Vehicular CO_2 Emission Forecasting using Time Series Analysis

Pushpander Kadian[1], Savita Ahlawat[2,*], Amit Choudhary[3], Supriya Choudhary[4], Sukriti Sinha[2]

[1]Govt Post Graduate Nehru College, Jhajjar, Haryana, India

[2]Maharaja Surajmal Institute of Technology, New Delhi, India

[3]University School of Automation and Robotics, GGSIPU, New Delhi, India

[4]Maharaja Surajmal Institute, New Delhi, India

E-mail: savita.ahlawat@gmail.com

Abstract - Global warming is posing serious problems for the world, and greenhouse gas emissions are a key contributor. In the year 2017, energy providers contributed 46% of worldwide Emissions of CO_2, showing great opportunity for a decrease. Greenhouse gases, particularly carbon dioxide (CO_2) production, are one of the primary causes of global warming, making it one of the world's biggest environmental challenges. For interim planning for the national aim to reduce Carbon footprints, more accuracy in short-term forecasts is essential. To make rational choices, rising climate change conventions require accurate predictions of participating countries' future emission growth paths. The most frequent way of forecasting is trend analysis, which is the act of formulating forecasts based on previous and present data. Forecasting models are becoming increasingly crucial in showing complicated correlations between large amounts of inaccurate information and unpredictable variables. The proposed work uses the Auto-Regressive Integrated Moving Average (ARIMA) model and Holt-Winters Exponential Smoothing (HWES) forecast model to predict carbon footprints in countries on an annual basis. This work intends to estimate time series data on emissions of CO_2 in countries throughout the world using quantitative tools. The researcher's grasp of CO_2 emissions projections will be boosted by this work. Furthermore, the findings of this study can be used by government bodies to establish strategic plans.

Keywords:- CO2 Emission, Global Warming, Forecasting, Time Series

1. Introduction

CO_2 emissions are a major contributor to climate change, and scientists agree that the average global surface temperature should not rise over 1.5 degrees Celsius higher than pre-industrial values. Gulf countries, for example, have long dominated the oil and gas business. As a result, forecasting CO_2 emissions in these oil-producing countries is essential. National governments will be able to change their climate policy with the help of such predictions. The purpose of the present work is to study global CO2 emissions from the year 1950 to 2020. The nations that emit 2% or more of global CO_2 emissions, as well as those that emit less than 2% of global CO_2 emissions in the atmosphere, is considered in the present work. The International Energy Agency (IEA) provided the data for this analysis, which estimates CO_2 emissions from the burning of coal, natural gas, oil, and other fuels, as well as toxic waste and non-renewable municipality trash. Future forecasts, on other hand, will be required to persuade developed economies to provide more financial help to emerging countries, in order to persuade the latter to compromise some of their economic objectives. Figure 1 represents the role of sectors like energy, waste, transport, industry, agriculture etc. in CO2 emission. It is evident from Figure 1 that transportation contributes a major part to CO2 emissions after the energy sector. Therefore, the goals of the present work are to determine which industries produce the most CO_2; to compare the effectiveness of all forecasting models to choose the best one; to predict future emissions of CO_2 using time series and the right model; to help in policy-making keeping the data availability in mind.

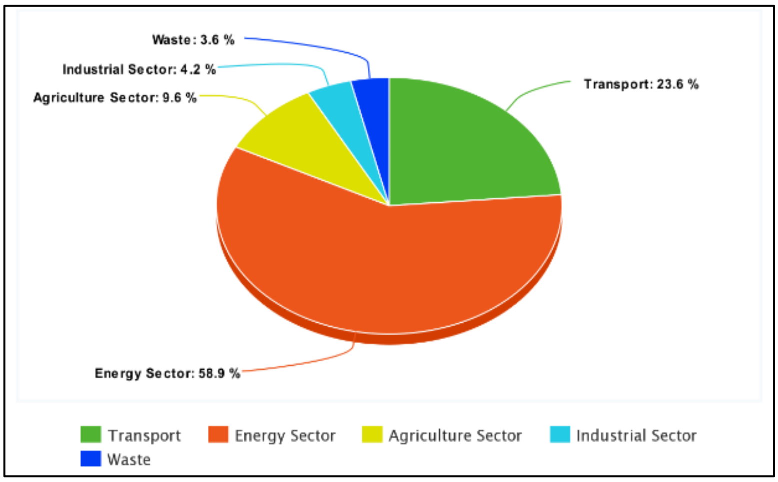


Fig 1: The contribution of transportation to India's CO_2 emissions

Forecasting models are critical tools used in a variety of scientific disciplines [23-24]. ANN models, which are highly complex and non-linear and can effectively estimate future emission values, are used in solving many real time based problem. A multilayer ANN model was used in past to forecast CO_2 emissions for high and low-producing countries. Several recent studies have also combined autoregressive integrated moving average (ARIMA), Holt-Winters exponential smoothing, and ANN approaches to estimate CO_2 emissions. ARIMA and other forecasting modelling techniques were used to evaluate, analyze, and evaluate CO_2 emissions projections based on time-series data [1].

The paper is organised as follows: Section 2 describes the techniques and methods used by researchers in the prediction of CO2 emission. Section 3 focuses on the theoretical background. Section 4 describes the proposed model working and discusses the findings of the model. In last, section 5 presents the conclusion and suggestions for future work.

2. Related Work

The researchers used a number of prediction models to estimate CO_2 emissions and other regions in their study [2]. In modelling linear and non-linear behaviours in the data set, the suggested hybrid model is different in that it combines the advantages of ARIMA and ANNs. Furthermore, the computational experience shows that the new combining model is more effective at creating more accurate predictions than previous tactics [3]. A hybrid technique that comprises both ARIMA and ANN models has been used in linear and non-linear modelling. Furthermore, the results suggest that combining the algorithms can improve the forecast efficiency of each model independently [4]. Forecast-TB is a R tool that examines the accuracy of several forecasting techniques in terms of time-series dataset properties. Forecast-TB tool also showed a basic time-series dataset that could be used to evaluate forecasting comparison analysis as a function of data set characteristics [5]. In addition, for the short-term prediction of Carbon dioxide emissions from energy use in 5 European countries, two proposed time series decomposition algorithms are presented [6]. Furthermore, a mixed approach based on artificial neural networks (ANN) and an agent-based framework for calculating carbon dioxide emissions from various power sources in Annaba have been built using real-world data. The report is based on gas and electricity statistics from Algeria's national energy corporation [7]. Grote et al. [8] utilised data from the Inductive Loop Detectors (ILD) of the Urban Transport Corporation (UTC). Emission models were developed using road traffic data. The data from car sensors, such as speed, speed, fuel flow, and mileage is also used in proposed models to analyse CO2 emission. The values were taken in 30-sec timeframe using the OBD connection. The forecast was based on an integrated model that combines the OBD and ILD models. The OBD model is used to set the settings for the ILD model. According to the findings, ILD data combined with vehicle type classification through an OBD model can be utilised to model Emissions of CO_2 in real-world traffic. In another work, Belaidi et al. used the ARIMA model, a statistical approach to forecast and assess Emissions of CO_2 in Algeria from the year 1963 to 2019 [9]. The information presented here is a time series of pollution produced by CO_2 emissions. Our method consists of four main steps: model identification, parameter estimate, validation, and ultimately, forecasting certain future values. The first phase is to identify the model; the second is to estimate the parameters; the third is to validate the model; and finally, the fourth step is to anticipate certain future values. The authors use theoretical results with the R programme. The ARIMA (1, 1, 2) model is known to be successful at forecasting future CO_2 emissions. Quast in [10] provides a deep learning architecture based on LSM (LSTM). LSTM is a variation of the Recurrent Neural Network (RNN) that remembers prior inputs as well as the current example, allowing it to efficiently exploit sequential correlations in data. The model does not require rigorous topic modelling and extraction because deep learning methods produce features on their own. To imitate real-life sensor disturbances, the model was tested on noisy data. It is possible to estimate non-trivial values like as CO_2 emissions, transmissions and drive train conditions, driver driving behaviour, unsafe driving, vehicle response time, and so on when vehicle-tracking data is combined with appropriate machine learning algorithms [11]. Such vehicle insights allow for accurate real-time vehicle monitoring systems. Because the tracking data is in the form of a time-series, traditional machine-learning techniques cannot be used directly. Traditional machine learning techniques that only train on the current case and have no memory of past outcomes are unable to detect sequential correlations, making them unsuitable for time-series forecasting. Zeng et al. investigated how the vehicle's fuel consumption was affected by the distance travelled [12]. Deep learning algorithms were used to predict fuel usage for some fixed routes. The researchers only used data collected at 60-second intervals to calculate speed and distance. The study's non-linear relationship demonstrates that elements other than vehicle speed and distance influence fuel use. They reported a relationship between performance and fuel flow. The researchers want to use a variety of OBD features to estimate the accurate flow of fuel in the future.

3. Theoretical Background

The philosophy for the numerous strategies covered in this work is outlined in this section. The SVM, ARIMA, and Neural Networks are all briefly described in this section keeping in mind the application of these approaches in accurate forecasting.

3.1 SVM

Because the SVM model is built on the structural risk minimization (SRM) concept, which prevents data overfitting, SVM has gained prominence in classification and regression problems [25]. Saleh et al. used the Support Vector Regression (SVR) technique of the SVM model to anticipate CO2 emissions from energy and coal production data [13]. The C and epsilon values for SVM were established by trial and error, and the ones with the lowest RMSE were chosen for the prediction. With a C value of 0.1 and an epsilon value of 0, the RMSE error for prediction was 0.004. In another work, Ju et al. focus on the importance of data preparation and standardisation in order to acquire accurate results [14]. The benefits of PSO over GA, such as its ease of installation, understanding, and the need for only a few parameters for adjustment were also discussed. They used SVM, which was improved by PSO-SVM, to anticipate precipitation. Years of metrological data from Nanjing Station were acquired and pre-processed before modelling, and the dimensionality was reduced using PCA (Principal Component Analysis). When MSE was used to compare the accuracy of SVM-PSO, SVM, SVM-GA (Genetic Algorithm), and Ant Colony SVM (ACSVM), SVM-PSO had the lowest MSE value and so had the best accuracy.

In [15], Jirong et al. showed the need to identify acceptable SVM parameters quickly. Therefore, they introduced GA (Genetic Algorithm) as a technique for optimising SVM because it takes less time. In that work, they forecast housing prices in China. The SVM-GA model was compared against the GM model using the absolute relative error (ABE) measure. It has been found that when compared to GM, GA-SVM had the least ABE value, proving its accuracy in predicting home values.

Using data from the World Bank from 1960 to 2016, Kunda et al. applied the SVM model to forecast CO_2 emissions in Zambia from the year 2017 to 2021 [16]. For comparisons on time series data, their study used the open-source data mining tool WEKA (Waikato Environment for Knowledge Analysis) and the SMOreg (Sequential minimal optimization regression) method. CO_2 emissions are expected to increase from 2017 to 2021, according to their future forecast.

3.2 ARIMA

ARIMA appears to be the most extensively used time series analysis model among scientists, and choosing the proper ARIMA model is crucial for generating precise time-series predictions. In one of the research, authors used the Grey Model (GM) and ARIMA model to forecast CO_2 emissions in Iran using data from British Petroleum [17]. They reported that in terms of predicting, GM beat the ARIMA model with reduced RMSE, MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error). Also, the future estimate made with GM for 2020 predicted a 66 per cent increase in emissions in their work. Accepting that the workforce is a major issue for everyone in today's era, Wang et al. used ARIMA to forecast employment based on supply-demand data from the China Human Resource market [18]. After maintaining the time series data constant and selecting a suitable ARIMA model, the quarter forecast for 2008 is completed, revealing that the most recruiting will happen in the third quarter of the year. They go on to claim that ARIMA isn't good for projecting the future because the errors will only get worse. In another work, Amin et al. used ARIMA and SVM short-term predictive modelling to ensure that the power system's operation remained stable and dependable [19]. SVM outperformed ARIMA models when it comes to forecasting short-term load since its RMSE and MAPE values were lower.

3.3 Neural Networks

Many researchers have used neural networks to forecast because they are flexible and make fewer implications about the input. In [20], Sheta et al. used data from natural gas, coal, global oil, and primary energy consumption to estimate CO_2 emissions using the Product Unit Neural Network (PUNN) and Neural Network Auto Regressive external (NNARX) models. The work reported that in terms of prediction, the PUNN model has been shown to be more efficient, with lower RMSE, MAE, and a high VAF (variance accounted for) value. Li et al. used an RBF neural network via MATLAB to execute time series of CO_2 emissions in China because of its advantages of fast learning and strong generalisation performance [21]. The model had a good predicting capacity and a low absolute inaccuracy when it comes to anticipating emissions. Yenidogan et al. proposed a time-series wavelet transform LSTM forecasting model hourly vehicle emissions of CO (Carbon Monoxide), HC (Hydrocarbons), and NO (Nitrogen Oxide) using geological data from the China Meteorological Administration from May to December 2017 [22]. The results showed that wavelet transform LSTM had better prediction performance than ARIMA, as measured by the RMSE and MAE performance measures.

4. Experimental Setup and Discussion of Results

4.1 ARIMA Process

Theory and applications of time-series analysis are becoming increasingly important in a range of industries. A time series model is a mathematical model that describes the pattern or fluctuation in time series data. The statistical properties of a stationary process are those that stay unchanged across time.

The autocorrelation coefficient (ρk) quantifies the association across successive observations of time series data and is expressed by the equation below:

$$\rho_k = \frac{\sum_{i=1}^{n-k}(Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

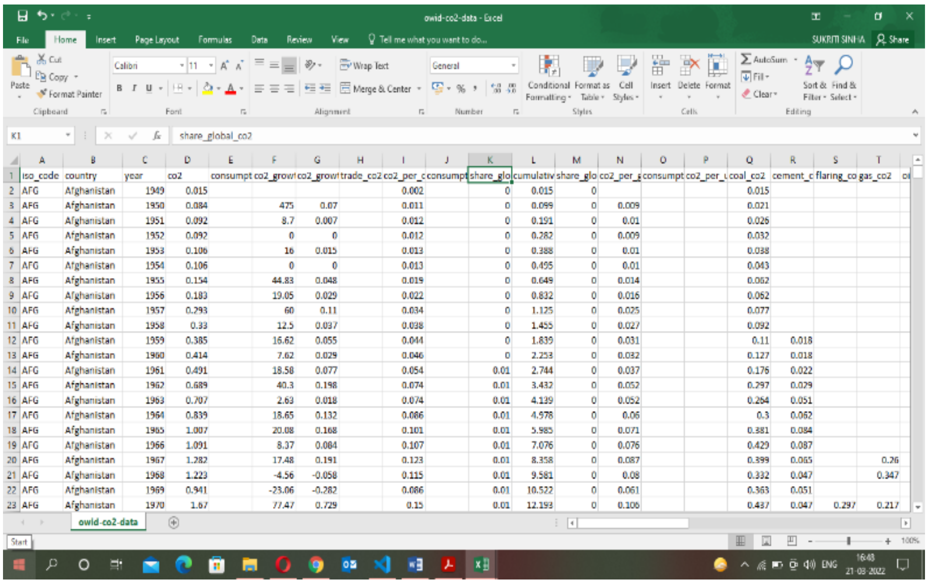
The equation below gives the ARIMA (p, d, q).

$$X_i = \mu + \frac{\theta(B)}{\varphi(B)}\varepsilon_i$$

Where i is the index of time; B is the backward shift operator; ϕ(B) is the auto-regressive operator; p is their order; μ is the mean term; d is the number of differentiations required to make the process stationary; θ(B) is the moving average operator; q is their order; ε_i is the independent error with mean 0 and variance $\sigma\varepsilon^2$. To understand how the data behaves, we begin time series analysis by charting and categorizing it using conventional plots and summary statistics. We will determine whether the data is stationary using the auto-correlogram and partial auto-correlogram; if not, stationary can be achieved by removing the differences between data values.

4.1.1. Data Analysis

The World Bank provided the data for this study, which included yearly CO_2 emissions throughout the world. Each year from 1950 - 2020 is represented by one observation.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	iso_code	country	year	co2	consumpt	co2_growt	co2_growt	trade_co2	co2_per_c	consumpt	share_glo	cumulativ	share_glo	co2_per_g	consumpt	co2_per_u	coal_co2	cement_c	flaring_co	gas_co2	oi
2	AFG	Afghanistan	1949	0.015					0.002		0	0.015	0				0.015				
3	AFG	Afghanistan	1950	0.084		475	0.07		0.011		0	0.099	0	0.009			0.021				
4	AFG	Afghanistan	1951	0.092		8.7	0.007		0.012		0	0.191	0	0.01			0.026				
5	AFG	Afghanistan	1952	0.092		0	0		0.012		0	0.282	0	0.009			0.032				
6	AFG	Afghanistan	1953	0.106		16	0.015		0.013		0	0.388	0	0.01			0.038				
7	AFG	Afghanistan	1954	0.106		0	0		0.013		0	0.495	0	0.01			0.043				
8	AFG	Afghanistan	1955	0.154		44.83	0.048		0.019		0	0.649	0	0.014			0.062				
9	AFG	Afghanistan	1956	0.183		19.05	0.029		0.022		0	0.832	0	0.016			0.062				
10	AFG	Afghanistan	1957	0.293		60	0.11		0.034		0	1.125	0	0.025			0.077				
11	AFG	Afghanistan	1958	0.33		12.5	0.037		0.038		0	1.455	0	0.027			0.092				
12	AFG	Afghanistan	1959	0.385		16.62	0.055		0.044		0	1.839	0	0.031			0.11	0.018			
13	AFG	Afghanistan	1960	0.414		7.62	0.029		0.046		0	2.253	0	0.032			0.127	0.018			
14	AFG	Afghanistan	1961	0.491		18.58	0.077		0.054		0.01	2.744	0	0.037			0.176	0.022			
15	AFG	Afghanistan	1962	0.689		40.3	0.198		0.074		0.01	3.432	0	0.052			0.297	0.029			
16	AFG	Afghanistan	1963	0.707		2.63	0.018		0.074		0.01	4.139	0	0.052			0.264	0.051			
17	AFG	Afghanistan	1964	0.839		18.65	0.132		0.086		0.01	4.978	0	0.06			0.3	0.062			
18	AFG	Afghanistan	1965	1.007		20.08	0.168		0.101		0.01	5.985	0	0.071			0.381	0.084			
19	AFG	Afghanistan	1966	1.091		8.37	0.084		0.107		0.01	7.076	0	0.076			0.429	0.087			
20	AFG	Afghanistan	1967	1.282		17.48	0.191		0.123		0.01	8.358	0	0.087			0.399	0.065		0.26	
21	AFG	Afghanistan	1968	1.223		-4.56	-0.058		0.115		0.01	9.581	0	0.08			0.332	0.047		0.347	
22	AFG	Afghanistan	1969	0.941		-23.06	-0.282		0.086		0.01	10.522	0	0.061			0.363	0.051			
23	AFG	Afghanistan	1970	1.67		77.47	0.729		0.15		0.01	12.193	0	0.106			0.437	0.047	0.297	0.217	

Fig 2: country-wise (annual) dataset of CO_2 Emissions in 1950 – 2020.

For our analysis, we used Exploratory Data Analysis (EDA) Technique. It is a way of studying datasets in order to summarise their key characteristics, which commonly involves the use of statistical graphics and other data visualisation techniques. EDA helps Data Scientists in a variety of ways:

1. Improving data comprehension.
2. Recognizing different data patterns.
3. Improving comprehension of the problem statement.

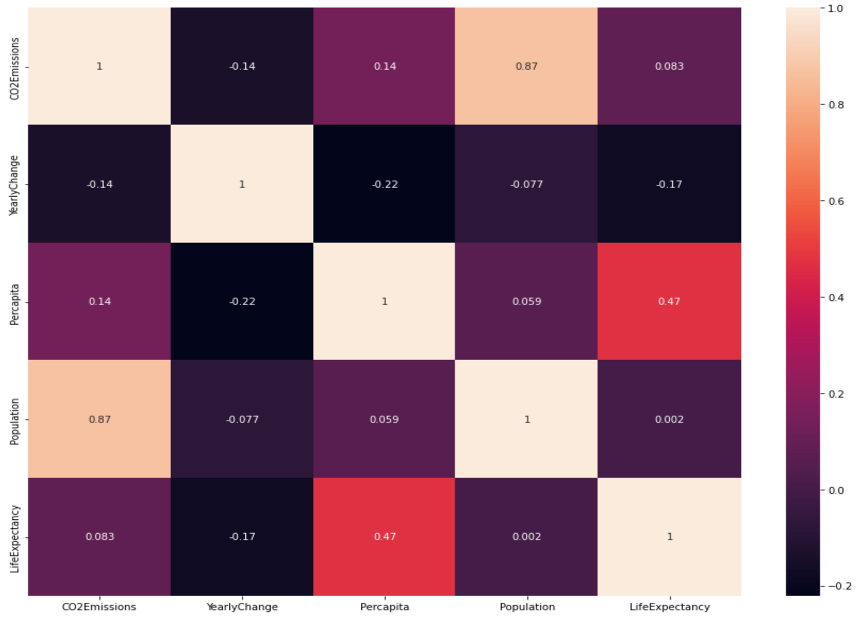


Fig 3: Heat-map depicting the correlation between each feature.

The dataset (country wise as shown in Fig 2) which contained the average value of the emissions over the range of years has been analysed. It has been observed that US was the most CO_2 emitting country (see Fig 2). In Figure 3, a heat map has been used to depict the correlation between all features signifying their importance in accurate forecasting. In the present work, we applied EDA techniques on the US dataset. Here we focused on post-1950 data for our global EDA above since some countries are missing significant data before that time as shown in Fig 4.

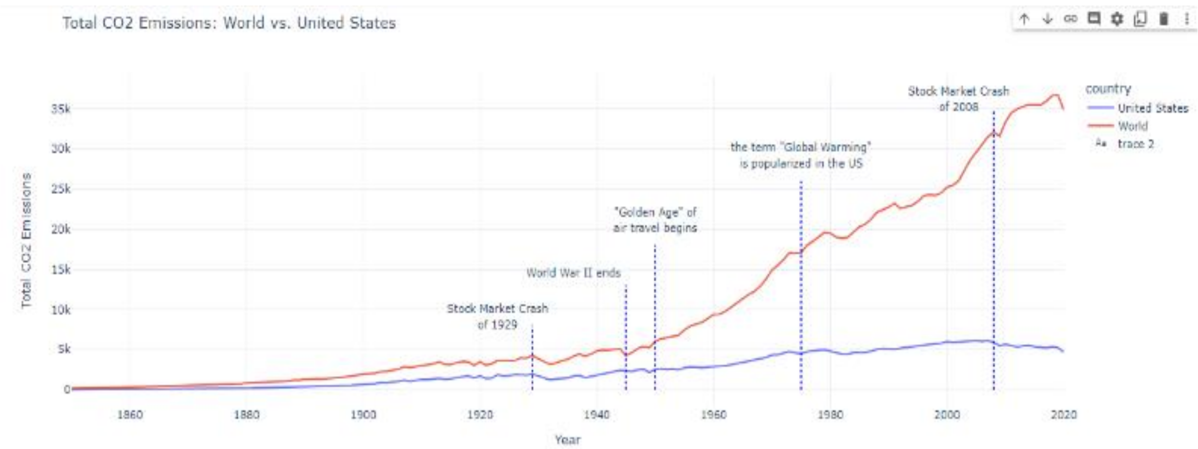


Fig 4: Graph depicting total CO_2 Emissions: World vs. United States

4.2 Modelling

The dataset was used to train various regression models, which were then compared. Finally, the study made use of SVM (Support Vector Machine), ARIMA, and Artificial Neural Networks (ANN). The remaining 10% of the data is forecasted using machine learning models that have been trained on 90% of the data.

4.2.1 Linear Regression

It is a supervised learning machine learning technique. It performs regression analysis. A target prediction value is calculated using independent factors. Usually it is used to predict and figure out how variables are related.

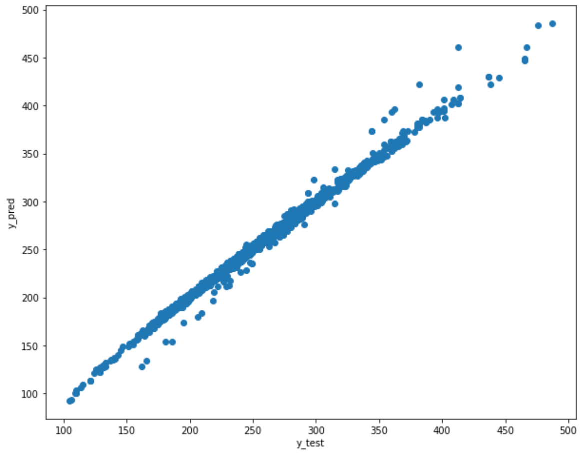


Fig 5: Plotted values of prediction made on the test data.

4.2.2 LASSO Regression

The Least Absolute Shrinkage and Selection Operator (LASSO) is a type of linear regression that is based on shrinkage. In shrinkage, data values are shrunk towards a central point, such as the mean. The LASSO technique encourages simple, sparse models (i.e. models with fewer parameters). The CO_2 emission using LASSO regression is shown in Fig 6.

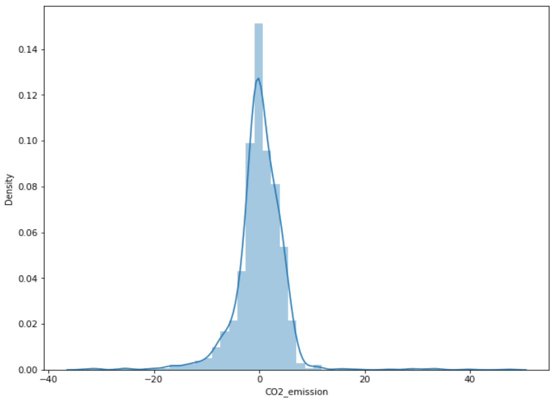


Fig 6: Density of CO_2 emission using lasso regression.

4.2.3 Ridge Regression

Ridge Regression is a multi-collinear data analysis tool used in multiple regression analysis. When multi-collinearity exists, least-square estimations are neutral, but their variances are high, therefore they may be far from the true value. The CO_2 emission using Ridge regression is shown in Fig 7.

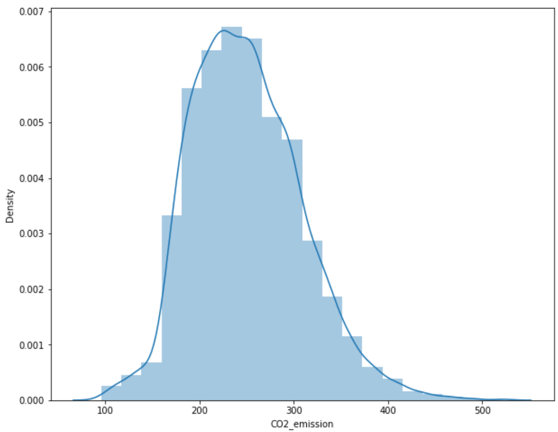


Fig 7: Density of CO_2 emission using Ridge regression.

4.2.4 Decision Tree Regression

A decision tree builds regression or classification techniques in the form of a tree structure. It breaks down a dataset into smaller and smaller segments over time while also building a decision tree. Finally, the outcome is a tree with decision nodes and leaf nodes. The results are shown in Table 1.

Table 1: Difference between predicted and real values of CO_2 emission using decision tree regression.

Predicted values	Real values	Difference
357.0	368.0	11.0
290.0	290.0	0.0
382.0	382.0	0.0
211.0	211.0	0.0
193.0	193.0	0.0
244.0	244.0	0.0
210.0	210.0	0.0
174.0	174.0	0.0
267.0	268.0	1.0
304.6	305.0	0.4

4.2.5 Random Forest

Random Forest Regression is a supervised learning method that uses the ensemble learning approach. Ensemble learning blends estimates from multiple machine learning techniques to produce a more precise forecast than a single set. The final outcome is shown in Table 2.

Table 2: Difference between predicted and real values of CO_2 emission using random forest.

Predicted values	Real values	Difference
358.83	368.0	9.17
290.59	290.0	0.59
3823.40	382.0	1.40
211.22	211.0	0.22
192.70	193.0	0.30
245.22	244.0	1.22
211.47	210.0	1.47
174.40	174.0	0.40
267.01	268.0	0.99
304.59	305.0	0.41

4.2.6 Support Vector Machine

The purpose of the Support Vector Machine (SVM) algorithm is to find a hyper-plane in N-dimensional space that differentiates across data points (N – the number of attributes). To divide the two groups of data points, we can choose from many hyperplanes. The output using the SVM model is shown in Table 3.

Table 3: Difference between predicted and real values of CO_2 emission using SVM.

Predicted values	Real values	Difference
363.55	368.0	4.45
294.18	290.0	4.18
383.37	382.0	1.37
210.62	211.0	0.38
192.43	193.0	0.57

248.72	244.0	4.72
210.99	210.0	0.99
174.29	174.0	0.29
266.68	268.0	1.32
304.31	305.0	0.69

4.3 Comparison

On comparing all the algorithms mentioned above, below is the cumulated result:

Table 4: Comparison of models

Models	RMSE_train	MAE_train	R2_train	RMSE_test	MAE_test	R2_test
Linear regression	4.962163	2.996013	0.992778	4.918261	2.979052	0.993042
Lasso regression	5.044863	3.063387	0.992535	5.010846	3.069379	0.992777
Ridge regression	4.962164	2.996319	0.992778	4.918237	2.979438	0.993042
Decision Tree	0.939813	0.318859	0.999741	3.719588	1.786436	0.996020
Random Forest	1.493768	0.878609	0.999346	3.172145	1.956885	0.997105
Simple Vector	5.348804	2.713308	0.991609	5.434088	2.803636	0.991506

4.4 Time series

There are several stages to take while studying time series data. You should first look for stationarity and autocorrelation. Stationarity is a metric for determining whether or not data exhibits structural patterns, such as seasonal trends.

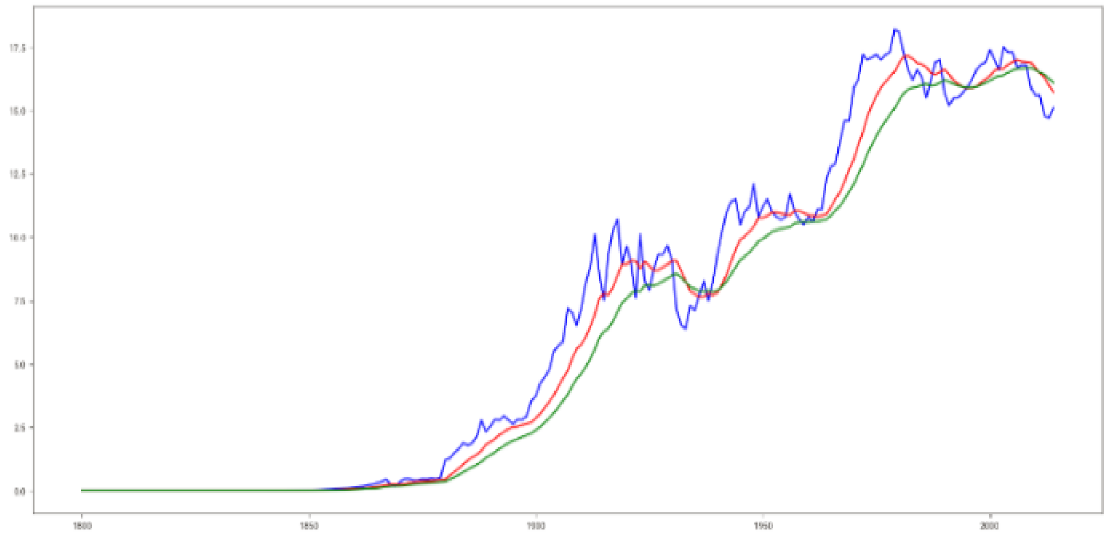


Fig 8: Simple Exponential Smoothing on dataset

When future values in a time series are linearly dependent on previous values, autocorrelation arises. Both of these assumptions are made by many commonly used time series analysis methods, thus you should check for them in time series data.

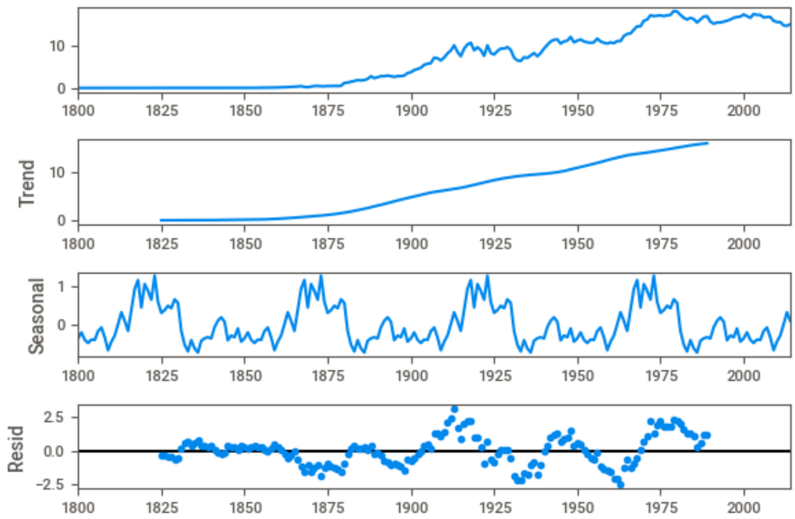


Fig 9: Seasonal decomposition and ADF test

Stationarity is assumed by the autoregressive integrated moving average (ARIMA) approach for predicting time series, for example. Furthermore, linear regression for time series forecasting presupposes no autocorrelation in the data. Before we begin these procedures, we must first determine if the data is suitable for analysis. There is a need to do trend decomposition and anticipate future values during a time series analysis. Decomposition helps in visualizing data trends, which is a terrific method to explain how they (trends) behave. Finally, forecasting helps in predicting future occurrences, which may help in making better decisions.

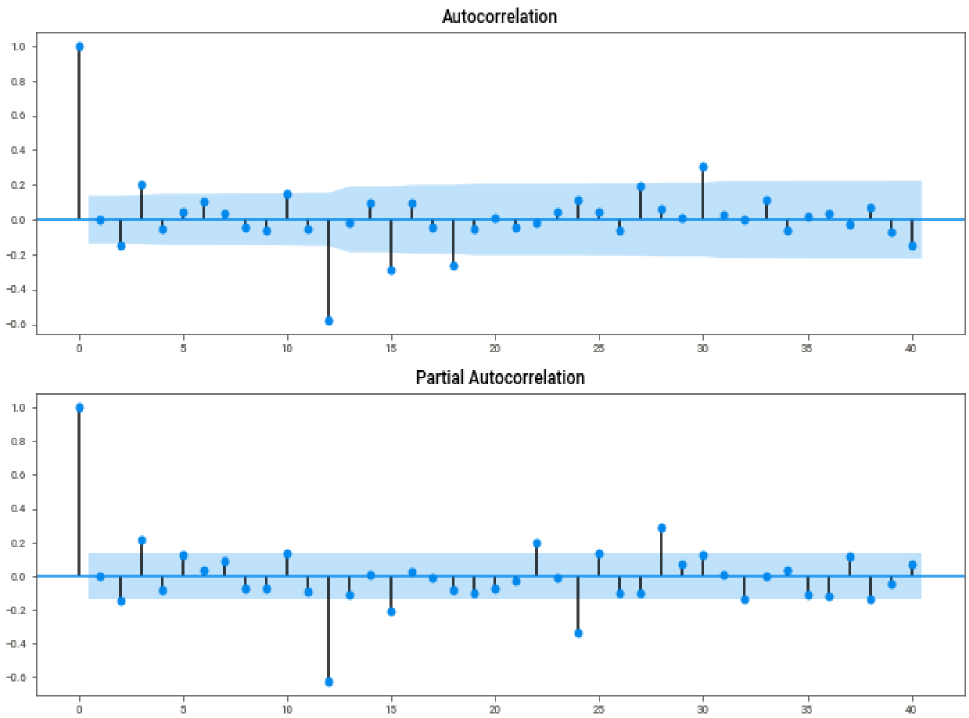


Fig 10: Autocorrelation

5. Conclusion

We have seen in this work that our simulated calculations on CO_2 emissions correlate very well with the World Bank's experimental data, allowing us to anticipate CO_2 emissions until 2030 and take the necessary actions in a reasonable timeframe. These findings support our ARIMA model's ability to provide highly accurate forecasts and show that it agrees well with experimental outcomes. This discovery could be extremely useful in forecasting disasters caused by an unanticipated increase in CO_2 emissions.

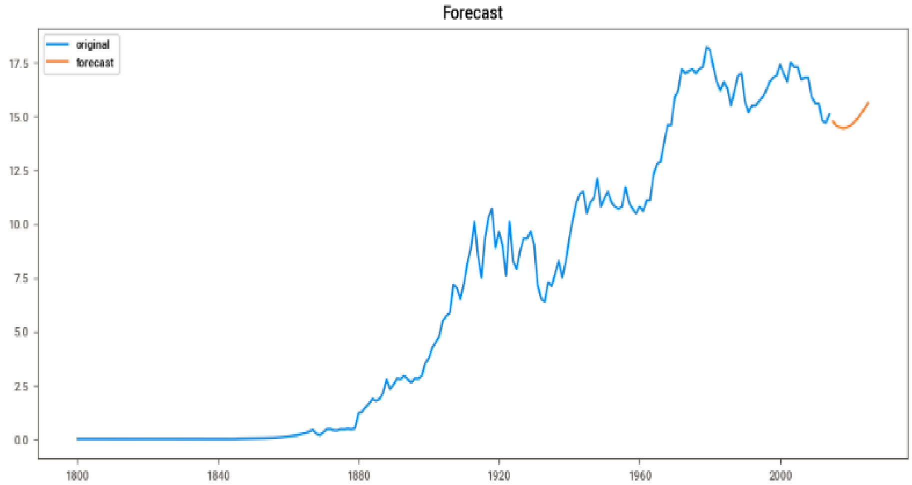


Fig 11: Forecasted CO_2 emission

REFERENCES

[1] S. Jiang, C. Yang, J. Guo, and Z. Ding, "ARIMA forecasting of China's coal consumption, price and investment by 2030," *Energy Sources, Part B: Economics, Planning and Policy*, vol. 13, no. 3, pp. 190–195, 2018.

[2] F. Jiang, X. Yang, and S. Li, "Comparison of forecasting India's energy demand using an MGM, ARIMA model, MGM-ARIMA model, and BP neural network model," *Sustainability*, vol. 10, no. 7, Article ID 2225, 2018.

[3] L. Wang, H. Zou, J. Su, L. Li, and S. Chaudhry, "An ARIMA-ANN hybrid model for time series forecasting," *Systems Research and Behavioral Science*, vol. 30, no. 3, pp. 244–259, 2013.

[4] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[5] N. D. Bokde, Z. M. Yaseen, and G. B. Andersen, "Forecast TB-an R package as a test-bench for time series forecasting—application of wind speed and solar radiation modeling," *Energies*, vol. 13, no. 10, Article ID 2578, 2020.

[6] N. D. Bokde, B. Tranberg, and G. B. Andresen, "Short-term CO_2 emissions forecasting based on decomposition approaches and its impact on electricity market scheduling," *Applied Energy*, vol. 281, Article ID 116061, 2021.

[7] S. E. Bouziane, M. T. Khadir, and J. Dugdale, "A collaborative predictive multi-agent system for forecasting carbon emissions related to energy consumption," *Multiagent and Grid Systems*, vol. 17, no. 1, pp. 39–58, 2021.

[8] M. Grote, I. Williams, J. Preston, and S. Kemp, "A practical model for predicting road traffic carbon dioxide emissions using inductive loop detector data." Transportation Research Part D: Transport and Environment, 63:809-825, 2018.

[9] M. Belaidi, A. Belguerna, O. Alami, "Comparative study between real data and theoretical results calculated by ARIMA model to predict and calculate C02 emissions", Vol.13 No. 01, 289-295, 2022.

[10]B. Quast, "rnn: a recurrent neural network in r", 2016.

[11]U. Shafi, A. Safi, A. R. Shahid, S. Ziauddin and M. Q. Saleem, "Vehicle remote health monitoring and prognostic maintenance system", Journal of Advanced Transportation, 8061514, Jan 2018

[12]W. Zeng, T. Miwa, Y. Wakita, and T. Morikawa, "Exploring trip fuel consumption by machine learning from GPS and CAN bus data", Journal of the Eastern Asia Society for Transportation Studies, 11:906–921, 2015.

[13]C. Saleh, N.R. Dzakiyullah and J.B. Nugroho, "Carbon dioxide emission prediction using support vector machine", IOP Conference Series: Materials Science and Engineering 114(1), 012148, 2016.

[14]J. Du, Y. Liu, Y. Yu and W. Yan, "A prediction of precipitation data based on Support Vector Machine and Particle Swarm Optimization (PSO-SVM) algorithms", Algorithms 10(2), 2017.

[15]G. Jirong, Z. Mingcang and J. Liuguan, "Housing price forecasting based on genetic algorithm and support vector machine", Expert Systems with Applications 38(4), 3383-3386, 2011.

[16]D. Kunda and H. Phiri, "An Approach for Predicting CO_2 Emissions using Data Mining Techniques", International Journal of Computer Applications 172(3), 7-10, 2017.

[17]M. R. Lotfalipour, M.A. Falahi and M. Bastam, "Prediction of CO_2 emissions in Iran using grey and ARIMA models", International Journal of Energy Economics and Policy 3(3), 229-237, 2013.

[18]X. Wang and Y. Liu, "ARIMA time series application to employment forecasting", Proceedings of 2009 4th International Conference on Computer Science and Education, ICCSE 2009 pp. 1124 – 1127, 2009.

[19]M. A. A. Amin and M.A. Hoque, "Comparison of ARIMA and SVM for short-term load forecasting", IEMECON 2019 - 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference pp. 205 – 210, 2019.

[20]A. F. Sheta, N. Ghatasheh and H. Faris, "Forecasting global carbon dioxide emission using auto-regressive with eXogenous input and evolutionary product unit neural network models", 2015 6th International Conference on Information and Communication Systems, ICICS 2015 pp. 182-187, 2015.

[21]S. Li, R. Zhou and X. Ma, "The forecast of CO_2 emissions in China based on RBF neural networks", 2010 2nd International Conference on Industrial and Information Systems, IIS 2010 1, 319-322, 2010.

[22]I. Yenidogan, A. Cayir, O. Kozan, T. Dag, and C. Arslan, “Bitcoin Forecasting Using ARIMA and PROPHET”, UBMK 2018 - 3rd International Conference on Computer Science and Engineering (September 2018), 621 – 624, 2018.

[23]D. Godara, A. Choudhary, and R. K. Singh, “Predicting Change Prone Classes in Open Source Software. In I. Management Association (Ed.), Research Anthology on Usage and Development of Open Source Software, pp. 653-675, IGI Global, 2021. https://doi.org/10.4018/978-1-7998-9158-1.ch034

[24]N. Bohra, V. Bhatnagar, A. Choudhary, S. Ahlawat, D. Sheoran, A. Kumari, “Popularity Prediction of Social Media Post Using Tensor Factorization”, Intelligent Automation & Soft Computing, Vol. 36(1), pp 205-221, 2023.

[25]S. Ahlawat and A. Choudhary, “Hybrid CNN-SVM Classifier for Handwritten Digit Recognition, Procedia Computer Science, Vol. 167, pp. 2554-2560, 2020.