

**HYDRO-BIOCHEMICAL PROCESSES AND RIVER WATER
QUALITY SIMULATION USING MULTIVARIATE STATISTICS
AND SOFT COMPUTING TECHNIQUES**

SAMEER ARORA



**DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI**

NOVEMBER 2021

© Indian Institute of Technology Delhi (IITD), New Delhi, 2021

**HYDRO-BIOCHEMICAL PROCESSES AND RIVER WATER
QUALITY SIMULATION USING MULTIVARIATE STATISTICS
AND SOFT COMPUTING TECHNIQUES**

by

SAMEER ARORA

Department of Civil Engineering

Submitted

**in fulfilment of the requirements of the degree of Doctor of Philosophy
to the**



INDIAN INSTITUTE OF TECHNOLOGY DELHI

NOVEMBER 2021

CERTIFICATE

This is to certify that the thesis entitled “**Hydro-biochemical processes and river water quality simulation using multivariate statistics and soft computing techniques**” submitted by **Sameer Arora (2013CEZ8073)** to the Indian Institute of Technology Delhi for the award of the degree of **Doctor of Philosophy** is a bonafide record of research work carried out by him under my supervision and guidance. The thesis work, in my opinion, has reached the requisite standard fulfilling the requirements for the degree of **Doctor of Philosophy**.

The results contained in this thesis have not been submitted, in part or full, to any other University or Institute for the award of any degree or diploma.

Prof. (Dr.) Ashok K. Keshari
Professor
Department of Civil Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016, India

ACKNOWLEDGMENTS

The research work is a collaboration of encouragement, motivation, and hard work. I have been inspired and motivated from various people and would like to express my gratitude to all of them. First and foremost, I would like to express my deepest gratitude to my guide and mentor Prof. Ashok K. Keshari, Department of Civil Engineering who stimulated me to work on this topic. His constant support and clear vision, inspired and motivated me throughout the journey. His resourcefulness and critical feedbacks have greatly assisted me in this work. The vast experience of my supervisor in the field of water resource engineering, enhanced my knowledge and confidence to perform in this field and undertaking research.

My thanks are also due to the student research committee (SRC) members, Prof. Manoj Datta (Chairperson), Prof Arvind Nema (Departmental Expert) and Prof. V Haridas (Subject Expert) for their valuable comments and suggestions.

I want to record my gratitude to the faculty members of Civil Engineering Department for their help and giving me an opportunity to carry out this research work. I would also like to appreciate the staff of Civil Engineering Department especially Mr. N R Gehlot for providing me helpful co-operation throughout the completion of this research work.

IIT Delhi become second home during this tenure and met with family members like friends since I joined here. I would like to acknowledge all PhD fellows with whom I shared social life and experience. I am also very grateful to my friends, Dr. Nitesh Patidar, Mr. Gopinadh Rongali and Mr. Amar Landage, for providing rejuvenating tea sessions and period of productive thinking.

Lastly, I am fortunate and grateful to my parents, Mrs. Savita Arora and Mr. J S Arora for everything that I have accomplished in life with their blessings, inspiration and support. Without their encouragement, support and sacrifice my study would not have been possible. I

am thankful to my sisters and wife for her continuous support in my ups and downs. Thanks to all of you, I really appreciate everything that you have done for me.

Sameer Arora

ABSTRACT

River dynamics involve composite interactions among flow, nutrients, transported material and channel type. Spatial and temporal variations are generally observed in biochemical parameters and concentration of dissolved chemicals in freshwater bodies, corresponding to its flow through different climatic, geographical and topographic environment. Discharge of domestic and industrial effluents from multiple sources accelerate the rate of variation in biochemical characterisation of water body. Inherent ability of water flow to assimilate organic load allow the self-purification of streams. Self-purification is an intricate phenomenon, comprising of physical, chemical and biological processes occur simultaneously to recover the native state of water body. The re-aeration capacity of the stream defines the net rate of oxygen transfer from atmosphere to air-water interface and depends significantly upon the degree of turbulence of the flowing stream, as the interface is constantly changing with the flow rate due to the mixing of the flowing water. The indirect measurement of re-oxygenation is based on the oxygen deficit of river, which is calculated as the variance between the dissolved oxygen content and the native oxygen concentration at specific temperature.

The assimilation capacity of the stream can be measured by estimating several indices like biota of the river, dissolved oxygen, temperature, and amount of flow, velocity, concentration of suspended and dissolved solids. Determination of re-aeration coefficient during re-oxygenation process in water body should depends on reasonable valuation of oxygen transfer rate across air-water interface. For years, several attempts have been made by researchers for envisaging the rate of oxygen transfer from defined physical and hydraulic properties of the river. However, the functional relationships between the hydraulic variables and the re-aeration coefficient are limited to specific boundary conditions. The relative

significance of variables used to design the re-aeration equations such as depth, velocity, slope, and width are also not found reliable in the literature and applicable to the streams free from any organic load. The several predictive re-aeration models have been developed and applied on confined study areas having tremendous variability of channel types causes significant difference in the results of each model.

The improvement in the estimation of an adequate re-aeration coefficient is a vital requirement to simulate the river water quality. The hierarchical agglomerative cluster analysis (HACA), principal component analysis (PCA) and structural equation modelling (SEM) are employed to examine the similarities in water quality parameters, critical parameters and their indirect relationships. The combination of these statistical approaches helps to identify the role of different parameters in variation of water quality individually and to identify the significant source of pollutants. In this study, 17 predictive re-aeration models have been compared to reveal their applicability to the study area. Yamuna River has been selected as the study area, receiving the high organic load from anthropogenically influenced multiple point sources, which significantly affects the flow rate of the river. The river receives the wastewater from 17 major and minor drains, which causes significant fluctuations in the river physio-chemical properties. The error in the predictive re-aeration models is estimated using MAE, MAPE, MSE and MME; however, none of the re-aeration model gives promising results with the estimated values of the re-aeration coefficient.

The artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and autoregressive integrated moving average (ARIMA) have been applied to develop a new model for the estimation of the re-aeration coefficient. The ten different configurations of input parameters were used to design various ANN and ANFIS models, and the performance of models are assessed using R , R^2 and RMSE. The model with least RMSE and highest R^2 from ANN and ANFIS is further integrated with ARIMA to improve the estimation of the re-aeration

coefficient. The ANFIS-ARIMA model improves the estimation of the re-aeration coefficient for the degraded river in comparison to predictive re-aeration models.

The simulation of river water quality is performed using QUAL2Kw integrated with the ANFIS-ARIMA model to identify the impact of newly developed model for the estimation of re-aeration coefficient. The sensitivity analysis of the DO and the BOD is also carried out to examine the variations in parameters concentration and related river characteristics. The study reveals that the developed ANFIS-ARIMA model improves the estimation of re-aeration coefficient significantly and can be used in the prediction of water quality. The study would be useful to the decision-makers while structuring the pollution abatement methods, developing river water management policies and defining the level of treatment that can be focused on the removal of sensitive contaminants or critical species that severely affect the river health.

बहुभिन्नरूपी सांख्यिकी और सॉफ्ट कंप्यूटिंग तकनीकों का उपयोग करते हुए

जल-जैव रासायनिक प्रक्रियाएं और नदी जल गुणवत्ता सिमुलेशन

सार

नदी की गतिशीलता में प्रवाह, पोषक तत्वों, परिवहन सामग्री और चैनल प्रकार के बीच समग्र परस्पर क्रिया शामिल है। विभिन्न जलवायु, भौगोलिक और स्थलाकृतिक वातावरण के माध्यम से इसके प्रवाह के अनुरूप, मीठे पानी के निकायों में जैव रासायनिक मानकों और भंग रसायनों की एकाग्रता में स्थानिक और अस्थायी भिन्नताएं आम तौर पर देखी जाती हैं। कई स्रोतों से घरेलू और औद्योगिक अपशिष्टों का निर्वहन जल निकाय के जैव रासायनिक लक्षण वर्णन में भिन्नता की दर को तेज करता है। कार्बनिक भार को आत्मसात करने के लिए जल प्रवाह की अंतर्निहित क्षमता धाराओं की आत्म-शुद्धि की अनुमति देती है। आत्म-शुद्धि एक जटिल घटना है, जिसमें भौतिक, रासायनिक और जैविक प्रक्रियाएं शामिल हैं जो जल निकाय की मूल स्थिति को ठीक करने के लिए एक साथ होती हैं। धारा की पुनः वातन क्षमता वायुमंडल से वायु-जल अंतरापृष्ठ में ऑक्सीजन स्थानांतरण की शुद्ध दर को परिभाषित करती है और बहने वाली धारा की अशांति की डिग्री पर महत्वपूर्ण रूप से निर्भर करती है, क्योंकि मिश्रण के कारण प्रवाह दर के साथ इंटरफ़ेस लगातार बदल रहा है बहता पानी। पुनः ऑक्सीकरण का अप्रत्यक्ष माप नदी के ऑक्सीजन की कमी पर आधारित है, जिसकी गणना विशिष्ट तापमान पर घुलित ऑक्सीजन सामग्री और देशी ऑक्सीजन एकाग्रता के बीच अंतर के रूप में की जाती है।

नदी के बायोटा, घुलित ऑक्सीजन, तापमान और प्रवाह की मात्रा, वेग, निलंबित और घुलित ठोस पदार्थों की सांद्रता जैसे कई सूचकांकों का अनुमान लगाकर धारा की आत्मसात क्षमता को मापा जा

सकता है। जल निकाय में पुनः ऑक्सीकरण प्रक्रिया के दौरान पुनः वातन गुणांक का निर्धारण वायु-जल इंटरफेस में ऑक्सीजन अंतरण दर के उचित मूल्यांकन पर निर्भर होना चाहिए। वर्षों से, शोधकर्ताओं द्वारा नदी के परिभाषित भौतिक और हाइड्रोलिक गुणों से ऑक्सीजन हस्तांतरण की दर की परिकल्पना के लिए कई प्रयास किए गए हैं। हालांकि, हाइड्रोलिक चर और पुनः वातन गुणांक के बीच कार्यात्मक संबंध विशिष्ट सीमा स्थितियों तक सीमित हैं। गहराई, वेग, ढलान और चौड़ाई जैसे पुनः वातन समीकरणों को डिजाइन करने के लिए उपयोग किए जाने वाले चर के सापेक्ष महत्व भी साहित्य में विश्वसनीय नहीं पाए जाते हैं और किसी भी कार्बनिक भार से मुक्त धाराओं पर लागू होते हैं। कई भविष्य कहनेवाला पुनः वातन मॉडल विकसित किए गए हैं और सीमित अध्ययन क्षेत्रों पर लागू किए गए हैं जिनमें चैनल प्रकारों की जबरदस्त परिवर्तनशीलता प्रत्येक मॉडल के परिणामों में महत्वपूर्ण अंतर का कारण बनती है।

नदी के पानी की गुणवत्ता का अनुकरण करने के लिए पर्याप्त पुनः वातन गुणांक के आकलन में सुधार एक महत्वपूर्ण आवश्यकता है। जल गुणवत्ता मापदंडों, महत्वपूर्ण मापदंडों और उनके अप्रत्यक्ष संबंधों में समानता की जांच करने के लिए पदानुक्रमित समूह विश्लेषण (एचएसीए), प्रमुख घटक विश्लेषण (पीसीए) और संरचनात्मक समीकरण मॉडलिंग (एसईएम) को नियोजित किया जाता है। इन सांख्यिकीय दृष्टिकोणों का संयोजन व्यक्तिगत रूप से पानी की गुणवत्ता की भिन्नता में विभिन्न मापदंडों की भूमिका की पहचान करने और प्रदूषकों के महत्वपूर्ण स्रोत की पहचान करने में मदद करता है। इस अध्ययन में, अध्ययन क्षेत्र में उनकी प्रयोज्यता को प्रकट करने के लिए 17 भविष्य कहनेवाला पुनः वातन मॉडल की तुलना की गई है। यमुना नदी को अध्ययन क्षेत्र के रूप में चुना गया है, जो मानवजनित रूप से प्रभावित कई बिंदु स्रोतों से उच्च जैविक भार प्राप्त कर रहा है, जो नदी की प्रवाह दर को महत्वपूर्ण रूप से प्रभावित करता है। नदी 17 बड़े और छोटे नालों से अपशिष्ट जल प्राप्त करती है, जिससे नदी के भौतिक-रासायनिक गुणों में महत्वपूर्ण उतार-चढ़ाव होता है। माध्य पूर्ण त्रुटि (एमएई), माध्य निरपेक्ष प्रतिशत त्रुटि (एमएपीई), औसत वर्ग त्रुटि (एमएसई) और माध्य गुणक त्रुटि (एमएमई) का उपयोग करके

भविष्य कहनेवाला पुनः वातन मॉडल में त्रुटि का अनुमान लगाया गया है; हालांकि, कोई भी रीएरेशन मॉडल पुनः वातन गुणांक के प्रेक्षित मूल्यों के साथ आशाजनक परिणाम नहीं देता है।

कृत्रिम तंत्रिका नेटवर्क (एएनएन), अनुकूली तंत्रिका फजी अनुमान प्रणाली (एएनएफआईएस) और स्वप्रतिगामी एकीकृत चलती औसत (एआरआईएमए) को पुनः वातन गुणांक के आकलन के लिए एक नया मॉडल विकसित करने के लिए लागू किया गया है। विभिन्न एएनएन और एएनएफआईएस मॉडल को डिजाइन करने के लिए इनपुट पैरामीटर के दस अलग-अलग कॉन्फिगरेशन का उपयोग किया गया था, और मॉडल के प्रदर्शन का मूल्यांकन आर, आर² और आरएमएसई का उपयोग करके किया जाता है। एएनएन और एएनएफआईएस से कम से कम आरएमएसई और उच्चतम आर² वाले मॉडल को पुनः वातन गुणांक के अनुमान में सुधार के लिए एआरआईएमए के साथ एकीकृत किया गया है। एएनएफआईएस - एआरआईएमए मॉडल भविष्य कहनेवाला पुनः वातन मॉडल की तुलना में अवक्रमित नदी के लिए पुनः वातन गुणांक के अनुमान में सुधार करता है।

पुनः वातन गुणांक के आकलन के लिए नए विकसित मॉडल के प्रभाव की पहचान करने के लिए एएनएफआईएस - एआरआईएमए मॉडल के साथ एकीकृत क्यूएल2केडब्ल्यू का उपयोग करके नदी के पानी की गुणवत्ता का अनुकरण किया जाता है। पैरामीटर एकाग्रता और संबंधित नदी विशेषताओं में भिन्नता की जांच के लिए डीओ और बीओडी का संवेदनशीलता विश्लेषण भी किया जाता है। अध्ययन से पता चलता है कि विकसित एएनएफआईएस - एआरआईएमए मॉडल पुनः वातन गुणांक के अनुमान में काफी सुधार करता है और इसका उपयोग पानी की गुणवत्ता की भविष्यवाणी में किया जा सकता है। प्रदूषण उपशमन विधियों की संरचना, नदी जल प्रबंधन नीतियों को विकसित करने और उपचार के स्तर को परिभाषित करने के दौरान निर्णय लेने वालों के लिए अध्ययन उपयोगी होगा जो संवेदनशील दूषित पदार्थों या महत्वपूर्ण प्रजातियों को हटाने पर केंद्रित हो सकता है जो नदी के स्वास्थ्य को गंभीर रूप से प्रभावित करते हैं।

TABLE OF CONTENT

CERTIFICATE	i
ACKNOWLEDGMENT	ii
ABSTRACT	iv
TABLE OF CONTENT	vii
LIST OF FIGURES	xii
LIST OF TABLES	xvii
LIST OF ABBREVIATIONS	xx
LIST OF NOMENCLATURE	xxiii
CHAPTER I INTRODUCTION	1
1.1 Hydro-Biochemistry of Riverine System.....	3
1.2 Research Motivation and Objectives	8
1.3 Organization of the Thesis	9
CHAPTER II LITERATURE REVIEW	11
2.1 Monitoring of Water Quality Parameters	12
2.2 Structural Equation Modelling of Water Quality Parameters.....	17
2.3 Estimation of Re-aeration Coefficient	19
2.4 Soft Computing Techniques in Water Resources	28
2.5 River Water Quality Modelling	38
2.6 Concluding Remarks.....	43
CHAPTER III METHODS AND METHODOLOGY	47

3.1	Development of Inter-relation among Water Quality Parameters	47
3.1.1	Identification of Similar Parameters using HACA	49
3.1.2	Pattern Recognition of Parameters using PCA	53
3.1.3	Analysis of Inter-relation through Model Hypothesis using SEM	57
3.2	Analysis of Re-aeration Models.....	61
3.2.1	Variable of Dissolved oxygen Balance Technique (DOBT)	62
3.2.2	Analysis of predictive models.....	63
3.3	Development of Neuro-fuzzy Models for Simulation of Re-aeration Processes.....	65
3.3.1	Re-aeration Simulation using ANN	66
3.3.2	Re-aeration Simulation using ANFIS	72
3.3.3	Re-aeration Modelling using Ensemble ANFIS-ARIMA Model	76
3.4	Simulation of River Water Quality	78
3.4.1	Selection of Model Parameters	80
3.4.2	Calibration and Validation of Model	81
3.5	Investigating Fluctuations in Simulated Parameters.....	84
CHAPTER IV STUDY AREA AND EXPERIMENTAL ANALYSIS		87
4.1	Description of Study Area	87
4.2	Geomorphology of Study Area.....	91
4.3	Sampling and Monitoring of Water Quality Parameters	95
4.4	Experimental Analysis of Water Quality Parameters	98
4.5	Conclusion	101

CHAPTER V RESULTS AND DISCUSSION: STATISTICAL ANALYSIS OF WATER	
QUALITY PARAMETERS	103
5.1 Spatio-Temporal Analysis of Water Quality	103
5.1.1 Spatial Variation of Water Quality Parameters	103
5.1.2 Temporal Variation of Water Quality Parameters	110
5.2 Clustering of Water Quality Parameters	121
5.2.1 Clustering of Parameters for Less Polluting Region.....	122
5.2.2 Clustering of Parameters for High Polluting Region.....	125
5.3 Classification of Water Quality Parameters.....	127
5.4 Inter-relation among Water Quality Parameters	139
5.5 Conclusions.....	147
CHAPTER VI RESULTS AND DISCUSSION: MODELLING RE-AERATION	
PROCESSES	149
6.1 Performance Evaluation of Predictive Re-aeration Models.....	149
6.1.1 Comparison of Predictive Re-aeration Models.....	150
6.2 Development of Re-aeration Model using Soft Computing Techniques.....	163
6.2.1 Estimation of Re-aeration Coefficient using ANN Models.....	164
6.2.2 Estimation of Re-aeration Coefficient using ANFIS Models	174
6.3 Ensemble ANFIS ARIMA Re-aeration Model.....	181
6.3.1 Performance Evaluation of Ensemble ANFIS-ARIMA Model.....	182
6.4 Conclusions.....	189

CHAPTER VII RESULTS AND DISCUSSION: SIMULATION OF YAMUNA RIVER	
WATER QUALITY	193
7.1 Simulation of River Water Quality using QUAL2Kw Integrated with Hybrid ANFIS-ARIMA Model	193
7.2 Configuration of Model	196
7.2.1 Model Input.....	196
7.2.2 River Reaches	196
7.2.3 Model Parameters	197
7.2.4 Calibration and Validation of Model	198
7.3 Performance Evaluation of QUAL2Kw Integrated with ANFIS-ARIMA Model for Water Quality Simulation.....	200
7.4 Sensitivity Analysis of Water Quality Parameters.....	207
7.4.1 Scenario 1.....	208
7.4.2 Scenario 2.....	209
7.4.3 Scenario 3.....	212
7.5 Conclusions.....	214
CHAPTER VIII SUMMARY AND CONCLUSIONS.....	215
8.1 Summary	215
8.2 Conclusions.....	218
8.3 Limitations and Future Scope of Work.....	224
REFERENCES	227
LIST OF PUBLICATION.....	259

BRIEF BIO-DATA OF THE AUTHOR.....263

LIST OF FIGURES

Figure 3.1	Methodology adopted for classification of parameters.....	50
Figure 3.2	Difference between before and after clustering of data points (Serban and Campan, 2008).....	51
Figure 3.3	Flow chart of processes of HACA.....	52
Figure 3.4	Schematic representation of variance of PCA, where PC1 and PC2 are first and second principal components, respectively, and λ_1 and λ_2 represent the eigenvectors of PC1 and PC2, respectively.....	54
Figure 3.5	Methodology adopted to perform the Principal Component Analysis (PCA).....	56
Figure 3.6	Schematic representation of the hypothesis of SEM model. MVs are the measured variables, LVs are the latent variables, where α is the error of measured parameter, β is the coefficient between latent and measured parameter, and ζ is the coefficient between latent variables.....	60
Figure 3.7a	Methodology adopted for designing the re-aeration models.....	67
Figure 3.7b	Structure of ANN and ANFIS techniques used in designing the re-aeration models	68
Figure 3.7c	Structure of ARIMA combined with the ANN or ANFIS to determine re-aeration coefficient	69
Figure 3.8	Structure of data processing of neurons in ANN.....	70
Figure 3.9	Architecture of feed forward ANN.....	72
Figure 3.10	Hybrid structure of ANFIS model, inference rules with AND and OR operators	74
Figure 3.11	Rule formation of different layers in ANFIS function.....	75

Figure 3.12	Mass and flow balance of river reach ‘i’.....	79
Figure 3.13	Flow balance and calibration of different reaches of study area.....	81
Figure 3.14	Methodology adopted for assessing the water quality using newly developed re-aeration model in QUAL2Kw.....	82
Figure 4.1	Inset map of India, showing location of Yamuna River in the Delhi stretch, in between Wazirabad and Okhla barrage.....	88
Figure 4.2	Wastewater drains joining Yamuna River in between Wazirabad and Okhla barrage.....	89
Figure 4.3	Land use land cover map of Delhi.....	91
Figure 4.4	Geomorphology of Delhi.....	92
Figure 4.5	Sampling photographs of Yamuna River.....	94
Figure 5.1	Variation in the water quality parameters at four sampling locations: Wazirabad Barrage (WAZ), ITO, Nizamuddin Bridge (NZM) and Okhla Barrage (OKH).....	106
Figure 5.2	Box-Whisker plots of spatial variations of water quality parameters for the Yamuna River.....	110
Figure 5.3	Average seasonal variation (summer) of water quality parameters at all the sampling locations	114
Figure 5.4	Average seasonal variation (monsoon) of water quality parameters at all the sampling locations	116
Figure 5.5	Average seasonal variation (winter) of water quality parameters at all the sampling locations	118
Figure 5.6	Box-Whisker plots of temporal variations of water quality parameters for Wazirabad and other sampling locations for Yamuna River	121

Figure 5.7	Cluster analysis of water quality parameters at Wazirabad.....	123
Figure 5.8	Cluster analysis of water quality parameters for integrated dataset.....	126
Figure 5.9	Scree plot and biplot of principal component analysis of water quality parameters at Wazirabad barrage.....	131
Figure 5.10	Scree plot and biplot of principal component analysis of water quality parameters at ITO.....	134
Figure 5.11	Scree plot and biplot of principal component analysis of water quality parameters at Nizamuddin bridge.....	136
Figure 5.12	Scree plot and biplot of principal component analysis of water quality parameters at Okhla barrage.....	139
Figure 5.13	SEM models to predict interrelation between the latent (dependent) and measured (independent) variables	146
Figure 6.1	The correlation between estimated and predicted values of re-aeration coefficient.....	151
Figure 6.2	MSE and MAPE found in predictive models.....	159
Figure 6.3	MAE and MME found in predictive models.....	160
Figure 6.4	Residuals plots of estimated and predicted values of re-aeration coefficient for predictive re-aeration models showing minimum error in evaluation matrices	161
Figure 6.5	Correlation of estimated and predicted re-aeration coefficient of four highly correlated predictive models.....	162
Figure 6.6	Plot of minimum error obtained in predictive models with different evaluation measures.....	163
Figure 6.7	Results of training, validation, testing and overall computation of ANN-1 model for the estimation of re-aeration coefficient	169

Figure 6.8	Results of training, validation, testing and overall computation of ANN-3 model for the estimation of re-aeration coefficient.....	170
Figure 6.9	Results of training, validation, testing and overall computation of ANN-5 model for the estimation of re-aeration coefficient.....	171
Figure 6.10	Results of training, validation, testing and overall computation of ANN-6 model for the estimation of re-aeration coefficient.....	172
Figure 6.11	Results of training, validation, testing and overall computation of ANN-8 model for the estimation of re-aeration coefficient.....	173
Figure 6.12	Comparative plots of training and testing data with the output obtained from all the designed ANFIS models (where output is re-aeration coefficient values)	180
Figure 6.13	Autocorrelation function of K_2 (1/day)	183
Figure 6.14	Partial autocorrelation function of K_2 (1/day)	184
Figure 6.15	Estimated and predicted values of re-aeration coefficient obtained from various ARIMA models.....	189
Figure 7.1	The mass and flow balance of the river reach 'i'.....	194
Figure 7.2	Plot of estimated re-aeration coefficient with predictive re-aeration models.....	201
Figure 7.3	Plot of estimated re-aeration coefficient with proposed ANFIS-ARIMA model.....	202
Figure 7.4	The calibration results of predicted DO and its comparison with the observed DO	203
Figure 7.5	The validation results of predicted DO and its comparison with the observed DO	204

Figure 7.6	The calibration results of predicted BOD and its comparison with the observed BOD	204
Figure 7.7	The validation results of predicted BOD and its comparison with the observed BOD	205
Figure 7.8	Scenario 1 for sensitivity analysis of DO and BOD at different confidence intervals	210
Figure 7.9	Scenario 2 for sensitivity analysis of DO and BOD at different confidence intervals	211
Figure 7.10	Scenario 3 for sensitivity analysis of DO and BOD at different confidence intervals	213

LIST OF TABLES

Table 2.1	Summary of predictive re-aeration models considered in the present study	24
Table 4.1	Water quality parameters and techniques used for analysis	97
Table 4.2	Average values of water quality parameters analysed in the study area for five years	100
Table 5.1	Proximity matrix of water quality parameters at Wazirabad barrage	124
Table 5.2	Agglomeration schedule of the cluster formed from water quality parameters at Wazirabad barrage	124
Table 5.3	Proximity matrix of water quality parameters for integrated dataset (Combination of ITO, Nizamuddin and Okhla).....	125
Table 5.4	Agglomeration schedule of the cluster formed from water quality parameters for integrated dataset (Combination of ITO, Nizamuddin and Okhla).....	127
Table 5.5	Eigenvalues of parameters at Wazirabad	129
Table 5.6	Eigenvectors of parameters at Wazirabad	129
Table 5.7	Correlation between parameters and factors at Wazirabad	129
Table 5.8	Squared cosines of parameters at Wazirabad	130
Table 5.9	Eigenvalues of parameters at ITO	132
Table 5.10	Eigenvectors of parameters at ITO	132
Table 5.11	Correlation between parameters and factors at ITO	133
Table 5.12	Squared cosines of parameters at ITO	133
Table 5.13	Eigenvalues of parameters at Nizamuddin	134
Table 5.14	Eigenvectors of parameters at Nizamuddin	134

Table 5.15	Correlation between parameters and factors at Nizamuddin	135
Table 5.16	Squared cosines of parameters at Nizamuddin	135
Table 5.17	Eigenvalues of parameters at Okhla	137
Table 5.18	Eigenvectors of parameters at Okhla	137
Table 5.19	Correlation between parameters and factors at Okhla	138
Table 5.20	Squared cosines of parameters at Okhla	138
Table 5.21	Covariance matrix of SEM model for Wazirabad	141
Table 5.22	Covariance matrix of SEM model for ITO	141
Table 5.23	Covariance matrix of SEM model for Nizamuddin	142
Table 5.24	Covariance matrix of SEM model for Okhla	143
Table 5.25	Performance analysis of SEM models for all the locations	147
Table 6.1	Summary of predictive re-aeration models	152
Table 6.2	Annual average values of estimated and predicted re-aeration coefficient obtained from different predictive models	154
Table 6.3	Average error observed from MSE, MAPE, MAE and MMR at all the sampling locations	156
Table 6.4	ANN and ANFIS models designed with combinations of different parameters	165
Table 6.5	Design criteria of ANN models	167
Table 6.6	Training, validation and testing results of different ANN models	168
Table 6.7	Design criteria of ANFIS models	175
Table 6.8	Training and testing results of various ANFIS models	177
Table 6.9	Configuration of various ARIMA models	185
Table 6.10	Statistical evaluation of ARIMA models	186
Table 7.1	Major calibration parameters rate constants for the Yamuna River	199

Table 7.2	The calibration and validation results of modelling of DO and BOD for Yamuna River.....	205
-----------	---	-----

LIST OF ABBREVIATIONS

DO	Dissolved Oxygen
BOD	Biochemical Oxygen Demand
CPCB	Central Pollution Control Board
USEPA	United States Environmental Protection Agency
DOBT	Dissolved oxygen balance technique
ANN	Artificial neural network
ANFIS	Adaptive neuro-fuzzy inference system
SVM	Support vector machine
ARIMA	Auto regressive integrated moving average
KNN	K-nearest neighbours
MLR	Multivariate linear regression
SEM	Structural equation modelling
PCA	Principal Component analysis
CA	Cluster analysis
DA	Discriminant analysis
FA	Factor analysis
COD	Chemical oxygen demand
TDS	Total dissolved solid
ANOVA	Analysis of variance
AI	Artificial Intelligence
GA	Genetic algorithm
FFBP	Feed forward back propagation
RBF	Radial basis function

ARMAX	Autoregressive–moving-average model with exogenous inputs
NARX	Nonlinear autoregressive network with exogenous inputs
NH ₃ -N	Ammonia nitrogen
WNN	Wavelet neural network
FNN	Feedforward neural network
RNN	Recurrent neural network
LM	Levenberg Marquardt
BR	Bayesian regularization
GDX	Gradient descent with momentum and adaptive learning rate
RMSE	Root mean square error
ARMA	Auto regressive moving average
DENFIS	Dynamic evolving neural fuzzy inference system
MLP	Multilayer perceptron
ANFIS-GP	Adaptive neuro-fuzzy inference system – Grid Partitioning
ANFIS-SC	Adaptive neuro-fuzzy inference system – Subtractive Clustering
FFA	Firefly optimization algorithm
R ²	Coefficient of determination
NSE	Nash-Sutcliffe Efficiency
WI	Willmott's Index
MAE	Mean Absolute error
SSA	Singular spectrum analysis
SA	Sensitivity analysis
MCS	Monte Carlo simulation
SAFE	Sensitivity analysis for everybody
HACA	Hierarchical aligned cluster analysis

FV	Feature Vector
NME	normalized mean error
SE	standard error
MME	mean multiplicative error
MSE	mean square error
MAPE	mean absolute percentage error
M-FIS	Mamdani fuzzy inference system
TS-FIS	Takagi-Sugeno fuzzy inference system
ACF	autocorrelation function
PACF	partial autocorrelation function
DJB	Delhi Jal Board

LIST OF NOMENCLATURE

D	Final DO deficit (mg/l)
D_0	Initial DO deficit (mg/l)
K_1	De-oxygenation coefficient (1/day),
K_2	Re-aeration coefficient (1/day),
K_3	Settling rate coefficient (1/day),
K_r	K_1+K_3
L_0	Initial BOD (mg/l)
L	BOD (mg/l)
t	time
c	oxygen concentration at time t
c_s	oxygen concentration at saturation point
H	depth (m)
V	velocity (m/s)
S	slope (1/S)
S_B	Sediment oxygen demand
L_x	longitudinal distance
F_r	Froude's number
E_d	Square Euclidean distance
Δ	Merging factor
λ	Eigenvector
σ_z^2	Variance of new variable
σ_{rs}^2	Variance of translated variable
SS	Square symmetric matrix

E	Square orthogonal covariance matrix
D_m	Diagonal matrix
MP	Measured parameter
LV	Latent variable
α	error of measured parameter
β	coefficient between latent and measured parameter
ζ	coefficient between latent variables
φ	Variance of latent variable
E_p	normalized error
μ_{A_i}	Membership function
w_i	Antecedent function
$\overline{w_i f_i}$	Coefficient of function
Q_{in}	Influent discharge from river
Q_{ab}	Effluent discharge from river
R_n	Reach of river