MATHEMATICAL MODELING OF BLACK CARBON IN THREE MAJOR COAL MINES OF INDIA

A Thesis

Submitted in partial fulfillment of the requirements for the

award of the degree of

DOCTOR OF PHILOSOPHY

in

Mathematics

By

Sidhu Jitendra Singh Makkhan

(Regd. No. 41400171)

Supervised By

Dr. Sachin Kaushal

Co-Supervised by

Dr. Kulwinder Singh Parmar

LOVELY PROFESSIONAL UNIVERSITY

PUNJAB

2020

DECLARATION

I, declare that this written submission entitled **"Mathematical Modeling of Black Carbon in Three Major Coal Mines of India"** represents my ideas in my own words and where others ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be the cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Name: Sidhu Jitendra Singh Makkhan

Regd. No.: 41400171

Date:

Place: Lovely Professional University, Punjab.

CERTIFICATE

This is to certify that the thesis entitled **"Mathematical Modeling of Black Carbon in Three Major Coal Mines of India"** submitted by Mr. Sidhu Jitendra Singh Makkhan to Lovely Professional University, Punjab for the award of the degree of Doctor of Philosophy is a bonafide record of research work carried out by him under our supervision. It is hereby certified that the declaration made by the candidate is correct to the best of our knowledge and that this thesis is fit to be considered for the award of the degree of Doctor of Philosophy.

Date:

Place: Lovely Professional University, Punjab.

Supervisor,
Dr. Sachin Kaushal,
Associate Professor,
Lovely Professional University,
Phagwara, Punjab.

Co-Supervisor,
Dr. Kulwinder Singh Parmar,
Assistant Professor,
I. K. Gujral Punjab Technical University,
Kapurthala, Punjab.

ABSTRACT

The main objective of this thesis entitled "**Mathematical Modeling of Black Carbon in Three Major Coal Mines of India**" is to analyze the black carbon concentration level in the three major coal mine regions of India namely; Raniganj, Jharia, and Bokaro. We have selected these three coal mines for our study since; Raniganj is the oldest coal mine of the country operating since 1774. Jharia has the largest coal reserves of about 19.4 billion tons and is also known throughout the world as the 'Burning coalfield' due to its underground fires from the past many decades. Bokaro is known as the 'Steel city of India' for which the Bokaro coalfields play a very crucial role by providing its continuous supply of coal. Thus these coal mines play a vital role in the growth and development of the region and the nation by contributing to its GDP and employment in the region. But on the other hand, continuous opencast mining in these regions has led to worsening the air pollution problem in the region of their location and the neighboring states especially the IGP (Indo-Gangetic Plain) region, which is a hot topic of research nowadays. We herewith the help of Mathematics aim to develop an efficient forecasting model for the prediction of concentration for black carbon in the three major coal mines of India by using Statistical, Time-series, Wavelet, and Soft computing techniques. The results and analysis are based on continuous time-series data of black carbon concentration for the past 38 years, 05 months from Jan 1980 to May 2018 over these sample sites. The data has been collected by NASA (National Aeronautics and Space Administration) via Giovanni website (http://nasa.gov/) and processing of the data is done by CSIR National Physical Laboratory, New Delhi. The work done in this study is divided into five chapters.

Chapter 1 provides a brief introduction of the topic along with the literature review; starting with the introduction on; role and importance of coal in India, coal mining activities, Black carbon with its impacts on the health of the nation and the climate, and discussing the present scenario of PM (Particulate matter) and Black carbon. This chapter also highlights the various Statistical techniques like central tendency

measures, skewness measures, kurtosis, coefficient of variation, correlation, and regression analysis used for studying the nature of the time-series of the three sample sites. It also gives a brief introduction about the various time-series models, wavelet analysis, Fourier transforms, wavelet transforms along with some important types of wavelets. Soft computing based artificial neural network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) used in the last chapter are discussed in detailed to make it convenient to apprehend and understand the work done. In the later part of the chapter, we have discussed the various model fitting parameters used for the selection of the models along with the motivation for the problem and a brief introduction to our sample sites with the respective time-series.

Chapter 2 begins with a formal introduction and review of the literature on statistical analysis. The statistical analysis of the problem is further carried out in three parts consisting of; basic statistical analysis, trend analysis, and correlation analysis of the problem. In the first part, basic statistical parameters as discussed above are used for analyzing the nature of the three time-series. In the second part, we have carried out a trend analysis of the problem after discussing the statistical parameters used like the Hurst exponent, Fractal dimension, predictability index, and trend percent for analyzing the behavior of the time-series. And in the third part, we have carried out correlation analysis using Karl Pearson's, Spearmen's, and Kendall's tau coefficient of correlation along with the cross-correlation function. The results conclude that for all the three time-series the nature of the distribution curve is leptokurtic and the shape of the distribution curve is close to a normal distribution. The trend pattern is observed to be anti-persistent for the three time-series and there is a high degree of correlation in the data values of the three sample sites. The three time-series data were found to follow a seasonal pattern.

Chapter 3 is based on the time-series analysis of the problem which is further carried out in two parts. The chapter begins with a review of the literature and an introduction to basic modeling concepts like the ACF, PACF, Box-Jenkins methodology; time-series models like the ARIMA, SARIMA. In the first part, the best-fitted time-series

model is developed using the popular Box-Jenkins methodology based ARIMA model. It was discovered that ARIMA $(1,0,1)(0,1,1)_{12,}$ model fitted very well to the data of the three time-series and it was used to obtain a time-series forecast of 5 years to the problem. In the second part, we have carried out an analysis of the volatility (conditional variance) using the GARCH models for the three time-series. The analysis was further carried out to find the best-fitted GARCH model to the problem and was used to obtain the forecast of conditional variance for the coming 5 years period. Model fitting parameters reveal that the models fitted very well to the time-series.

Chapter 4 deals with the development of a Wavelet-ARIMA coupled model and comparison of its results with the time-series ARIMA model to decide the best-fitted model for forecasting among the two. The chapter starts with a review of the literature followed by a detailed discussion about the Fourier Transform, Wavelet transforms, and Wavelet-ARIMA coupled approach. The coupled model was developed after a suitable wavelet decomposition of the time-series for the three sample sites by Daubechies wavelet db8, level 3. This decomposition led the original time-series $s(t)$ to be decomposed into three detailed parts d_1, d_2, d_3 and an approximation part a_3 such that $s = a_3 + d_1 + d_2 + d_3$ for all the time-series. These parts were further treated as individual sub-time-series and modeled with the ARIMA approach. The results obtained were then summed to reconstruct the time-series signal. The results obtained by the application of the ARIMA and Wavelet-ARIMA approach were then compared with each other. Based on the model fitting parameters it was found that the coupled approach gave superior results than the individual ARIMA model approach. The coupled model was then used to obtain a forecast for the coming 5 years period.

Chapter 5 presents the soft computing analysis of the problem and development of the hybrid ANNFWCM model termed as the 'Artificial Neural Network Fuzzy Wavelet Conjugation model'. This model is formed by the conjugation of ANN, wavelets, and fuzzy-interface system (FIS). The chapter starts with the review of the literature, an introduction to various components of soft computing followed by machine learning

algorithms and their important types. After a detailed discussion of ANN and ANFIS models along with their various components, the hybrid ANNFWCM model is developed. The time-series data of Raniganj was decomposed with the help of wavelet decomposition and the individual sub-time-series were then treated with the ANFIS model. Gauss membership function along with the Mamdani technique was used for building the Fuzzy inference system (FIS). The FIS is then treated with ANN's Feedforward neural network (FFNN) based backpropagation approach with 1000 Epochs for training the signals. After appropriate training, the signals were tested with the actual observations. The results obtained finally conclude that the hybrid model performed far better than the time series ARIMA and Wavelet-ARIMA coupled approach and can be considered as an ideal tool for the prediction of black carbon concentration over the coal mine regions.

ACKNOWLEDGEMENTS

This Ph.D. thesis is the culmination of a few years of fervent learning and research experience. Throughout these years, I have been fortunate to interact with many people who have influenced me greatly. One of the pleasures of finally finishing is this opportunity to thank them. This thesis is a reflection of not only my work but also the input of many, whose contributions I would like to acknowledge.

Firstly, I would like to express my sincere gratitude to my respected Advisor **Dr. Sachin Kaushal**, Associate Professor, Department of Mathematics, Lovely Professional University, Punjab, India and my respected Co-Advisor **Dr. Kulwinder Singh Parmar**, Assistant Professor, I. K. Gujral Punjab Technical University, Punjab, India for their unconditional support, constant inspiration and great patience, without which this whole work would never been accomplished. I truly appreciate and value their esteemed guidance and encouragement from the beginning to the end of this research work. Whatever words that I can write here to thank them shall be less than what my heart feels in regards to research guides like them.

I convey my heartiest regards to **Mr. Ashok Mittal**, Chancellor of Lovely Professional University, Punjab. I am immensely grateful to our respected Head of Faculty **Dr. Loviraj Gupta**, our dynamic Head of School **Dr. Ramesh Thakur**, and Head of Mathematics Department **Dr. Kulwinder Singh**.

I also thank the entire faculty members of the Mathematics Department, Lovely Professional University, Punjab of Interdisciplinary Research for their unhesitant aid whenever they were approached.

I am also thankful to **Dr. Kirti Soni**, Principal Scientist, National Physical Laboratory, CSIR, New Delhi for supporting and providing me valuable data for my research.

I am grateful to all the members of the Evaluation **(RAC)** committee for their fruitful discussions, suggestions, and directions.

I am extremely thankful to **Dr. Rekha** and **RDP** committee members for their valuable support and suggestions.

I am very grateful to my friends and colleagues especially **Dr. Sukhdev Singh** and **Mr. Sarbjit Singh** (Research scholar, GNDU) for the valuable time of learning I spent with them.

I am also grateful to my working institute 'Sri Guru Angad Dev College, Khadur sahib' and it's principal '**Dr. Kanwaljit Singh**' for providing me the necessary support for my research work.

Last but not the least; I would like to express my profound thanks to my parents, brother, wife, and my little son. It is their love, encouragement, and support that have supported me throughout the accomplishment of this Ph.D. program during tough times which I was passing through. Mere thanks would not suffice their sacrificed inspirations in completion of this work.

Sidhu Jitendra Singh Makkhan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

BC – Black Carbon

PM – Particulate Matter

GW – Gigawatt

CO – Carbon Monoxide

CO_2 – Carbon Dioxide

SO_2 – Sulphur Dioxide

SO_x – Oxides of Sulphur

NO_x – Oxides of Nitrogen

PI - Predictability Index

MFs - Membership Functions

FT – Fourier Transform

WT – Wavelet Transform

GBP – Geosphere Biosphere Programme

AQI - Air Quality Index

CBM – Coal Bed Mining

NPL - National Physical Laboratory (India)

GBD – Global Burden of Disease

GDP – Gross Domestic Product

LPG – Liquefied Petroleum Gas

TPPs – Thermal Power Plants

VOCs – Volatile Organic Compounds

IGP – Indo Gangetic Plain

SPM – Suspended Particulate Matter

WHO – World Health Organisation

DWT – Discrete Wavelet Transform

CWT – Continuous Wavelet Transform

STFT – Short-Time Fourier Transform

CPCB - Central Pollution Control Board

CMRI - Central Mining Research Institute (India)

ISRO - Indian Space Research Organisation

CSIR – Council of Scientific and Industrial Research

NASA - National Aeronautics and Space Administration

NCAP - National Carbonaceous Aerosols Program

DALYs – Disability Adjusted Life-Years

NAAQs – National Ambient Air Quality Standards

AR – Autoregressive

MA – Moving average

ARMA – Autoregressive Moving Average

ARIMA – Autoregressive Integrated Moving Average

SARIMA – Seasonal Autoregressive Integrated Moving Average

ARMAX – Autoregressive Moving Average with Exogenous Inputs

ARCH - Autoregressive Conditional Heteroscedasticity

GARCH – Generalized Autoregressive Conditional Heteroscedasticity

EGARCH – Exponential GARCH

GARCH-M – GARCH in Mean

AGARCH – Asymmetric GARCH

IGARCH - Integrated GARCH

NGARCH - Nonlinear GARCH

GJR - Glosten, Jagannathan, and Runkle Model

APARCH - Asymmetric Power ARCH Model

NAR – Non-linear Autoregressive

TAR - Threshold Autoregressive

NMA - Nonlinear Moving Average

ANN - Artificial Neural Network

FIS - Fuzzy Inference System

ANFIS - Adaptive Neuro-Fuzzy Inference System

FFNN - Feedforward Neural Network

RNN - Recurrent Neural Network

CNN - Convolutional Neural Network

RBNN - Radial Basis Neural Network

RBF - Radial Basis Function

MLP - Multi-Layered Perceptron

RMSE - Root Mean Square Error

MSE - Mean Square Error

MAPE - Mean Absolute Percentage Error

MAE - Mean Absolute Error

ACF – Autocorrelation Function

PACF - Partial Autocorrelation Function

AIC - Akaike Information Criterion

BIC - Bayesian Information Criterion

LLF – Log-Likelihood Function

CCF - Cross-Correlation Function

DET - Daily Effective Temperature

COVID – Coronavirus Disease

SPARK - Spatial Pattern Recognition via Kernels

MODIS – Moderate-Resolution Imaging Spectroradiometer

LSSVM - Least Square Support Vector Machine

ANNFWCM - Artificial Neural Network Fuzzy Wavelet Conjugation Model

Chapter 1 Introduction

"Coal kills- An assessment of death and disease caused by India's dirtiest energy source" (Source: Goenka and Guttikunda 2014[100])

In this chapter, we introduce the topic and point towards the reasons for the development and motivation of the research topic. A brief explanation of the growth of the air pollution problem due to the coal industry in India, the impact of black carbon on the health of nation and climate, review of literature is done. The use of various research techniques like statistical analysis, time series models, fractal analysis, GARCH models, wavelet analysis, soft computing techniques and coupled approaches are used for obtaining accurate forecasts of black carbon concentration over the major coal mines of India to estimate the future black carbon, air pollution level due to coal mines.

1.1 Role and Importance of Coal

In the present scenario, India as such being the third largest and one of the fastest-growing economies of the world, with over more than one billion of the population it becomes very difficult to catch up with ever-increasing demand for energy with limited available sources in the country **(Hubacek et al. 2007[118], Goenka and Guttikunda 2014[100], Agarwal et al. 2016[12])**. Natural resources play a vital role in the growth and meeting the increasing demand of energy for every nation **(Shih 2009[253], Zou et al. 2016[305])**. The majority of the population living in villages uses livestock dung, agricultural wastes, wood as the major source of energy **(Gautam et al. 2009[97], Kaygusuz 2011[133])**. While urban areas use natural gas, LPG, electricity, petroleum products, coal for their domestic, transportation, and industrial use **(Sovacool 2011[266]).**

Also, India is the fourth largest consumer of electricity in the world and more than 400 million of the population suffers from regular power cuts across the country due to unreliable power sources **(Jha et al. 2014[128], Tripathi et al. 2016[281])**. In this age of modernization, there is a continuous increase in the demand and use of these natural resources to meet the energy requirement of electricity for Industries, Cement and Steel

production units, Oil refineries, households, etc. Coal, hydropower, biofuels, nuclear power is the major sources of energy in India **(Singh and Gu 2010[255])**. Specifically, for countries like India where more than 75% of the total energy demand is accomplished from imported oil and gas **(Sharma et al. 2019[251])**.

In such, Coal plays a key role in fulfilling this ever-increasing demand, as India being the world's third-largest producer of coal after China and USA **(Raghuvanshi et al. 2006[219])**. It is the most widely available fossil fuel in the country and in the world. Coal is safe, reliable and has a relatively low cost; out of the total installed power capacity in 2014 of 250GW, nearly 148 GW comes from coal which is more than 59% of the total electricity production as given in Figure 1.1(a, b) **(Central Electrical Authority report, June 2014[52]; Jan 2018[51])**. As per the planning commission of India **(Planning Commission, 2002-2007[212]),** the total coal reserves in India are 245.69 billion tones out of which only 91.63 billion tones (approximately 38%) are a proven resource and only 21% of this is extractable. In addition to this coal mining is one of the major industries in India contributing to the GDP as well as employment to several million of the workforce **(Santra et al. 2014[245])**.

Coal mining is one of the major industries in India that contributes a big role in the economic progress of the nation **(Chaulya and Chakraborty 1995[59])**. The total coal consumption is expected to increase by 2-3 times from 660 million tons/year in 2014 to nearly 1800 million tons/year in 2030 **(Goenka and Guttikunda 2014[100])**. Thermal power plants (TPPs) are the major consumers of coal which consumes about 75% of the total coal generated in the country **(Singh et al. 2013[257])**. To meet this huge demand for consumption the total coal generation capacity is expected to increase from 159 GW in 2014 to 450 GW in 2030, which is nearly 3 times more than the present capacity **(Goenka and Guttikunda 2014[100])**. Also, coal prices are more stable than compared with oil and natural gas prices. Thus due to this growing demand for coal year by year has led to a continuous increase of new coal-fired power plants as shown in Figure 1.2 **(Goenka and Guttikunda 2014[100])**, which on the other side led to worsening the air

pollution problem day by day. A Schematics of Coal-fired thermal power plants working resulting in the production of electricity, fly-ash, air pollution is represented in Figure 1.3.

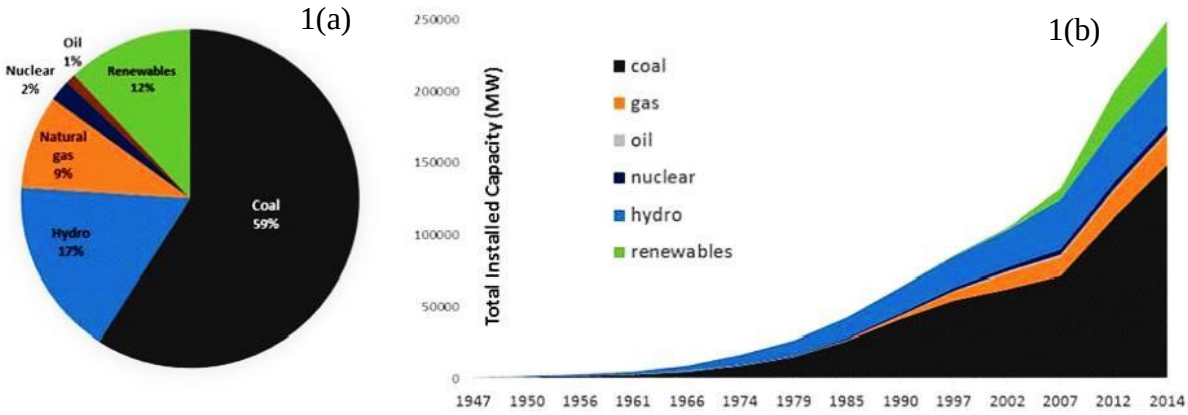


Figure 1.1(a) Sources of electricity production in India based on the installed capacity and **1.1(b)** Growth of installed power generation in India **(Source: Central Electrical Authority report, Jun 2014[52])**

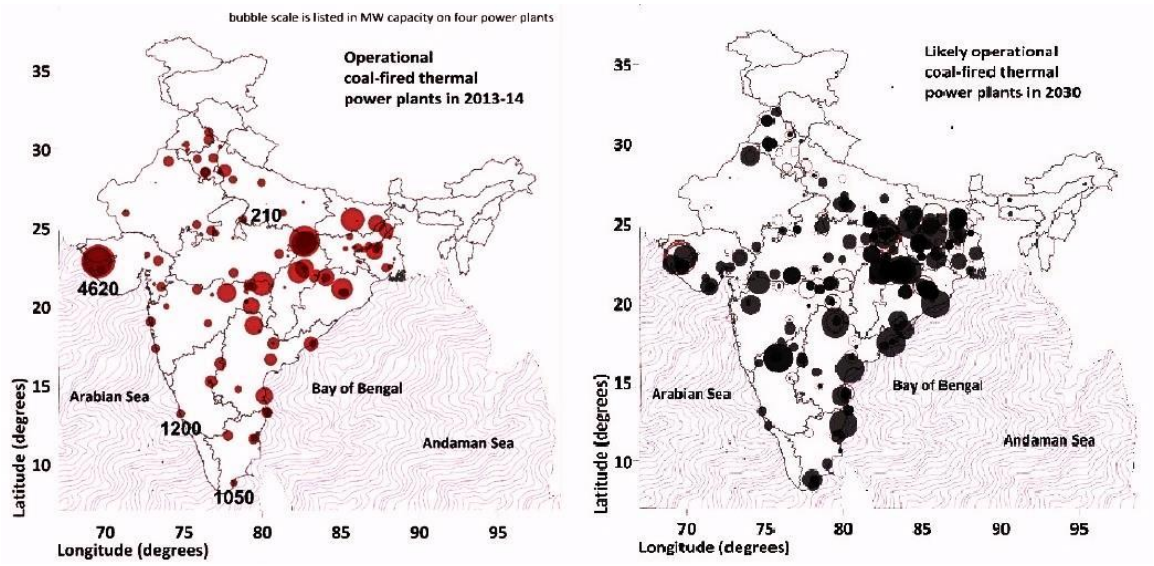


Figure 1.2 Coal-fired plants of India in 2013-14 and likely to be operational until 2030 **(Source: Goenka and Guttikunda 2014[100])**

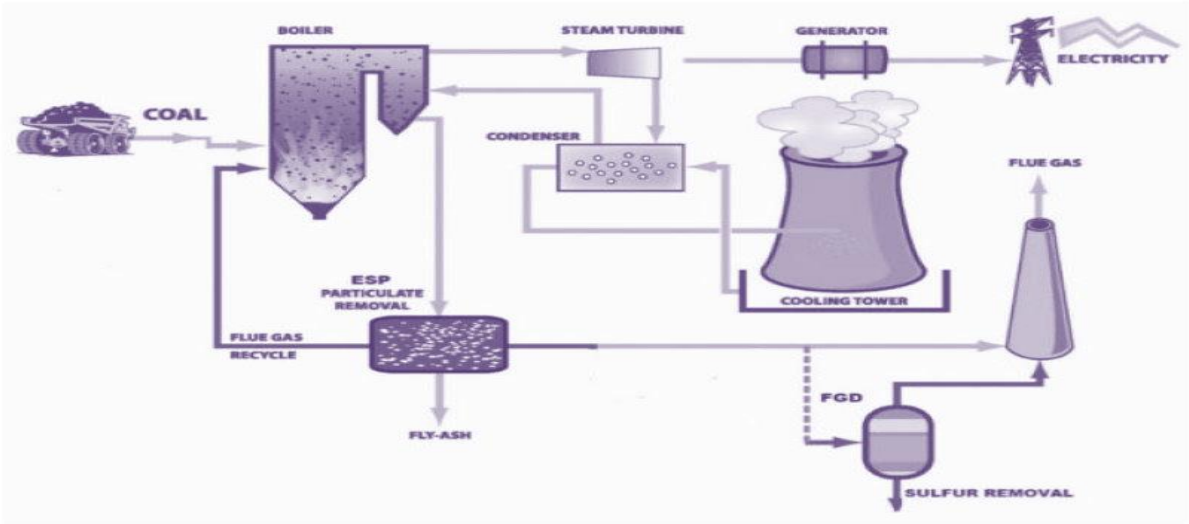


Figure 1.3 Schematics of Coal-fired thermal power plant

1.1.1 Coal Mining

Coal mining is mainly carried by two main methods open cast mining and underground mining. In India, more than 92% of coal mining is done by opencast mining **(Sahu et al. 2017[242])** as shown in Table 1.1. Opencast mining is mainly preferred in India since it is economical than underground mining, has lower extraction risk, and generates more production but the major limitation associated with it is that this leads to more air pollution than compared to underground mining **(Sahu et al. 2017[242])**. The first coal mine established in India was in 1774, in the Raniganj region of West Bengal **(Prasad 1992[215])**, after this, there was gradually very slow progress in the growth of the coalfields, far after in 1856 coal mines were established in Jharia, Bokaro and Karanpura in Jharkhand. During that period the production of coal continued to grow at a very steady phase and so India had to import coal, despite having huge coal deposits. During 1880, the annual import of coal was 600,000 tons which continued up to 1902 **(Gadgil 1971[95])**. At that time the main consumer of coal was Railways **(Gadgil 1971[95])**. Although things had changed with time in terms of demand, consumption, production of coal. But, the mineworkers have continued to live in a very poor and miserable working condition in terms of poverty and poor health conditions **(Kumaramangalam 1973[144],**

Kejriwal 2006[134]). Figure 1.4 shows a poor condition of women coal laborers working in the opencast Jharia coalfield of India.

Figure 1.4 Laborers working at the opencast coal fields in India

Table 1.1 Total coal production due to opencast mining in India and the percentage of total production due to opencast mining **(Source: Coal Controller Organization of India, Ministry Of Coal, 2019[63]).**

Year	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
Production (Million tons)	532.042	532.694	539.950	556.402	565.765	612.435	639.230	657.868	675.400	728.718
% of total production by opencast mining	89.0	89.70	90.38	90.62	91.22	92.09	92.74	93.26	93.81	94.17

1.1.2 Impacts of Coal Mining

As per a report from the Central pollution control board **(CPCB 2009[68])**, air pollution in coal mines is leading to spontaneous emissions of various other dangerous gases like black carbon, methane, carbon monoxide, oxides of nitrogen, and sulphur dioxide. Along with all these air pollutants like PM and VOCs are also released from coal mining regions while performing various activities like drilling, blasting, coal handling plants, stockyards, haul roads, transport roads, workshop exposed to overburden dumps, exposed pit faces, overburden loading, and unloading operations produce pollution **(CMRI 1998[53])**. The volume of these pollutants is expected to get at least double during this period from 2014 to 2030, polluting the environment to the greatest extent and having severe health impacts **(Goenka and Guttikunda 2014[100])**.

Major causes of air pollution in urban areas are industrialization, vehicular emission, brick kilns, oil refineries, etc. while in rural areas the major contributor is the burning of biofuels such as coal, wood, cow dung, and agricultural wastes **(Saud et al. 2011[247])**. $PM_{2.5}$ emission is expected to increase in India by nearly 25% in 2030, whereas nitrogen oxides (NO_x) and sulphur dioxide (SO_2) emissions will increase by 3-5 times, VOCs emission is expected to rise to 30% **(Purohit et al. 2010[216])**. The diurnal average $PM_{2.5}$ concentrations level in major Indian cities was observed to be $100-550\mu g/m^3$ in 2014 (Pant et al. 2016 **(Pant et al. 2016[195])**, while in rural parts of India was at $125\mu g/m^3$ which was twice the limit prescribed by CPCB **(Agarwal et al. 2012[13])**. Air pollution is a global problem and the main reason for ill health in India have a direct link with minor physiologic disturbances to major and multiple health disorders notably leading to premature deaths along with various types of cardiovascular problems like heart attack, strokes; respiratory problems, birth disorders and various types of cancer **(Bascom et al. 1996[31], Stieb et al. 2012[270])**.

The DALYs due to air pollution in Delhi has nearly doubled from 0.34 to 0.75 million during the years 1995 to 2015 and that in Mumbai has increased from 0.34 to 0.51 million. There have been 80,665 premature adult deaths during this period in Delhi

and Mumbai due to air pollution **(Borwankar 2017[42])**. As per the Global burden of disease (GBD) report for 1990-2010, more than 200 health risks causing premature deaths are related to air pollution. The topmost of these causes include dietary risks, high blood pressure, household air pollution, smoking, ambient PM pollution. There were more than 20 million asthma cases in 2011-12 from exposure due to air pollution and which is expected to grow to 42.7 million by 2030 and total pre matured deaths are expected to rise from 186,500 to 229,500 per year in 2030 due to coal-fired power plants **(Goenka and Guttikunda 2014[100])**. The monetary cost due to these health impacts exceeds Rs. 16,000 to Rs. 23,000 crores per year **(Goenka and Guttikunda 2014[100])**. According to the **World Health Organization (WHO) report 2009[291]**, due to air pollution approximately 3.1 million people in developing countries, die prematurely each year.

1.2 Black Carbon

Black carbon (BC) is a black colored sooty particle released during combustion activities of biomass, biofuels and fossil fuels from kitchens, vehicles, aircraft emission, heavy engines, brick kilns, industries, oil refineries, coal-fired thermal power plants, steel plants, petroleum industries, forest fires, burning of agricultural wastes **(Rehman et al. 2011[231])**. In indoor conditions, it is mainly released due to the cooking and burning of fuels like wood, coal, animal manure, residues of crops **(Andreae and Crutzen 1997[21])**. Black carbon is made up of pure or organic carbon in several linked forms **(Conny and Slater 2002[65])**, is the main product for combustion, BC is also known as soot. It is used in the manufacturing of plastics, paints and inks; it is also used as a color pigment. It is also widely used as a diesel oxidant in experiments. It is used in reinforcing fillers in tires or other rubber products. It can absorb plant nutrients from the soil and significantly contributes to its fertility when used in proper proportion. It compromises a significant portion of particulate matter, which is an air pollutant **(Bove et al. 2019[43])**.

Since BC is in form of particle thus it is not greenhouse gas and is referred as a component of particulate matter, the particles emitted during combustion are classified as

particulate matter (PM) in terms of their sizes those smaller than 10 micrometers are termed as PM_{10} or $PM_{2.5}$ means particles smaller than 2.5 micrometers **(Janssen et al. 2012[125], Begum et al. 2013[33])**. The size of black carbon is very small compared to that of even $PM_{0.1}$, which has severe impacts on human health and climate **(Schwarz et al. 2008[248], Brzezina et al. 2020[45]).** A pictorial representation of the cycle for black carbon emission due to different sources in the environment and its effects like absorption of incoming solar radiations leading to surface warming, melting of glaciers, affecting cloud formation and rainfall precipitation is shown in Figure 1.5.

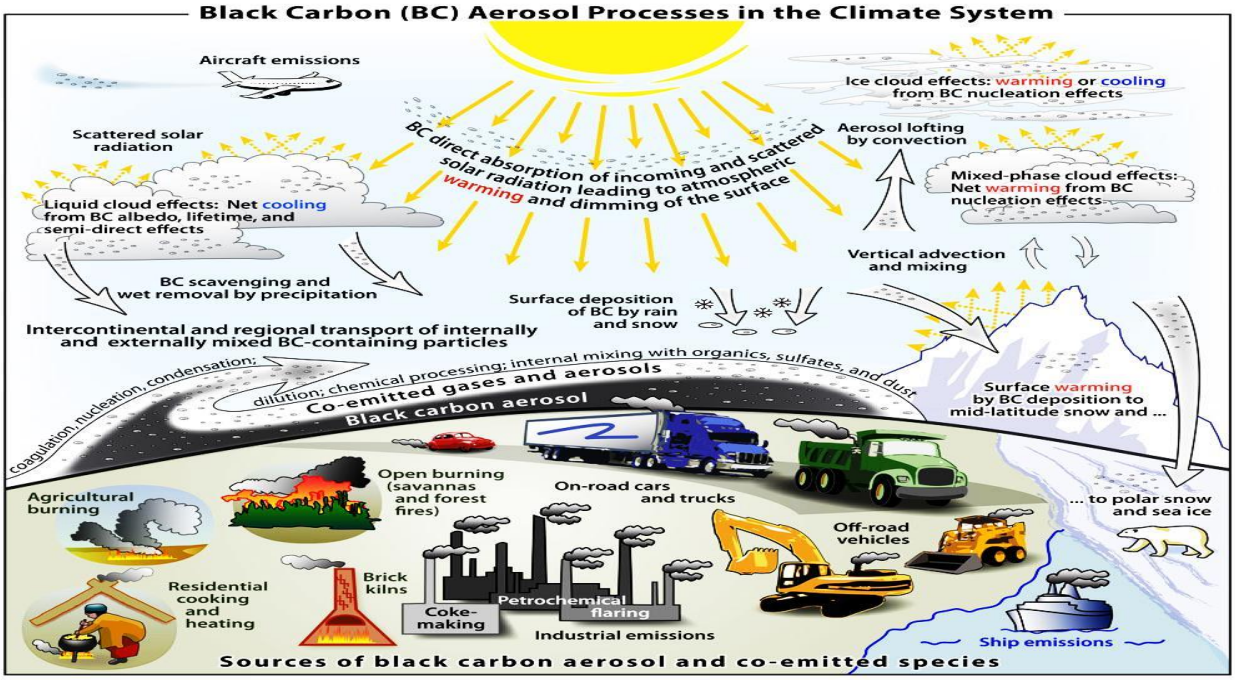


Figure 1.5 Cycle of black carbon in the environment

1.2.1 Impact of BC on Health of Nations

Carbonaceous aerosols have received great attention from the researchers recently due to its severe impact on human health **(Japar et al. 1986[126], Liousse et al. 1996[147], Tzanis and Varotsos 2008[283])**, agriculture **(Chameides et al. 1999[56])**, air quality, and health effects **(Jansen et al. 2005[124], Highwood and Kinnersley 2006[111])**. Compared to underground mining, open-cast mining is a more threatening problem for air pollution. It

is a global problem that has a negative impact on human health such as Inhalation of BC leads to problems related to respiratory such as chronic bronchitis and asthma, lung disease, damage to eye sights, cardiovascular disease **(Penner et al. 1993[210], Varotsos 2005[285])**. It also leads to multiple health disorders like lung cancer, heart attack, stroke, premature births and deaths; birth disorders, tuberculosis, etc. **(Novakov et al. 2005[189])**. It gets mixed with air, water, and soil. Thus black carbon enters into the food chain and human body. One of the main sources of air pollution in India is due to fugitive emission of black carbon along with various other harmful gases from the coal mines released due to various mining activities **(CMRI 1998[53])**.

Fly ash is one of the main content released during these activities, which is very difficult to handle. According to a study, the ash collection from coal-fired thermal power plants ranges from 70% to 80% of the total ash present in coal **(Sahu et al. 2009[241], Senapati 2011[249])**. This ash if properly retained can be used in various construction activities, cement and tile manufacturing, brick kilns. High levels of suspended particulate matter (SPM) primarily like black carbon increase respiratory diseases such as chronic bronchitis and asthma cases while emissions of gases like Methane, which is 23 times more potent Green House Gas than carbon dioxide is released during mining **(Crosson 2008[69])**, leads to health issues and global warming. The pollution here leads to dangerous health issues and also leads to lung cancer, low visibility, tuberculosis, detrition of air quality, and equipment, leading to an increase in the maintenance cost around the mines **(Gautam et al. 2018[98])**.

1.2.2 Impact of BC on Climate

In the present era, black carbon is the major contributor to environmental and climatic changes after carbon dioxide (CO_2) leading to global warming and melting of glaciers, thus it is a matter of great concern. Black carbon is the major absorber of solar heat radiation in the atmosphere, BC leads to the heating of the Earth's atmosphere as it results into the reduction in incoming short-wave solar radiation at the Earth's surface **(Horvath 1993[112], Jacobson 2001[122])**, and thus leading to the change in the

temperature of the troposphere, which affects the microphysical properties of the cloud's and thus affecting the rainfall mechanisms **(Menon et al. 2002[162])**. Black carbon particles are transported to distant places by air and water to deposit on the ice on these glaciers and snow-covered mountain peaks, as it absorbs the heat radiations from the sun which causes the melting of ice **(Xu et al. 2009[292])**. Black carbon deposits on snow or ice this leads to an increase in the absorption of heat in the Arctic, poles, and also mountain ranges like the Himalayan glaciers which may be contributing factors for the rapid melting of ice in these regions **(Gautam et al. 2018[98])**. Black carbon also has severe impacts on weather patterns, rainfall, cloud's lifetime, and cloud formations **(Tao et al. 2012[276])**.

Although the anthropogenic causes of black carbon emission are globally distributed, it is mostly concentrated at the tropics as the solar radiation is greatest at the tropics **(Ramanathan and Carmichael 2008[223])**. Due to air and water, it is transported to distant places which in turn gets mixed with other aerosols leading to the formation of brown clouds extending up to 3 to 5 km **(Hsu et al. 2003[114], Ramanathan et al. 2007[224])**. Due to the high absorption of heat, it reduces the surface albedo (the ability to reflect sunlight) and capacity to form widespread brown clouds, which after carbon dioxide is the second-highest patron for global heating leading to melting of ice in the Artic **(Ramanathan and Carmichael 2008[223])**. The warming impact of BC per unit of mass, on climate is 460-1500 times stronger than CO_2. BC directly absorbs incoming and scattered solar radiation leading to atmospheric warming. BC is also leading to the melting of the Himalayan glaciers at a rapid rate **(Ramanathan and Carmichael 2008[223])**. Darker the soot the more is its heat-absorbing tendency. The soot from biofuels and fossil fuels are darker than that from biomass combustion. About 25-30% of the world's total BC emission comes from India and China **(Mohajan 2014[168])**. It also warms the air, affects the precipitation, rainfall pattern, cloud formation, and disturbs the ecosystem.

1.2.3 Black Carbon Variation with Particulate Matter Level in the Current Scenario

As per a study by the Government of India for environmental health **(Centre for Environmental Health, Government of India, 2017[54]),** the annual standard level of $PM_{2.5}$ set up by WHO is of 10 $\mu g / m^3$, but in over 70% of the Indian subcontinent, the level is beyond this limit. The annual standard level of $PM_{2.5}$ set up by CPCB is 40 $\mu g / m^3$, but the results obtained by satellite data and ground monitoring sources reveal that in the Indian subcontinent the range is 10-100 $\mu g / m^3$ **(Dey et al. 2012[81], Pant et al. 2015[196])**, the concentration in the northern states is comparatively higher than the southern states **(Pant et al. 2016[195], Tiwari et al. 2015[277])**. The dominating cities in tier 1 among these were the metro cities namely, New Delhi, Mumbai, Hyderabad, and Kolkata having the highest $PM_{2.5}$ levels as 40-81 $\mu g / m^3$. The average diurnal $PM_{2.5}$ level over Delhi was found to be at 300 $\mu g / m^3$ which was higher than the permissible standard limit of 60 $\mu g / m^3$ **(Das et al. 2015[71], Pant et al. 2015[196], Centre for Environmental Health, Government of India, 2017[54]).** Among the tier 2 were the cities of Raipur, Kanpur, Agra, and the rural locations of the IGP regions near the coal mines **(Massey et al. 2013[159], Pant et al. 2016[195], Begam et al. 2017[32], Rana et al. 2019[225])**.

A high concentration of pollutants and comparatively inferior technology, along with low wind speeds and low mixing heights has caused low atmospheric dilution rate leading to low dispersion of pollutants and higher pollutant over the IGP region than the rest of India **(Guttikunda et al. 2014[103], Guttikunda and Jawahar 2014[104], Attri 2008[26])**, further enhancing the pollutants levels in the IGP region as shown in Figure 1.6. The Indian Space Research Organisation (ISRO) has taken an initiative to study the behavior and effects of atmospheric pollution in the Indo-Gangetic Plains (IGP) **(Indian Space Research Organisation, 1994[121])**. The program is known as the ISRO-GBP (ISRO Geosphere-Biosphere Programme) it is having a check over 37 surface locations in India to measure the concentration of black carbon. The pollution from thermal and coal power plants, vehicles emission and from the burning of agricultural wastes have

polluted the IGP to a great extent as seen in Figure 1.6 **(Subbaraya et al. 2000[274], Guttikunda and Jawahar 2014[104])**, also the dust along with pollution traveling through winds from Thar desert, Iran, Afghanistan, and Pakistan leads to the imparts of aerosols during summers. In the ISRO-GBP study, measurements of the carbonaceous aerosol are to be calculated which is the main contributor to pollution in the IGP region **(Ram et al. 2010[222])**.

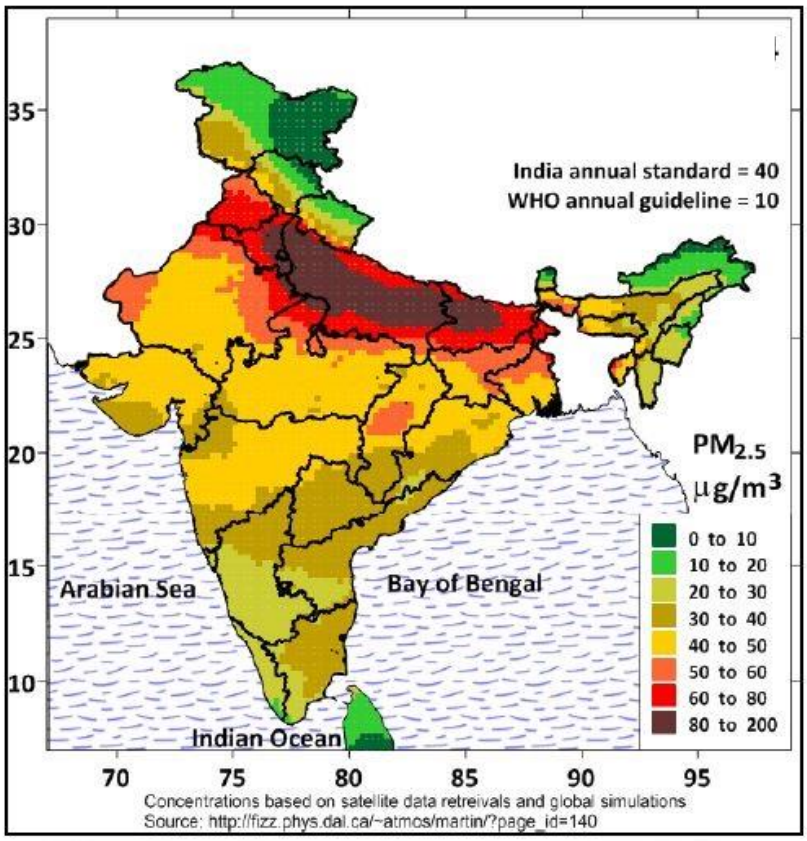


Figure 1.6 Annual average $PM_{2.5}$ concentrations $\mu g / m^3$ due to coal-fired thermal power plants in India **(Source: Goenka and Guttikunda 2014[100])**

In developing countries, mainly India and China the black carbon emission amounts to 25-35 percent to that of the entire world as observed by the **National Carbonaceous Aerosols Program (NCAP) report 2011[181] (Srivastava et al. 2014[267])**. India especially is the major sufferer as the Himalayas which have a great

influence on the climatic conditions of India has a high level of black carbon deposited on its snow peaks **(Menon et al. 2010[163])**. Thus due to the above discussed severe impact of black carbon on human life and environment both nationally and globally, the future study of black carbon is very important for framing national and international policies for prediction of the level of black carbon emission in the future. Coal mines region being one of the major sources for BC emission, we need to develop a new mathematical forecasting model having minimum prediction error using mathematical techniques like the statistical analysis, time series analysis, wavelets transform and soft computing methods like ANN and ANFIS, to discuss and find solutions to this problem at the national and global level.

1.3 Statistical Analysis

Statistical analysis is generally used in the field of modeling for understanding the nature of the distribution curve, depiction of meaningful interpretation, finding the missing values, and obtaining the forecasts of future values for the distribution curve of the data. The advantages of these are that it is simple to understand, time, and cost-efficient lays the basic foundation for modeling. Statistical analysis consists of Mean, Median, Mode, Standard deviation, Kurtosis, Skewness, coefficient of variation, Correlation, and Regression **(Parmar and Bhardwaj 2013[201, 203])**. Mathematically these are described as:

1.3.1 Mean: The average value of the sample is explained by mean, it is also termed as the Arithmetic mean. For a discrete series, $\{y_t\}, t = 1, 2, 3, \ldots$ it is obtained by summing all the values of the series and dividing it by the number of observations. It is calculated as,

$$Mean = \bar{Y} = \sum_{k=1}^{n} \frac{y_k}{n} \tag{1.1}$$

1.3.2 Median: The positional average or middle value of the discrete series $\{y_t\}, t = 1, 2, 3, \ldots$ is called the median. Its value depends on the number of observations present in the series. It is calculated as,

$$Median = \begin{cases} \dfrac{\left(\dfrac{r}{2}\right)^{th} term + \left(\dfrac{r}{2}+1\right)^{th} term}{2}, if \quad r \quad is \quad even \\ \left(\dfrac{r+1}{2}\right)^{th} term \qquad , if \quad r \quad is \quad odd \end{cases} \tag{1.2}$$

1.3.3 Mode: The value with the maximum frequency or maximum occurrence in the series is termed as the mode. A distribution is termed as unimodal if it has a single mode, bimodal for having two modes and trimodal if having three modes.

1.3.4 Standard Deviation: It is a measure of dispersion, which is widely used for measuring the variation and scattering among the data values. It is the square of the variance also termed as the root mean square deviation, for a discrete series $\{y_t\}, t = 1, 2, 3, \ldots$ it is usually calculated as,

$$\sigma = \sqrt{\frac{\sum_{k=1}^{n} \left(y_k - \bar{y}\right)^2}{n}} \tag{1.3}$$

1.3.5 Range: It is the most basic measure of dispersion which is used to find the scatteredness of the distribution. If *H* and *L* represents the highest and the lowest value among the observed values then we calculate range as,

$$Range = H - L \tag{1.4}$$

1.3.6 Coefficient of Variation: It is a measure of dispersion used to measure the variation of the observed values about the mean of the distribution; it can also be used for comparing the variation among the two series. It is calculated by the ratio of the standard deviation with the mean of the distribution. Mathematically this can be expressed as,

$$CV = \frac{\sigma}{\mu} \tag{1.5}$$

1.3.7 Kurtosis: It is used to determine the shape of the distribution curve, depending on the moments (μ_r) of the distribution. It determines the peakedness or flatness of the curve classifying it as platykurtic, mesokurtic, and leptokurtic relative to the normal distribution. For a discrete series, $\{x_t\}, \{y_t\}, t = 1, 2, 3, \ldots$ it is usually calculated as,

$$\gamma = \frac{\mu_4}{\mu_2^2} \tag{1.6}$$

Where, $$\mu_r = \frac{\sum_{k=1}^{n} (y_k - \bar{y})^r}{n} \tag{1.7}$$

1.3.8 Skewness: It a measure for lack of symmetry of the sample, lower the skewness indicates that the data values are away from skewness. The skewness can be positive, negative or undefined, if the skewness is positive then the data values are skewed towards the right, if it is negative then it indicates data is skewed towards the left and if the value is zero then it indicates that the data values are symmetrical and the data curve follows a normal distribution. It is calculated as,

$$Skewness = \frac{3(Mean - Median)}{\sigma} = \frac{(Mean - Mode)}{\sigma} \tag{1.8}$$

1.3.9 Correlation Analysis: It measures the degree to which two or more than two variables are in relation to each other along with the direction of relationship or association between them. The value of the coefficient of correlation lies between -1 to 1, closer the value to 1 or -1 indicates a high degree of correlation in the positive and negative direction respectively. The value zero of correlation indicates that there does not exist, any relationship between the two series. The sample correlation coefficient for two series $\{x_t : x \in X\}, \{y_t : y \in Y\}, t = 1, 2, 3, \ldots$ can be calculated as,

$$r = \frac{\sum (x_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sqrt{\sum (x_t - \bar{x}_t)^2} \sqrt{\sum (y_t - \bar{y}_t)^2}} \tag{1.9}$$

Here, $\bar{x}_t, \bar{y}_t$ denotes the mean of X and Y **(Parmar and Bhardwaj 2013[201, 203])**.

1.3.10 Regression Analysis: In regression, a mathematical relationship is established between two or more variables. This technique is used for predicting the value of the dependent variable (parameter) by using the required values of the independent variables also termed as predictors. After developing the equation of the regression line it can further be used for modeling and predicting the future or missing values. These are of two types:

(a) Simple Regression Analysis: It is used when two variables are in consideration, i.e. when we are dealing with a single independent variable. Regression equation of a line with dependent variable Y in terms of the independent variable X is defined as,

$$Y = b_{yx}X + C \qquad (1.10)$$

Where C is the constant of integration and b_{yx} = regression coefficient = $r \times \frac{\sigma_y}{\sigma_x}$ is called the slope of the regression line **(Parmar and Bhardwaj 2014[202]).**

(b) Multiple Regression Analysis: It is an extension of linear regression analysis which is used when two or more than two independent variables $\left(X_1, X_2, X_3, ..., X_n\right)$ are under consideration while determining the value of the dependent variable *Y*. It is widely used in financial inferences and econometrics. The equation of multiple regression can be expressed as,

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + ... + b_nX_n \qquad (1.11)$$

Where a is the constant which is termed as the y-intercept at zero time and $b_1, b_2, b_3, ..., b_n$ are the slope coefficients of the regression line.

Regression Analysis is a basic technique used for mathematical forecasting and curve-fitting by determining the relationship among the variables but the limitations associated with this is that it is based on the linearity condition and the results obtained are errorness when the data is scattered or consists of outliers.

1.4 Time Series Analysis

The evidence of health effects and various other consequences of pollutants on human health and environment are completely based on cross-sectional studies and time series analysis. Thus the major need of the present era is making time series analysis of the problem and developing accurate forecast with minimum forecasting errors. A time series is a sequential set of data points measured over successive time intervals arranged in proper chronological order. Mathematically described as, vectors $x(t), t = 0,1,2,3,...$ where *t*, is the time variable representing the time elapsed. A time series is also distinguished as discrete and continuous depending upon observations measured based on the time instance. If the time series is measured at equally spaced intervals of time such as hourly, daily, weekly, monthly, quarterly, yearly then it is termed as discrete time series. By combining the time series over a specific interval a continuous time series can be converted into a discrete-time series.

The main aim of time series modeling is to carefully collect data from a reliable source and observing these past values after choosing a proper model for a given time series which, best fits it. These time series models are further used to make predictions and forecasts based on past observations and generate future values. In this proper care is to be taken while fitting an appropriate model based on the time-series data available. Thus accurate forecasting for the time series depends on appropriate model fitting. Due to this various models are developed in literature from time to time which can best fit the data to get an efficient output and forecasting accuracy.

The most widely used stochastic models of time series are namely the Auto-Regressive (AR) model, Moving Average (MA) model, Autoregressive Moving Average (ARMA) model, Autoregressive Integrated Moving Average (ARIMA) model, and Box-Jenkins methodology. The basic assumptions for implementations of these models are the linearity condition and it follows a particular distribution like the normal distribution. The most popular of these is the ARIMA model along with the Box Jenkins methodology due to its independence to represent several variables with simplicity for a given time series.

These methods are largely used **(Su et al. 2011[273], Sachindra et al. 2013[239], Underwood 2013[284], Parmar and Bhardwaj 2014[202], Diodato et al. 2014[83])** for time series modeling. To study fluctuations and making good forecasts modeling of time series is very beneficial in the decision-making of climatic conditions and estimation of future data and extending it to real lifetime series **(Soltani et al. 2007[262])**. The autoregressive (AR) model, moving average (MA) model, and the combination of two Autoregressive integrated moving average (ARIMA) models have been used in various studies related with air pollutant modeling **(Ballester et al. 2002[29], Abdel-Aziz and Frey 2003[2], Chelani and Devotta 2006[60], Liang et al. 2009[146], Chattopadhyay and Chattopadhyay 2009[58], Portnov et al. 2009[213], Abish and Mohanakumar 2013[5])**. There are various important time series models **(Ratnadip and Agrawal 2013[228])** such as:

1.4.1 Autoregressive AR(*p*) Model: This is a basic time-series model in which we study the future behavior of the time series based on past observations. This model is based on the linear correlation among the data values at different lags. This process is similar to a multiple regression model, we calculate the predicted value by taking a linear combination of *p* past observations with a random error and a constant term. Mathematically it can be expressed as,

$$y_t = c + \sum_{k=1}^{p} \varphi_k y_{t-k} + \varepsilon_t \tag{1.12}$$

Here the integer *p* is called the order of the model, y_t is the predicted value and ε_t is the random error at time *t*, $\varphi_k \left(k = 1, 2, 3, ..., \mathrm{p}\right)$ are the parameters of the model and *c* is a constant.

1.4.2 Moving Average MA(*q*) Model: This is another basic model for time-series analysis in which the model uses the past errors as the prediction variables by taking a linear combination of *q* such past errors with a random error at time *t* and the mean of the series to obtain the future values. Mathematically written as,

$$y_t = \mu + \sum_{k=1}^{q} \theta_k \varepsilon_{t-k} + \varepsilon_t \qquad (1.13)$$

Here $\theta_k \left(k = 1, 2, 3, ..., \mathrm{q}\right)$ are the parameters of the model and q is called the order of the model. μ is the mean of the series, y_t the actual value at time t, ε_t random error and c is a constant.

1.4.3 Autoregressive Moving Average ARMA(*p*,*q*) Model: The Autoregressive (AR) and the Moving average (MA) are effectively combined to form the Autoregressive Moving average (ARMA), model. Mathematically it is represented as,

$$y_t = c + \varepsilon_t + \sum_{k=1}^{p} \varphi_k y_{t-k} + \sum_{k=1}^{q} \theta_k \varepsilon_{t-k} \qquad (1.14)$$

Symbols have their usual notations. ARMA models are employed using lag operators also known as the backshift operator, which is defined as, $Ly_t = y_{t-1}$. Lag polynomials used to represent the ARMA model are as follows,

$$\text{AR}(p) \text{ model: } \varepsilon_t = \varphi\left(L\right) y_t,$$

$$\text{MA}(q) \text{ model: } y_t = \theta(L)\varepsilon_t,$$

$$\text{ARMA}(p,q): \varphi\left(L\right) y_t = \theta(L)\varepsilon_t,$$

$$\text{Here } \varphi\left(L\right) = 1 - \sum_{k=1}^{p} \varphi_k L^k \text{ and } \theta(L) = 1 + \sum_{k=1}^{q} \theta_k L_k$$

1.4.4 Assumptions of Time series Analysis

1.4.4.1 Principle of Parsimony: It refers to the selection of a simpler time-series model by using the minimum number of parameters for representing a systematic structure of the time-series in place of complex ones. It leads to the selection of a model containing a

fewer number of parameters then the basic time-series models like the AR and MA model, thus simplifying the forecasting process.

1.4.4.2 Stationarity: If the statistical properties of a time series such as mean, median, mode, variance, and covariance do not change over time then such a time-series is called stationary otherwise it is called non-stationary.

1.4.4.3 Linearity: If every data point X_t of the time-series can be expressed as a linear combination of past observations, future values, or differences then such a time-series is said to be linear.

1.4.4.4 Differencing: First order differencing refers to the difference between the consecutive observations of a time-series while the second-order differencing refers to the difference between the values obtained by the first differencing approach. It helps to remove non–stationarity, trends, and fluctuation from the mean of a time series data and thus making it convenient to use the time series models.

However, even-though having numerous applications the limitation of the time-series analysis approach is to tackle nonlinear data because the time-series models use the linear correlation and the coefficient of correlation which are based on the linearity of the data. To remove this limitation of the time-series methods, various other models that are capable of nonlinearity were developed, compromising of Artificial neural network (ANN), Fuzzy, ANFIS, Wavelets **(Nayak et al. 2004[182], Partal and Kisi 2007[205], Yenigun and Ecer 2013[294])**.

1.5 Wavelet Analysis

Wavelet analysis was initially used in the field of seismology to study seismic waves as it contained the time domain which was lacked by the Fourier analysis. Thereafter, wavelets have been extensively used in various fields of research including the field of mathematical modeling to obtain good time-series forecast. Wavelet transforms are used to decompose a time-series signal into further component and generate information in

both frequency and time domain from them to get more accurate results. Wavelet transforms are capable of effectively handling of non-stationarity and non-linearity in the data which the Fourier transforms lacked. Wavelet transform is considered as an extension to Fourier transform to overcome the limitations of the Fourier transform. It was only possible with the help of wavelets that oil extractors could use it for identifying oil traces; signals processors were able to transmit signals through wires for telecommunication, and in future its application is expected to increase to weather prediction, diagnosing cancer and various other diseases, fingerprint tracking, computer animation, and imaging. Wavelets have proven to be more powerful in finding solutions for computational problems, pattern recognition, de-noising data, and data compression.

1.5.1 Fourier Transforms

In 1807 a French mathematician Jean Baptiste Joseph Fourier founded that sine and cosine functions could effectively be used to express most of the functions. The Fourier series of a function $f(t)$ is represented as,

$$f(t) = C_0 + C_1 e^{i\omega t} + C_{-1} e^{-i\omega t} + \ldots\ldots\ldots\ldots + C_n e^{in\omega t} + C_{-n} e^{-in\omega t} + \ldots = \sum_{n=-\infty}^{\infty} C_n e^{in\omega t} \qquad (1.15)$$

Where, $C_1 = \frac{(a_1 - ib_1)}{2}, C_{-1} = \frac{(a_1 + ib_1)}{2}$

Therefore, $C_1 e^{i\omega t} + C_{-1} e^{-i\omega t} = a_1 \cos \omega t + b_1 \sin \omega t \qquad (1.16)$

As $e^{i\omega t} = \cos \omega t + i \sin \omega t, e^{-i\omega t} = \cos \omega t - i \sin \omega t$

Thus equation (1.15) becomes,

$$f(t) = a_0 + a_1 \cos \omega t + b_1 \sin \omega t + \ldots\ldots\ldots + a_n \cos n\omega t + b_n \sin n\omega t + \ldots\ldots \qquad (1.17)$$

Where ω is the angular frequency and a_0, a_n, b_n are given as,

$$\left.\begin{aligned} a_0 &= \frac{1}{T}\int_0^T f(t)dt \\ a_n &= \frac{2}{T}\int_0^T f(t)\cos n\omega t\, dt, n = 1,2,3,..... \\ b_n &= \frac{2}{T}\int_0^T f(t)\sin n\omega t\, dt, n = 1,2,3,..... \end{aligned}\right\} \quad (1.18)$$

Here, $T = 2\pi$ is the period of the sine and the cosine series.

The Fourier transform is an extension to Fourier series it occurs when the period of the function is strengthened to infinity. It decomposes a function into its frequency components, for a function $f(t)$ its Fourier transform is given by,

$$\hat{f}(\omega) = F\{f(t)\} = \int_{-\infty}^{\infty} f(t)e^{-2i\pi\omega t}dt \quad ,\omega \in \Re \quad (1.19)$$

The equation given in (1.19) is also called the analysis part of the Fourier transform. The quantity of frequency present in the original function is obtained by the magnitude of its Fourier transform. Inverse Fourier transform of a function on the other side is used to generate the original function from the frequency domain. The Inverse Fourier transform of the function $\hat{f}(\omega)$ is given by,

$$f(t) = \overline{F}\left\{\hat{f}(\omega)\right\} = \int_{-\infty}^{\infty} \hat{f}(\omega)e^{2i\pi\omega t}d\omega \ ,t \in \Re \quad (1.20)$$

The equation given in (1.20) is known as the synthesis part of the Fourier transform. In signal processing the value of phase is obtained by $Arg\left(\hat{f}\right)$, the energy spectrum is given by $\left|\hat{f}\right|$ while $\hat{f}$ denotes the spectrum of f, Fourier and Inverse Fourier transform is continuous and linear. In a Fourier transform the non-periodic structures of a

signal could be resolved into a compound scale of frequencies. The structure of a periodic wave can be represented in a much closer and a rigorous manner with the help of Fourier transforms than the conventional idea **(Santoso et al. 2000[244])**.

But this expansion had a limitation that it contained only the frequency resolution and not the time resolution i.e. the frequencies could be predicted without the knowledge of the time at which it occurred. Although Fourier analysis being the best technique for dealing with stationary data the limitation with this is that it cannot be efficiently used for non-stationary data. In the later twentieth century, it was realized that a signal could be decomposed into further components to divide the information in the signal into frequency and time domains which were not pure sine waves **(Zhang et al. 2018[299])**.

Thus due to the limitations of Fourier transforms and the capability of Wavelet transforms to handle the information in a signal both in time and frequency domain, it was preferred over the Fourier transforms. Also while dealing with non-stationary time series data (signal), wavelet analysis proved to be a more effective tool than the Fourier transforms. A signal could also be decomposed into different components at various resolution levels with the help of wavelets which could further be used for obtaining more accurate time-series forecasts **(Abdulqader 2018[3])**. Wavelet transform turned out to be better than the Fourier transform as during computation they were less affected by small errors.

1.5.2 Wavelet Transforms

The origin of wavelet transform in literature is considered to be from the work of a Hungarian mathematician 'Alfred Haar' in 1910 **(Haar 1910[105])** which later on led to Haar wavelet. But the concept of wavelets was established in 1981 after the work of a geophysicist Jean Morlet when he was searching for a suitable method to find underground oil reserves through the seismic sound wave. As sound waves travel through different mediums with different speeds with the help of which geologist could predict the medium of the surface beneath. This process was represented as a mathematical

problem and then solved by engineers with the help of Fourier transforms **(Cohen and Kovacevic 1996[64])**. Morlet created components of the signals which he further shifted and compressed in time which he called the Morlet wavelets. He found that after shifting, compressing, and dilating a signal different wavelets could be created by a single wave which he termed as mother wavelet. Morlet worked along with Grossman to confirm that waves could be reconstructed from their decomposed wavelets components. They were the first to introduce the term 'wavelet' while working together in 1982 **(Morlet et al. 1982[175], Grossman and Morlet 1984[102])** and they found that compared to Fourier transform the wavelet transform performed better and were capable of efficient error handling.

Till the end of 1984, the only existing orthogonal wavelet was the Haar wavelet, later in 1985, Yves Meyer constructed another orthogonal wavelet which was named the Meyer wavelet and discovered the relation among the existing wavelets **(Meyer 1992[165], 1993[166], 1998[167])**. Later Stephane Mallat combined wavelets with computer applications and introduced a new concept of multi-resolution analysis in image processing **(Mallat 1989[155], 1999[154])**. Daubechies in 1987 introduced a new type of wavelet called Daubechies wavelets in series of papers **(Daubechies 1988[74, 76] 1992[73, 75], 1993[72])** which could be executed with the help of simple digital filtering ideas of short digital filters and were orthogonal like Meyer's wavelet, smooth and simple to program and use as in Haar wavelets. The theory of wavelets got a new shape with the help of Daubechies wavelets as a practical implementation with a minimum mathematical training that could be efficiently used for programming **(Daubechies 1990[77])**.

Wavelet is a small wave-like family of oscillatory functions consisting of finite energy for a period of time obtained due translations and dilations of mother wavelet, it is defined as **(Sahay and Srivastava 2014[240])**,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), a \neq 0, a, b \in R \qquad (1.21)$$

Where 'a' is called the scaling parameter and it denotes the degree of compression and 'b' the parameter for translation.

Wavelets are basically mathematical functions by means of which complex functions can be expressed in a simple way. They break down data into its frequency components and so each and every part the signal could be studied in a more precise manner **(Behera and Mehra 2013[34], Mei et al. 2020[161])**. The wavelet series expansion for a function $f(x)$ is expressed as,

$$f(x) = \sum_{i,j=-\infty}^{\infty} a_{ij}\psi_{ij}(x) \tag{1.22}$$

Where a_{ij} are the wavelet coefficients and $\psi_{ij}(x), i, j \in Z$ is obtained from the function $\psi \in L^2(R)$ by dilation and translation such that,

$$a_{ij} = \left[W_\psi f\right]\left(2^{-i}, j2^{-i}\right) \tag{1.23}$$

$$\psi_{ij}(x) = 2^{i/2}\psi\left(2^i x - j\right) \tag{1.24}$$

These parameters i, j are integers that play the role of the frequency and time are used to generate the basis by dilating and shifting the variations of the mother wavelet; hence it leads to the generation of the multi-channel multi-resolution approach.

The continuous wavelet transform of any function $f(t)$ is given by,

$$T_\psi f(m,n) = \frac{1}{\sqrt{m}} \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-n}{m}\right)dt = \left\langle f, \psi_{m,n} \right\rangle \tag{1.25}$$

Where m, n denotes the parameter for the scale of frequency and the spatial time and ψ^* is the complex conjugate of the function $\psi(t)$.

A wavelet transform decomposes a signal into various components which contain information about the signal at various scales of time and frequency. The information present in the signal does not change due to the wavelet transform. The frequency and time depiction of a signal is obtained with the help of a wavelet transform. It is based on the multi-resolution technique in which different resolutions can be examined at different frequencies.

A function $\psi \in L^1 \cap L^2$ is called the mother wavelet if it can generate the wavelet family after compression, dilation, and translation. It is a fast decaying oscillating wavelet of finite length. A mother wavelet can be used as a prototype to generate all the basis function for a wavelet family. It is similar to an orthogonal function and is orthogonal to a polynomial of degree less than or equal to *n*. It has $n+1$ moments equal to zero,

i.e. $$\int_R \psi(t) t^r dt = 0, \quad p = 0,1,2,...,n \qquad (1.26)$$

A decomposition of the original signal of the mother wavelet is performed by the wavelet transform into a weighted set of scaled wavelets functions. Various families of wavelets with distinct shapes are constructed using the mother wavelet which can be used as a basis function. A few of them are described below:

1.5.3 Some Important Types of Wavelets

1.5.3.1 Haar wavelet: It was introduced in 1909 by 'Alfred Haar' a Hungarian mathematician. Haar wavelet consists of a family of "rescaled square-shaped" functions, which was later known as db1, a specific kind of Daubechies wavelet. This wavelet consisted of a function in which a short term positive pulse was succeeded by a short term negative pulse. It was the first introduced wavelet which was orthogonal. Haar wavelet function $\psi(t)$ is defined as,

$$\psi(t) = \begin{cases} 1 & 0 \le t \le 1/2 \\ -1 & 1/2 \le t \le 1 \\ 0 & elsewhere \end{cases} \qquad (1.27)$$

Diagrammatically it is represented as,

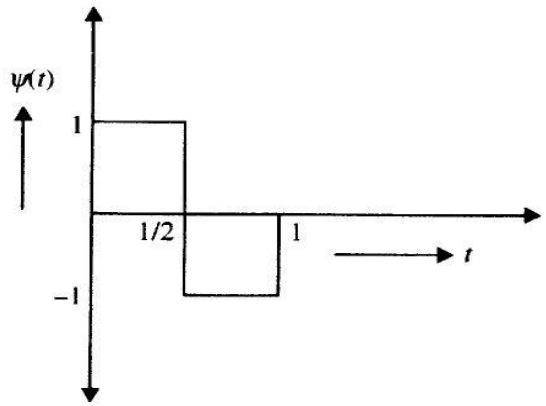


Figure 1.7 Haar wavelet function

While the Haar scaling function $\varphi(t)$ is given by,

$$\varphi(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & otherwise \end{cases} \qquad (1.28)$$

Diagrammatically it is represented as,

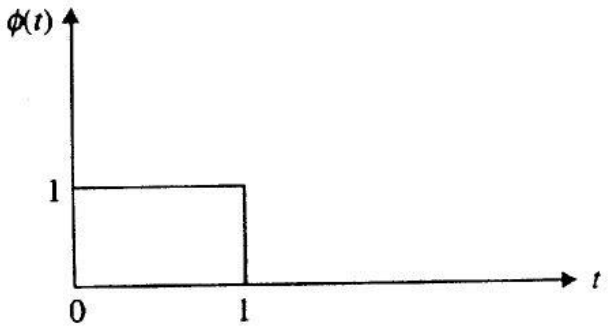


Figure 1.8 Haar scaling function

The Haar function on real line *R* for a pair of integers m, n is given by,

$$\psi_{m,n}(t) = 2^{m/2}\psi\left(2^m t - n\right), t \in R \tag{1.29}$$

These functions are pairwise orthogonal given as,

$$\int_R \psi_{m_1,n_1}(t)\psi_{m_2,n_2}(t)dt = \delta_{m_1,m_2}\delta_{n_1,n_2} \tag{1.30}$$

Where $\delta_{i,j}$ represents the Kronecker delta function while $\psi_{m,n}(t)$ is the corresponding wavelet function. The biggest limitation associated with the Haar wavelet is that it neither differentiable nor continuous.

1.5.3.2 Morlet wavelet: Morlet wavelet is also known as Gabor wavelet, it was introduced in the year 1984 by Jean Morlet when he was trying to generate components which were localized in space while studying seismic waves. It does not consist of any scaling function. These are also termed as, 'wavelets of constant shape' since the wavelets retain a constant shape even after the components were shifted, compressed or dilated. It is also described as a sine wave curbed with a Gaussian and obtained by the product of the Gaussian window with a complex exponential function. It has a close relation to natural vision, hearing, and used in non-stationary time-series signals such as transcription of music waves and various medicinal applications. The assumptions of Morlet wavelet are that the signals are stationary and sinusoidal. It is defined as,

$$\psi(t) = e^{-t^2/2\sigma^2}e^{2i\pi ft} \tag{1.31}$$

Where σ is the width of the Gaussian function given by $\sigma = \frac{n}{2\pi f}$.

Diagrammatically it is represented as,

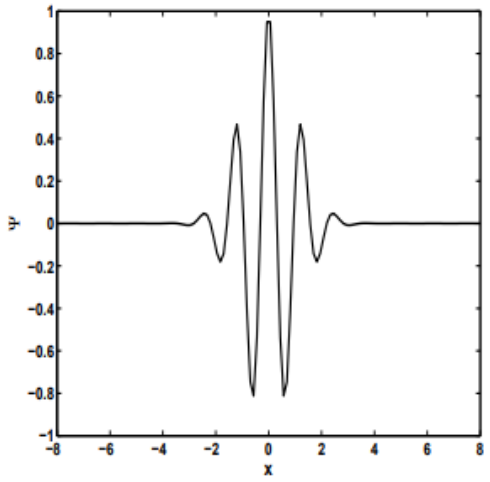


Figure 1.9 Morlet wavelet function

1.5.3.3 Meyer wavelet: Meyer discovered a new family of wavelet in 1985, which was easy to handle with wavelet transform and preserved the orthogonality property. Its wavelet function is given by,

$$\psi(\omega)=\begin{cases}0 & ,if\ |\omega|\notin\left[\frac{2\pi}{3},\frac{8\pi}{3}\right]\\ (2\pi)^{-1/2}\,e^{i\omega/2}\sin\left(\frac{\pi}{2}v\left(\frac{3}{2\pi}|\omega|-1\right)\right) & ,if\ \frac{2\pi}{3}\le|\omega|\le\frac{4\pi}{3}\\ (2\pi)^{-1/2}\,e^{i\omega/2}\cos\left(\frac{\pi}{2}v\left(\frac{3}{4\pi}|\omega|-1\right)\right) & if\ \frac{4\pi}{3}\le|\omega|\le\frac{8\pi}{3}\end{cases} \tag{1.32}$$

While it's scaling function is given by,

$$\varphi(\omega)=\begin{cases}0 & ,if\ |\omega|>\frac{4\pi}{3}\\ (2\pi)^{-1/2} & ,if\ |\omega|\le\frac{2\pi}{3}\\ (2\pi)^{-1/2}\cos\left(\frac{\pi}{2}v\left(\frac{3}{2\pi}|\omega|-1\right)\right) & if\ \frac{2\pi}{3}\le|\omega|\le\frac{4\pi}{3}\end{cases} \tag{1.33}$$

This wavelet is infinitely differentiable and it ensures orthogonal analysis.

Diagrammatically the scaling and wavelet functions are represented as,

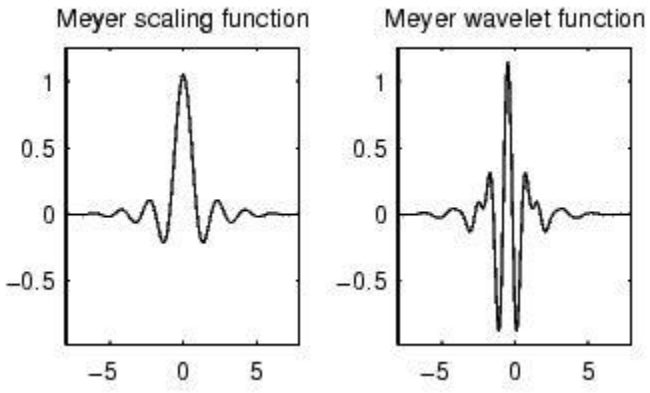


Figure 1.10 Meyer scaling and wavelet function

1.5.3.4 Daubechies wavelet: Ingrid Daubechies invented a family of discrete wavelets that were orthogonal compactly supported and orthonormal named as Daubechies wavelet. They are represented as *dbN* (*db1*, *db2*… *db10*) where *N* is the order of the wavelet. The most basic and shortest wavelet is *db1* which is also termed as the Haar wavelet. The best time resolution wavelet among the family of Daubechies wavelet is *db4*. The Daubechies wavelets have *N* vanishing moments and are mostly asymmetrical. Orthogonal multi-resolution is generated by the scaling function of each wavelet. For given support, it has the maximum number of vanishing moments. There is a different scaling function for each member of this family of wavelet which further leads to orthogonal multi-resolution analysis. It is simple to use and most effective in dealing with spike, self-similarity problems, and random time-series. The construction of graphs for Daubechies wavelet is based on a cascade algorithm, for distinct types; there are different wavelet and scaling functions.

Diagrammatically the wavelet functions are represented as,

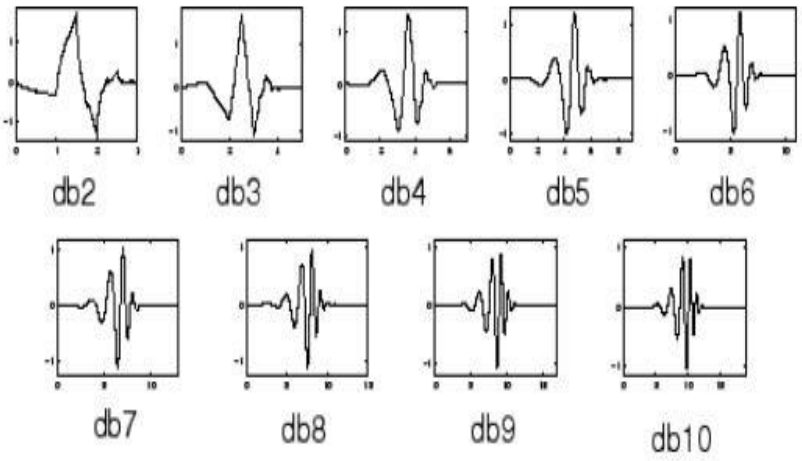


Figure 1.11 Daubechies wavelet function

1.5.3.5 Mexican Hat wavelet: It is also known as the Ricker wavelet and is derived from the second-order derivative of a function similar to the Gaussian function. The Gaussian function is given as,

$$\psi(x) = e^{-x^2}\left(1 - x^2\right) \tag{1.34}$$

Its wavelet function is,

$$\psi(t) = \left(\frac{2}{\sqrt{3}}\pi^{-1/4}\right) * e^{-t^2/2} * \left(1 - t^2\right) \tag{1.35}$$

The wavelet function is proportional to the derivative of the second-order of the Gaussian function and n moments of the nth order derivative of the Gaussian function vanishes. The scaling function does not exist for this wavelet. From a real-valued mother wavelet function, it can be easily created by adjusting the signals of the analytic function of the signal.

Diagrammatically it is represented as,

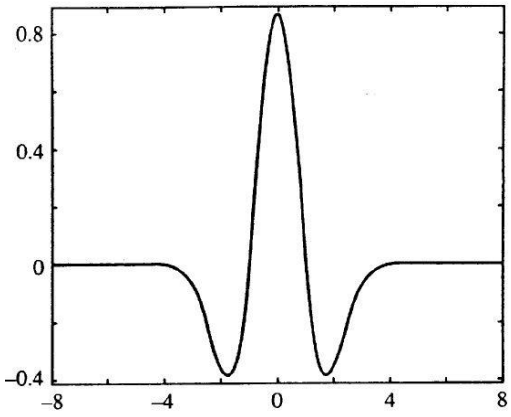


Figure 1.12 Mexican Hat wavelet function

1.5.3.6 Biorthogonal wavelet: This wavelet was constructed by Cohen in 1992; it consists of symmetric and linear phase property which is useful in image and signal reconstruction. It is popular in sub-band filtering in which precise reconstruction and symmetry are discordant if the similar FIR is used for decomposition and reconstruction. Thus different filters are used for decomposition and reconstruction in this wavelet. It can be used in constructing symmetric wavelet functions. These are not fundamentally orthogonal while the related wavelet transform is invertible. Compactly supported symmetric wavelets are derived from the biorthogonal basis. For any function, f biorthogonal basis can be formed with the help of two basis functions $\psi_{i,j}$ and $\tilde{\psi}_{i,j}$ derived from two mother wavelets such that,

$$f(x) = \sum_{i \in Z} \sum_{j \in Z} \tilde{b}_{ij} \psi_{ij} \tag{1.36}$$

$$\tilde{b}_{ij} = \int_{-\infty}^{\infty} f(t) \tilde{\psi}_{i,j}(t) dt \tag{1.37}$$

A single template Biorthogonal wavelet $\psi(t)$ generates the Biorthogonal basis, if $\left\{\tilde{\psi}_{i,j} : i, j \in Z\right\}$ denotes the Biorthogonal basis of the signal then,

$$\left\langle \tilde{\psi}_{i,j}, \psi_{k,l} \right\rangle = \delta_{i-k}\delta_{j-l}, \quad i, j, k, l \in Z \tag{1.38}$$

Whereas the wavelets coefficients $\left\{a_i(j)\right\}_{i,j\in Z}$ are obtained by,

$$a_i(j) = \left\langle f, \tilde{\psi}_{i,j} \right\rangle_{L^2}, i, j \in Z \tag{1.39}$$

In these two filters are used successively as shown in Figure 1.13, the one shown on the left-hand side is used for the decomposition part while the other on the right-hand side is used for the reconstruction part. For different types, different wavelet and scaling functions are used. Diagrammatically these are represented as,

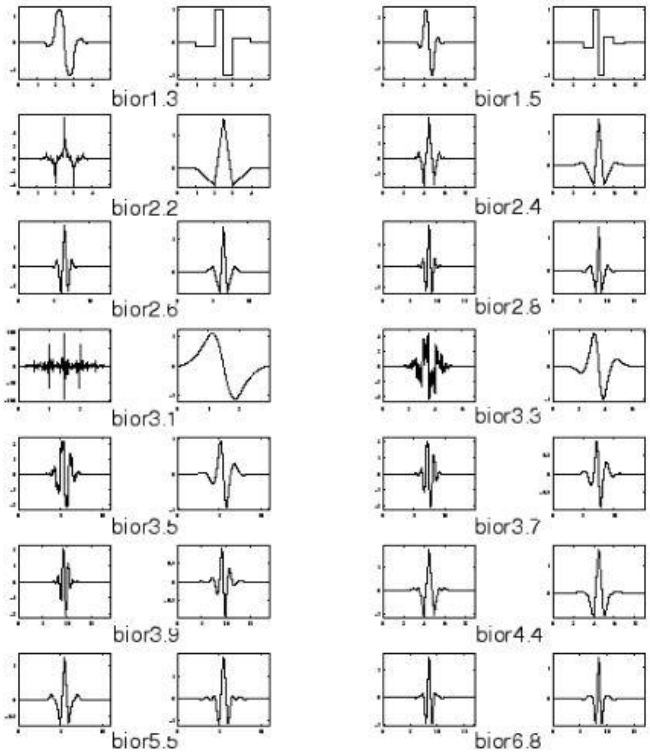


Figure 1.13 Biorthogonal wavelet function

1.5.3.7 Coiflet wavelet: This wavelet was developed by Ingrid Daubechies and Ronald Coifman; these are a type of discrete wavelet which is usually expressed as coif*N*, where *N* is the order of the wavelet. The scaling function has 2*N*-1 moments while the wavelet

function has 2*N* moments which vanish, it has a support of length 6*N*-1. These wavelets are more symmetrical than the Daubechies wavelets. For distinct types of Coiflet wavelet, there are different wavelet and scaling functions. Diagrammatically the wavelet functions are represented as,

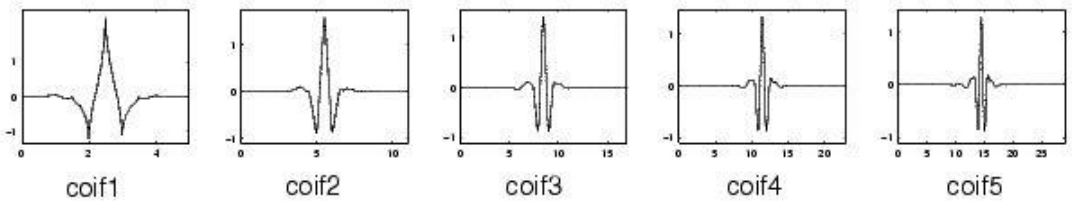


Figure 1.14 Coiflet wavelet function

1.5.3.8 Symlets wavelet: These wavelets were developed by Daubechies as a reform to *db* family of wavelet which is close to symmetrical wavelets as opposite to Daubechies wavelets. They are denoted as *symN*, where *N* is the order. They have the least size of support for a fixed number of moments equal to zero which is the same as half the size of the support. For distinct types of Symlets wavelet, there are different wavelet and scaling functions. Diagrammatically the wavelet functions are represented as,

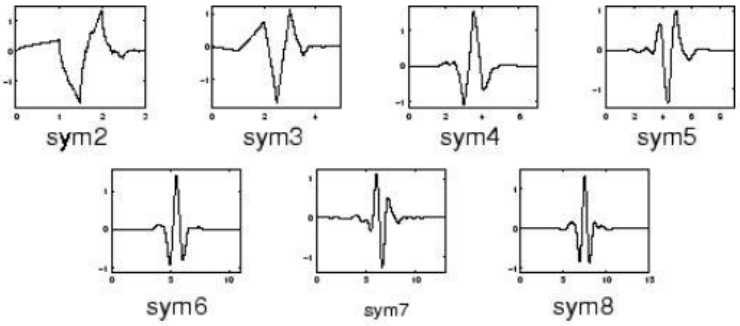


Figure 1.15 Symlets wavelet function

1.5.3.9 Orthogonal wavelet: The concept of Orthogonal wavelet was given in 1985 by Meyer, a wavelet whose wavelet transforms is orthogonal is called an orthogonal wavelet. A wavelet $\psi \in L^2$ is orthogonal if it is defined as,

$$\left\langle \psi_{i,j}(t), \psi_{p,q}(t) \right\rangle = \int_{-\infty}^{\infty} \psi_{i,j}(t) \psi_{p,q}(t)\, dt = \begin{cases} 0 & i \neq p,\ j \neq q \\ 1 & i = p,\ j = q \end{cases} \tag{1.40}$$

This means the inner product of a mother wavelet with itself is unity while with some other wavelet obtained by dilation and translation is zero. The first known orthogonal wavelet was the Haar wavelet and the second was Meyer wavelet. This wavelet uses the same filter for the decomposition and reconstruction process. The scaling function φ and the wavelet function ψ are related to each other as,

$$\frac{1}{2}\psi\left(\frac{x}{2}\right) = \sum_{k \in Z} a_k \varphi(x-k) \tag{1.41}$$

Orthogonal wavelets are obtained by orthogonal basis; a basis is called orthogonal if the scaling function φ satisfies,

$$\left\langle \varphi_l(t), \varphi_m(t) \right\rangle = 0 \tag{1.42}$$

1.5.4 Scaling and Translation of a Wavelet: A signal is represented by its orthonormal basis function which is generated with the help of translations and dilations (scaling) of a wavelet function. Translations and dilations of a signal help to extract the information present in the signal both in the frequency and time domain. In scaling the frequency is changed while in translation the central position of the wavelet changes. Diagrammatically this represented as,

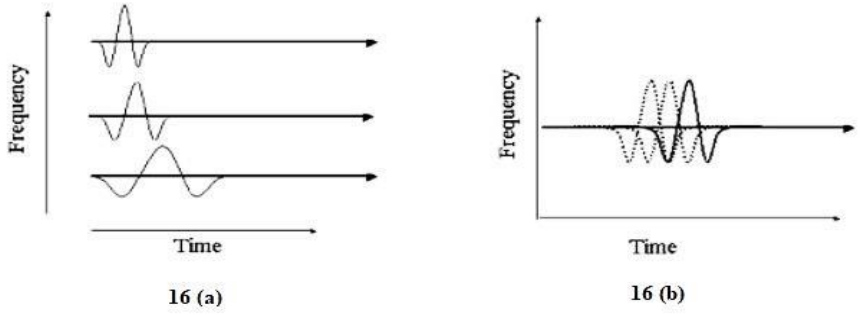


Figure 1.16 (a) Scaling of a wavelet, **1.16 (b)** Translation of a wavelet

For a wavelet defined as, $\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), a \neq 0, a, b \in R$

Here a and b denotes the scaling and the translation parameter. For the function $\phi(t)$ translation for different values of t are $\phi(t-1), \phi(t+1), \phi(t-2), ...$ or in general, $\phi(t-r)$ while scaling for the function $\phi(t)$ are given as, $\phi(2t), \phi(3t), \phi(4t), ...$ or in general $\phi(kt)$.

1.6 Soft Computing

Soft computing also known as Machine language is a branch of Computational intelligence that uses genetic algorithms, neural network, probability theory, and Fuzzy logic. It is used for obtaining solutions to complex computational problems that otherwise are very difficult to solve. It helps to solve problems as human mind considering approximations, uncertainty, imprecision, and partial truth. There are various types of soft computing techniques used in the field of Mathematical forecasting but the most popular among these are the Artificial Neural Network (ANN) and the Adaptive Neuro-Fuzzy Inference System (ANFIS).

1.6.1 Artificial Neural Network (ANN)

ANN is also termed as a neural network or nonlinear statistical data modeling methods. ANNs are used as a prominent alternative mathematical tool in the field of mathematical modeling and forecasting of time series data with nonlinear properties, due to its powerful pattern recognition and classification abilities. It establishes a functional relationship between the data values even though it is difficult to describe it by any other method for time series having a sufficient number of data values. The most important advantage of ANN is that it can learn from observing the data sets. The ANN-based tools are ideal and cost-effective methods for obtaining solutions using distributions and computing functions. To arrive at the solution, ANN uses data samples than the whole

data sets which make it cost and time-efficient. These are relatively simple mathematical models, which are used to improve the present data analysis technologies. ANNs have been used to solve a wide variety of problems in the fields of science, industry, and business.

Initially biologically inspired but now it has been applied in various fields especially in the case of time series forecasting and predictions as it can learn and generalize results from past experience **(Hung et al. 2009[119])**. Firstly, **Hu 1964[115]** introduced the realization of the ANN approach in weather prediction. It is used as a substitute for the modeling of a time series, it is based on human brain functioning and has gained great popularity during the last few years. Salient features of ANNs are that it is data-driven and self-adaptive, they can be used for nonlinear time series modeling unlike the stochastic time series models, ANN makes no previous assumptions pertaining to the data distribution and it can model greatly nonlinear relationships **(Nagendra and Khare 2006[177])** and finally, ANNs are universal function approximators. To give high accuracy in prediction for the data ANNs use parallel processing of the information. It can further be applied in situations when the data is incomplete, erroneous, or fuzzy. Correlation between the objective values and the input values is identified and studied by ANN. It is based upon the network between neurons and nodes, which are interconnected with each other. ANN uses an algorithm at different neurons to improve their weights **(Hsu et al. 1995[113], Jeong et al. 2012[127], Moosavi et al. 2014[174])**.

ANNs are three-layered namely input, hidden and the output layer which consists of several neurons. ANNs are used for a variety of tasks like classification, function approximation, data processing, filtering, clustering, compression, robotics, regulations, decision making, etc. After receiving information from a source the neuron gathers it and then conducts a nonlinear operation on the obtained value giving the final outputs **(Bas and Boyaci 2007[30])**. The main advantage of ANN is that without any mathematical model it gets trained from the input and output data values by recognizing the patterns in the series **(Yeon et al. 2009[295])**. Numerous intelligent systems have been developed

using ANN to solve a variety of problems like pattern recognition, prediction, control system, associative memory, and optimization inspired by biological neural networks. Depending on the algorithm of ANN there are various types of ANN, the most popular among these are discussed below:

1.6.1.1 Feedforward Neural Network (FFNN) – It is the most widely used ANN, in this the information travels in the forward direction before reaching the output node without flowing in loops. There are three layers in the architecture of FFNN viz. the input, hidden, and output layer. The number of neurons in the first layer depends on the number of inputs **(Farajzadeh et al. 2014[92])**. Its main aim is to obtain output values close to the targeted values by reducing the estimation error **(Nourani et al. 2013[186])**. In this, there is no backpropagation while the hidden layer may or may not exist. These are commonly used in case of noisy data as in speech recognition and computer vision; depending on the number of layers they are further classified as Single layer and Multilayer perceptron's **(Heghedus et al. 2019[110])**. A perceptron sums the inputs and consists of an activation function, weights, and summation processor. These are used in systems dealing with the reorganization of speech and objects **(Nourani et al. 2017[184])**.

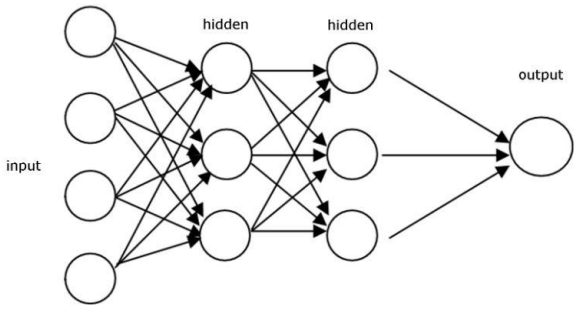


Figure 1.17 Feed-forward neural network architecture

1.6.1.2 Recurrent Neural Network (RNN) – In the form of loops, it involves a recurrence of operations**,** in order to obtain the forecast, the information of the output layer is fed back into the input layer. Each neuron stores the information of the previous time lag, making it act as a memory cell **(Luongvinh and Kwon 2005[149])**. Neurons

work on the front propagation method; in case of the wrong prediction, we use the error correction method with backpropagation. There are two types of such ANNs the first one with infinite impulse while the other with finite impulse. The former cannot be unrolled and is a directed cyclic graph while the later can be replaced by a strictly feed-forward neural network and can be unrolled as it is a directed acyclic graph. For obtaining forecasts of nonlinear time series, the recurrent network is preferably used as in language generation, text, and speech reorganization **(Williams and Zipser 1989[290], Heghedus et al. 2019[110])**.

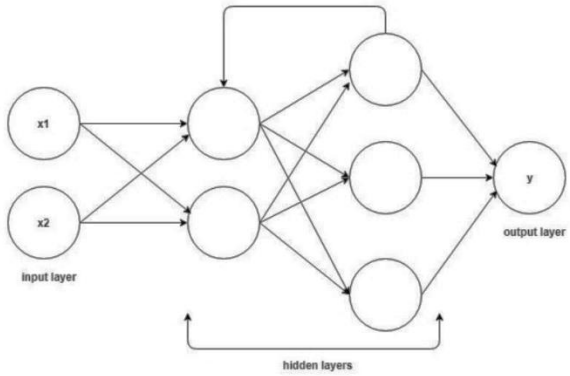


Figure 1.18 Recurrent neural network architecture

1.6.1.3 Modular Neural Network – In this, the neural network diminishes to a single layer and possibly to components that can be managed more easily. While performing the computational process to obtain the output, the network functions independently without interacting with each other, thus enabling the process to perform faster as the networks are not dependent on each other. These independent neural networks act as modules that perform the computations of the inputs. All the modules are connected with each other, which are further consigned with a specific function. It reduces the complexity by breaking down the larger computations into smaller components which leads to a reduction in the computation speed **(Happel and Murre 1994[108])**.

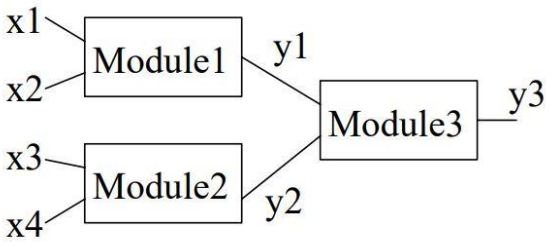


Figure 1.19 Modular neural network architecture

1.6.1.4 Convolutional Neural Network (CNN) – Being similar to a feed-forward network it consists of one or more layers of convolution interconnected with each other working on the principle of variations in the MLPs. CNN shows very effective results as it contains a lesser number of parameters and is much deeper since the convolutional operation on the input is performed in the layers **(Yao et al. 2017[293])**. This reduces the time and the data required for obtaining the output. The information from the visual field is processed by the neurons in the convolutional layer. These are mainly used in the application of artificial intelligence functioning as the human brain and object recognition. They are also applied in image classification, signal processing, paraphrase detection, and semantic parsing **(Lawrence et al. 1997[145])**.

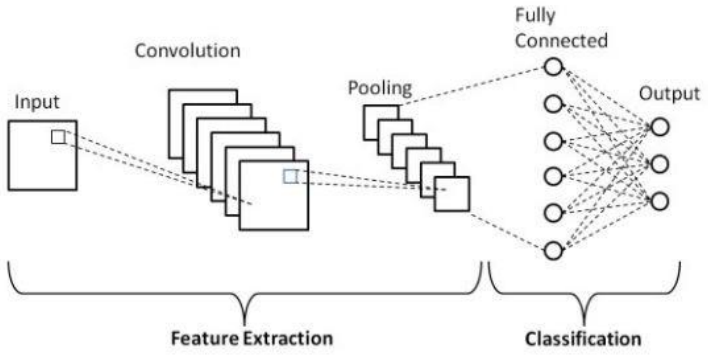


Figure 1.20 Convolutional neural network architecture

1.6.1.5 Radial Basis Function Neural Network (RBNN) – It is used as an alternative to the sigmoid transfer function. In this, the output obtained by uniting the signals with the

radial basis function is further taken into consideration in the output layer. Mostly a Gaussian function is taken as an RBF, in this Euclidean distance from the center of each neuron to a point is considered, for each neuron the radius may differ which are determined during training along with the weights applied to the RBF outputs, coordinates of the center and the number of nodes in the hidden layer. As in the MLPs, the benefit of RBF is that it avoids the local minima as the parameters in the output layer are formed by linearly mapping of the hidden layers **(Adamowski and Sun 2010[8])**. The selection for the number of Basis functions must not be more than the total data points in the input dataset. When each data point is associated with an RBF it leads to the Gaussian and Support vector machine process. It is used in a power restoration system which has increased in complexity and size with time and it is needed to immediately restore the power **(Aksoy and Dahamsheh 2009[16], Nourani and Farboudfam 2019[188])**.

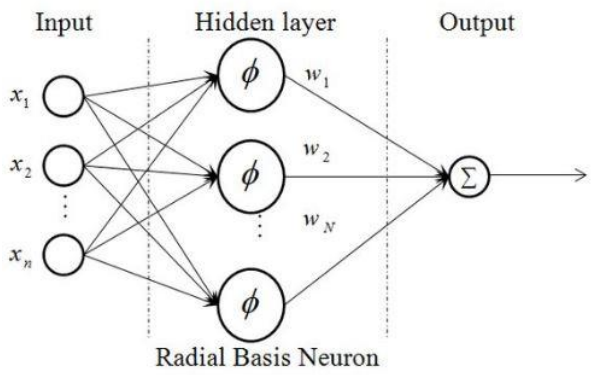


Figure 1.21 Radial basis function neural network architecture

1.6.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANN along with the fuzzy interface system (FIS) forms the ANFIS which can be used for estimation and forecasting of the time series, the results obtained by ANFIS are more reliable and accurate **(Jang 1993[123])**. It maps input characteristics functions to input membership functions which are related to rules and it is further related to a set of output characteristics. Again the output characteristics are mapped to output membership functions which results in a unique output or a decision related to the output **(Jang**

1993[123]). The important components of fuzzy systems are fuzzifier, fuzzy database, and defuzzifier. The fuzzy database consists of two major parts the fuzzy rule base and the inference structure **(Karmakar and Mujumdar 2006[130], Akrami et al. 2013[15], Sahay and Srivastava 2014[240])**.

Like ANN, Fuzzy Interface System (FIS) is also a data-driven, soft computing model, which represents fuzzy if-then rules which are crucial to developing by a crisp parameter model **(Jeong et al. 2012[127], Parmar and Bhardwaj 2015[199])**. In the Mamdani technique, the whole data is divided into two parts, training and testing and the model results compared with the actual data **(Hung et al. 2009[119], Chen and Chang 2010[61]**). The use of fuzziness makes the model more consistent and results towards accuracy **(Aksoy et al. 2004[17], Toprak et al. 2004[279], Toprak et al. 2009[278], Hung et al. 2009[119], Chen and Chang 2010[61])**.

1.6.2.1 ANFIS Architecture

If *x* and *y* represents the two inputs variables and *f* represents the output as shown in Figure 1.22 **(Parmar and Bhardwaj 2015[199], Dalkilic and Hashimi 2020[70])**.

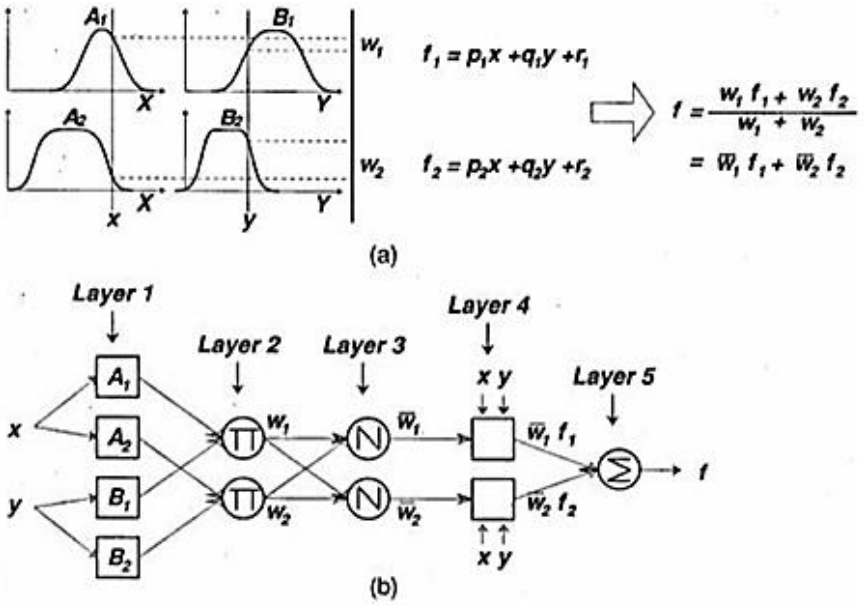


Figure 1.22 Neuro-Fuzzy architecture

Rule 1 If x is A_1 and y is B_1 then $f_1 = p_1 x + q_1 y + r_1$

Rule 2 If x is A_2 and y is B_2 then $f_2 = p_2 x + q_2 y + r_2$

Layer 1 Every node *i* in this layer is a square node with a node function, $O_i^1 = \mu A_i(x)$.

Where x represents input at node *i*, and A_i is the linguistic label related to this node function. So, O_i^1 is the membership function of A_i, which expresses the degree to which the given *x* satisfies the A_i. Mostly, it is to supposed $\mu A_i(x)$ to be bell-shaped and minimum value to 0, maximum value to be 1, such as the Generalized Bell function

$$\mu A_i(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{b_i}} \tag{1.43}$$

Or Gaussian function

$$\mu A_i(x) = e^{\left[-\left(\frac{x-c_i}{a_i}\right)^2\right]} \tag{1.44}$$

Where $\{a_i, b_i, c_i\}$ the set of parameters, the bell-shaped function is behaving as the values of parameters set changes. It exhibits various forms of membership functions on linguistic labels A_i. Parameters in this layer are referred to as premise parameters.

Layer 2 Every node in this layer is a circle node. It multiplies the input signals and sends the product out. For instance

$$O_i^2 = w_i = \mu A_i(x) \times \mu B_i(x), i = 1, 2, 3, \ldots \tag{1.45}$$

Each node output represents the firing strength of a rule.

Layer 3 Every node in this layer is a circle node labeled N. The i^{th} node measures the ratio of the i^{th} rule's firing strength to the sum of all rule's firing strengths

$$O_i^3 = w_i = \frac{w_i}{w_1 + w_2}, i = 1, 2, 3, \ldots \tag{1.46}$$

For convenience, outputs of this layer will be called normalized firing strength.

Layer 4 Every node *i* in this layer is a square node with a node function

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i \left(p_i x + q_i y + r_i \right), i = 1, 2, 3, ... \quad (1.47)$$

Where the output of layer 3 is represented by $\bar{w}_i$ and $\left(p_i, q_i, r_i \right)$ is the set of parameters.

Layer 5 Circular node is the single node in this layer labeled, which calculates the entire The output is expressed as the summation of all incoming signals i.e.

$$O_i^5 = \sum \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, i = 1, 2, 3, ... \quad (1.48)$$

This is working equivalent to type 3 FIS (Fuzzy Interface System).

1.7 Model Fitting Parameters: Mathematical forecasting is a branch of Mathematical modeling, its importance lies in generating good time-series forecasts having minimum forecasting errors for a given time-series data. There are various mathematical models used for generating these forecasts, but the success of Mathematical forecasting lies in the selection of the model which is best suited for a given time-series model yielding optimal results. For this purpose, the following model fitting parameters are used;

1.7.1 Error Measures: To check and compare the best-fitted model, the most widely used error measures are Root mean square error (RMSE), Mean square error (MSE), Mean absolute percentage error (MAPE) and Mean absolute error (MAE) respectively **(Makkhan et al. 2020[152])**. The least the values of these forecasting errors the better will be the model fitted, the accurate will be the forecast. These error measures are given by the following results **(Kumar et al. 2015[141])**:

1.7.1.1 Root Mean Square Error (RMSE): It is the coefficient of error representing the standard deviation of the difference of actual values of the data from the values predicted by the time-series model also termed as the residual values, it is used to determine the amount of spreadness of the values from the line of best fit for a model and to determine the accuracy of the forecasted values. Its value is non-negative; the closer the value to zero indicates better the results. Mathematically it is given by,

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(x_i - \bar{x}_i \right)^2} \quad (1.49)$$

Where n denotes the time period and e_i denotes the error of forecasting.

1.7.1.2 Mean Square Error (MSE): It is also known as Mean square deviation, it is calculated as the mean of the square of the error. It is the square of the RMSE; its value is non-negative, closer the value to zero indicates better the results. It is given by,

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}_i\right)^2 \qquad (1.50)$$

1.7.1.3 Mean Absolute Percentage Error (MAPE): It is also known as Mean absolute percentage deviation; it is calculated as the mean ratio of the error with the observed value multiplied by 100. Mathematically it is given by,

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - \bar{x}_i}{x_i}\right| * 100 \qquad (1.51)$$

1.7.1.4 Mean Absolute Error (MAE): It is the mean of the absolute error values. Its value is non-negative; the closer the value to zero indicates better the results. It is calculated as,

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \bar{x}_i\right| \qquad (1.52)$$

Where x_i and $\bar{x}_i$ represents the observed and the forecasted values of the time series, n represents the number of observations.

1.7.1.5 Relative Error: It is the ratio of absolute error to the absolute actual value.

It is calculated as,

$$\in = \frac{\left|Actual - Predicted\right|}{\left|Actual\right|} \qquad (1.53)$$

When multiplied by 100 it leads to the percentage error.

After developing the model, a comparison between the actual values and the forecasted values is made, to check the accuracy of the model fitted to the data.

1.7.2 R-squared and Stationary R-squared Values: These are measures for the goodness of fit for a time series model; they are used as a coefficient of determination of a model. It explains the effect of the amount of variation of one variable on the other and to what extent the movements of observation in the series are correlated. It aims at finding the line of best fit for the set of data points (observations). The value of R-square ranges from 0 to 1 while that of stationary R-squared ranges from -∞ to 1, higher values indicate that the model considered is better than the baseline model.

1.7.3 Autocorrelation and Partial Autocorrelation Function (ACF and PACF)

Autocorrelation refers to the amount of correlation of a time series with its past or lagged values while ACF plots are used to find the correlation between the present data points with those at different lags and thus it determines the correlation of a time series with itself **(Abish and Mohanakumar 2013[5])**. Generally, either the AR or MA approach is used while developing an ARIMA model but in a few cases both approaches are also used. In most cases, if at lag 1 the autocorrelation is positive then the AR approach is used but if it is negative then the MA approach is used.

Autocorrelation at lag *k* is given as,

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{1.54}$$

Where $\gamma_k = Cov(y_t, y_{t-k})$ and γ_0 is the auto-covariance at lag 0 also called the unconditional variance.

On the other hand, partial autocorrelation between two data values refers to the relationship between y_t and y_{t-k} after controlling and removing all the linear dependence between $y_1, y_2, y_3, ..., y_{t-k+1}$, i.e. it finds a correlation of the residuals with the lagged values. These residuals might contain some hidden information, thus with the help of PACF, we can obtain a good correlation. The residuals can be modeled by the next lag and this lag can be set as a feature while modeling. If we consider too many

features it can lead to multi-collinearity in the time series. Thus with PACF, we hold the relevant features only. It is denoted by $\phi_{k,k}$.

$$\phi_{k,k} = \frac{Cov\left(y_i, y_{i-k} \left| y_{i-1}, y_{i-2}, ..., y_{i-k+1}\right.\right)}{\sqrt{Var\left(y_i \left| y_{i-1}, y_{i-2}, ..., y_{i-k+1}\right.\right) * Var\left(y_{i-k} \left| y_{i-1}, y_{i-2}, ..., y_{i-k+1}\right.\right)}} \tag{1.55}$$

Both ACF and PACF are used for proper model estimation; they describe the relationship between the observations of a time series **(Parmar and Bhardwaj 2015[200], Pandey et al. 2019[194])**.

1.7.4 AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) – AIC was formed in 1973 by Hirotsugu Akaike whereas BIC was developed by Gideon E. Schwarz in 1978. Both of these are penalized-likelihood criteria that are used for estimating model parameters while fitting or choosing the best model. AIC is also termed as a measure of the goodness of fit; it is nearly the distance between the fitted likelihood function for the model and true likelihood function of the data. It aims to find the best prediction model with the least error. Smaller the value of AIC indicates that the model under consideration is best fitted which is closer to reality and will have the least forecasting error than the others. While the approximation of the posterior probability of a function is called the BIC under certain Bayesian setups, lower the BIC means the model under consideration to be a true model. Compared to the AIC penalty for additional parameters is greater than in BIC, free parameters are penalized more strongly in BIC than AIC. The difference between AIC and BIC is that AIC is very useful for making asymptotically equivalent to cross-validate while BIC is good for consistent estimation.

They are usually calculated by the formula $-2\log L + kp$, where L represents the log-likelihood function which is again used as a measure of goodness of fit for a model, the numbers of parameters are represented by *p*, *k* is mostly taken as 2 while determining AIC and $logp$ for BIC **(Shetty et al. 2018[252])**.

1.8 Motivation for Selection of the Problem

Air is the most basic need for life to exist on this planet Earth, we can live without food for a few weeks and water for a few days but we cannot live without air for even a few seconds. In the present age of modernization due to the growth of Industrialization, urbanization, automobiles, coal-fired thermal power plants, and various factories the quality of air has degraded year by year **(Japar et al. 1986[126], Liousse et al. 1996[147], Tzanis and Varotsos 2008[283])**. Many cities of India are among the list of severely polluted cities of the world with highest air quality Index, the situation is getting worst day by day even after some remedial measures and the masses living in these cities are severely affected with no light of hope to fight against this devil of air pollution **(Massey et al. 2013[159], Pant et al. 2016[195], Begam et al. 2017[32], Rana et al. 2019[225])**.

Thousands of peoples living in the country have to lose their lives each year and lakhs of them have become prone to several deadly diseases related to cardiovascular, respiratory, cancer, eye sights, and birth **(Penner et al. 1993[210], Varotsos 2005[285])**. Black carbon is one of the major components of air pollution, having a severe impact on the health of the nation and the climate **(Horvath 1993[112], Jacobson 2001[122])**. The major sources of black carbon emission and air pollution in India are coal mining regions situated in the eastern part of the country leading to high emission of air pollutants as can be seen from Figure 1.23 **(Goenka and Guttikunda 2014[100])**, where an extensive amount of open cast mining is performed **(Coal Controller Organization of India, Ministry Of Coal, 2019[63])**. So we have selected the three major coal mines of India situated in this region of high emission of air pollutant, $PM_{2.5}$ namely, Raniganj, Jharia, and Bokaro coal mines of India.

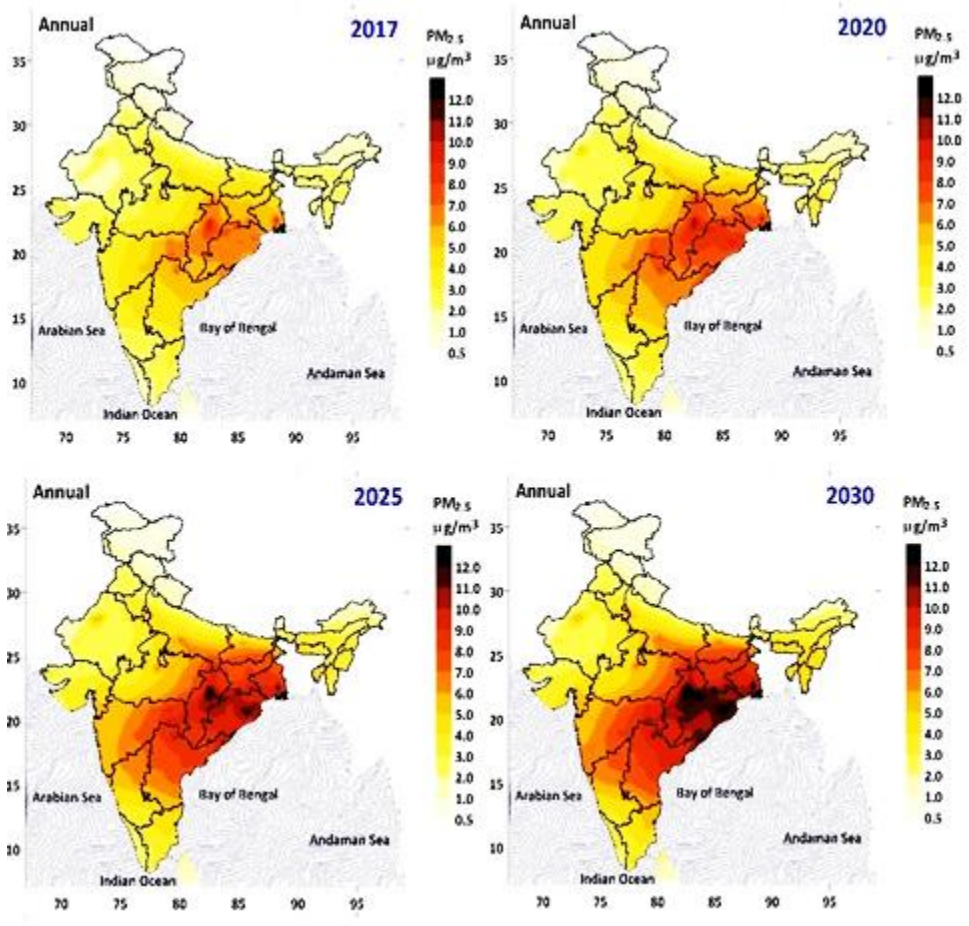


Figure 1.23 $PM_{2.5}$ emission concentrations in India **(Source: Goenka and Guttikunda 2014[100])**

1.9 Sample Sites

The sample sites for our current study include the coal mine regions of India namely, 'Raniganj' which is the oldest coal mine, 'Jharia' which has the largest coal reserves, and 'Bokaro' the steel city. Raniganj (23° 40' N 87° 05' E) in West Bengal, Jharia (23° 50' N 86° 33' E) and Bokaro (23° 46' N 85° 55' E) in Jharkhand are the major coal mines as shown in Figure 1.24 **(Srivastava 1988[268], Chandra et al. 2000[57], Bhattacharjee 2017[39])**. Raniganj coalfield was established in India in 1774 and was the first operational coalfield in the country, after this there was gradually very slow progress in the growth of the coalfields and far after in 1856 coal mines were established in Jharia and Bokaro in

Jharkhand **(Srivastava and Mitra 1995[269])**. With an area of 443.50 km^2 (171.24 square miles) Raniganj coalfield is the largest coalfield in West Bengal located in the eastern part of the Damodar valley coalfields, spread across Indian states of West Bengal and Jharkhand **(Srivastava and Mitra 1995[269])**. The Raniganj basin is elongated semi-elliptical lying between the river Ajoy and Damodar **(Ghosh 2002[99])**. This coalfield is encountered with southerly dipping beds and east-west trends.

While Jharia coalfield is located in the Dhanbad district of Jharkhand and covers an area of about 280 km^2 (110 square miles) and has coal reserves of about 19.4 billion tons **(Indian Bureau of Mines 2019[120])**. It has been actively accompanying coal mining activities from the past century and more than 85.55 km^2 area of it is used for CBM activities. The Jharia coalfield is sickle-shaped and the axis of its basin is approximately trending east-west and dipping in the west **(Saraf et al. 1995[246])**. The Bokaro coalfield is one of the most important coal basins in the Damodar Valley in the state of Jharkhand, well known for metallurgical coal of substantial thickness and superior quality. Structurally, this coalfield is in the shape of an elongated synclinal graben, preceding and cropping out successively towards the west and east **(Singh et al. 2013[257], Mahato et al. 2014[150])**. It is divided into two major segments, 'East Bokaro' (208 km^2) and the 'West Bokaro' (259 km^2) by Lugu hill (975 m), these include deposits of the Karharbari, Barakar, and Raniganj **(Singh et al. 2013[257])**. These three coalfields are on the eastern extremity of the Damodar valley basin and this area is known for its high rank in producing coal of superior quality which can be used as a clean fuel **(Paul et al. 2018[209])**.

Air pollution from these coal mines has affected the neighboring states especially the IGP region. The air pollution situation is worst in Jharia since underground fires have continuously been burning in the coalfield from nearly the past 100 years affecting lakhs of people. It is one of the world's oldest and widely spread mine fire. Bokaro Steel Plant along with various other industries depending on the coal from these coal mines are also contributing to air pollution from these regions. These coalfields cause a high emission of

anthropogenic CO_2 with enhanced coalbed methane and black carbon **(Srivastava 1988[268], Chandra et al. 2000[57], Bhattacharjee 2017[39])**.

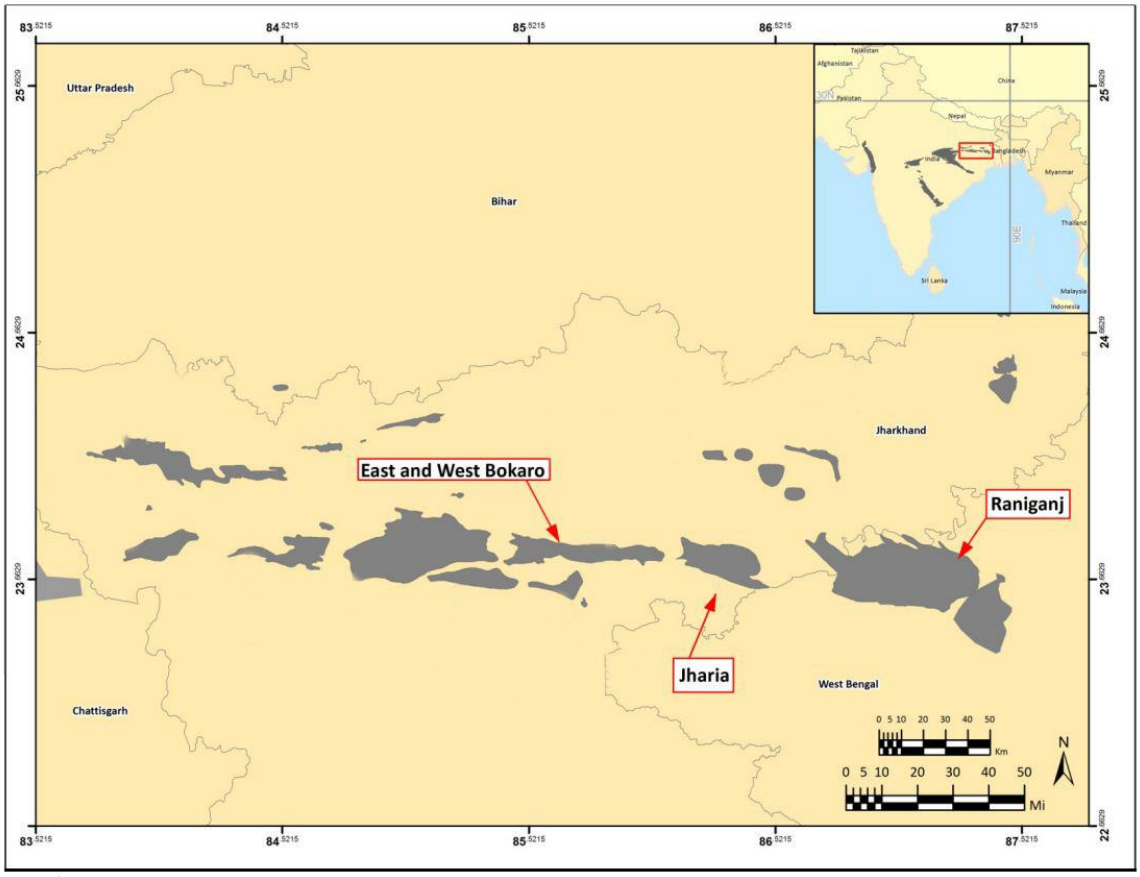


Figure 1.24 Coal mines of Bokaro, Jharia, and Raniganj

A continuous time-series data of black carbon concentration for the past 38 years, 05 months from Jan 1980 to May 2018 has been obtained. The data is collected by NASA (National Aeronautics and Space Administration) via Giovanni website (http://nasa.gov/) and processing of the data is done by CSIR National Physical Laboratory, New Delhi as shown in Figure 1.25. The observations and results obtained are expressed in aerosol volume (magnitude) as $e^{-11}\ kgm^{-3}$ units.

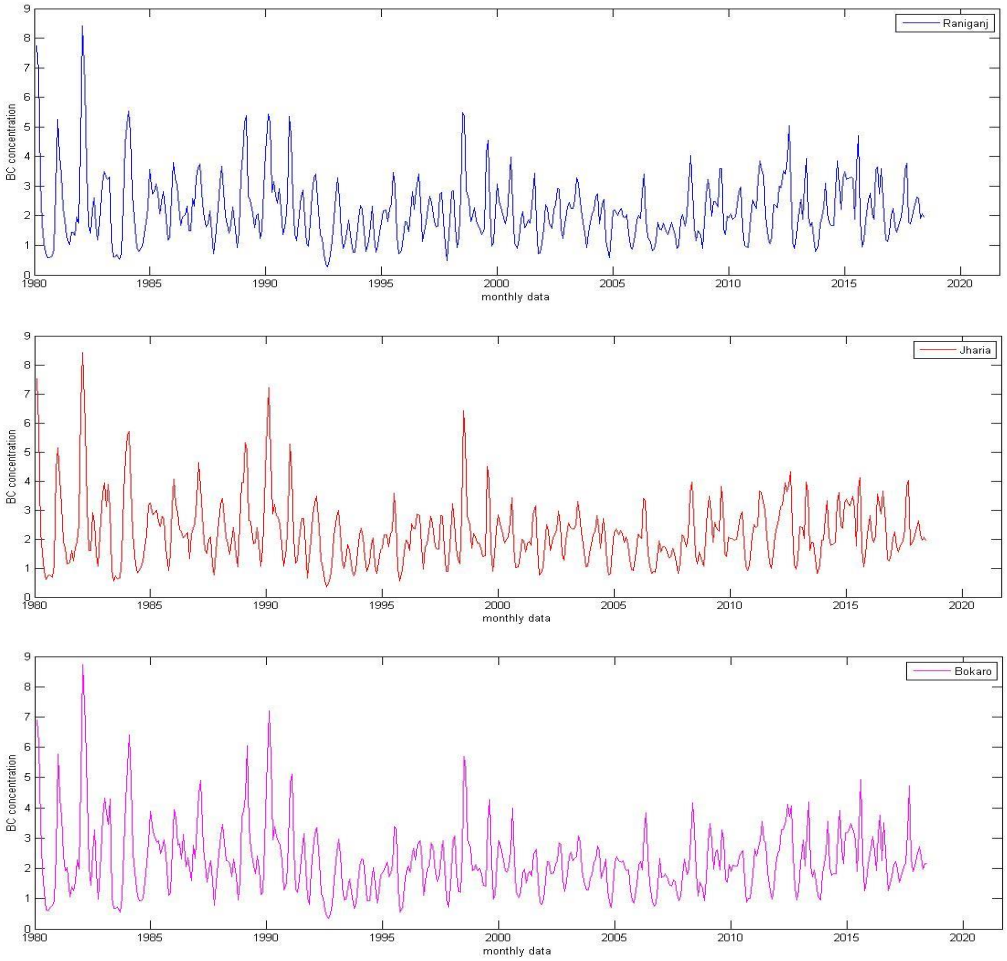


Figure 1.25 Time series of Raniganj, Jharia, and Bokaro

Chapter 2 Statistical Analysis of Black Carbon

Statistical analysis of a time series data is an important initial step taken to get an insight into the data for understanding and studying its past behavior for making future predictions of the time series. It involves the study of various statistical parameters also termed as descriptive statistics, such as the central tendency measures primarily like the mean, mode, and median. Mean is the most important central tendency measure which is representative of the entire series. It calculates the average of the entire time series, giving an idea of the value for the parameter to be studied. It is calculated by dividing the sum of all the time-series observations by the total number of data values. It is an important statistical tool as it is used in almost all studies dealing with statistics and is further used for obtaining various other statistical measures. Mode represents the highest occurring parameter value of the time series which has the maximum frequency of occurrence in the time series.

Median on the other hand gives an idea about the middle value of the time series data about which the data is scattered or can be divided into two equal parts. If mean, mode and median coincide then the data is said to follow normal or a symmetrical distribution. To measure the variability or spreadness of the data, dispersion measures like standard deviation, variance, skewness, range, and coefficient of variation are used. If the data is not symmetrical then it is called skewed data which may be skewed either to the left or the right depending on the relationship between the mean, median, and mode. For determining the shape of the distribution curve kurtosis is used. It measures the amount of heaviness of the tail of a distribution from that of the normal distribution. Depending on the number of extreme values and heaviness at the tail of the distribution, kurtosis is further categorized as mesokurtic, platykurtic, and leptokurtic. The trend analysis technique is used to determine the future movement of our parameter and behavior of the time series; it examines the past values and predicts the future. Based on trend analysis we determine whether the time series follows persistent, anti-persistent, or Brownian motion.

In this chapter, we study the nature of the trend and study the behavior of the time-series over the three sample sites using statistical measures like central tendency measures, dispersion, skewness, regression analysis, Hurst exponent, fractal dimension, predictability index, trend percent, and correlation analysis. To analyze our parameter i.e. BC concentration, we have obtained the values of these statistical parameters along with the trend, regression, and correlation analysis for the sample sites. Correlation analysis is further used to determine the amount of similarity or dependency among the time series for the three sample sites **(Pasanen and Holmstrom 2017[207])**. To obtain a clearer picture of the situation we have used three different coefficients of correlation along with Cross-correlation function (CCF) and time-series plots. The results have been represented by the Figures (2.1-2.5) and Tables (2.1-2.4) along with a bar chart depiction of the observed results.

2.1 Literature Review

A case study of the Qiantang River, China was done by **Su et al. 2011[273]**; time-series measurement for 13 variables at 41 monitoring locations during 1996 to 2004 was used to describe trends and temporary trends using non-parametric statistical tests in terms of four parameters to understand the reasons of water pollution. **Yenigun and Ecer 2013[294]** carried out the trends analysis in the Euphrates basin located in southeastern Anatolia, Turkey for observing climatic and hydrologic parameters using statistical analysis and overlay mapping technique, the results were beneficial in forecasting and decision-making for the management of water resources. A comparison of statistics was made from 1985 to 2009 and correlated with the period 1904–1984 by **Underwood 2013[284]** using daily effective temperatures (DET) by examining the wind speed and temperature.

Parmar and Bhardwaj 2013[203] carried out an analysis of water quality parameters at Harike Lake, the junction of Sutlej and Beas rivers of Punjab (India) using statistical methods, based on the parameter and monthly variation of water quality index compared with the WHO standards, results concluded that the water of the lake was very

badly polluted and was unfit for industrial and household use. At Hathnikund Bridge of river Yamuna, **Parmar and Bhardwaj 2015**[200] carried out a trend analysis of water quality of Yamuna using predictability index, fractal, lag, and Hurst exponent. It was observed that water quality degrades at Agra, Mazawali, Nizamuddin and was good at Juhikha. Inorganic ions of $PM_{2.5}$ soluble in water in Wuhan were studied by **Huang et al. 2016**[117] using correlation and seasonal variations to find the average mass of concentration for the ions and determine the major components of $PM_{2.5}$. Correlation analysis between black carbon and carbon monoxide was conducted by **Mok et al. 2017**[172]; results represented a high degree of correlation during summer and spring season. Considering the long term data of black carbon concentration from 2008-2015 over the coastal region of Preila, Lithuania, **Davuliene et al. 2019**[78] observed that there exists a positive trend, however, this behavior was not consistent and a seasonal pattern was observed in BC concentration.

Fiero et al. 2019[94] presented a review paper based on clinical trials in lung cancer data of US Food and Drug Administration from Jan 2008 to Dec 2017 considering the statistical analysis of patient-reported outcomes. For spatially fixed transcriptomic research, **Sun et al. 2020**[275] used statistical analysis method SPARK for recognizing genes demonstrating spatial expression shapes and showed that it was more powerful than the existing methods. **Ruan et al. 2020**[236] in a novel study dealing with the recent pandemic used statistical graphical tools like a bar graph, pie graph, and Box whisker graph to represent the results indicating the reasons for causes of death based on symptoms, age factor due to novel coronavirus by collecting a data of patients in China along with their recovery period. Considering a long term time series data for the concentration of black carbon from 1980 to 2018, **Makkhan et al. 2020**[152] performed a statistical and correlation analysis. Results predicted that the sample sites of the coal mines considered for the study represented a high degree of correlation and followed a similar pattern of emission in future predictions.

2.2 Statistical Analysis

For studying the behavior and relation of the time-series for black carbon concentration over our sample sites, statistical analysis parameters like central tendency measures, dispersion measures, skewness measures, correlation analysis, regression analysis, and trend analysis have been used.

2.2.1 Correlation Analysis: It is used to measure the degree of dependency or association of two or more variables with themselves. The value -1 or 1 signifies a high degree of correlation among the variables in the positive or negative direction respectively. The value is zero if there is no correlation between the associated variables. The commonly used important types of the coefficient of correlation are **(Singh et al. 2004[256], Parmar and Bhardwaj 2013[201, 203], Lu et al. 2014[148])**:

2.2.1.1 Karl Pearson's Coefficient of Correlation: It is the most widely used coefficient of correlation it is given by the formula,

$$\rho_{xy} = r = \frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{\left(E\left(X^2\right) - \left(E(X)\right)^2\right)\left(E\left(Y^2\right) - \left(E(Y)\right)^2\right)}} \qquad (2.1)$$

Here, σ_x, σ_y are the standard deviation of X and Y while $E(X), E(Y)$ and $E(XY)$ represents the expected values.

2.2.1.2 Spearmen Rank Correlation: It is used to measure the strength and direction of two ranked variables. It is represented by r_s and given as

$$r_s = 1 - \frac{6\sum_i a_i^2}{n\left(n^2 - 1\right)} \qquad (2.2)$$

Here, $a_i = x_i - y_i$, is the differences between the ranks of two samples and *n* denotes the number of samples.

2.2.1.3 Kendall's tau Correlation: It is a non-parametric hypothesis test used to measure the ordinal association between two variables *X* and *Y*. If the value is zero then

the variables are statistically independent of each other and if the value is 1 then they are highly correlated. It calculated as,

$$\tau = \frac{C - D}{C + D} \qquad (2.3)$$

Where, *C* and *D* represent the number of concordant and discordant pairs, which are used to measure the association between the pairs of ordinal observations.

Concordant pairs: If the members of the observations are ranked in the same direction.

Discordant pairs: If the members of the observations are ranked in opposite direction.

2.2.2 Regression Analysis: It is commonly used for predicting the value of the dependent variable (parameter) in a time series by considering the values of the independent variable and taking other variables as fixed after developing the equation of the regression line. This modeling technique can also be used to find out the missing values and analyze the present variables of the sample. It is used to understand the variation in the dependent variable due to frequent changes in the independent variables. Regression equation of the line is written as,

$$x = a + b_{xy} y \qquad (2.4)$$

Where a is the constant of integration;

Regression coefficient = $b_{xy} = \frac{\sigma_x}{\sigma_y} \times r$

Here σ_x, σ_y are the standard deviations for the two series and r is the coefficient of correlation calculated as given in equation (2.1),

2.2.3 Trend Analysis: It is a statistical technique used to forecast the future movement of a time-series on basis of the past observations. It is very helpful in predicting the future trends of time-series dealing with pollutants, groundwater, rainfall precipitation, stock exchange, and growth cycles **(Makkhan et al. 2020[153])**. There are various techniques used for trend analysis but the mathematical techniques have proved to be very efficient in studying trends. The mathematical trend parameters used in this study are:

2.2.3.1 Hurst Exponent (*H*): It is a statistical measure used in time series analysis for predictability; it determines the relative direction for a time series. Its value varies from 0 to 1; we determine the trend pattern of a time series based on these values of *H*. The trend in the reinforcing series is said to exhibit a 'persistent behavior' if *H >0.5* with positive autocorrelation in the time series, while if *H<0.5* then the trend follows an 'anti-persistent behavior' also called 'mean reversion' along with negative autocorrelation values, in this the trend in the series is subsequent such as, a rise followed by a fall and then again by a rise. If *H= 0.5* then the series is random also termed as a Brownian time series, and it is hard to predict as in this case, it is very hard to establish any relationship between the future and the past values of the series. Time series analysis can lead to more accurate forecasts if the value of *H* obtained is closer to 0 or 1, this indicates the larger strength of the trend in the series and thus the model leads to a more accurate forecast of the time series. Mathematically it is calculated as,

$$H = 0.5 * \left|1 - b_{yx}\right| \tag{2.5}$$

The power law of decay $p(k) = Ck^{-\alpha}$ can also be used to derive the value of the Hurst exponent *H*. Here $p(k)$ represents the autocorrelation function at *lag k* related to α by the relation,

$$H = 1 - \frac{\alpha}{2} \tag{2.6}$$

2.2.3.2 Fractal Dimension (*D*): It is a statistical measure that is used to find the presence of fractals in a time series as we zoom in further to higher scales. If the value of *D<1.5* then the time series behavior is called persistent, if *D>1.5* then anti persistent behavior exists. A time series is said to follow a Brownian motion if *D=1.5* which leads to unpredictability. It is calculated as,

$$D = 2 - H \tag{2.7}$$

2.2.3.3 Predictability Index (*PI*): The ability to obtain accurate forecasts with least forecasting error is termed as predictability index. The value of *PI* lies between 0 and 1. Higher the *PI* the least is the gap between the predicted and the actual series. Thus it

enables us to determine the future attainment of a time series; based on the value of *D* it is calculated as,

$$PI = 2 * |D - 1.5| \quad (2.8)$$

2.2.3.4 Trend Percent: It is the amount of variation in the value of the parameter over a comprehensive period; it is used to determine the future movement of the time series considering the past observations. Trend can both be downward and upward, it is calculated as,

$$Trend\% = \frac{a * N * 100}{\bar{x}} \quad (2.9)$$

Where *a* represents the slope of the regression line, *N* is the number of data values (monthly) and $\bar{x}$ denotes the average of the time series.

2.3 Results and Discussion

2.3.1 Statistical Analysis using Central Tendency, Dispersion, and Skewness

Raniganj (23° 40' N 87° 05' E)

The measure of the central tendency viz. mean, mode and the median value of BC concentration are at 2.192876, 1.447137, and 1.983193. Since all these values are close to each other and near to 2, thus the data distribution curve nearly follows close to a normal distribution and is nearly symmetrical. The value of standard deviation and skewness are at 1.113655 and 1.507609 indicating that the data points are distributed close to each other along with the mean. Since the value of the mean is greater than the median therefore the distribution curve is moderately skewed towards the right. The curve is leptokurtic as indicated by the value of kurtosis at 4.192889 in Table 2.1.

Jharia (23° 50' N 86° 33' E)

As shown in Table 2.1 the mean, mode, and median value of BC concentration stand at 2.232444, 1.548387and 2.059908 all near to 2 as shown in the table indicating that the data nearly exhibit a normal distribution and the distribution curve is close to

symmetrical. The value of standard deviation is 1.130772 and skewness as 1.560077 indicating moderately positive skewness of the data towards the right with data values near to each other. The shape of the distribution curve is leptokurtic as the value of kurtosis is 4.245786.

Bokaro (23° 46' N 85° 55' E)

The value of the mean, mode, and median value of BC concentration are 2.269679, 1.842105, and 2.078582 all near to 2 as shown in Table 2.1 indicates that the data curve exhibit nearly normal distribution and is closely symmetrical. The small value of standard deviation along with skewness at 1.142112 and 1.529152 respectively indicate that the data values are closely distributed with the mean and the data is positively skewed moderately towards the right. The curve is leptokurtic as the value of kurtosis is greater than 3.

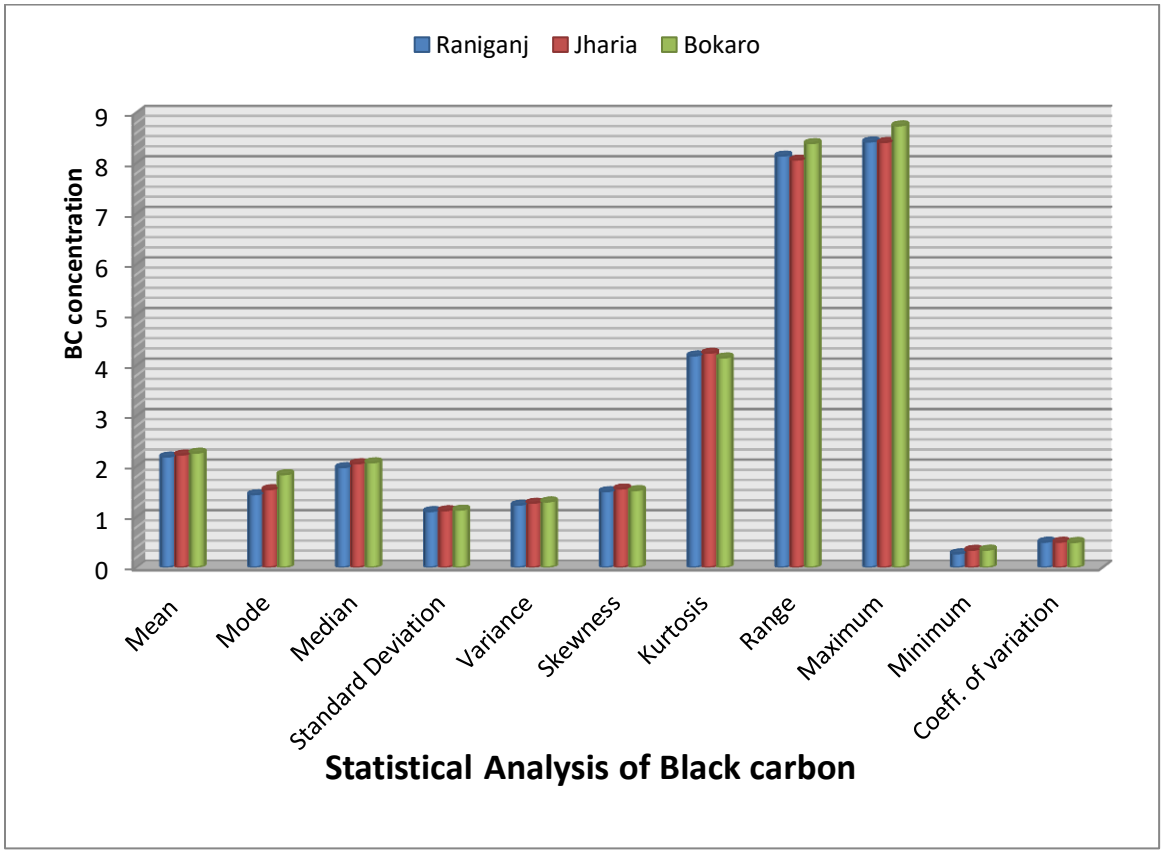


Figure 2.1 Statistical analysis of black carbon at Raniganj, Jharia, and Bokaro

Table 2.1 Statistical analysis of black carbon

	Raniganj	Jharia	Bokaro
No. of observations	461	461	461
Mean	2.192876	2.232444	2.269679
Mode	1.447137	1.548387	1.842105
Median	1.983193	2.059908	2.078582
Standard Deviation	1.113655	1.130772	1.142112
Variance	1.240226	1.279	1.304419
Skewness	1.507609	1.560077	1.529152
Kurtosis	4.192889	4.245786	4.152666
Range	8.145595	8.065878	8.391813
Maximum	8.423128	8.414977	8.742690
Minimum	0.277533	0.349099	0.350877
Coeff. of variation	0.507851	0.506517	0.503204

In general, nearly the same statistical results are observed at all the three sample sites, thus the distribution curve follows the same pattern and trends. The small value of standard deviation, skewness, coefficient of correlation, and closeness in the value of mean, mode and median indicates that the distribution curve nearly follows normal distribution as they are closely scattered near each other, and the shape of the curve is leptokurtic in all the three cases.

2.3.2 Statistical Analysis using Trend Analysis

To develop a comparison for the amount of correlation among the three-time series with their past observations we perform a trend analysis of the time series for the three sample sites for this we obtain the regression equation of lines taking two locations at a time using equation (2.4). Six regression equations of the line for site-wise comparison of the three sample sites are obtained as shown in Table 2.2. Regression coefficients of the line

obtained from these are further used for trend analysis by finding the values of Hurst exponent, fractal dimension, and predictability index. In trend analysis, R square is used to calculate the extent of variability among the two parameters or time series. Its value ranges between 0 and 1, closer the value to 1 in all the cases, indicates the forecasts obtained will be with minimum forecasting errors, and thus the dependent series can be accurately predicted from the independent series. The value of R square in all the cases is near to 0.9 which indicates that there is a huge correlation in the time series data among the three sample sites.

The value of Hurst exponent *H* obtained using equation (2.5) is less than 0.5 and that of fractal dimension *D* obtained using equation (2.7) is greater than 1.5 in all the cases. Thus it can be concluded that the time series exhibit a mean reversion or an anti-persistent behavior. In this case, an anti-persistent trend pattern exists in the time series in which after a fall there is considerable rise for a period which is again followed by a fall for some period and there is a regular repetition of this behavior. Thus this time-series can be efficiently modeled to generate a good forecast leading to accurate prediction results which are shown by the value of the predictability index formed using equation (2.8) which is close to 1 in all the cases. The monthly trend percentage value for the sample sites of Raniganj, Jharia, and Bokaro is 10.5%, 14.45%, and 18.28% respectively as shown in Table 2.3. Thus for the time series of all the three locations, it can be easily seen that compared to the previous time the trends are positive with a sufficient amount of variation, and the highest trend is observed at Bokaro.

Table 2.2 Fractal analysis for the cross-correlation among the three sites

Comparative study of sites	Raniganj with Jharia	Jharia with Raniganj	Jharia with Bokaro	Bokaro with Jharia	Bokaro with Raniganj	Raniganj with Bokaro
Regression equation	y = 0.9742x + 0.096	y = 0.945x + 0.0833	y = 0.9628x + 0.1203	y = 0.8988x + 0.2188	y = 0.9443x + 0.0496	y = 0.9912x + 0.0667
R square (r^2)	R^2 = 0.9206	R^2 = 0.9206	R^2 = 0.9086	R^2 = 0.8652	R^2 = 0.9379	R^2 = 0.8899
Reg.coeff.(b_{yx})	0.9742	0.945	0.9628	0.8988	0.9443	0.9912
Hurst exp. *(H)* *(abs)*	0.0129	0.0275	0.0186	0.0506	0.02785	0.0044
Fractal *(D)*	1.9871	1.9725	1.9814	1.9494	1.97215	1.9956
Predictability index *(PI)*	0.9742	0.945	0.9628	0.8988	0.9443	0.9912

Table 2.3 Trend percent analysis of black carbon of the major coal mines

Site	Raniganj	Jharia	Bokaro
Trend%	10.51131	14.45501	18.28012

2.3.3 Statistical Analysis using Correlation

For further analysis, a correlation comparison of the three time-series has been established using Karl Pearson's, spearmen's rho, and Kendall's tau correlation coefficients. Correlation comparison between these time-series is done by considering them two at a time; the values of the correlation coefficients obtained are shown in Table 2.4. Using Karl Pearson's, spearmen's rho, and Kendall's tau correlation coefficients, maximum values of the correlation coefficient are observed between Raniganj and Bokaro as 0.968, 0.969, and 0.856. For Raniganj and Jharia the values are 0.959, 0.960, and 0.834, between Jharia and Bokaro they are 0.953, 0.948, and 0.813. This shows that there is a high degree of correlation in the positive direction among the three sites as the correlation coefficient values are very close to 1.

Table 2.4 Comparison of coefficients of Correlation

CORRELATION COEFFICIENT'S	KARL PEARSON'S			SPEARMEN'S RHO			KENDALL'S TAU_B		
Sample sites	Raniganj	Jharia	Bokaro	Raniganj	Jharia	Bokaro	Raniganj	Jharia	Bokaro
Raniganj	1	.959	.968	1	.960	.969	1	.834	.856
Jharia	.959	1	.953	.960	1	.948	.834	1	.813
Bokaro	.968	.953	1	.969	.948	1	.856	.813	1

To get a clearer picture of the similarity among the time series data we have constructed a cross-correlation function (CCF) plot by considering two locations successively as depicted in Figure 2.2 at corresponding lags. The results are tested at a 5% level of significance showing maximum positive correlation at lag zero, lag 12, and lag 24. A comparison of correlation among the three time-series is further made by plotting them together taking two at a time as given in Figures 2.3, 2.4, and 2.5. As the time-series are nearly overlapping each other this shows that a significant amount of correlation between Raniganj and Bokaro followed by that between Raniganj and Jharia and then between Jharia and Bokaro.

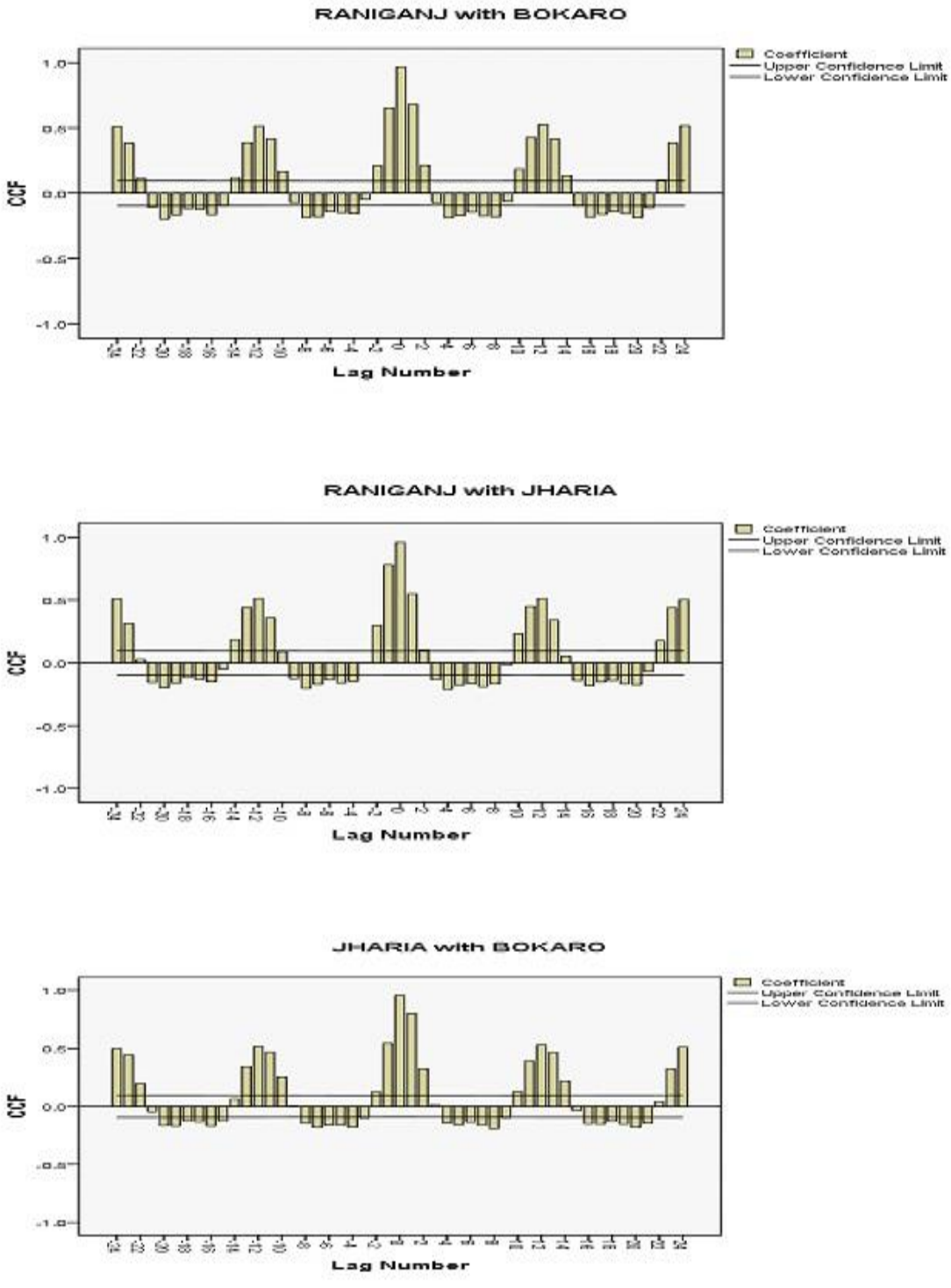


Figure 2.2 Cross-correlation function plots

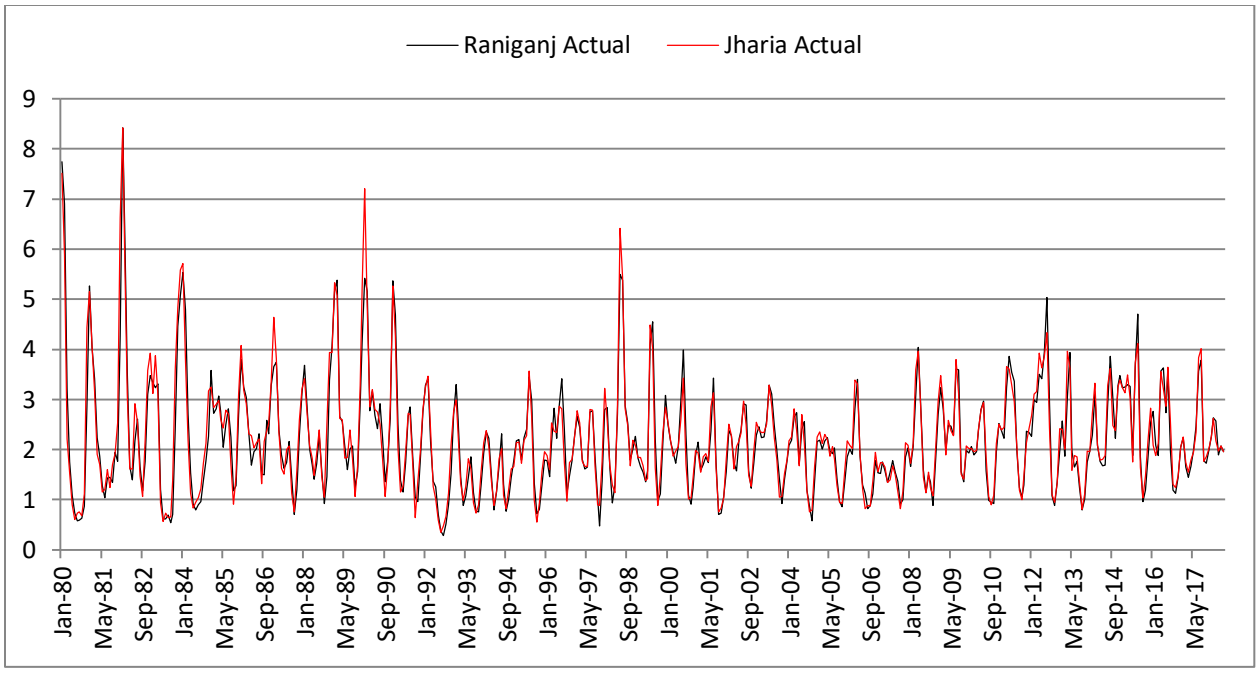


Figure 2.3 Time series of Raniganj and Jharia

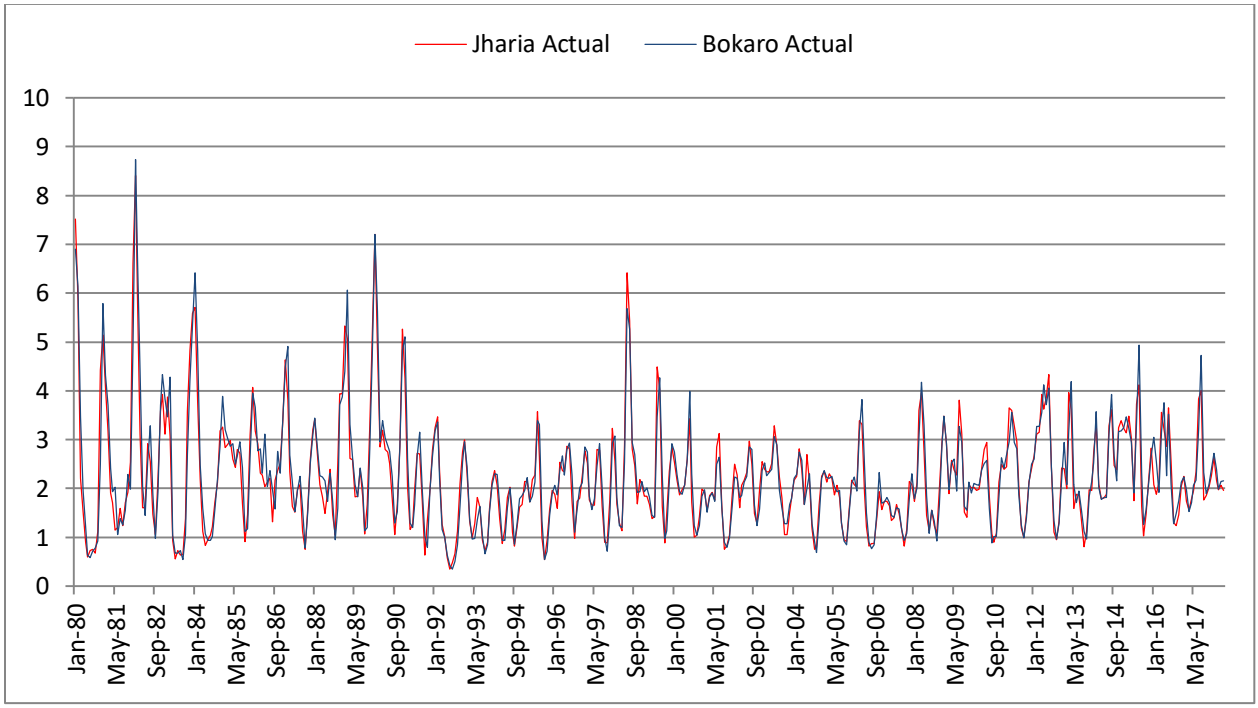


Figure 2.4 Time series of Jharia and Bokaro

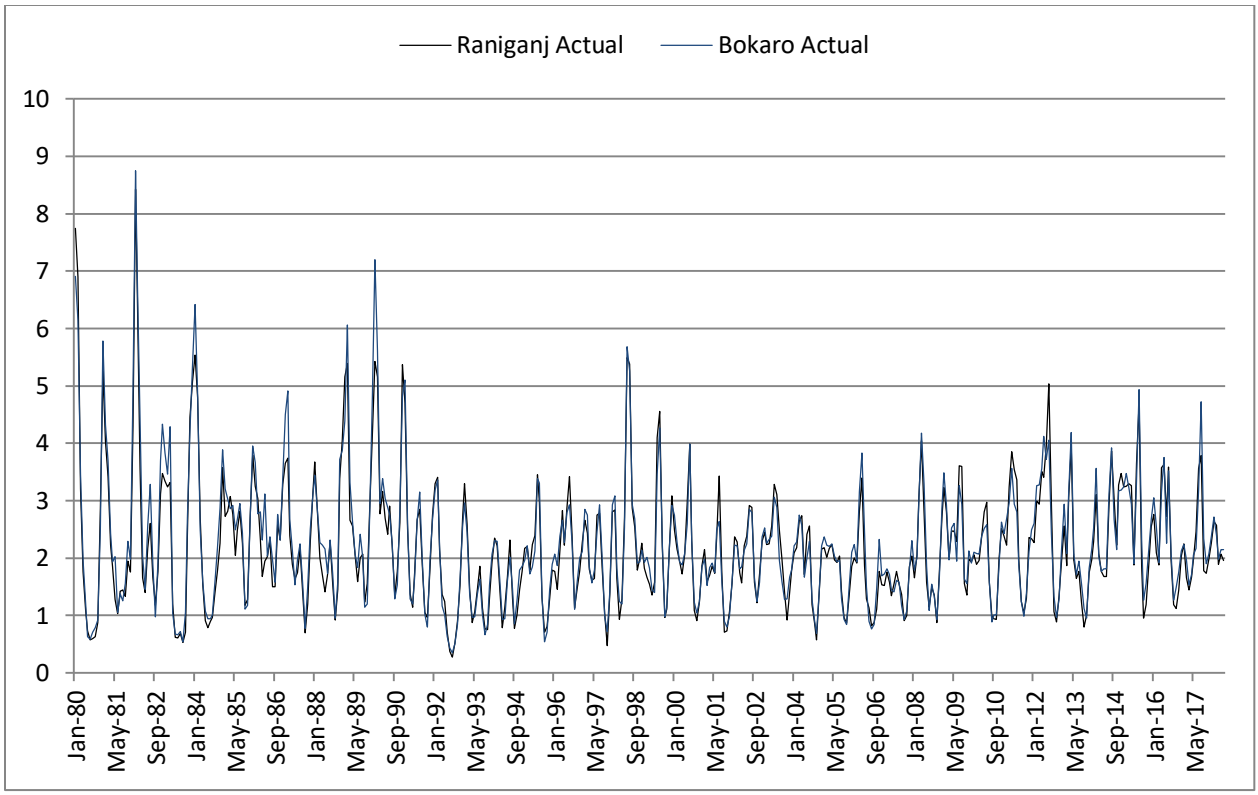


Figure 2.5 Time series of Raniganj and Bokaro

2.4 Conclusion: A long term time series data for the amount of concentration of black carbon in the three major coal mine regions of India is considered for the study. The results have been tested and compared at 95% confidence intervals to determine the nature and behavior of time series at the three locations. For this statistical analysis, along with trend analysis consisting of statistical parameters like trend percent, predictability Index (PI), Hurst exponent, fractals, coefficient of determination, equation of regression line, regression coefficient and correlation analysis is performed. The results of statistical analysis obtained shows that the value of the mean, mode, and median for all the sample sites are very close to each other. Thus the distribution curve follows a normal distribution, low values of standard deviation, variance, skewness, and coefficient of variation shows that the data is not skewed and the values are closely scattered near each other. The shape of the distribution curve is leptokurtic in all three cases as can be seen from the value of kurtosis.

In trend analysis, the value of Hurst exponent is close to 0, the fractal dimension is close to 2, and predictability index is close to 1 depending on these values we can predict that for all the cases anti-persistent behavior is followed. Thus the time series display a huge tendency of reverting towards its long term mean value. Due to this every rise is followed by a fall for some time and this behavior continues to repeat further. As Brownian motion is not displayed by the results for all the cases thus different time series models can be effectively used for obtaining prediction for these time series. The time series exhibit positive trends at sample sites of Raniganj, Jharia, and Bokaro having a monthly trend percentage of 10.5%, 14.45%, and 18.28% respectively. Thus at Bokaro, there exists the highest trend and a sufficient amount of variation compared to the previous time than the other two locations. To determine the degree of association among the data values of the three sample sites we have performed correlation analysis considering Karl Pearson, Spearmen's rank, and Kendall's tau coefficients of correlation. Results display a high degree of correlation between the observed values among the sample sites with the maximum between Raniganj and Bokaro observed at lag zero, 12 and 24 showing that the data values of black carbon emission are seasonally correlated with each other.

The results so obtained will help us to develop an insight into the behavior of the time series over the coal mines of India by studying the irregular patterns in it with the help of statistical, fractals, and trend analysis enabling us to forecast and understand their future behavior. A high degree of correlation among the three sample sites shows that the black carbon emissions follow nearly a same pattern over these sample sites as shown by results obtained, thus nearly similar preventive measures are required to be taken by the government and some effective pollution control policies concerned with coal mining are needed to be framed to curtail the growing BC concentration in these areas.

Chapter 3 Time Series Modeling of Black Carbon

A model is a miniature illustration of something; Modeling is an art of developing models. Mathematical modeling refers to framing and developing a model based on a real-life situation using mathematical techniques. Mathematical modeling is used by mathematicians, economists, scientists, and engineers to develop a horizon and obtain a forecast for a real-life problem. Time series analysis has been extensively used by mathematicians in the field of mathematical modeling. Time series refers to a sequential set of data points measured over successive intervals of time. Various time series models have been developed and used by researchers for obtaining the forecasts of the time series and predict the future values based on previously observed values such as the AR, MA, ARMA, ARIMA, SARIMA, etc. Of all these time series models the most widely used is the Box Jenkins based ARIMA model and its different variations, these are based on the assumption of linearity and stationarity. The linear time series models are readily used due to its simplicity in understanding and application. Whereas for time-series displaying non-linear behavior, non-linear time series models like NAR, TAR, NMA, and the GARCH family of models like the GARCH, EGARCH, GARCH-M, GJR, IGARCH, APARCH, AGARCH, NGARCH, etc. are very popular.

These time series models have been extensively used by researchers in solving and discussing various real-life problems like forecasting the level of pollutants, groundwater prediction, economics, financial time series, electricity demand and consumption, stock exchanges, epidemics, etc. **(Zhang and Wang 2013[300], Farajzadeh et al. 2014[92], Kumar et al. 2015[141], Tularam and Saeed 2016[282], Shakti et al. 2017[250], Pandey et al. 2019[194])**. Time series models like the AR and MA model fail to model volatility or change in variance of a time series over a period of time. Various models dealing with heteroskedasticity like the ARCH, GARCH, EGARCH, and GARCH-M model have been used by researchers in time series for modeling and forecasting volatility **(Capobianco 1995[48], Caporin 2003[50], Nasr et al. 2010[180], De Capitani 2012[79], Gao et al. 2012[96], Dimitrakopoulos and Tsionas 2019[82]).**

In this chapter, for obtaining the time series forecast of black carbon, we have used the Box Jenkins based SARIMA model and attained a time-series forecast for the coming 5 years period for our sample sites using IBM SPSS software. Also, a forecast of conditional volatility is obtained using the GARCH, EGARCH, and GARCH-M models by Num-XL software. A volatility forecast is also obtained for a period of 5 years from Jun 2018 to May 2023 using these GARCH models. A comparison is developed among these conditional volatility models, so as to find the best-fitted model for our sample sites. For an efficient model fitting measure of goodness of fit like the ACF, PACF, R-square, AIC, BIC, and Error measures like RMSE, MSE, MAPE, and MAE are used so as to choose the appropriate time series and volatility model **(Chelani and Devotta 2006[60], Shetty et al. 2018[252], Ayele et al. 2020[27])**. After selecting the best fitted time series ARIMA model-based on Box Jenkins methodology and GARCH model, we have obtained a forecast of time series data values and the conditional volatility.

3.1 Literature Review

Ballester et al. 2002[29] used FIR neural networks, multilayer perceptron, and ARMAX model for obtaining 24hr prediction of ozone concentration in Spain, using a time-series data of over 4 years from 1996 to 1999. The accuracy of the predictions due to these models led to an effective warning among the European Union. Study of hourly emission of Nitrogen oxides for a coal power plant in Charlotte was conducted by **Abdel-Aziz and Frey 2003[2]**; multivariate time series models were used for predictions of ambiguity in hourly emissions and the relationship between the emissions was analyzed. A total of 330 ARCH models were compared with each other on the basis of conditional variance by **Hansen and Lunde 2005[107]** comparative results showed that GARCH(1,1) outperformed the other GARCH models in analyzing exchange rates while considering return and exchange rate data. A study based on ARMA model selection was performed by **Qian and Zhao 2007[217]** considering a random sample generation algorithm and the maximum likelihood method using parameter estimation.

For predicting the SO_2 level at five stations in Tehran from 2000-2005 **Hassanzadeh et al. 2009[109]** used time series analysis. High emission levels were

recorded during autumn-winter and lowest in spring-summer. For the prediction of the concentration of the submicron particle in Hangzhou, China; **Jian et al. 2012**[129] used the ARIMA model. The results suggested that while obtaining forecast of these particles meteorological factors acted as significant predictors. **Byun and Cho 2013**[46] used GARCH type models to forecast the volatility of carbon future prices and examined their volatilities forecasts. ARIMA modeling and trend analysis were used by **Narayanan et al. 2013**[179] on rainfall data for western India. An increasing trend was observed at the two stations of Bikaner and Ajmer, for the change in rainfall practical significance was done in terms of percentage changes. ARIMA model was used to obtain a forecast of pre-monsoon showing that there was a significant rise in the pre-monsoon rainfall.

Time series ARIMA model was used by **Zhao and Wang 2014**[302] to obtain the forecast of crude oil considering a long term data from 1970 to 2006, results showed that the model gave optimal forecasts. ARIMA model and statistical methods were used to forecast future values for the quality of water parameters at the Yamuna River in Delhi by **Parmar and Bhardwaj 2014**[202]. In the southern, eastern, and south-eastern coal fields of India, **Soni et al. 2014**[263] carried out long-term trend analysis and studied aerosol optical depth (AOD) variability from Mar. 2000 to Dec. 2012 using statistical methods and terra-MODIS (Moderate-Resolution Imaging Spectro-radiometer). An increasing trend in AOD was observed in the coalfields of Godavari Valley (32 %), Korba (5.0 %), and Raniganj (7.31 %) in the study. Considering a time series data of 13 years **Soni et al. 2015**[265] obtained a forecast from Jan. 2013 to Dec. 2015 using the ARIMA model, aerosol optical depth (AOD550 nm), and Terra MODIS, over 11 coal mining sites of India for the assessment and management of air quality.

Zeitlberger and Brauneis 2016[298] used GARCH models for studying the behavior and modeling return series for carbon emission in the European Union ETS. ARMA-GARCH approach was used by **Razali and Mohamad 2018**[230] to study the problems for volatility in prices of pepper and crude palm oil in Malaysia and it was observed that GARCH (1,1) was the best-fitted model to obtain accurate forecasts. The climatology of long-term aerosol over the central Indo-Gangetic Plain (IGP) was done by

Kumar et al. 2018[142] using the ARIMA model; the study indicated different patterns in the distribution of AOD over the IGP. In Seoul, South Korea, the time-series SARIMA model was used by **Alsharif et al. 2019**[19] for obtaining accurate prediction of the daily and monthly average global solar radiation considering a data of 37 years (1981-2017). **Ceylan 2020**[55] used the best fitted ARIMA models to study the epidemiological trend of COVID-19 outbreak in Italy, Spain, and France, it was concluded that ARIMA model could be used as the best forecasting tool to predict the outcomes of the pandemic.

3.2 Time-series Analysis

For obtaining a time-series forecast for the black carbon concentration over the three sample sites the Box-Jenkins based ARIMA methodology has been used. For choosing the best-fitted time-series model, modeling parameters like ACF, PACF, R-square, Error measures like RMSE, MSE. MAPE and MAE are considered.

3.2.1 Partial Correlation Analysis: The partial correlation (PC) is an estimation of direction and strength of two variables measured in terms of the amount of correlation between two variables after controlling the effect of other variables. For e.g. $\rho_{XY.Z}$ is a correlation between $X\{x_i\}$ and $Y\{y_i\}$ for *n* controlling variables $Z=\{z_1,z_2,...,z_n\}$, it is defined as **(Parmar and Bhardwaj 2014**[202]**),**

$$\rho_{XY.Z}=\frac{\rho_{XY}-\rho_{XZ}\rho_{YZ}}{\sqrt{1-\rho_{XZ}^2}\sqrt{1-\rho_{YZ}^2}} \tag{3.1}$$

3.2.2 Autocorrelation and Partial Autocorrelation Functions (ACF and PACF*)*: The relation of the variables in the sample can also be determined by ACF and the PACF; these are calculated at consecutive time lags for the purpose of forecasting and modeling. For the sample $\{x(t),t=0,1,2,...\}$ at lag k, the Auto-covariance is defined as,

$$\gamma_k=Cov(x_t,x_{t+k})=E\left[(x_t-\mu)(x_{t+k}-\mu)\right] \tag{3.2}$$

At lag *k* the Auto-correlation coefficient is defined as, $\rho_k=\frac{\gamma_k}{\gamma_0}$, where μ is the mean of sample i.e. $\mu=E[x_i]$. γ_0 represents the variance of the sample at lag zero, ρ_k is dimensionless, and independent of the scale of measurement. Also, $-1\le\rho_k\le1$.

Relation of ACF and PACF for first three lags is defined as below,

$$\left.\begin{aligned} PACF(1) &= ACF(1) \\ PACF(2) &= \frac{ACF(2)-\left(ACF(1)\right)^2}{1-\left[ACF(1)\right]^2} \\ PACF(3) &= \frac{-2\left(ACF(1)\right)ACF(2)-\left[ACF(1)\right]^2 ACF(3)}{1+2\left[ACF(1)\right]^2 ACF(2)-\left[ACF(2)\right]^2-2\left[ACF(1)\right]^2} \end{aligned}\right\} \quad (3.3)$$

The non-randomness in the data is calculated by using the autocorrelation function and if the data are not random then it is used for the identification of a suitable time series model **(Parmar and Bhardwaj 2014[202]).**

3.2.3 Box Jenkins Methodology

Box Jenkins Methodology was developed by statisticians 'George Box' and 'Gwilym Jenkins' **(Box and Jenkins 1976[44])** in the year 1970 to design a mathematical time-series model for obtaining forecast of a given time series data. This methodology is used to build an ARIMA model satisfying the parsimony principle which best fits a given time series. Time series is basically categorized as stationary and non-stationary, the one in which the statistical terms such as mean, mode, median, variance, autocorrelation, etc. are all constant and time-independent are called stationary while the other in which these are time-dependent are called non-stationary. In a stationary series, the ACF decays fairly or quickly over time lags while in the case of non-stationary the decay is very slow. Stationary time series can be modeled by using AR(p), MA(q), ARMA(p,q) model while for modeling the non-stationary time series it is converted to stationary time series with Box Jenkins methodology by using the difference operator d of the ARIMA (p,d,q).

The main aim of this methodology is to develop an appropriate model that can be used to obtain a highly accurate time series forecast with minimum forecasting error. It does not assume the time series to follow a specific pattern, in fact, it is based on three steps namely model identification, parameter estimation, and diagnostic checking to select the best parsimonious model **(Chelani and Devotta 2006[60], Parmar and Bhardwaj 2015[200])**. Diagrammatically it can be represented as:

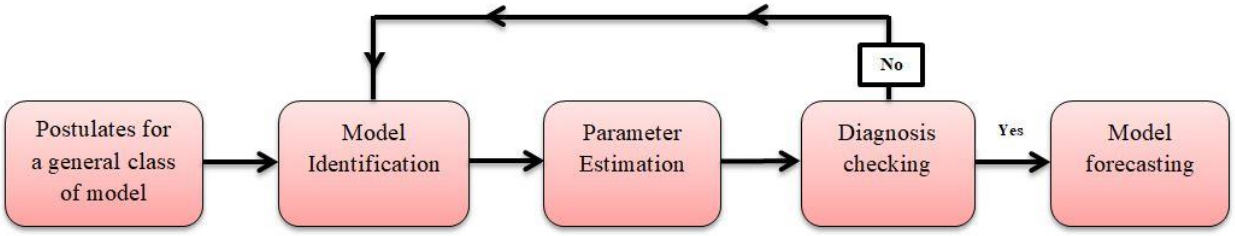


Figure 3.1 Diagrammatic representation of Box Jenkins Methodology

For selecting the best model this process is repeated several times until optimal model parameters values are obtained. The most widely used statistical techniques for the selection of optimal models are the AIC, BIC, R-squared values, and error measures like RMSE, MSE, MAPE, MAE.

3.2.4 ARIMA Model Analysis

ARMA model has successfully been used to obtain forecasts of stationary time series dealing with various types of data. ARMA model is developed by the sum or a combination of Autoregressive (AR) and Moving Average (MA). ARMA model has been successively used to study time series forecasts **(Chen et al. 1995[62], Cortez et al. 2004[67], Mohammadi et al. 2006[170], Rojas et al. 2008[234], Pappas et al. 2010[197], Benmouiza and Cheknane 2016[35], Ratnam et al. 2019[229]).** But most of the time series dealing with real-life situations exhibit a non-stationary behavior this may be due to man-made changes or sudden changes in the surrounding for eg. seasonal variations, trends, forest fires, recession, riots, epidemics, war, etc **(Aminzadeh 2009[20],**

Chattopadhyay and Chattopadhyay 2009[58], Contreras-Reyes and Palma 2013[66], Al-Gounmeein and Ismail 2020[18]).

To overcome this limitation of the ARMA model, the Box Jenkins methodology based ARIMA(p,d,q) model was developed which uses the differencing parameter *d* to convert a non-stationary time series to a stationary time series **(Nury et al. 2017[190])**. The first order differencing refers to the difference between the current and the previous data values while second differencing refers to the differencing between the values of the first differencing, the process is repeated until the non-stationarity in the series is reduced. To check stationarity the ACF (Autocorrelation function) and PACF (Partial Autocorrelation function) plots are most readily used. These can further be used to identify the values of the parameters *p* and *q* **(Parmar and Bhardwaj 2015[200], Shakti et al. 2017[250], Wang et al. 2018[288], Singh et al. 2019[258], Benvenuto et al. 2020[36])**. The ARIMA(p,d,q) model is given in general form as **(Soni et al. 2015[265], Nury et al. 2019[191]),**

$$\varphi(L)(1-L)^{d} y_t = \theta(L)\varepsilon_t, \quad \text{i.e.,}$$
$$\left(1-\sum_{i=1}^{p}\varphi_i L^i\right)(1-L)^{d} y_t = \left(1+\sum_{j=1}^{q}\theta_j L^j\right)\varepsilon_t \qquad (3.4)$$

Here $\varphi(L)$ corresponds to the Autoregressive (AR) part, $\theta(L)$ to the Moving average (MA) part, and *d* corresponds to the differencing parameter.

3.2.5 Seasonal ARIMA

ARIMA models are also capable of modeling the seasonal time series data with the help of seasonal ARIMA usually written as ARIMA$(p,d,q)(P,D,Q)_m$ where(p,d,q) refers to the non-seasonal part, (P,D,Q) refers to the seasonal part of the model and m is the number of observations available each year. The terms of the seasonal part are very similar to the non-seasonal components but the later involves backshifts of the seasonal part. A time-series is said to be seasonal if its pattern changes regularly and repeats itself

after a time period. A seasonal ARIMA is formed by the inclusion of some seasonal terms in the ARIMA model. The modeling procedure for both the seasonal and non-seasonal data are the same except in this we need to select the seasonal AR and MA terms as well as the non-seasonal components of the model. In seasonal ARIMA we use the AR models if the seasonal auto-correlation is positive, in this case of a purely seasonal AR model, the PACF falls to zero while the ACF decays slowly. While we use the MA model if the seasonal auto-correlation is negative, in case of a purely seasonal MA model, ACF cuts off to zero and vice versa for the PACF decays gradually **(Hamzacebi 2008[106], Abish and Mohanakumar 2013[5], Farajzadeh et al. 2014[92], Graham and Mishra 2017[101], Al-Gounmeein and Ismail 2020[18]**). The general equation of seasonal ARIMA is written as,

$$\Phi\left(B^m\right)\varphi\left(B\right)\nabla_m^D\nabla^d X_t = \Theta\left(B^m\right)\theta\left(B\right)Z_t \tag{3.5}$$

Or,

$$\left(1-\varphi_1 B\right)\left(1-\Phi_1 B^m\right)\left(1-B\right)\left(1-B^m\right)y_t = \left(1+\theta_1 B\right)\left(1+\Theta_1 B^m\right)\varepsilon_t \tag{3.6}$$

And,

$$\left.\begin{aligned}\left(1-B\right)x_t &= x_t - x_{t-1}\\ \left(1-B^{12}\right)x_t &= x_t - x_{t-12}\end{aligned}\right\} \tag{3.7}$$

Where φ_1, θ_1 represents the non-seasonal autoregressive and moving average parameters, Φ_1, Θ_1 represents the seasonal autoregressive and moving average parameters whereas, $\left(1-B^{12}\right)$ is the seasonal differencing and $\left(1-B\right)$ is the non-seasonal differencing.

3.3 Volatility Forecast using GARCH Models

A time series is said to be volatile if it represents unusual rises and falls in it, such behavior of a time series is called heteroskedasticity. In the case of heteroskedasticity, the

volatility of a time series changes periodically with time, and irregular patterns of the variations are displayed by the variable and they tend to cluster instead of forming a linear pattern. Volatility is also sometimes referred to as the unpredictability or unevenness in the market, it is a statistical measure of dispersion for a given time series and is often referred to as market index **(Mohamadi et al. 2017[169], Oh and Lee 2018[192], Mao et al. 2020[156])**. For comparing the conditional variances or volatility we first obtain the log return series for our time series and then apply the GARCH models.

3.3.1 Log Returns: For a time series $\{x_t, t = 1,2,3,...\}$, the log-returns series $\{P_t\}$ is defined as,

$$P_t = \log\left(\frac{x_t}{x_{t-1}}\right) \tag{3.8}$$

Since this series is unit-less therefore it can be easily compared with other series. The basic equation for volatility models is given by, $P_t = \sigma_t Z_t$, here $\{\sigma_t\}$ represents a non-negative stochastic process called the volatility process and $\{Z_t\}$ represents the asymmetric sequence of $i.i.d.$(independent, identically distributed) both of which are independent for each fixed t **(Abdalla 2012[1], Arachchi 2018[24])**.

3.3.2 Autoregressive Conditional Heteroskedasticity (ARCH) Model: It is notably used to find time-varying conditional variance or volatility over a period of time in a time-series. ARCH(p) model given by **Engle 1982[91]** are usually expressed as,

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i P_{t-i}^2 \tag{3.9}$$

In this conditional variance depends on the past values of the process, but ARCH models consist of several parameters. In order to get rid of high ARCH order, the Generalized ARCH (GARCH) model was developed containing few parameters in which the conditional variance is a linear combination of previous conditional variances and

squared innovations. These models are based on the basic assumption of Heteroscedasticity or Heteroskedasticity in which the variances are not constant **(Antonakakis and Darby 2013[23], Mohamadi et al. 2017[169], Arachchi 2018[24])**.

3.3.3 Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model: It is being used in this era of research as a useful tool in modeling volatility of such time series **(Storti and Vitale 2003[271], Storti 2008[272], Caporin and Lisi 2010[49], Doguwa and Omotosho 2012[85], Orhan and Koksal 2012[193], Katsiampa 2017[131], Arachchi 2018[24], El Jebari and Hakmaoui 2018[89], Trapero et al. 2018[280], Virginia et al. 2018[287], Dritsaki 2019[88], Ayele et al. 2020[27]).** In these models, we try to establish a relationship between the current and the past observations. The Generalized ARCH (GARCH(*p,q*)) model which was introduced by **Bollerslev 1986[41]** is mathematically written as,

$$\sigma_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i P_{t-i}^2 + \sum_{j=1}^{q} \gamma_j \sigma_{t-j}^2 \qquad (3.10)$$

Here $\beta_i's$ and $\gamma_j's$ represents non-negative parameters of the model if $q=0$ then it reduces to the ARCH(*p*) model. In the GARCH model, the numbers of parameters are reduced from infinitely many to only a few. P_t are conditional variances of the residual terms **(Nelson 1991[183], Abdalla 2012[1], Arachchi 2018[24], El Jebari and Hakmaoui 2018[89])**.

3.3.4 Exponential GARCH (EGARCH) Model: It is an asymmetric GARCH model which captures asymmetric responses for variances varying with time. It is represented as,

$$In\left(\sigma_t^2\right) = \beta_0 + \sum_{i=1}^{p} \beta_i g\left(Z_{t-i}\right) + \sum_{j=1}^{q} \gamma_j In\left(\sigma_{t-j}^2\right) \qquad (3.11)$$

A positive conditional variance independent of signs of the estimated parameters in the model is always produced having no restrictions **(Dritsaki 2019[88])**.

3.3.5 GARCH in Mean (GARCH-M) Model: It was developed by **Engle et al. 1987**[90] allows the conditional mean to depend on its own conditional variance. It is represented as,

$$r_t = \mu + c\sigma_t^2 + P_t \tag{3.12}$$

$$\sigma_t^2 = \omega + \alpha P_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3.13}$$

Where σ_t^2 represents the conditional variance and μ, c denotes mean and a constant. Estimation of the parameter is the initial step while fixing the GARCH model; it is commonly done by the maximum likelihood estimation (MLE) using the student's *t*-test, normal and generalized error distribution for better results.

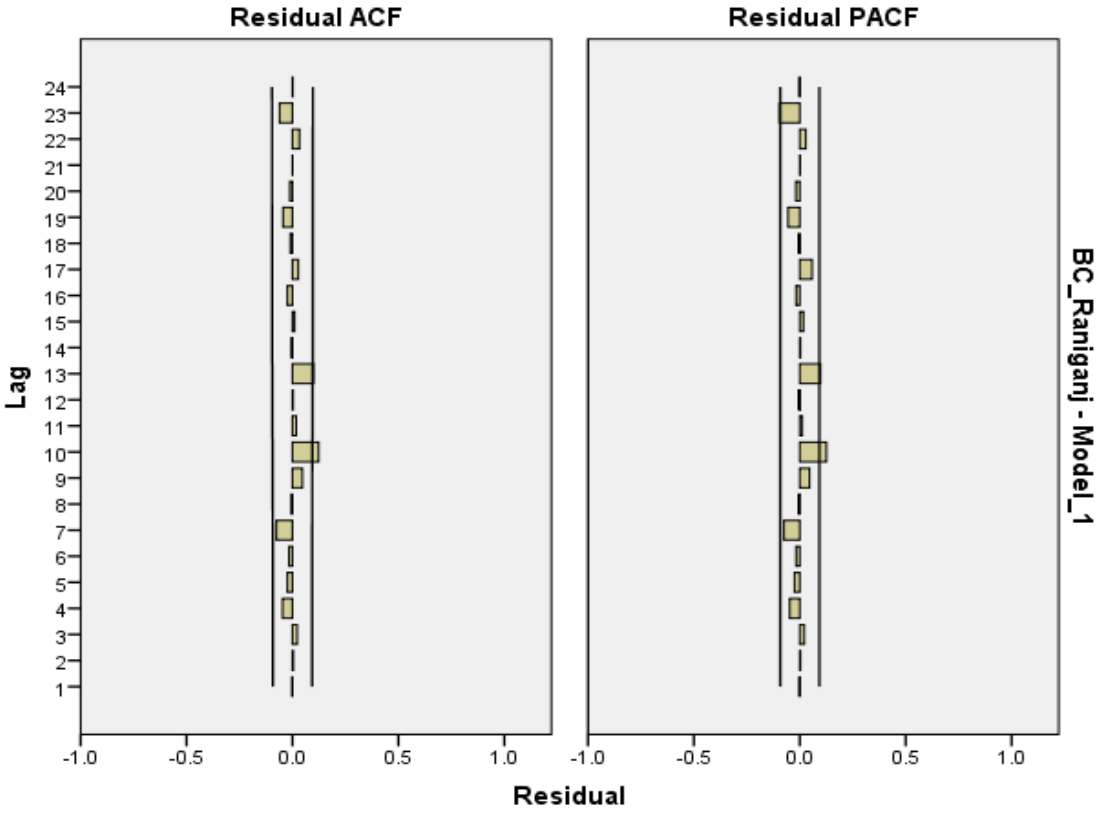
GARCH (1,1), EGARCH (1,1), and GARCH-M (1,1) models have been used for obtaining the volatility forecasts for a period of 5 years based on the data values. These are the most popular and readily used GARCH models, due to its mean-reverting property; it has an advantage over the EWMA method. A comparison is also developed among these three GARCH models, so as to find the best-fitted GARCH model for the three sample sites.

3.4 Time-series Forecasting of Black Carbon

Before selecting an appropriate time series model we check the stationarity of our time series data. If data is stationary then the forecast can be obtained using the ARMA model, but in case of non-stationary time series data ARMA model cannot yield good forecasts, then we introduce the Box Jenkins methodology based ARIMA model. In this, we use the differencing parameter d as discussed earlier in the ARIMA(p,d,q) methodology to convert our non-stationary data series into a stationary one. The stationarity is then determined or checked by plotting the ACF and PACF plots as shown in Figure 3.2 for the three sample locations, these plots are also beneficial for selecting the order of AR and MA models as they reflect the relation between the data values at different

consecutive lags. Using the above-mentioned techniques we have used the IBM SPSS software for training and testing our data and selecting an appropriate model.

After examining the values of the model fitting parameters like R-square, Error measures like the RMSE, MSE, MAPE, MAE, AIC, normalized BIC values, it was discovered that at 95% confidence limits with 460 degrees of freedom time series ARIMA $(1,0,1)(0,1,1)_{12}$ model was the best-fitted model. As discussed in the previous chapter that the data exhibited a seasonal behavior, thus seasonal ARIMA was discovered to be the best-fitted model. The values of the coefficient of determination i.e. R-squared and stationary R^2 which are used as the measures for the goodness of fit, were near to 1 indicating that this model fitted very well to the data. Also, the value of the error measures obtained is very small and near to zero. This model was further used for obtaining the forecast of the black carbon concentration for the coming 5 years period from Jun 2018 to May 2023.



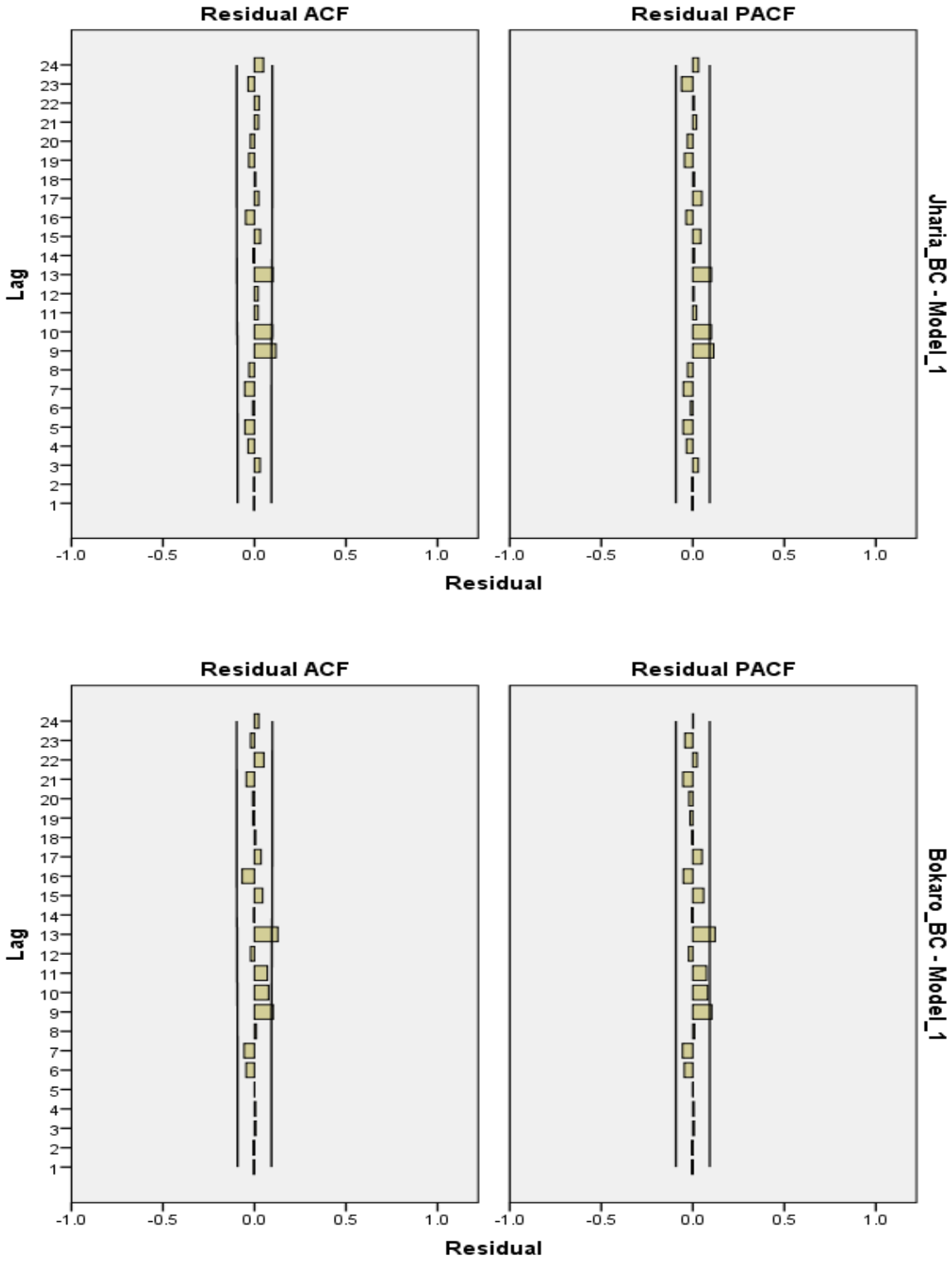


Figure 3.2 ACF and PACF plots for the sample locations

It can be seen from these ACF and PACF plots that after the application of the time series ARIMA $(1,0,1)(0,1,1)_{12}$ model, the data values represented a stationary behavior. Thus it can be further used for testing and training phases to obtain accurate time-series forecasts. The results obtained for the model fitting parameters for the three sample sites are discussed below.

Raniganj (23° 40' N 87° 05' E)

The values of coefficient of determination i.e. stationary R-square and R-squared values are 0.703 and 0.63, both these values are close to 1. The value of root mean square error (RMSE) is 0.64 and those of remaining error measures like MAPE, and MAE are very small. The normalized BIC value is -0.851. Thus it can be concluded that the predicted model performed better than the baseline model and can be efficiently used for time series forecasting. The mean predicted, upper confidence limit (UCL), lower confidence limit (LCL) and residual values are observed to be 2.334806, 4.740448, 0.980680, and 0.004330.

Jharia (23° 50' N 86° 33' E)

The values of coefficient of determination i.e. stationary R-square and R-squared values are 0.678 and 0.624, both these values are close to 1. The value of root mean square error (RMSE) is 0.662 and those of remaining error measures like MAPE, and MAE are very small. The normalized BIC value is -0.783. Thus it can be concluded that the predicted model performed better than the baseline model and can be efficiently used for time series forecasting. The mean predicted, upper confidence limit (UCL), lower confidence limit (LCL), and residual values are observed to be 2.368480, 4.730589, 1.019490, and 0.004394.

Bokaro (23° 46' N 85° 55' E)

The values of coefficient of determination i.e. stationary R-square and R-squared values are 0.663 and 0.619 both these values are close to 1. The value of root mean square error (RMSE) is 0.677 and those of remaining error measures like MAPE, and MAE are very

small. The Normalized BIC value is -0.739. Thus it can be concluded that the predicted model performed better than the baseline model and can be efficiently used for time series forecasting. The mean predicted, upper confidence limit (UCL), lower confidence limit (LCL), and residual values are observed to be 2.426113, 4.910465, 1.023607, and 0.004353.

Table 3.1 Time series analysis of black carbon

Fit Statistic	Raniganj	Jharia	Bokaro
Stationary R-squared	0.703	0.678	0.663
R-squared	0.630	0.624	0.619
RMSE	0.640	0.662	0.677
MAPE	20.900	20.820	21.359
MAE	0.428	0.450	0.461
Normalized BIC	-0.851	-0.783	-0.739
Degree of Freedom	460	460	460

Table 3.2 Forecasted values for the sample locations

ARIMA (1,0,1) $(0,1,1)_{12}$ Model	Mean Predicted value	LCL value	UCL value	Residual value
Raniganj	2.334806	0.980680	4.740448	0.004330
Jharia	2.368480	1.019490	4.730589	0.004394
Bokaro	2.426113	1.023607	4.910465	0.004353

Thus ARIMA $(1,0,1)(0,1,1)_{12}$ model fitted very well to the data. The high value of the coefficient of determination and small value of the error measures along with the small value of normalized BIC indicate that the forecasted time series model is very close to the baseline model. This is also represented in Figure 3.3, 3.4, and 3.5, as the actual time series and the predicted time series coincide with each other. A forecast of the time series for the next 5 years starting from Jun 2018 to May 2023 is also obtained using this

model. The numerical result of predicted value, upper confidence limit (UCL), lower confidence limit (LCL), normalized BIC, and residual values for each of the sample sites obtained is also represented in these figures.

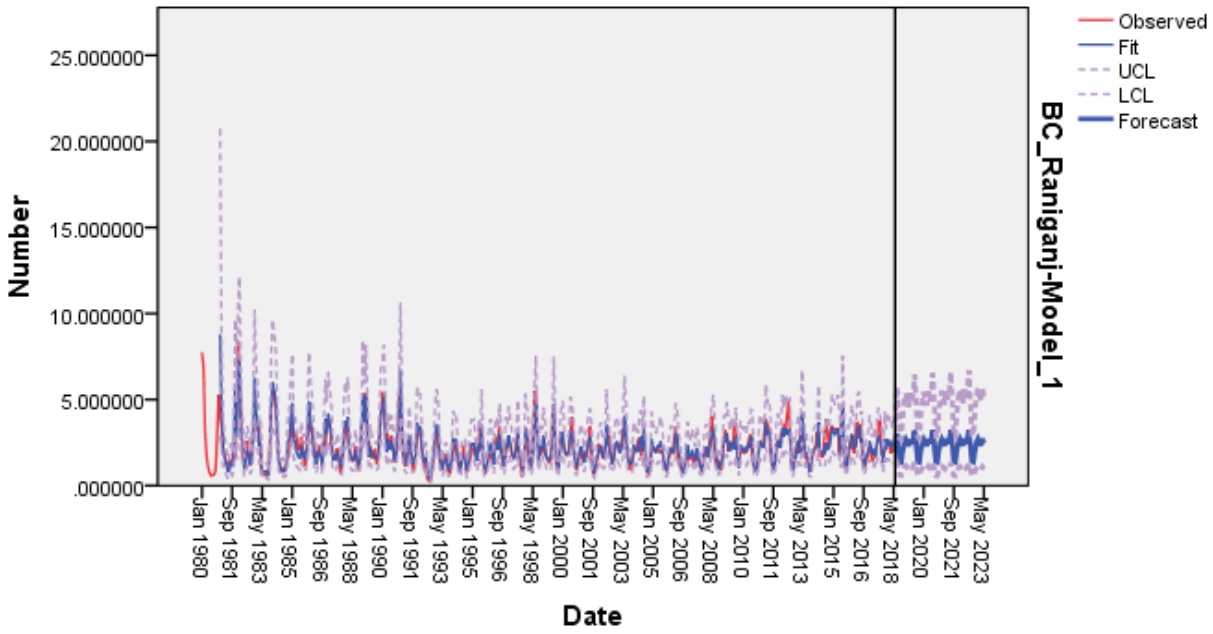


Figure 3.3 Time series and ARIMA forecast of Raniganj

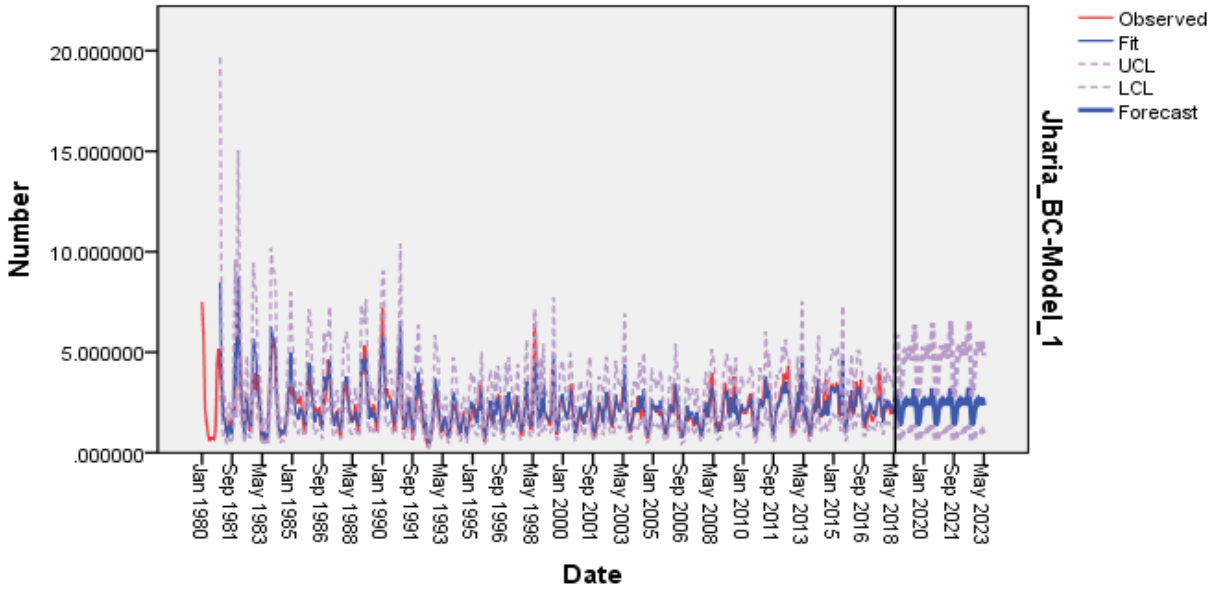


Figure 3.4 Time series and ARIMA forecast of Jharia

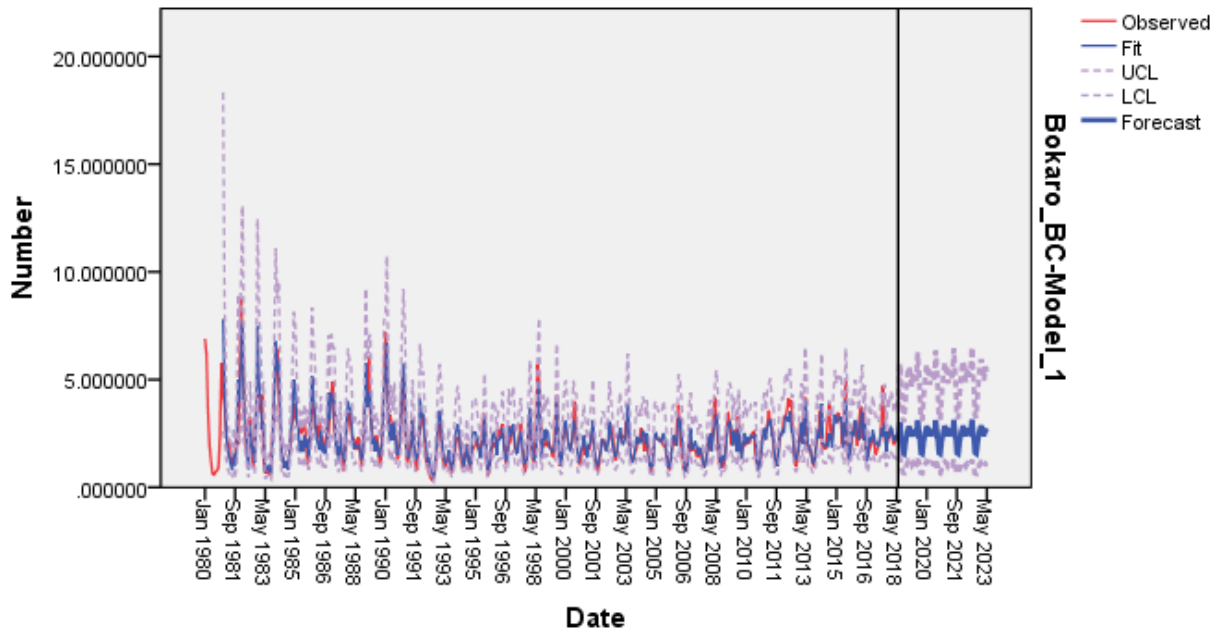


Figure 3.5 Time series and ARIMA forecast of Bokaro

The observed time series data curve is represented by red color and the fitted data values curve is represented by blue color. Since both the fitted and observed curves almost coincide, thus it can diagrammatically be shown that the model fitted gave quite reliable results with minimum forecasting error.

3.5 Forecasting Volatility

The summary statistics of the log return series for the three sample sites consisting of mean, median, standard deviation, skewness, kurtosis, quartile values, minimum and maximum values are shown in Table 3.3, nearly same results are seen at these three sample sites. The log return series is observed to follow a normal distribution and is platykurtic. The diagrammatic representation of the log-returns series; weighted moving average (WMA), exponential weighted moving average (EWMA) of the monthly data values for the three sample sites is shown in Figure 3.6, 3.7, and 3.8 from which it can be seen that the WMA and the EWMA are within the control limits. Figure 3.9, 3.10, and 3.11 represent the autocorrelation function (ACF) and partial autocorrelation function

(PACF) plots tested at 95% confidence limits for the log return series of Raniganj, Jharia, and Bokaro, which shows that the log return series is stationary.

Table 3.3 Summary statistics for the Log return time series of Raniganj, Jharia, and Bokaro

Sample sites	Raniganj	Jharia	Bokaro
Average	-0.00298	-0.00293	-0.00253
Standard Deviation	0.378808	0.377368	0.371401
Skewness	-0.04233	0.014307	-0.06196
Kurtosis	0.023822	0.27506	0.044631
Median	0.014112	0	0.018525
Minimum	-1.01663	-1.18801	-1.41182
Maximum	1.228164	1.379668	0.993001
1st Quartile	-0.25191	-0.23299	-0.23535
3rd Quartile	0.25069	0.24867	0.231786

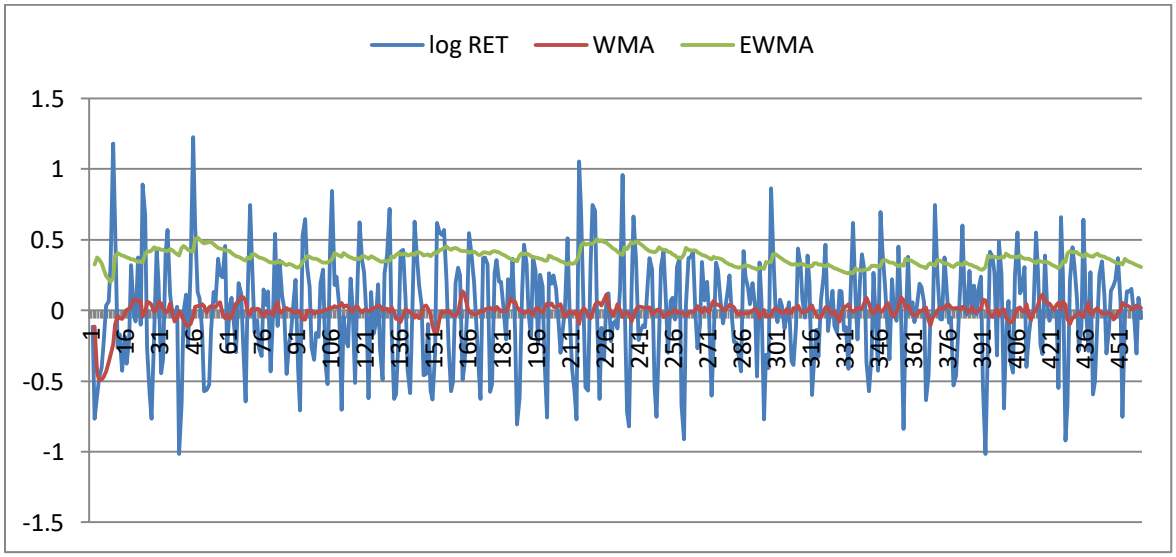


Figure 3.6 Log return, Weighted moving average (WMA) and Exponential (EWMA) of Raniganj

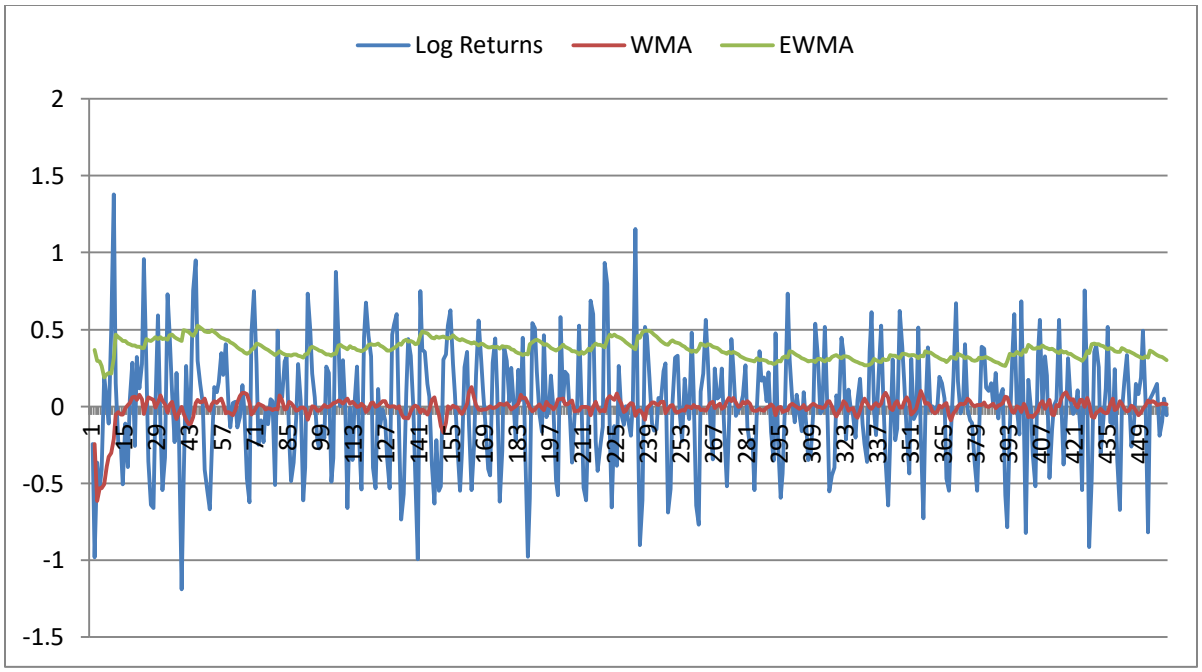


Figure 3.7 Log return, Weighted moving average (WMA) and Exponential (EWMA) of Jharia

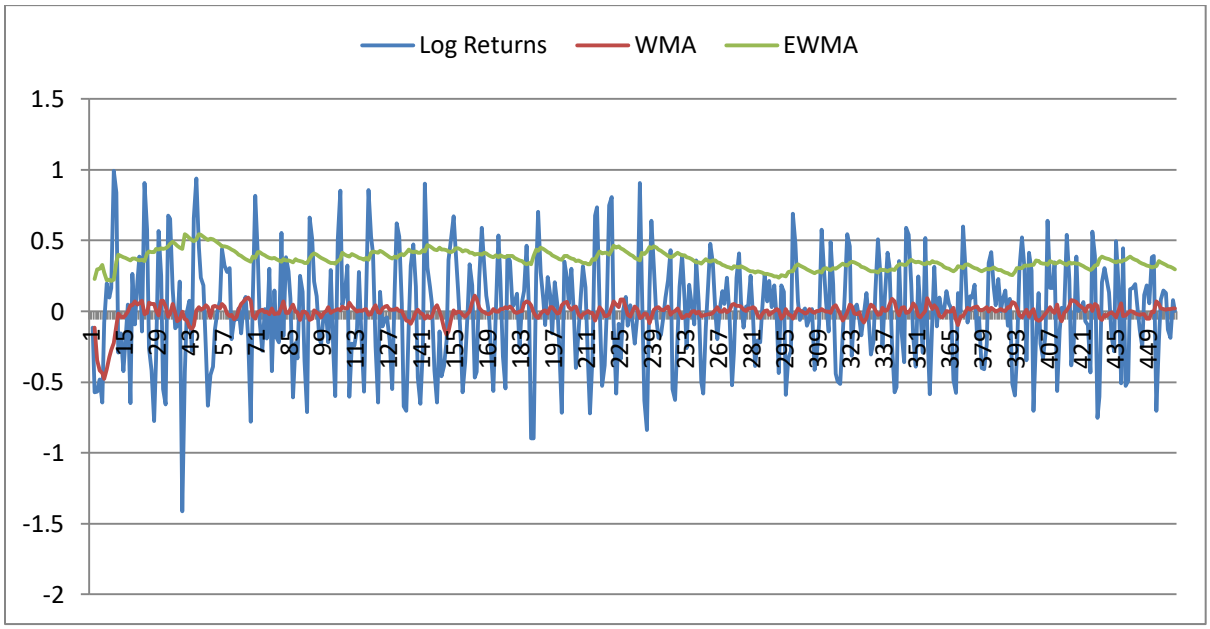


Figure 3.8 Log return, Weighted moving average (WMA) and Exponential (EWMA) of Bokaro

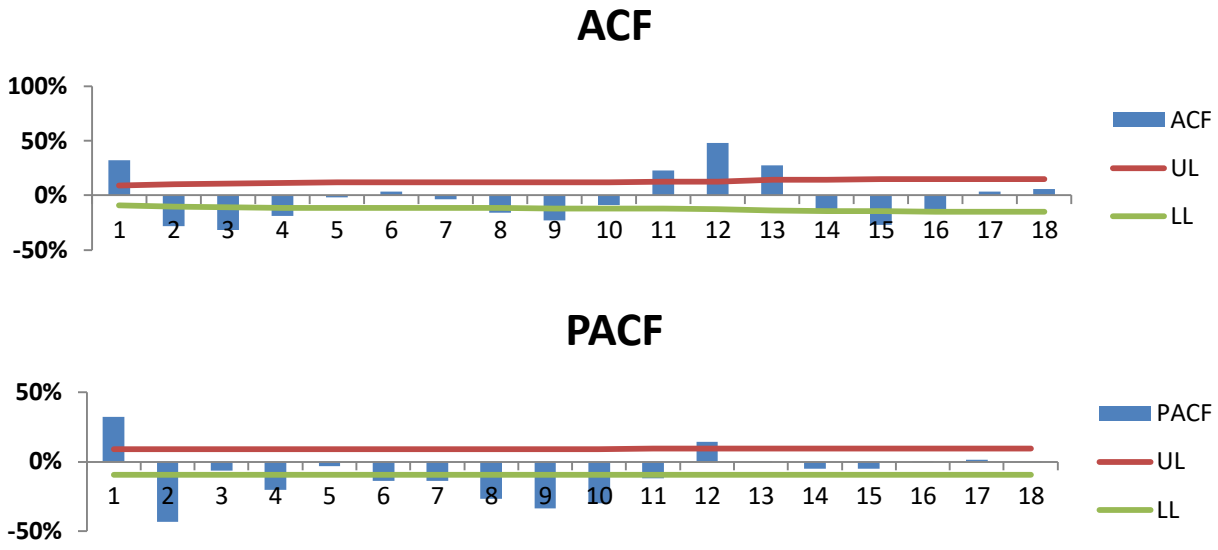


Figure 3.9 ACF and PACF plots for Log return series of Raniganj

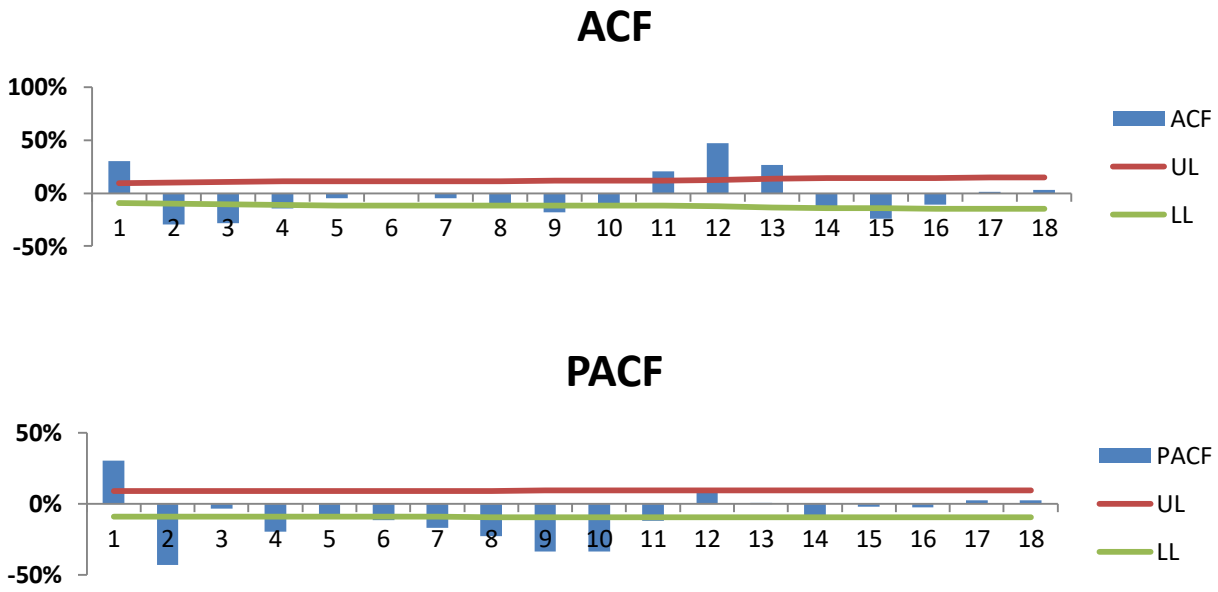


Figure 3.10 ACF and PACF plots for Log return series of Jharia

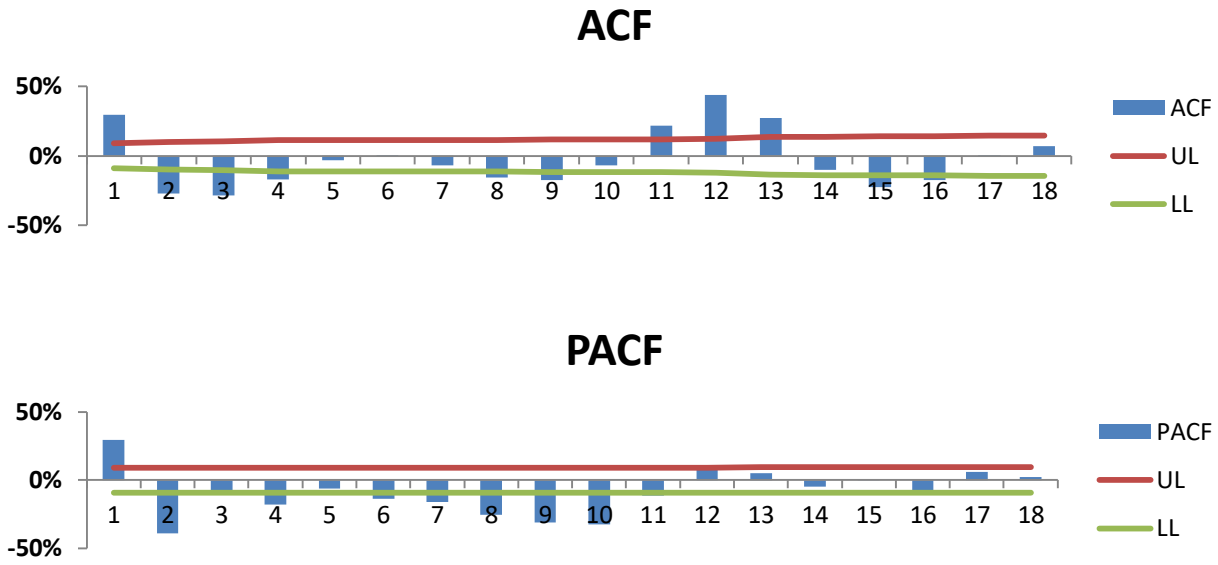


Figure 3.11 ACF and PACF plots for Log return series of Bokaro

Table 3.4, 3.5 and 3.6 below, represents the values of the parameters $\mu, \lambda, \beta_1, \alpha_0, \alpha_1, \gamma_1$ obtained for the EGARCH(1,1), GARCH(1,1), and the GARCH-M(1,1) models for each of the three locations respectively.

Table 3.4 Parameter values of GARCH models for Raniganj

EGARCH(1,1)		GARCH(1,1)		GARCH-M(1,1)	
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\boldsymbol{\mu}$	0.003368	μ	-0.00217	μ	-0.35495
$\boldsymbol{\alpha_0}$	-1.70396	α_0	0.096466	λ	0.9314
$\boldsymbol{\alpha_1}$	0.215867	α_1	0.129779	α_0	0.084189
$\boldsymbol{\gamma_1}$	-0.49521			α_1	0.118673
$\boldsymbol{\beta_1}$	0.217669	β_1	0.198737	β_1	0.29252

Table 3.5 Parameter values of GARCH models for Jharia

EGARCH(1,1)		GARCH(1,1)		GARCH-M(1,1)	
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\boldsymbol{\mu}$	0.018249	μ	0.001396	μ	-0.3057
$\boldsymbol{\alpha_0}$	-2.00681	α_0	0.086119	λ	0.812645
$\boldsymbol{\alpha_1}$	0.208517	α_1	0.116592	α_0	0.085805
$\boldsymbol{\gamma_1}$	-1.05338			α_1	0.115084
$\boldsymbol{\beta_1}$	0.063982	β_1	0.281998	β_1	0.283322

Table 3.6 Parameter values of GARCH models for Bokaro

EGARCH(1,1)		GARCH(1,1)		GARCH-M(1,1)	
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\boldsymbol{\mu}$	-0.00125	μ	-0.00366	μ	-0.00253
$\boldsymbol{\alpha_0}$	-1.47542	α_0	0.08019	λ	9.11E-11
$\boldsymbol{\alpha_1}$	0.348861	α_1	0.159661	α_0	0.118692
$\boldsymbol{\gamma_1}$	-0.17894			α_1	0.073913
$\boldsymbol{\beta_1}$	0.402079	β_1	0.263029	β_1	0.064384

As seen in Table 3.7, the error values viz. RMSE, MSE, MAPE, MAE for each of the models EGARCH(1,1), GARCH(1,1), and GARCH-M(1,1) is very less and close to zero for all the locations i.e. Raniganj, Jharia, and Bokaro. Also, the values of R-square for all of these models at the sample sites are very close, all near to 0.7 as given in Table

3.7. Also as observed from Table 3.8, the minimum value of AIC (Akaike information criterion) for Raniganj is 411.81 which exists for GARCH(1,1) with LLF (log-likelihood function) at -202.9 and volatility at 37.9% depicting that GARCH(1,1) is among the best-fitted model than the other two for Raniganj. For Jharia, the minimum value of AIC is 406.28 for EGARCH(1,1) with LLF at -198.14, and volatility value is 37.4%, thus EGARCH(1,1) is among the best-fitted model for Jharia location followed by GARCH(1,1). Also for Bokaro, the least value of AIC is 392.06 with LLF at -193.03 and volatility being 37.2% depicting that for Bokaro, GARCH (1,1) is among the best-fitted model.

Table 3.7 Error and R-square comparison at the three locations using GARCH models

	Raniganj			Jharia			Bokaro		
Model	EGARCH(1,1)	GARCH(1,1)	GARCHM(1,1)	EGARCH(1,1)	GARCH(1,1)	GARCHM(1,1)	EGARCH(1,1)	GARCH(1,1)	GARCHM(1,1)
RMSE	0.003	0.004	0.005	0.001	0.005	0.005	0.008	0.006	0.001
MSE	1.421E-05	1.892E-05	2.597E-05	2.745E-06	2.552E-05	2.564E-05	7.378E-05	3.613E-05	3.733E-06
MAPE	0.006	0.008	0.009	0.002	0.009	0.009	0.0171	0.011	0.003
MAE	0.002	0.003	0.003	0.001	0.003	0.003	0.006	0.004	0.001
R-Square	0.734	0.751	0.766	0.701	0.764	0.764	0.774	0.770	0.714

Table 3.8 Volatility and Model fitting

	Raniganj			Jharia			Bokaro		
	AIC	LLF	Volatility	AIC	LLF	Volatility	AIC	LLF	Volatility
EGARCH(1,1)	414.63	-202.31	0.3757	**406.28**	-198.14	0.3741	393.08	-191.54	0.3675
GARCH(1,1)	**411.81**	-202.90	0.3790	410.33	-202.16	0.3784	**392.06**	-193.03	0.3726
GARCH-M(1,1)	412.24	-202.12	0.3781	410.82	-201.41	0.3776	396.38	-194.19	0.3711

Table 3.9 Volatility forecast and Regression equation of the line for the forecast of volatility

Sample sites	Raniganj		Jharia		Bokaro	
	Volatility forecast for coming 5 yrs	**Regression equation of the line for the forecast**	**Volatility forecast for coming 5 yrs**	**Regression equation of the line for the forecast**	**Volatility forecast for coming 5 yrs**	**Regression equation of the line for the forecast**
EGARCH(1,1)	0.37243	y = 0.0002x + 0.3669	0.37283	y = 8E-05x + 0.3705	0.35913	y = 0.0005x + 0.3451
GARCH(1,1)	0.37504	y = 0.0002x + 0.3682	0.37358	y = 0.0003x + 0.3655	0.36687	y = 0.0003x + 0.3572
GARCH-M(1,1)	0.37322	y = 0.0003x + 0.365	0.37282	y = 0.0003x + 0.3647	0.36954	y = 9E-05x + 0.3667

The average volatility forecast for the coming 5 years and the equation of regression line for the forecasted volatility from Jun 2018 to May 2023 is obtained for the three locations using the EGARCH(1,1), GARCH(1,1) AND GARCH-M(1,1) model respectively as shown in Table 3.7. Thus the above results indicate that these models fitted very well to the time series data of black carbon concentration over the coal mines in India and they can be effectively used to obtain the forecasts of conditional variances and GARCH(1,1) being the best-fitted model out of these models.

3.6 Conclusion

As discussed above for all the three sample sites viz. Raniganj, Jharia, and Bokaro the following conclusions can be derived from the results obtained above,

3.6.1 Time Series Analysis

- Time series ARIMA (1,0,1) $(0,1,1)_{12}$ prediction model fitted very well to the data for all the sample sites.
- The application of this model resulted in very small values of error measures like RMSE, MAPE, MAE.

- The coefficient of determination values, namely R-squared and stationary R-square was close to 1, which are considered very good.
- Diagrammatic representation of the results is also given which shows that the original and the forecasted time series almost coincide. Thus the applied model resulted in very small error values.
- The numerical result of predicted value, upper confidence limit (UCL), lower confidence limit (LCL), normalized BIC, and residual values are obtained.
- A time-series forecast for the next 5 years starting from Jun 2018 to May 2023 is also obtained using this model.

3.6.2 Analysis of Volatility or Conditional Variances

- Small values of the errors viz. RMSE, MSE, MAPE, MAE, and high value of R-square which are all close to 0.7 for all the models i.e. EGARCH(1,1), GARCH(1,1) and GARCH-M(1,1) indicates that these models of conditional volatility fitted very well.
- Considering the value of AIC (Akaike information criterion) it is clearly seen that,

  a) For Raniganj, GARCH(1,1) is the best-fitted volatility model with least AIC value as 411.81, volatility as 37.9% and LLF at -202.9.

  b) For Jharia, EGARCH(1,1) is the best-fitted volatility model with the least AIC value as 406.28, volatility as 37.4%, and LLF at -198.14.

  c) Also for Bokaro, the GARCH(1,1) is the best-fitted volatility model with the least AIC value as 392.06, volatility as 37.2%, and LLF at -193.03.

- A forecast of average conditional volatility for the coming 5 years of the data from Jun 2018 to May 2023 along with the regression equation of the line of forecasted volatility is obtained using these GARCH models
- These models can be efficiently used as an effective tool in obtaining forecast of conditional variances for the coal mine regions to study the volatility for the concentration of pollutants along with other parameters.

Chapter 4 Wavelet Analysis of Black Carbon

Wavelets Analysis is a very hot topic in all the areas of research nowadays around the globe especially in the area of prediction modeling. Wavelets are mathematical functions, which are capable of writing down complex functions in an easy way. They break down data into its frequency components and so we can study each and every part of it in a more precise manner as it is scaled for our convenience. It can be termed as a tool to decompose signals and trends as a function of time. **Grossman and Morlet 1984[102]** gave the concept of wavelets which was developed thereafter by, **Meyer 1986[164]**, **Daubechies 1988[74, 76]**, and **Mallat 1989[155]** who further led to the development of orthogonal wavelet functions. However, until the end of the 1980s, the use of wavelets increased in the field of engineering. They are used in place Fourier Transforms as wavelets give more freedom to work with **(Parmar and Bhardwaj 2013[203], Kisi et al. 2017[137]).**

The Fourier Transforms and many other techniques were incapable of treating non-stationary signals due to the absence of any temporary information and also since in this frequency resolution could be predicted without the knowledge of the time at which it occurred. Since the information present in the signal is preserved by Fourier transforms in the frequency domain but not in terms of time domain, thus it is a function of time only and not of frequency. Thus Fourier transforms can be used efficiently only while dealing with a stationary time series. Short-time Fourier Transform (STFT) can treat non-stationary signals but the major drawback of this was with the change in the resolution the frequency derived assumed a constant value. At all the frequencies a constant resolution is provided by the short term Fourier transform while for wavelet transform we have a different resolution for different frequencies. Also, the wavelet transforms could handle the information in a signal both in time and frequency domain, so it was preferred over the Fourier transforms. A signal could also be decomposed into different components at various resolution levels with the help of wavelets which could further be used for obtaining more accurate time-series forecasts, wavelet transform gave high time resolution at high frequencies and high-frequency resolution at low frequencies

(Grossmann and Morlet 1984[102], Quiroz et al. 2011[218]). Due to the occurrence of trends in most of the time series dealing with real-life situations they exhibit non-stationary behavior. So wavelet transforms are readily used over Fourier transforms as it a function of time and frequency, more realistic and flexible of using in case of non-stationary data, having few variables than the Fourier transforms. Information present in the signal is preserved by wavelet transforms.

Wavelet methods are a very useful tool in analyzing a time series **(Can et al. 2005[47], Dokmen and Aslan 2013[86])** as it is used for removing noise from statistical data which is the most important job in data analysis. Wavelets are used in various important fields as quantum physics, artificial intelligence, image processing, visual recognition, mathematics, engineering, finance, economics, computer application, image, and signal processing **(Addison 2002[9], Zhao and Qi 2008[304], Bodyanskiy and Vynokurova 2013[40], Singh and Mohapatra 2019[261])**. It is considered as an important tool in studying the different features of biomedical applications like protein structure, DNA patterns, along with meteorological data, financial time series analysis, fractal analysis and studying variations or predictions of time series dealing with natural phenomena's like pollutant concentration, groundwater level, temperature along with various other fields **(Hu and Nitta 1996[116], Andreo et al. 2006[22], Kisi and Shiri 2011[139], Rajaee 2011[221], Wang et al. 2018[288], Saadaoui and Rabbouch 2019[238], Singh et al. 2020[259])**.

In this chapter, for obtaining the time-series forecast we have used the Box-Jenkins methodology based ARIMA model for all the three sample sites using MATLAB software. After this, a wavelet-ARIMA coupled approach is developed and applied to the time-series data. For this first the signals are decomposed using Daubechies wavelets (db8) level 3 and are further treated with suitable ARIMA model then the reconstructed the wavelet signals are used to obtain the results. The results obtained by time-series ARIMA and coupled wavelet-ARIMA models are then compared with each other for identifying the best-fitted model.

4.1 Literature Review

Grossman and Morlet 1984[102] found that an arbitrary square-integrable real-valued function (Hardy function) could be easily decomposed into a proper family of square-integrable wavelets of constant structure. If the admissibility condition of the wavelet is satisfied then the integral transform obtained is self-reciprocal and isometric. Later **Daubechies 1988**[74] developed an approach of multi-resolution analysis with various algorithms in reconstruction and image decomposition. He constructed an orthonormal basis for compactly supported wavelets, having arbitrarily high regularity. The order of which increased linearly with the support width and construction followed by the synthesis of distinct paths. Based on wavelet transforms, **Aksoy et al. 2004**[17] discovered a wind speed data generation scheme and related it with the present wind speed generation methods of Weibull and normally distributed independent random numbers, autoregressive models of first and second-order along with Markov chain of the first-order were used. Observations conclude that the approach based on wavelets was a better alternative to present methods.

Seasonal variations of atmospheric waves and characteristics of gravity waves (GWs) over Istanbul were analyzed by **Can et al. 2005**[47] using wavelets considering data from 1993 to 1997 for daily radiosonda of Istanbul in the lower stratosphere and troposphere. Variations were observed and compared in terms of pressure heights, air temperature, and deviations from mean values for annual, seasonal, monthly, and daily basis. **Quiroz et al. 2011**[218] with the help of rain gauge daily rainfall data and normalized difference vegetation index (NDVI) introduced a new approach by combining wavelet transforms and Fourier Transform for improvement of daily rainfall estimation of NDVI. For checking the nature of trends and hydrological variations from 1954–2008 in southern Ontario and Quebec at seven meteorological stations and eight flow stations **Nalley et al. 2012**[178] used wavelet transforms to construct a new trend detection method. For studying the annual, monthly, seasonal trends discrete wavelet transform and Mann–Kendall trend tests were used.

Several data-driven models like ANN, ANFIS, Wavelet-ANN, and Wavelet-ANFIS models were correlated for predicting groundwater level at various time periods, to analyze their results for prediction by **Moosavi et al. 2013**[173]. Results concluded that Wavelet-ANFIS models gave more reliable and accurate results than the others. In Golcuk, Turkey data for 13-months was studied using wavelets and then compared by **Dokmen and Aslan 2013**[86] with the results from laboratory analysis to study various water quality parameters. The samples were studied for pH values, nitrate, and chlorine for deterring their suitability for domestic use, drinking, and irrigation, seasonal variations were observed with wavelet analysis. Using Haar wavelets at level 3 water quality parameter such as PH, FC, TC, TKN, AMM, WT, DO, BOD, and COD were studied by **Parmar and Bhardwaj 2013**[203] at Nizamuddin bridge-mid stream (Delhi) of Yamuna River in India using last 10 years data.

A new forecasting wavelet coupled model was formed by **Parmar and Bhardwaj 2015**[200] for forecasting the monthly water quality parameter of COD at Nizamuddin station, New Delhi of river Yamuna. The performance of the wavelet coupled model was compared with the results obtained by combining regression models, artificial neural networks, classical neuro-fuzzy, and wavelets. Observations show that the wavelet coupled model produced better forecasting than the others. At Janakpuri, Nizamuddin, and Shahazada Bagh in Delhi, wavelet and statistical analysis of air pollutants for a period of more than 20 years from 1987 to 2010 in India was studied by **Soni et al. 2017**[264], results showed that the mean concentration of SO_2 decreased for both residential (Janakpuri, Nizamuddin) as well as industrial (Shahzada Bagh) area, whereas NO_2 increased but it was under the prescribed limits of National Ambient Air Quality Standards (NAAQS), discrete wavelet analysis of air pollutants using the Daubechies wavelet (level 5) was also used.

For forecasting, the short term flow of traffic **Zhang et al. 2018**[299] proposed a new hybrid multivariate model based on wavelets and time series SARIMA model termed as WSARIMAX which led to better forecasts compared to the other methods.

Salazar et al. 2019[243] used Haar discrete wavelet analysis for modeling the pollution of ozone (O_3) in Chile and obtain an hourly forecast of ozone using the coupled approach. Results obtained by the wavelet approach were then compared with the ARIMA model. Wavelet analysis was used over a sub-divisional rainfall data of India from 1871 to 2016 **Paul et al. 2020**[208] for decomposing the signals and then treated with ARIMA and ANN models for obtaining the prediction of rainfall, a forecast comparison was done among the ARIMA, Wavelet-ARIMA, and Wavelet-ANN and was observed that the latter outperformed the other two. **Singh et al. 2020**[259] used the wavelet ARIMA approach to forecast the deaths in five most affected countries due to novel coronavirus (COVID-19) and obtained one month ahead forecast for the situation.

4.2 Wavelet Transform

The waveform is a wave-like function that oscillates to zero, bounded in both frequency and time duration having finite energy determined in time. This property and capability of wavelet help us to examine the signals in time-variations. Fourier Transform is used for converting signals to a continuous series of sine/cosine waves or a combination of both the waves, having a fixed amplitude, frequency, and an infinite duration of propagation. On the other hand, signals can be converted to a series of wavelets using wavelet transforms (WT) **(Beylkin et al. 1991**[37]**)**. As many signals in reality (images, data signal, music, etc.) are of constant duration and have different variations in frequency. Thus a better approximation of the real-world signals can be constructed with rough edges by using WT **(Addison 2005**[10]**)**. For investigating a signal with finite energy function and short duration of propagation we use wavelet analysis, this helps to describe the function and transform the signal under examination into a new depiction which presents a clearer picture about the signal this transformation of the signal is called wavelet transform **(Kisi and Shiri 2011**[139]**)**.

The transform is then performed at several scales of the wavelet and at different points of the signal to complete the transform plane. With the help of WT, we can examine different resolutions of the signal at different frequencies by using multi-

resolution techniques. Wavelets are the basis functions of the wavelet transforms. As given in the equation below, the basis function represented by $\psi(t)$ is called the mother wavelets of the translated and dilated form $\psi_{a,b}(t)$ having the relation **(Rashid et al. 2015[227], Baghanam et al. 2019[28])**.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), a \in \mathbb{R}^+, b \in \mathbb{R}$$

Here, a is the scale of frequency and b is the time or spatial position. For any function, $f(t)$ the wavelet transform is given as,

$$T_\psi f(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt = \langle f, \psi_{a,b}\rangle \quad (4.1)$$

T_ψ denotes the wavelet transform

Properties of wavelet transform: The wavelet transform for a function $\psi(t)$ satisfying the following conditions,

i) The major property of wavelets is the admissibility and the regularity conditions i.e. the admissibility condition is satisfied by the square-integrable function $\psi(t)$, mathematically given as **(Maheswaran and Khosa 2012[151])**,

$$\int_0^{\infty} \frac{|\hat{\psi}(\omega)|}{\omega} d\omega < \infty \quad (4.2)$$

Here $\hat{\psi}(\omega)$ denotes the Fourier transform of $\psi(t)$

ii) In order to impose the admissibility, the condition $\hat{\psi}(0) = 0$ is taken for granted i.e. it vanishes at the zero frequency. This gives rise to the zero mean condition given as,

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0 \qquad (4.3)$$

This illustrates that a wavelet must have a band-pass like a spectrum. A zero at the zero frequency means that the average value of the wavelet in the time domain must be zero. It can be used for first analyzing and then rebuilding the signal without loss of information **(Singh et al. 2019[258])** and therefore it must be oscillatory. In other words, $\psi(t)$ must be a wave **(Parmar and Bhardwaj 2012[198], Singh and Mohapatra 2019[261])**.

iii) The wavelet function is of unit energy i.e. the square norm is 1, mathematically explained as,

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1 \qquad (4.4)$$

In addition to the above, the wavelets also satisfy the orthogonality property, in this information carried by the wavelets are completely independent of each other. This makes wavelets to be more convenient for use. The transform is said to be a continuous wavelet transform if the process is performed in a continuous and smooth fashion and discrete wavelet transform when the position and the scale changes in discrete stages.

4.2.1 Discrete Wavelet Transforms

Discrete Wavelet Transform (DWT) is a procedure applied for decomposing multi-resolution analysis, pyramidal coding, sub-band coding; speech signal coding, and discrete-time signals **(Kisi and Shiri 2011[139])**. Thus with the help of digital filtering techniques, a time-scale representation of the digital signal is obtained and at various scales with different limits of frequencies, a signal is passed for examination **(Maheswaran and Khosa 2012[151])**. It breaks down a signal into various groups (vectors) of coefficient and different coefficients vectors containing information about characteristics of the sequence at different scales **(Saadaoui and Rabbouch 2019[238])**.

Wavelet analysis employs a prototype function called mother wavelet $f(t)$, which has a null mean, drops in an oscillatory way, and data are displayed via superposition of scaled and translated versions of the pre-specified mother wavelet. It can be calculated as **(Parmar and Bhardwaj 2013[201, 203])**,

$$DWT(m,t) = \frac{1}{\sqrt{a^m}} \sum_n x(n) f\left(\frac{t - nba^m}{a^m}\right) \quad (4.5)$$

Where a and b are scaling and translation function of integer variable m; t is an integer variable which refers to a point of the input signal; n is the discrete-time index; $x(t)$ is a given signal and $f(t)$ is the mother wavelet. Mallat-tree decomposition is used for calculating the discrete wavelet transform by consecutive high and low pass filtering of the time domain signal of discrete nature **(Baghanam et al. 2019[28])**.

The importance of DWT is that it associates the discrete-time filters with the continuous-time multi-resolution. Till we achieve the preferred level of accuracy the decimation and filtering process is carried on, the measurement of the signal is determined by the optimal number of levels.

4.2.2 Continuous Wavelet Transforms

The Continuous Wavelet Transform (CWT) of a function $x(t)$ w.r.t. to a function $\psi(t)$ is provided by equation

$$W_{(a,b)} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(t - \frac{b}{a}\right) dt \quad (4.6)$$

Where $\psi(t)$ is the mother wavelet, $x(t)$ is the signal to be approximated, and a, b are real's where a gives the amount of time scaling and b is the dilution variable. All the wavelet functions used in the transformation are derived from the mother wavelet through translation and scaling (compression or dilation). The mother wavelet used to generate all the basis functions is designed based on some desired characteristics associated with that

function. The translation parameter τ relates to the location of the wavelet function as it is transferred through the signal. Thus, it corresponds to the time information in the Wavelet Transform **(Andreo et al. 2006[22])**. The scale parameter *s* is defined as |1/frequency| and corresponds to frequency information. Scaling either compresses or dilates (expands) the signal. Small scales (high frequencies) compress the signal and provide global information about it while large scales (low frequencies) dilate the signal and give detailed information hidden in the signal **(Parmar and Bhardwaj 2012[198])**.

4.3 Wavelet-ARIMA Coupled Approach

It is difficult to obtain accurate time-series forecasts of time series dealing with non-stationarity, due to the presence of trend and noise in the data. Thus, in order to reduce the noise from the data we use a wavelet-ARIMA model coupled approach over our time series data to get a better forecast with minimum forecasting errors. To use the wavelet-ARIMA forecast we first decide the number of decompositions required, by using, **(Nury et al. 2017[190], Soni et al. 2017[264])**

$$L = Int\left[\log N\right] \tag{4.7}$$

Where the numbers of decompositions are denoted by *L* and *N* denotes the length of the time series. The following steps are carried out in this process of modeling **(Kumar et al. 2015[141], Wang et al. 2018[288])**,

a) **Wavelet Decomposition-** In wavelet analysis a signal is decomposed using wavelet transform into low and high frequencies. After determining the decomposition layer and the scale function, we decompose the original signal $\{X_t\}$ into the detailed coefficients $D_1, D_2, D_3, ..., D_k$ along with an approximation coefficient A_k such that the time-series coefficient X_t can be expressed as, $X_t = D_1 + D_2 + D_3 + ... + D_k + A_k$

b) **Suitable Model Construction-** After wavelet decomposition, a suitable time series model is constructed for the decomposition and approximate coefficients.

Depending on the stationarity of the coefficients we use the ARMA model for stationary series and if the coefficients display non-stationarity, we use the differencing techniques and choose a suitable ARIMA model. Thus, using these implemented results and sample quantities to further obtain the forecasts denoted as, $\bar{D}_j, 1 \le j \le k$ and $\bar{A}_k$.

c) **Accumulation and Reconstruction of Signal-** After decomposition and applying the suitable model on the decomposed and approximation coefficients we find $\bar{D}_k$ and $\bar{A}_k$ these obtained values are further combined and reconstructed to obtain $\left\{\bar{X}_t\right\}$ as $\bar{X}_t = \bar{D}_1 + \bar{D}_2 + \bar{D}_3 + ... + \bar{D}_k + \bar{A}_k$ which can then be used for ARIMA forecasting and obtain quite reliable results.

4.4 Results and Discussion

The wavelet decomposition technique is used for obtaining detailed and approximate coefficients for various levels of resolution. MATLAB software is used for the purpose of wavelet decomposition and choosing the appropriate model after processing the data through the testing and training phase. In this study for Raniganj, Jharia, and Bokaro sample sites, we have used (Daubechies wavelets) db8 level 3 for decomposing and reconstructing the wavelet signal. The detailed and approximate coefficients are shown in Figure 4.1, 4.2, and 4.3 below, where *s* represents the value of original time series data, a_3 represents the approximation, and d_1, d_2, d_3 represents the detailed parts of decomposition such that, $s = d_1 + d_2 + d_3 + a_3$. These decomposed parts are further treated with time series ARMA and ARIMA models to obtain the predictive results which are later summed up for attaining the final results for the wavelet-time series-coupled approach **(Rahman and Hasan 2014[220])**.

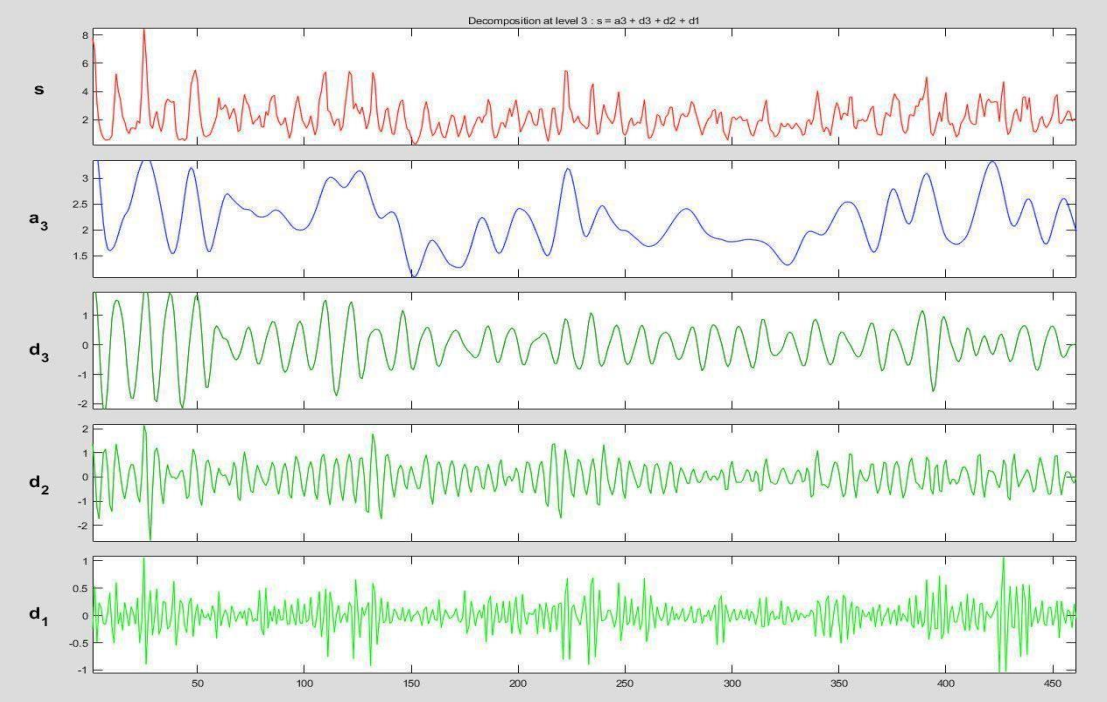


Figure 4.1 Wavelet decomposition of time-series data for Raniganj

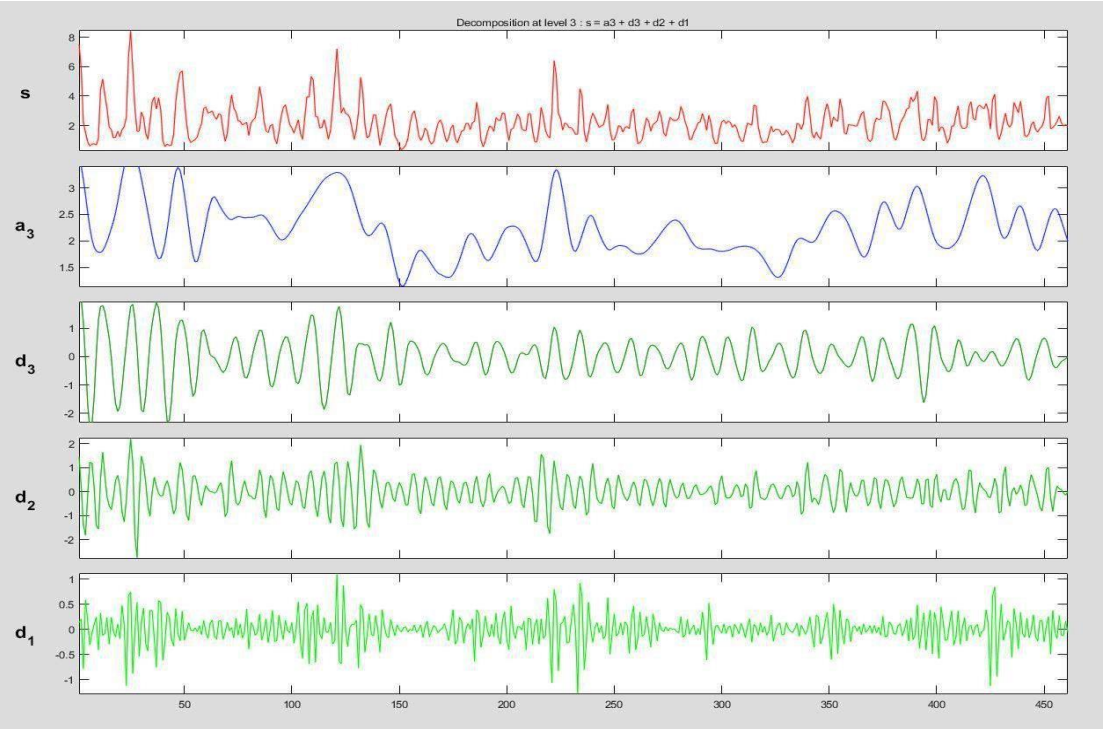


Figure 4.2 Wavelet decomposition of time-series data for Jharia

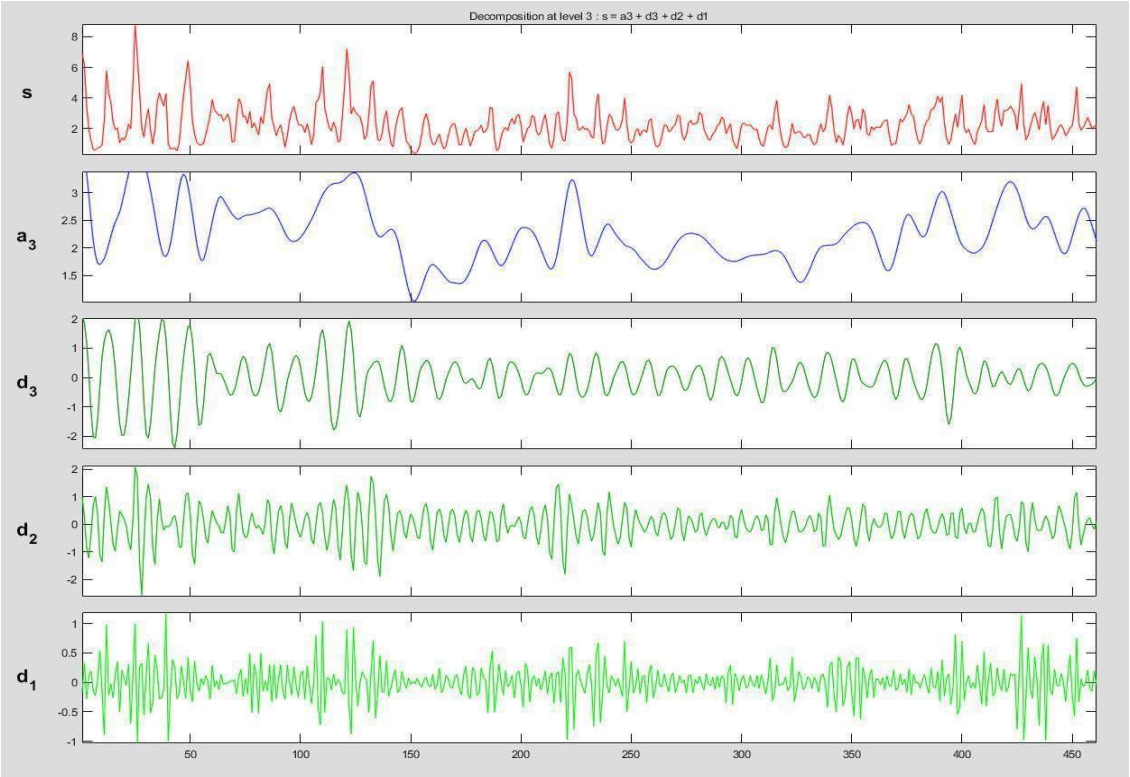


Figure 4.3 Wavelet decomposition of time-series data for Bokaro

4.4.1 ARIMA and Wavelet-ARIMA Model Fitting and Forecasting

For constructing and selecting the best model with minimum forecasting errors and accurate forecasted values we have divided the time series data set into two parts of 33 years, 5 months (from Jan 1980 to May 2013) for training and 5 years data (from Jun 2013 to May 2018) for the testing phase so as choose an appropriate model for the data. The validity of the results is tested with the help of AIC, BIC, error measures like RMSE, MSE, MAPE, and MAE.

Raniganj (23° 40' N 87° 05' E): Time series ARIMA (5, 0, 8) model fitted very well to the time series data of Raniganj (training phase) and was used to obtain the time series forecast (testing phase) which on comparison with the original values gave the minimum error measures of RMSE, MSE, MAPE, and MAE observed as 0.850067, 0.722614, 0.315779 and 0.678623 respectively. After this Daubechies wavelets db8, level 3 was

used for decomposing and reconstructing the wavelet signal as shown in Figure 4.1. Here for the approximation part A3, ARIMA (2, 0, 8) model was used while for the decomposition parts D1, D2 and D3 the best fitted were the ARIMA (2, 0, 1), ARIMA (2, 0, 1) and ARIMA (5, 0, 3) respectively. The signal was then reconstructed by summing the values of D1, D2, D3, and A3. The forecasted values of 5 years of data when compared with the original values gave the values of RMSE, MSE, MAPE, and MAE as 0.82512, 0.68082, 0.31332, and 0.64809. Diagrammatic representation and comparison of this are shown in Figure 4.4 and 4.5. Thus, the forecasts obtained by Wavelet-ARIMA coupled approach gave better results than the time series ARIMA model.

Jharia (23° 50' N 86° 33' E): Time series ARIMA (10, 0, 7) model fitted very well to the time series data of Jharia (training phase) and was used to obtain the time series forecast (testing phase) which on comparison with the original values gave the minimum error measures of RMSE, MSE, MAPE, and MAE observed as 0.78964, 0.62353, 0.25946 and 0.62224 respectively. After this Daubechies wavelets db8, level 3 was used for decomposing and reconstructing the wavelet signal as shown in Figure 4.2. Here for the approximation part A3, ARIMA (2, 0, 9) model was used while for the decomposition parts D1, D2 and D3 the best fitted were the ARIMA (2, 0, 1), ARIMA (2, 0, 1) and ARIMA (2, 0, 3) respectively. The signal was then reconstructed by summing the values of D1, D2, D3, and A3. The forecasted values of 5 years of data when compared with the original values gave the values of RMSE, MSE, MAPE, and MAE as 0.75688, 0.57286, 0.29361, and 0.62480. Diagrammatic representation and comparison of this are shown in Figure 4.6 and 4.7. Thus, the forecasts obtained by Wavelet-ARIMA coupled approach gave better results than the time series ARIMA model.

Bokaro (23° 46' N 85° 55' E): Time series ARIMA (9, 0, 1) model fitted very well to the time series data of Bokaro (training phase) and was used to obtain the time series forecast (testing phase) which on comparison with the original values gave the minimum error measures of RMSE, MSE, MAPE and MAE observed as 0.82260, 0.67667, 0.24484 and 0.59600 respectively. After this Daubechies wavelets db8, level 3 was used for

decomposing and reconstructing the wavelet signal as shown in Figure 4.3. Here for the approximation part A3, ARIMA (2, 0, 8) model was used while for the decomposition parts D1, D2 and D3 the best fitted were the ARIMA (2, 0, 1), ARIMA (2, 0, 1) and ARIMA (2, 0, 5) respectively. The signal was then reconstructed by summing the values of D1, D2, D3, and A3. The forecasted values of 5 years of data when compared with the original values gave the values of RMSE, MSE, MAPE, and MAE as 0.58985, 0.34793, 0.19950, and 0.48409. Diagrammatic representation and comparison of this are shown in Figure 4.8 and 4.9. Thus the forecasts obtained by Wavelet-ARIMA coupled approach gave better results than the time series ARIMA model.

The results obtained after applying the time-series ARIMA and wavelet-ARIMA coupled approach have been shown in Table 4.1. Thus it observed that the wavelet time series-coupled model performed better than the time series ARIMA model and hence can be used as an effective model for time series forecasting, with minimum forecasting errors.

Table 4.1 Comparison of ARIMA and Wavelet-ARIMA coupled models

Sample sites	**Raniganj (23° 40' N 87° 05' E)**		**Jharia (23° 50' N 86° 33' E)**		**Bokaro (23° 46' N 85° 55' E)**	
Model	ARIMA (5,0,8) model	Wavelet-ARIMA coupled model	ARIMA (10,0,7) model	Wavelet-ARIMA coupled model	ARIMA (9,0,1) model	Wavelet-ARIMA coupled model
RMSE	0.850067	**0.82512**	0.78964	**0.75688**	0.82260	**0.58985**
MSE	0.722614	**0.68082**	0.62353	**0.57286**	0.67667	**0.34793**
MAPE	0.315779	**0.31332**	0.25946	**0.25361**	0.24484	**0.19950**
MAE	0.678623	**0.64809**	0.62224	**0.61480**	0.59600	**0.48409**

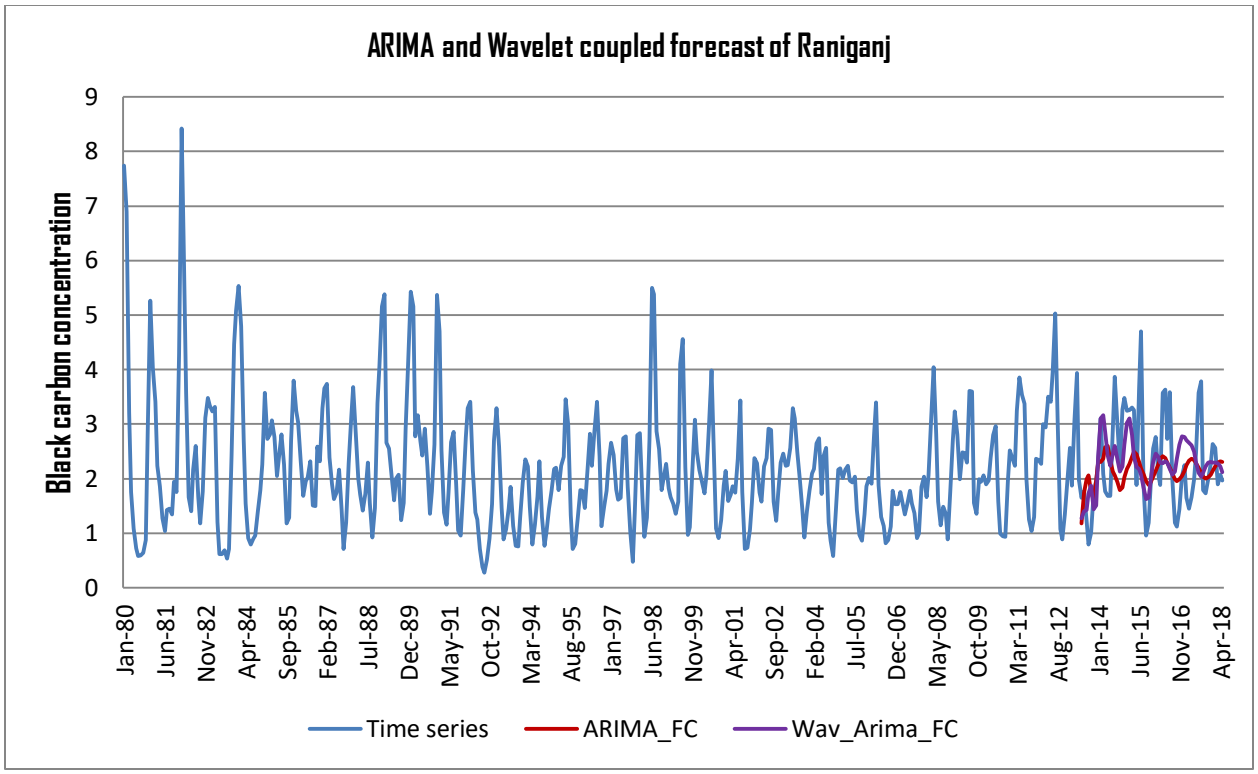


Figure 4.4 ARIMA and Wavelet-ARIMA forecast of time series data for Raniganj

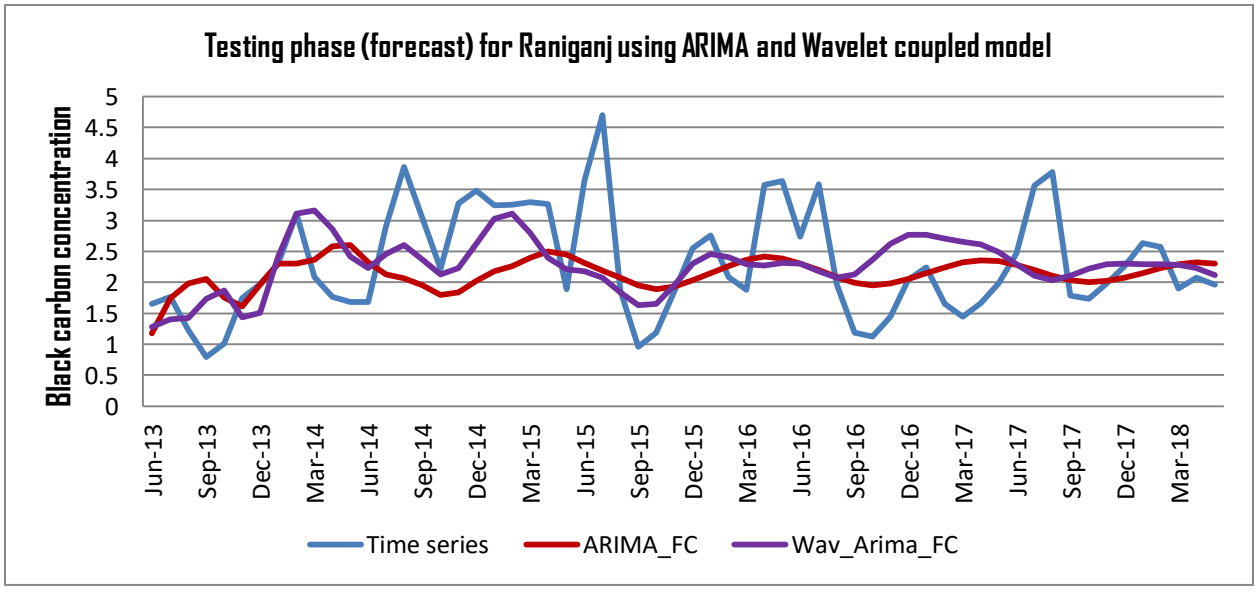


Figure 4.5 Testing phase (forecast) for Raniganj using ARIMA and Wavelet coupled model

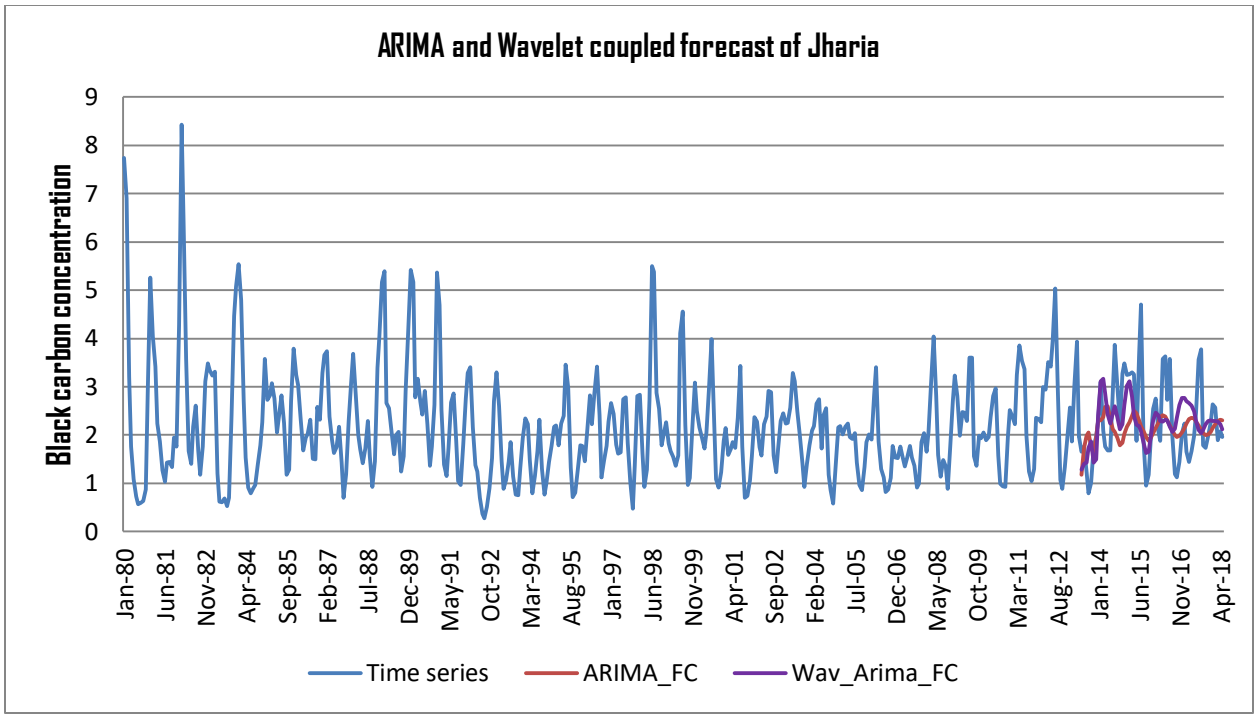


Figure 4.6 ARIMA and Wavelet-ARIMA forecast of time series data for Jharia

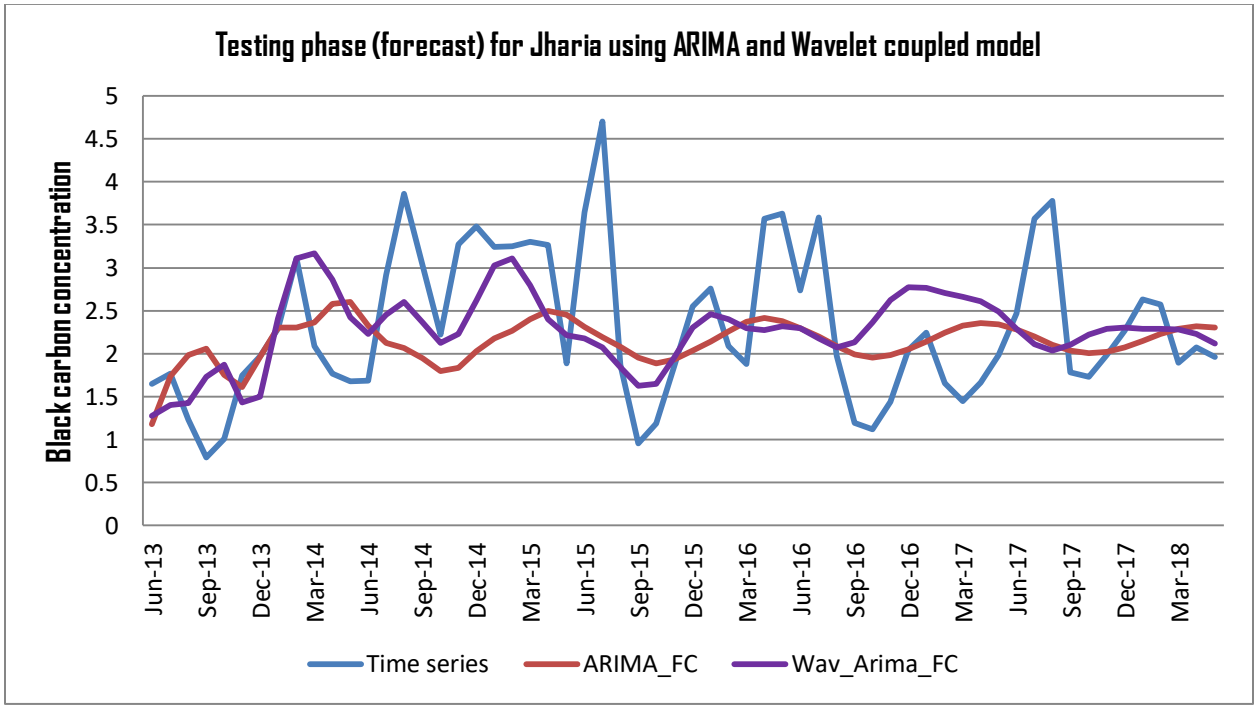


Figure 4.7 Testing phase (forecast) for Jharia using ARIMA and Wavelet coupled model

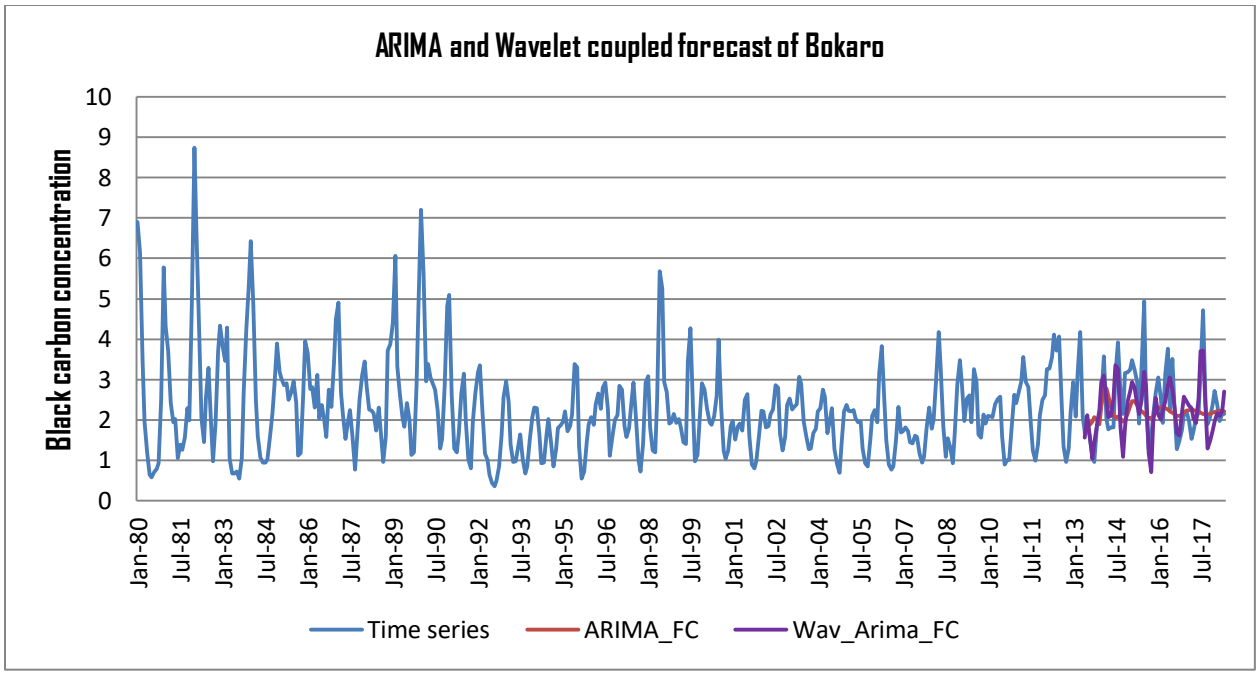


Figure 4.8 ARIMA and Wavelet-ARIMA forecast of time series data for Bokaro

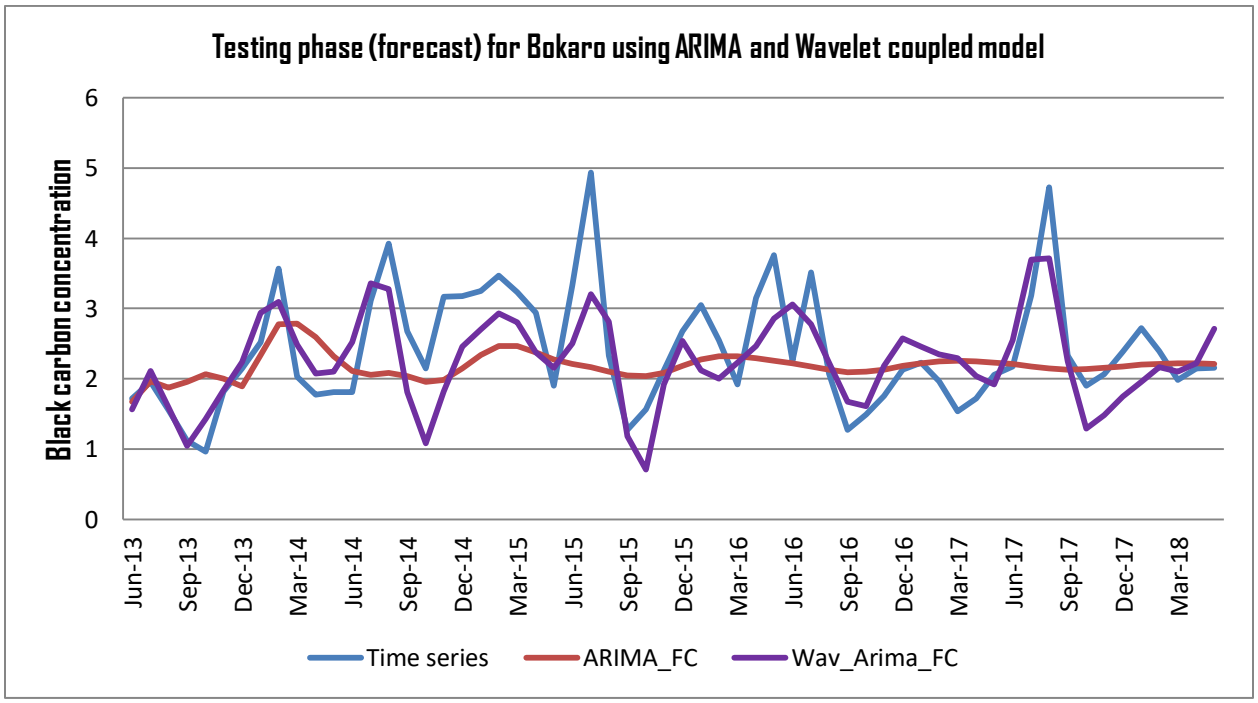


Figure 4.9 Testing phase (forecast) for Bokaro using ARIMA and Wavelet coupled model

4.5 Conclusion

Considering the long term time series data of black carbon concentration over the major coal mines of India viz. Raniganj, Jharia, and Bokaro time-series ARIMA and wavelet time-series coupled approach is developed and used for obtaining the forecasts. As seen in Table 4.1, the results of the forecast obtained by the Wavelet-ARIMA coupled approach are better than the simple time-series ARIMA model. Diagrammatic representation and comparison of the original time series along with forecasted values are obtained by time series ARIMA and wavelet-ARIMA coupled model as shown in Figure 4.4, 4.6, and 4.8. Thus the wavelet time series ARIMA coupled approach can be effectively used for obtaining the forecast of time series dealing with pollutant concentration over the coal mine regions. This study is very beneficial for NGO's as well as government agencies, so as to frame new policies and preventive measures to curtail and keep a check on growing pollutant concentrations, which is the major problem faced by all the developing countries.

Chapter 5 Soft Computing Analysis of Black Carbon

Soft computing means designing computing techniques with the help of computer-based programs for achieving optimization. **Zadeh 1981[296]** introduced the concept of soft computing based on human brain functioning. He initiated that soft computing is the combination of Probabilistic computing, Genetic computing, Fuzzy logic, Evolutionary, and Neuro-computing. It is used for modeling a variety of real-life problems of complex nature in business and industry having noisy data of unknown patterns with no future information which are difficult to handle by the conventional methods **(Rojas et al. 2008[234])**. Soft computing is also used in intelligent systems by machine learning to include in it the features like learning, thinking, and adaptation along with examining, and improving. These systems are developed to have functioned like the human brain with efficiency to learn and improve with the environment. Thus with the help of soft computing techniques, it is aimed to develop machine learning based on mathematical programs or models which are cost-efficient, easy to adapt, close to real predictions with least error **(Pentos et al. 2020[211])**. The prominent characteristics of machine learning are:

- Developing an adaptable algorithm which can adjust with the dynamic nature
- Different methods can be used to perform a variety of tasks
- A simple algorithm can perform better than a vast mathematical model
- It is time-saving and cost-efficient
- It can efficiently handle non-linear real-life data

In this chapter, with the conjugation of ANN, wavelets, and fuzzy-interface system (FIS), a new mathematical model called the Artificial Neural Network Fuzzy Wavelet Conjugation model (ANNFWCM) is constructed and used in the forecasting of black carbon concentration in the coal mine regions of India.

5.1 Literature Review

Jang 1993[123] presented the ANFIS method, which is a combination of a fuzzy inference system along with ANN. ANFIS architecture was used to obtain prediction in case of

nonlinear time series, it combines the stipulated input-output data pairs along with the input-output mapping based on human knowledge (in the form of fuzzy if-then rules). For recognizing the parameters and structure of three-layer feed-forward ANNs, a new method of linear least square simplex (LLSSIM) was discovered by **Hsu et al. 1995[113],** presenting the future behavior of nonlinear hydrologic watersheds. It was shown that a very fine depiction of the rainfall-runoff relationship was provided by the nonlinear ANN model approach than the time series linear ARMAX (autoregressive moving average with exogenous inputs) model. For the prediction of wind speed in Jeddah, Saudi Arabia **Mohandes et al. 1998[171]** used a neural network and showed that it outperformed the autoregressive model based on RMSE and prediction graph.

Abraham and Nath 2001[6] used artificial neural network (ANN) and an evolving fuzzy neural network (EFuNN) trained by backpropagation (BP) algorithm and scaled conjugate gradient algorithm (CGA) for forecasting demand of electricity in Australia along with the Box–Jenkins autoregressive integrated moving average (ARIMA) model, results depict that the neuro-fuzzy system led to better forecasts. **Maqsood et al. 2005[158]** used radial basis function network (RBFN) model to examine its applicability in weather analysis for daily 24 hr prediction of weather in southern Saskatchewan, Canada as it performed better than the Hopfield model (HFM), Elman recurrent neural network (ERNN) and multi-layered perceptron (MLP) network to obtain accurate forecasts.Considering data for nitrogen dioxide emission and dispersion over the Delhi city, **Nagendra and Khare 2006[177]** used ANN-based models to examine two air-quality control regions (AQCRs).

For the prediction of rainfall and management of floods in Bangkok, Thailand, **Hung et al. 2009[119]** used a generalized feed-forward ANN model using hyperbolic tangent transfer function and obtained a good forecast for rainfall showing that ANN forecasts had superiority over the others models. A quantitative and qualitative monthly precipitation forecast using ANFIS and the forward selection method was made by **Jeong et al. 2012[127]**, this rainfall prediction helped in determining future month's rainfall condition, results obtained were compared with those obtained from some weather

agency of Korea. To make the ANFIS method more realistic Modified ANFlS (MANFlS) structure was discovered by **Akrami et al. 2013**[15] with regards to the computing epoch, parameters like Signal to Noise Ratio (SNR), Root Mean Absolute Error (RMAE), Correlation Coefficient (R2) and Root Mean Square Error (RMSE) were used for modeling nonlinear systems. It was observed that the Modified ANFIS gave better and more realistic rainfall prediction with small errors and simplicity in calculations than the conventional ANFIS model.

For predicting the flow of two Indian rivers namely the Kosi and Gandak which are unsafe during monsoon a wavelet transform-genetic algorithm-neural network model (WAGANN) was developed by **Sahay and Srivastava 2014**[240], on comparison WAGANN models were discovered to be superior to the autoregressive models (ARs) and GA-optimized ANN models (GANN). **Nourani et al. 2014**[187] presented a review on the use of hybrid modeling, the convenience of combined models, are in future use of these models in hydrology for forecasting many key mechanisms of the hydrological cycle and highlighting the importance of the capability of AI methods in forecasting. Optimized Taguchi method was used by **Moosavi et al. 2014**[174] for assessing distinct factors influencing the working of Wavelet-ANN and Wavelet-ANFIS hybrid models at several stages and at last, it was concluded that the Wavelet-ANFIS model outperformed the best fitted Wavelet-ANN model. For the management of water resources and prediction of chemical oxygen demand (COD), a water quality parameter of Yamuna River near Nizamuddin station, **Parmar and Bhardwaj 2015**[199] developed a coupled model using wavelets, fuzzy and neural networks model for making a monthly forecast for the river water and compared the results with other data-driven models.

In forecasting SO_2 concentration at Janakpuri, Nizamuddin, and Shahzadabad, located in Delhi, India, **Kisi et al. 2017**[137] used the least square support vector machine (LSSVM), multivariate adaptive regression splines (MARS) and M5 Tree model to the monthly data for a better forecast. For determining the accuracy in forecasting the monthly stream-flow of Gilgit river basin **Muhammad et al. 2019**[176] used five soft computing methods namely the adaptive neuro-fuzzy inference system with subtractive

clustering (ANFIS-SC), adaptive neuro-fuzzy inference system with grid partition (ANFIS-GP), generalized regression neural network (GRNN), radial basis neural network (RBNN) and feed-forward neural network (FFNN) along with seasonal autoregressive integrated moving average (SARIMA) considering the interaction between the streamflow and temperature, ANFIS-SC, and RBNN gave more accurate results than the others. To predict and model the outbreak of COVID-19 **Ardabili et al. 2020**[25] used soft computing and machine learning technique namely the ANFIS and multi-layered perceptron (MLP) to discuss the situation and predictions in five nations, the study indicated that for modeling the outbreak soft computing could be used as an effective tool.

5.2 Soft Computing

Soft computing is a group of computing techniques like Artificial neural network (ANN), Fuzzy Logic, Probabilistic Reasoning, Evolutionary Computation, or a combination of these as shown in Figure 5.1 below for obtaining efficient and realistic solutions with simplicity **(Rao and Raju 2011**[226]**)**.

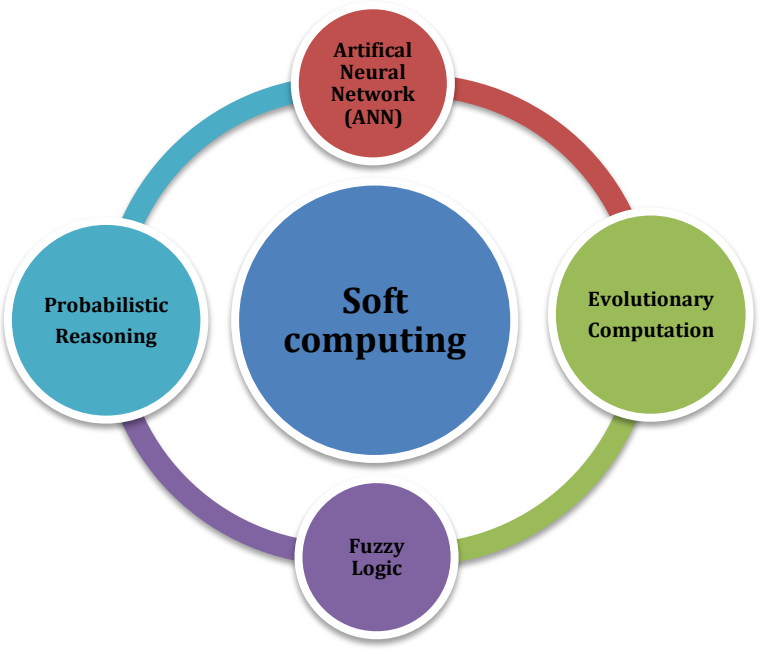


Figure 5.1 Main Components of Soft Computing

The models formed by the combination of these are called the hybrid models for example; the combination of the ANN and the fuzzy interface system is called the Adaptive Neuro-Fuzzy Inference System (ANFIS) model.

5.2.1 Evolutionary Computation – These algorithms are stimulated by the natural evolution which uses the trial and error technique for optimization. In this, the initial values are used in the iterative process to get more exact values. This was initially called a genetic algorithm, which includes the applications and fundamentals of computational methods. These are optimization and examining techniques that are capable of obtaining the optimal solution for the problem. It involves the use of computation for solving multi-dimensional problems of practical application and simulations. Some of these methods stimulated by Evolutionary algorithms for modeling biological structures are namely recombination, survival, natural selection, reproduction, mutation, etc. **(Ardabili et al. 2020[25])**.

5.2.2 Probabilistic Reasoning – It is a technique of demonstrating information where the idea of probability is applied to specify the ambiguity in the information. In order to exploit the structure with deductive logic to handle uncertainty and combine it with the probability theory, probabilistic reasoning is used. The concept of fuzziness and fuzzy reasoning are used considering the uncertainty. It aims at finding deriving results with the help of probabilistic expressions which is an extension of the conventional logic truth tables. The limitation of this is that it combines the computational complications of their logical and probabilistic mechanisms. A common example of this is the Bayesian approach; it is also used in cases of artificial intelligence where the outputs are errorness, a large number of predicates, and uncertainty in predicates **(Zhang et al. 2011[301])**.

5.2.3 Artificial Neural Networks (ANN) – ANN are computational models or algorithms based on machine learning which were developed based on the functioning of the human brain and biological neurons which are used to approximate functions, it consists of three layers namely input, hidden and output

layer. Each of these layers consists of a number of nodes and neurons with weights assigned to them which are capable of performing simple operations to compute its output from its input **(Relvas and Miranda 2018[232])**. ANNs are adaptive systems that are used in modeling of the constantly changing population, segmentation, classification **(Mao and Aggarwal 2001[157])**. ANN has been widely used in real-life problems such as forecasting precipitation, droughts, stream-flow, groundwater level, pollutant concentration, etc. **(Kisi 2009[140], Adamowski and Sun 2010[8], Zhang et al. 2011[301], Sachindra et al. 2013[239], Siddiquee and Hossain 2015[254], Ribeiro et al. 2017[233], Ventura et al. 2019[286])**.

5.2.4 Fuzzy Logic – It is used to obtain an intermediate solution between two sets of a possible solution is expected to be obtained. It is a generalization of Boolean logic and involves the probability theory although being different from it. Fuzzy logic deals with situations between truths and false when there occur a partial truth. In the Fuzzy inference system (FIS), based on the fuzzy rules input variables are mapped onto the output variables, such rules are established by the principle of modeling. It involves the use of human reasoning to deal with different facts based situations. It is widely used in the field of science, business, finance, technology, and modeling for controlling the input and output variables **(Toprak et al. 2009[278])**.

5.3 Machine Learning Algorithms

Machine learning originated from artificial intelligence and later on was classified into neural networks. These are programs based on past observations that are trained to improve their efficiency. The learning here aims to develop an algorithm by training and experience which can predict the output values based on the past data values as a function that maps the input to the output **(White 1989[289])**. These are mostly used in problems where it is difficult to obtain predictions using conventional algorithms like computer

vision and email filtering **(Ardabili et al. 2020[25])**. There are four main types of machine learning algorithms as discussed below **(Dey 2016[80])**:

5.3.1 Supervised Learning – These operations are performed under direct supervision based on a set algorithm, for operating the algorithms strict boundaries are fixed and sample data values are labeled. In this, we select a function after training an algorithm that best fits the input data and can be used to predict the output i.e. in this we try to model the relationship between the input and target to obtain the desired output. The training process is carried out and practiced until we achieve the desired level of accuracy for the output. The main aim of this is to obtain a forecast of the unseen or unavailable data based on the available labeled data. Regression and classification are the two main processes of supervised learning, in classification process labeling of incoming data values are done based on the previous data values and then the algorithms are manually trained for recognizing certain kinds of objects and classifying them while in the regression process pattern are recognized and forecasts are obtained. Commonly used supervised learning algorithms include Neural networks, Support vector machines (SVM), Logistical regression, Linear regression, Nearest neighbor, Naive Bayes, Decision trees, Gradient boosted trees and Random forest which are further used for stock trading, retail commerce, trend and price forecasting **(Kisi and Parmar 2016[138])**.

5.3.2 Unsupervised Learning – This does not involve direct supervision as the results to be predicted are unknown. Another major difference between this from supervised learning is that it uses unlabeled data values. Dimensionality reduction and Clustering are the main steps of unsupervised learning algorithms. In dimensionality reduction to remove noise from the signal, we refine the information by using Machine learning algorithms. While in clustering we develop clusters of the data based on their internal structure to study the data to split it into meaningful sets based on their internal patterns without any previous

information. The most common types of unsupervised learning algorithms are Association rule, PCA (Principal Component Analysis), t-Distributed Stochastic Neighbor embedding, k-means clustering. It is mainly used in increasing efficiency, detecting patterns, extracting valuable insights, exploring the structure of signals, provide service to the customer according to the available information. Its application can further be extended for the identification of patterns for better results and efficient target achievement.

5.3.3 Semi-supervised Learning – It is formed by the combination of both supervised and unsupervised learning. For grouping the data it uses the clustering process while for identification of the data it uses the classification process. In this only a few data samples are labeled, this leads to a partially trained data set that performs the task of labeling the unlabeled data leading into pseudo-labeled data. Thus a different algorithm consisting of labeled and pseudo-labeled data is formed. These are often used in cases where the training labels are imprecise, limited, and noisy since it can be performed over unlabeled data so they are cheaper leading to bigger and efficient training sets with accuracy in learning. It is used in speech and image reorganization, content aggregation systems and crawling engines, Healthcare, and legal industries where labels are absent.

5.3.4 Reinforcement Learning – After interacting with the incoming data and training the labeled data a self-sustained system is formed which improves its efficiency by using the exploration technique. In this, the machine learning algorithm based on the present situation the agent decides the future steps by studying its nature to get optimal rewards. On performing specific tasks based on certain algorithms there exist reward signals which act as navigation tools. There occur two main types of reward signals called the positive and the negative reward signal. In a positive reward signal, a particular sequence of action is inspired while in a negative reward signal it is penalized until the algorithm becomes correct. Thus the system tries to minimize the negative rewards while it tries to maximize the positive rewards. Reinforcement machine learning aims at collecting information

from the environment during training for performing actions according to the rewards and leaning based on past experiences. This is also known as Machine learning artificial intelligence. The commonly used machine learning algorithm includes Asynchronous Actor-Critic Agents, Deep Adversarial Networks, Monte-Carlo Tree Search, Temporal Difference, and Q-Learning. These are used in chat boxes for dialogue generation, computer games, natural language processing; detection of collision, Modern NPCs, driverless cars, technical and marketing operations.

5.4 Basic Concepts of Neural Networks – A few basic concepts of the neural network like the perceptron, activation function, Sigmoid or Logistic activation function, backpropagation that helps in the functioning of a neural network are discussed below:

5.4.1 Perceptron - The output value is obtained by a linear combination of the inputs of nodes multiplied with the weights. At each i^{th} neuron, the incoming signals $x_1, x_2, x_3, ..., x_n$ are summed with the weights $w_1, w_2, w_3, ..., w_n$ such as $\sum_{j=1}^{n} w_{ij} x_j$ followed by loading this with an activation function f to obtain the output value $x_i = f\left(\sum_{j=1}^{n} w_{ij} x_j\right)$ here n denotes the number of neurons in the previous layer **(Abghari et al. 2012[4])**. Each neuron calculates the current value x_i by the above rule which can be diagrammatically represented as,

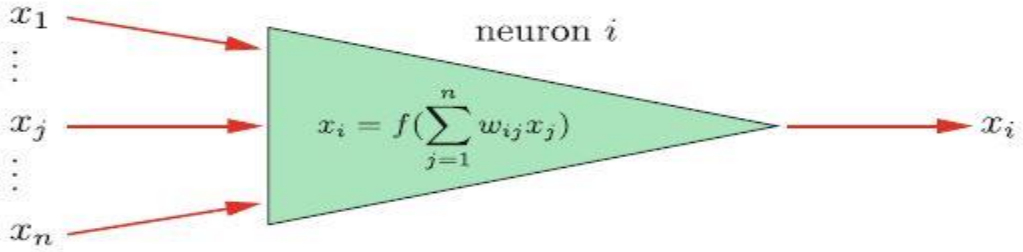


Figure 5.2 A simple perceptron

5.4.2 Activation Function - A function that can determine the output from the node is termed as an activation function. The range of activation function is from 0 to 1 or -1 to 1. Depending on the nature of the function, these are classified as linear activation and non-linear activation function.

5.4.3 Sigmoid or Logistic Activation Function – It is a differentiable and bounded function defined over all real input values. It is a strictly increasing monotonic function having an S-shaped curve; a common example of this is the logistic function which is mathematically written as, $f(x) = \frac{1}{1+e^{-x}}$ with the domain of Real numbers. It is the most commonly used activation function. As the output of this function lies between 0 and 1 which is similar to the range of the probability prediction of many models. The output value at the nodes maintains the order relation; some other examples of the sigmoid function are the Hyperbolic tangent function, Error function, generalized logistic function, Gudermannian function, Smooth-step function and the Arctangent function **(Bhardwaj et al. 2020[38])**.

5.4.4 Backpropagation – For supervised learning, it is one of the extensively used algorithms for training feed-forward neural networks. **Rumelhart et al. 1986[237]** were the first to apply and publicize the use of backpropagation in neural networks. It was observed that backpropagation was capable of solving various complex nature neural networks more efficiently than the other approaches to learning. Backpropagation aims at training an MLP neural network and such that it can learn from the suitable interior demonstrations to permit it to learn from any arbitrary mapping input to output **(Kim and Singh 2015[136])**. It gives a clear explanation of how the change in biases and weights can lead to a complete change in the behavior of the ANN.

Backpropagation is an algorithm that is used for calculating with respect to the weights, the gradient of the loss function, and not on discovering its application. The gradient calculation in the delta rule is generalized by

backpropagation it originates from the incremental delta rule. In this, the approximation error is calculated after determining the output by forwarding propagation for training and based on the results these the weights are altered backward from layer to layer **(Mao and Aggarwal 2001[157])**. This entire process is repeated until there is no change in the output value. It may have more than two neuron layers of a network for adjusting the weights at the nodes; the sigmoid function is used as an activation function over the sum of the weighted inputs **(Kavzoglu and Mather 2003[132])**.

5.5 Artificial Neural Network (ANN) - This was first introduced by **Warren McCulloch and Walter Pitts 1943[160]**; they developed an algorithm based computational model for neural networks called threshold logic. **Frank Rosenblatt 1958[235]** proposed a single layer perceptron; which was planned to model human brain functioning to recognize objects and process visual data. The learning capabilities and pattern-matching properties of ANN endorsed them to be used for many problems which otherwise would have been very difficult or impossible to solve by other statistical methods and standard computational techniques **(Kavzoglu and Mather 2003[132])**. It was then decided that in addition to getting perceptions of the functionality of the human brain, ANNs can further be used as an independent tool. Thus till the 1980s, ANNs was started being used for various purposes, ANN works by establishing a connection between different nodes called neurons arranged in different layers interconnected with each other and assigned with randomized weights capable of solving simple mathematical calculations **(Luongvinh and Kwon 2005[149])**.

These nodes identify the relation between the targeted values and the input data sets **(Faruk 2010[93])**. Input signals are received by these neurons organized in different layers which in turn convert them into a single output. This output is further transferred to the other layer where the same process is carried again until finally, we get a single output **(Moosavi et al. 2013[173])**.

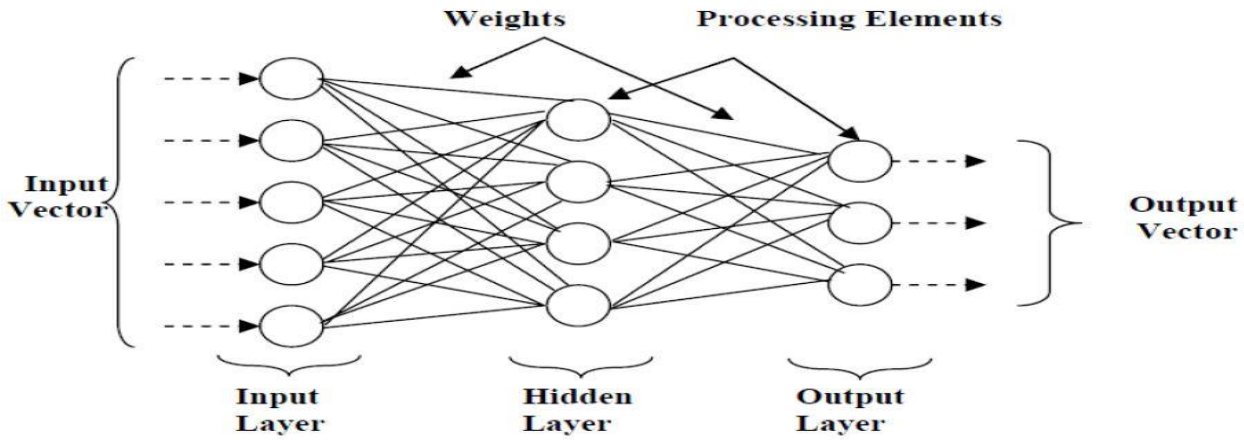


Figure 5.3 A three-layered Artificial Neural Network

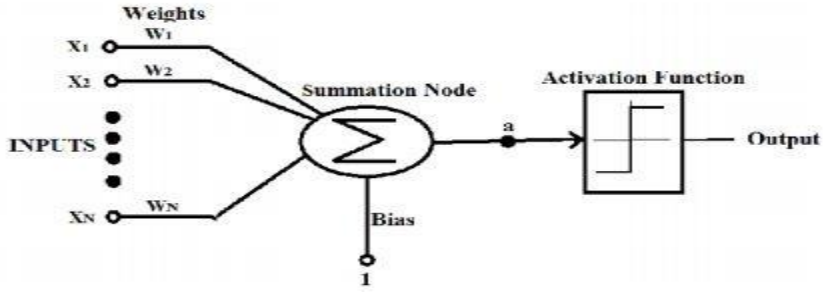


Figure 5.4 Working principle of neural network

In an ANN signals pass through the nodes at which processing of information occurs, these nodes are arranged in a sequential pattern interconnected with each other in different layers and each having weights assigned to them **(Rojas et al. 2008[234])**. Each neuron consists of an activation function that is used to determine the output signal of the sum of weighted input signals. ANNs develop and identify the relationship between the data values by recognizing the pattern between them which otherwise is difficult to predict **(Yeon et al. 2009[295])**. They are universal function, approximators, self-adaptive, and data-driven algorithms which learn and train themselves based on past capabilities. ANNs have been used by researchers in various fields like mathematical forecasting, science, finance, industry, business **(Adamowski and Chan 2011[7], Doucoure et al. 2016[87], Kim et al. 2016[135])**.

5.6 Fuzzy Inference System (FIS) - Fuzzy Inference System is based on the set of fuzzy reasoning and fuzzy rules which is used for modeling uncertainty **(Karmakar and Mujumdar 2006[130])**. The fuzzy rules act as an important tool in making conclusions about the fuzziness and uncertainty in the results of the problems dealing with control and modeling **(Chen and Chang 2010[61])**. The logic in which fuzzy input values are mapped to crisp outcomes is termed as fuzzy logic. Everything existing in nature is of approximate behavior this highlights the importance of fuzzy logic. In the absence of precise or complete information, fuzzy logic acts as an initiator. Fuzzy logic and fuzzy set theory are used in handling vague, imprecise, indeterministic, and incomplete information of the storage and database in fuzzy systems **(Zhao et al. 2020[303])**. The fuzzy set theory aims at finding solutions and building methodology for problems that are complex in nature and difficult to deal with conventional techniques **(Toprak et al. 2009[278])**. It has application in various real-life problems dealing with classification of data, recognition of pattern, and expert systems for modeling uncertain patterns like operation research, decision making, clustering, process control, and retrieval of information **(Zeinalnezhad et al. 2020[297])**.

5.7 Adaptive Neuro-Fuzzy Inference System (ANFIS) Architecture: - ANFIS on the other side is formed by the combination of neural networking (ANNs) and fuzzy interface system (FIS) based on the fuzzy if-then rule thus it exhibits the benefits of both the proficiencies in an integrated structure **(Nayak et al. 2004[182])**. The fuzzy interface system has been extensively used for obtaining forecasts of time series, leading to more accurate and reliable predictions **(Poul et al. 2019[214])**. In this input, characteristic functions are mapped into input membership functions by it which is related to the output properties which further have a relation with the output membership function resulting in a single decision or output **(Adedeji et al. 2020[11])**. The membership functions in ANFIS are different from the conventional fuzzy systems and can be multidimensional and nonlinear.

In ANFIS appropriate membership functions (MFs) are used along with the fuzzy if-then rules to build a learning algorithm of the neural network and thus it can be used to

approximate nonlinear functions or as a universal function estimator **(Chen and Chang 2010[61], Do 2020[84])**. The approximation of the unknown function of the data set for the production of ANFIS networks is done using data-driven procedures **(Moosavi et al. 2013[173])**. It aims at improving the proficiency of the fuzzy systems and thus it has been effectively used for pattern recognition, control processes, task classification problems **(Karmakar and Mujumdar 2006[130])**. It consists of various similar models used for obtaining fuzzy rules from an input-output data set and validating the systematic approach as proposed by Mamdani. Fuzzifier, fuzzy database, and Defuzzifier are the important stages of fuzzy modeling as shown in Figure 5.5 below **(Toprak et al. 2009[278], Kumar 2020[143])**.

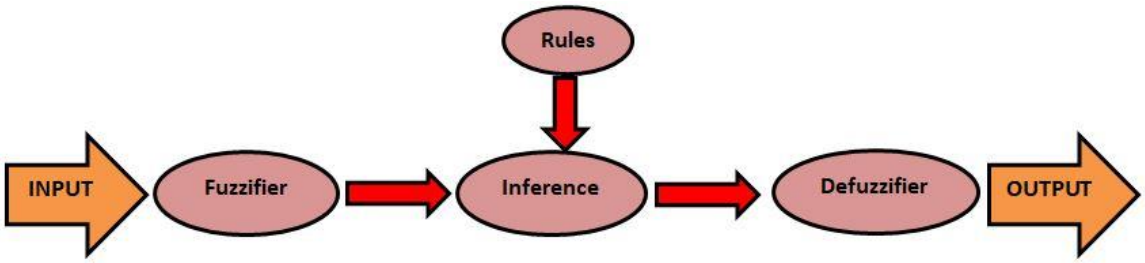


Figure 5.5 Working process of Fuzzy modeling

Along with backpropagation, the technique feed-forward neural network (FFNN) is used for getting better predictions. Various time series models were mostly preferred by researchers to obtain the predictions of several real life time series data but most of these models fail to handle nonlinear problems as these were based on the linear correlation **(Nayak et al. 2004[182])**. In the meanwhile, soft computing techniques like the artificial neural network (ANN) and adaptive neuro-fuzzy interface system (ANFIS) have proved to be an efficient tool in handling such nonlinear problems capable of modeling complex patterns **(Nourani and Andalib 2015[185], Partovian et al. 2016[206], Ahmed et al. 2019[14], Parmar et al. 2019[204])**.

5.8 Results and Discussion

For the development of an accurate model for obtaining the forecast of black carbon concentration over the coal mines, a new hybrid model is developed by the conjugation

of wavelets, ANN, and FIS coupled approach termed as the wavelet-ANFIS model or the 'Artificial Neural Network Fuzzy Wavelet Conjugation model' termed as the hybrid ANNFWCM model. It is developed and selected after training and testing the values of the dataset. The self-learning ability of the neural network is displayed in this wavelet-ANFIS model. Error measures namely the Mean absolute error (MAE) and the relative error$(\in)$are used for determining the results of the model. Time series of Raniganj coal mine is selected for this study, as it is the oldest coal mine of the country and there is a huge correlation among the time-series of the three sample sites.

5.8.1 Wavelet Decomposition

For analysis, the time-series signal of the Raniganj sample site is decomposed utilizing the wavelet decomposition technique. While dealing with spike and random series, Daubechies wavelets are the most suitable, and with the increase in the order the smoothness increases. Thus we are using the best fitted (Daubechies wavelets of order 8) db8 level 3 for decomposing the time-series signal. This decomposition led to the generation of the three detailed parts D1, D2 and D3 and an approximation part A3 which are further treated with the ANFIS model considering these four components as individual sub-time-series, such that the original time-series signal S=A3+D1+D2+D3 is decomposed as shown in Figure 5.6

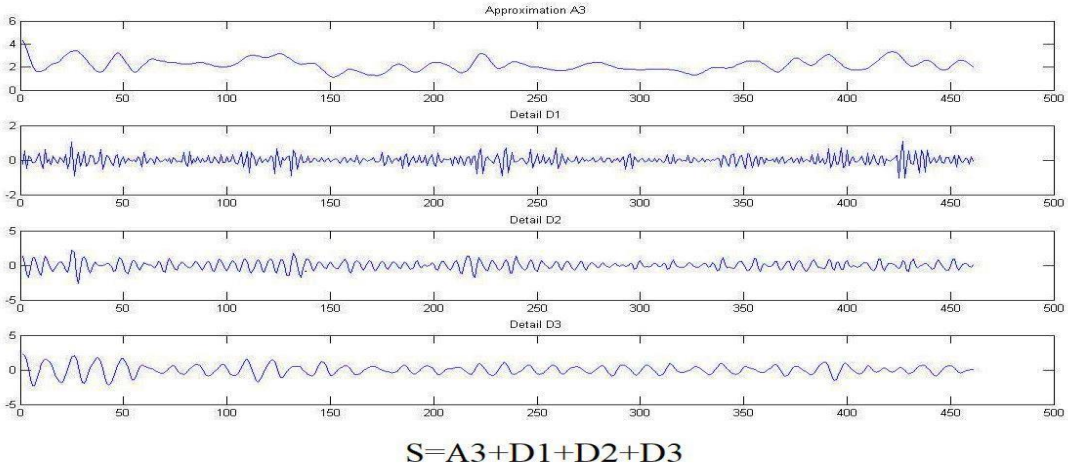


Figure 5.6 Wavelet decomposition of time-series of Raniganj

5.8.2 The Adaptive Neuro-Fuzzy Inference System (ANFIS) Model

For the application of the ANFIS model, firstly the sub-time-series signals obtained after decomposition namely, A3, D1, D2, and D3 are divided into three parts, input, testing, and training phases. As such that the first three observations are taken as the input for the machine algorithm, 400 observations were taken for the training phase while 59 are considered for the testing phase. The detailed and approximation parts are further treated with the Gauss membership function along with the Mamdani technique for the Fuzzy Inference (FIS). Mamdani system is the most popular FIS (Fuzzy inference system) having wide acceptance, the output of this a fuzzy set that is defuzzified to get a final output. Gauss membership function (*gaussmf*) is a form of the fuzzy membership function. It is defined as, $f\left(x,\sigma,\mu\right)=e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, it calculates the fuzzy membership values and leads to the improvement in the robustness and reliability of the system.

ANN-based backpropagation technique forming the basis of the three-layered feed-forward neural network is used with 1000 Epochs and zero tolerance level for training the algorithm. Same as the number of inputs it consists of three input nodes with sigmoid activation function along with a single output. After suitable training of the sub-time-series signals of A3, D1, D2, and D3 we obtain the training errors in terms of mean absolute error (MAE) and the relative error$\left(\in\right)$as shown in Table 5.1. The quiver and surface views of training obtained for A3, D1, D2, and D3 are displayed in Figure 5.7. After suitable training, the training output values along with the actual data values are plotted on a graph to check the efficiency of training. The graph of the actual data values verses the training output values obtained by the ANFIS of A3, D1, D2 and D3 are displayed in Figure 5.8, which depicts that after training the trained output signal fitted very well with the original time-series signal.

This trained ANFIS model is then passed through the testing phase; from this, we obtain 59 data values. These values obtained from the testing phase are then compared with the actual data values of the time-series to obtain the Mean absolute error (MAE)

and the relative error $(\in)$ as shown in Table 5.1. Figure 5.9, gives the graphical representation of the reconstructed modeled trained (output) signal with the actual time-series signal. Figure 5.10 represents the results of the reconstructed modeled tested (output) signal with the actual time-series signal. These figures display that the wavelet-ANFIS model gave the output (both trained and tested) very close to the actual time-series.

Table 5.1 Error measures of Training and Testing phase

Error Measures	**Mean absolute error (MAE)**	**Relative error** $(\in)$
Training phase	0.0484	0.0223
Testing phase	0.0382	0.0164

As observed in Table 5.1 the error values are very close to zero, thus the hybrid wavelet-ANFIS model fitted very well to the time-series data. The goodness of testing can be seen as the testing error is smaller compared to the training error. To make a comparison of the wavelet-ANFIS hybrid model with the time-series ARIMA and the wavelet-ARIMA coupled model developed in the previous chapter, we henceforth compare the results of error measure obtained by all the models for Raniganj sample site. These results are displayed in Table 5.2 from which it can be clearly seen that the error value namely the MAE of the hybrid wavelet-ANFIS model is much smaller than the time-series ARIMA and the wavelet-ARIMA coupled approach.

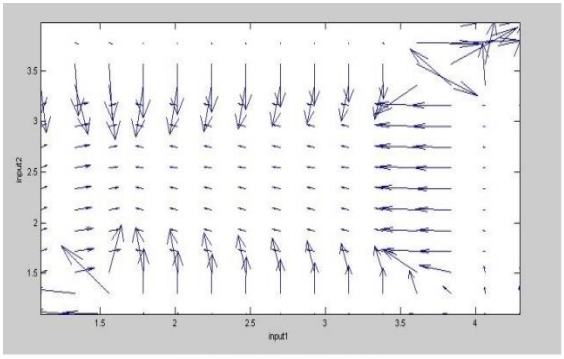


Fig. 5.7(a) A3 training Quiver view

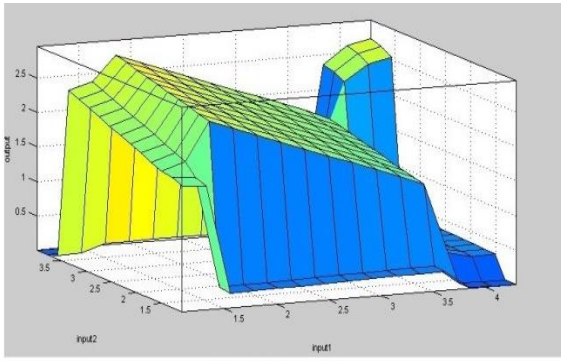


Fig. 5.7(b) A3 training Surface view

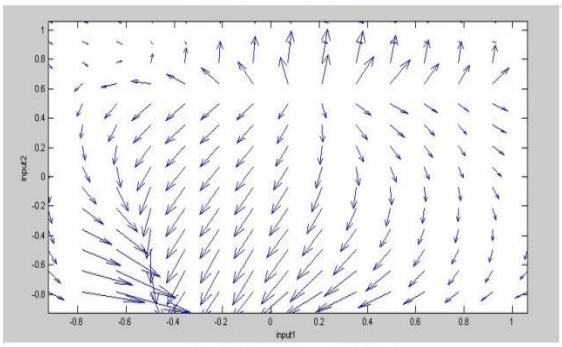


Fig. 5.7(c) D1 training Quiver view

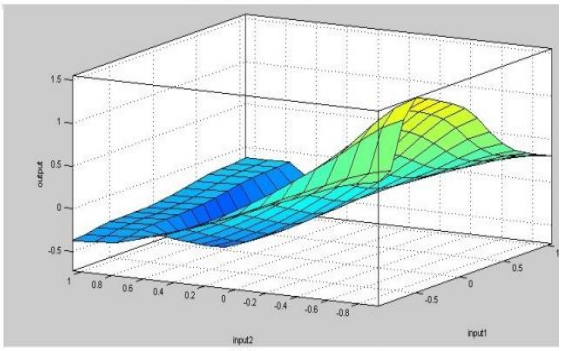


Fig. 5.7(d) D1 training Surface view

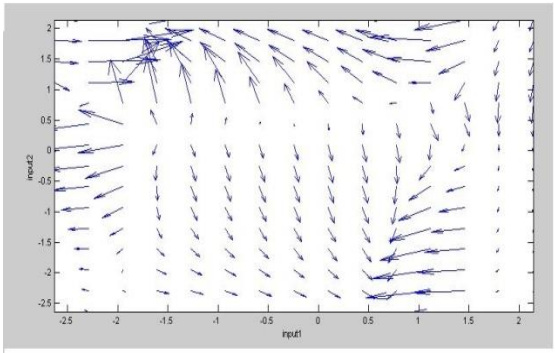


Fig. 5.7(e) D2 training Quiver view

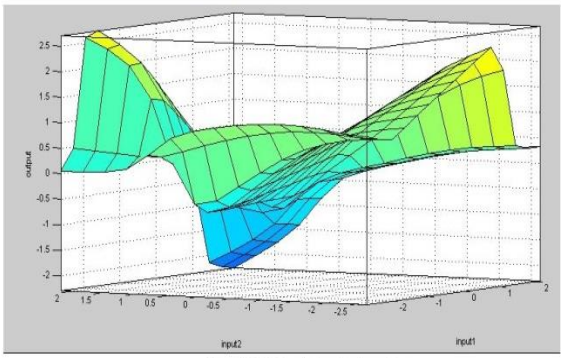


Fig. 5.7(f) D2 training Surface view

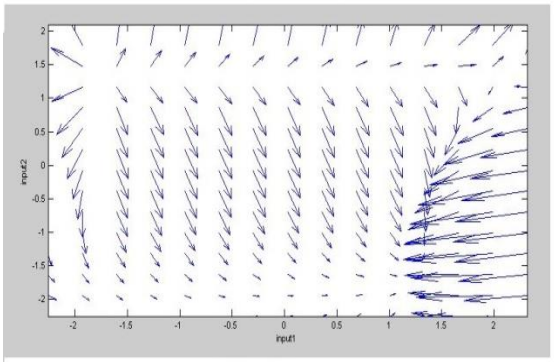


Fig. 5.7(g) D3 training Quiver view

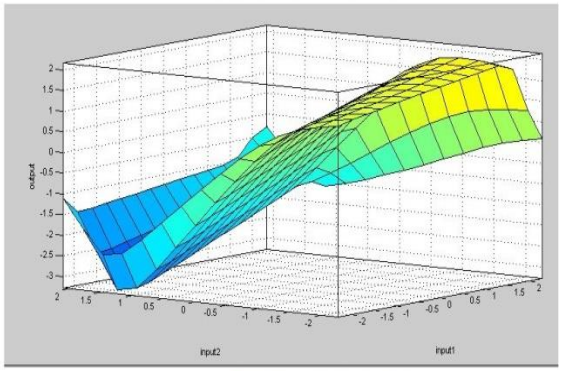


Fig. 5.7(h) D3 training Surface view

Figure 5.7 (a-h) A3, D1, D2, and D3 Quiver and Surface view

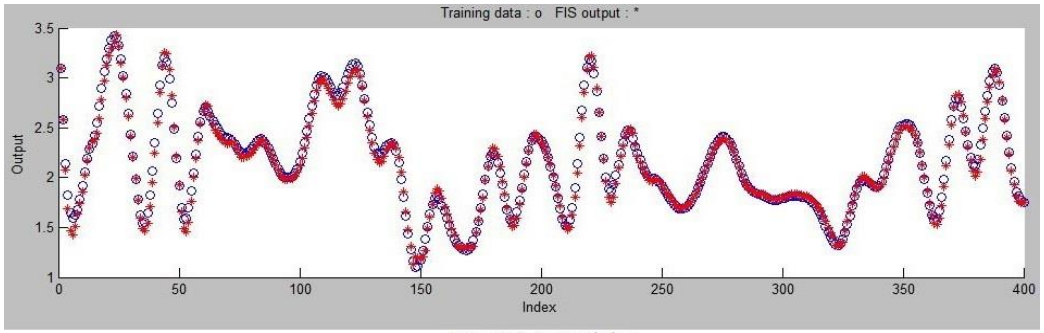


Fig. (a) A3 Training

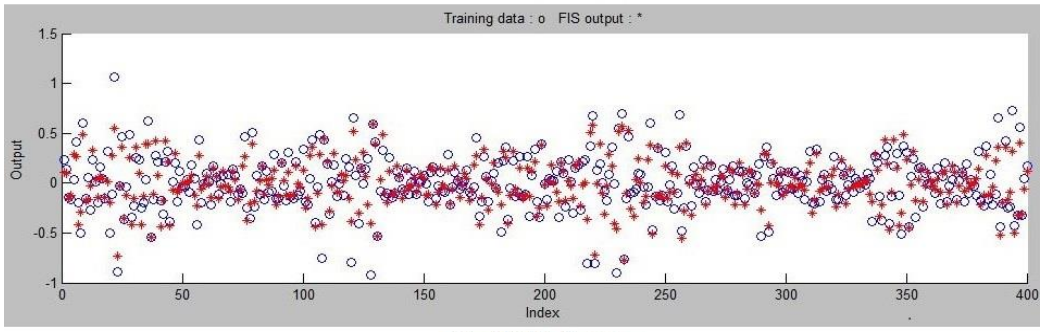


Fig. (b) D1 Training

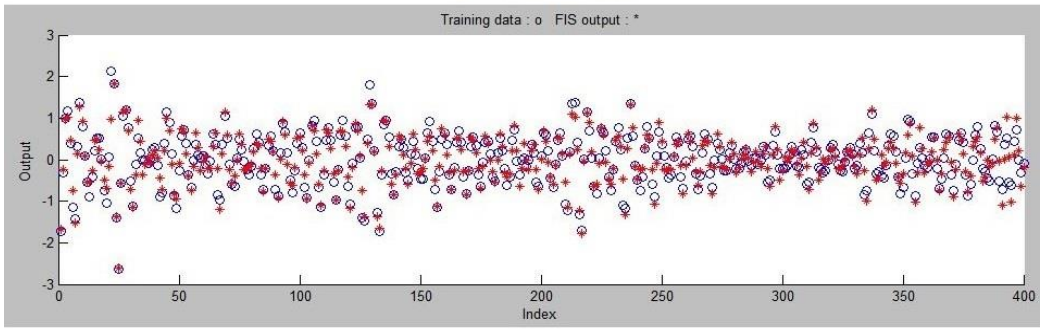


Fig. (c) D2 Training

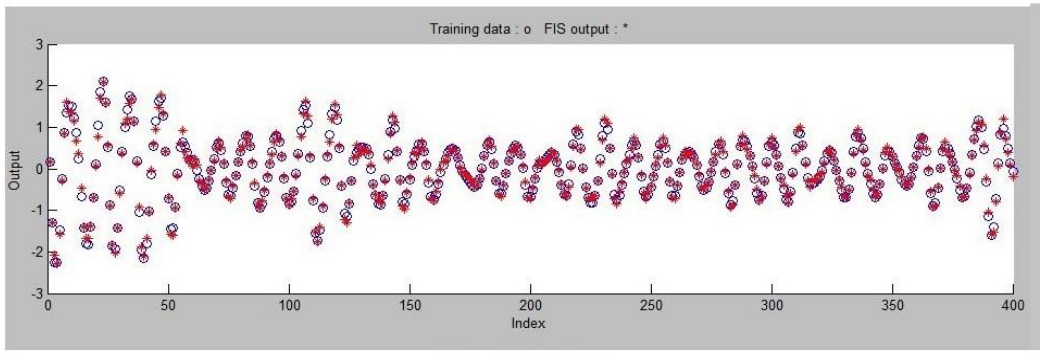


Fig. (d) D3 Training

Figure 5.8 (a-d) A3, D1, D2, D3 Training

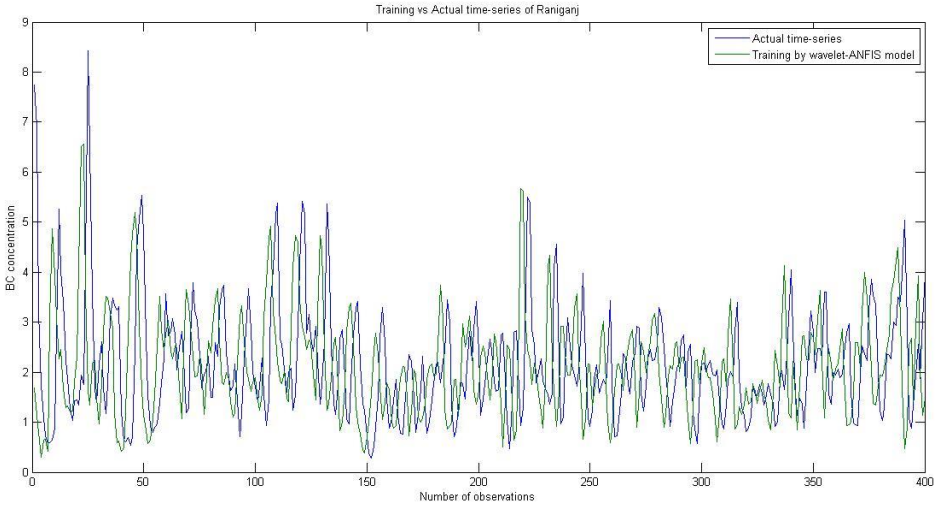


Figure 5.9 Actual time-series signal vs trained signal by wavelet-ANFIS model

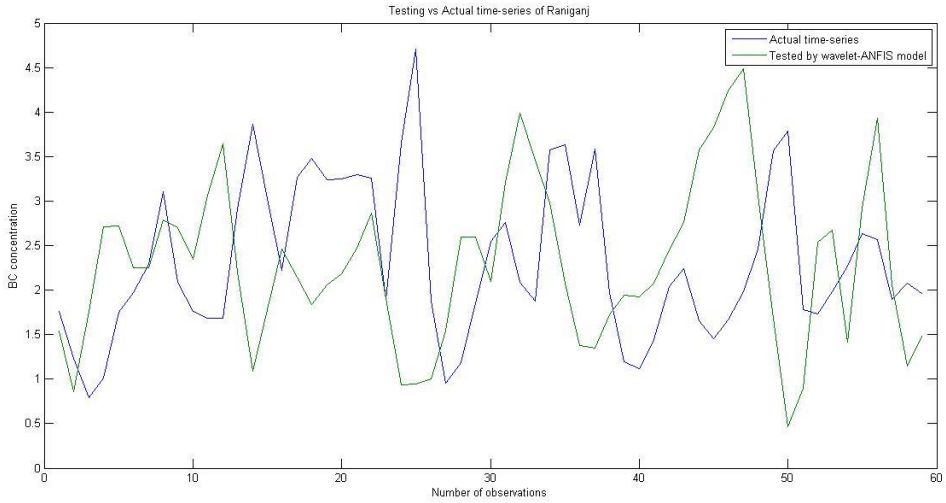


Figure 5.10 Actual time-series signal vs tested signal by wavelet-ANFIS model

Table 5.2 Error comparison of the time-series of Raniganj

Error Measure	Time-series ARIMA model	Wavelet-ARIMA coupled model	Hybrid Wavelet-ANFIS model
MAE	0.678623	0.64809	**0.0382**

5.9 Conclusion

In this study, for the time-series data of Raniganj the hybrid wavelet-ANFIS approach has been developed for the analysis of the best-fitted model. After suitable training and testing the model for the data values, it is observed that the hybrid wavelet-ANFIS approach fitted very well to the time-series data. Based on the model fitting parameter namely the MAE, it is observed that the hybrid wavelet-ANFIS approach performed better than the time-series ARIMA and wavelet-ARIMA coupled approach as shown in Table 5.2. Thus we conclude that the hybrid wavelet-ANFIS model is the best-fitted model and can be effectively used as an efficient forecasting model with minimum forecasting error for our study. This is a noble study as it can be used as an effective tool to obtain the future forecasts of black carbon concentration in the coal mine regions and help the government agencies to frame suitable policies and preventive measures for the future of these regions.

5.10 Future Scope of the Study

The following mathematical models could be developed in future and may be tested for accuracy for the prediction of pollutants like black carbon over the coal mine regions in India using the results discussed above as the base for their study.

- ARIMA-GARCH coupled approach.
- Hybrid models could be formed by coupling of various other models like DENFIS (Dynamic Evolving Neural Fuzzy Inference System), SVM (Support Vector Machine), Tree Models.

BIBLIOGRAPHY

[1] Abdalla SZ. Modelling exchange rate volatility using GARCH models: Empirical evidence from Arab countries. International Journal of Economics and Finance. 2012 Mar;4(3):216-229.

[2] Abdel-Aziz A, Frey HC. Development of hourly probabilistic utility NO_x emission inventories using time series techniques: Part II—multivariate approach. Atmospheric Environment. 2003 Dec 1;37(38):5391-5401.

[3] Abdulqader QM. Annual Forecasting Using a Hybrid Approach. General Letters in Mathematics. 2018;4(2):86-95.

[4] Abghari H, Ahmadi H, Besharat S, Rezaverdinejad V. Prediction of daily pan evaporation using wavelet neural networks. Water Resources Management. 2012 Sep 1;26(12):3639-3652.

[5] Abish B, Mohanakumar K. A stochastic model for predicting aerosol optical depth over the north Indian region. International Journal of Remote Sensing. 2013 Feb 20;34(4):1449-1458.

[6] Abraham A, Nath B. A neuro-fuzzy neural network approach for modeling electricity demand in Victoria. Appl. Soft Comput. 2001 Aug;1(2):127-138.

[7] Adamowski J, Chan HF. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology. 2011 Sep 15;407(1-4):28-40.

[8] Adamowski J, Sun K. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology. 2010 Aug 20;390(1-2):85-91.

[9] Addison PS. The illustrated wavelet transform handbook: Introductory theory and applications in science and engineering, medicines and finance. Institute of Physics Publishing, Bristol and Philadelphia; 2002. 362 p.

[10] Addison PS. Wavelet transforms and the ECG: a review. Physiological measurement. 2005 Oct 8;26(5):R155-R199.

[11] Adedeji PA, Akinlabi S, Madushele N, Olatunji OO. Hybrid adaptive neuro-fuzzy inference system (ANFIS) for a multi-campus university energy consumption forecast. International Journal of Ambient Energy. 2020 Jan;31:1-0. https://doi.org/10.1080/01430750.2020.1719885

[12] Agarwal A, Lubet A, Mitgang E, Mohanty S, Bloom DE. Population aging in India: Facts, issues, and options. In Population Change and Impacts in Asia and the Pacific, Springer, Singapore. 2016 Aug. p. 289-311.

[13] Agarwal R, Awasthi A, Singh N, Gupta PK, Mittal SK. Effects of exposure to rice-crop residue burning smoke on pulmonary functions and Oxygen Saturation level of human beings in Patiala (India). Science of the Total Environment. 2012 Jul 1;429:161-166.

[14] Ahmed AN, Othman FB, Afan HA, Ibrahim RK, Fai CM, Hossain MS, Ehteram M, Elshafie A. Machine learning methods for better water quality prediction. Journal of Hydrology. 2019 Nov 1;578:124084. https://doi.org/10.1016/j.jhydrol.2019.124084

[15] Akrami SA, El-Shafie A, Jaafar O. Improving rainfall forecasting efficiency using modified adaptive neuro-fuzzy inference system (MANFIS). Water Resources Management. 2013 Jul 1;27(9):3507-3523.

[16] Aksoy H, Dahamsheh A. Artificial neural network models for forecasting monthly precipitation in Jordan. Stochastic Environmental Research and Risk Assessment. 2009 Oct 1;23(7):917-931.

[17] Aksoy H, Toprak ZF, Aytek A, Ünal NE. Stochastic generation of hourly mean wind speed data. Renewable Energy. 2004 Nov 1;29(14):2111-2131.

[18] Al-Gounmeein RS, Ismail MT. Forecasting the Exchange Rate of the Jordanian Dinar versus the US Dollar Using a Box-Jenkins Seasonal ARIMA Model. Computer Science. 2020;15(1):27-40.

[19] Alsharif MH, Younes MK, Kim J. Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea. Symmetry. 2019 Feb;11(2):240.

[20] Aminzadeh MS. Sequential and non-sequential acceptance sampling plans for autocorrelated processes using ARMA (p, q) models. Computational Statistics. 2009 Feb 1;24(1):95-111.

[21] Andreae MO, Crutzen PJ. Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry. Science. 1997 May 16;276(5315):1052-1058.

[22] Andreo B, Jiménez P, Durán JJ, Carrasco F, Vadillo I, Mangin A. Climatic and hydrological variations during the last 117–166 years in the south of the Iberian Peninsula, from spectral and correlation analyses and continuous wavelet analyses. Journal of Hydrology. 2006 Jun 15;324(1-4):24-39.

[23] Antonakakis N, Darby J. Forecasting volatility in developing countries nominal exchange returns. Applied Financial Economics. 2013 Nov 1;23(21):1675-1691.

[24] Arachchi AK. Comparison of Symmetric and Asymmetric GARCH Models: Application of Exchange Rate Volatility. American Journal of Mathematics and Statistics. 2018;8(5):151-159.

[25] Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U, Rabczuk T, Atkinson PM. Covid-19 outbreak prediction with machine learning. Available at SSRN 3580188. 2020 Apr 19. https://doi.org/10.1101/2020.04.17.20070094

[26] Attri SD. Atlas of hourly mixing height and assimilative capacity of atmosphere in India. India Meteorological Department, New Delhi; 2008. 113 p.

[27] Ayele AW, Gabreyohannes E, Edmealem H. Generalized Autoregressive Conditional Heteroskedastic Model to Examine Silver Price Volatility and Its Macroeconomic Determinant in Ethiopia Market. Journal of Probability and Statistics. 2020 May 25;2020(1):1-10.

[28] Baghanam AH, Nourani V, Keynejad MA, Taghipour H, Alami MT. Conjunction of wavelet entropy and SOM clustering for multi-GCM statistical downscaling. Hydrology Research. 2019 Feb 1;50(1):1-23.

[29] Ballester EB, Valls GC, Carrasco-Rodriguez JL, Olivas ES, del Valle-Tascon S. Effective 1 day ahead prediction of hourly surface ozone concentrations in eastern Spain using linear models and neural networks. Ecological Modelling. 2002 Oct 15;156(1):27-41.

[30] Baş D, Boyacı İH. Modeling and optimization II: comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. Journal of Food Engineering. 2007 Feb 1;78(3):846-854.

[31] Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. Health effects of outdoor air pollution. American Journal of Respiratory and Critical Care Medicine. 1996;153(1):3–50.

[32] Begam GR, Vachaspati CV, Ahammed YN, Kumar KR, Reddy RR, Sharma SK, Saxena M, Mandal TK. Seasonal characteristics of water-soluble inorganic ions and carbonaceous aerosols in total suspended particulate matter at a rural semi-arid site, Kadapa (India). Environmental Science and Pollution Research. 2017 Jan 1;24(2):1719-1734.

[33] Begum BA, Hopke PK, Markwitz A. Air pollution by fine particulate matter in Bangladesh. Atmospheric Pollution Research. 2013 Jan 1;4(1):75-86.

[34] Behera R, Mehra M. Integration of barotropic vorticity equation over spherical geodesic grid using multilevel adaptive wavelet collocation method. Applied Mathematical Modelling. 2013 Apr 1;37(7):5215-5226.

[35] Benmouiza K, Cheknane A. Small-scale solar radiation forecasting using ARMA and nonlinear autoregressive neural network models. Theoretical and Applied Climatology. 2016 May 1;124(3-4):945-958.

[36] Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M. Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in Brief. 2020 Feb 26;29:105340.

[37] Beylkin G, Coifman R, Rokhlin V. Fast wavelet transforms and numerical algorithms I. Communications on Pure and Applied Mathematics. 1991 Mar;44(2):141-183.

[38] Bhardwaj S, Chandrasekhar E, Padiyar P, Gadre VM. A comparative study of wavelet-based ANN and classical techniques for geophysical time-series forecasting. Computers & Geosciences. 2020 Feb 29;138:104461.

[39] Bhattacharjee S. India's Coal Story from Damodar to Zambezi. SAGE Publications Pvt. Ltd India, First Edition; 2017 Apr 3. 288 p.

[40] Bodyanskiy Y, Vynokurova O. Hybrid adaptive wavelet-neuro-fuzzy system for chaotic time series identification. Information Sciences. 2013 Jan 20;220:170-179.

[41] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics. 1986 Apr 1;31(3):307-327.

[42] Borwankar V. "Air pollution killed 81,000 in Delhi & Mumbai, cost Rs 70,000 crore in 2015." Times of India [Internet]. 2017; Available from: http://timesofindia.indiatimes.com/city/mumbai/air-pollution-killed-81000-in-delhi-mumbai-cost-rs-70000-crore-in-2015/articleshow/56656252.cms

[43] Bové H, Bongaerts E, Slenders E, Bijnens EM, Saenen ND, Gyselaers W, Van Eyken P, Plusquin M, Roeffaers MB, Ameloot M, Nawrot TS. Ambient black carbon particles reach the fetal side of human placenta. Nature Communications. 2019 Sep 17;10(1):1-7.

[44] Box GE. GM Jenkins Time Series Analysis: Forecasting and Control Wiley. New York; 1976. 712 p.

[45] Brzezina J, Köbölová K, Adamec V. Nanoparticle Number Concentration in the Air in Relation to the Time of the Year and Time of the Day. Atmosphere. 2020 May 19;11(5):523.

[46] Byun SJ, Cho H. Forecasting carbon futures volatility using GARCH models with energy volatilities. Energy Economics. 2013 Nov 1;40(C):207-221.

[47] Can Z, Aslan Z, Oguz O, Siddiqi AH. Wavelet transform of metrological parameter and gravity waves. Annales Geophysicae. 2005 Mar 30;23(3):659–663.

[48] Capobianco E. Misspecifying GARCH-M Processes. Complex Systems. 1995 Oct 23;9(6):477-490.

[49] Caporin M, Lisi F. Misspecification tests for periodic long memory GARCH models. Statistical Methods and Applications. 2010 Mar 1;19(1):47-62.

[50] Caporin M. Identification of long memory in GARCH models. Statistical Methods and Applications. 2003 Dec 1;12(2):133-151.

[51] Central Electricity Authority, Jan 2018 (http://www.cea.nic.in/reports/committee/nep/nep_jan_2018.pdf)

[52] Central Electricity Authority, June 2014 (http://www.cea.nic.in/reports/monthly/dpd_div_rep/village_electrification.pdf)

[53] Central Mining Research Institute (CMRI). Determination of Emission Factor for Various Opencast Mining Activities, GAP/9/EMG/MOEF/09, Environmental Management Group, Dhanbad, India.1998.

[54] Centre for Environmental Health, Government of India. Air pollution and health in India: A review of the current evidence and opportunities for the future. 2017 (https://www.ceh.org.in/wp-content/uploads/2017/10/Air-Pollution-and-Health-in-India.pdf)

[55] Ceylan Z. Estimation of COVID-19 prevalence in Italy, Spain, and France. Science of the Total Environment. 2020 Apr 22;729:138817.

[56] Chameides WL, Yu H, Liu SC, Bergin M, Zhou X, Mearns L, Wang G, Kiang CS, Saylor RD, Luo C, Huang Y. Case study of the effects of atmospheric aerosols and regional haze on agriculture: an opportunity to enhance crop yields in China through emission controls?. Proceedings of the National Academy of Sciences. 1999 Nov 23;96(24):13626-13633.

[57] Chandra D, Singh RM, Singh MP. Text Book of Coal (Indian Context). Tara Book Agency, Varanasi. 2000. 400 p.

[58] Chattopadhyay G, Chattopadhyay S. Autoregressive forecast of monthly total ozone concentration: A neurocomputing approach. Computers & Geosciences. 2009 Sep 1;35(9):1925-1932.

[59] Chaulya SK, Chakraborty MK. Perspective of new national mineral policy and environmental control for mineral sector. In Proceedings of national seminar on status of mineral exploitation in India, New Delhi, India; 1995. p. 114-123.

[60] Chelani AB, Devotta S. Air quality forecasting using a hybrid autoregressive and nonlinear model. Atmospheric Environment. 2006 Mar 1;40(10):1774-1780.

[61] Chen HW, Chang NB. Using fuzzy operators to address the complexity in decision making of water resources redistribution in two neighboring river basins. Advances in Water Resources. 2010 Jun 1;33(6):652-666.

[62] Chen JF, Wang WM, Huang CM. Analysis of an adaptive time-series autoregressive moving average (ARMA) model for short-term load forecasting. Electric Power Systems Research. 1995 Sep 1;34(3):187-196.

[63] Coal Controller Organization of India, Kolkata, Ministry Of Coal, Government of India. Provisional Coal Statistics 2018-19. (http://www.coalcontroller.gov.in/writereaddata/files/download/provisionalcoalstat/ProvisionalCoalStat2018-19.pdf) 2019 Nov; 96 p.

[64] Cohen A, Kovacevic J. Wavelets: The mathematical background. Proceedings of the IEEE. 1996 Apr;84(4):514-522.

[65] Conny JM, Slater JF. Black carbon and organic carbon in aerosol particles from crown fires in the Canadian boreal forest. Journal of Geophysical Research: Atmospheres. 2002 Jun 16;107(D11):AAC 4-1-AAC 4-12.

[66] Contreras-Reyes JE, Palma W. Statistical analysis of autoregressive fractionally integrated moving average models in R. Computational Statistics. 2013 Oct 1;28(5):2309-2331.

[67] Cortez P, Rocha M, Neves J. Evolving time series forecasting ARMA models. Journal of Heuristics. 2004 Jul 1;10(4):415-429.

[68] CPCB (Central Pollution Control Board), India, National Ambient Air Quality Standards (NAAQS). Gazette notification, New Delhi. 2009.

[69] Crosson E. A cavity ring-down analyzer for measuring atmospheric levels of methane, carbondioxide, and water vapor. Applied Physics B. 2008 Sep 1;92(3):403-408.

[70] Dalkiliç HY, Hashimi SA. Prediction of daily streamflow using artificial neural networks (ANNs), wavelet neural networks (WNNs), and adaptive neuro-fuzzy inference system (ANFIS) models. Water Supply. 2020 June 1;20(4):1396-1408.

[71] Das R, Khezri B, Srivastava B, Datta S, Sikdar PK, Webster RD, Wang X. Trace element composition of $PM_{2.5}$ and PM_{10} from Kolkata–a heavily polluted Indian metropolis. Atmospheric Pollution Research. 2015 Sep 1;6(5):742-750.

[72] Daubechies I. Different perspectives on wavelet. Proceeding on the Symposia on Applied Mathematics. American Mathematical Society. 1993 Jan 11;47, 206 p.

[73] Daubechies I. Introduction to the special issue. IEEE Trans. Inform. Theory. 1992;38(2):529-531.

[74] Daubechies I. Orthonormal bases of compactly supported wavelets. Communications on pure and applied mathematics. 1988 Oct;41(7):909-996.

[75] Daubechies I. Ten lectures on wavelets (SIAM, Philadelphia, 1992). MR 93e. 1992;42045.

[76] Daubechies I. Time-frequency localization operators: a geometric phase space approach. IEEE Transactions on Information Theory. July 1988;34(4):605-612.

[77] Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE transactions on information theory. 1990 Sep;36(5):961-1005.

[78] Davuliene L, Sakalys J, Dudoitis V, Reklaite A, Ulevicius V. Long-term black carbon variation in the South-Eastern Baltic Region in 2008–2015. Atmospheric Pollution Research. 2019 Jan 1;10(1):123-133.

[79] De Capitani L. Interval estimation for the Sharpe Ratio when returns are not iid with special emphasis on the GARCH (1, 1) process with symmetric innovations. Statistical Methods & Applications. 2012 Nov;21(4):517-537.

[80] Dey A. Machine learning algorithms: a review. International Journal of Computer Science and Information Technologies. 2016;7(3):1174-1179.

[81] Dey S, Di Girolamo L, van Donkelaar A, Tripathi SN, Gupta T, Mohan M. Variability of outdoor fine particulate ($PM_{2.5}$) concentration in the Indian Subcontinent: A remote sensing approach. Remote Sensing of Environment. 2012 Dec 1;127:153-161.

[82] Dimitrakopoulos S, Tsionas M. Ordinal-response GARCH models for transaction data: A forecasting exercise. International Journal of Forecasting. 2019 Oct 1;35(4):1273-1287.

[83] Diodato N, Guerriero L, Fiorillo F, Esposito L, Revellino P, Grelle G, Guadagno FM. Predicting monthly spring discharges using a simple statistical model. Water resources management. 2014 Mar 1;28(4):969-978.

[84] Do QH. Development of MI-ANFIS-BBO Model for Forecasting Crude Oil Price. In Reliability and Statistical Computing. Springer International Publishing; 2020. p. 167-191.

[85] Doguwa SI, Omotosho BS. Understanding the dynamics of inflation volatility in Nigeria: A GARCH Perspective. CBN Journal of Applied Statistics. 2012;3(2):51-74.

[86] Dökmen F, Aslan Z. Evaluation of the parameters of water quality with wavelet techniques. Water resources management. 2013 Nov 1;27(14):4977-4988.

[87] Doucoure B, Agbossou K, Cardenas A. Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. Renewable Energy. 2016 Jul 1;92:202-211.

[88] Dritsaki C. Modeling the Volatility of Exchange Rate Currency using GARCH Model. Economia Internazionale/International Economics. 2019;72(2):209-230.

[89] El Jebari O, Hakmaoui A. GARCH Family Models vs EWMA: Which is the Best Model to Forecast Volatility of the Moroccan Stock Exchange Market?. Revista de Métodos Cuantitativos para la Economía y la Empresa. 2018;26:237-249.

[90] Engle RF, Lilien DM, Robins RP. Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica: journal of the Econometric Society. 1987 Mar 1;55(2):391-407.

[91] Engle RF. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica. 1982;50(4):987-1008.

[92] Farajzadeh J, Fard AF, Lotfi S. Modeling of monthly rainfall and runoff of Urmia lake basin using "feed-forward neural network" and "time series analysis" model. Water Resources and Industry. 2014 Sep 1;(7-8):38-48.

[93] Faruk DÖ. A hybrid neural network and ARIMA model for water quality time series prediction. Engineering Applications of Artificial Intelligence. 2010 Jun 1;23(4):586-594.

[94] Fiero MH, Roydhouse JK, Vallejo J, King-Kallimanis BL, Kluetz PG, Sridhara R. US Food and Drug Administration review of statistical analysis of patient-reported outcomes in lung cancer clinical trials approved between January, 2008, and December, 2017. The Lancet Oncology. 2019 Oct 1;20(10):e582-e589.

[95] Gadgil, DR. The Industrial Evolution of India in Recent Times, 1860-1939. Oxford University Press, New Delhi; 1971. 362 p.

[96] Gao Y, Zhang C, Zhang L. Comparison of GARCH Models based on Different Distributions. JCP. 2012 Aug;7(8):1967-1973.

[97] Gautam R, Baral S, Herat S. Biogas as a sustainable energy source in Nepal: Present status and future challenges. Renewable and Sustainable Energy Reviews. 2009 Jan 1;13(1):248-252.

[98] Gautam S, Patra AK, Sahu SP, Hitch M. Particulate matter pollution in opencast coal mining areas: a threat to human health and environment. International Journal of Mining, Reclamation and Environment. 2018 Feb 17;32(2):75-92.

[99] Ghosh SC. The Raniganj Coal Basin: An example of an Indian Gondwana rift. Sedimentary Geology. 2002 Feb;147(1):155-176.

[100] Goenka D, Guttikunda S. Coal Kills: An Assessment of Death and Disease caused by India's Dirtiest Energy Source. Conservation Action Trust (India) and Urban Emissions (India); 2014. 20 p.

[101] Graham A, Mishra EP. Time series analysis model to forecast rainfall for Allahabad region. Journal of Pharmacognosy and Phytochemistry. 2017;6(5):1418-1421.

[102] Grossmann A, Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis. 1984 Jul;15(4):723-736.

[103] Guttikunda SK, Goel R, Pant P. Nature of air pollution, emission sources, and management in the Indian cities. Atmospheric environment. 2014 Oct 1;95:501-510.

[104] Guttikunda SK, Jawahar P. Atmospheric emissions and pollution from the coal-fired thermal power plants in India. Atmospheric Environment. 2014 Aug 1;92:449-460.

[105] Haar A. On the theory of orthogonal function systems. Mathematische Annalen. 1910;69(3):331-371.

[106] Hamzaçebi C. Improving artificial neural networks' performance in seasonal time series forecasting. Information Sciences. 2008 Dec 1;178(23):4550-4559.

[107] Hansen PR, Lunde A. A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. Journal of applied econometrics. 2005 Dec;20(7):873-889.

[108] Happel BLM, Murre JM. The design and evolution of modular neural network architecture. Neural Networks. 1994 Nov;7(6-7):985-1004.

[109] Hassanzadeh S, Hosseinibalam F, Alizadeh R. Statistical models and time series forecasting of sulfur dioxide: a case study Tehran. Environmental Monitoring and Assessment. 2009 Aug 1;155(1-4):149-155.

[110] Heghedus C, Segarra S, Chakravorty A, Rong C. Neural Network Architectures for Electricity Consumption Forecasting. International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data. 2019 Jul 14;776-783.

[111] Highwood EJ, Kinnersley RP. When smoke gets in our eyes: The multiple impacts of atmospheric black carbon on climate, air quality and health. Environment International. 2006 May 1;32(4):560-566.

[112] Horvath H. Atmospheric light absorption—A review. Atmospheric Environment. Part A. General Topics. 1993 Feb 1;27(3):293-317.

[113] Hsu KL, Gupta HV, Sorooshian S. Artificial neural network modeling of the rainfall-runoff process. Water Resources Research. 1995 Oct;31(10):2517-2530.

[114] Hsu NC, Herman JR, Tsay SC. Radiative impacts from biomass burning in the presence of clouds during boreal spring in southeast Asia. Geophysical Research Letters. 2003 Mar;30(5):1224.

[115] Hu MJ. Application of the adaline system to weather forecasting (Doctoral dissertation, Department of Electrical Engineering, Stanford University) 1964.

[116] Hu ZZ, Nitta T. Wavelet analysis of summer rainfall over North China and India and SOI using 1891-1992 data. Journal of the Meteorological Society of Japan. Ser. II. 1996 Dec 25;74(6):833-844.

[117] Huang T, Chen J, Zhao W, Cheng J, Cheng S. Seasonal variations and correlation analysis of water-soluble inorganic ions in $PM_{2.5}$ in Wuhan, 2013. Atmosphere. 2016 Apr;7(4):49.

[118] Hubacek K, Guan D, Barua A. Changing lifestyles and consumption patterns in developing countries: A scenario analysis for China and India. Futures. 2007 Nov 1;39(9):1084-1096.

[119] Hung NQ, Babel MS, Weesakul S, Tripathi NK. An artificial neural network model for rainfall forecasting in Bangkok Thailand. Hydrol. Earth Syst. Sci. (2009); 13(8):1413–1425.

[120] Indian Bureau of Mines, Ministry of Mines, Government of India. Indian Minerals Yearbook 2018 (Part-III: Mineral Reviews) Coal and Lignite, 57th Edition; Jul 2019. 33 p.

[121] Indian Space Research Organisation. Scientific results from ISRO Geosphere Biosphere Programme 1994. Bangalore, India: Indian Space Research Organisation, Dept. of Space. 1994. p. 42-94.

[122] Jacobson MZ. Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols. Nature. 2001 Feb;409(6821):695-697.

[123] Jang JS. ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics. 1993 May;23(3):665-685.

[124] Jansen KL, Larson TV, Koenig JQ, Mar TF, Fields C, Stewart J, Lippmann M. Associations between health effects and particulate matter and black carbon in subjects with respiratory disease. Environmental Health Perspectives. 2005 Dec;113(12):1741-1746.

[125] Janssen NA, Gerlofs-Nijland ME, Lanki T, Salonen RO, Cassee F, Hoek G, Fischer P, Brunekreef B, Krzyzanowski M. (2012) Health Effects Of Black Carbon. World Health Organization Regional Office for Europe. (https://www.euro.who.int/__data/assets/pdf_file/0004/162535/e96541.pdf). 2012; 86 p.

[126] Japar SM, Brachaczek WW, Gorse Jr RA, Norbeck JM, Pierson WR. The contribution of elemental carbon to the optical properties of rural atmospheric aerosols. Atmospheric Environment. 1986 Jan 1;20(6):1281-1289.

[127] Jeong C, Shin JY, Kim T, Heo JH. Monthly precipitation forecasting with a neuro-fuzzy model. Water Resources Management. 2012 Dec 1;26(15):4467-4483.

[128] Jha IS, Sen S, Kumar R. Smart grid development in India-A case study. Eighteenth National Power Systems Conference (NPSC) IEEE. 2014 Dec 18;1-6.

[129] Jian L, Zhao Y, Zhu YP, Zhang MB, Bertolatti D. An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China. Science of the Total Environment. 2012 Jun 1;426:336-345.

[130] Karmakar S, Mujumdar PP. Grey fuzzy optimization model for water quality management of a river system. Advances in Water Resources. 2006 Jul 1;29(7):1088-1105.

[131] Katsiampa P. Volatility estimation for Bitcoin: A comparison of GARCH models. Economics Letters. 2017 Sep 1;158:3-6.

[132] Kavzoglu T, Mather PM. The use of backpropagating artificial neural networks in land cover classification. International Journal of Remote Sensing. 2003 Jan 1;24(23):4907-4938.

[133] Kaygusuz K. Energy services and energy poverty for sustainable rural development. Renewable and sustainable energy reviews. 2011 Feb 1;15(2):936-947.

[134] Kejriwal BK. The Mines Act, 1952, Lovely Prakashan, Dhanbad. 2006.

[135] Kim S, Kisi O, Seo Y, Singh VP, Lee CJ. Assessment of rainfall aggregation and disaggregation using data-driven models and wavelet decomposition. Hydrology Research. 2016 Feb;48(1):99-116.

[136] Kim S, Singh VP. Spatial disaggregation of areal rainfall using two different artificial neural networks models. Water. 2015 Jun;7(6):2707-2727.

[137] Kisi O, Parmar KS, Soni K, Demir V. Modeling of air pollutants using least square support vector regression, multivariate adaptive regression spline, and M5 model tree models. Air Quality, Atmosphere & Health. 2017 Sep 1;10(7):873-883.

[138] Kisi O, Parmar KS. Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. Journal of Hydrology. 2016 Mar 1;534:104-112.

[139] Kisi O, Shiri J. Precipitation forecasting using wavelet-genetic programming and wavelet-neuro fuzzy conjunction models. Water resources management. 2011 Oct 1;25(13):3135-3152.

[140] Kişi Ö. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. Journal of Hydrologic Engineering. 2009 Aug;14(8):773-782.

[141] Kumar J, Kaur A, Manchanda P. Forecasting the time series data using ARIMA with wavelet. Journal of Computer and Mathematical Sciences. 2015 Aug;6(8):430-438.

[142] Kumar M, Parmar KS, Kumar DB, Mhawish A, Broday DM, Mall RK, Banerjee T. Long-term aerosol climatology over Indo-Gangetic Plain: Trend, prediction and potential source fields. Atmospheric Environment. 2018 May 1;180:37-50.

[143] Kumar S. Estimation capabilities of biodiesel production from algae oil blend using adaptive neuro-fuzzy inference system (ANFIS). Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 2020 Apr 2;42(7):909-917.

[144] Kumaramangalam M. Coal industry in India: Nationalisation and tasks ahead. New Delhi: Oxford & IBH Publishing Company; 1973. 77 p.

[145] Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks. 1997 Jan;8(1):98-113.

[146] Liang WM, Wei HY, Kuo HW. Association between daily mortality from respiratory and cardiovascular diseases and air pollution in Taiwan. Environ Res. 2009 Jan;109(1):51–58.

[147] Liousse C, Penner JE, Chuang C, Walton JJ, Eddleman H, Cachier H. A global three dimensional model study of carbonaceous aerosols. Journal of Geophysical Research: Atmospheres. 1996 Aug 27;101(D14):19411-19432.

[148] Lu WX, Zhao Y, Chu HB, Yang LL. The analysis of groundwater levels influenced by dual factors in western Jilin Province by using time series analysis method. Applied Water Science. 2014 Sep 1;4(3):251-260.

[149] Luongvinh D, Kwon Y. Behavioral modeling of power amplifiers using fully recurrent neural networks. In IEEE MTT-S International Microwave Symposium Digest. 2005 Jun 17;1-4.

[150] Mahato MK, Singh PK, Tiwari AK. Evaluation of metals in mine water and assessment of heavy metal pollution index of East Bokaro Coalfield area, Jharkhand, India. Int J Earth Sci Eng. 2014 Aug;7(04):1611-1618.

[151] Maheswaran R, Khosa R. Comparative study of different wavelets for hydrologic forecasting. Computers & Geosciences. 2012 Sep 1;46:284-295.

[152] Makkhan SJ, Parmar KS, Kaushal S, Soni K. Correlation and time-series analysis of black carbon in the coal mine regions of India: a case study. Modeling Earth Systems and Environment. 2020 Feb 4;6(1):659-669.

[153] Makkhan SJ, Parmar KS, Kaushal S, Soni K. Fractal Analysis of Black Carbon in the Coal Mine Regions of India. In Journal of Physics: Conference Series, IOP Publishing. 2020 May 1;1531(1):012072.

[154] Mallat S. A wavelet tour of signal processing. 2nd ed. San Diego, CA: Academic Elsevier; 1999 Sep 14. 620 p.

[155] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE transactions on pattern analysis and machine intelligence. 1989 Jul;11(7):674-693.

[156] Mao H, Zhu F, Cui Y. A generalized mixture integer-valued GARCH model. Statistical Methods & Applications. 2020 Sept;29(3):527-552.

[157] Mao PL, Aggarwal RK. A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network. IEEE Transactions on Power Delivery. 2001 Oct;16(4):654-660.

[158] Maqsood I, Khan MR, Huang GH, Abdalla R. Application of soft computing models to hourly weather analysis in southern Saskatchewan, Canada. Engineering Applications of Artificial Intelligence. 2005 Feb 1;18(1):115-125.

[159] Massey DD, Kulshrestha A, Taneja A. Particulate matter concentrations and their related metal toxicity in rural residential environment of semi-arid region of India. Atmospheric Environment. 2013 Mar 1;67:278-286.

[160] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics. 1943 Dec;5(4):115-133.

[161] Mei Z, Zhang W, Zhang L, Wang D. Real-time multistep prediction of public parking spaces based on Fourier transform–least squares support vector regression. Journal of Intelligent Transportation Systems. 2020 Jan 2;24(1):68-80.

[162] Menon S, Hansen J, Nazarenko L, Luo Y. Climate effects of black carbon aerosols in China and India. Science. 2002 Sep 27;297(5590):2250-2253.

[163] Menon S, Koch D, Beig G, Sahu S, Fasullo J, Orlikowski D. Black carbon aerosols and the third polar ice cap. Atmospheric Chemistry and Physics. 2010 Apr 15;10(LBNL-3596E):4559–4571.

[164] Meyer Y. Ondelettes, function splines, et analyses gradues. Lectures given at the Mathematics Department, University of Torino. 1986.

[165] Meyer Y. Wavelets and Operators: Volume 1. Cambridge university press, Cambridge, UK. 1992.

[166] Meyer Y. Wavelets, Algorithms and Applications. SIAM Publications Philadelphia; 1993.

[167] Meyer Y. Wavelets, vibrations and scalings. American Mathematical Society; 1998.

[168] Mohajan HK. Greenhouse Gas Emissions of China. Journal of Environmental Treatment Techniques. 2014 Feb 20;1(4):190-202.

[169] Mohamadi S, Amindavar H, Hosseini SM. ARIMA-GARCH modeling for epileptic seizure prediction. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA. 2017 Mar 5;994-998.

[170] Mohammadi K, Eslami HR, Kahawita R. Parameter estimation of an ARMA model for river flow forecasting using goal programming. Journal of Hydrology. 2006 Nov 30;331(1-2):293-299.

[171] Mohandes M, Rehman S, Halawani TO. Estimation of global solar radiation using artificial neural networks. Renewable Energy. 1998 May 1;14(1-4):179-184.

[172] Mok J, Park SS, Lim H, Kim J, Edwards DP, Lee J, Yoon J, Lee YG, Koo JH. Correlation analysis between regional carbon monoxide and black carbon from satellite measurements. Atmospheric Research. 2017 Nov 1;196:29-39.

[173] Moosavi V, Vafakhah M, Shirmohammadi B, Behnia N. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. Water Resources Management. 2013 Mar 1;27(5):1301-1321.

[174] Moosavi V, Vafakhah M, Shirmohammadi B, Ranjbar M. Optimization of wavelet-ANFIS and wavelet-ANN hybrid models by Taguchi method for groundwater level forecasting. Arabian Journal for Science and Engineering. 2014 Mar 1;39(3):1785-1796.

[175] Morlet J, Arens G, Fourgeau E, Glard D. Wave propagation and sampling theory-Part I: Complex signal and scattering in multilayered media. Geophysics. 1982 Feb;47(2):203-221.

[176] Muhammad Adnan R, Yuan X, Kisi O, Yuan Y, Tayyab M, Lei X. Application of soft computing models in streamflow forecasting. In Proceedings of the Institution of Civil Engineers Water Management. 2019 June;172(3):123-134.

[177] Nagendra SS, Khare M. Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. Ecological Modelling. 2006 Jan 10;190(1-2):99-115.

[178] Nalley D, Adamowski J, Khalil B. Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). Journal of Hydrology. 2012 Dec 19;475:204-228.

[179] Narayanan P, Basistha A, Sarkar S, Kamna S. Trend analysis and ARIMA modelling of pre monsoon rainfall data for western India. Comptes Rendus Geoscience. 2013 Jan 1;345(1):22-27.

[180] Nasr AB, Boutahar M, Trabelsi A. Fractionally integrated time varying GARCH model. Statistical Methods & Applications. 2010 Aug 1;19(3):399-430.

[181] National Carbonaceous Aerosols Programme (NCAP): Science Plan. Black carbon research initiative. Ministry of Earth Sciences, Government of India. 2011. 40 p.

[182] Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS. A neuro-fuzzy computing technique for modeling hydrological time series. Journal of Hydrology. 2004 May 31;291(1-2):52-66.

[183] Nelson DB. Conditional heteroskedasticity in asset returns: A new approach. Econometrica. 1991 Mar 1;59(2):347-370.

[184] Nourani V, Andalib G, Dąbrowska D. Conjunction of wavelet transform and SOM-mutual information data pre-processing approach for AI-based Multi-Station nitrate modeling of watersheds. Journal of Hydrology. 2017 May 1;548:170-183.

[185] Nourani V, Andalib G. Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. Journal of Mountain Science. 2015 Jan 1;12(1):85-100.

[186] Nourani V, Baghanam AH, Adamowski J, Gebremichael M. Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. Journal of Hydrology. 2013 Jan 7;476:228-243.

[187] Nourani V, Baghanam AH, Adamowski J, Kisi O. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. Journal of Hydrology. 2014 Jun 6;514:358-377.

[188] Nourani V, Farboudfam N. Rainfall time series disaggregation in mountainous regions using hybrid wavelet-artificial intelligence methods. Environmental Research. 2019 Jan 1;168:306-318.

[189] Novakov T, Menon S, Kirchstetter TW, Koch D, Hansen JE. Aerosol organic carbon to black carbon ratios: Analysis of published data and implications for climate forcing. Journal of Geophysical Research: Atmospheres. 2005 Nov 16;110:D21205.

[190] Nury AH, Hasan K, Alam MJ. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. Journal of King Saud University Science. 2017 Jan 1;29(1):47-61.

[191] Nury AH, Sharma A, Marshall L, Mehrotra R. Characterising uncertainty in precipitation downscaling using a Bayesian approach. Advances in Water Resources. 2019 Jul 1;129:189-197.

[192] Oh H, Lee S. On score vector-and residual-based CUSUM tests in ARMA–GARCH models. Statistical Methods & Applications. 2018 Aug 10;27(3):385-406.

[193] Orhan M, Köksal B. A comparison of GARCH models for VaR estimation. Expert Systems with Applications. 2012 Feb 15;39(3):3582-3592.

[194] Pandey PK, Tripura H, Pandey V. Improving Prediction Accuracy of Rainfall Time Series By Hybrid SARIMA–GARCH Modeling. Natural Resources Research. 2019 Jul 1;28(3):1125-1138.

[195] Pant P, Guttikunda SK, Peltier RE. Exposure to particulate matter in India: A synthesis of findings and future directions. Environmental Research. 2016 May 1;147:480-496.

[196] Pant P, Shukla A, Kohl SD, Chow JC, Watson JG, Harrison RM. Characterization of ambient $PM_{2.5}$ at a pollution hotspot in New Delhi, India and inference of sources. Atmospheric Environment. 2015 May 1;109:178-189.

[197] Pappas SS, Ekonomou L, Karampelas P, Karamousantas DC, Katsikas SK, Chatzarakis GE, Skafidas PD. Electricity demand load forecasting of the Hellenic power system using an ARMA model. Electric Power Systems Research. 2010 Mar 1;80(3):256-264.

[198] Parmar KS, Bhardwaj R. Analysis of Water parameters using Haar wavelet (level 3). Int J Curr Eng Technol. 2012;2(1):166-171.

[199] Parmar KS, Bhardwaj R. River water prediction modeling using neural networks, fuzzy and wavelet coupled model. Water Resources Management. 2015 Jan 1;29(1):17-33.

[200] Parmar KS, Bhardwaj R. Statistical, time series, and fractal analysis of full stretch of river Yamuna (India) for water quality management. Environmental Science and Pollution Research. 2015 Jan 1;22(1):397-414.

[201] Parmar KS, Bhardwaj R. Water quality index and fractal dimension analysis of water parameters. International Journal of Environmental Science and Technology. 2013 Jan 1;10(1):151-164.

[202] Parmar KS, Bhardwaj R. Water quality management using statistical analysis and time-series prediction model. Applied Water Science. 2014 Dec 1;4(4):425-434.

[203] Parmar KS, Bhardwaj R. Wavelet and statistical analysis of river water quality parameters. Applied Mathematics and Computation. 2013 Jun 15;219(20):10172-10182.

[204] Parmar KS, Makkhan SJ, Kaushal S. Neuro-fuzzy-wavelet hybrid approach to estimate the future trends of river water quality. Neural Computing and Applications. 2019 Dec 1;31(12):8463-8473.

[205] Partal T, Kişi Ö. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. Journal of Hydrology. 2007 Aug 15;342(1-2):199-212.

[206] Partovian A, Nourani V, Alami MT. Hybrid denoising-jittering data processing approach to enhance sediment load prediction of muddy rivers. Journal of Mountain Science. 2016 Dec 1;13(12):2135-2146

[207] Pasanen L, Holmström L. Scale space multiresolution correlation analysis for time series data. Computational Statistics. 2017 Mar 1;32(1):197-218.

[208] Paul RK, Paul AK, Bhar LM. Wavelet-based combination approach for modeling sub-divisional rainfall in India. Theoretical and Applied Climatology. 2020 Feb 1;139(3-4):949-963.

[209] Paul S, Ali M, Chatterjee R. Prediction of compressional wave velocity using regression and neural network modeling and estimation of stress orientation in Bokaro Coalfield, India. Pure and Applied Geophysics. 2018 Jan 1;175(1):375-388.

[210] Penner JE, Eddleman H, Novakov T. Towards the development of a global inventory for black carbon emissions. Atmospheric Environment. Part A. General Topics. 1993 Jun 1;27(8):1277-1295.

[211] Pentoś K, Pieczarka K, Lejman K. Application of Soft Computing Techniques for the Analysis of Tractive Properties of a Low-Power Agricultural Tractor under Various Soil Conditions. Complexity. 2020 Jan 10;2020. https://doi.org/10.1155/2020/7607545

[212] Planning Commission. Government of India; 2007. Planning Commission Government of India Tenth Five Year Plan, New Delhi. 2002-2007.

[213] Portnov BA, Dubnov J, Barchana M. Studying the association between air pollution and lung cancer incidence in a large metropolitan area using a kernel density function. Socio Econ. Plan.Sci. 2009 Aug; 43(3):141-150.

[214] Poul AK, Shourian M, Ebrahimi H. A comparative study of MLR, KNN, ANN and ANFIS models with wavelet transform in monthly stream flow prediction. Water Resources Management. 2019 Jun 15;33(8):2907-2923.

[215] Prasad BM. Second World War and Indian Industry, 1939-45: A Case Study of the Coal Industry in Bengal and Bihar. Anamika Prakashan; 1992. p. 232-234.

[216] Purohit P, Amann M, Mathur R, Gupta I, Marwah S, Verma V, Bertok I, Borken-Kleefeld J, Chambers A, Cofala J, Heyes C. GAINS ASIA: Scenarios for cost-effective control of air pollution and greenhouse gases in India. IIASA, Laxenburg, Austria. Nov 2010. 62 p. http://pure.iiasa.ac.at/id/eprint/9379/

[217] Qian G, Zhao X. On time series model selection involving many candidate ARMA models. Computational Statistics & Data Analysis. 2007 Aug 15;51(12):6180-6196.

[218] Quiroz R, Yarlequé C, Posadas A, Mares V, Immerzeel WW. Improving daily rainfall estimation from NDVI using a wavelet transform. Environmental Modelling & Software. 2011 Feb 1;26(2):201-209.

[219] Raghuvanshi SP, Chandra A, Raghav AK. Carbon dioxide emissions from coal based power generation in India. Energy Conversion and Management. 2006 Mar 1;47(4):427-441.

[220] Rahman MJ, Hasan MA. Performance of wavelet transform on models in forecasting climatic variables. InComputational Intelligence Techniques in Earth and Environmental Sciences Springer, Dordrecht; 2014. p. 141-154. http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-94-017-8642-3_8

[221] Rajaee T. Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. Science of the total environment. 2011 Jul 1;409(15):2917-2928.

[222] Ram K, Sarin MM, Tripathi SN. A 1 year record of carbonaceous aerosols from an urban site in the Indo-Gangetic Plain: Characterization, sources, and temporal variability. Journal of Geophysical Research: Atmospheres. 2010 Dec 27;115:D24313.

[223] Ramanathan V, Carmichael G. Global and regional climate changes due to black carbon. Nature Geoscience. 2008 Apr;1(4):221-227.

[224] Ramanathan V, Ramana MV, Roberts G, Kim D, Corrigan C, Chung C, Winker D. Warming trends in Asia amplified by brown cloud solar absorption. Nature. 2007 Aug;448(7153):575-578.

[225] Rana A, Jia S, Sarkar S. Black carbon aerosol in India: a comprehensive review of current status and future prospects. Atmospheric Research. 2019 Apr 1;218:207-230.

[226] Rao KK, Raju GS. An overview on soft computing techniques. In International Conference on High Performance Architecture and Grid Computing. Communications in Computer and Information Science, vol 169. Springer, Berlin, Heidelberg. 2011 Jul 19, p. 9-23. https://doi.org/10.1007/978-3-642-22577-2_2

[227] Rashid MM, Beecham S, Chowdhury RK. Statistical downscaling of CMIP5 outputs for projecting future changes in rainfall in the Onkaparinga catchment. Science of the Total Environment. 2015 Oct 15;530:171-182.

[228] Ratnadip A, Agrawal RK. An Introductory Study on Time Series Modeling and Forecasting. LAP LAMBERT Academic Publishing, Germany ISBN: 978-3-659-33508-2; 2013 Feb 26. 67 p.

[229] Ratnam DV, Otsuka Y, Sivavaraprasad G, Dabbakuti JK. Development of multivariate ionospheric TEC forecasting algorithm using linear time series model and ARMA over low latitude GNSS station. Advances in Space Research. 2019 May 1;63(9):2848-2856.

[230] Razali JB, Mohamad AM. Modeling and forecasting price volatility of crude palm oil and sarawak black pepper using ARMA and GARCH model. Advanced Science Letters. 2018 Dec 1;24(12):9327-9330.

[231] Rehman IH, Ahmed T, Praveen PS, Kar A, Ramanathan V. Black carbon emissions from biomass and fossil fuels in rural India. Atmospheric Chemistry & Physics Discussions. 2011 Apr 1;11(4):7289-7299.

[232] Relvas H, Miranda AI. An urban air quality modeling system to support decision-making: design and implementation. Air Quality, Atmosphere & Health. 2018 Aug 1;11(7):815-824.

[233] Ribeiro A, Seruca I, Durão N. Improving organizational decision support: Detection of outliers and sales prediction for a pharmaceutical distribution company. Procedia Computer Science. 2017 Jan 1;121:282-290.

[234] Rojas I, Valenzuela O, Rojas F, Guillén A, Herrera LJ, Pomares H, Marquez L, Pasadas M. Soft computing techniques and ARMA model for time series prediction. Neurocomputing. 2008 Jan 1;71(4-6):519-537.

[235] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review. 1958 Nov;65(6):386-408.

[236] Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Medicine. 2020 May;46(5):846-848.

[237] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986 Oct;323(6088):533-536.

[238] Saâdaoui F, Rabbouch H. A wavelet-based hybrid neural network for short-term electricity prices forecasting. Artificial Intelligence Review. 2019 Jun 1;52(1):649-669.

[239] Sachindra DA, Huang F, Barton A, Perera BJ. Least square support vector and multi-linear regression for statistically downscaling general circulation model outputs to catchment streamflows. International Journal of Climatology. 2013 Apr;33(5):1087-1106.

[240] Sahay RR, Srivastava A. Predicting monsoon floods in rivers embedding wavelet transform, genetic algorithm and neural network. Water Resources Management. 2014 Jan 1;28(2):301-317.

[241] Sahu SK, Bhangare RC, Ajmal PY, Sharma S, Pandit GG, Puranik VD. Characterization and quantification of persistent organic pollutants in fly ash from coal fueled thermal power stations in India. Microchemical Journal. 2009 May 1;92(1):92-96.

[242] Sahu V, Dewangan P, Mishra R, Jhariya DC. Opencast Coal Mining At Large Depth In India-Challenges Ahead. World Journal of Engineering Research and Technology. 2017 Apr 6;3(3):201-211.

[243] Salazar L, Nicolis O, Ruggeri F, Kisel'ák J, Stehlík M. Predicting hourly ozone concentrations using wavelets and ARIMA models. Neural Computing and Applications. 2019 Aug 1;31(8):4331-4340.

[244] Santoso S, Grady WM, Powers EJ, Lamoree J, Bhatt SC. Characterization of distribution power quality events with Fourier and wavelet transforms. IEEE Transactions on Power Delivery. 2000 Jan;15(1):247-254.

[245] Santra S, Bagaria N. Labour productivity in coal mining sector in India: with special to major coal mining states. Researchjournali's Journal of Human Resource. 2014 Jan 20;2(1):1-14. http://mpra.ub.uni-muenchen.de/53519/

[246] Saraf AK, Prakash A, Sengupta S, Gupta RP. Landsat-TM data for estimating ground temperature and depth of subsurface coal fire in the Jharia coalfield, India. International Journal of Remote Sensing. 1995 Aug 1;16(12):2111-2124.

[247] Saud T, Mandal TK, Gadi R, Singh DP, Sharma SK, Saxena M, Mukherjee A. Emission estimates of particulate matter (PM) and trace gases (SO_2, NO and NO_2) from biomass fuels used in rural sector of Indo-Gangetic Plain, India. Atmospheric Environment. 2011 Oct 1;45(32):5913-5923.

[248] Schwarz JP, Gao RS, Spackman JR, Watts LA, Thomson DS, Fahey DW, Ryerson TB, Peischl J, Holloway JS, Trainer M, Frost GJ, et al. Measurement of the mixing state, mass, and optical size of individual black carbon particles in urban and biomass burning emissions. Geophysical Research Letters. 2008 Jul;35(13):L13810.

[249] Senapati MR. Fly ash from thermal power plants–waste management and overview. Current Science. 2011 Jun 25;100(25):1791-1794.

[250] Shakti SP, Hassan MK, Zhenning Y, Caytiles RD, SN IN. Annual Automobile Sales Prediction Using ARIMA Model. International Journal of Hybrid Information Technology. 2017;10(6):13-22.

[251] Sharma S, Saxena A, Saxena N. Unconventional Resources in India: The Way Ahead. Springer International Publishing. 2019 Sept 16;VII, 82 p. doi: 10.1007/978-3-030-21414-2.

[252] Shetty DK, Sumithra, Ismail B. Hybrid SARIMA-GARCH Model for Forecasting Indian Gold Price. Research Review International Journal of Multidisciplinary. 2018 Aug 7;3(8):263-269.

[253] Shih WC. Energy security, GATT/WTO, and regional agreements. Natural Resources Journal. 2009 Apr 1;49(2):433-484.

[254] Siddiquee MS, Hossain MM. Development of a sequential Artificial Neural Network for predicting river water levels based on Brahmaputra and Ganges water levels. Neural Computing and Applications. 2015 Nov 1;26(8):1979-1990.

[255] Singh J, Gu S. Biomass conversion to energy in India—A critique. Renewable and sustainable energy reviews. 2010 Jun 1;14(5):1367-1378.

[256] Singh KP, Malik A, Mohan D, Sinha S. Multivariate statistical techniques for the evaluation of spatial and temporal Environ Sci Pollut Res variations in water quality of Gomti River (India)—a case study. Water Res. 2004;38(18):3980-3992.

[257] Singh PK, Tiwari AK, Mahato MK. Qualitative assessment of surface water of West Bokaro Coalfield, Jharkhand by using water quality index method. Int J Chem Tech Res. 2013;5(5):2351-2356.

[258] Singh S, Parmar KS, Kumar J, Kaur J. ARIMA-Wavelet Coupled Approach for Time Series Analysis. International Journal of Scientific Research and Review. 2019 Mar;7(3):3743-3756.

[259] Singh S, Parmar KS, Kumar J, Makkhan SJ. Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. Chaos, Solitons & Fractals. 2020 May 11;135:109866.

[260] Singh SK, Bajpai VK, Garg TK. Measuring productivity change in Indian coal-fired electricity generation: 2003-2010. International Journal of Energy Sector Management. 2013 Apr 5;7(1):46-64.

[261] Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renewable Energy. 2019 Jun 1;136(C):758-768.

[262] Soltani S, Modarres R, Eslamian SS. The use of time series modeling for the determination of rainfall climates of Iran. International Journal of Climatology: A Journal of the Royal Meteorological Society. 2007 May;27(6):819-829.

[263] Soni K, Kapoor S, Parmar KS. Long-term aerosol characteristics over eastern, southeastern, and south coalfield regions in India. Water, Air, & Soil Pollution. 2014 Jan 1;225(1):1832.

[264] Soni K, Parmar KS, Agrawal S. Modeling of air pollution in residential and industrial sites by integrating statistical and Daubechies wavelet (level 5) analysis. Modeling Earth Systems and Environment. 2017 Sep 1;3(3):1187-1198.

[265] Soni K, Parmar KS, Kapoor S. Time series model prediction and trend variability of aerosol optical depth over coal mines in India. Environmental Science and Pollution Research. 2015 Mar 1;22(5):3652-3671.

[266] Sovacool BK. Conceptualizing urban household energy use: Climbing the "Energy Services Ladder". Energy Policy. 2011 Mar 1;39(3):1659-1668.

[267] Srivastava AK, Bisht DS, Ram K, Tiwari S, Srivastava MK. Characterization of carbonaceous aerosols over Delhi in Ganga basin: seasonal variability and possible sources. Environmental Science and Pollution Research. 2014 Jul 1;21(14):8610-8619.

[268] Srivastava AK. Coal Mining Industry in India. Deep & Deep Publications, New Delhi. 1988 Jan 1. 167 p.

[269] Srivastava VK, Mitra D. Study of drainage pattern of Raniganj Coalfield (Burdwan District) as observed on Landsat-TM/IRS LISS II imagery. Journal of the Indian Society of Remote Sensing. 1995 Dec 1;23(4):225-235.

[270] Stieb DM, Chen L, Eshoul M, Judek S. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. Environmental Research. 2012 Aug 1;117:100-111.

[271] Storti G, Vitale C. BL-GARCH models and asymmetries in volatility. Statistical Methods and Applications. 2003 Feb 1;12(1):19-39.

[272] Storti G. Modelling asymmetric volatility dynamics by multivariate BL-GARCH models. Statistical Methods and Applications. 2008 May 1;17(2):251-274.

[273] Su S, Li D, Zhang Q, Xiao R, Huang F, Wu J. Temporal trend and source apportionment of water pollution in different functional zones of Qiantang River, China. Water Research. 2011 Feb 1;45(4):1781-1795.

[274] Subbaraya BH, Jayaraman A, Krishnamoorthy K, Mohan M. Atmospheric aerosol studies under ISRO's Geosphere Biosphere Programme. J. Indian Geophys. Union. 2000;4(1):77-90.

[275] Sun S, Zhu J, Zhou X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. Nature Methods. 2020 Feb;17(2):193-200.

[276] Tao WK, Chen JP, Li Z, Wang C, Zhang C. Impact of aerosols on convective clouds and precipitation. Reviews of Geophysics. 2012 Jun;50(2):1-62.

[277] Tiwari S, Hopke PK, Thimmaiah D, Dumka UC, Srivastava AK, Bisht DS, Rao PS, Chate DM, Srivastava MK, Tripathi SN. Nature and sources of ionic species in precipitation across the Indo Gangetic Plains, India. Aerosol and Air Quality Research. 2015;16(4):943-957.

[278] Toprak ZF, Eris E, Agiralioglu N, Cigizoglu HK, Yilmaz L, Aksoy H, Coskun HG, Andic G, Alganci U. Modeling monthly mean flow in a poorly gauged basin by fuzzy logic. Clean–Soil, Air, Water. 2009 Jul;37(7):555-564.

[279] Toprak ZF, Sen Z, Savci ME. Comment on "Longitudinal dispersion coefficients in natural channels". Water Research. 2004 Jul;38(13):3139-3143.

[280] Trapero JR, Cardos M, Kourentzes N. Empirical safety stock estimation based on kernel and GARCH models. Omega. 2018 May 8;000:1-13.

[281] Tripathi L, Mishra AK, Dubey AK, Tripathi CB, Baredar P. Renewable energy: An overview on its contribution in current energy scenario of India. Renewable and Sustainable Energy Reviews. 2016 Jul 1;60:226-233.

[282] Tularam GA, Saeed T. Oil-price forecasting based on various univariate time-series models. American Journal of Operations Research. 2016 May 10;6(3):226-235.

[283] Tzanis C, Varotsos CA. Tropospheric aerosol forcing of climate: a case study for the greater area of Greece. International Journal of Remote Sensing. 2008 May 1;29(9):2507-2517.

[284] Underwood FM. Describing seasonal variability in the distribution of daily effective temperatures for 1985–2009 compared to 1904–1984 for De Bilt, Holland. Meteorological Applications. 2013 Dec;20(4):394-404.

[285] Varotsos C. Airborne measurements of aerosol, ozone, and solar ultraviolet irradiance in the troposphere. Journal of Geophysical Research: Atmospheres. 2005 May 16;110(D9). doi: 10.1029/2004JD005397

[286] Ventura LM, de Oliveira Pinto F, Soares LM, Luna AS, Gioda A. Forecast of daily $PM_{2.5}$ concentrations applying artificial neural networks and Holt–Winters models. Air Quality, Atmosphere & Health. 2019 Mar 11;12(3):317-325.

[287] Virginia E, Ginting J, Elfaki FA. Application of GARCH Model to Forecast Data and Volatility of Share Price of Energy (Study on Adaro Energy Tbk, LQ45). International Journal of Energy Economics and Policy. 2018;8(3):131-140.

[288] Wang X, Zhang N, Chen Y, Zhang Y. Short-term forecasting of urban rail transit ridership based on ARIMA and wavelet decomposition. In AIP Conference Proceedings 2018 May 23;1967(1):040025.

[289] White H. Learning in artificial neural networks: A statistical perspective. Neural Computation. 1989 Dec;1(4):425-464.

[290] Williams RJ, Zipser D. A learning algorithm for continually running fully recurrent neural networks. Neural Computation. 1989 Jun;1(2):270-280.

[291] World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organization; 2009.

[292] Xu B, Cao J, Hansen J, Yao T, Joswia DR, Wang N, Wu G, Wang M, Zhao H, Yang W, Liu X. Black soot and the survival of Tibetan glaciers. Proceedings of the National Academy of Sciences. 2009 Dec 29;106(52):22114-22118.

[293] Yao S, Hu S, Zhao Y, Zhang A, Abdelzaher T. Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In Proceedings of the 26th International Conference on World Wide Web. 2017 Apr 3;351-360.

[294] Yenigun K, Ecer R. Overlay mapping trend analysis technique and its application in Euphrates Basin, Turkey. Meteorological Applications. 2013 Dec;20(4):427-438.

[295] Yeon IS, Jun KW, Lee HJ. The improvement of total organic carbon forecasting using neural networks discharge model. Environmental Technology. 2009 Jan 1;30(1):45-51.

[296] Zadeh LA. Utilization of fuzzy evidence in decision analysis. In ORSA/TIMS Joint Meeting Houston 1981 Oct.

[297] Zeinalnezhad M, Chofreh AG, Goni FA, Klemeš JJ. Air pollution prediction using semi experimental regression model and Adaptive Neuro-Fuzzy Inference System.

Journal of Cleaner Production. 2020 Jul 10;261:121218. doi: 10.1016/j.jclepro.2020.121218

[298] Zeitlberger AC, Brauneis A. Modeling carbon spot and futures price returns with GARCH and Markov switching GARCH models. Central European Journal of Operations Research. 2016 Mar 1;24(1):149-176.

[299] Zhang H, Wang X, Cao J, Tang M, Guo Y. A multivariate short-term traffic flow forecasting method based on wavelet analysis and seasonal time series. Applied Intelligence. 2018 Oct 1;48(10):3827-3838.

[300] Zhang J, Wang C. Application of ARMA model in ultra-short term prediction of wind power. International Conference on Computer Sciences and Applications 2013 Dec 14;361-364.

[301] Zhang X, Liang F, Yu B, Zong Z. Explicitly integrating parameter, input, and structure uncertainties into Bayesian Neural Networks for probabilistic hydrologic forecasting. Journal of Hydrology. 2011 Nov 9;409(3-4):696-709.

[302] Zhao CL, Wang B. Forecasting crude oil price with an autoregressive integrated moving average (ARIMA) model. Advances in Intelligent Systems and Computing. 2014 Oct;211:275-286.

[303] Zhao F, Yang M, Batista JR. Using Fuzzy Models and Time Series Analysis to Predict Water Quality. International Journal of Intelligent Systems and Applications. 2020 Apr 1;12(2):1-10.

[304] Zhao Y, Qi C. Study on openness of Chinese stock market based on wavelet analysis. International Conference on Business Intelligence and Financial Engineering (BIFE 2008) 2008 Jan. p. 736-741.

[305] Zou C, Zhao Q, Zhang G, Xiong B. Energy revolution: From a fossil energy era to a new energy era. Natural Gas Industry. 2016 Jan;3(1):1–11.

APPENDICES

Appendix I:

List of Publications (Published Papers)

1. Makkhan SJ, Parmar KS, Kaushal S, Soni K. Correlation and time-series analysis of black carbon in the coal mine regions of India: a case study. Modeling Earth Systems and Environment. 2020 Feb 4;6(2):659–669 *(SCI Indexed, Springer)*.
2. Makkhan SJ, Parmar KS, Kaushal S, Soni K. Fractal analysis of black carbon in the coal mine regions of India. In Journal of Physics: Conference Series, IOP Publishing. 2020 May 1;1531(1):p. 012072. *(SCOPUS Indexed).*
3. Parmar KS, Makkhan SJ, Kaushal S. Neuro-fuzzy-wavelet hybrid approach to estimate the future trends of river water quality. Neural Computing and Applications. 2019 Dec 1;31(12):8463-8473. *(SCI Indexed, Springer)*.
4. Makkhan SJ, Parmar KS, Kaushal S, Soni K. Statistical and time-series analysis of black carbon in the major coal mines of India. Jnanabha. 2020 Dec;50(2):30-37. (*UGC-CARE List Journal, VPI*).

Appendix II:

Communicated Papers (In SCI/UGC approved journals)

1. Forecasting Volatility Using GARCH Models in the Coal Mines of India.
2. Wavelets and Time Series Analysis of Black Carbon in the Major Coal Mines of India.

Appendix III:

National and International Conferences/Workshops/Seminars

1. Participated in 21st Annual Conference of Vijnana Parishad of India (VPI) on "Modeling, Optimization and Computing for Technological and Sustainable Development (MOCTSD-2019)" at SRMIST, Delhi NCR Campus Ghaziabad on

April 26-28, 2019 and presented a paper entitled “Statistical and Time Series Analysis of Black Carbon in the Major Coal Mines of India”.

2. Participated in a National Seminar sponsored by CSTT, MHRD, Government of India on “Applications of Technical Terminology in Modeling, Optimization and Computing in Science & Technology” at SRMIST, Ghaziabad on April 26-28, 2019 and presented a paper entitled “भारत के प्रमुख कोयला खानो में ब्लैक कार्बन का सांख्यिकीय और समय श्रृंखला विश्लेषण”.
3. Participated in International Conference sponsored by DRDO on “Functional Materials and Simulation Techniques (ICFMST-2019)” at Chandigarh University, Gharuan, Mohali on June 7-8, 2019 and presented a paper entitled “A comparison of Black Carbon using the Time series and Correlation analysis along the Coal Mines of India”.
4. Participated in International Conference on “Recent Advances in Fundamental and Applied Sciences (RAFAS-2019)” held at Lovely Professional University, Phagwara on November 5-6, 2019 and presented a paper entitled “Fractal Analysis of Black carbon in the Coal mine regions of India”.
5. Participated in a National Seminar sponsored by CSTT, MHRD, Government of India on “Applications of Technical Terminology in Modeling, Optimization and Computing in Science & Technology” at SRMIST, Bareilly, Uttar Pradesh on November 15-16, 2019 and presented a paper entitled “Forecasting Volatility using GARCH models in the Coal mines of India”.
6. Participated in one week ‘National Centre for Mathematics TEW’ workshop sponsored by TIFR, Mumbai on “Computation Methods for Solving Differential Equations” at I.K. Gujral Punjab Technical University, Jalandhar on December 25-30, 2019.