



OPEN ACCESS

EDITED BY

Guochang Fang,
Nanjing University of Finance and Economics,
China

REVIEWED BY

Yan Tan,
Yulin Normal University, China
Jiyi XU,
Anhui University of Technology, China

*CORRESPONDENCE

Sheng Fang
✉ rwzp_aust@126.com
Meiling Zhao
✉ meiling.zhao.cn@outlook.com

RECEIVED 23 December 2024

ACCEPTED 27 February 2025

PUBLISHED 25 March 2025

CITATION

Shao J, Fang S, Zhao M, Qian W and
Wang C (2025) Understanding the
environmental health implications of tourism
on carbon emissions in China.
Front. Public Health 13:1550395.
doi: 10.3389/fpubh.2025.1550395

COPYRIGHT

© 2025 Shao, Fang, Zhao, Qian and Wang.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Understanding the environmental health implications of tourism on carbon emissions in China

Jinhua Shao¹, Sheng Fang^{1*}, Meiling Zhao^{2*}, Wanxin Qian¹ and
Cai Wang¹

¹School of Humanities and Social Sciences, Anhui University of Science and Technology, Huainan, China, ²College of Tourism, Hainan Tropical Ocean University, Sanya, China

Tourism development is important for the formulation of the national carbon reduction policy. China has put forward the goals of carbon peaking and carbon neutrality. Studying the impact of China's tourism industry on carbon emissions is of great significance in scientifically formulating emission reduction policies and helping China to realize its carbon reduction goals. In this study, we simulate the complex relationship between the tourism industry and carbon emissions in China using machine learning models. This study is the first to employ interpretable machine learning to analyze the impact of the tourism industry on carbon emissions in China. Our findings demonstrate that sparrow search algorithm and random forest (SSA-RF) hybrid model can model the relationship between carbon emissions and tourism factors with low error. The expansion of the tourism industry positively contributes to the increase in carbon emissions. Our study highlights the need to consider tourism factors when formulating national carbon reduction policy.

KEYWORDS

tourism industry, carbon emission, machine learning, decoupling index, environmental health

1 Introduction

Sea level rise and extreme weather events caused by global warming are a serious threat to the sustainable development of society (1, 2), and reducing carbon emissions is essential for controlling global warming. China's role in mitigating global change and reducing emissions has become even more important after the United States withdrew from the Paris Agreement (3, 4). China is the world's largest carbon emitter (5), and the Chinese Government is actively taking measures to reduce emissions and has formulated targets for such reductions (6, 7). China proposes to achieve carbon peaking by 2030 and carbon neutrality by 2050. These targets are not only significant for China, but also play an important role in slowing down global warming.

According to a report released by the International Energy Agency, carbon emissions from the tertiary industry grew at a rate of up to 50.43% from 2011 to 2018, which is an important contributor to the growth of carbon emissions, and the growth rate of carbon emissions from the tertiary industry is much larger than that of the primary and secondary industries. Tourism is an important part of the tertiary industry (8), and the 2021 United Nations Climate Change Conference formally passed the Glasgow Declaration, which emphasized practical and strong measures to help achieve the global commitment to zero carbon emissions by 2050. Most of the analysis of the factors affecting carbon emissions in the tourism industry uses decomposition analysis (9–11), which includes structural decomposition analysis (SDA) and index decomposition analysis (IDA). Using decomposition

method, total domestic tourists' economic yield proved to be the largest driver of carbon emission growth in China's tourism industry, and total international tourists' economic yield also contributed to the growth of carbon emissions in China's tourism industry. Energy consumption intensity and energy consumption structure, on the other hand, are important factors in slowing down the carbon emissions of the tourism industry (9). There are other factors that have also contributed the growth of carbon emissions in China's tourism industry, such as population and economic development (12). The factors influencing carbon emissions from the tourism industry are relatively well researched (13).

In recent years, more and more scholars have begun to focus on how tourism affects a region's total carbon emissions, but there are no conclusive findings on how tourism affects a region's total carbon emissions. Existing studies show that tourism reduces carbon emissions in Europe, but in Asia it contributes to the growth of regional carbon emissions. According to Lee and Brahmaire (14), the tourism industry in Europe exerts a significant negative impact on carbon emissions. At the global scale, most studies show that tourism contributes to global carbon emissions (15). Air transport and visitor spending have hindered Australia's progress toward achieving net-zero emissions (16). The relationship between tourism and regional carbon emissions is complex (17, 18). Tourism development has an obvious role to play in boosting a region's economy, but tourism development may either promote or inhibit regional carbon emissions. Therefore, it is necessary to analyze the impact of tourism development on regional carbon emissions before deciding whether or not to mitigate carbon emissions by developing tourism. Currently, analyzing the impact of tourism on the region's carbon emissions also mainly uses decomposition analysis. Visas et al. (19) analyzed the impact of tourism on energy consumption in the BRICS using decomposition method, the results reveal that the tourism promoted the energy consumption.

Over the past several years, the influence of artificial intelligence (AI) has expanded rapidly (20), and AI has been widely used in establishing carbon emission prediction models (21), with numerous studies demonstrating that AI models outperform traditional approaches in carbon emission forecasting (22). This is due to the fact that AI models are better able to capture the complex relationship between carbon emissions and influencing factors than traditional regression models. Despite the high prediction accuracy of the AI models, the models are poorly interpretable and have difficulty explaining the effects of variables on carbon emissions. Many cutting-edge studies in the fields of chemistry (23), biology, and materials science (24) use machine learning models to explain the results of their studies, which are referred to as interpretable machine learning models in research. Interpretable machine learning models are not only a data fitting tool but also an analytical tool in research. We believe that interpretable machine learning techniques have great potential to reveal the complex relationship between tourism and carbon emissions. Although many previous papers have examined the relationship between carbon emissions and tourism, there are few studies using interpretable machine learning, and we are interested in filling a gap. Overall, the main goal of this paper is to analyze the impact of tourism factors on carbon emissions using interpretable ML, emphasizing that the development of tourism should also focus on the impact on global warming.

2 Methodology

2.1 Data sources

China's carbon emissions data for 2000–2019 is from BP's World Energy Statistics. The data on international tourism revenue (ITR), number of international tourist arrivals (ITA), foreign exchange earnings from international tourism transportation (ITT), total domestic tourism expenditure (DTE), and number of domestic tourists (DT) are sourced from the National Bureau of Statistics of China. We processed some of the missing data using autoregressive integrated moving average (ARIMA) method. ARIMA is a widely used statistical model for time series analysis and forecasting, and it is also widely used in data interpolation (25). Compared to other interpolation methods such as long short-term memory (LSTM) and conditional generative network (CNN), the ARIMA model is highly interpretable, whereas the internal mechanisms of LSTM and CNN are complex, and the LSTM and CNN models require a large amount of data to achieve good prediction accuracy. For this study, the data size is small, so we chose ARIMA for missing data.

2.2 Random forest

Random forest (RF) is a commonly used machine learning algorithm for regression prediction (26), which achieves high accuracy by constructing multiple decision trees and aggregating their prediction results. Compared to conventional deep learning algorithms, RF does not require large amounts of data and trains quickly. It has fewer hyperparameters, which are easier to adjust, whereas deep learning requires extensive time and data to train hyperparameters. Given the small scale of data in this study, using deep learning algorithms could lead to overfitting. This study uses the RF method to model the relationship between carbon emissions and the tourism industry in China. The training data for each base learner in RF is obtained through the bootstrap method, which involves randomly selecting a subset of the overall features, with the rest of the data constituting the out-of-bag samples (OOB) sample. The randomness of the algorithm is demonstrated in two ways: firstly, the training data for each tree is selected by random sampling; secondly, the tree features selected are also chosen randomly. This approach not only prevents model overfitting, but also enhances the differences between individual decision trees by adding randomness to the training process. With these steps, the constructed forest can improve the prediction accuracy by averaging the predictions of the trees as shown in the Equation 1:

$$\bar{f} = \frac{1}{n} \sum_{i=1}^n T(x, O_i) \quad (1)$$

Where \bar{f} is the average, $T(x, O_i)$ is the output of each tree and n represent the trees.

2.3 Optimization algorithm

The particle swarm algorithm (PSO), an optimization algorithm developed by simulating the foraging behavior of birds, is a simple and

easy-to-implement optimization algorithm that is widely used in training machine learning models for hyperparameter optimization. The algorithm proposes the concept of particles to simulate birds in a flock, where particles learn and exchange information among themselves to achieve a globally optimal search. Each particle in the particle swarm algorithm has its own position and velocity, and the velocity of each example is updated according to Equations 2–4:

$$v_j^{(i+1)} = wv_j^{(i)} + c_1r_1(pbest_j - x_j^{(i)}) + c_2r_2(gbest_j - x_j^{(i)}) \quad (2)$$

$$v_{\min} \leq v_j^{(i+1)} \leq v_{\max} \quad (3)$$

$$x_j^{(i+1)} = x_j^{(i)} + v_j^{(i+1)} \quad j = 1, 2, \dots, n \quad (4)$$

where $v_j^{(i)}$ is the velocity of particle j at the i -th iteration, and is w the inertia weight. c_1 and c_2 are the learning factors, which control how close the particle is to its own optimal position $pbest_j$ and the global optimal position $gbest_j$. r_1 and r_2 is a random number between 0 and 1 to increase the randomness of the algorithm. $x_j^{(i)}$ is the position of the particle at the i -th iteration. v_{\min} and v_{\max} are the minimum and maximum velocities of the particle. Equation 4 is the formula for particle position update.

The whale optimization algorithm (WOA) is a swarm optimization algorithm proposed by simulating the feeding behavior of humpback whales. The whale optimization algorithm consists of three steps surrounding the prey, bubble net attack method and prey search.

The behavior of surrounding prey is simulated by Equations 5–8:

$$\bar{D} = \left| \bar{C} \cdot \bar{X}^*(t) - \bar{X}(t) \right| \quad (5)$$

$$\bar{X}(t+1) = \bar{X}^*(t) - \bar{A} \cdot \bar{D} \quad (6)$$

$$\bar{A} = 2\bar{a} \cdot \bar{r} - \bar{a} \quad (7)$$

$$\bar{C} = 2 \cdot \bar{r} \quad (8)$$

Where $\bar{X}^*(t)$ is the optimal solution (prey position), and $\bar{X}(t)$ is the location of the whale. \bar{A} and \bar{C} are coefficient vectors that change dynamically as the algorithm iterates. \bar{a} is a factor that decreases linearly from 2 to 0, and \bar{r} is a random forest between 0 and 1. When $|\bar{A}|$ is less than 1, the whale moves closer to the prey, and when $|\bar{A}|$ is greater than 1, the whale moves away from the prey, thus performing a global search.

The bubble net attack is the algorithm and core to simulate the bubble net hunting behavior of whales (Equation 9).

$$\bar{X}(t+1) = \bar{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \bar{X}^*(t) \quad (9)$$

where \bar{D}' denotes the current distance between the whale and the prey, b is a constant, and l is a random number between -1 and 1 .

To avoid a local optimal solution, the whale population performs a global search by randomly selecting another whale and moving away from it (Equation 10):

$$\bar{X}(t+1) = \bar{X}_{rand} - \bar{A} \cdot \bar{D} \quad (10)$$

where $\bar{X}(t+1)$ is the location of a randomly selected whale.

Sparrow search algorithm (SSA) is an intelligent optimization algorithm proposed based on the foraging and anti-predator behavior of sparrows. In the foraging process of sparrows, they are divided into finders and joiners. The discoverers are responsible for providing foraging areas and directions for the entire sparrow flock, while the joiners rely on the guidance of the discoverers for food. In addition, when a sparrow flock becomes aware of danger, it triggers anti-predatory behavior. Discoverers have better adaptations and therefore finders can gain a larger search range than joiners. The finder's location update method can be described as follows (Equation 11):

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp\left(-\frac{i}{s \cdot t_{\max}}\right) & Y_2 < AN \\ X_{i,j}^t + ZD & Y_2 \geq AN \end{cases} \quad (11)$$

where $X_{i,j}^{t+1}$ is the position of the sparrow i in dimension j and t_{\max} is the maximum number of iterations. s is a random number between 0 and 1. Y_2 denotes the warning value, AN denotes the safety value, the value range of Y_2 and AN is 0–1 and 0.5–1, respectively. Z is a random number with normal distribution. D is the $1 \times d$ matrix, d is the dimension of the optimization problem, and all elements of the matrix are 1. When $Y_2 < AN$, it indicates that there are no predators in the range. The finder can perform a wide search operation. When $Y_2 \geq AN$, it indicates that a predator has been found. At this time, all sparrows need to quickly fly to other safe places. The method of updating joiner locations is described as follows (Equation 12):

$$X_{i,j}^{t+1} = \begin{cases} Z \cdot \exp\left(\frac{X_w - X_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| A^+ \cdot D & else \end{cases} \quad (12)$$

where X_p is the current best position of the discoverer, and X_w is the worst position of the current discoverer. A is a $1 \times d$ matrix where each element of the matrix is randomly assigned a value of 1 or -1 and $A^+ = A^T(AA^T)^{-1}$, n the number of sparrows in the flock. When $i > n/2$, it means that the joiner i has a low fitness value and is very hungry and needs to fly elsewhere to forage for food. When aware of the danger, sparrow populations perform an anti-predatory behavior, the mathematical expression for which is (Equation 13):

$$X_{i,j}^{t+1} = \begin{cases} X_b^t + \beta |X_{i,j}^t - X_b^t| & f_i > f_g \\ X_{i,j}^t + G \left[\frac{|X_{i,j}^t - X_w^t|}{(f_i - f_w) + \varepsilon} \right] & f_i = f_g \end{cases} \quad (13)$$

Where X_b is the global optimal position, β is the parameter controlling the step size and obeys a normally distributed random number with mean 0 and variance 1. G is a random number between -1 and 1 , f_i is the fitness value of the sparrow i , and f_g and f_w are the global optimum and global worst value. f_i is the smallest constant that avoids a denominator of 0 in equation.

2.4 SHapley Additive exPlanation

The aim of interpretable machine learning is to learn how the model make predictions, and find the relationship between the input and output, more importantly, to answer the question of which input feature is important for driving prediction results. Model interpretable method can be divided into two categories, local interpretable method and global interpretable method. Global interpretation can explain the importance distribution of features across the entire dataset, while local interpretation allows for a detailed analysis of the contribution of each feature to the prediction of a particular observation. SHapley Additive exPlanation (SHAP) enables not only local interpretation but also global interpretation, and SHAP makes use of the Shapley value concept in game theory, and by accurately calculating the contribution of each feature to the model output, it directly demonstrates the weight and influence of each feature in the model prediction (27). For feature i in the feature set S , the Shapley value is calculated as follows (Equation 14):

$$\Phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S)) \quad (14)$$

where N denotes the set of all features, and S is any subset of features that does not contain feature i . $|S|$ is the number of features in the set S . $v(S)$ is the contribution of the feature set S to the predicted output of the model. $v(S \cup \{i\})$ is the contribution of the feature set $S \cup \{i\}$ containing feature i to the predicted output of the model.

2.5 Hybrid model

SHAP is model-agnostic, so it can be applied to a wide range of machine learning models. The calculation of SHAP values is not directly related to the accuracy of the model, but the performance of the model may indirectly affect the stability of SHAP values. If the accuracy of model is high, it means that the model is performing well in capturing the relationship between features and outputs. As a result, the SHAP value is more reflective of the true impact of the features on the model’s predictions. In other words, in high accuracy models, SHAP values are more interpretable and results are more credible. Therefore, in this paper, the RF model is first optimized using the swarm optimization algorithm (PSO, WOA and SSA) to improve the prediction accuracy of the RF model and to ensure the

stability of the subsequent SHAP analysis. We optimized the hyperparameters (n_estimators and max_features) of the RF model using PSO and SSA, and three hybrid model (PSO-RF, WOA-RF and SSA-RF) are established. The flowchart of three hybrid models is shown in Figure 1.

2.6 Tapio decoupling index

The carbon emission decoupling index is an index that measures the relationship between economic growth and carbon emissions (28). It aims to evaluate the relationship between economic growth and greenhouse gas emissions in a country or region, i.e., whether economic growth is accompanied by a corresponding reduction in carbon emissions (29). Currently, the Tapio model is mostly used in the analysis of decoupling relationship. Compared with the previous OECD model (30), Tapio can more comprehensively reflect the multiple and complex relationships between environmental pressure and economic development, which can better help policy makers to develop feasible strategies to achieve a balance between economic growth and environmental protection. According to the definition given Tapio, the tourism carbon emission decoupling index can be calculated as follows (Equation 15):

$$T = \frac{\% \Delta CE}{\% \Delta E} = \frac{(CE_t - CE_{t-1}) / CE_{t-1}}{(E_t - E_{t-1}) / E_{t-1}} \quad (15)$$

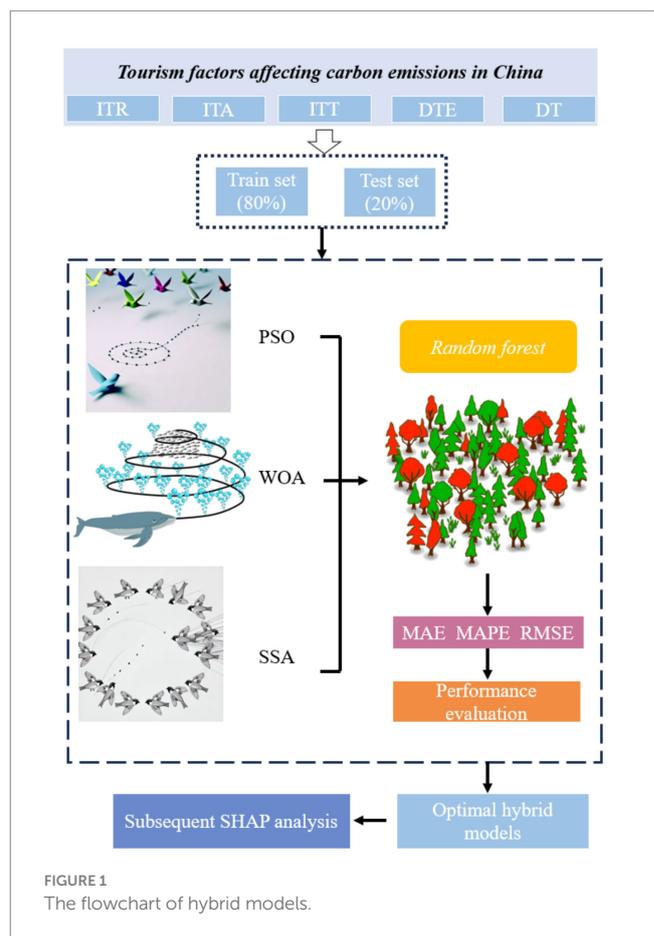


FIGURE 1 The flowchart of hybrid models.

TABLE 1 Decoupling index classification criteria.

State		%ΔCE	%ΔE	T
Decoupling	Strong decoupling	< 0	> 0	T < 0
	Weak decoupling	> 0	> 0	0 ≤ T < 0.8
	Recessive decoupling	< 0	< 0	T > 1.2
Coupling	Expansive coupling	> 0	> 0	0.8 ≤ T ≤ 1.2
	Recessive coupling	< 0	< 0	0.8 ≤ T ≤ 1.2
Negative decoupling	Expansive negative decoupling	> 0	> 0	T > 1.2
	Weak negative decoupling	< 0	< 0	0 ≤ T < 0.8
	Strong negative decoupling	> 0	< 0	T < 0

Where CE is the carbon emission of China, and E is the revenue from tourism industry. ΔCE and ΔE are the change rates of carbon emissions and tourism revenues in China. The relationship between carbon emissions and tourism revenue can be divided into three major types of states: coupling, decoupling and negative decoupling, as shown in Table 1.

3 Results and discussion

3.1 Optimal hybrid model selection

Based on existing research and data availability (31–34), we use international tourism revenue (ITR), number of international tourist arrivals (ITA), foreign exchange earnings from international tourism transportation (ITT), total domestic tourism expenditure (DTE), and number of domestic tourists as the input to the RF model, and China’s carbon emissions are the output of the model. N_estimators and max_features are two important parameters of the RF model. Increasing the value of n_estimators usually improves the stability and performance of the model, as more trees means that the model can fit the data better. More trees also increase training time and memory consumption. Max_features values usually increase the diversity of the model, thus reducing the risk of overfitting. Small max_features value may result in a decrease in the performance of each tree, affecting the overall model performance. Conventional grid search is time-consuming, so in this study, three swarm optimization algorithms were used to optimize the parameters of RF to obtain highly accurate models. When training the model, we use 80% of the dataset as the train set and 20% of the dataset as the test set. The division of the train and test sets is random. Figure 2 shows the prediction results of the three hybrid models and the unoptimized RF model on the test set.

As shown in Figure 2, the prediction results of the SSA-RF and WOA-RF models have the smallest deviation from the actual value of carbon emissions, which indicates that the two models are better able to simulate the relationship between carbon emissions and tourism indicators. The PSO-RF has large errors in predicting carbon emission values for 2001 and 2006, but the deviation between the model predictions and the actual values is small when predicting carbon emissions for 2015 and 2017. Whereas RF is the model with the largest deviation in prediction results among the four models, the model does not work well in each year of the test set.

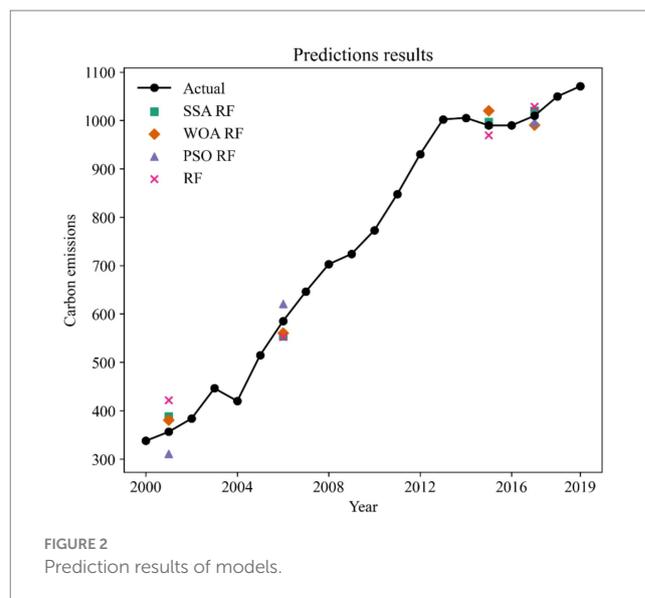


FIGURE 2 Prediction results of models.

In order to quantitatively compare the prediction accuracy of different machine learning models, we choose Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) as the evaluation metrics of model performance. Evaluation metrics are calculated based on the prediction results of different machine learning prediction models, and the evaluation metrics are compared to measure the strengths and weaknesses of the prediction models. The formula for calculating the evaluation metrics are as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |E_p - E_a| \tag{16}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|E_p - E_a|}{|E_a|} \times 100\% \tag{17}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_p - E_a)^2} \tag{18}$$

In Equations 16–18, E_p and E_a are the predicted and actual values of machine learning models. The calculation results of evaluation metrics are shown in Table 2.

MAE represents the mean absolute difference between the predicted value and the actual value. According to the MAE, the prediction accuracy is ranked from highest to lowest as SSA-RF > PSO-WOA-RF > PSO-RF > RF. The prediction accuracies of the hybrid models are all greatly improved compared to the single RF. The prediction accuracies of models SSA-RF, WOA-RF, and PSO-RF are improved by 1, 2, and 3%. An important feature of MAE is that it is insensitive to outliers because it only focuses on the mean of the absolute error, so we need other metrics to evaluate the model comprehensively. According to the MAPE, the prediction accuracy ranked from highest to lowest is SSA-RF > WOA-RF > PSO-RF > RF. SSA-RF has the smallest MAPE value indicating that it has the best

prediction performance. SSA-RF is the optimal model can also be found based on the value of RMSE. In Figure 2. We can see that both SSA-RF and WOA-RF possess good prediction results, but based on the evaluation metrics we are able to determine that SSA-RF is the best prediction model.

3.2 The influence of tourism factors on China's carbon emissions

To analyze the influence of tourism industry on carbon emissions in China, scatter plots of SHAP values are plotted as shown in Figure 3.

This figure of SHAP values shows the impact of different features on the carbon emissions. In the figure of SHAP values, the horizontal axis indicates the SHAP value (effect on carbon emission) and the vertical axis indicates the different features (DT, ITT, DTE, ITR, ITA). The color from red to blue indicate the feature values from high to low. Negative values of SHAP indicate that the feature has a decreasing impact on the carbon emissions, while positive values indicate an increasing impact. The SHAP values for DT, ITT, DTE, and ITR are uniformly distributed across both the negative and positive sides of the scale, indicating that these features can either

positively or negatively influence model predictions. The uniform variation of SHAP values with changes in DT, ITT, DTE, and ITR suggests a linear effect of these features on carbon emissions in China. This uniformity implies that as the values of these features increase or decrease, their impact on the model's prediction of carbon emissions consistently follows a linear trend, either enhancing or reducing the predicted emissions. On the other hand, the SHAP values for international tourist arrivals (ITA) are predominantly found in the positive region. This concentration in the positive spectrum indicates that ITA significantly contributes to increasing the predicted value of the model's output, suggesting a positive relationship between ITA and carbon emissions in China.

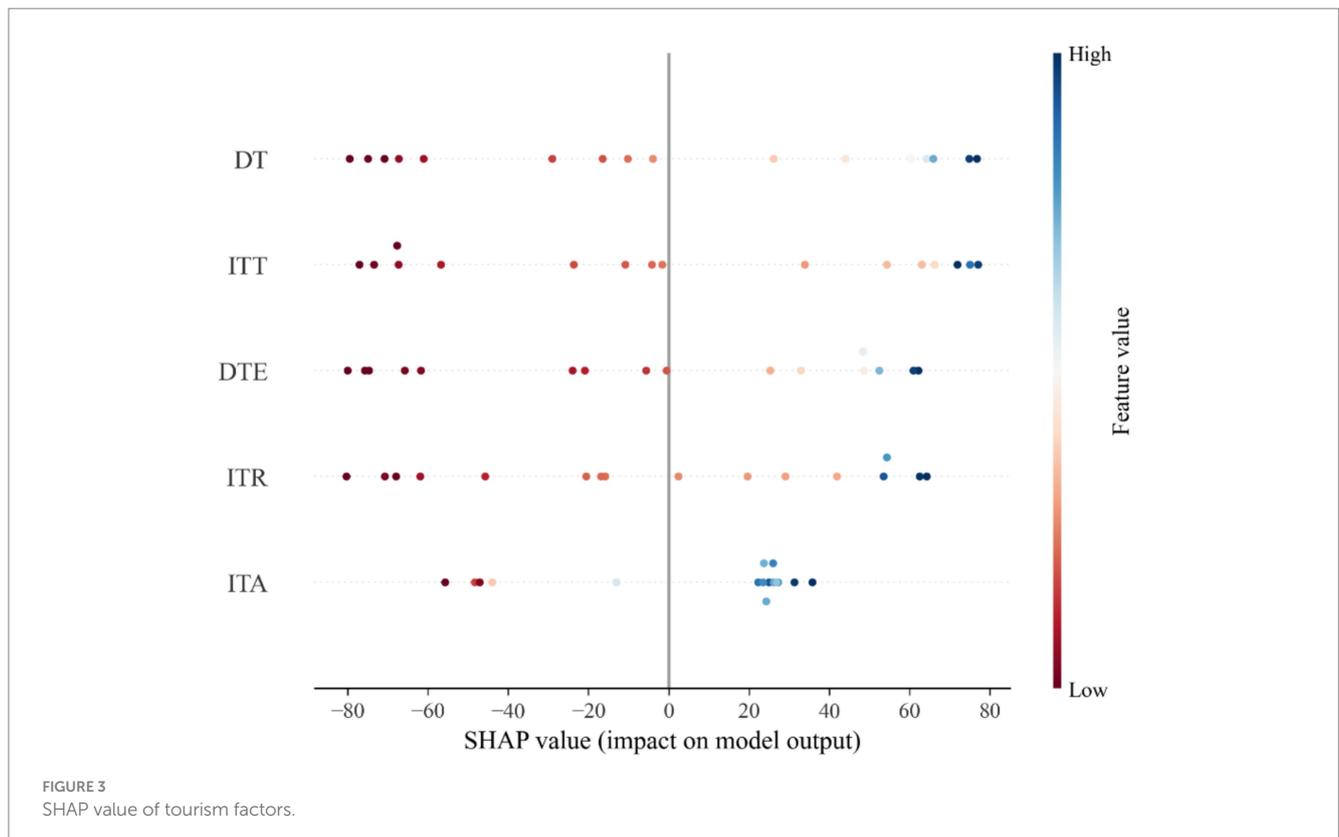
Figure 4 shows the average SHAP values for DT, ITT, DTE, ITR and ITA. The largest values for DT and ITT indicate that DT and ITT drive the growth of carbon emissions in China, and the smallest SHAP value for ITA indicates that ITA has the least impact on the growth of carbon emissions in China. The smallest impact of ITA on China's carbon emissions can also be observed in Figure 3, where the SHAP value of ITA is closest to the 0-axis.

3.3 Decoupling analysis between tourism industry and carbon emissions

According to the Equation 14, the decoupling index of carbon emissions and tourism in China from 2000 to 2019 is calculated, and the results are shown in Figure 5. Weak decoupling in 2001 and 2002 indicates that tourism revenue is growing faster than carbon emissions, but carbon emissions are still increasing. The shift from strong negative decoupling to strong decoupling between 2003 and

TABLE 2 Evaluation metrics of models.

	SSA-RF	WOA-RF	PSO-RF	RF
MAE	20.21	24.69	24.29	34.19
MAPE	4	4.01	5.12	6.94
RMSE	23.12	25.02	29.70	38.96



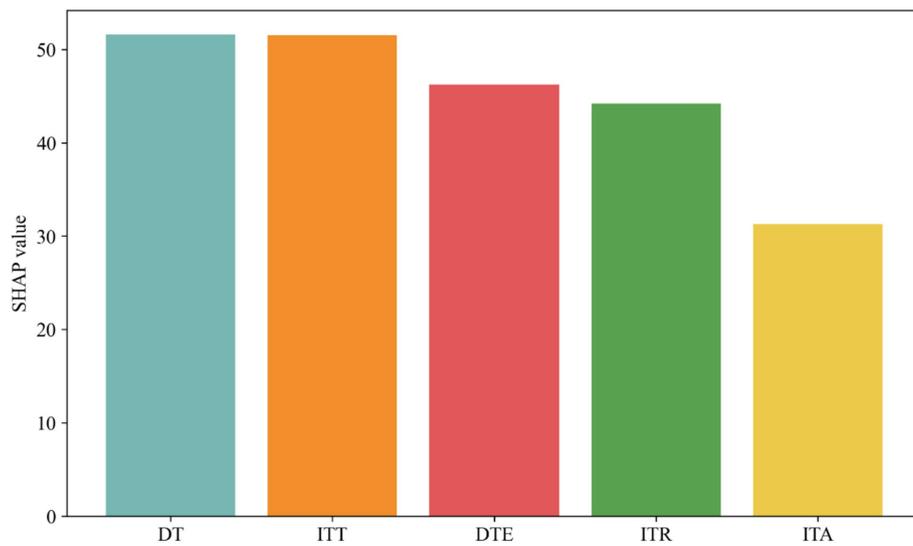


FIGURE 4 SHAP value of DT, ITT, DTE, ITR, ITA.

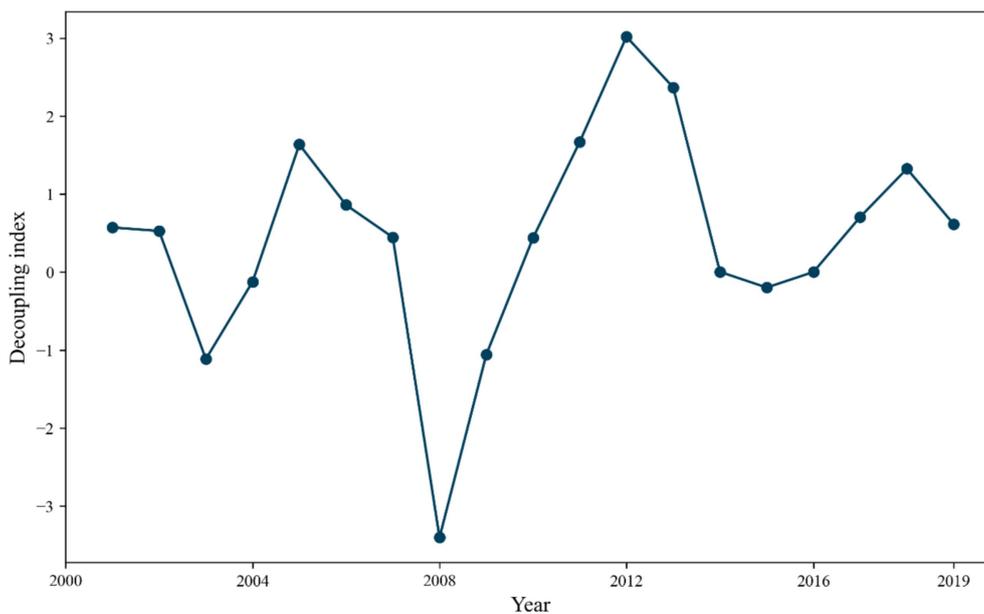


FIGURE 5 Decoupling index from 2000 to 2019.

2004 suggests that economic policies or environmental protection measures during this period promoted economic growth while reducing carbon emissions. Both decoupling and coupling states existed from 2005 to 2009, and in 2009, the decoupling state was a strong negative decoupling, illustrating the fact that tourism revenues were decreasing at the same time that carbon emissions were increasing, and that the policies of this period seriously threatened the sustainability of society.

From 2010 to 2019, the decoupling state shows a steady weak decoupling, and the growth rate of tourism revenue is greater than the increase rate of carbon emissions in this period, but policy guidance

should also be strengthened to focus on environmental protection while boosting tourism revenue.

4 Conclusions and policy recommendations

4.1 Conclusion

This paper firstly compares the results of different machine learning models in predicting China's carbon emissions, and then

analyzes the influence of international tourism revenue, number of international tourist arrivals, foreign exchange earnings from international tourism transportation, total domestic tourism expenditure, and number of domestic tourists on China's carbon emissions using SHAP values. Finally, the decoupling index is used to analyze the decoupling relationship between tourism revenue and carbon emissions in China. The main conclusions are as follows:

SSA-RF is able to capture the complex relationship between tourism factors and China's carbon emissions, and the predictions of SSA-RF model have the lowest MAE, MAPE and RMSE. Domestic tourists and international tourism transportation have the largest SHAP values, and these two factors have the greatest impact on changes in China's carbon emissions. According to the SHAP value analysis, the growth of China's tourism industry shows a linear relationship with the increase in China's carbon emissions, i.e., the expansion of the tourism industry is directly proportional to the increase in carbon emissions. Over the past decade, China's tourism revenues have continued to grow, while at the same time the trend of increasing carbon emissions has slowed down, showing a decoupling between tourism revenues and carbon emissions.

4.2 Policy implications

Considering the role of tourism in contributing to the growth of China's carbon emissions, China should strongly promote a low-carbon transition in the tourism industry. Firstly, the accounting of carbon emissions from tourism should be strengthened, and the tourism industry should specify clear goals and paths for low-carbon development. Actively develop low-carbon and environmentally friendly tourism products and strengthen low-carbon and environmental protection education during tourists' travels.

Transportation plays an extremely important role in the low-carbon transition of tourism, as confirmed in our study. Governments should encourage the use of public transportation by tourists, for example, by installing bicycle rental stations and walking paths in tourist destinations, and by improving orientation information to help tourists make better use of these services. In regional tourism, resources can be shared through multi-regional cooperation to jointly promote low-carbon transportation strategies, such as the construction and operation of cross-regional public transportation systems.

The transportation process for foreigners traveling into China also has a significant impact on China's carbon emissions, and carbon emissions from air transits can be reduced by cooperating on direct flights to popular tourist destinations. Increase multilingual signage on public transportation and actively guide foreign tourists to use public transportation.

References

- Shu EG, Porter JR, Hauer ME, Sandoval Olascoaga S, Gourevitch J, Wilson B, et al. Integrating climate change induced flood risk into future population projections. *Nat Commun.* (2023) 14:7870. doi: 10.1038/s41467-023-43493-8
- Marsooli R, Lin N, Emanuel K, Feng K. Climate change exacerbates hurricane flood hazards along US Atlantic and gulf coasts in spatially varying patterns. *Nat Commun.* (2019) 10:3785. doi: 10.1038/s41467-019-11755-z
- Shan Y, Huang Q, Guan D, Hubacek K. China CO2 emission accounts 2016–2017. *Sci Data.* (2020) 7:54. doi: 10.1038/s41597-020-0393-y
- Terhaar J, Frölicher TL, Aschwanden MT, Friedlingstein P, Joos F. Adaptive emission reduction approach to reach any global warming target. *Nat Clim Chang.* (2022) 12:1136–42. doi: 10.1038/s41558-022-01537-9
- Friedlingstein P, et al. Global carbon budget 2019. *Earth Sys Sci Data.* (2019) 11:1783–838. doi: 10.5194/essd-11-1783-2019
- Li L, Zhang Y, Zhou T, Wang K, Wang C, Wang T, et al. Mitigation of China's carbon neutrality to global warming. *Nat Commun.* (2022) 13:5315. doi: 10.1038/s41467-022-33047-9

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

JS: Conceptualization, Writing – original draft, Writing – review & editing. SF: Conceptualization, Writing – review & editing. MZ: Data curation, Writing – review & editing. WQ: Formal analysis, Writing – review & editing. CW: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by National Social Science Fund of China (No. 23BF084), Research Project of Anhui University of Science & Technology (No. 2023yjrc30), Anhui Provincial Social Science Innovation and Development Research Project (No. 2024CXQ052) and Hainan Tropical Ocean University (No. RHDRC202211).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

7. Shan Y, Guan D, Hubacek K, Zheng B, Davis SJ, Jia L, et al. City-level climate change mitigation in China. *Sci Adv.* (2018) 4:eaq0390. doi: 10.1126/sciadv.aaq0390
8. Lenzen M, Sun YY, Faturay F, Ting YP, Geschke A, Malik A. The carbon footprint of global tourism. *Nat Clim Chang.* (2018) 8:522–8. doi: 10.1038/s41558-018-0141-x
9. Luo F, Moyle BD, Moyle CLJ, Zhong Y, Shi S. Drivers of carbon emissions in China's tourism industry. *J Sustain Tour.* (2020) 28:747–70. doi: 10.1080/09669582.2019.1705315
10. Zha J, Fan R, Yao Y, He L, Meng Y. Framework for accounting for tourism carbon emissions in China: an industrial linkage perspective. *Tour Econ.* (2021) 27:1430–60. doi: 10.1177/1354816620924891
11. Ma H, Liu J, Xi J. Decoupling and decomposition analysis of carbon emissions in Beijing's tourism traffic. *Environ Dev Sustain.* (2022) 24:5258–74. doi: 10.1007/s10668-021-01657-w
12. Mou D. Tourism energy consumption estimation and driving factors of carbon emissions based on LMDI and panel data models. *J Comput Methods Sci Eng.* (2024) 24:1839–49. doi: 10.3233/JCM-230007
13. Zha J, Dai J, Ma S, Chen Y, Wang X. How to decouple tourism growth from carbon emissions? A case study of Chengdu, China. *Tour Manag Perspect.* (2021) 39:100849. doi: 10.1016/j.tmp.2021.100849
14. Lee JW, Brahmasrene T. Investigating the influence of tourism on economic growth and carbon emissions: evidence from panel analysis of the European Union. *Tour Manag.* (2013) 38:69–76. doi: 10.1016/j.tourman.2013.02.016
15. Sun Y-Y, Gossling S, Zhou W. Does tourism increase or decrease carbon emissions? A systematic review. *Ann Tour Res.* (2022) 97:103502. doi: 10.1016/j.annals.2022.103502
16. Sun Y-Y, Gossling S, Babakhani N. Macro-scale decarbonisation of tourism: insights from Australia. *J Sustain Tour.* (2023) 1-25:1–25. doi: 10.1080/09669582.2023.2242608
17. Gössling S, Balas M, Mayer M, Sun Y-Y. A review of tourism and climate change mitigation: the scales, scopes, stakeholders and strategies of carbon management. *Tour Manag.* (2023) 95:104681. doi: 10.1016/j.tourman.2022.104681
18. Segarra V, Brida JG, Cárdenas-García PJ. On the relationships between tourism demand, carbon dioxide emissions and economic growth: a literature review. *J Policy Res Tour, Leis Events.* (2024) 6:1–42. doi: 10.1080/19407963.2024.2331792
19. Visas H, Rehan R, Ul-Haq J, Cheema AR, Khanum S, Hye QMA. Does tourism increase energy consumption in BRICS countries? *J Tour Manag Res.* (2023) 10:94–106. doi: 10.18488/31.v10i1.3402
20. Jiang Z, Liu J, Leung AK. Tracking spatial heterogeneity of local degree of saturation in unsaturated soils at different scales via X-ray computed tomography. *J Rock Mech Geotech Eng.* (2025). doi: 10.1016/j.jrmge.2024.12.004
21. Zhao J, Kou L, Wang H, He X, Xiong Z, Liu C, et al. Carbon emission prediction model and analysis in the Yellow River basin based on a machine learning method. *Sustain For.* (2022) 14:6153. doi: 10.3390/su14106153
22. Zhao Y, Liu R, Liu Z, Liu L, Wang J, Liu W. A review of macroscopic carbon emission prediction model based on machine learning. *Sustain For.* (2023) 15:6876. doi: 10.3390/su15086876
23. Schütt KT, Arbabzadah F, Chmiela S, Müller KR, Tkatchenko A. Quantum-chemical insights from deep tensor neural networks. *Nat Commun.* (2017) 8:13890. doi: 10.1038/ncomms13890
24. Schmidt J, Marques MR, Botti S, Marques MA. Recent advances and applications of machine learning in solid-state materials science. *Npj Comput Mater.* (2019) 5:83. doi: 10.1038/s41524-019-0221-0
25. Luo H, Wang C, Li C, Meng X, Yang X, Tan Q. Multi-scale carbon emission characterization and prediction based on land use and interpretable machine learning model: a case study of the Yangtze River Delta region, China. *Appl Energy.* (2024) 360:122819. doi: 10.1016/j.apenergy.2024.122819
26. Noh S, Lee S. Forecasting meteorological drought conditions in South Korea using a data-driven model with lagged global climate variability. *Sustain For.* (2024) 16:6485. doi: 10.3390/su16156485
27. Aras S, Van MH. An interpretable forecasting framework for energy consumption and CO₂ emissions. *Appl Energy.* (2022) 328:120163. doi: 10.1016/j.apenergy.2022.120163
28. Liu F, Kang Y, Guo K. Is electricity consumption of Chinese counties decoupled from carbon emissions? A study based on Tapio decoupling index. *Energy.* (2022) 251:123879. doi: 10.1016/j.energy.2022.123879
29. Meng H, Hu Y, Dong Z. The spatiotemporal decoupling relationship between economic development, energy consumption, and carbon dioxide emissions in Xinjiang Province from 2006 to 2020. *Sustainability.* (2024) 16:6421. doi: 10.3390/su16156421
30. Ruffing K. Indicators to measure decoupling of environmental pressure from economic growth. *Sustain Indic.* (2007) 67:211.
31. Katircioglu ST. International tourism, energy consumption, and environmental pollution: the case of Turkey. *Renew Sust Energ Rev.* (2014) 36:180–7. doi: 10.1016/j.rser.2014.04.058
32. Ng TH, Lye CT, Lim YS. A decomposition analysis of CO₂ emissions: evidence from Malaysia's tourism industry. *Int J Sustain Dev World Ecol.* (2016) 23:266–77. doi: 10.1080/13504509.2015.1117534
33. Malik MAS, Shah SA, Zaman K. Tourism in Austria: biodiversity, environmental sustainability, and growth issues. *Environ Sci Pollut Res.* (2016) 23:24178–94. doi: 10.1007/s11356-016-7609-x
34. Lin Y-H, Nitivattananon V. A low-emission pathway for tourism passenger transport in small and medium-sized urban destinations—case of Yilan, Taiwan. *Int J Tour Cities.* (2024) 4:43. doi: 10.1108/IJTC-02-2024-0043