Resour Environ Econ, 2024, 5(1): 379-398
DOI: 10.25082/REE.2023.01.002

RESEARCH ARTICLE

The performance of mixed and penalized effects models in predicting the value of the ecological footprint of tourism

Ahmad Roumiani[1*] **Omid Akhgari**[2]

[1] Department of Geography, Faculty of Letters and Humanities, Ferdowsi University of Mashhad, Mashhad, Iran
[2] Department of Statistics, Amin University, Tehran, Iran

Correspondence to: Ahmad Roumiani, Department of Geography, Faculty of Letters and Humanities, Ferdowsi University of Mashhad, Mashhad, Iran;
Email: roumiani.ah@mail.um.ac.ir

Received: January 7, 2024;
Accepted: March 15, 2024;
Published: March 19, 2024.

Citation: Roumiani A and Akhgari O. The performance of mixed and penalized effects models in predicting the value of the ecological footprint of tourism. *Resour Environ Econ*, 2024, **5**(1): 379-398.
https://doi.org/10.25082/REE.2023.01.002

Abstract: In recent decades, the issue of ecological footprint (EF) in the world has become a serious anxiety among environmental stakeholders. This anxiety is more in top tourism attracting countries. The purpose of this research is the performance of mixed and penalized effects models in predicting the value of the EF of tourism in the top eight countries of tourism destinations. The World Bank and Global Footprint Network databases have been used in this study. Penalized regression and MCMC models have been used to estimate the EF over the past 19 years (2000-2018). The findings of the research showed that the amount of ecological footprint in China, France and Italy is much higher than other countries. In addition, based on the results, a slight improvement in the performance of penalized models to linear regression was observed. The comparison of the models shows that in the Ridge and Elastic Net models, more indicators were selected than Lasso, but Lasso has a better predictive performance than other models on ecological footprint. Therefore, the use of penalized models is only slightly better than linear regression, but they provide the selection of appropriate indices for model parsimoniousness. The results showed that the penalized models are powerful tools that can provide a significant performance in the accuracy and prediction of the EF variable in tourism attracting countries.

Keywords: tourism growth, variable selection, top tourism-attracting countries

1 Introduction

Tourism development is now considered one of the most important industries in the 21st century [1]. The industry plays an essential role in employment and income, expanding exports, and promoting cultural and environmental values in developed and developing countries [2, 3]. According to Roumiani et al. (2022) [4], tourism contributed \$8.8 trillion to the global economy in 2018, which generated about 10.4% of GDP and 10% of world employment. In addition, the industry attracted approximately \$ 941 billion in investment in 2018, and this investment will increase by 4.2% over the next decade, equivalent to \$ 1.489 billion in 2029 [5].

Over the past few decades, tourism development studies show that tourism has caused economic activity. On the other hand, it has increased environmental pollution (greenhouse gas emissions), climate change, environmental degradation, excessive exploitation of natural resources and increased energy consumption [6–11]. In addition, As the number of tourist increases, the use of electricity per person, transportation, air conditioning/heating, and energy services will increase and expand the ecological footprint [12].However, tourism developments affect the ecological footprint and provide grounds are a concern for economists, social scientists, academics, environmental researchers, government policymakers, and other stakeholders [13–20].

Economic variables (investment, GDP, imports and exports) [21, 22]. energy variables (oil, coal and gas) [23]. variables such as hotels, restaurants, road transport [4] and tourist expectations and demandsare the main influencing factors of the tourism development on ecological footprint. Martín-Cejas and Sánchez (2010) [12] reported that transportation (94%), place of residence (4%) and other service activities (2%) are among tourism-related factors which consume the most significant amount of energy. Hotels alone account for 70.8% of electricity consumption. According to the CTCI Foundation (2004) survey conducted in Taiwan, 0.32% of total electricity consumption in hotels. However, all researchers agree that the relationship between the use of energy and the tourism industry in the amount of greenhouse gas emitted is positive and significant [24]. Thus, a tourist is expected to use local transportation, accommodation, and food on his or her international travel [25]. which in turn consumes non-renewable energy and fossil fuels [26]. Accordingly, the discussion of a cause-and-effect relationship between environmental degradation and tourism development introduced the term“ecological tourism footprint” [24]. Furthermore, the EF is a composite indicator of human demand for natural resources and provides a better understanding regarding economic activity and environmental degradation [27].

Ecological footprint calculation can measure different biologically productive lands, water consumption, population density, energy consumption, and resources [27]. Itcan contribute to the future of humanity in environmental protection and sustainable development by providing legal judgments in practice [28]. Therefore, assessing the ecological tourism footprint can play a crucial role as a solution to strengthen and preserve the environment, reduce greenhouse gas (GHG) emissions, mitigate the harmful effects of human activities and address critical environmental challenges such as climate change and environmental damage [29, 30].

This research evaluates the effect of tourism developments on the EF in the top tourist countries using MCMC and penalized models. Past studies have utilized econometric and regression (OLS) models focusing on the effect of different tourism development activities on the EF [31–36]. Theoretically, econometric and regression models increase the sample error in predicting variables, due to high alignment and cannot estimate the interpretability and accuracy of the prediction at the same time. According to Cherlin et al. (2018) [37] such variables have poorer predictability overall in the experiment. Therefore, to overcome the shortcomings above, the use of penalized regression methods, also called contraction or regularization methods, has become popular [38]. Researchers such as Hoerl and Kennard (1970) [39], Tibshirani (1996) [40] and Zou and Hasti (2005) [41] popularized the use of penalized regression (PR) models with the introduction of a slight bias in model prediction to deal with linear regression drawbacks and decrease the variance of estimates and enhance predictions. Ridge regression, for example, has been proposed as a possible solution for estimating and selecting independent variables which are highly aligned, and it provides a more accurate estimate of ridge parameters [42]. Lasso regression is used to select a subset of variables and has better predictive accuracy than other regression models and helps to increase the interpretation capability of the model [43–45]. Hence, penalized models, due to their capability to overcome multiple alignments, select appropriate variables, and reduce research costs as a result,can be effective methods for improving tourism and ecological footprint research studies.

The current research aimed to find the similarities between the performance of MCMC and penalized regression models and their ability to predict tourism change in 8 countries. We used two train and test data sets to achieve the research goal. In addition, this study provides excellent insights for researchers and tourism stakeholders. Secondly, we use indicators such as number of international tourist arrivals, total import costs, cost of passenger transport, cost of travel, number of departures, cost of revenue from total exports, which have economically significant effects on the ecological footprint. Third, few studies of Penalized Models have been used to enquire into the impacts of tourism development on ecological emission. Hence, this study focuses on MCMC and Penalized Models and seeks to answer the following fundamental question; which of the MCMC, OLS, Ridge, Lasso, and Elastic Net models can be more accurate in predicting EF?

1.1 Theoretical Foundations

There is a lot of research on tourism development and ecological footprint, but the EF of tourism in top tourist destination countries has been neglected. Therefore, the theoretical framework of this research is made up of three parts, the first part of which examines the history of the EF of tourism and its importance from the point of view of experts. In the second part, the correlation between tourism growth indicators and the EF is discussed.The third part includes the critical view of researchers on the effects of tourism Growth on the environment.

1.2 Ecological footprint of tourism

In 1996, Wackernagel and Rees used the term EF of tourism for the first time [46]. In 2002, Hunter categorized the ecological footprint of tourism and applied its function in tourism planning. After that, the organization of the World Summit on Sustainable Development in Schroeder & Lovell (2012) [47], the conventions of the Climate Change Organization

(UNFCCC) in 1990; Kyoto Protocol (1999), Copenhagen (2009), Duha (2012), Paris, France (2015) and Poland (2018) mentioned the development of tourism as one of the main sectors of increasing the EF [48]. In addition, according to the report of the World Tourism Organization (2013), this industry has caused about 8% of the total CO2 emissions and 14% of the total global greenhouse gas emissions [49].

Many researches on the ecological footprint of tourism have been conducted by researchers such as [49, 50]. They stated that the development of tourism increases the EF. In addition, Gossling et al. (2002) [50] believe that about 97.5% of the footprints of European tourists were caused by air travel.

Cadarso et al. (2016) concluded that the inclusion of tourism investment; It increases the carbon footprint by about 34%. Khan and Hou (2021) [21] empirically examine the International Energy Agency in 30 countries from 1995 to 2015 and conclude that panel co-analysis shows long-run relationships between economic variables and environmental sustainability. The experimental study of Işik et al. (2022) [51] in USMCA countries showed that there is convergence of EF with a value of 48.08% in the second category and divergence in the first category. These results indicate common environmental policies-actions among USMCA countries to reduce and stop their environmental degradation; and to understand which strategies-actions to use in the case of convergent or divergent ecological footprints. Therefore, using the concept of the ecological footprint of tourism at the national level, as a measure for human use of environmental resources (renewable), and at the global level, it evaluates the consumption of biological environmental resources of humanity with 20.6 billion hectares [52].

1.3 Relationship and dynamics of tourism indicators with ecological footprint

Tourism is a trillion-dollar industry that includes about 7% of global exports and has a important influence to global gross domestic product (GDP). International tourism arrivals and receipts have grown by 3-5% annually, outpacing the growth of international trade. In 2016, respectively, it exceeded 1.2 trillion US dollars, and in 2018, it created one tenth of all jobs in the world [49, 53]. In developing countries, tourism has been mentioned as a strategy in creating employment and increasing foreign exchange earnings [54]. In Mediterranean countries, it is known as one of the key elements in economic activities. So that they account for about 30% of the world's tourism arrivals and about a third of the total world tourism income [29].

Therefore, the cause-and-effect relationship between tourism and socio-economic development has been reported. With a 1% increase in tourism growth; 0.051 percent (GPD) and 2.647 percent foreign investment increases [21]. However, excessive emphasis on tourism development has increased the ecological footprint and it has caused concerns among the researchers of the countries of the world. For example, Kongbuamai et al. (2020) [55] used different economic parameters (investment, GDP, trade openness) to investigate the correlation between tourism development indicators and ecological footprint. Muchapondwa & Stage (2013) [56] used income generation indicators, the number of international tourists and travel items. Also, based on the report of 141. Dwyer et al. (2010) [57], a significant relationship between tourism growth rate and ecological footprint has been estimated. In Turkey, Godil et al. (2020) [31] stated that there is a positive relationship between tourism development, globalization and financial development with ecological footprint. Ansari & Villanthenkodath (2022) [53] concluded that tourism revenues are negatively related to environmental degradation. The results also show that economic growth, energy intensity and urbanization reduce environmental quality in the long term.

Liu et al. (2022) [58] investigated the correlation between the ecological footprint and tourism development indicators using a regression model for the years 1980-2011. In their experimental study in China, Lin et al. (2018) [59] concluded that the footprint of tourists' purchases and the footprint of tourists' traffic, negatively affect environmental quality. According to this research, the growth of tourism not only increases the ecological footprint, but also increases human performance on the earth's ecosystem, conflicts between natural resources, pollution and excessive consumption of environmental resources [60].

1.4 The views of critics on the effects of tourism development indicators on ecological footprint

The trend of environmental consequences in the world shows that different views have been expressed in relation to the development of tourism, but they have not provided a general perception with the palaning for the society about the evolution of tourist destinations and ecological footprint. Therefore, in this research, the views of researchers from various angles have been criticized on this issue [61].

Researchers such as Alola et al. (2021) [62], Chakraborty (2021) [60], Telfer and Sharpley (2015) [63] believe that most tourism beneficiaries are only looking for economic income and do not pay attention to the issue of ecological footprint. Bohdanowicz (2006) [64] believes that the environmental challenges of tourism are the result of bad policies and planning of the government and management institutions in the world. Dasgupta De Cian (2018) [65] stated that there is an interactive relationship between powerful political parties and tourism environmental protection in the countries of the world, but governments have not paid serious attention to this issue. Environmentalists consider the growth in the sum of tourists as a key factor in expanding the ecological footprint. Economists state that extensive investment in tourism infrastructure leads to an increase in the ecological footprint [66].

Croall (1995) [67] stated that the growth of tourism affects our environmental landscape. Akinboade et al. (2010) [68] stated that United Nations studies have shown that tourism has brought about social and environmental changes in countries. However, even tourism development projects often ignore the global environmental aspects of travel. Tourism development may be locally sustainable (in the sense that it poses minimal threat to local ecosystems through land conversion, trampling, species collection, *etc.*), but it may not be a stable world in most cases [55].

1.5 Area of the study

Some countries attract a higher rate of tourist trips than other countries due to their various attractions. The countries studied in this research include China, Italy, Germany, France, America, Thailand, Turkey and Mexico. Therefore In 2019, 1.459 billion international tourists arrived worldwide, which was a 3.7% increase compared to 2018 [51, 62]. The share of these countries in 2019; 508 million was international tourism. Figure 1 shows the number of tourist arrivals in these countries between 2000 and 2018. This figure shows that all countries are faced with a wide change in tourism. For example, France has attracted the most tourism in all periods. In 2006, the United States made a huge leap in attracting tourists.

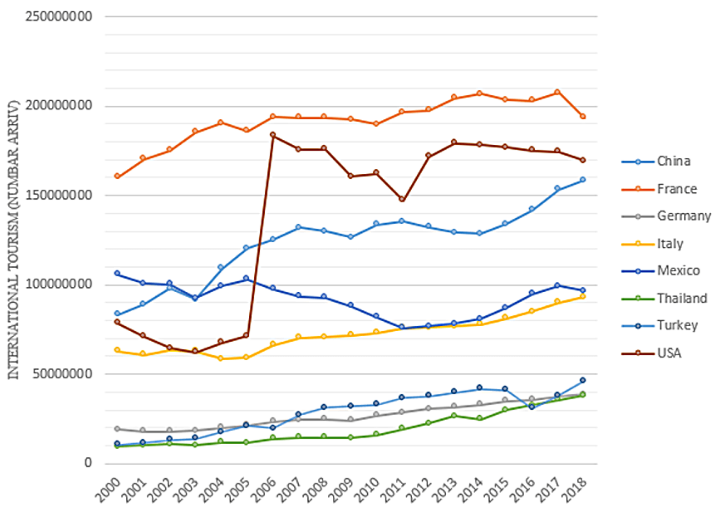

Figure 1 The growth trend of the number of tourists (2000-2018) in the top tourist destination countries

2 Materials and methods

2.1 Descriptive statistics and data preparation

The data needed to examine tourism development indicators and ecological footprint from (2000 to 2018) was extracted from (https://databank.worldbank.org) and (https://data.footprintnetwork.org). We used the dependent variable (output) to help make effective decisions in the top tourist attraction countries. That is, ecological footprint (EF) and 10 indicators of tourism development (predictors) mentioned in Table 1.

Table 1 Study indicators

ID	Variable	Unit	Sources
X1	Period of years	-	World bank
X2	International tourism, number of arrivals	10^6	World bank
X3	International tourism, expenditures (% of total imports)	10^6	World bank
X4	International tourism, expenditures (current US$)	10^6	World bank
X5	International tourism, expenditures for passenger transport items (current US$)	10^7	world bank
X6	International tourism, number of departures	number of departures	world bank
X7	International tourism, receipts (% of total exports	10^6	World bank
X8	International tourism, receipts (current US$)	10^8	World bank
X9	International tourism, receipts for passenger transport items (current US$)	10^8	World bank
X10	International tourism, receipts for travel items (current US$)	10^8	World bank
Y	Ecological footprint	global hectares (gha)	Global Footprint Network

Some descriptive information of the set of variables is presented in Table 2. In this table, the predictor variables are compared in terms of Mean, TSE, CV, and Range. In total, variable y is 2.338 in terms of average, 2.000 in terms of TSE, 36.050 in terms of CV, and 2.872 in terms of Range.

Table 2 Descriptive statistics of indicators

Indicators	Mean	SE	CV (%)	Range
X1	4.500000e+00	5.400000e-02	51.086	7.000000e+00
X2	2.009000e+03	1.290000e-01	0.274	1.800000e+01
X3	8.681247e+07	1.449087e+06	71.289	2.024190e+08
X4	5.906000e+00	9.000000e-02	65.068	1.842400e+01
X5	2.753693e+10	1.021963e+09	158.501	1.724270e+11
X6	1.217400e+01	2.110000e-01	74.180	3.845800e+01
X7	2.421854e+04	4.329800e+02	76.354	5.424373e+04
X8	2.533645e+08	9.767297e+06	164.642	1.391463e+09
X9	6.916900e+01	2.740000e-01	16.902	4.637900e+01
X10	5.041000e+00	4.200000e-02	35.655	7.531000e+00
Y	2.338000e+00	2.000000e-02	36.050	2.827000e+00

Note: CV: Coefficient of variation, known as relative standard deviation (RSD), defines as CV%=$\sigma/\mu\times 100$

Table 2 shows the descriptive statistics of the predictor variables of each country. The normality of the residuals (errors) of the linear regression model and their autocorrelation were investigated using the Shapiro-Wilk and Durbin-Watson tests. This statistical test to consider the correlation between residuals is defined as follows:

$$D = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \tag{1}$$

Where D is in the range of 1.5–2.5, it indicates no correlation between the residues. After fitting the initial model, multicollinearity was examined using VIF (Variance Inflation Factor [69] as follows:

$$VIF = \frac{1}{1 - R_i^2} \tag{2}$$

Finally, the validity of the initial model is shown (*e.g.*, Figure 2).

2.2 Statistical methods and model configuration

If we consider a standard multiple linear regression model as follows: $y = 1\beta_0 + X\beta + e$ that y is a response variable vector, $X = x_{i1}, \ldots, x_{ip}$ is a predictor variables matrix, β_0 is the intercept, $\beta = \beta_1, \ldots, \beta_p$ is a regression coefficients vector, and e is an error terms vector, assuming normal distribution $e \sim N\ (0,\ \sigma_e^2)$. In this case, β_0 and β_s coefficient values be estimated by minimizing the residual sum of squares (RSS) [70] as:

$$\hat{\beta}_0, \hat{\beta}_{(\mathrm{OLS})} \stackrel{\text{def}}{=} \arg \min_{\beta_0, \beta \in \mathbb{R}^p} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 \tag{3}$$

And a PR coefficients is expressed as:

$$\hat{\beta}_0, \hat{\beta}_{(\mathrm{PR})} \stackrel{\text{def}}{=} \arg \min_{\beta_0, \beta \in \mathbb{R}} \underbrace{\sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2}_{\text{Loss function}} + \underbrace{\lambda \mathrm{P}(\lambda, \beta)}_{\text{Penalty function}} \tag{4}$$

Here, to prevent overfitting and control the penalty function's shrinkage amount of contraction, the hyper-parameter λ tunes the equation. In fact, bias-variance trade-off is set by this hyper parameter. Its amount is directly related to bias and inversely associated with variance, *i.e.*, with increasing lambda, bias increases and variance decreases.

By applying an l_2-norm penalized least squares criterion [*i.e.*, P (λ, β) = λ $\|\beta\|$ l_2] on the linear regression coefficients [39], the RR estimates are obtained as follows:

$$\widehat{\beta}_0, \widehat{\beta}_{(\mathrm{RR})} \stackrel{\mathrm{def}}{=} \arg \min_{\beta_0, \beta \epsilon \mathbb{R}^p} \left[\Sigma_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \|\beta_j\|_2^2 \right] \tag{5}$$

In RR, the shrinkage value is tuned so that no variable is exactly zero and only reduces their variance, so the estimates are biased.

In another case of PR, the values of the coefficients are obtained by applying the Lasso constraint (*i.e.*, an -$l_1$norm penalized least-squares as P (λ, β) $=$ λ $\|\beta\|\, l_1$) [40]. An important feature of Lasso is that it allows the coefficients to be exactly zero, thus selecting the variable. If we consider x_{ij} to be standardized so that $\frac{\sum_i x_{ij}}{N} = 0$ and $\frac{\sum_i x_{ij}{}^2}{N} = 1$, then the lasso coefficients are estimated as follows:

$$\widehat{\beta}_0, \widehat{\beta}_{(\mathrm{LASSO})} \stackrel{\mathrm{def}}{=} \arg \min_{\beta_0, \beta \in \mathbb{R}^p} \left[\Sigma_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \|\beta_j\|_1 \right] \tag{6}$$

The EN method is another mode of PR that uses a combination of two penalties applied in RR and Lasso on the coefficients:

$$\widehat{\beta}_0, \widehat{\beta}_{(\mathrm{EN})} \stackrel{\mathrm{def}}{=} \arg \min_{\beta_0, \beta \in \mathbb{R}^p} \left[\sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} \left(\frac{1}{2}(1-\alpha)\|\beta_j\|_2^2 + \alpha\|\beta_j\|_1 \right) \right] \tag{7}$$

Where $0 \leq \ \alpha \leq \ 1$ is a penalty weight. If α is equal to 1, EN functions like Lasso but modifies how it deals with high correlated variables [70].

2.3 Cross validation and parameter optimization

To ensure the main distribution of the indicators, the data were measured in two levels of training and testing through minimum and maximum normalization. 70% of the data was used to train the models and the rest was used for testing. In this study, the performance of the models was evaluated using the 10-fold cross-validation method. The entire data was randomly divided into ten equal subsets. A subset of the validation set was considered for model testing and the remaining k-1 subset was used for model training.

This method reduces the dependence of the performance on the test-training set and reduces the variance of the performance measures and confirms that the results are free from any sampling bias. The optimal value of λ minimizes the percentage of cross-validation prediction error in the training set. This λ value is automatically determined using the cv. glmnet function. With a default of 10 times cross-validation, the cv. glmnet function sets the optimal λ value to provide the simplest model. The appropriate model lies in an optimal λ standard error, *i.e.*, lambda.1se. If λ equals lambda.1se we have a simpler model than λ equals lambda.min, but it may be slightly less accurate. The selection of λ values for each Fi fold was done with the following cross-validation technique:

$$\epsilon_{n,\lambda} \stackrel{\mathrm{def}}{=} (y_n - \widehat{\beta}_{\mathcal{PR}}^{\mathcal{F}_i,\lambda} x_n)^2 \quad \forall_n \epsilon \mathcal{F}_i \tag{8}$$

that $\hat{\beta}_{PR}^{F_i,\lambda}$ was estimated on $D - F_i$ (here, $D = \{x_n, y_n\}$ and F_i is the data not including the ith fold). ith fold was used as a test set and the rest of the data as the training set. So, λ was chosen as:

$$\lambda^{x-val} = \arg\min_{\lambda} \frac{1}{N} \sum_{n=1} N_{\epsilon\ F_i,\lambda} \tag{9}$$

2.4 Performance evaluation

With the introduction of new data, the behavior of the models was evaluated using the following criteria: MSE (mean square error), RMSE (root mean square error), and R-square values (coefficient of determination). Considering n as the total number of observations.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{11}$$

After reviewing the penalized models, the multi-level MCMC model method has been introduced to estimate its parameters.

2.5 MCMC model

The main features of multilevel data are their statistical grouping. Many statistical inference methods have been expressed about the parameters of multi-level models, which are usually done in two ways: frequency and Bayesian. Meanwhile, Bayesian analysis in two-level models was first performed by Seltzer (1993) [71] using the Gibbs algorithm. Seeger (2004) [72] also used Gibbs sampling for multilevel logistic models. As is customary in Bayesian methods, to estimate the parameters, one needs the prior distribution, the likelihood function and the calculation of the posterior distribution. Since the posterior distributions of some parameters in multilevel models do not have a closed form, the use of the MCMC method will be inevitable.

Therefore, Bayesian inference was used to fit the two-level model and estimate its parameters in this research. In addition, due to the nature of multilevel models, Gibbs sampling algorithm was used more than other algorithms. The main reason for that is the closed form for the complete conditional distributions of the parameters involved in the model. Therefore, for this purpose, the prior distribution of the model is considered to be ignorant, and according to the theoretical literature and as the posterior distribution of the model parameters is not closed, one of the Markovian chain Monte Carlo (MCMC) algorithms called the Gibbs sampling algorithm was used and their equations are discussed below:

$$y_{ij} = \beta_{0j} + \beta_{01} year_j + \sum_{k=2}^{10} \beta_{0k} x_k + \epsilon_{ij} \tag{12}$$

$$\beta_{0j} = \beta_{00} + u_{0j} \tag{13}$$

By placing relation (2) in (1), the final model will be obtained as follows.

$$y_{ij} = \beta_{00} + \beta_{01}\ year_j + \sum_{k=2}^{10} \beta_{0k}\ x_k + u_{0j} + \epsilon_{ij} \tag{14}$$

In which ϵ_ij is the error of the first level (measurement), u_0j is the error of the second level (sampling), Year is the variable of the second level, x is the variable of the first level and β = (β_00, β_010) is the coefficients of the model, which is an 11-dimensional vector. In this model, u_0j has a normal distribution with zero mean and variance σ_u^2 and ϵ_ij also has a normal distribution with zero mean and variance σ_e^2, and u_0j and ϵ_ij are uncorrelated.

2.6 Gibbs sampling algorithm for model (3)

Assume that the prior distributions of the parameters of this model are:

$$\beta \ \propto \ N_{10}\,(0,\ S)\,, \quad \sigma_u^2 \sim IG\,(a_u, b_u)\ , \quad \sigma_e^2 \sim IG\,(a_e, b_e) \tag{15}$$

In which S is the 10×10 covariance matrix for β vector, IG is the inverse gamma distribution. It is proved that σ_e^2.

It should be noted that the mentioned analyzes were performed in the R 4.0 program (R Core Team, 2020) using different packages.

3 Results

Figure 2 shows that based on the diagram (Residuals vs Fitted) there is a linear relationship between tourism changes and ecological footprint. It also shows that the points around and close to the regression line are estimated and the equal variance assumption is reasonable. Here, the corresponding residuals versus the proportional diagram for the simple linear regression model of the ecological footprint set is considered as the response and the level of change and tourism developments are considered as predictors. Therefore, the residuals are displayed on the y-axis and the fitted values on the x-axis. Also, to evaluate the distribution of the data set, the Q-Q normal diagram has been used, which is one of the most important assumptions in R statistical tests. Figure 2 shows that the data distribution is linear and the remaining average is almost completely on the reference line. Therefore, a set of data used in the research has a normal distribution. The (Scale-Location) chart shows how to analyze the spatial scale chart of the data. In other words, this figure shows that the residuals are equally spread over the range of input (predictor) variables and are standardized between the fitted values and the square root of the residuals. Therefore, Figure 2 shows the assumption of equal variance and that the distribution points are spread randomly along the horizontal line. This means that the plot shows a linear fit to the data. The graph (Residuals vs Leverage) shows the mean squared residuals against the Cook distance. The findings in this figure show that the data are between Cook's intervals. The points located near or outside the red curves are considered outliers.

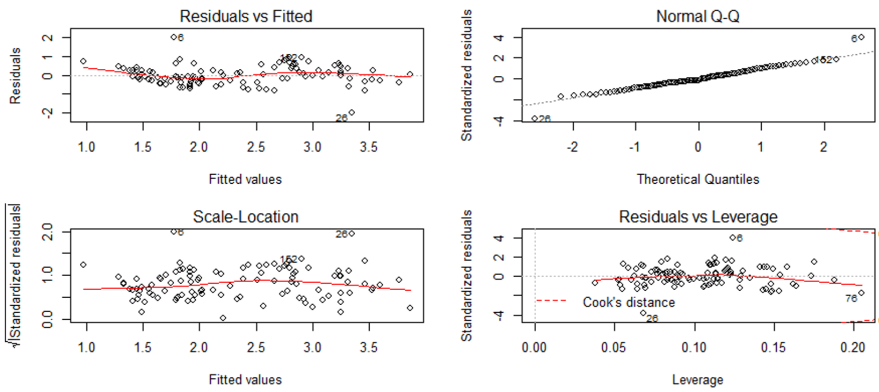

Figure 2 Image, Residuals *vs* Fitted and Normal Q-Q and Scale-Location and Residuals *vs* Leverage of ecological footprint in the top tourism attracting countries

Ecological Footprint is a resource review tool that helps different countries understand their environmental assessment and provides them with the necessary data to manage resources and create a safe and resilient future. In the last few decades, this approach was considered as a strong communication tool in the field of measuring environmental, economic and political systems. Table 3 shows the characteristics of the ecological footprint index in the top 8 tourism attracting countries. This table shows that the average ecological footprint in China, France and Italy is 3.682, 3.251 and 3.002 respectively. For example, the ecological footprint survey in China during the years 2007 to 2017 shows that during the 11-year study period, China's EF per capita has gradually increased and in fact, it increased from 2.3027 square meters in 2007 to 2.9837 square meters in 2017 [73].

Other research in this country shows that China's environmental per capita is 3.71 hectares. China's total ecological deficit is (-3,435.62), which is the largest deficit in the world. With the expansion of the rapid growth of the population of 1.4 billion people and the rapid growth of the economy, people's income and as a result their consumption has increased, and it has created the ground for inefficient use of environmental resources and maximum pressure on it, and environmental risks [74–76]. In France, approximately 70% of the country's ecological footprint is caused by household energy consumption (24%), transportation (23%) and food (22%) and the rest is related to other goods (15%) and services (16%). . This leads to deforestation, reduction of fish stocks, drought, water shortage, soil erosion, loss of biodiversity and climate change (Report Lautre deficit de la France, 2018). The ecological footprint in this country has been decreasing from 2008 to 2015 and increased by +5% between 2015 and 2018. This is primarily due to the increase in greenhouse gas emissions in the transportation, construction and power generation sectors, and the second case is the greater use of fossil fuels between 2014 and 2016 [77, 78]. The lowest amount of ecological footprint is related to the countries of Mexico, Thailand and Turkey, each with an average of 1.297, 1.610 and 1.745, respectively. In Mexico, about 56% of the earth's surface is under the influence of human activities. Its use is not evenly distributed in the regions of this country; the lowest values are in the arid regions of the north and northwest and the tropical regions of the southeast.

While the greatest values are along the coast of the Gulf of Mexico and from there within the east-west corridor that follows the transverse volcanic ranges of Mexico. The high plateau associated with the distribution of low and high ecological footprint areas in different regions forms a complex mosaic: Mexico's generally protected deserts have some highly altered agricultural and industrial areas; while many well-preserved footprints still remain in highly altered areas [79, 80].

Also, in Thailand in 2011, Thailand Greenhouse Gas Management Organization (General

Organization: TGO) in collaboration with National Metals and Materials Technology Center, Thailand (MTEC) and National Science and Technology Development Agency (NSTDA), under the supervision of the Ministry of Science and Technology, promoted the development carbon footprint project. In fact, the Carbon Footprint Project, "TGO" launched a suitable strategy for effective management of greenhouse gas emissions, supporting the economy, protecting the environment, and the economy of the environment and society [81]. In Turkey, the researchers conducted by various researchers such as Destek et al. (2021) [82] have shown that the process of industrialization in this country has led to a reduction in carbon emissions and has not had significant effects on the ecological footprint. Therefore, industrialization has had few effects on the environment due to the use of advanced technologies.

Table 3 Descriptive statistics of ecological footprint variables in top tourism attraction countries

Countries	ID	Mean±SE	SE	CV (%)	R	Skewness	Kurtosis	Minimum	Maximum
France	1	3.251	0.070	7.43	1.12	1.200	2.245	14.350	4.592
USA	2	2.372	0.065	9.56	0.92	-0.400	-0.308	17.434	15.131
China	3	3.682	0.031	2.93	0.37	0.245	-0.917	7.320	5.795
Italy	4	3.002	0.095	10.93	0.86	-0.185	-1.709	6.836	5.387
Mexico	5	1.297	0.027	7.33	0.29	0.172	-1.479	4.157	3.808
Turkey	6	1.745	0.021	4.18	0.23	0.092	-1.346	4.692	4.035
Germany	7	1.749	0.025	4.92	0.34	0.743	-0.011	9.642	8.969
Thailand	8	1.610	0.039	8.36	0.53	0.864	0.251	3.851	3.309

3.1 Evaluation of tourism development indicators on ecological footprint using MCMCglmm model

To estimate the parameters of the model, the standard deviation of these estimates and the estimation of random effects have been calculated using MCMC methods. For the number of simulations, it was based on 1,300,000 times of sampling with the burn-in stage of 300,000 times. Then, one out of every 100 is selected from the remaining 1,000,000 Gibbs samples. Therefore, finally, 10,000 samples of the desired parameters are available. The results of this model (*i.e.*, the mean of the posterior distribution, the lower and upper limits of the 95% confidence limits of the posterior distribution, the effective sampling size and the pi value for MCMC) are presented in Table 4. In this table, the effects of tourism development indicators are considered as "fixed effects". The noteworthy point in these analyzes is that the effect of some tourism development indicators was not significant, which actually led us to use penalized regression models in these MCMC models.

Table 4 Random structure of indicators

	Post. Mean	l-95% CI	u-95% CI	Eff. Samp	pMCMC
(Intercept)	7.377	-8.831	22.510	9193	0.360
Year	-0.002	-0.010	0.006	10000	0.606
X1	0.003	-0.018	0.023	10000	0.815
X2	0.046	0.016	0.075	10000	0.003
X3	0.003	0.000	0.006	9251	0.071
X4	-0.007	-0.018	0.003	10000	0.166
X5	-0.013	-0.027	0.000	10000	0.052
X6	0.048	0.028	0.068	10000	0.000
X7	0.726	-0.035	1.308	10000	0.057
X8	-0.601	-0.908	-0.209	10000	0.000
X9	0.048	-0.007	0.100	9503	0.081

As can be seen, based on the level of significance (PMCMC), all variables except the two variables x_1 and year, the width from the origin, are significant. Therefore, the variance of the first and second level effects is mentioned in Table 5.

Table 5 Variance of first and second level effects

	Post. Mean	l-95% CI	U-95% CI	Eff. Samp
σ_u^2	23.630	0.123	71.020	10000
σ_e^2	0.018	0.014	0.023	9803

In the case of approximation convergence, the approximation effect plot shows how the MCMC chain is progressing. Therefore, we were interested in monitoring the process of chain movement for estimators. Figure 3 shows the effect diagram of fixed effects estimation. In models that reach convergence (like what we have reached in this study), the distribution usually has a peak that is more or less skewed (long right tail). In these graphs, which are time series (graphs on the left), the samples did not remain in a certain range, so the posterior average of the total samples is shown correctly, and also the length of burn-in and total sampling were favorable.

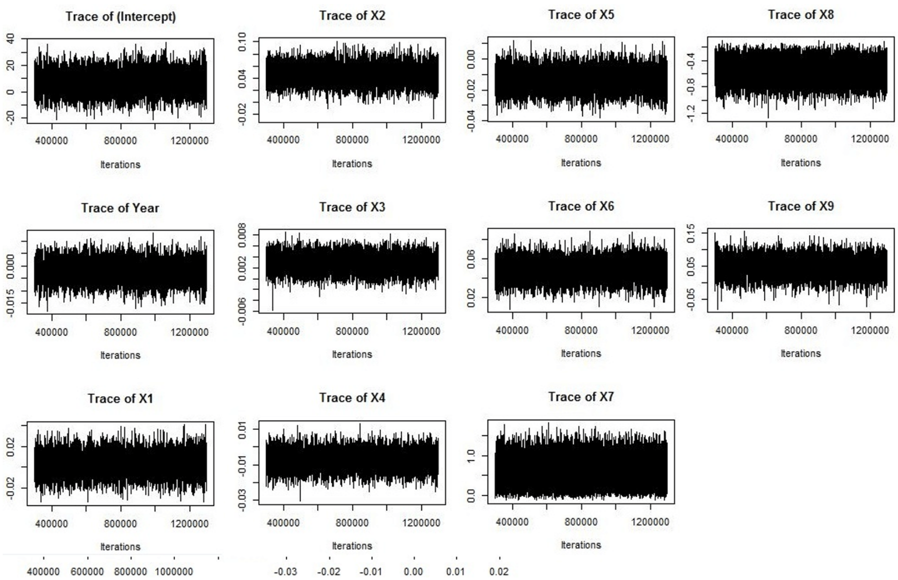

Figure 3 Plot of model fixed effects estimates

Regarding the accuracy of the method in estimating the parameters, according to the smooth graphs of the posterior densities, Figure 4 shows that the parameters have been estimated acceptably, which indicates the accuracy of the method in estimating the indices.

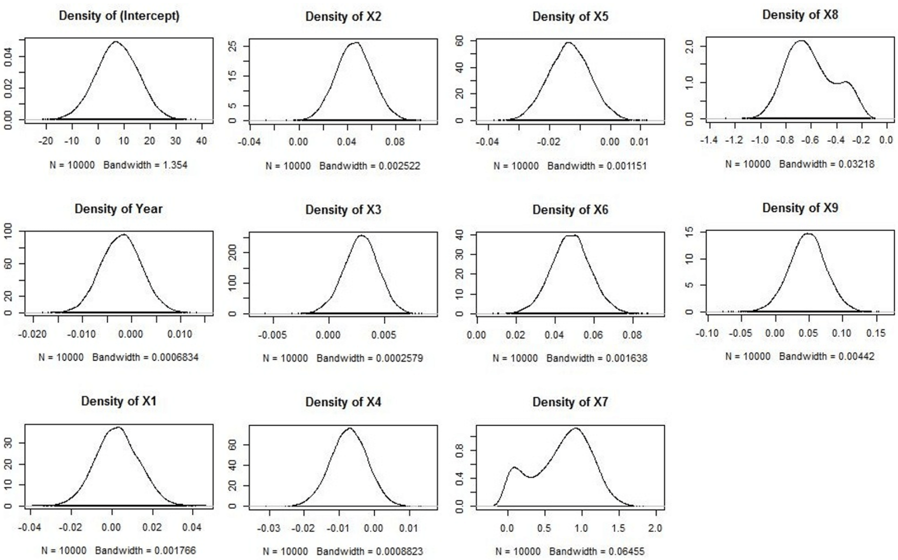

Figure 4 Posterior density diagram of model fixed effects estimates

In this part, it refers to the countries as a random effect and the unit refers to the years under study, *i.e.*, about 19 years. According to what was stated, this model (mean of posterior distribution, lower and upper limit of 95% confidence limits of posterior distribution, effective sampling size and pi value for (MCMC) is presented in Figure 5. Therefore, in this figure, the country effect is "random" and the unit shows the residual error effect.

Autocorrelation diagrams of 10,000 Gibbs samples were drawn, which indicates the independence of these coefficients due to the close to zero value of this coefficient. Therefore, Figure 6 shows that for each parameter it shows the autocorrelation as a function of the distance between the samples. If the interval is 0, the autocorrelation is one because the correlation of a variable with itself is one. However, as the distance between samples increases, the autocorrelation decreases.

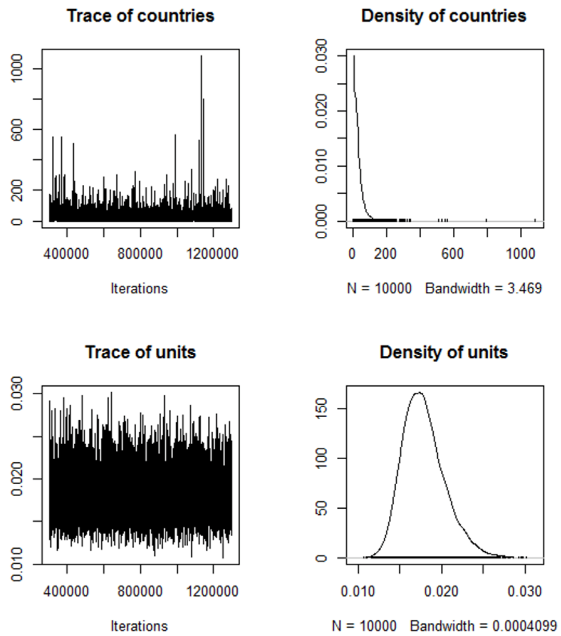

Figure 5 Estimates chart and density function of random effects estimates of the model

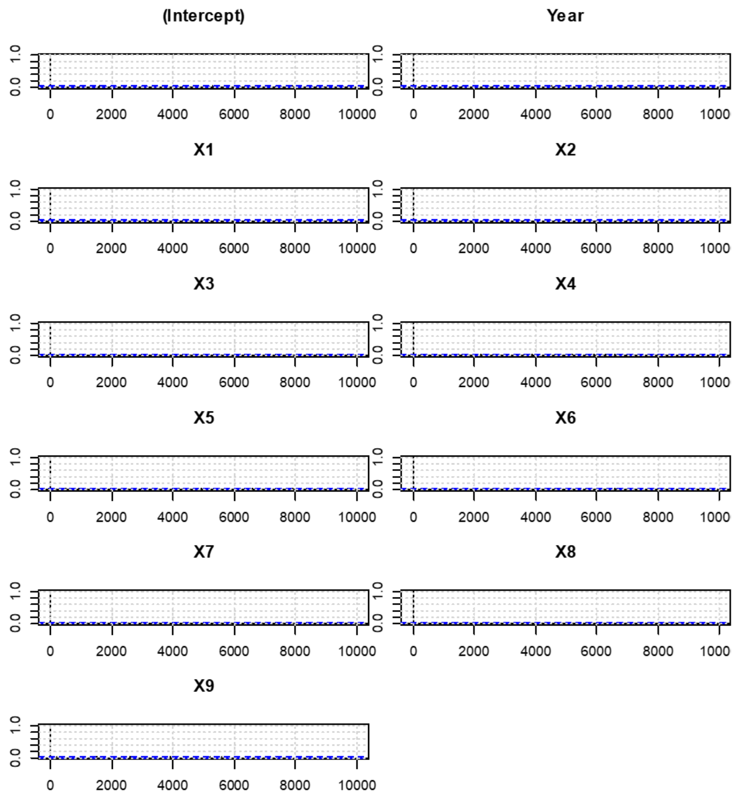

Figure 6 Internal correlation diagram of indicators

We also see that the interception parameter has much lower autocorrelation than other parameters. which shows a very high intra-group correlation of observations and this confirms the correctness of using the mentioned model. Therefore, in Figure 6, the correlation value is equal to 0.999. The value of intraclass correlation (ICC) is equal to

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} = 0.999 \tag{16}$$

Which shows a very high intra-class correlation of observations and this confirms the correctness of using the mentioned model.

In order to analyze the sensitivity of the mentioned model, Gelman-Robin method was used. The graphs in Figure 7 show the conformity of parameter estimates in different conditions. According to the Golman and Rubin criterion, less than 1.010 was obtained for 9 parameters, and the Golman diagram for all 9 parameters reached the convergence condition, *i.e.*, less than 1.1, after 1,200,000 repetitions. Therefore, based on the comments of Gelman and Rubin (1992) [83] and Brooks and Gelman (1998) [84], they suggest that diagnostic Rc values greater than 1.2 for each of the model parameters should indicate a lack of convergence. But in practice, a stricter rule than Rc < 1.1 is often suggested to declare convergence, which in test statistic it is called the scaling factor. The closer this coefficient is to 1, the better the convergence of chains. In practice, values below 1.1 are acceptable and values below 1.02 are good. In the graphs below, the downscaling is shown for bins of increasing size (1 to 0.12, *etc.*), thus showing how the downscaling factor has developed over time. Also, the 97.5% confidence interval is shown with a red dashed line, which shows the attention of the x-axis of the main indicators of the samples before thinning.

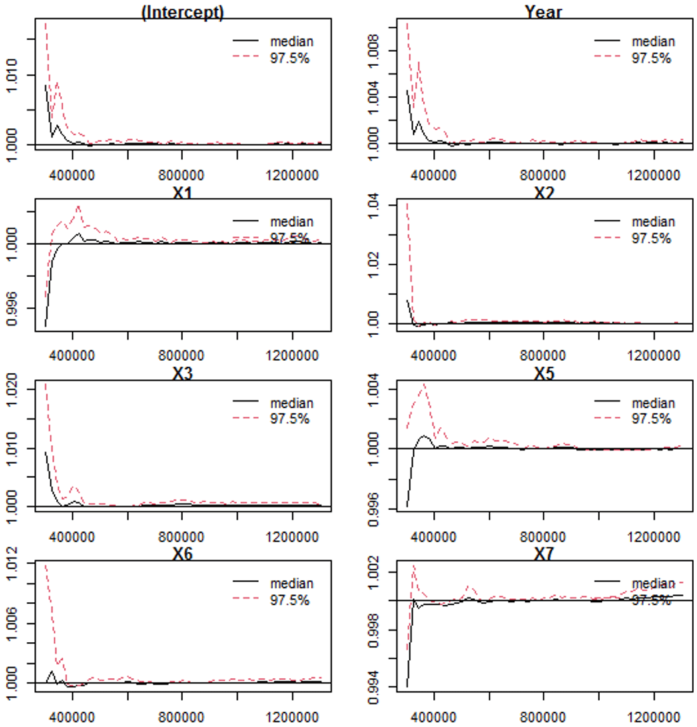

Figure 7 Golman diagrams of parameters for running 1,200,000 Monte Carlo samples

It is noteworthy that MCMC models also led us to use penalized regression models. Therefore, it is expected that Lasso regression with the "variable selection" feature can help simplify the issue.

3.2 Predictive evaluation of penalized models

To choose the appropriate model in predicting the ecological footprint of tourism and estimating the mean square error (MSE) from linear regression models, Rich; lasso and elastic net were used and we compared them with each other. Our cross-validation consisted of randomly splitting the data into a training and testing set, adjusting the model parameters into the training and prediction sets.

The summary of RMSE scores is presented in Table 6. So that it is 0.578 for linear regression, 0.334 for ridge, 0.333 for lasso and 0.337 for elastic net. Therefore, in the Lasso regression model, three indicators have been removed, but it shows the same value compared to other ridge and sticknet models. Because in addition to good out-of-sample prediction performance, Lasso was able to select variables by reducing the coefficients to zero and thus increasing interpretability. Although elastic network can also perform variable selection, it tends to select more variables and does not perform better than Lasso despite being a more complex and robust model.

Table 6 Comparison of predictive validity for the ecological footprint index in top tourism destination countries

Variable	Method	RMSE	R^2	Acc-R
EF	OLS	0.578	0.62	0.81
	RR	0.334	0.81	0.91
	Lasso	0.333	0.82	0.91
	EN	0.337	0.81	0.91

The reality may be much more complex than a proposed model. Hence, there is no certainty that the commonly studied events are simple enough to be approximated by human-understandable models. “Everything should be as simple as possible, but no simpler - ALBERT EINSTEIN.” and stable RMSE values.

As Figure 8 shows, the upper part of the graphs shows the CV 10 times mean square error (MSE) in the values of λ and the lower part shows the coefficient values of the predictors against the contraction parameters. The first and second vertical dashed lines correspond to the value of λ with the minimum MSE and the maximum value of λ within a standard error of the minimum MSE. T indicates how much the coefficients can be constrained while still predicting with maximum accuracy. In the figures shown, the y-axis represents the MSE values, while the upper x-axis represents the number of predictors. Therefore, a total of 152 data in 11 indicators were used to identify the optimal smoothing parameter λ from several models. Therefore, we reduced the data noise, which makes the model more accurate in detecting real signals.

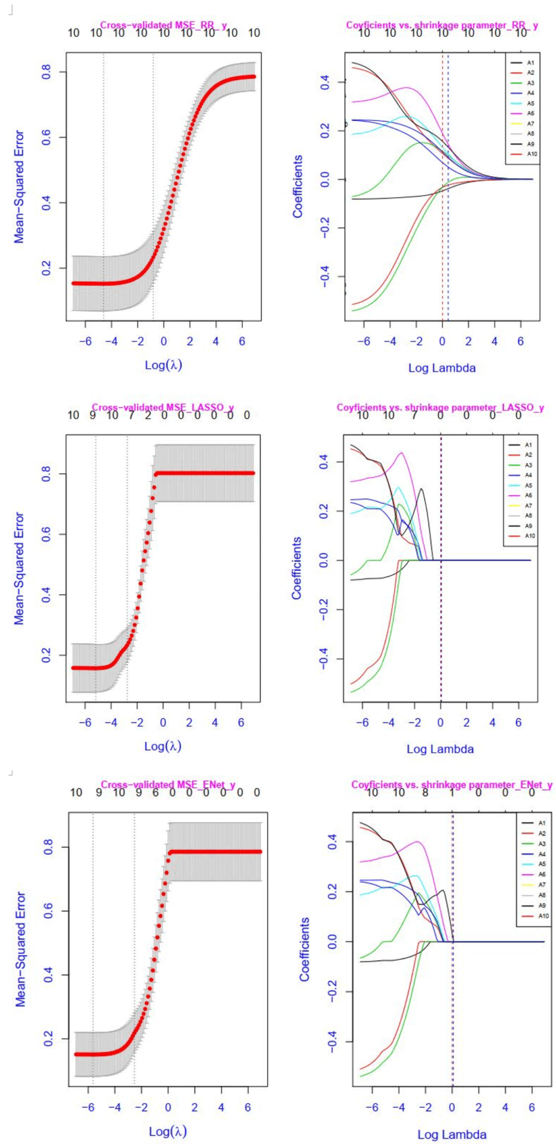

Figure 8 Equal CV mean error (RMSE) in λ values for EF

For this reason, the indicators of the top 8 tourism attracting countries were used to determine the optimal smoothing parameter λ. As shown in Figure 8, by default, the number of dots is 10 and alpha is (0.1 and 0.5). RR reduced the coefficient of the variables almost to zero, but did not reduce any of them to zero, and all variables remained in the model. In fact, RR has reduced correlated variables relative to each other until one is positive and the other is negative. As shown in Figure 6, the improvement of the model and automatic variable selection is shown by increasing the values of λ through the Lasso method. Our output shows a decrease in the number of predictor variables in the Lasso model when $\log(\lambda) \rightarrow$ -5. These variables are likely to have a strong relationship with other variables and cause their coefficients to swell. The advantage of EN is that it allows adjustment via RR with the Lasso variable selection feature.

The graphs of the RMSE test with the value of λ and the practical aspects of the PR model are shown in Figure 8 for EF.

4 Discussion

In recent years, ecological footprint problems in the world have become a serious concern among government managers, economists and environmentalists. This concern is more in the top tourism attracting countries, which attracted about 508 million tourists in 2019. Reports show that between 2018 and 2019, the growth rate of tourism in the United States of America decreased by 0.6% and in the rest of the countries, including France, by..., in China, by 4.9%, in Italy by 4.8%, in Turkey 11.9%, Mexico 9%, Thailand 4.3% and Germany 1.8% increased tourism attraction (https://en.wikipedia.org/wiki/World_Tourism_rankings). These 8 countries have taken a very large share of the gross national product (GNP) and international trade through tourism. On the other hand, despite attracting tourism, they face environmental challenges and problems.

For example, based on the theoretical support of the results of this research in these countries, the international transportation sector is one of the important factors influencing the ecological footprint of the world. Changes and developments in the flow of tourism increase the arrival and departure of tourists, respectively, and direct more transportation services [85]. Therefore, transportation, air and sea travel are one of the most important producers of greenhouse gas emissions, because the primary needs of transportation fuel used by air, road, railway and water are provided from fossil energy sources [22]. Therefore, the use of fossil energy in the top tourism attracting countries, in the creation of ecological footprints can lead to various consequences at the global level, such as global warming, polar ice melting, sea level rise, flooding of some farms, and an increase in tsunamis [86]. On the other hand, with the increase in the number of tourists and investment in all kinds of infrastructure services such as accommodation, hotels, restaurants, airports, ports, roads, railways, and telecommunications, these services will increase and play an essential role in expanding the ecological footprint [87].

Therefore, considering the difference in terms of (economic, institutional, technological, infrastructure, human capital and environmental awareness) Which exists between the most visited countries, five developed countries (France, China, the United States, Italy and Germany) and three developing countries. (Mexico, Turkey and Thailand) [88], countries may apply different alternative policies to manage their ecological footprint.

In line with the answer to the goal raised in this research, mixed and penalized effects models were used. This study first investigated the estimation of research indicators with MCMC models. Based on the findings of Table 4 and 5 and Figure 7, very high intra-group correlation between the indicators was observed, which showed the correctness of using the model. Also, one of the oldest and simplest algorithms that perform better than fancy and complex models is the linear regression model (OLS). But OLS methods can only determine the average effects of the influencing factors on the dependent variable and have serious defects in the analysis when the predictor groups are highly related and produce unstable results. Therefore, to fix some of the weaknesses of this method, penalized regression models such as ridge, lasso, and elastic net were used to improve the performance of the variables. The advantage of these models is for two reasons. Because it can manage multiple lines of the model and select the model as well. The use of these methods provides the possibility of testing a large number of predictive variables. In fact, it introduces biases in the estimation of models and reduces the mean square error of the satisfying variable.

In this research, we have shown how ML algorithms are more reliable in estimating production process parameters than classical statistical models. The use of ML models can help to plan the ecological footprint of countries attracting tourism, to reduce both economic cost and biological degradation. In this research, it was shown that a wide range of tourism developments, including import and export costs, costs of transportation items, incomes and the number of tourist visits, affect the ecological footprint and the evidence of the relationship of the positive effects of these indicators on the ecological footprint was presented. In addition, we have shown how the changes and transformations of tourism have increased the ecological footprint in the top tourism attracting countries. Also, the results of using the models showed that the ecological footprint in China, France and Italy was higher than other countries and researchers like [73, 74, 78] confirmed it in their research. On the other hand, according to the report of the World Tourism Organization, these countries ranked 1st, 4th and 5th in terms of attracting international tourism [89]. Therefore, the analysis of tourism data can be an added value to pay attention to the ecological footprint in the top countries of tourism attraction. In addition, using cross-validation, we have shown that the use of the "non-standard" approach of the lasso regression method is more accurate in predicting the ecological footprint compared to the "standard" linear regression. Also, choosing Lasso regression comes with the caveat that variables not selected by the model (with zero coefficients) can still be predictors, especially when there is a high correlation between selected and unselected variables.

Table 5 shows, MSE, RMSE models were better than OLS model in forecasting performance of new data. And the results obtained are consistent with the studies and literature in this field. By studying and estimating the determinants of individual tourism costs using Skate Elastic Net, Giambona and Grassini (2020) [90] concluded that there are significant effects of tourism spending on some accommodation facilities such as expensive houses, personal experiences from past visits, doing specific activities, accommodation and seasonality. They suggested that choosing rational indicators can help managers to take preventive and protective measures.

Also, using traditional methods such as OLS and Ridge, Fan et al. (2016) [91] stated that they could not meet the interpretability and forecasting accuracy at the same time. They stated that the use of Lassone's method can not only solve the above problems, but also reduce the calculation complexity. Hence, by using the employed indicators, the findings provide empirical support from previous studies. Therefore, the use of penalized models such as lasso plays a very important role in estimating the factors that determine the selection of indicators to support the decision-making process [92]. Dorugade (2014) [93] suggested that LASSO, unlike Ridge and OLS, in addition to selecting variables, makes it easier to interpret the regression model. Also, Zhou et al. (2011) [94] used logistic regression in the process of choosing between tourism and other goods.

The results indicate the existence of the effects of tourism on the ecological footprint in the top countries of tourism attraction, Adedoyin et al. (2021) [24] suggested that the development of tourism can provide the basis for environmental destruction. In addition, the increase in tourist arrivals increases the ecological footprint in both the long and short term. There is a positive and negative relationship between tourism development indicators and ecological footprint. The literature trends show that the development of tourism will lead to environmental destruction and environmental pollution such as CO2 and PM2 emissions and will increase the challenges of the ecological footprint [95].

Therefore, in general, it can be said that among the contraction methods, Lasso is more economical in selecting variables and has performed better. It means that this model was generous in the selection of variables. Also, ridge regression is very appropriate when there are a large number of predictor variables whose coefficients are non-zero and are extracted from the normal distribution. That is, this model showed that there is a large number of variables or severe multiple collinearities. The Elastic Net model showed that it performs variable selection and regularization at the same time, and grouping and variable selection are key roles in this model, and it uses two regression models, Ridge and Lasso.

5 Conclusion

Evaluating the effects of tourism development on the ecological footprint in the top tourism attraction countries can be of great importance. This research has been written with the help of tourism development indicators to highlight the importance of tourism indicators and explain the ecological footprint. This effort has opened a new way to discuss the effects of tourism development on the ecological footprint using mixed and penalized effects models. The main findings from the previous analysis show that the ecological footprint is able to address serious issues. But it is possible when all the policies related to the tourism policy of a country are well explained. Investigating the effects of tourism development on the ecological footprint is a means for the top tourism attracting countries to pay attention to tourism policy so that they can use appropriate methods to reduce environmental damage. Examining the statistical and empirical view of tourism development indicators and its effects on the ecological footprint in top tourism destination countries is of great importance. This article was written with a statistical reflection and criticism of the perspective of tourism development in order to highlight the importance of tourism indicators and explain the ecological footprint. This effort has opened a new way in discussing the effects of tourism development on the ecological footprint using MCMC and penalized models. The findings from previous analyzes show that dealing with the ecological footprint is a serious issue, but it is possible when the entire tourism policy of a country is well explained. The review of this article is a means for the countries of the world to pay attention to the top tourism destinations, so that they can use appropriate methods to reduce the environmental damage.

In this study, we presented the performance of using MCMC and penalized models due to the effects of tourism development indicators on the ecological footprint of top tourism destination countries. We have shown that trying to apply a penalized regression model to the ecological footprint can indeed be useful. Most of these methods perform variable selection, and this is a process in which the number of independent variables for prediction is reduced. These are well-known methods that can provide very accurate predictions. The results showed that the use of all statistical methods along with other existing modeling methods are useful and competitive.

The prediction results of the model produced by the above techniques were compared using root mean square error (RMSE) and coefficient of determination (R2). Based on the RMSE results, the prediction accuracy for the EF values obtained from all the competing models are very close, but the RMSE and R2 results have shown that the sequential LASSO model has performed better than the other competing models. Although this data set demonstrates the simplicity and potential superiority of the LASSO model, LASSO is closely adapted to the training data, and it requires ingenuity to use its extensive flexibility in choosing the estimation method to achieve high-accuracy predictions. In addition, Elastic Net and LASSO can play an important role in tourism and ecological footprint studies that have a large number of parameters. In such cases, these techniques are used to analyze the changes and evolutions of tourism on the ecological footprint, they are the best choice for modeling and forecasting this type of research.

Ethical approval

This paper does not contain any studies with human participants or animals performed by any of the authors.

Competing interests

The authors declare that they have no conflict of interest

Availability of data and materials

The datasets used and/or analyses during the current study are available from the corresponding author on reasonable request.

References

[1] Miller DS. Abrupt new realities amid the disaster landscape as one crisis gives way to crises. Worldwide Hospitality and Tourism Themes. 2021, 13(3): 304-311.
https://doi.org/10.1108/whatt-02-2021-0037

[2] Lee CC, Chen MP. Ecological footprint, tourism development, and country risk: International evidence. Journal of Cleaner Production. 2021, 279: 123671.
https://doi.org/10.1016/j.jclepro.2020.123671

[3] Nepal R, Indra al Irsyad M, Nepal SK. Tourist arrivals, energy consumption and pollutant emissions in a developing economy–implications for sustainable tourism. Tourism Management. 2019, 72: 145-154.
https://doi.org/10.1016/j.tourman.2018.08.025

[4] Roumiani A, Shayan H, Sharifinia Z, et al. Estimation of ecological footprint based on tourism development indicators using neural networks and multivariate regression. Environmental Science and Pollution Research. 2022, 30(12): 33396-33418.
https://doi.org/10.1007/s11356-022-24471-x

[5] WTTC. Travel & tourism economic impact 2017 European Union LCU. London: World Travel & Tourism Council, 2019.
https://travelbi.turismodeportugal.pt

[6] Cavalheiro MB, Joia LA, Cavalheiro GM do C. Towards a Smart Tourism Destination Development Model: Promoting Environmental, Economic, Socio-cultural and Political Values. Tourism Planning & Development. 2019, 17(3): 237-259.
https://doi.org/10.1080/21568316.2019.1597763

[7] Chen HS, Hsieh T. An environmental performance assessment of the hotel industry using an ecological footprint. Journal of Hospitality Management and Tourism. 2011, 2(1): 1-11.

[8] Dogan E, Aslan A. Exploring the relationship among CO 2 emissions, real GDP, energy consumption and tourism in the EU and candidate countries: Evidence from panel models robust to heterogeneity and cross-sectional dependence. Renewable and Sustainable Energy Reviews. 2017, 77: 239-245.
https://doi.org/10.1016/j.rser.2017.03.111

[9] Gokmenoglu KK, Eren BM. The role of international tourism on energy consumption: empirical evidence from Turkey. Current Issues in Tourism. 2019, 23(9): 1059-1065.
https://doi.org/10.1080/13683500.2019.1574723

[10] Işik C, Doğan E, Ongan S. Analyzing the Tourism–Energy–Growth Nexus for the Top 10 Most-Visited Countries. Economies. 2017, 5(4): 40.
https://doi.org/10.3390/economies5040040

[11] Robaina-Alves M, Moutinho V, Costa R. Change in energy-related CO2 (carbon dioxide) emissions in Portuguese tourism: a decomposition analysis from 2000 to 2008. Journal of Cleaner Production. 2016, 111: 520-528.
https://doi.org/10.1016/j.jclepro.2015.03.023

[12] Rendeiro Martín-Cejas R, Pablo Ramírez Sánchez P. Ecological footprint analysis of road transport related to tourism activity: The case for Lanzarote Island. Tourism Management. 2010, 31(1): 98-103.
https://doi.org/10.1016/j.tourman.2009.01.007

[13] Abid M, Schilling J, Scheffran J, et al. Climate change vulnerability, adaptation and risk perceptions at farm level in Punjab, Pakistan. Science of The Total Environment. 2016, 547: 447-460.
https://doi.org/10.1016/j.scitotenv.2015.11.125

[14] Anwar S, Hamilton J. Tourism into the Future—Towards 2020, and Beyond. Tourism Recreation Research. 2005, 30(3): 77-85.
https://doi.org/10.1080/02508281.2005.11081489

[15] Daigle C, Vasseur L. Is It Time to Shift Our Environmental Thinking? A Perspective on Barriers and Opportunities to Change. Sustainability. 2019, 11(18): 5010.
https://doi.org/10.3390/su11185010

[16] Enfedaque ND, Martínez JMD. El impacto económico de los desastres naturales. eXtoikos. 2014, 15: 99-101.

[17] Kattumuri R. Technology as a Driver for Governance by the People for the People. Governance and Governed. Published online 2018: 155-168.
https://doi.org/10.1007/978-981-10-5963-6_9

[18] Rutty M, Scott D. Will the Mediterranean Become "Too Hot" for Tourism? A Reassessment. Tourism and Hospitality Planning & Development. 2010, 7(3): 267-281.
https://doi.org/10.1080/1479053x.2010.502386

[19] Shakouri B, Khoshnevis Yazdi S, Ghorchebigi E. Does tourism development promote CO2emissions? Anatolia. 2017, 28(3): 444-452.
https://doi.org/10.1080/13032917.2017.1335648

[20] Wang B, Zheng X, Zhang H, et al. Bacterial community responses to tourism development in the Xixi National Wetland Park, China. Science of The Total Environment. 2020, 720: 137570.
https://doi.org/10.1016/j.scitotenv.2020.137570

[21] Khan A, Bibi S, Ardito L, et al. Revisiting the Dynamics of Tourism, Economic Growth, and Environmental Pollutants in the Emerging Economies—Sustainable Tourism Policy Implications. Sustainability. 2020, 12(6): 2533.
https://doi.org/10.3390/su12062533

[22] Koçak E, Ulucak R, Ulucak ZŞ. The impact of tourism developments on CO2 emissions: An advanced panel data estimation. Tourism Management Perspectives. 2020, 33: 100611.
https://doi.org/10.1016/j.tmp.2019.100611

[23] Bildirici ME, Bakirtas T. The relationship among oil, natural gas and coal consumption and economic growth in BRICTS (Brazil, Russian, India, China, Turkey and South Africa) countries. Energy. 2014, 65: 134-144.
https://doi.org/10.1016/j.energy.2013.12.006

[24] Adedoyin FF, Alola UV, Bekun FV. On the nexus between globalization, tourism, economic growth, and biocapacity: evidence from top tourism destinations. Environmental Science and Pollution Research. 2021, 29(17): 24995-25005.
https://doi.org/10.1007/s11356-021-17651-8

[25] Gössling S, Peeters P. Assessing tourism's global environmental impact 1900–2050. Journal of Sustainable Tourism. 2015, 23(5): 639-659.
https://doi.org/10.1080/09669582.2015.1008500

[26] Becken S, Simmons DG. Chapter 13. Tourism, Fossil Fuel Consumption and the Impact on the Global Climate. Tourism, Recreation and Climate Change. Published online December 31, 2005: 192-206.
https://doi.org/10.21832/9781845410056-015

[27] Ulucak R, Lin D. Persistence of policy shocks to Ecological Footprint of the USA. Ecological Indicators. 2017, 80: 337-343.
https://doi.org/10.1016/j.ecolind.2017.05.020

[28] Bello FG, Lovelock B, Carr N. Enhancing community participation in tourism planning associated with protected areas in developing countries: Lessons from Malawi. Tourism and Hospitality Research. 2016, 18(3): 309-320.
https://doi.org/10.1177/1467358416647763

[29] Galli A, Iha K, Halle M, et al. Mediterranean countries' food consumption and sourcing patterns: An Ecological Footprint viewpoint. Science of The Total Environment. 2017, 578: 383-391.
https://doi.org/10.1016/j.scitotenv.2016.10.191

[30] Cooper JA, McCullough BP. Bracketing sustainability: Carbon footprinting March Madness to rethink sustainable tourism approaches and measurements. Journal of Cleaner Production. 2021, 318: 128475.
https://doi.org/10.1016/j.jclepro.2021.128475

[31] Godil DI, Sharif A, Rafique S, et al. The asymmetric effect of tourism, financial development, and globalization on ecological footprint in Turkey. Environmental Science and Pollution Research. 2020, 27(32): 40109-40120.
https://doi.org/10.1007/s11356-020-09937-0

[32] Hunter C, Shaw J. The ecological footprint as a key indicator of sustainable tourism. Tourism Management. 2007, 28(1): 46-57.
https://doi.org/10.1016/j.tourman.2005.07.016

[33] Johnson PA. Exploring the ecological footprint of tourism in Ontario (Master's thesis, University of Waterloo), 2003.

[34] Lin W, Li Y, Li X, et al. The dynamic analysis and evaluation on tourist ecological footprint of city: Take Shanghai as an instance. Sustainable Cities and Society. 2018, 37: 541-549.
https://doi.org/10.1016/j.scs.2017.12.003

[35] Nathaniel SP, Barua S, Ahmed Z. What drives ecological footprint in top ten tourist destinations? Evidence from advanced panel techniques. Environmental Science and Pollution Research. 2021, 28(28): 38322-38331.
https://doi.org/10.1007/s11356-021-13389-5

[36] Ozturk I, Al-Mulali U, Saboori B. Investigating the environmental Kuznets curve hypothesis: the role of tourism and ecological footprint. Environmental Science and Pollution Research. 2015, 23(2): 1916-1928.
https://doi.org/10.1007/s11356-015-5447-x

[37] Cherlin S, Howey RAJ, Cordell HJ. Using penalized regression to predict phenotype from SNP data. BMC Proceedings. 2018, 12(S9).
https://doi.org/10.1186/s12919-018-0149-2

[38] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer New York, 2009.
https://doi.org/10.1007/978-0-387-84858-7

[39] Hoerl AE, Kennard RW. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics. 1970, 12(1): 55-67.
https://doi.org/10.1080/00401706.1970.10488634

[40] Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996, 58(1): 267-288.
https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

[41] Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2005, 67(2): 301-320.
https://doi.org/10.1111/j.1467-9868.2005.00503.x

[42] Friedman J, Hastie T, Tibshirani R. glmnet: Lasso and elastic-net regularized generalized linear models. R package version. 2009, 1(4): 1-24.

[43] Chen J, de Hoogh K, Gulliver J, et al. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. Environment International. 2019, 130: 104934.
https://doi.org/10.1016/j.envint.2019.104934

[44] Van der Kooij A, Meulman J. Regularization with ridge penalties, the lasso, and the elastic net for regression with optimal scaling transformations. Submitted for publication, 2008.

[45] Yu L, Ma X, Wu W, et al. A novel elastic net-based NGBMC(1,n) model with multi-objective optimization for nonlinear time series forecasting. Communications in Nonlinear Science and Numerical Simulation. 2021, 96: 105696.
https://doi.org/10.1016/j.cnsns.2021.105696

[46] Wackernagel M, Yount JD. Environmental Monitoring and Assessment. 1998, 51(1/2): 511-529.
https://doi.org/10.1023/a:1006094904277

[47] Schroeder H, Lovell H. The role of non-nation-state actors and side events in the international climate negotiations. Climate Policy. 2012, 12(1): 23-37.
https://doi.org/10.1080/14693062.2011.579328

[48] Grasso M, Giugni M. The Routledge Handbook of Environmental Movements. Routledge, 2021.
https://doi.org/10.4324/9780367855680

[49] Lenzen M, Sun YY, Faturay F, et al. The carbon footprint of global tourism. Nature Climate Change. 2018, 8(6): 522-528.
https://doi.org/10.1038/s41558-018-0141-x

[50] Gössling S, Hansson CB, Hörstmeier O, et al. Ecological footprint analysis as a tool to assess tourism sustainability. Ecological Economics. 2002, 43(2-3): 199-211.
https://doi.org/10.1016/s0921-8009(02)00211-2

[51] Işık C, Ongan S, Bulut U, et al. Reinvestigating the Environmental Kuznets Curve (EKC) hypothesis by a composite model constructed on the Armey curve hypothesis with government spending for the US States. Environmental Science and Pollution Research. 2021, 29(11): 16472-16483.
https://doi.org/10.1007/s11356-021-16720-2

[52] Venetoulis J, Talberth J. Refining the ecological footprint. Environment, Development and Sustainability. 2007, 10(4): 441-469.
https://doi.org/10.1007/s10668-006-9074-z

[53] Ansari MA, Villanthenkodath MA. Does tourism development promote ecological footprint? A nonlinear ARDL approach. Anatolia. 2021, 33(4): 614-626.
https://doi.org/10.1080/13032917.2021.1985542

[54] Antonakakis N, Dragouni M, Eeckels B, et al. The Tourism and Economic Growth Enigma: Examining an Ambiguous Relationship through Multiple Prisms. Journal of Travel Research. 2017, 58(1): 3-24.
https://doi.org/10.1177/0047287517744671

[55] Kongbuamai N, Zafar MW, Zaidi SAH, et al. Determinants of the ecological footprint in Thailand: the influences of tourism, trade openness, and population density. Environmental Science and Pollution Research. 2020, 27(32): 40171-40186.
https://doi.org/10.1007/s11356-020-09977-6

[56] Muchapondwa E, Stage J. The economic impacts of tourism in Botswana, Namibia and South Africa: Is poverty subsiding? Natural Resources Forum. 2013, 37(2): 80-89.
https://doi.org/10.1111/1477-8947.12007

[57] Dwyer L, Forsyth P, Spurr R, et al. Estimating the carbon footprint of Australian tourism. Journal of Sustainable Tourism. 2010, 18(3): 355-376.
https://doi.org/10.1080/09669580903513061

[58] Liu Y, Sadiq F, Ali W, et al. Does tourism development, energy consumption, trade openness and economic growth matters for ecological footprint: Testing the Environmental Kuznets Curve and pollution haven hypothesis for Pakistan. Energy. 2022, 245: 123208.
https://doi.org/10.1016/j.energy.2022.123208

[59] Lin W, Li Y, Li X, et al. The dynamic analysis and evaluation on tourist ecological footprint of city: Take Shanghai as an instance. Sustainable Cities and Society. 2018, 37: 541-549.
https://doi.org/10.1016/j.scs.2017.12.003

[60] Chakraborty A. Can tourism contribute to environmentally sustainable development? Arguments from an ecological limits perspective. Environment, Development and Sustainability. 2020, 23(6): 8130-8146.
https://doi.org/10.1007/s10668-020-00987-5

[61] Jafari J. Research and scholarship: the basis of tourism education. Journal of tourism studies. 1990, 1(1): 33-41.

[62] Alola AA, Eluwole KK, Lasisi TT, et al. Perspectives of globalization and tourism as drivers of ecological footprint in top 10 destination economies. Environmental Science and Pollution Research. 2021, 28(24): 31607-31617.
https://doi.org/10.1007/s11356-021-12871-4

[63] Telfer DJ, Sharpley R. Tourism and Development in the Developing World. Routledge, 2015.
https://doi.org/10.4324/9781315686196

[64] Bohdanowicz P. Environmental awareness and initiatives in the Swedish and Polish hotel industries—survey results. International Journal of Hospitality Management. 2006, 25(4): 662-682.
https://doi.org/10.1016/j.ijhm.2005.06.006

[65] Dasgupta S, De Cian E. The influence of institutions, governance, and public opinion on the environment: Synthesized findings from applied econometrics studies. Energy Research & Social Science. 2018, 43: 77-95.
https://doi.org/10.1016/j.erss.2018.05.023

[66] Sajjad F, Noreen U, Zaman K. Climate change and air pollution jointly creating nightmare for tourism industry. Environmental Science and Pollution Research. 2014, 21(21): 12403-12418.
https://doi.org/10.1007/s11356-014-3146-7

[67] Croall J. Preserve or destroy: tourism and the environment. Calouste Gulbenkian Foundation, 1995.

[68] Akinboade OA, Braimoh LA. International tourism and economic development in South Africa: a Granger causality test. International Journal of Tourism Research. 2009, 12(2): 149-163.
https://doi.org/10.1002/jtr.743

[69] Myers RH, Myers RH. Classical and modern regression with applications, Duxbury press Belmont, 1990.

[70] Waldmann P, Mészáros G, Gredler B, et al. Evaluation of the lasso and the elastic net in genome-wide association studies. Frontiers in Genetics. 2013, 4.
https://doi.org/10.3389/fgene.2013.00270

[71] Seltzer MH. Sensitivity Analysis for Fixed Effects in the Hierarchical Model: A Gibbs Sampling Approach. Journal of Educational Statistics. 1993, 18(3): 207-235.
https://doi.org/10.3102/10769986018003207

[72] SEEGER M. GAUSSIAN PROCESSES FOR MACHINE LEARNING. International Journal of Neural Systems. 2004, 14(02): 69-106.
https://doi.org/10.1142/s0129065704001899

[73] Guo J, Ren J, Huang X, et al. The Dynamic Evolution of the Ecological Footprint and Ecological Capacity of Qinghai Province. Sustainability. 2020, 12(7): 3065.
https://doi.org/10.3390/su12073065

[74] Bi M, Xie G, Yao C. Ecological security assessment based on the renewable ecological footprint in the Guangdong-Hong Kong-Macao Greater Bay Area, China. Ecological Indicators. 2020, 116: 106432.
https://doi.org/10.1016/j.ecolind.2020.106432
[75] Zhang X, Xu L, Chen Y, et al. Emergy-based ecological footprint analysis of a wind farm in China. Ecological Indicators. 2020, 111: 106018.
https://doi.org/10.1016/j.ecolind.2019.106018
[76] Wang Y, Shen H, Ye S, et al. Being rational and emotional: An integrated model of residents' support of ethnic tourism development. Journal of Hospitality and Tourism Management. 2020, 44: 112-121.
https://doi.org/10.1016/j.jhtm.2020.05.008
[77] Kenner D. Inequality of overconsumption: The ecological footprint of the richest. Global Sustainability Institute, 2015, 1-18.
[78] Pata UK, Yilanci V. Financial development, globalization and ecological footprint in G7: further evidence from threshold cointegration and fractional frequency causality tests. Environmental and Ecological Statistics. 2020, 27(4): 803-825.
https://doi.org/10.1007/s10651-020-00467-z
[79] González-Abraham C, Ezcurra E, Garcillán PP, et al. The Human Footprint in Mexico: Physical Geography and Historical Legacies. Silva LCR, ed. PLOS ONE. 2015, 10(3): e0121203.
https://doi.org/10.1371/journal.pone.0121203
[80] Madlener R, Sunak Y. Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management? Sustainable Cities and Society. 2011, 1(1): 45-53.
https://doi.org/10.1016/j.scs.2010.08.006
[81] Chomkhamsri K, Mungcharoen T, Yuvaniyama C. 10-year experience with the Thai national LCI database: case study of "refinery products." The International Journal of Life Cycle Assessment. 2016, 22(11): 1760-1770.
https://doi.org/10.1007/s11367-016-1160-3
[82] Destek MA, Sinha A. Renewable, non-renewable energy consumption, economic growth, trade openness and ecological footprint: Evidence from organisation for economic Co-operation and development countries. Journal of Cleaner Production. 2020, 242: 118537.
https://doi.org/10.1016/j.jclepro.2019.118537
[83] Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992, 7(4).
https://doi.org/10.1214/ss/1177011136
[84] Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics. 1998, 7(4): 434-455.
https://doi.org/10.1080/10618600.1998.10474787
[85] Dogan E, Seker F, Bulbul S. Investigating the impacts of energy consumption, real GDP, tourism and trade on CO2emissions by accounting for cross-sectional dependence: A panel study of OECD countries. Current Issues in Tourism. 2015, 20(16): 1701-1719.
https://doi.org/10.1080/13683500.2015.1119103
[86] Spencer A. Climate Change and the Sustainable Development of Tourism Within the Caribbean. Travel and Tourism in the Caribbean. Published online October 5, 2018: 27-46.
https://doi.org/10.1007/978-3-319-69581-5_2
[87] Endo K. Foreign direct investment in tourism—flows and volumes. Tourism Management. 2006, 27(4): 600-614.
https://doi.org/10.1016/j.tourman.2005.02.004
[88] Chan APC, Darko A, Olanipekun AO, et al. Critical barriers to green building technologies adoption in developing countries: The case of Ghana. Journal of Cleaner Production. 2018, 172: 1067-1079.
https://doi.org/10.1016/j.jclepro.2017.10.235
[89] World Tourism Organization, 2019.
https://www.unwto.org/tourism-statistics
[90] Giambona F, Grassini L. Tourism attractiveness in Italy: Regional empirical evidence using a pairwise comparisons modelling approach. International Journal of Tourism Research. 2019, 22(1): 26-41.
https://doi.org/10.1002/jtr.2316
[91] Fan GF, Peng LL, Hong WC, et al. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. Neurocomputing. 2016, 173: 958-970.
https://doi.org/10.1016/j.neucom.2015.08.051
[92] Saayman M, Saayman A. Determinants of Spending: an Evaluation of Three Major Sporting Events. International Journal of Tourism Research. 2011, 14(2): 124-138.
https://doi.org/10.1002/jtr.841
[93] Dorugade AV. New ridge parameters for ridge regression. Journal of the Association of Arab Universities for Basic and Applied Sciences. 2014, 15(1): 94-99.
https://doi.org/10.1016/j.jaubas.2013.03.005
[94] Zhou J, Yu L, Gao X, et al. Plasma MicroRNA Panel to Diagnose Hepatitis B Virus–Related Hepatocellular Carcinoma. Journal of Clinical Oncology. 2011, 29(36): 4781-4788.
https://doi.org/10.1200/jco.2011.38.2697
[95] Xiangyu S, Jammazi R, Aloui C, et al. On the nonlinear effects of energy consumption, economic growth, and tourism on carbon footprints in the USA. Environmental Science and Pollution Research. 2021, 28(16): 20128-20139.
https://doi.org/10.1007/s11356-020-12242-5