The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2024

Volume 27, Pages 99-107

IConTech 2024: International Conference on Technology

Towards Net-Zero Emissions in OECD Countries: Forecasting AI by Machine Learning Methods

Serdar Celik
Ostim Technical University

Abstract: Achieving net-zero emissions is a paramount objective for Organisation for Economic Co-operation and Development (OECD) countries in combating climate change and fostering sustainable development. This study provides an overview of the strategies, opportunities, and challenges facing OECD countries in their transition towards achieving net-zero emissions. This study delineates the OECD's commitment to ambitious climate targets, including the overarching goal of achieving carbon neutrality by 2040. While some existing models provide reasonably accurate predictions of CO_2 emissions, the model presented in this study offer improved prediction capabilities for all OECD countries. This study successfully predicts historical emissions, current emissions, and future emissions from 1990 to 2022 using Machine Learning (ML) methodology. The study forecasts global CO_2 emissions for all OECD countries from 2022 to 2042 (near future) using prediction models using SARIMA (Seasonal Autoregressive Integrated Moving Average) based on ARIMA (Autoregressive Integrated Moving Average). The primary aim is to compare these models and identify the most effective one for predicting the transition to net-zero emissions for all OECD countries. These predictions highlight that policymakers should thoroughly evaluate the measures and strategies to promote a transition to net-zero emissions and reduce the levels of CO_2 emissions. Furthermore, it highlights the potential co-benefits of transitioning to a low-carbon economy, including improved air quality, enhanced energy security, and job creation.

Keywords: Environmental sustainability, Carbon emissions, Net-zero emissions, Machine learning (ML),

Introduction

One of the 21st century's biggest problems is global warming and related climate change, which is accepted by scientific circles to be human-induced. Climate change has a high potential to cause major environmental disasters, hunger, war, and thirst in the future. Therefore, it is very important for the whole world to analyze and plan the factors that trigger climate change. One of the most important factors that trigger change is greenhouse gases(Zhong & Haigh, 2013). This study aims to estimate CO_2 emissions from greenhouse gases that cause climate change. The assessment and estimation of CO_2 emissions is important for global warming (Florides & Christodoulides, 2009) and climate change (Solomon et al., 2009). The estimations are performed for OECD countries. The OECD is an international organization of 38 countries that brings together countries with market economies and democratic structures for economic cooperation. Its members are Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States. OECD countries generally have very large economies and require high amounts of energy due to their high living standards. According to 2022 data, OECD countries have 17% of the world population and 59% of GDP. At the same time, 54% of the world's total production is made by this group of countries Figure 1. Therefore, it is of vital importance to examine high CO_2 emitting countries such as the OECD. The CO_2 estimates obtained at the end of the study can be used in climate scenarios and can be used in policy-making

- This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

© 2024 Published by ISRES Publishing: www.isres.org

decisions and in the implementation of global climate protocols. Estimation of CO_2 emissions is important to quantify and tackle irreversible climate change (Solomon et al., 2009).

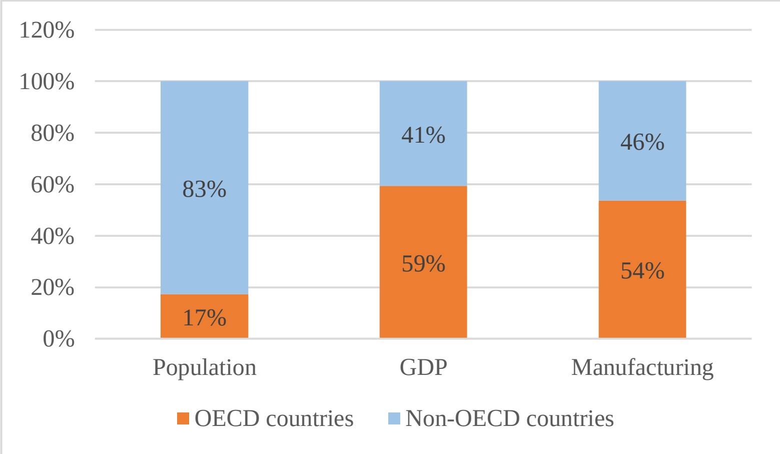

Figure 1. World and OECD economic outlook

Artificial intelligence studies are now used in classification, association and prediction for many different types of problems and produce successful results. Machine learning is a subset of artificial intelligence systems that make predictions by learning time-dependent historical data. In the literature, there are classical time series methods and statistical methods that estimate the CO_2 emissions of countries. Statistical time series methods are complex and model the change of data over time. While statistical models are mathematically complex and use many extra parameters, machine learning methods make more understandable predictions from the information they learn. In this study, the future CO_2 emissions of 38 countries were predicted by using machine learning method, which makes predictions by learning time-dependent historical data.

In this study, CO_2 emissions of 38 countries between 1990 and 2022 are used. 80% of the data is used for training and the remaining 20% is used for testing. Evaluation metrics are used when the prediction results need to be compared. Performance metrics are based on how close the predicted value in the test data is to the true value. The performance of the method is compared with statistical functions such as Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE) and Mean Square Error (MSE).

In the literature, many studies have been conducted for CO_2 emission estimation. To date, there are some studies such as (Iania et al., 2022) that model the global CO_2 emission footprint. Most of the existing studies have a regional context, for example (Nyoni & Bonga, 2019) used the Box-Jenkins ARIMA approach on time series data on CO_2 emission in India from 1960 to 2017 and proposed five policy solutions to improve environmental conditions based on the forecasts. (Qian et al., 2020) studied CO_2 emission forecasting of Changxing county in China, while (Tanania et al., 2020) analyzed and predicted CO_2 emissions of India using dataset from 1995 to 2018. (Zuo et al., 2020), collected CO_2 emission data from different provinces of China and proposed an integrated model, LSTM-STRIPAT, to predict emissions in 2020. (Shabani et al., 2021) estimated local CO_2 emission in Iran using machine learning and neural network-based modeling. Furthermore, (Rolnick et al., 2023), provides a good overview of CO_2 emission and related issues but lacks in building a CO_2 emission model for a global situation. A number of modeling approaches have been attempted by various authors from different perspectives to estimate CO_2 emissions. (Kadam & Vijayumar, 2018) provides insight into a CO_2 emission estimation model using machine learning. A new hybrid model using combined principal component analysis (PCA) is built for China based on data from 1978 to 2014 (Sun & Sun, 2017). In addition, a forecasting model for CO_2 emissions in China based on multiple linear regression analysis is also studied (Libao et al., 2017).

The aim of this study is to model accurate CO_2 emission behavior for the past, present and near future. First, this study uses current data together with historical data. Second, it uses existing time series-based machine learning models to develop the CO_2 emission prediction model. The model selection process is different in that it obtains the best model and relevant parameters. The model developed based on the selected parameters becomes accurate (less prone to error). As a result, near-future forecasts become accurate when compared to the actual CO_2 emissions of that time. Initially, models based on the available time series were selected; the ML model optimization algorithm was developed, which selects the best model for each CO_2 data set in order to obtain the best possible forecast. Finally, this study estimates the CO_2 emission footprint for 20 years (from 2023 to 2042) as an example.

Method

This paper uses data on CO_2 emissions of OECD countries for the period 1990-2022. Primary data are taken from this repository (Ritchie & Roser, 2024). The SARIMA model was used to develop forecasting models for the estimation of CO_2 emission behavior. The dependent variable in this study is the total amount of CO_2 emissions and the independent variable is the year. In this study, the data is divided into training data and test data; 80% of the CO_2 emission data is used for learning and 20% is used for testing. Training data were used in the CO_2 emission forecasting process of the model and test data were used to determine the accuracy of CO_2 emission forecasting. Data from December 1990 to December 2020 were considered as training data and data from December 2021 to December 2022 were used as test data for the model. A similar training-test separation was performed for each OECD country. Validation and comparison were then performed to evaluate our forecasting models and model results.

Machine Learning

Machine learning is based on creating algorithms based on data on a specific topic, updating the outputs as new data becomes available, and using statistical analysis to predict the results. There are many different types of algorithms under machine learning that enable computers to learn. Computers extract a model from the learned data using various functions and statistical methods and can predict, predict or classify new data according to this model. There are many different machine learning methods in the literature. These methods perform differently depending on the type of data. Therefore, it is difficult to say that one machine learning method is superior to others. There are three basic stages in machine learning. These are:

- Preparation of data: The first stage is the preparation of the right data, and the data must be prepared with great care in order to reach the right results. At this stage, the data is made ready for processing by finding outliers, normalizing the data, etc. Depending on the type of problems, both numerical and symbolic (nominal) data can be processed.
- Training: The second phase is to find the most appropriate model and train the data prepared in the first phase with this model. First, a machine learning method suitable for the problem is selected and the most appropriate model is created from the data by training. In order to find the appropriate model, as many models as possible should be built and tested. A part of the data in the training set is separated and the model is validated and it is decided whether the model is suitable or not.
- Test: The last stage is the performance test. For this, the model is tested using the model created with data other than the training data used to create the model. This data is called the test set. During the testing process, the performance of machine learning with the test set is measured with metrics such as accuracy rate, number of false positives, number of true positives.

Machine learning is used for classification, regression and clustering, but it is also frequently used for time-series analysis. Machine learning and time series try to make new predictions based on information from the past. This is a type of supervised learning in machine learning. In this study, time-series analysis was performed using SARIMA, a machine learning method known to give successful results.

ARIMA is an acronym that stands for autoregressive integrated moving average. Specifically, it is a model that uses the dependent relationship between an observation and a set of lagged observations, AR (autoregression), I (integrated), the use of differencing of raw observations and MA (moving average) to make the time series stationary. SARIMA, an extension of ARIMA, is known for its ability to incorporate seasonality into forecasts. SARIMA is a seasonal autoregressive integrated moving average with seasonality as well as trend. These algorithms are robust in handling both stationary and non-stationary time series data. Forecasting with ARIMA or SARIMA usually involves three main steps: specifying the temporal model, estimating the parameters and diagnostic checking. A standard notation is used for ARIMA (p, d, q), where parameters are replaced by integer values to quickly indicate the specific ARIMA model being used. Where p is the number of lagged observations included in the model, also called the lag order, d is the number by which the raw observations differ, also called the degree of differencing, and q is the size of the moving average window, SARIMA is represented as SARIMA (p, d, q) $(P, D, Q)^S$, where "P", "D" and "Q" correspond to the seasonal autoregressive, seasonal difference and seasonal moving average terms respectively. "S" denotes the seasonal parameter. The researcher found that S=19 gives the best seasonality parameter during the ARIMA model optimization. As a result, the ARIMA model was transformed into SARIMA model and presented as SARIMAX. To evaluate the model, we use prediction metrics such as mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE), and mean absolute deviation (MAD). To generate different models, the ARIMA algorithm was executed repeatedly using the optimization algorithm developed by the author.

Performance Metrics for Forecasts

After the predictions are made, it is necessary to quantify the prediction success of machine learning methods on the data set. Performance metrics have been developed for these purposes. Prediction metrics such as mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE) and mean absolute deviation (MAD) are used to evaluate the model (Sammut & Webb, 2010). MAPE and MSE are the most commonly used functions for prediction performance evaluation. However, the presence of two other evaluation criteria in the study will enable future studies to compare the results with the results obtained. (Lewis, 1982) evaluated predictions with MAPE values below 10% as "very good". For the sake of simplicity and completeness, MAPE scores are presented in this article to compare with model accuracy. To generate different models, the ARIMA algorithm was executed repeatedly using the optimization algorithm developed by the author. After checking efficiency issues, the most efficient model was used.

$$MAPE = \frac{\sum_{t=1}^{n} |(y_t - \hat{y}_t)/y_t|}{n} 100 \tag{3.3}$$

$$MSE = \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n} \tag{3.4}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n}} \tag{3.5}$$

$$MAD = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n} \tag{3.6}$$

Having successfully developed the model, it was time to create visual representations of the modeling results. Figure 2 is sufficient as evidence to validate the models presented in this paper. Figure 2 shows the forecast behavior of emissions for current (1990-2022) and future (2022 and beyond) years. The pink shades in Figure 2 are confidence intervals (upper and lower bounds of the forecasts). Furthermore, for a better understanding of the forecast values, Table 1 presents the modeling error and accuracy parameters found during model development.

Results and Discussion

A series of significant findings emerged from the forecast of near and far-future CO_2 emissions utilizing an exact model. Initial validation of the model was performed by comparing its predictions against observed data for the near future. Subsequently, projections were made for CO_2 emissions in the far future, yielding results summarized in Table 1. The model's efficacy in predicting near-future CO_2 emissions was assessed, providing empirical validation for its reliability. Notably, the model generated predictions for specific years in the near future, offering valuable insights that may bolster the development of models for forecasting distant future emissions. The forecasting models show a decreasing trend in CO_2 emissions over time and this trend is consistent with historical data showing seasonal variations. This observation is in line with previous forecasts and proves the reliability of the established models. Analysis of historical data, as shown in Figure 3, reveals a decrease in CO_2 emission rates in 2023, mirroring similar fluctuations in other years characterized by both increases and decreases. In contrast, the methodology used in this study exhibits remarkable precision in determining estimated CO_2emission values.

As a result, the developed models are less sensitive to errors. This agreement between the estimated CO_2 emission data and the accuracy metrics indicates a convergence, increasing confidence in the reliability of both the methodology and the estimated CO_2 emissions. Therefore, the methodological approach and the resulting CO_2 emission estimates deserve to be considered as accurate and reliable tools for future use.

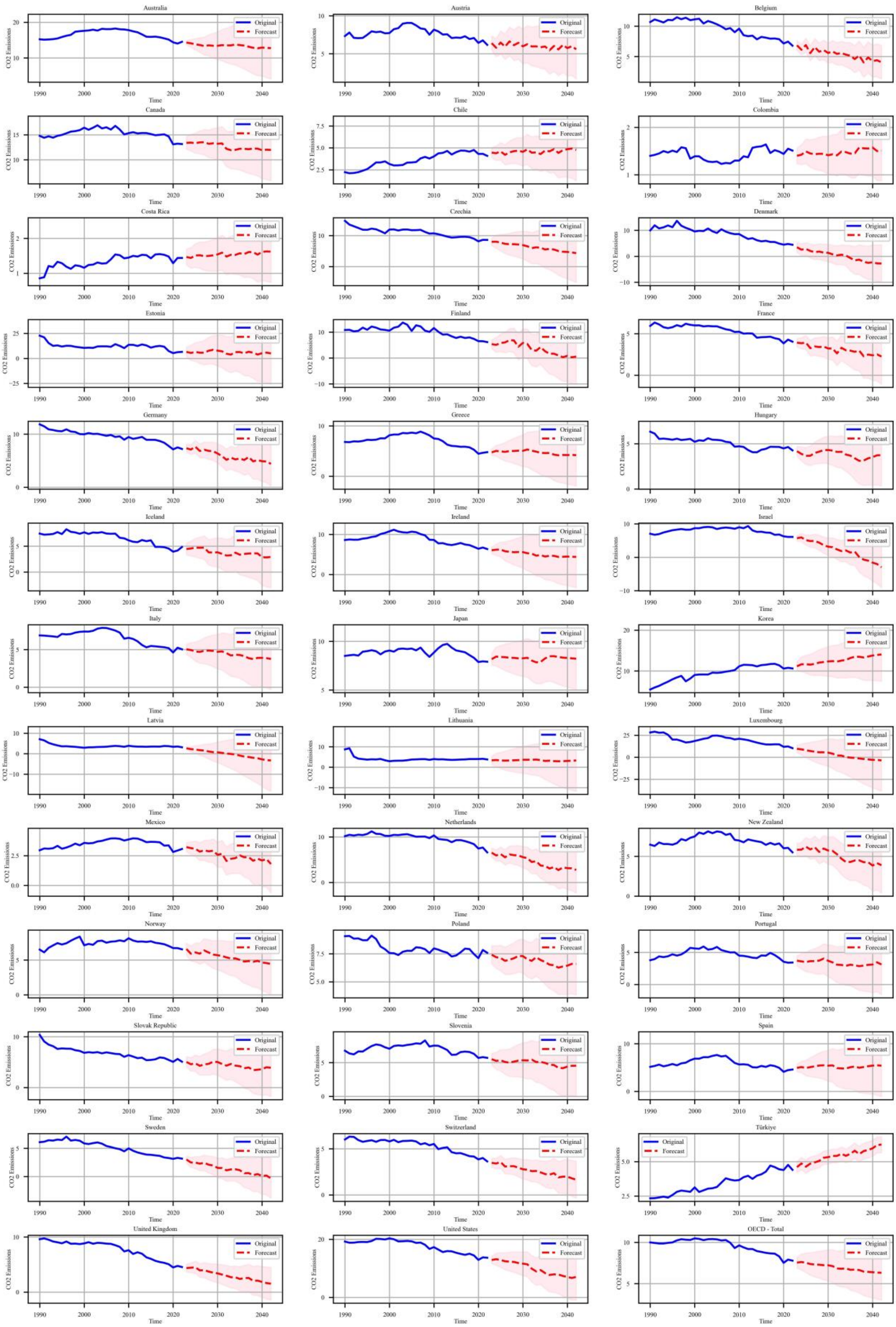

Figure 2. Emissions estimates for OECD countries

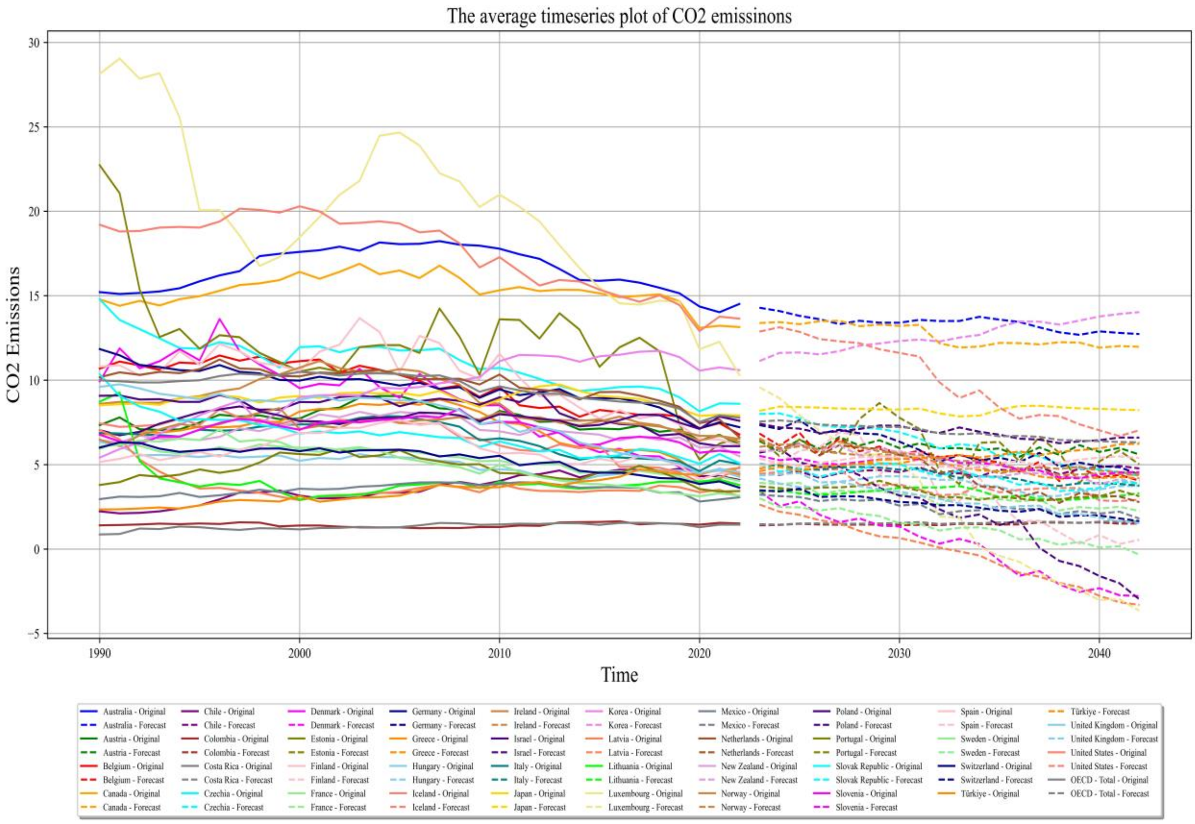

Figure 3. OECD countries emission estimates for 2023-2042

The forecasting model shows that CO_2 emissions have been decreasing over time except for some countries. This phenomenon can be seen in all time periods except for some countries as shown in Figure 3 of the 20-year forecast values of OECD countries. Minimizing CO_2 emissions should be considered by every nation. Looking at the OECD countries collectively, it is seen that the amount of CO_2 emissions has decreased. The situation is different for Turkey, Costa Rica and South Korea. In these three countries, the amount of CO_2 emissions is generally increasing until 2019 and it is seen in the study that this increase will continue in the coming years if CO_2 emissions are not directed to non-alternative energy sources.

All models in this article were developed in Python. The evaluation criterion values calculated by all models in this article are presented in Table 1. Taking into account seasonality and trends in the data, the SARIMA model provides a forecast consistent with the downward trend seen. The SARIMA model highlights the importance of considering seasonality in the forecasting context. Looking at the table results, a downward trend is predicted in the next twenty years compared to today, indicating that countries will undergo structural changes. The projected decrease is relatively significant and does not decrease for every country. This model implies that OECD countries will reach net zero emissions in the near future.

Conclusion

Carbon emissions, the greenhouse effect, climate change and devastating environmental problems have become the most important issues of today's world. If the increase in the amount of carbon dioxide in OECD countries and the world is not stopped or reduced, seasonal shifts, very high temperatures, floods and floods due to heavy rainfall will continue to increase. The underlying reason behind the increase in CO_2 emissions is the world's ever-increasing demand for energy. Apart from fossil fuels such as oil, natural gas and coal, the main sources of energy are nuclear energy, hydroelectric energy and sustainable energy. With the industrial revolution and the overuse of fossil fuels, CO_2 emissions, one of the most important greenhouse gases that cause climate change, have doubled in the last century. In addition to global warming, which is a very big problem, this increase in CO_2 emissions may cause much bigger problems such as water wars and migrations in the coming years. For these reasons, countries should make various plans to reduce their CO_2 emissions. Among these plans, countries should turn to renewable energy sources instead of using fossil fuels to reduce CO_2 emissions.

Table 1. OECD CO2 emission outlook to 2042

Countries/Years	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	MAPE	RMSE	MSE	MAD
Australia	14.29	14.09	13.81	13.61	13.32	13.51	13.40	13.40	13.57	13.50	13.50	13.75	13.58	13.44	13.12	12.83	12.68	12.89	12.80	12.73	18.96%	3.40	11.58	3.22
Austria	6.39	5.79	6.47	5.94	6.63	6.13	6.44	5.97	6.25	5.96	5.97	5.86	5.95	5.46	6.02	5.57	6.15	5.75	5.98	5.59	20.70%	1.79	3.21	1.64
Belgium	6.84	6.12	6.88	5.55	6.55	5.79	6.06	5.58	5.82	5.44	5.56	5.31	5.13	4.58	5.16	4.02	4.85	4.25	4.39	4.05	38.05%	3.36	11.29	3.31
Canada	13.38	13.43	13.31	13.48	13.51	13.19	13.29	13.21	13.27	12.18	11.93	11.99	12.20	12.19	12.11	12.24	12.22	11.92	12.01	11.97	16.72%	2.68	7.18	2.58
Chile	4.46	4.39	4.60	4.21	4.49	4.63	4.68	4.55	4.83	4.45	4.48	4.29	4.70	4.61	4.85	4.43	4.73	4.85	4.92	4.77	14.17%	0.66	0.44	0.52
Colombia	1.40	1.42	1.49	1.45	1.44	1.44	1.44	1.41	1.43	1.41	1.49	1.48	1.45	1.45	1.58	1.56	1.55	1.58	1.50	1.51	8.67%	0.13	0.02	0.12
Costa Rica	1.47	1.44	1.51	1.52	1.49	1.50	1.51	1.54	1.59	1.48	1.53	1.53	1.57	1.55	1.61	1.61	1.54	1.61	1.63	1.62	7.36%	0.13	0.02	0.10
Czechia	8.00	8.03	7.74	7.33	7.20	7.21	7.11	6.92	6.55	5.95	6.21	6.09	5.50	5.71	5.48	5.02	4.78	4.72	4.55	4.33	39.42%	3.99	15.91	3.97
Denmark	3.44	2.54	2.82	2.03	1.59	1.81	1.39	1.34	0.73	0.32	0.60	0.29	-0.69	-1.59	-1.31	-2.10	-2.54	-2.32	-2.74	-2.79	105.97%	7.05	49.70	7.02
Estonia	6.50	5.50	6.49	5.75	5.74	7.69	8.64	7.77	7.11	4.67	3.85	6.29	6.34	5.20	6.64	5.69	3.90	5.19	5.83	5.01	43.17%	5.79	33.55	5.20
Finland	5.35	5.07	5.80	6.02	6.90	6.76	4.35	5.83	5.45	3.09	2.83	4.22	2.60	1.66	1.67	0.98	0.32	0.83	0.29	0.55	65.82%	6.02	36.19	5.91
France	3.92	3.85	3.93	3.25	3.45	3.39	3.44	3.22	3.21	2.63	3.29	2.99	2.95	2.82	2.95	2.28	2.48	2.39	2.50	2.25	38.43%	1.91	3.65	1.89
Germany	7.34	7.10	7.63	6.83	7.06	6.85	6.79	6.36	5.86	5.16	5.57	5.25	5.41	5.14	5.69	4.88	5.12	4.90	4.85	4.41	33.91%	3.06	9.35	3.00
Greece	4.75	5.04	4.93	4.85	5.00	5.03	5.08	5.06	5.32	5.08	4.89	4.66	4.60	4.62	4.30	4.16	4.24	4.21	4.25	4.19	28.58%	2.43	5.92	2.12
Hungary	4.18	3.88	3.66	3.68	3.99	4.13	4.30	4.31	4.25	4.10	4.11	3.81	3.69	3.40	3.12	3.14	3.39	3.52	3.72	3.74	19.25%	1.09	1.18	0.94
Iceland	4.45	4.56	4.72	4.69	4.72	3.81	3.77	3.82	3.52	3.19	3.24	3.75	3.35	3.50	3.58	3.62	3.60	2.88	2.83	2.92	36.40%	2.33	5.44	2.22
Ireland	6.05	6.18	6.21	5.79	5.67	5.56	5.60	5.58	5.31	5.20	4.70	4.85	4.48	4.66	4.65	4.38	4.41	4.47	4.42	4.38	37.66%	3.29	10.85	3.18
Israel	5.70	5.94	5.13	4.85	4.87	4.37	3.62	3.16	3.03	2.35	1.84	2.05	1.42	1.69	0.07	-0.71	-1.01	-1.60	-2.02	-2.98	77.67%	6.09	37.04	5.83
Italy	5.03	4.93	4.74	4.69	4.83	4.86	4.82	4.70	4.75	4.52	4.21	4.31	4.28	4.12	3.85	3.73	3.89	3.90	3.85	3.75	28.18%	1.94	3.77	1.81
Japan	8.18	8.44	8.40	8.37	8.32	8.30	8.26	8.26	8.32	8.03	7.85	7.90	8.25	8.47	8.48	8.40	8.33	8.30	8.26	8.22	7.99%	0.88	0.78	0.73
Korea	11.13	11.61	11.64	11.52	11.72	12.04	12.17	12.32	12.40	12.29	12.52	12.68	13.16	13.47	13.46	13.29	13.52	13.77	13.91	14.03	18.43%	2.06	4.23	1.94
Latvia	2.63	2.21	2.02	1.71	1.52	1.07	0.75	0.65	0.39	0.08	-0.15	-0.38	-0.93	-1.41	-1.64	-2.01	-2.25	-2.77	-3.15	-3.31	107.40%	4.15	17.24	3.74
Lithuania	3.38	3.50	3.29	3.22	3.27	3.39	3.46	3.63	3.64	3.64	3.72	3.46	3.06	3.18	2.97	2.90	2.94	3.07	3.14	3.30	13.01%	0.60	0.36	0.50
Luxembourg	9.57	8.93	7.93	7.02	6.32	5.60	5.58	5.22	4.07	1.92	1.86	0.12	-0.40	-0.77	-1.44	-2.04	-2.44	-3.02	-2.98	-3.66	92.20%	15.80	249.70	15.76
Mexico	3.22	3.13	3.07	2.84	2.94	2.86	2.91	2.57	2.66	2.05	2.24	2.29	2.52	2.36	2.38	2.07	2.24	2.10	2.21	1.80	30.36%	1.15	1.32	1.10
Netherlands	6.55	5.94	6.25	5.53	6.15	6.00	5.93	5.57	5.34	4.50	4.67	3.72	3.63	3.07	3.38	2.73	3.31	3.16	3.11	2.77	51.52%	4.76	22.69	4.71
New Zealand	5.82	5.83	6.21	5.85	6.13	5.54	5.97	5.72	5.68	5.25	4.68	4.29	4.28	4.47	4.48	4.25	4.25	3.83	4.14	3.87	28.36%	2.01	4.05	1.97
Norway	6.55	5.90	6.08	5.90	6.36	6.06	5.78	5.67	5.56	5.37	5.20	5.22	5.01	4.77	4.84	4.74	4.87	4.64	4.55	4.43	27.99%	2.12	4.50	2.08
Poland	7.42	7.20	7.12	6.87	6.98	7.10	7.31	7.29	7.02	6.87	7.20	6.95	6.77	6.53	6.48	6.26	6.37	6.44	6.60	6.59	10.79%	0.92	0.85	0.84
Portugal	3.71	3.61	3.50	3.64	3.65	3.73	4.04	3.70	3.45	3.09	3.05	2.91	3.09	2.98	2.86	2.95	3.07	3.11	3.46	3.12	26.55%	1.39	1.94	1.27
SlovakRepublic	5.02	4.58	4.76	4.27	4.55	4.66	5.08	5.02	4.64	4.31	4.63	4.32	4.20	3.80	3.94	3.42	3.50	3.61	3.99	3.88	28.03%	1.75	3.05	1.69
Slovenia	5.51	5.26	5.35	5.10	5.00	5.15	5.30	5.35	5.32	5.45	5.09	5.11	4.97	4.70	4.65	4.24	4.18	4.41	4.53	4.55	28.16%	2.07	4.27	1.99
Spain	4.93	5.11	4.91	5.10	5.26	5.43	5.48	5.40	5.47	4.88	4.83	4.79	5.08	5.21	4.96	5.12	5.29	5.41	5.48	5.39	15.76%	1.31	1.72	0.99
Sweden	3.00	2.48	2.47	2.28	2.42	2.09	1.95	1.52	1.49	1.09	1.26	1.28	1.11	0.57	0.60	0.24	0.43	0.09	0.16	-0.34	72.61%	3.00	9.00	2.99
Switzerland	3.48	3.41	3.57	3.08	3.11	3.12	2.95	2.78	2.73	2.59	2.60	2.45	2.27	2.20	2.36	1.91	1.98	1.98	1.81	1.65	47.68%	2.33	5.43	2.32
Türkiye	4.60	4.86	4.63	4.85	4.94	5.05	5.28	5.32	5.41	5.41	5.57	5.42	5.56	5.79	5.54	5.76	5.85	5.95	6.20	6.23	38.42%	1.49	2.23	1.47
UnitedKingdom	4.35	4.43	4.50	3.94	3.98	3.72	3.53	3.40	3.18	2.97	2.70	2.69	2.39	2.54	2.56	2.06	2.08	1.85	1.64	1.52	56.60%	3.85	14.82	3.79
United States	12.87	13.13	12.86	12.43	12.30	12.19	11.83	11.61	11.39	9.88	8.94	9.39	8.37	7.72	7.94	7.87	7.38	7.03	6.69	7.03	39.78%	6.40	40.98	6.37
OECD - Total	7.55	7.62	7.52	7.37	7.31	7.29	7.23	7.22	7.12	6.84	6.80	6.84	6.67	6.69	6.64	6.46	6.42	6.38	6.34	6.32	24.23%	2.30	5.30	2.25

Accurate near real-time forecasting of CO_2 emissions of OECD countries is of great importance for governments' efforts to control and mitigate climate change. This paper focuses on evaluating the forecasting capabilities of the SARIMAX machine learning model for near real-time annual CO_2 emission forecasting of OECD countries. The study uses a series of real-time datasets of OECD countries covering the period from 1990 to 2022. Four evaluation criteria, namely MSE, RMSE, MAD and MAPE, are analyzed to identify the most appropriate model for future forecasts and to compare their forecasting performance. By building the ML model that takes into account the reduction of CO_2 emissions, some remarkable results are obtained that can help to understand CO_2 emissions among OECD countries. The study shows that the SARIMA (1, 1, 0, 19) model is not only stable but also the most appropriate model for forecasting the annual total CO_2 of OECD countries for the next 20 years. The model predicts that the total annual CO_2 emissions of OECD countries will average 9,310 t CO_2 by 2042. Whatever the estimate we obtain, it can practically reflect the actual CO_2 emission. This is an important warning signal, especially with regard to climate change and global warming. The results of this study are very important for the governments of OECD countries, especially when it comes to medium and long-term planning. Since global warming affects the world on a large scale, the study results can be applied to every country. The contribution of the study is to save human lives by controlling CO_2 emission.

Recommendations

This study presents a CO_2 emission prediction model based on a wealth of valuable information compiled from past data. The findings of this research enrich our understanding of CO_2 emission dynamics, aiding in the formulation of reduction policies and effectively monitoring emission trends. The ability to predict CO_2 emissions plays a crucial role in shaping reduction policies and evaluating their effectiveness.

The methodological framework and resulting solutions represent significant contributions to relevant research areas. Exciting opportunities for future research in CO_2 emission tracking are evident, with proposed policy recommendations for OECD countries including emission taxes, carbon taxes, and incentives for sustainable energy. These proposals stand out as effective strategies for reducing greenhouse gas emissions. While laying a valuable foundation for future research on CO_2 emissions, this study also provides practical guidance to policymakers and industries, contributing to the advancement of a sustainable future.

Scientific Ethics Declaration

The author declares that the scientific ethical and legal responsibility of this article published in EPSTEM Journal belongs to the author.

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Technology (www.icontechno.net) held in Alanya/Turkey on May 02-05, 2024.

References

Bai, L., Wang, J., Ma, X., & Lu, H. (2018). Air pollution forecasts: An overview. *International Journal of Environmental Research and Public Health*, *15*(4), 780.

Iania, L., Algieri, B., & Leccadito, A. (2022). Forecasting total energy's CO2 emissions. *Available at SSRN 4124979*.

Kadam, P., & Vijayumar, S. (2018). Prediction model: CO2 emission using machine learning. In 2018 *3rd International Conference for Convergence in Technology (I2CT)* (pp.1–3). IEEE.

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth Scientific.

Libao, Y., Tingting, Y., Jielian, Z., Guicai, L., Yanfen, L., & Xiaoqian, M. (2017). Prediction of CO 2 emissions based on multiple linear regression analysis. *Energy Procedia*, *105*, 4222–4228.

Nyoni, T., & Bonga, W. G. (2019). *Prediction of CO2 emissions in India using ARIMA Models* (SSRN scholarly paper 3346378). Retrieved from https://papers.ssrn.com/abstract=3346378

Qian, Y., Sun, L., Qiu, Q., Tang, L., Shang, X., & Lu, C. (2020). Analysis of CO2 drivers and emissions forecast in a typical ındustry-oriented county: Changxing county, China. *Energies*, *13*(5), 1212.

Ritchie, H., & Roser, M. (2024). CO_2 emissions dataset: Our sources and methods. *Our World in Data*. Retrieved from https://ourworldindata.org/co2-dataset-sources

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., … Bengio, Y. (2023). Tackling climate change with machine learning. *ACM Computing Surveys*, *55*(2), 1–96.

Sammut, C., & Webb, G. I. (Eds.). (2010). Mean squared error. In *Encyclopedia of machine learning* (pp. 653–653). Springer.

Shabani, E., Hayati, B., Pishbahar, E., Ghorbani, M. A., & Ghahremanzadeh, M. (2021). A novel approach to predict CO2 emission in the agriculture sector of Iran based on ınclusive multiple model. *Journal of Cleaner Production*, *279*, 123708.

Solomon, S., Plattner, G.-K., Knutti, R., & Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, *106*(6), 1704–1709.

Sun, W., & Sun, J. (2017). Prediction of carbon dioxide emissions based on principal component analysis with regularized extreme learning machine: The case of China. *Environmental Engineering Research*, *22*(3), 302–311.

Tanania, V., Shukla, S., & Singh, S. (2020). Time series data analysis and prediction of CO2 emissions. *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 665–669.

Zhong, W., & Haigh, J. D. (2013). The greenhouse effect and carbon dioxide. *Weather*, *68*(4), 100–105.

Zuo, Z., Guo, H., & Cheng, J. (2020). An LSTM-STRIPAT model analysis of China's 2030 CO_2 emissions peak. *Carbon Management*, *11*(6), 577–592.

Author Information

Serdar Celik
Ostim Technical University, Ankara, Turkey
Contact e-mail: *serdar.celik@ostimteknik.edu.tr*

To cite this article:

Celik, S. (2024). Towards net-zero emissions in OECD countries: Forecasting AI by machine learning methods. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 27,* 99-107.