Multiple-Forecasting Approach: A Prediction of CO_2 Emission from the Paddy Crop in India

Pawan Kumar Singh (pawan.singh@thapar.edu)
Thapar Institute of Engineering and Technology https://orcid.org/0000-0002-4822-886X
Alok Kumar Pandey
Banaras Hindu University
Sahil Ahuja
Thapar Institute of Engineering and Technology
Ravi Kiran
Thapar Institute of Engineering and Technology

Research Article

Keywords: SARIMAX, Holt-Winters, Support Vector Regression, CO2, Regression

Posted Date: September 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-767860/v1

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on November 29th, 2021. See the published version at https://doi.org/10.1007/s11356-021-17487-2.

Abstract

This paper compares four prediction methods namely Random Forest Regressor (RFR), SARIMAX, Holt-Winters (H-W), and the Support Vector Regression (SVR) to forecast the total CO_2 emission from the paddy crop in India. The major objective of this study is to compare these four models to suggest an effective model to predict the total CO_2 emission. Data from 1961 to 2018 has been categorised into two parts: training and test data. The study forecasts total CO_2 emission from paddy crop in India from 2019 to 2025. A comparison of mean absolute percentage error (MAPE) and the mean square error (MSE), highlights the differences in accuracy among the four models. The mean absolute percentage error (MAPE) and the mean square error (MSE) for the four methods are: RFR (MAPE: 5.67; MSE: 549900.02), SARIMAX (MAPE:1.67; MSE:70422.35), H-W (MAPE:0.75; MSE:16648.58), and SVR (MAPE: 0.91; MSE: 17832.4). The values of MAPE and MSE with the Holt-Winters (H-W) and the Support Vector Regression (SVR) is relatively low as compared to SARIMAX and RFR. On the basis of these results, it can be inferred that H-W and SVR were found suitable models to forecast the total CO_2 emission from paddy crop. Holt-Winters the model predicted 14364.97 for the year 2025 and SVR predicted 13696.67 for the year 2025. These predictions can be used by the decision-maker to build a suitable policy for future studies. For further research, this approach can be contrasted with other approaches, such as the Neural Network or other forecasting methods, using more important datasets to train the model to achieve better forecast accuracy.

1. Introduction

Maize, Paddy, and Wheat are the three major crops and these make more than fifty per cent of food intake of the human population. Wheat is the highest cultivated crop (214 million hectares annually), followed by rice (154 million hectare annually) and maize (140 million hectares annually). As far as consumption is concerned, human beings consume 85 per cent of rice, 72 percent of wheat and 19 per cent of maize. Rice being the main food for more than fifty percent of the population, is grown worldwide. In the global market, the United States sell half of its rice production annually. There are two rice varieties, i.e. long grain rice and short-grain rice grown in the three regions in the South USA and one region in California respectively (*USDA ERS - Rice*,n.d., 2020).

World trade in rice in the financial year 2021 is expected to be 45.6 million tonnes, with an increase of 0.8 million from the previous estimate and 2% more than the previous year. India accounts for much of the upward revision in the global export outlook, while Bangladesh accounts for much of the upward revision of global imports. Moreover much of Bangladesh's increased imports is primarily projected to be supplied by India (*USDA ERS - Rice*, n.d.), (*Rcs-21a.Pdf*, n.d.).

With the rapid increment in paddy production in India, emission of CO_2 from paddy crop has increased dramatically, and it is a biggest issue for India to reduce CO_2 emission for sustainable development of India. Under the anaerobic state of submerged soils of flooded paddy fields, methane is emitted and much of it escapes from the soil into the atmosphere (FAO, 2020).

Crop and fodder residues consist of nitrogen oxide (N_2O) left at the agriculture field, a source of direct and indirect emission of nitrogen is one of the vital components of greenhouse gases.

The formation N_2O is due to the process of nitrification and de-nitrification is deposited after the leaching process and re-deposition/volatilization (FAOSTAT, FAO, 2020). Rising GHG in the atmosphere increases temperature and is a major cause of concern for the whole world. The two leading gases which cause global warming are methane (15 %) and nitrous oxide (5 %) (Watson et al., 1996). The concentration of these two gases is rising by 3.0 % and 0.22 % annually in the atmosphere (Battle et al., 1996). The production of irrigated lowland rice plays a pivotal role in feeding the rapidly growing population of the world, but at the same time it has been contributing emission of CH_4 and N_2O. It is claimed that the use of inorganic fertilizer with flooded water is contributing to this anoxic situation (*Kanno et al., 1997*). In future, due to rapid development and environmental concerns, the agriculture sector may have to face intensified competition from other sectors for the use of water (Mancosu et al. 2015).

The High water-intensive crops like lowland rice put a stretch on the availability of irrigated water (Kima et al.2014.). Therefore in a sustainable climate zone, the primary challenge is to provide food with a limited water supply to the growing population. Through various studies, it has been observed that field production methods such as fertiliser use, time of irrigation and properties of soil play a dominant role in the emission of greenhouse gases (Wu et al. 2014). However, the tradeoff for GHG emissions in the rice production process remains possible without overlapping the production of the two gases. The waterlogged state of the continuous flooded rice production environment creates an anoxic atmosphere beneficial to methane production by anaerobic methanogenic archaea (Cai et al., 1999). At the same time, the emission of N_2O is related to microbial de-nitrification and nitrification processes that are largely dependent on the degree of soil anaerobicity and nitrate content (Suddick et al., 2011). However, if the emission of CO_2 from paddy will continue, the climate of India will be in a dangerous situation soon and it certainly has an effect on eco-system and human being.

To solve this problem, it is very important to reduce or stop the emission of CO_2 from paddy by applying sustainable method of paddy production. Carbon emissions prediction can provide a scientific basis for the proposition measures of emission reduction. Therefore, in the present paper, multivariable prediction models based on SARIMAX, Random Forest Regressor, Holt-Winters, and Support Vector Regression (SVR) are proposed direct and indirect, and total for carbon emissions from the paddy crops in India. The method of ADF is introduced to know the stationary and non-stationarity in CO_2 emission time-series data. All models (SARIMAX, Random Forest Regressor, and Holt-Winters) of CO_2 emission in this paper show excellent performance in forecasting carbon emissions from paddy in India. However, the performance of Holt-Winters and SVR methods in forecasting carbon emissions from paddy in India is relatively better. The experimental results highlight that Holt-Winters method and Support Vector Regression (SVR) yield higher prediction accuracy than the SARIMAX and Random Forest Regressor (RFR) model for prediction of CO_2 emissions from the paddy crop in India. This paper has applied and suggested Holt-Winters and SVR for forecasting the emission of CO_2 from Paddy crop in India. The accurate prediction of the CO_2 emission from paddy will help India formulate a reasonable policy related to CO_2 emissions control from paddy. The Holt-Winters algorithm smoothes out a time series and helps to uses the data to predict interest areas. Exponential smoothing allocates exponentially diminishing weights and values against historical data to reduce past data weight value (Ferbar Tratar & Strmčnik, 2016).

SVR is a useful and scalable method that allows the user to overcome the shortcomings of the distribution properties of the underlying variables, the geometry of the data and the common issue of model overfitting(Awad & Khanna, 2015). The SVR prediction model was used to predict of carbon dioxide emissions (Zhu et al., 2020). (Shuai Yang et al., 2018) tried to predict the CO_2 emission using SVR of Chongqing from 1997 to 2015. (Zhu et al., 2020) have used SVR model to predict the emission of CO_2 and evaluated the sensitivity of Gross Domestic Production, urbanization rate, population, the structure of energy consumption, industrial structure, and energy intensity influencing the carbon dioxide emission in China. (Saleh et al., 2016) have applied the Support Vector machines to predict carbon (CO_2) emission expenditure. Some other methods also have been using time to time to predict GHG emission from different industries. (Safa et al., 2016) have used Artificial neutral network (ANN) and multiple linear regression model (MLR) to simulate CO_2 emission from wheat farms in New Zealand (Canterbury region). ANN was also used to predict output energy and GHS emissions in potato production in Iran by (Khoshnevisan et al., 2014). (Lehuger et al., 2011) conducted a study to forecast and mitigate the net greenhouse gas emission of crop rotation in Western Europe. (Blagodatsky & Smith, 2012) made an effort for the better mechanistic greenhouse gas emission prediction from the soil. A model was developed to predict CO_2 emission based on the revised version Stochastic Impacts by Regression on Population, Affluence, and Technology Model and used to simulate energy-related CO_2 emissions in five scenarios (Qian et al., 2020). (Wang & Yang, 2018) have used multi-variable prediction models based on GM(1, N) and SVM are proposed for carbon emission from the manufacturing industry in Chongqing. (Marjanović et al., 2016) have developed the Extreme Learning Machine to predict GDP based on CO_2 emission. The ELM results have been compared with genetic programming and ANN. Eventually, this study is demonstrated that EML can be utilised effectively in the application of GDP forecasting. Min & Rulík, 2020, used tillage management practices on CO_2 fluxes on an experimental rice paddy field in Myanmar. (Lin et al., 2011) applied the grey forecasting model to predict future CO_2 emissions in Taiwan from 2010 to 2012. The De-nitrification and Decomposition (DNDC) model used to calibrated and validated for the field experiments and also used to evaluate the ability to simulate methane (CH_4), carbon dioxide (CO_2), and nitrous oxide (N_2O) emissions with various management practices in Indian rice fields (Pathak et al.,2005).

2. Data, Models Training Validation, And Testing

Green House Gases emissions are given as direct, indirect and cumulative by area, regions and particular categories, with the global coverage compared to the period 1961-present (with annual updates) and forecasts for 2030 and 2050, for the gases N_2O and CO_2eq, by crop and N residue content (FAOSTAT, FAO,2020).

The emission of CO_2 (direct, indirect, and total) data was collected from the open-source *Food and Agriculture Organization* (FAO,2020) for 1961 to 2018.

Thr trend of direct, indirect, and total emission of CO_2 was found similar (figure 1), therefore we have analysed only total emission of CO_2 in gigagrams (Gg.). In this study, CO_2 emission data was divided into training data and test data where training data covers from 1961 to 2015 and, test data covers for 2016 to 2018.

On the basis of training and test data, this study conducted long-term forecasts for seven years (2019 to 2025). Python software was used to run SARIMAX and random forest model for prediction. To calculate error in the model testing and actual data was compared for the period 2016 to 2018. In interpreting the predictions, we have used the forecasting accuracy measure MAPE (Mean Absolute Percentage Error) (Kim & Kim, 2016).

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N} \left| \frac{A_t - F_t}{At} \right| \qquad (1)$$

The Mean Square Error is

$$\text{MSE} = \frac{\sum_{t=1}^{n}(At - Ft)^2}{N} \qquad (2)$$

Where, At = Actual values at data time t, and Ft=forecast value at data time t

SARIMAX: The SARIMAX method is an extension of the SARIMA model, improved by adding exogenous variables to improve its forecasting efficiency.

Hence, the model is called Seasonal ARIMA with an Exogenous Factor (i.e. SARIMAX), is commonly expressed mathematically as follows:(Vagropoulos et al., 2016).

$$\varphi_p(B)\varphi_p(B^s)\nabla^d\nabla_s^D y_t = \beta_k x_{k,}\ t + \theta_q(B)\ominus_Q(B^s)\mathcal{E}_t \qquad (3)$$

Where k, t, x ' is the vector, including the kth explanatory input variables at time t and β^k is the k^{th} exogenous input variable's coefficient value. The stationarity and invertibility conditions are as of ARMA methods.

Random Forest Regressor (RFR): Random Forest Regressor is a bagging based supervised learning algorithm used in this paper for time series prediction. The algorithm is based on random sampling with replacement, i.e. bagging the data and a set of decision trees are created.

These decision trees have the problem of getting overfitted easily therefore, the random forest has a regularizing effect by the inclusion of randomness.

Decision Trees have been made by recursively splitting the dataset so that there is maximum information gain where Information content is represented by:

$$I = -\sum p_i log_2(p_i) \quad (4)$$

where is the probability of a particular category of values

Decision Tree reaches to a particular prediction by using a set of decisions learnt from the data and parsing the data through the tree until a leaf node is reached. Random Forest considers several decision tree weak learners and accumulate their independent predictions by averaging them as depicted in the figure 2.

Holt-Winters: Holt-Winters prediction method has been categorized into Multiplicative Holt-Winters, and Additive Holt-Winters.

The equation of Multiplicative Holt-Winters is described as (Ferbar Tratar & Strmčnik, 2016).

$$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \quad (5)$$
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \quad (6)$$
$$s_t = \gamma \frac{y_t}{(\ell_t)} + (1-\gamma)s_{t-m} \quad (7)$$
$$F_{t+m} = (L_t + b_{t+m})\, S_{t-s+m} \quad (8)$$

The component form for the additive method is:

$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \quad (9)$$
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \quad (10)$$
$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \quad (11)$$
$$F_{t+m} = (L_t + b_{t+m})\, S_{t-s+m} \quad (12)$$

Where: $y_{t\,=}$ value of y at t, s= the seasonal length, and m=the amount of data to be predicated.

Support Vector Regression Method

Due to small samples with non-linear features, the support vector machine approach is used to evaluate the selected results. SVM is one of the mathematical theory-based machine learning approaches suggested by Cortes and Vapnik in 1995 (Cortes & Vapnik, 1995).

It offers excellent results for small samples, non-linear and high-dimensional pattern recognition (Cui et al., 2008).

The support vector machine, also called SVR for regression, supports vector thinking and the Lagrange multiplier process, and data can be evaluated. The basic function of the Support vector regression is given by (Cortes & Vapnik, 1995), (Cui et al., 2008),(Chang & Lin, 2011), (Cherkassky & Ma, 2004), (Baydaroğlu & Koçak, 2014):

$$f(x) = \omega.\oint(x) + b \quad (13)$$

3. Results

Decomposition of the time series record of direct emission of CO2 from paddy in India categorised in trend, Seasonality, and random noise. Stat model has been used for seasonal decomposition and the time series frequency, which is the periodicity of the data from 1961 to 2018.

Hence the model is:

Additive model=Trend + Seasonality +Random Noise.   (14)

Forecasting test and error in the model: The analysis highlights that there has been continuing enhancement in direct, indirect, and total emission of CO_2 from paddy crop from 1961 to 2018. In India, it is caused by the conventional method of paddy production. Rice is cultivated in wet, waterlogged soils. Farmers historically flood rice paddies during the growing season-a process known as continuous flooding-providing optimal conditions for microbes that emit significant quantities of methane. Simple improvements to agricultural practices will dramatically minimise these methane emissions, while also reducing the amount of water used during the planting season. Sustainable crop yield practices that minimise water consumption and the need for fertilisers will monitor rising methane emissions and ensure the livelihoods of millions of smallholder rice farmers.

The objective of the current paper is to predict the data for the emission of CO_2 through SARIMAX, Random Forest Regressor, Holt-Winters, and Support Vector Regression (SVR), and compare the results of these four models to find out the most effective. Actual and predicted results of RFR, SARIMAX, SVR, and H-W are shown with the help of figure 3-6 respectively.

At the initial stage, SARIMAX and Random Forest Regressor (RFR) techniques were applied to predict the CO_2 emission from 2019 to 2025 and the basis of training (1961-20115) and tested data (2016-2018). Holt-Winters and SVR models were also applied to compare the efficacy for CO_2 emission from paddy crop in India.

Table 1: Results of fitting direct emission of CO_2 from paddy from 2016 to 2018 in India.

Test	Actual	RFR	APE	SARIMAX (0,1,1)	APE	H-W	APE	SVR	APE
2016	12488.28	11594.3	7.16	12505	0.13	12468.05	0.16	12636.13	1.18
2017	12778.5	12275.2	3.94	12539.7	1.87	12994.86	1.69	12753.97	0.19
2018	13047.98	12275.2	5.92	12655.6	3.01	12995.78	0.4	12871.81	1.35
		MAPE	5.67	MAPE	1.67	MAPE	0.75	MAPE	0.91
		MSE	549900.02	MSE	70422.35	MSE	16648.58	MSE	17832.4

For checking the accuracy of different methods, a comparison of mean absolute percentage error and mean square error from the test data results has been made. Table 1 and 2 presents the comparison of accuracy value, and the forecasting results of direct emission of CO_2 from paddy crop in India for the period 2019 to 2025 of each method respectively. Table 1 shows that the actual CO_2 emission from paddy forecasted CO_2 emission, mean absolute percentage error (MAPE), mean square error (MSE) obtained from random forest regressor (RFR), SARIMAX, Holt-Winters (H-W), and Support Vector Regression (SVR) models. The results were estimated and compared using MAPE and MSE for the RFR, SARIMAX, RFR, and H-W. Table 1 shows that MAPE and MSE values with H-W (0.75;16648.58) and SVR (0.91; 17832.4) were relatively low compared to RFR and SARIMAX. It shows that the comprehensive performance of H-W and SVR models is much better than that of the SARIMAX and RFR with the total emission of CO_2 from paddy crop in India. The Support Vector Regression (SVR) prediction model has been using to predict of carbon dioxide emissions (Zhu et al., 2020), in Chongqing from 1997 to 2015 (Shuai Yang et al., 2018)(These line seams to repeated).

The Holt-Winters models are one of the most popular forecasting algorithms (Trull et al., 2020) and were applied to estimate the trend in overall emissions of organic water contaminants, as well as the exposure of the textile industry to pollution for the top polluters in Eastern Europe, i.e. Poland and Romania (Paraschiv et al., 2015). It has ensured the best predicting values for long-term thermal load forecasting and weekly short-term heat load forecasting (Ferbar Tratar & Strmčnik, 2016).

Finally, we forecasted the total emission of CO_2 from paddy in India with all the above four models from 2019 to 2025 in table 2.

Table 2: Forecast for the total emission of CO_2 from Paddy crop in India

Test	RFR	SARIMAX	H-W	SVR
2019	12307.2	12793.3	13378.38	12989.65
2020	12307.2	12895.4	13115.06	13107.48
2021	11776.9	13016.1	13240.67	13225.32
2022	12307.2	13134	13387.51	13343.16
2023	12307.2	13249.4	13953.16	13460.99
2024	12307.2	13367.6	13954.15	13578.83
2025	12307.2	13484.9	14364.97	13696.67

Prediction of the total emission of CO_2 by Random Forest Regressor (RFR): The amounts of modelling results by RFR model from 2019 to 2025 are shown in Table 2. The total emission of CO_2 from paddy crop in India is predicted to stagnant around 12307.2 gigagrams (Gg.) from 2019 to 2025 by RFR model. But the accuracy results of the RFR model are not good, the MAPE value of the RFR model is 5.67 (MSE:549900.02) which is relatively high compared to SARIMAX, H-W, and SVR models.

Prediction of the total emission of CO_2 by SARIMAX: The values of modelling estimates of SARIMAX model from 2019-2025 are shown in Table 2. The total emission of CO_2 from paddy crop in India will be exceeding 13000 gigagrams (Gg.) in 2021, and it is predicted 13016.1 (Gg.) in 2021 and also predicted to increase slowly to 13484.9 (Gg.) in 2025. The MAPE value of SARIMAX is recorded at 1.67 (MSE:70422.35).

Prediction of the total emission of CO_2 by Holt-Winters (H-W): The values of forecasting results of the Holt-Winters (H-W) model from 2019 to 2025 shown in Table 2. This prediction of the total CO_2 emission from paddy crop in India is predicted 13378.38; 13115.06; 13240.67; 13387.51; 13953.16; 13954.15, and 14364.97 from 2019-2025 respectively given in table 2. The MAPE value for 2019-2025 was 0.75 percent (MSE: 16648.58) to validate the data. Notably, this model (H-W) is predicted to exceed fourteen thousand by 2025, a major concern to be taken in the centre to reduce CO_2 emission from paddy in India. Reduction of CO_2, CH_4 emission is important to prevent the average temperature of the atmosphere (Minamikawa & Sakai, 2006), (Tokida et al., 2010). Climate change is primarily an issue with so much carbon dioxide (CO_2) in the atmosphere. This carbon overload is caused mainly by burning fossil fuels such as gas, oil, and coal or by chopping down and burning trees (Kazmeyer, 2018), (UCSUSA, 2017).

Prediction of the total emission of CO_2 by the Support Vector Regression (SVR): The Support Vector Regression (SVR) prediction model forecasted total emission of CO_2 from the paddy crop in India for 2018-2025 based on training and test data. The value of SVR model shows that total CO_2 emission from paddy in India will be reached 13696.63 (Gg.) in 2025. We obtained the mean absolute percentage error (MAPE) and the mean square error (MSE) 0.91, and 17832.4, respectively.

4. Discussion

The management of CO_2 emission from the paddy fields is very important in mitigating climate change and maintaining a sustainable environment. Recently, a study suggests that reducing GHG emissions will help prevent around 3 million premature loss of lives by 2100. Exposure to pollution in the air is caused to 7 million live deaths worldwide yearly (United Nations, 2020). However, even as paddy is the most widely cultivated and consumed cereal in the Asia continent and the third most consumed grain worldwide. There is a great paucity of information on the emission of CO_2 gases from paddy fields. As a result, there are differences in evidence on paddy's role in driving the emission of CO_2 and the continuous enhancement is implicated in paddy production due to the increasing population.

Consequently, there is various prediction model to forecast CO_2 emission, and it is vital to recommend an appropriate prediction model to forecast CO_2 emission linking less error from paddy fields. It would be possible to utilise time-series data of paddy fields to associate links with CO_2 emission. This study was undertaken in which the SARIMAX, RFR, H-W, and SVR was employed. These techniques were considered because they have the capability to make predictions of CO_2 emission.

5. Conclusion

It was seen in results that there was an excellent out-put of SVR and Holt-Winter(H-W) as compared to SARIMAX and RFR. SVR and H-W's experimental findings were 0.91 and 17832.4 (MAPE & MSE) and 075 and 16648.58 (MAPE & MSE), respectively. The prediction of CO_2 emission from paddy fields for 2025 is obtained at 13696.67 (SVR) and 14364.97 (H-W) in India. This SVR and Holt-Winters can be applied to estimate the total emission of CO_2 from paddy fields in India. In this way, the model will contribute to filling India's vast data gaps in quantum CO_2 emissions from paddy fields. The same prediction method can be used to model nitrous oxide and methane emissions in other parts of the world that will be equally profitably based on evidence from India. Comparative analysis of this evidence on CO_2 emissions from paddy fields would provide insight into paddy production's relative effect on the expansion or the reduction of CO_2 emissions.

6. Policy Implication And Further Suggestion For Research:

The prediction of CO_2 emission is difficult in Indian paddy fields due to its socio-economic status, different soil quality and environment. Also, variation in fertiliser management seeds management plays an important role in the emission of CO_2 from paddy fields. The actual contribution of the agriculture sector in India to CO_2 emission from paddy fields can only be addressed by suggesting an appropriate prediction model. It will predict the emission of CO_2 in the coming year and suggest a baseline from which future emission must be reduced and mitigation strategies. The major challenge worldwide is to produce more food and fiber to meet the increasing requirement for a nine billion population by 2050. Climate change is one of a big problem worldwide, and the rapid emission of CO_2 from the paddy field is also contributing very significantly (Ahmad et al., 2009), (Lohan et al., 2018),(Oo et al., 2018),(Abbasi et al., 2019). The decision-maker can use the results based on the prediction to build policies for future studies. For further research, this approach can be contrasted with other techniques, such as the Neural Network or other forecasting methods, using more important datasets to train the model to achieve better forecast accuracy. This study is conducted for academic and research purposes only, and the forecasts for the future are based on the premise that the restrictive circumstances will remain.

Declarations

Acknowledgements: The authors express their gratitude to the editorial board and reviewer for the efforts for suggestion and reviewing this paper. The authors also appreciate the editor for his cooperation during the review process.

Data availability: The data that support the findings of this study are openly available on request.

Author contribution: Conceptualization, introduction and results writing, and data analysis: Pawan Kumar singh and Sahil Ahuja. Literature review: Alok Kumar Pandey, Ravi Kiran.

Ethical approval: Not applicable.

Consent to participate: Not applicable.

Consent to publish: All authors have read and approved this manuscript.

Competing interests: The authors declare no competing interests.

Funding: The authors did not received any funding from any organization to conduct this study.

References

1. Abbasi, T., Abbasi, T., Luithui, C., & Abbasi, S. A. (2019). Modelling Methane and Nitrous Oxide Emissions from Rice Paddy Wetlands in India Using Artificial Neural Networks (ANNs). *Water*, *11*(10), 2169. https://doi.org/10.3390/w11102169
2. Ahmad, S., Li, C., Dai, G., Zhan, M., Wang, J., Pan, S., & Cao, C. (2009). Greenhouse gas emission from direct seeding paddy field under different rice tillage systems in central China. *Soil and Tillage Research*, *106*(1), 54–61. https://doi.org/10.1016/j.still.2009.09.005
3. Awad, M., & Khanna, R. (2015). Support Vector Regression. In M. Awad & R. Khanna (Eds.), *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers* (pp. 67–80). Apress. https://doi.org/10.1007/978-1-4302-5990-9_4
4. Battle, M., Bender, M., Sowers, T., Tans, P. P., Butler, J. H., Elkins, J. W., Ellis, J. T., Conway, T., Zhang, N., Lang, P., & Clarket, A. D. (1996). Atmospheric gas concentrations over the past century measured in air from firn at the South Pole. *Nature*, *383*(6597), 231–235. https://doi.org/10.1038/383231a0
5. Baydaroğlu, Ö., & Koçak, K. (2014). SVR-based prediction of evaporation combined with chaotic approach. *Journal of Hydrology*, *508*, 356–363. https://doi.org/10.1016/j.jhydrol.2013.11.008
6. Blagodatsky, S., & Smith, P. (2012). Soil physics meets soil biology: Towards better mechanistic prediction of greenhouse gas emissions from soil. *Soil Biology and Biochemistry*, *47*, 78–92. https://doi.org/10.1016/j.soilbio.2011.12.015
7. Cai, Z.-C., Xing, G.-X., Shen, G.-Y., Xu, H., Yan, X.-Y., Tsuruta, H., Yagi, K., & Minami, K. (1999). Measurements of CH_4 and N_2 0 emissions from rice paddies in Fengqiu, China. *Soil Science and Plant Nutrition*, *45*(1), 1–13. https://doi.org/10.1080/00380768.1999.10409320
8. Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 27:1-27:27. https://doi.org/10.1145/1961189.1961199
9. Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, *17*(1), 113–126. https://doi.org/10.1016/S0893-6080(03)00169-2
10. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018
11. Cui, B., Xue, T., & Yang, K. (2008). Vehicle Recognition Based on Support Vector Machine. *2008 International Symposium on Intelligent Information Technology Application Workshops*, 443–446. https://doi.org/10.1109/IITA.Workshops.2008.23
12. *FAOSTAT*. (n.d.). Retrieved January 27, 2021, from http://www.fao.org/faostat/en/#data/GA
13. Ferbar Tratar, L., & Strmčnik, E. (2016). The comparison of Holt–Winters method and Multiple regression method: A case study. *Energy*, *109*, 266–276. https://doi.org/10.1016/j.energy.2016.04.115
14. *Greenhouse Gas Emissions from Cotton Field under Different Irrigation Methods and Fertilization Regimes in Arid Northwestern China*. (n.d.). Retrieved January 27, 2021, from https://www.hindawi.com/journals/tswj/2014/407832/
15. *Greenhouse gas emissions from Indian rice fields: Calibration and upscaling using the DNDC model—NASA/ADS*. (n.d.). Retrieved January 27, 2021, from https://ui.adsabs.harvard.edu/abs/2005BGeo....2..113P/abstract
16. *Is CO2 Bad for the Planet?* (n.d.). Retrieved January 29, 2021, from https://sciencing.com/co2-bad-planet-4876.html
17. *Janetos et al. - 1997—Climate Change 1995 Impacts, Adaptations and Miti.pdf*. (n.d.). Retrieved January 27, 2021, from https://library.harvard.edu/sites/default/files/static/collections/ipcc/docs/36__WGIISAR_FINAL.pdf

18. Khoshnevisan, B., Rafiee, S., Omid, M., Mousazadeh, H., & Rajaeifar, M. A. (2014). Application of artificial neural networks for prediction of output energy and GHG emissions in potato production in Iran. *Agricultural Systems*, *123*, 120–127. https://doi.org/10.1016/j.agsy.2013.10.003
19. Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003
20. Lehuger, S., Gabrielle, B., Laville, P., Lamboni, M., Loubet, B., & Cellier, P. (2011). Predicting and mitigating the net greenhouse gas emissions of crop rotations in Western Europe. *Agricultural and Forest Meteorology*, *151*(12), 1654–1671. https://doi.org/10.1016/j.agrformet.2011.07.002
21. Lin, C.-S., Liou, F.-M., & Huang, C.-P. (2011). Grey forecasting model for CO2 emissions: A Taiwan study. *Applied Energy*, *88*(11), 3816–3820. https://doi.org/10.1016/j.apenergy.2011.05.013
22. Lohan, S. K., Jat, H. S., Yadav, A. K., Sidhu, H. S., Jat, M. L., Choudhary, M., Peter, J. K., & Sharma, P. C. (2018). Burning issues of paddy residue management in north-west states of India. *Renewable and Sustainable Energy Reviews*, *81*, 693–706. https://doi.org/10.1016/j.rser.2017.08.057
23. Marjanović, V., Milovančević, M., & Mladenović, I. (2016). Prediction of GDP growth rate based on carbon dioxide (CO2) emissions. *Journal of CO2 Utilization*, *16*, 212–217. https://doi.org/10.1016/j.jcou.2016.07.009
24. *Methane emission from rice paddy fields in all of Japanese prefecture | SpringerLink*. (n.d.). Retrieved January 27, 2021, from https://link.springer.com/article/10.1023/A:1009778517545
25. Minamikawa, K., & Sakai, N. (2006). The practical use of water management based on soil redox potential for decreasing methane emission from a paddy field in Japan. *Agriculture, Ecosystems & Environment*, *116*(3), 181–188. https://doi.org/10.1016/j.agee.2006.02.006
26. Nations, U. (n.d.). *Key Findings*. United Nations; United Nations. Retrieved February 23, 2021, from https://www.un.org/en/climatechange/science/key-findings
27. Oo, A. Z., Sudo, S., Inubushi, K., Mano, M., Yamamoto, A., Ono, K., Osawa, T., Hayashida, S., Patra, P. K., Terao, Y., Elayakumar, P., Vanitha, K., Umamageswari, C., Jothimani, P., & Ravi, V. (2018). Methane and nitrous oxide emissions from conventional and modified rice cultivation systems in South India. *Agriculture, Ecosystems & Environment*, *252*, 148–158. https://doi.org/10.1016/j.agee.2017.10.014
28. Paraschiv, D., Tudor, C., & Petrariu, R. (2015). The Textile Industry and Sustainable Development: A Holt–Winters Forecasting Investigation for the Eastern European Area. *Sustainability*, *7*(2), 1280–1291. https://doi.org/10.3390/su7021280
29. Qian, Y., Sun, L., Qiu, Q., Tang, L., Shang, X., & Lu, C. (2020). Analysis of CO2 Drivers and Emissions Forecast in a Typical Industry-Oriented County: Changxing County, China. *Energies*, *13*(5), 1212. https://doi.org/10.3390/en13051212
30. *Rcs-21a.pdf*. (n.d.). Retrieved January 27, 2021, from https://www.ers.usda.gov/webdocs/outlooks/100237/rcs-21a.pdf?v=4975.2
31. Safa, M., Nejat, M., Nuthall, P., & Greig, B. (2016). Predicting CO2 Emissions from Farm Inputs in Wheat Production using Artificial Neural Networks and Linear Regression Models. *International Journal of Advanced Computer Science and Applications*, *7*(9). https://doi.org/10.14569/IJACSA.2016.070938
32. Saleh, C., Dzakiyullah, N. R., & Nugroho, J. B. (2016). Carbon dioxide emission prediction using support vector machine. *IOP Conference Series: Materials Science and Engineering*, *114*, 012148. https://doi.org/10.1088/1757-899X/114/1/012148
33. Shuai Yang, Yu Wang, Wengang Ao, Yun Bai, & Chuan Li. (2018). Prediction and Analysis of CO2 Emission in Chongqing for the Protection of Environment and Public Health. *International Journal of Environmental Research and Public Health*, *15*(3), 530. https://doi.org/10.3390/ijerph15030530

34. Suddick, E. C., Steenwerth, K., Garland, G. M., Smart, D. R., & Six, J. (2011). Discerning Agricultural Management Effects on Nitrous Oxide Emissions from Conventional and Alternative Cropping Systems: A California Case Study. In *Understanding Greenhouse Gas Emissions from Agricultural Management* (Vol. 1072, pp. 203–226). American Chemical Society. https://doi.org/10.1021/bk-2011-1072.ch012

35. Tokida, T., Fumoto, T., Cheng, W., Matsunami, T., Adachi, M., Katayanagi, N., Matsushima, M., Okawara, Y., Nakamura, H., Okada, M., Sameshima, R., & Hasegawa, T. (2010). Effects of free-air CO_2 enrichment (FACE) and soil warming on CH_4 emission from a rice paddy field: Impact assessment and stoichiometric evaluation. *Biogeosciences*, *7*(9), 2639–2653. https://doi.org/10.5194/bg-7-2639-2010

36. Trull, O., García-Díaz, J. C., & Troncoso, A. (2020). Initialization Methods for Multiple Seasonal Holt–Winters Forecasting Models. *Mathematics*, *8*(2), 268. https://doi.org/10.3390/math8020268

37. *USDA ERS - Rice*. (n.d.). Retrieved January 27, 2021, from https://www.ers.usda.gov/topics/crops/rice/

38. Nations, U. (2020). *Key Findings*. United Nations; United Nations. Retrieved February 23, 2021, from https://www.un.org/en/climatechange/science/key-findings

39. Vagropoulos, S. I., Chouliaras, G. I., Kardakos, E. G., Simoglou, C. K., & Bakirtzis, A. G. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. *2016 IEEE International Energy Conference (ENERGYCON)*, 1–6. https://doi.org/10.1109/ENERGYCON.2016.7514029

40. Wang, Y., & Yang, S. (2018). The Prediction of CO2 Emissions from Manufacturing Industry Based on GM(1,N) Model and SVM in Chongqing. *2018 International Conference on Sensing,Diagnostics, Prognostics, and Control (SDPC)*, 85–89. https://doi.org/10.1109/SDPC.2018.8664935

41. *Water | Free Full-Text | Improving Irrigated Lowland Rice Water Use Efficiency under Saturated Soil Culture for Adoption in Tropical Climate Conditions*. (n.d.). Retrieved January 27, 2021, from https://www.mdpi.com/2073-4441/6/9/2830

42. *Why Does CO2 get more attention than other gases? | Union of Concerned Scientists*. (n.d.). Retrieved January 29, 2021, from https://www.ucsusa.org/resources/why-does-co2-get-more-attention-other-gases

43. Zhu, C., Wang, M., & Du, W. (2020). Prediction on Peak Values of Carbon Dioxide Emissions from the Chinese Transportation Industry Based on the SVR Model and Scenario Analysis. *Journal of Advanced Transportation*, *2020*, 1–14. https://doi.org/10.1155/2020/8848149

Figures

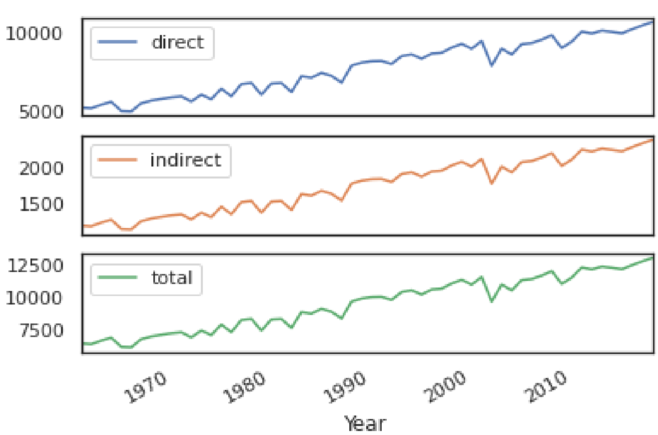

Figure 1

Direct, Indirect and Total emission of CO2 from Rice, Paddy crop in India

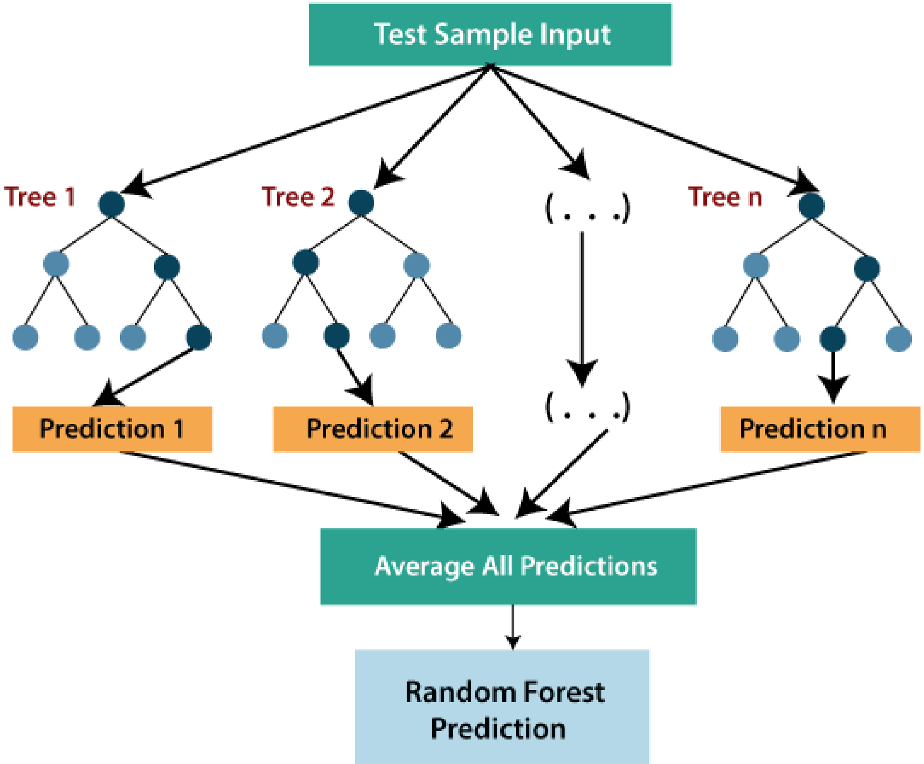

Figure 2

Random Forest model

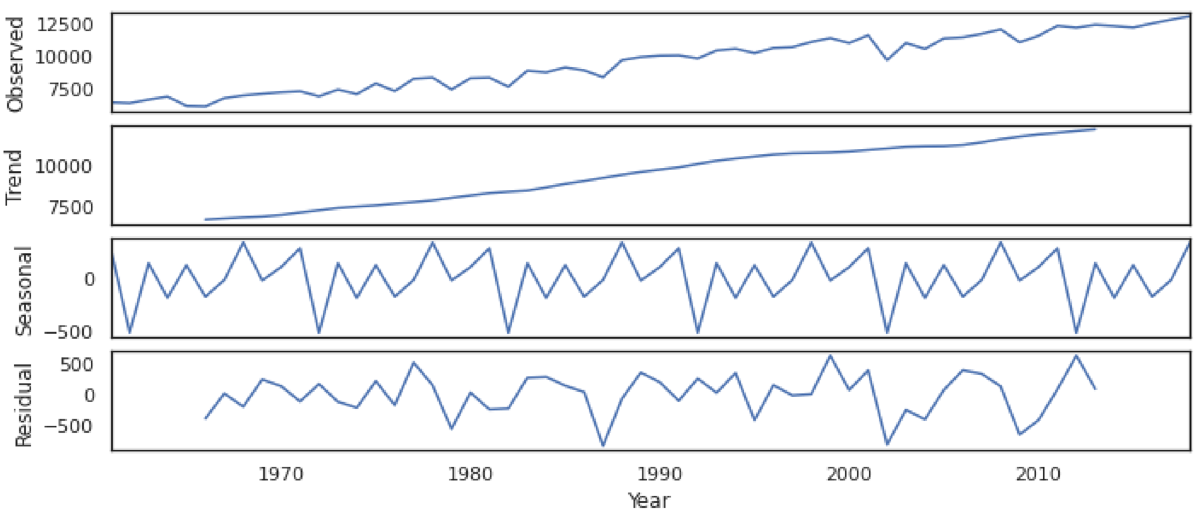

Figure 3

A caption was not provided with this version

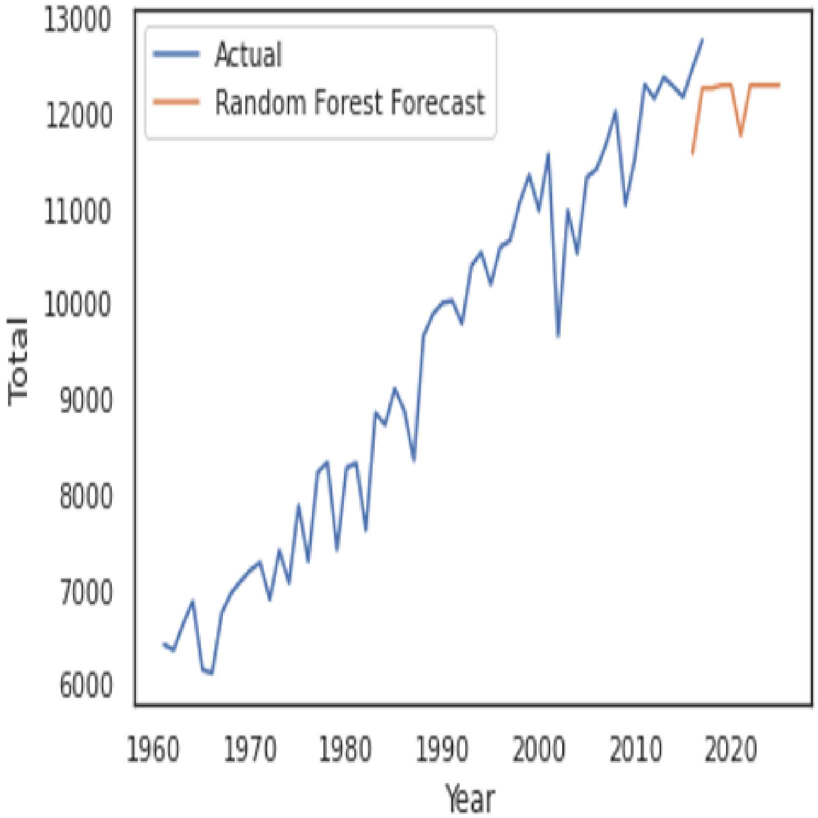

Figure 4

Figure 3 in paper: RFR prediction results

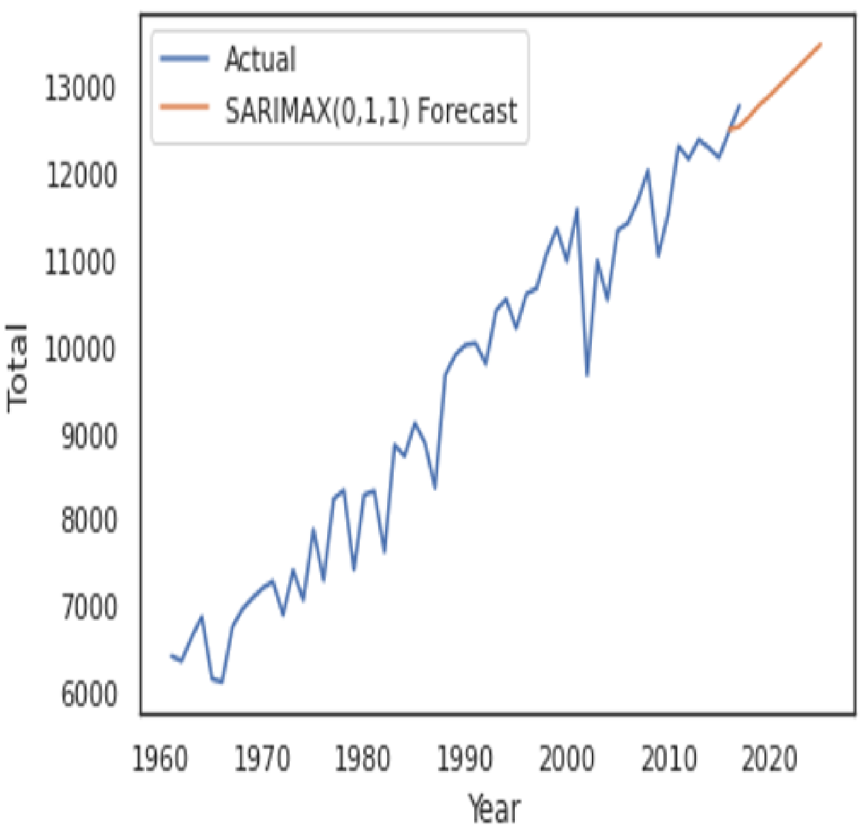

Figure 5

Figure 4 in paper: SARIMAX prediction results

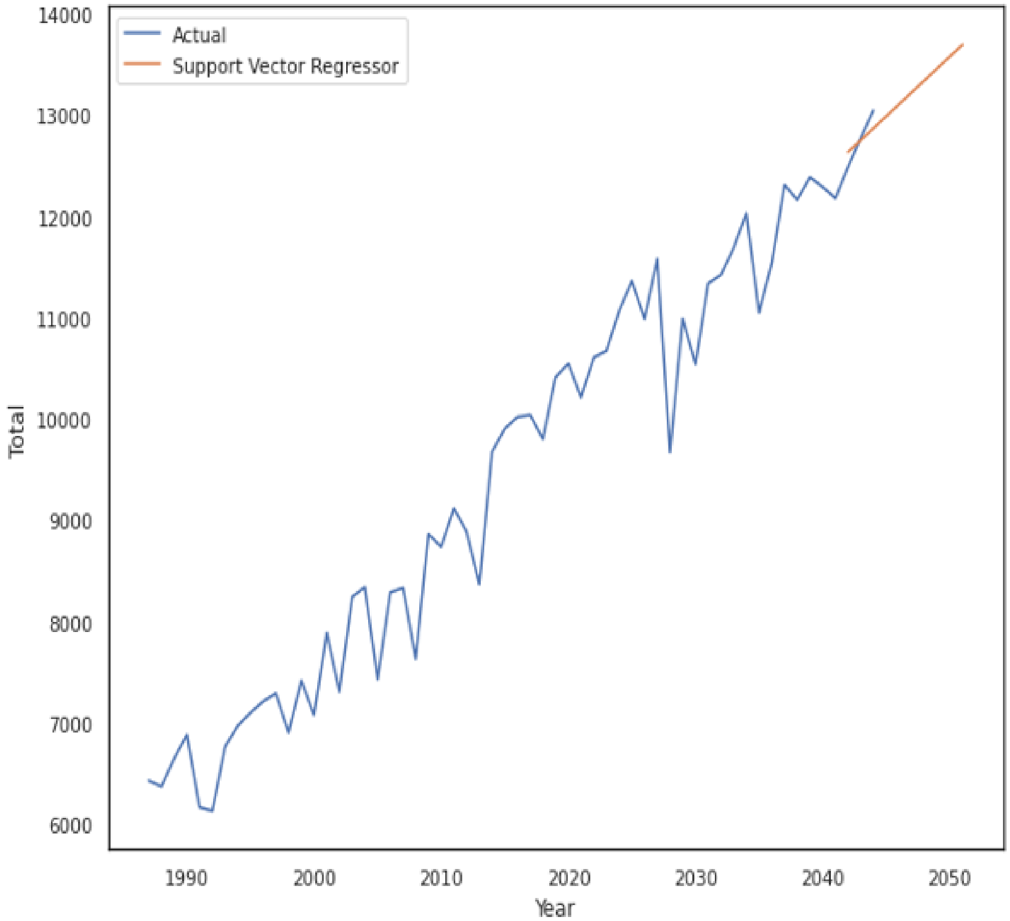

Figure 6

Figure 5 in paper: SVR prediction results

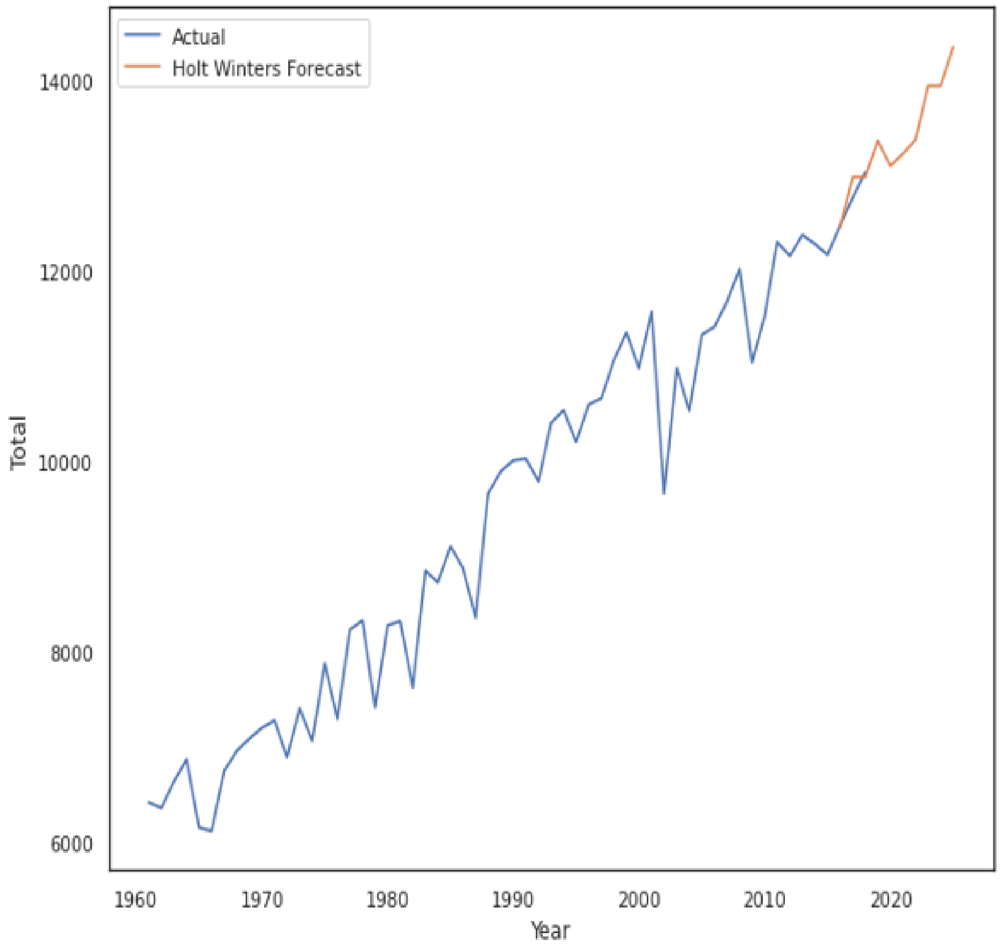

Figure 7

Figure 6 in paper: H-W prediction results