the CO2 emission in the literature. Besides, in many countries, there is an enormous effort to decrease the damage to the natural environment. For this reason, to use energy as efficiently as possible a new metric can be considered to measure the amount of unit energy per unit of GDP which defined as energy intensity. High energy intensity value means the high cost for converting this energy to GDP. Countries in (OECD) have been decreased energy intensity considerably in the years between 1990 and 2015 which is an indication of a great effort to reduce the CO_2 emission. Figure 1b shows the total energy intensity of OECD countries as a function of years.

All in all such efforts are needed. The cause of emission mostly depend on national sources of countries and therefore shows a great variety in different geographical regions. Also, some costly and long-term investment is required. For this reason, OECD, EU, even governments are supposed to construct radical policies related to the reduction of CO_2 emission.

For many reasons, such as advising to policymakers, for future planning, etc., in the literature, there is an increasing amount of work which aims to find the effects and also the relation between some parameters and CO_2 emissions, which are mainly nonlinear. However, the lack of driving mechanisms of such a process forces to use soft programming techniques to reach a significant result. Also increasing power and advances in technology allow to generate and collect data at an incredible rate. This situation results in having an extensive data set in size and dimension. This truth has an important effect to build new and more efficient computational methods than ever.

To that end in this work to investigate the CO_2 emission by considering the energy intensity of OECD countries, machine learning algorithm ELM is used. In the next section, the details of the method and the results obtained by using this method are introduced.

2. Model

In this paper, it is aimed to find out the relationship between energy intensity and the amount of CO_2 emission. In the literature conventional neural network algorithm is frequently used in many research for estimation and building a correlation between parameters of the systems [9]. Because of updating all parameters at every iteration in the neural network algorithm, the time used for calculation is very high. Also in many cases, there is no analytical approach to the driving mechanism of the process. For this reason, soft computing techniques introduce a unique way to get the solution. Therefore to find out the relation between the energy intensity and the amount of CO2 emission, Extreme Learning Machine (ELM) [8] algorithm, based on Single-Layer Neural Networks, is used in this work.

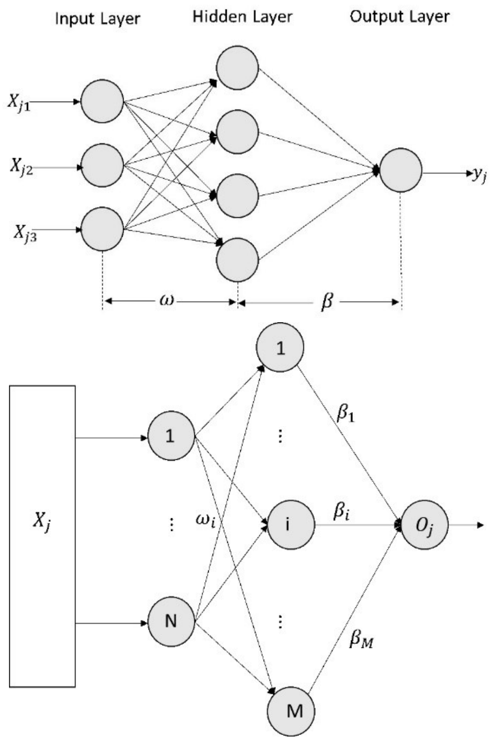


Figure 2. A schematics representation of ELM.

ELM algorithm was first introduced in [8] for training Single-Layer Feed forward Neural Networks (SLFN) and since used in many researches [10,11] because of its success and fast learning speed. A schematic representation is given in Figure 2.

For N arbitrary distinct samples (x_i ,y_i) the output of an SLFN which have M hidden neurons can be mathematically modelled as,

$$f_M(x) = \sum_{i=1}^{M} \beta_i \, g_i(x_j) = \sum_{i=1}^{M} \beta_i \, g(w_i \, . x_j + b_i) = O_j, \qquad j = 1, \dots, N$$

Here b_i is the threshold of the ith hidden nodes, $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector between ith hidden node and input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$ is the weight vector between output and hidden nodes.

In its original article [8], these equations are written in an compact form as

$$\boldsymbol{H}\beta = \boldsymbol{T}$$

where

$$\boldsymbol{H} = \begin{bmatrix} g(w_1\boldsymbol{x_1} + b_1) & \cdots & g(w_M\boldsymbol{x_1} + b_M) \\ \vdots & \cdots & \vdots \\ g(w_1\boldsymbol{x_N} + b_1) & \cdots & g(w_M\boldsymbol{x_N} + b_M) \end{bmatrix}_{N*M}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{M*m} \quad \text{and} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_M^T \end{bmatrix}_{N*m}$$

$$\beta = \mathbf{H}^{\dagger}\boldsymbol{T}$$

H is called the hidden layer output matrix of the neural network.

In the calculations to understand how well future samples are likely to be predicted by the model, the coefficient of determination is used which provides a measure. Having n different samples, for the ith sample, if the true value of the variable is y_i and its predicted value is, the coefficient of determination is defined as,

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}$$

here $\overline{y}$ is

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

The value of the coefficient of determination between 0 and 1 means that the independent variable can be predicted, 0 means cannot be predicted and 1 means can be predicted without error from the dependent variable.
For each estimation, we also calculated the mean absolute error, MAE, over n sample, is defined in Equation below, which is a measure of the average of the absolute errors

$$MAE(y,\widehat{y}) = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

The results are introduced in the next section.

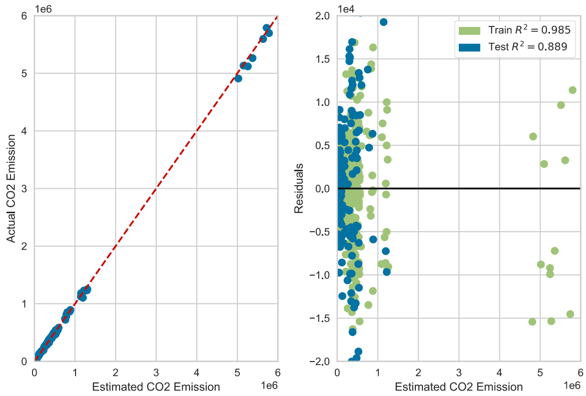
3. Results and discussion

In this work, ELM algorithm is used to estimate the amount of CO2 emission from energy intensity values for OECD countries. The energy intensity value is used as an input parameter and amount of CO2 emission is estimated as output. The results including the coefficient of determinations and MAE are introduced in Table 1.

Table 1. Coefficient of determination and Mean Absolute Error for ELM.

Method	R^2 Test	R^2 Train	Mean Absolute Error
ELM	0.8891	0.9847	8947

As it is seen, the value of coefficient of determination indicates that energy intensity, input value, is a suitable variable to estimate the CO_2 emission, output value, and ELM algorithm perform well for this estimation. Fig.3 shows the graphical representation of prediction of CO2 emission by considering the energy intensity under ELM algorithm. In the graphic, the left figure shows the prediction and the right one shows the residuals. The residuals, difference between the observed value and the estimated value, are also considered in this work to have a sense of how accurate our model is by relating the amount of CO2 emission with energy intensity. The countries in OECD have two main behaviors about CO2 emission which is seen clearly in Figure 3a that some have low-level emission, some have high but no middle. For this reason, residual plot is affected this having lack of some data region but still give good results.



(a) (b)

Figure 3. (a) Estimation of CO_2 emission by using ELM and (b) Residuals of estimation.

References

[1] I. Ozturk and A. Caravci, CO2 emissions, energy consumption and economic growth in Turkey. Renew Sustain Energy Rev 2010;14:3220–5.

[2] BP. BP statistical review of world energy June 2016; 2016. http://www.bp.com/statisticalreview.

[3] SJ. Davis, K. Caldeira and HD. Matthews, Future CO2 emissions and climate change from existing energy infrastructure. Science 2010;329:1330–3.

[4] P.R. Ehrlich and J.P. Holdren, Impact of population growth. Science 1971, 3977, 1212–1217.

[5] A. Shi, The impact of population pressure on global carbon dioxide emissions, 1975–1996: Evidence from pooled cross-country data. Ecol. Econ. 2003, 1, 29–42.

[6] M. Wang and C. Feng, Decomposition of energy-related CO2 emissions in China: an empirical analysis based on provincial panel data of three sectors. Appl Energy 2017;190:772–87.

[7] B. Lin and H. Liu, CO2 emissions of China's commercial and residential buildings: Evidence and reduction policy. Build Environ 2015;92:418–31.

[8] www.worlddatabank.com

[9] J. Long, L. Shuai, H. Bin and L. Mei, A survey on projection neural networks and their applications Applied Soft Computing, Volume 76, 2019, Pages 533-544

[10] G.B. Huang, Q.Y. Zhu and C.K. Siew, Extreme learning machine:Theory and applications, Neurocomputing 70 (2006a) 489501

[11] S. Gang and D. Qun, A novel double deep ELMs ensemble system for time series forecasting , Knowledge-Based Systems, Volume 134, 2017, Pages 31-49.

[12] K. Marius, Y. Yang, L. Caihong, C. Yanhua and L. Lian Mixed kernel based extreme learning machine for electric load forecasting Neurocomputing, Volume 312, 2018, Pages 90-106.

Current Address: RECEP ERYİĞİT: Ankara University, Computer Engineering Department, Gölbaşı Ankara Turkey
E-mail: eryigit@eng.ankara.edu.tr
Orcid ID: https://orcid.org/0000-0002-4282-6340
Current Address: SEMRA GÜNDÜÇ: Ankara University, Computer Engineering Department, Gölbaşı Ankara Turkey
E-mail: gunduc@.ankara.edu.tr
Orcid ID: https://orcid.org/0000-0002-3811-9547