ENSEMBLE
A bi-lingual peer reviewed academic journal
ISSN: 2582 - 0427 (Online)
Vol. 2, No. 2
September, 2020
http://www.ensembledrms.in

Article Type: Research Article
Article Ref. No.: 200215107N4AASA
https://doi.org/10.37948/ensemble-2020-0202-a001

A STATISTICAL ANALYSIS AND PREDICTION OF CARBON DIOXIDE EMISSION IN HIMACHAL PRADESH, PUNJAB AND HARYANA STATES OF INDIA

Kartick Mondal[1], Sudipta Sinha[2] ✉, Pijush Basak[3], D.P. Goswami[4]

Abstract:

The universe is facing heavily the evil effects of Global warming, which is a burning issue today. With the advancement of civilization, new industries have been set-up to enhance production as per the need of society along with the population explosion. India is familiar as a developing country in the world, and at present, India has a significant contribution to the augmentation of Greenhouse Gases in the environment, which induces global warming. Global warming has serious effects of worsening the environment. It causes the melting of ice, extinction of species, prevalance of several fatal diseases, loss of biodiversity etc. In this paper, we develop a statistical model that analyzes and predicts the trend of emission of Carbon dioxide in Himachal Pradesh, Punjab, and Haryana. For this purpose, we have considered the dataset of 21 years of gas emission and tried to fit a fourth-degree polynomial curve by the least square method. We have tried to sketch a comparative scenario of carbon-dioxide emissions in the chosen states along with the forecast of the long-term evolutionary trend of the emission. The Instantaneous Rate of Change (IROC)analysis has been employed for this purpose. The efficacies of the model have been tested by residual analysis, coefficient of determination R^2, and adjusted R^2.

Article History: Submitted on 15 Feb 2020 | Accepted on 07 Jun 2020 | Published online on 01 May 2021

Keywords: Global warming, Green House Gases (GHG), Mathematical model, Regression analysis, Instantaneous Rate of Change (IROC).

1.0 Introduction:

Greenhouse gases are the natural part of the atmosphere of the earth, which absorbs and re-radiates the sun's warmth to maintain the earth's temperature at a level essential to protect life (Basak et al, 2015). However, due to indiscriminate human activities, the amount of the gases has been increased to an alarming stage and the Earth is heating beyond control. Environmentalists and scientists are reiterating again and again about the existence of human being. Climate is changing beyond our perception. The term 'Global warming' and 'Climate Change' are

1 [Author] Assistant Professor, PRMS Mahavidyalaya, Baragari, Jambani, Bankura, West Bengal, INDIA
2 [Author] ✉ [Corresponding Author] Assistant Professor, Burdwan Raj College, Burdwan, West Bengal, INDIA. E-mail: sudipta.sinha08@gmail.com
3 [Author] Assistant Professor, Dream Institute of Technology, Thakurpukur, Kolkata, West Bengal, INDIA
4 [Author] Assistant Professor, Netaji Subhas Engineering College, Kolkata, West Bengal, INDIA

complementary. Global warming refers to gradual increase in average surface temperature of the Earth. Greenhouse gases are one of the causes for it, whereas the word Climate Change is used in a broader sense (Howarth, 2014). The latter indicates long-term changes in climatic conditions, for example, changes in rainfall and average temperature etc., due to global warming (Dewrite,2018). Khansis et al. (2005) discussed the relation between GHG emission and human diseases. According to the Indian Network for Climate Change Assessment (INCCA) report pressed by Montek Singh Ahluwalia in the year 2007, India ranked fifth behind the USA, China, European Union, and Russia with respect to aggregate Green House Gas (GHG) emissions in the world. The report also points out that the 2007 emissions of the USA and China are almost four times than that of India. The report mainly enlightens on emissions from different sectors such as Energy, Industry, Land-use, Land-use Change, Forestry (LULUF) and Waste. The net GHG emissions from India, that is, emissions with LULUCF, are reported to be 1727.71 million tons of CO_2 equivalents in 2007. Out of this, CO_2 emissions were 1221.76 million tons. An analytical study regarding the dynamics of CO_2 emission was done by Asadoorian et al. (2006). Mondal et al (2019[10,11]) developed models in the aspects of CO_2 emission in different States of the eastern region of India.

This work is an effort to formulate a mathematical model for the emission of carbon dioxide (CO_2) in three selected states of the northern part of India, namely, Himachal Pradesh, Punjab, and Haryana. It is based on historical data of about 21 years (Basak et al.,2014[2,3]). Among these, data of 18 years is utilized to formulate the proposed model and the data of the remaining three years is used to validate the model. A statistical method involving the least-squares method and regression analysis have been employed for this purpose. An attempt has been made to study a comparative scenario of carbon dioxide emission in these chosen states along with the forecast in the long term or short term evolutionary trend of the emission of the same utilizing the Instantaneous Rate of Change (IROC) analysis.

2.0 Materials and Methods:

2.1 Method of Least Squares

We formulated a fourth-degree polynomial model for analyzing CO_2 emission of Himachal Pradesh, Punjab and Haryana. For this purpose, we have followed the work of Ghosal et al. (2008), Jin et al. (2010), Nandi et al (2014), Tokos et al. (2009) for generating the model of CO_2 emission. The proposed fourth degree polynomial model for emission of the gas is

$$Y = \alpha + \beta x + \gamma x^2 + \delta x^3 + \mu x^4$$

(Equation 1)

where Y is emission of carbon dioxide in '000 Metric ton of carbon and x represents time in years.

We have observed that several researchers (Basak et al etc) have used up to third degree (3rd) polynomial but none of them have used 4th degree polynomial. We have tried to fit the curve with the fourth-degree polynomial to increase the accuracy.
With the help of the given data $\{(x_1, y_1), (x_2, y_2), \dots \dots \dots . (x_n, y_n)\}$ we may define the error associated by

$$\omega(\alpha, \beta, \gamma, \delta, \mu) = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i - \gamma x_i^2 - \delta x_i^3 - \mu x^4)^2$$

(Equation 2)

is a function of five variables $\alpha,\ \beta,\ \gamma\ ,\ \delta$ and μ. For minimizing the error and estimating corresponding α, β, δ and μ multivariate calculus is used to calculate

$$2^{nd} \text{ order } \frac{\partial\omega}{\partial\alpha}=0,\ \frac{\partial\omega}{\partial\beta}=0,\ \frac{\partial\omega}{\partial\gamma}=0,\ \frac{\partial\omega}{\partial\delta}=0,\ \frac{\partial\omega}{\partial\mu}=0$$

Differentiating (2) partially and equating to zero we get

$$-2\sum\left(y_i-\alpha-\beta x_i-\gamma x_i^2-\delta x_i^3-\mu x_i^4\right)=0$$

$$-2\sum x_i(y_i-\alpha-\beta x_i-\gamma x_i^2-\delta x_i^3-\mu x_i^4)=0$$

$$-2\sum x_i^2(y_i-\alpha-\beta x_i-\gamma x_i^2-\delta x_i^3-\mu x_i^4)=0$$

$$-2\sum x_i^3(y_i-\alpha-\beta x_i-\gamma x_i^2-\delta x_i^3-\mu x_i^4)=0$$

$$-2\sum x_i^4(y_i-\alpha-\beta x_i-\gamma x_i^2-\delta x_i^3-\mu x_i^4)=0$$

Now, the corresponding normal equations are:

$$\sum y_i=n\alpha+\beta\sum x_i+\gamma\sum x_i^2+\delta\sum x_i^3+\mu\sum x_i^4$$

$$\sum y_i x_i=\alpha\sum x_i+\beta\sum x_i^2+\gamma\sum x_i^3+\delta\sum x_i^4+\mu\sum x_i^5$$

$$\sum x_i^2 y_i=\alpha\sum x_i^2+\beta\sum x_i^3+\gamma\sum x_i^4+\delta\sum x_i^5+\mu\sum x_i^6$$

$$\sum x_i^3 y_i=\alpha\sum x_i^3+\beta\sum x_i^4+\gamma\sum x_i^5+\delta\sum x_i^6+\mu\sum x_i^7$$

$$\sum x_i^4 y_i=\alpha\sum x_i^4+\beta\sum x_i^5+\gamma\sum x_i^6+\delta\sum x_i^7+\mu\sum x_i^8$$

For the given set of points (x_i, y_i); (i=1, 2,…,n), the equations can be solved for $\alpha,\ \beta,\ \gamma,\ \delta$ and μ to find the estimated $\alpha^*\beta^*, \gamma^*, \delta^*$ *and* μ^*. It has been found that in all the cases, the value of the 2nd order derivatives evolves to be positive at the points $\alpha,\ \beta,\ \gamma,\ \delta$ and μ.These satisfy the minimization criteria of the function ω.

Thus, the 4th degree fitted polynomial of carbon dioxide emission is estimated as

$$Y=\alpha^*+\beta^* x+\gamma^* x^2+\delta^* x^3+\mu^* x^4$$

(Equation 3)

A footnote to show the condition for minimization criteria is incorporated at the bottom of the page[1].

2.2 Instantaneous rate of change of emission

To compute the rate of change of the gas, the derivative of equation (3) is presented as

$$\frac{dY}{dx}=\beta^*+\gamma^* x+\delta^* x^2+\mu^* x^3$$

(Equation 4)

[1]To show minimization criteria of a function of several variables, we determine the first derivatives with respect to the variables $\alpha, \beta, \gamma, \delta$ *and* μ and equal to zero. We have shown this condition in the text. Now to fulfil the minimization criteria, we have to find the 2nd order derivatives of the function with respect to the variables stated above and have to check whether these are greater than zero or not. These must be greater than zero and for the sufficient condition we also have to examine the Hessian matrix.
Mathematically,

$$\frac{\partial\omega}{\partial\alpha}=0;\ \frac{\partial\omega}{\partial\beta}=0;\ \frac{\partial\omega}{\partial\gamma}=0;\ \frac{\partial\omega}{\partial\delta}=0;\ \frac{\partial\omega}{\partial\mu}=0;$$

$$\frac{\partial^2\omega}{\partial\alpha^2}>0\ ;\frac{\partial^2\omega}{\partial\beta^2}>0\ ;\frac{\partial^2\omega}{\partial\gamma^2}>0\ ;\frac{\partial^2\omega}{\partial\delta^2}>0\ ;\frac{\partial^2\omega}{\partial\mu^2}>0$$

The equation (4) is utilized to predict the emitted CO_2 at a particular year.

3.0 Quality of the Proposed Model & Prediction:

Equation (4) may be used for estimating the emission of CO_2 for medium or short terms. The goodness of fit is estimated with different statistical tools like R^2, R^2 (adjusted) and residual analysis.

3.1 Coefficient of Determination (R^2)

The coefficient of determination(R^2) provides an overall measure of how well the model fits. The general definition of the coefficient of determination is

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}}$$

where $SS_{tot} = \sum_i (y_i - \bar{y})^2$,

$$SS_{reg} = \sum_i (f_i - \bar{y})^2 \quad ,SS_{reg} = \sum_i (y_i - f_i)^2$$

Here, SS_{tot} = Total sum of squares (proportional to the sample variance),

SS_{reg} = the regression sum of squares or the explained sum of square and SS_{err} = the sum of squares of residuals, also called the residual sum of square. y_i and f_i are observed and estimated values of CO_2 emission.

3.2 Adjusted Coefficient of Determination

The adjusted R^2 is defined as:

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}$$

$$= R^2 - (1 - R^2)\frac{p}{n-p-1}$$

where p is the total number of regressors in the model excluding the constant term and n is the sample size. It also helps us to understand how good the model fit with the observed data.

4.0 Results & Discussion:

4.1 Comparative CO_2 emission

For quantification of the magnitude of the CO_2 emission in the states of Himachal Pradesh, Punjab and Haryana, a comparative bar diagram of the emission data is given below (Fig.1).

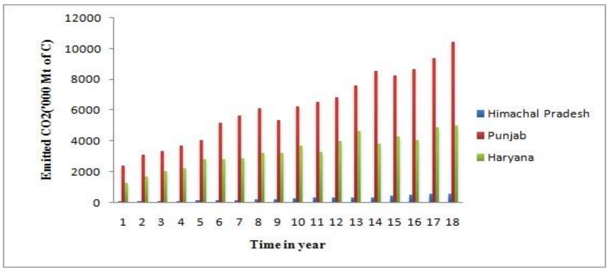

Fig.1: Comparative bar diagram of CO_2 emission.

In all the states, the increasing tendency of emission is observed whereas the emission of Punjab is followed by Haryana and Himachal Pradesh. The emission is very high in Punjab and very low in Himachal Pradesh. The emission pattern of the Punjab and Himachal Pradesh is mostly increasing, but for Haryana, both increasing and decreasing pattern of emission can be observed.

4.2 Fourth degree polynomial model formulation

The polynomials representing emission of CO_2 for the states of Himachal Pradesh (HP), Punjab and Haryana are estimated as follows

$$Y(HP) = 1.6573858097700201 \times 10^{10} - 3.3247154667943 \times 10^{7}x$$

$$+25011.340276606108x^2 - 8.36291790572811x^3 + 0.0010486521318882529x^4$$

$$Y(Punjab) = 1.453699556868005 \times 10^{12} - 2.930051012726789 \times 10^{9}x + 2.214645798759466 \times 10^{6}x^2 - 743.9576636241314x^3 + 0.0937175372590689x^4$$

$$Y(Haryana) = 5.896096865022722 \times 10^{11} - 1.1896066461363006 \times 10^{9}x + 900048.4717402834x^2 - 302.64847040384217x^3$$

$$+0.03816234302018486x^4$$

where Y is the emission of CO_2 in '000 Mt. of carbon and x represents time in years (1980 is considered as the first year)

A graphical representation of the observed data and the fitted model for emission of CO_2 for the states of Himachal Pradesh, Punjab and Haryana are compared in Fig. 2(a), 2(b) and 2(c) respectively.

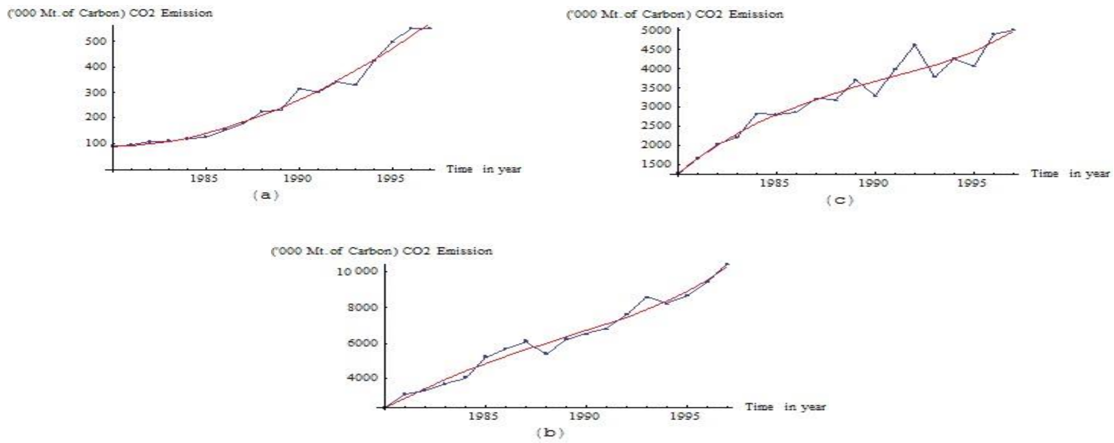

Fig. 2: CO_2 emission in (a) Himachal Pradesh; (b) Punjab and (c) Haryana

It is seen from the graphical representation that the emission of CO_2 in Punjab is maximum and in Himachal Pradesh, it is minimum. In all three cases, emission growth is rapid. The rapid growth reflects the growth of emissions from industry, agriculture, forestry and waste products etc. The goodness of fit of the model is tested with different statistical tools such as residual analysis, coefficient of determination R^2 and adjusted R^2.

We have used software Mathematica 7.0 to solve the normal equations but any numerical analysis software can be used for the same.

4.3 Calculations of R^2& R^2 (adjusted)

The numerical values of R^2 and R^2 (adjusted) are listed in Table 1.

Table 1: Numerical value of R^2 and R^2 (adjusted).

State	R^2	R^2 (adjusted)
Himachal Pradesh	0.982861	0.977587
Punjab	0.979637	0.973371
Haryana	0.946329	0.929815

The magnitude of R^2 and R^2 (adjusted) indicates that the chosen model performs well to represent the emission of CO_2 in the respective zones.

4.4 Instantaneous Rate of Change (IROC)

IROC may be utilized to predict the future emission of the gas at any particular year. The IROC of CO_2 emission for the states of Himachal Pradesh (HP), Punjab and Haryana can be presented as

$$\frac{dY}{dx}(\text{HP}) = -3.3247154667943 \times 10^7 + 50022.680553212216x$$
$$-25.088753717184332\,x^2 + 0.0041946085275530115\,x^3$$

(Equation 5.a)

$$\frac{dY}{dx}(\text{Punjab}) = -2.930051012726789 \times 10^9 + 4429291.597518932x$$
$$-2231.8729908723944x^2 + 0.3748470149036275x^3$$

(Equation 5.b)

$$\frac{dY}{dx}(\text{Haryana}) = -1.1896066461363006 \times 10^9 + 1800096.9434805668x$$
$$-907.94541121115266x^2 + 0.1526493720807394x^3$$

(Equation 5.c)

where x represents time in years(1980 as the $1st\ Year$)

The graphical visualization of the above expressions 5(a-c) is given below in Fig 3(a-c).

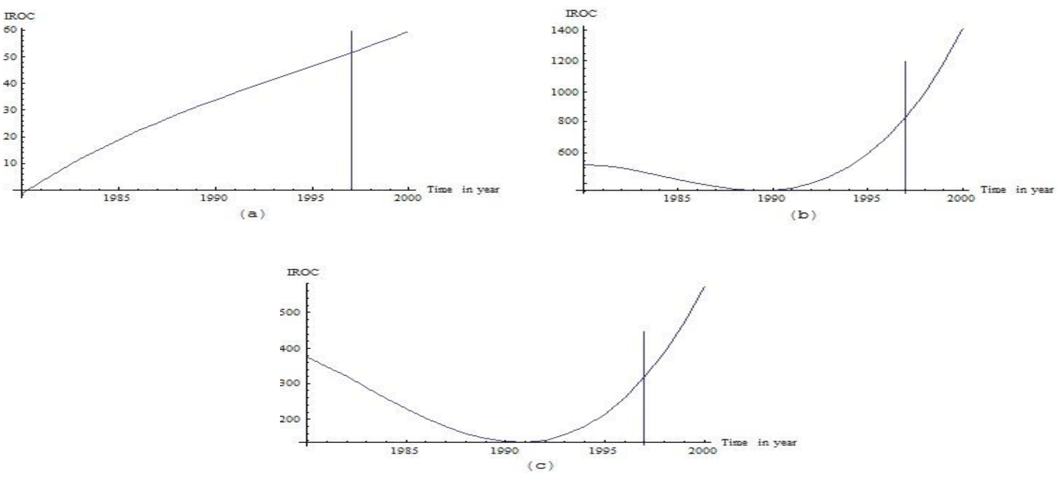

Fig. 3: IROC of (a) Himachal Pradesh; (b) Punjab and (c) Haryana.

- In Himachal Pradesh, the IROC is negative before the year of 1981and an increasing trend is followed after that.
- Decreasing rate of emission tendency has been visualized at the beginning but after the year of 1989, the rapidly increasing trend in the rate of emission continues for the state of Punjab.
- In the state of Haryana, decreasing rate of emission trend is followed up to the year 1991, and rapidly increasing rate of emission tendency starts from the year 1992.
- In all three cases, the vertical line indicates the completion of 18 years of estimated IROC values and using it, one can predict the future IROC values for CO_2 emission.

4.5 *Residual analysis*

Table 2: Residual analysis of CO_2 emission for the state of Himachal Pradesh

Year	Real data ('000 Mt. of carbon)	Model data ('000 Mt. of carbon)	Residuals
1980	88.75	92.07	3.32
1981	93.77	92.87	−0.89
1982	107.79	98.18	−9.6
1983	109.47	107.7	−1.76
1984	118.68	121.15	2.47
1985	125.38	138.26	12.88
1986	153.42	158.81	5.39
1987	177.5	182.59	5.09
1988	222.71	209.41	−13.29
1989	228.78	239.12	10.34
1990	316.68	271.58	−45.09
1991	301.62	306.7	5.08
1992	343.48	344.39	0.91
1993	331.76	384.59	52.83
1994	422.39	427.27	4.88
1995	498.16	472.43	−25.73
1996	547.55	520.08	−27.46
1997	549.65	570.27	20.62
Standard Deviation of the residuals			20.77
Mean of residuals			0.00003

Table 3: Residual analysis of CO_2 emission for the state of Punjab

Year	Real data ('000 Mt. of carbon)	Model data ('000 Mt. of carbon)	Residuals
1980	2377.12	2400.33	23.21
1981	3114.93	2920.25	−194.68
1982	3332.82	3428.54	95.71
1983	3694.72	3918.29	223.56
1984	4026.68	4384.83	358.15
1985	5185.4	4825.75	−359.64
1986	5638.13	5240.88	−397.24
1987	6109.5	5632.30	−477.19
1988	5370.01	6004.35	634.34
1989	6225.66	6363.60	137.94
1990	6544.86	6718.89	174.02
1991	6830.78	7081.28	250.5
1992	7629.28	7464.12	−165.15

1993	8572.63	7882.98	−689.65
1994	8235.86	8355.68	119.82
1995	8646.74	8902.31	255.56
1996	9408.62	9545.17	136.55
1997	10434.64	10308.87	−125.77
Standard Deviation of the residuals			331.86
Mean of residuals			0.0026

Table 4: Residual analysis of CO_2 emission for the state of Haryana

Year	Real data ('000 Mt. of carbon)	Model data ('000 Mt. of carbon)	Residuals
1980	1258.36	1287.04	28.67
1981	1651.86	1649.61	−2.24
1982	2009.99	1983.22	−26.77
1983	2207.57	2286.81	79.23
1984	2817.07	2560.25	−256.81
1985	2788.61	2804.36	15.74
1986	2860.82	3020.81	159.98
1987	3229.41	3212.23	−17.17
1988	3192.99	3382.17	189.18
1989	3698.06	3535.07	−162.98
1990	3307.49	3676.31	368.81
1991	3983.13	3812.16	−170.96
1992	4625.08	3949.84	−675.24
1993	3786.60	4097.44	310.84
1994	4262.36	4264.02	1.66
1995	4068.34	4459.51	391.17
1996	4901.59	4694.79	−206.8
1997	5007.92	4981.62	−26.29
Standard Deviation of the residuals			251.65
Mean of residuals			0.0014

For all the three states, the result of statistical analysis reflects that a good model is identified.

Table 5: Himachal Pradesh

Year	Observed Data	Estimated Data
1998	488.22	623.081
1999	682.98	678.59
2000	659.12	736.923

Table 6: Punjab

Year	Observed Data	Estimated Data
1998	9392.09	11220.2
1999	11012.55	12308.3
2000	10845.72	13604.3

Table 7: Haryana

Year	Observed Data	Estimated Data
1998	5532.66	5332.71
1999	5250.84	5761.67
2000	5460.46	6283.03

4.6 Future predictions

Using the proposed model, the emission of CO_2 is predicted for the years 1998, 1999, and 2000 for all of the above states and the result is compared with the actual emission data. The result is shown in the Table No. 2, 3 and 4 respectively.

5.0 Conclusion:

Mathematical models using the least-squares method have been developed for the characterization of the behavioural pattern of CO_2 emission for the three states in India. It has been observed that the pattern of growth is not uniform; it depends on time and zone. The proposed model is found to be in good agreement with the CO_2 emission data in the considered time span. The analytical expressions for emission have been tested with the statistical procedures R^2 and R^{2}-adjusted to identify the efficacy of our mathematical expression, and each of them reflects the reality very well.

A theoretical basis for future researches regarding CO_2 emission in different regions in India may be obtained from this study and the model may be utilized for functional planning and strategic applications for the reduction of appalling global warming in the near future.

References:

Asadoorian,M.O.,Sarofim,M.C.,Reilly,J.M.,Paltsev,S.&Forest,C.C.,(2006).Historical Anthropogenic Emissions Inventories for Greenhouse Gases and Major Criteria Pollutants. MIT Joint Program on the Science and Policy of Global Changes, Technical Note No. 8.

Basak,P.&Nandi,S.,(2014). An analytical study of emission dynamics of carbon dioxide in India. IOSR Journal of Applied Chemistry,1,16-21.

Basak,P&Nandi,S.A.,(2014).Statistical analysis and prediction of carbon dioxide emission in some eastern and northern states of India. International Journal of Environmental Sciences,4 (5), 956-967.

Basak, P. & Nandi, S.(2015). Emission dynamics of some green house gases and criteria pollutants in Indian perspective, International Journal on Recent and Innovation Trends in Computing and Communication ,3(2),001-007.

Dewitte, S.&Clerbaux, N.,(2018). Decadal Changes of Earth's Outgoing Longwave Radiation.Remote Sens., 10(10), 1539.

Ghoshal,T&Bhattacharyya,R.,(2008).State level carbon dioxide emissions of India 1980-2000, Contemporary Issues and Ideas in Social Sciences.

Howarth,R.W.,(2014). A bridge to nowhere: methane emissions and the green house gas footprint of natural gas. Energy Science & Engineering, 2(2),47-60.

Jin,R., Tian, L.,Qian J. & Liu Y.(2010),The Dynamic evolutionary analysis on carbon emissions in Yangtze delta. International Journal of Nonlinear Science,10(3), 259-263.

Khansis,A.A., &Nettleman,M.D.,(2005). Global warming and infectious disease.Archieves of Medical Research ,36(6), 689-696.

Mondal,K., Basak,P.&Sinha,S.,(2019). Comparative analysis of carbon dioxide emission in West Bengal and India, International Journal for Research in Engineering Application and Management,5(2),348-352.

Mondal,K., Basak,P.&Sinha,S.,(2019).A Mathematical Analysis and Prediction of Carbon Dioxide Emission in Assam, West Bengal and Orissa, India , 12(5), 19-25.

Nandi,S.&Basak,P.,(2014).Emission of carbon dioxide from different attributes in India: A mathematical study. Journal of Applied Chemistry,1, 06-10.

Tokos,C.P.&Xu,Y.,(2009). Modelling carbon dioxide emissions with a system of differential equations. Non-linear analysis: Theory, Methods and Applications,71(12),1182-1197.

Acknowledgment

1. The CO_2 emission data is exported from the paper of Ghoshal and Bhattacharyya, 2008.
2. We went through the report 'Carbon Intensity- Himachal Pradesh, 2012-13' of 'Department of Environment, Science & Technology, Government of Himachal Pradesh'.