RESEARCH ARTICLE

Multiple forecasting approach: a prediction of CO2 emission from the paddy crop in India

Pawan Kumar Singh1 · Alok Kumar Pandey2 · Sahil Ahuja3 · Ravi Kiran1

Received: 8 August 2021 / Accepted: 8 November 2021 / Published online: 29 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

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 CO2 emission from the paddy crop in India. The major objective of this study is to compare these four models and suggest an effective model for the prediction of total CO2 emission. Data from 1961 to 2018 has been categorised into two parts: training and test data. The study forecasts total CO2 emission from paddy crops 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: 549,900.02), SARIMAX (MAPE: 1.67; MSE:70,422.35), H-W (MAPE:0.75; MSE:16,648.58), and SVR (MAPE: 0.91; MSE: 17,832.4). The values of MAPE and MSE with the Holt-Winters (H-W) and the support vector regression (SVR) are relatively low as compared to SARIMAX and RFR. Based on these results, it can be inferred that H-W and SVR were found suitable models to forecast the total CO2 emission from paddy crops. Holt-Winters model predicted 14,364.97 for the year 2025, and SVR predicted 13,696.67 for the year 2025. The decision-maker can use these predictions to build a suitable policy for the future. This approach can be contrasted with other forecasting methods, such as the neural network, and train the model to achieve better forecast accuracy.

Keywords SARIMAX · Holt-Winters · Support vector regression · CO2 · Regression

Introduction
Maize, paddy, and wheat are the three major crops, and these make more than fifty percent of the food intake of the human population. Wheat is the highest cultivated crop (214 million ha annually), followed by rice (154 million ha annually) and maize (140 million ha annually). As far as consumption is concerned, human beings consume 85 percent of rice, 72 percent of wheat, and 19 percent of maize. Rice being the main food for more than fifty percent of the population is grown worldwide. In the global market, the USA sells 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, 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, 2021).

With the rapid increment in paddy production in India, the emission of CO2 from paddy crop has increased dramatically. Reducing CO2 emissions is one of India’s biggest issues for sustainable development. 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 (N2O) 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 N2O is due to the process of nitrification. De-nitrification is deposited after the leaching process and re-deposition/volatilisation (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 CH4 and N2O. It is claimed that the use of inorganic fertiliser with flooded water is contributing to this anoxic situation (Kanno et al., 1997). In the 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 fertiliser use, irrigation times, and soil properties 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 continuously flooded rice production environments creates an anoxic atmosphere beneficial to methane production by anaerobic methanogenic archaea (Cai et al., 1999). At the same time, the emission of N2O 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 CO2 from paddy continues in India, it will create a dangerous situation, and soon it will certainly affect the ecosystem and human beings.

Hence, it is essential to reduce or stop the emission of CO2 from paddy by applying a 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 as direct and indirect and total for carbon emissions from the paddy crops in India. The method of ADF is introduced to know the stationarity and non-stationarity in CO2 emission time series data. In the paper, all models (SARIMAX, random forest regressor, and Holt-Winters) of CO2 emission 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 the Holt-Winters method and support vector regression (SVR) yield higher prediction accuracy than the SARIMAX and random forest regressor (RFR) model for predicting CO2 emissions from the paddy crop in India. This paper has applied and suggested Holt-Winters and SVR for forecasting the emission of CO2 from the paddy crop in India. The accurate prediction of the CO2 emission from paddy will help India formulate a reasonable policy related to CO2 emissions control from paddy. The Holt-Winters algorithm smooths out a time series and helps to use the data to predict interest areas. Exponential smoothing allocates diminishing weights and values against historical data to reduce past data weight values (Ferbar Tar & 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 carbon dioxide emissions (Zhu et al., 2020). Shuai Yang et al. (2018) tried to predict the CO2 emission using SVR of Chongqing from 1997 to 2015. Zhu et al. (2020) have used SVR model to predict the emission of CO2 and evaluated the sensitivity of gross domestic production, urbanisation 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 (CO2) emission expenditure. Some other methods also have been using time to time to predict GHG emissions from different industries. Safa et al. (2016) have used an artificial neutral network (ANN) and multiple linear regression model (MLR) to simulate CO2 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 rotations in Western Europe. Blagodatsky and Smith (2012) made an effort for the better mechanistic greenhouse gas emission prediction from the soil. A model was developed to predict CO2 emission based on the
revised version Stochastic Impacts by Regression on Population, Affluence, and Technology model and used to simulate energy-related CO2 emissions in five scenarios (Qian et al., 2020). Wang and Yang (2018) have used multivariable prediction models based on GM(1, N) and SVM for carbon emission from the manufacturing industry in Chongqing. Marjanović et al. (2016) have developed the extreme learning machine to predict GDP based on CO2 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 and Rulík (2020) used tillage management practices on CO2 fluxes on an experimental rice paddy field in Myanmar. Lin et al. (2011) applied the grey forecasting model to predict future CO2 emissions in Taiwan from 2010 to 2012. The Denitrification and Decomposition (DNDC) model was used by Pathak et al. (2005) to calibrate and validate the field experiments and also used to evaluate the ability to simulate methane (CH4), carbon dioxide (CO2), and nitrous oxide (N2O) emissions with various management practices in Indian rice fields. MK and V (2020) used machine learning to better predict the relationship between CO2 emission and economic growth. Long et al. (2018) evaluated major determinants of CO2 emission intensity from Chinese agriculture during 1997–2014. A nonlinear approach was applied to know the causality among carbon emissions from energy consumption, agriculture, fertiliser, and cereal food production in Pakistan (Koondhar et al., 2021). Prediction of CO2 emissions was conducted in applying ARIMA models in India (Nyoni & Bonga, 2019). Rahman and Hasan (2017) applied ARIMA forecasting approach to predict carbon dioxide emissions in Bangladesh. Xu et al. (2021) has used a hybrid ARIMA-LSTM prediction model based on the SPEI for drought prediction.

Data

Greenhouse gas 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 N2O and CO2eq, by crop and N residue content (FAOSTAT, FAO,2020).

The emission of CO2 (direct, indirect, and total) data was collected from the open-source Food and Agriculture Organization (FAO,2020) from 1961 to 2018.

Models training validation and testing

The trend of direct, indirect, and total emission of CO2 was found similar (Fig. 1). Therefore, we have analysed only the total emission of CO2 in gigagrams (Gg). In this study, CO2 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.

The present study conducted long-term forecasts for 7 years (2019 to 2025) based on training and test data. Python software was used to run SARIMAX and random forest model for prediction. To calculate error in the model, testing and actual data were compared for 2016 to 2018. In interpreting the predictions, we have used the forecasting accuracy measure MAPE (mean absolute percentage error) (Kim & Kim, 2016).

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{A_i - F_i}{A_i} \right|
\]

(1)

The mean square error is.

\[
MSE = \frac{\sum_{i=1}^{n} (A_t - F_t)^2}{N}
\]

(2)

where \(A_t\) = actual values at data time \(t\) and \(F_t\) = forecast value at data time \(t\).

ML models such as ANN, GM, and SVM are intended to make the most precise predictions possible, whereas statistical models such as H-W, SVR, SARIMAX, and RFR are intended to infer associations between variables. Models in machine learning are defined simply by their performance, operating as a BlackBox. On the other hand, statistical modelling is more concerned with identifying correlations between variables and their significance and making predictions.

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) and commonly expressed mathematically as follows (Vagropoulos et al., 2016):

![Fig. 1 Direct, indirect, and total emission of CO2 from rice paddy crop in India](image-url)
where \( k, t, \) and \( x' \) are the vector, including the \( k \)th explanatory input variables at time \( t \) and \( \beta_k \) 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 creating a set of decision trees.

These decision trees have the problem of getting overfitted easily; therefore, the random forest has a regularising 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:

\[
I = - \sum p_i \log_2(p_i)
\]

where \( p_i \) is the probability of a particular category of values. The decision tree reaches to a particular prediction by using a set of decisions learned 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 Fig. 2 (Fig. 3).

**Holt-Winters**

Holt-Winters prediction method has been categorised into multiplicative Holt-Winters and additive Holt-Winters.

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

\[
\hat{y}_t = a \frac{y_t}{S_{t-m}} + (1 - a)(\hat{y}_{t-1} + b_{t-1})
\]

\[
b_t = \beta^* + (\hat{y}_t - \hat{y}_{t-1})(1 - \beta^*)b_{t-1}
\]

\[
S_t = \gamma \left( \frac{y_t - \hat{y}_t}{\hat{y}_{t-1}} - b_{t-1} \right) + (1 - \gamma)S_{t-m}
\]

\[
F_{t+m} = (L_t + b_{t+m})S_{t-s+m}
\]

The component form for the additive method is:

\[
\hat{y}_t = a(y_t - S_{t-m})(1 = a)(\hat{y}_{t-1} + b_{t-1})
\]

**Support vector regression method**

Due to small samples with nonlinear 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, nonlinear, 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):

\[
f(x) = \omega \int p(x) + b
\]
Results

Decomposition of the time series record of direct emission of CO₂ from paddy in India is categorised into 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:

\[
\text{Additive model} = \text{Trend} + \text{seasonality} + \text{Random Noise}
\]

(14)

The additive model is useful when the seasonal variation is relatively constant over time. The multiplicative model is useful when the seasonal variation increases over time.

Forecasting test and error in the model

The analysis highlights that there has been continuing enhancement in direct, indirect, and total emission of CO₂ from paddy crops 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 and reduce the 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 CO₂ 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 Figs. 4, 5, and 6, respectively.

At the initial stage, SARIMAX and random forest regressor (RFR) techniques were applied to predict the CO₂ emission from 2019 to 2025 and the basis of training (1961–2011) and tested data (2016–2018). Holt-Winters and SVR models were also applied to compare the efficacy for CO₂ emission from paddy crops in India Fig. 7.

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. Tables 1 and 2 compare accuracy values and the forecasting results of direct emission of CO₂ from paddy crop in India for the period 2019 to 2025 of each method, respectively. Table 1
and Fig. 8 show that the actual CO2 emission from paddy forecasted CO2 emission, mean absolute percentage error (MAPE), and 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 also shows that MAPE and MSE values with H-W (0.75; 16,648.58) and SVR (0.91; 17,832.4) were relatively low compared to RFR and SARIMAX. Holt-Winters is a method for modelling three characteristics of a time series: the average value, the slope (trend) across time, and the cyclical repeating pattern (seasonality).

On the other hand, SVR is a machine learning model that aims to minimise the error by discovering a function that places more of the original points inside the tube created by the hyperplane and its border while lowering slack. The comparison was carried out to gather the best characteristics from both models and create a hybrid model. 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 CO2 from paddy crops in India. However, the accuracy level of Holt-Winter is relatively high as compared to any other model used in the current study, including SVR. The Holt-Winters model is one of the most popular forecasting algorithms (Trull et al., 2020) and was 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 CO2 from paddy in India with all the above four models from 2019 to 2025 in Table 2, but the finding of the current study suggests that H-W should be used to predict CO2 emission from paddy crops in India for more accurate results.

**Prediction of the total emission of CO2 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 CO2 from paddy crop in India is predicted to be stagnant around 12,307.2 Gg from 2019 to 2025 by RFR model. But, the accuracy results of the RFR model are not good, and the MAPE value of the RFR model is 5.67 (MSE: 549,900.02), which is relatively high compared to that of the SARIMAX, H-W, and SVR models.

**Prediction of the total emission of CO2 by SARIMAX**

The values of modelling estimates of SARIMAX model from 2019 to 2025 are shown in Table 2. The total emission of CO2 from paddy crop in India will be exceeding 13,000 Gg in 2021, and it is predicted to be 13,016.1 Gg in 2021 and also predicted to increase slowly to 13,484.9 Gg in 2025. The MAPE value of SARIMAX is recorded at 1.67 (MSE: 70,422.35).

**Prediction of the total emission of CO2 by Holt-Winters (H-W)**

The values of forecasting results of the Holt-Winters (H-W) model from 2019 to 2025 are shown in Table 2. The predictions of the total CO2 emission from paddy crop in India are 13,378.38; 13,115.06; 13,240.67; 13,387.51; 13,953.16; 13,954.15; and 14,364.97 from 2019 to 2025, respectively, given in Table 2. The MAPE value for 2019–2025 was 0.75 percent (MSE: 16,648.58) 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 CO2 emission from paddy in India. Reduction of CO2, CH4 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 (CO2) 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 CO2 by the support vector regression (SVR)

The support vector regression (SVR) prediction model forecasted total emission of CO2 from the paddy crop in India for 2018–2025 based on training and test data. The value of SVR model shows that total CO2 emission from paddy in India will reach 13,696.63 Gg in 2025. We obtained the mean absolute percentage error (MAPE) 0.91 and 17,832.4, respectively.

Discussion

Rice is the most important food crop for consumption worldwide. It is a staple crop for the majority of the population in the world. Rice is being cultivated in an area of 154 million ha worldwide (FAO, 2012), and the demand of use of rice will increase by around 24% in the next 20 years (Van Nguyen & Ferrero, 2006). However, the cultivation of rice crops adds to the most prominent greenhouse gases (GHGs) emission, mainly CH4 and N2O. Rice cultivation attributes around 30% and 11% of global agricultural CH4 and N2O
Fig. 7  H-W prediction results

Table 1  Results of fitting direct emission of CO2 from paddy from 2016 to 2018 in India

|       | Actual   | RFR      | APE | SARIMAX (0,1,1) APE | H-W APE | SVR APE | MAPE   | MSE    |
|-------|----------|----------|-----|---------------------|---------|---------|--------|--------|
| 2016  | 12,488.28| 11,594.3 | 7.16| 12,505              | 0.13    | 12,468.05| 0.16   | 12,636.13|
| 2017  | 12,778.5 | 12,275.2 | 3.94| 12,539.7            | 1.87    | 12,994.86| 1.69   | 12,753.97|
| 2018  | 13,047.98| 12,275.2 | 5.92| 12,655.6            | 3.01    | 12,995.78| 0.4    | 12,871.81|

MAPE 5.67  MAPE 1.67  MAPE 0.75  MAPE 0.91
MSE 549,900.02  MSE 70,422.35  MSE 16,648.58  MSE 17,832.4
emissions, respectively (NEPIS|US EPA, 2006). Anastasi et al. (1992) estimated that methane emissions from rice crops would be increased to 145 Tg year$^{-1}$ by 2025. Production of rice has been identified as a major source of GHGs emissions, such as CH$_4$, N$_2$O, and CO$_2$ (Zou et al., 2005; Babu et al., 2005; Linquist et al., 2012). Carbon dioxide (CO$_2$) emissions from paddy crops are a subject of discussion among policymakers. It is not possible to refuse that the idea of sustainable development stimulates their enquiry into this arena. The principal purpose of this work is to use the Holt-Winters (H-W) techniques to predict carbon dioxide emissions from paddy crop cultivation in India. The management of CO$_2$ emissions 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 causes 7 million live deaths worldwide yearly (United Nations, 2020). However, paddy is the most widely cultivated and consumed cereal in Asia 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 are various prediction models 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 were employed. These techniques were considered because they have the capability to make predictions of CO$_2$ emission. Furthermore, the result highlights that the change from paddy cultivation to other crop cultivation could be a possible way to reduce carbon dioxide emissions without sacrificing the loss of agriculture crops.

### Conclusion

It was seen in the results that there was an excellent output of SVR and Holt-Winters (H-W) as compared to SARIMAX and RFR. SVR and H-W’s experimental findings were 0.91 and 17,832.4 (MAPE & MSE) and 0.75 and 16,648.58 (MAPE & MSE), respectively. The prediction of CO$_2$ emission from paddy fields for 2025 is obtained at 13,696.67 (SVR) and 14,364.97 (H-W) in India. These SVR and Holt-Winters can be applied to estimate the total emission of CO$_2$ from paddy fields in India. However, the accuracy level of Holt-Winter is relatively high as compared to any other model used in the current study.

![Graph](image.png)

**Fig. 8** MSE and MAPE values
including SVR. Hence, the H-W should be used to predict CO2 emission from paddy crops. In this way, the model will contribute to filling India’s vast data gaps in quantum CO2 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 profitable based on evidence from India. Comparative analysis of this evidence on CO2 emissions from paddy fields would provide insight into paddy production’s relative effect on the expansion or the reduction of CO2 emissions.

Policy implication and further suggestion for research

The prediction of CO2 emission is difficult in Indian paddy fields due to its socio-economic status, different soil quality, and environment. Also, variation in fertiliser management and seeds management plays an important role in the emission of CO2 from paddy fields. The actual contribution of the agriculture sector in India to CO2 emission from paddy fields can only be addressed by suggesting an appropriate prediction model. It will predict the emission of CO2 in the coming year and suggest a baseline from which future emissions must be reduced and mitigation strategies. The major challenge worldwide is to produce more food and fibre to meet the increasing requirement for a nine billion population by 2050. Climate change is one of the big problems worldwide, and the rapid emission of CO2 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.

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.

Author contribution Conceptualisation, introduction, results writing, and data analysis: Pawan Kumar Singh and Sahil Ahuja. Literature review and analysis: Alok Kumar Pandey, Ravi Kiran. All authors have read and approved this manuscript.

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

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Competing interests The authors declare no competing interests.

References

Abbasi T, Abbasi T, Luiithui C, Abbasi SA (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

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

Alexandratos, N. (n.d.). World agriculture towards 2030/2050: the 2012 revision. 154.

Anastasi C, Dowding M, Simpson VJ (1992) Future CH4 emissions from rice production. Journal of Geophysical Research: Atmospheres 97(D7):7521–7525. https://doi.org/10.1029/92JD00157

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

Babu YJ, Li C, Frohling S, Nayak DR, Datta A, Adhya TK (2005) Modelling of methane emissions from rice-based production systems in India with the denitrification and decomposition model: field validation and sensitivity analysis. Curr Sci 89(11):9

Baydaroğlu Ö, Koçak K (2014) SVR-based prediction of evaporation including chaotic approach. J Hydrol 508:356–363. https://doi.org/10.1016/j.jhydrol.2013.11.008

Blagodatsky S, Smith P (2012) Soil physics meets soil biology: towards better mechanistic prediction of greenhouse gas emissions from soil. Soil Biol Biochem 47:78–92. https://doi.org/10.1016/j.soilbio.2011.12.015

Cai Z-C, Xing G-X, Shen G-Y, Xu H, Yan X-Y, Tsuruta H, Yagi K, Minami K (1999) Measurements of CH4 and N2O emissions from rice paddies in Fengqiu. China Soil Science and Plant Nutrition 45(1):1–13. https://doi.org/10.1080/00380768.1999.10409320

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

Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw 17(1):113–126. https://doi.org/10.1016/S0893-6080(03)00169-2

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297. https://doi.org/10.1007/BF00994018

Cui B, Xue T, Yang K (2008) Vehicle recognition based on support vector machine. International Symposium on Intelligent Information Technology Application Workshops 2008:443–446. https://doi.org/10.1109/IITA.Workshops.2008.23

Document DisplayNEPISUS EPA. (2021). Retrieved October 9, 2021, from https://nepis.epa.gov/Exe/ZyNET.exe/2000ZLSG.TXT?
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.jcous.2016.07.009

Methane emission from rice paddy fields in all of Japanese prefecture. SpringerLink. (n.d.). Retrieved January 27, 2021, from https://link.springer.com/article/https://doi.org/10.1038A.1009778517545

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. Agr Ecosyst Environ 116(3):181–188. https://doi.org/10.1016/j.agee.2006.02.006

AshinNishan MK, Muhammad Ashiq V (2020) Role of energy use in the prediction of CO2 emissions and economic growth in India: evidence from artificial neural networks (ANN). Environ Sci Pollut Res 27(19):23631–23642. https://doi.org/10.1007/s11356-020-08675-7

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

Nynø, T., & Bonga, W. G. (2019). Prediction of CO2 emissions in India using ARIMA models (SSRN Scholarly Paper ID 3346378). Social Science Research Network. https://papers.ssrn.com/abstract=3346378

Oo AZ, Sudo S, Inubushi K, Mano M, Yamamoto A, Ono K, Osawa T, Hayashida S, Patra PK, 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. Agr Ecosyst Environ 252:148–158. https://doi.org/10.1016/j.agee.2017.10.014

Paraschiv D, Tudor C, Petraru 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

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

Rcs-21a.pdf. (n.d.). Retrieved January 27, 2021, from https://www.ers.usda.gov/webdocs/outlooks/100237/rcs-21a.pdf?v=4975.2

Rahman A, Hasan MM (2017) Modeling and forecasting of carbon dioxide emissions in Bangladesh using autoregressive integrated moving average (ARIMA) models. Open J Stat 7(4):560–566. https://doi.org/10.4236/ojs.2017.74038

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

Saleh C, Dzakiyullah NR, Nugroho JB (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

Shuai Yang Yu, Wang WA, Bai Y, Li C (2018) Prediction and analysis of CO2 emission in Chongqing for the protection of environment and public health. Int J Environ Res Public Health 15(5):530. https://doi.org/10.3390/ijerph15050350

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

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 CO₂ enrichment (FACE) and soil warming on CH₄ 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

Trull O, García-Díaz JC, Troncoso A (2020) Initialization methods for multiple seasonal Holt-Winters forecasting models. Mathematics 8(2):268. https://doi.org/10.3390/math8020268

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

Vagropoulos SI, Chouliaras GI, Kardakos EG, Simoglou CK, Bakirtzis AG (2016) Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. IEEE International Energy Conference (ENERGYCON) 2016:1–6. https://doi.org/10.1109/ENERGYCON.2016.7514029

Van Nguyen N, Ferrero A (2006) Meeting the challenges of global rice production. Paddy Water Environ, 4(1):1–9. https://doi.org/10.1007/s10333-005-0031-5

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

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

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

Xu D, Zhang Q, Ding Y, Zhang D (2021) Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-021-15325-z

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. J Adv Transp 2020:1–14. https://doi.org/10.1155/2020/8848149

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.