1-26-2023 10:15 AM

Modeling and multi-objective optimization of wastewater treatment process

Navneet Kaur, *The University of Western Ontario*

Supervisor: Ray, Ajay, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Chemical and Biochemical Engineering

Abstract

Wastewater is water that has already been used and requires treatment before releasing it into natural water bodies like lakes and rivers. Wastewater treatment is the process of removing impurities from wastewater. In this treatment process, the impurities are removed and converted to effluent. This effluent is returned to the water cycle with minimum impact on the environment. Conventional treatment plants consist of three stages: primary, secondary, and tertiary treatment. Treatment of wastewater is quite complicated because of the number of stages involved in this process. Most wastewater treatment plants are operated manually therefore at times it becomes difficult for operators to maintain desired effluent quality. Modeling and simulation technique is suggested to improve and predict the performance of the plant.

Mechanistic models involve fundamental equations to model the process, since wastewater treatment is a complicated process artificial neural network is suggested to model the process. This is a data-driven approach that identifies patterns between input and output data. This type of technique is called black-box modeling. The past four years' data is analyzed and used to predict the effluent quality of the plant. The effluent quality is measured in terms of four major pollutants namely biochemical oxygen demand, suspended solids, total phosphorus, and ammonia. 70% of the total data was used for training purposes and 30% was used for validation purposes. The correlation coefficient between the modeled values and actual values was around 0.97.

To minimize the concentration of the pollutants in the effluent stream multi-objective optimization is suggested. A genetic algorithm is used to solve multi-objective optimization of the treatment plant. An equalization tank or buffer system is suggested to counterbalance the fluctuating flow and composition of influent to the treatment plant. The decision variables associated with this process are the temperature of the influent stream, total sewage flow, biochemical oxygen demand, suspended solids, pH, total phosphorus, and ammonia of the influent stream. The optimizer was able to minimize the concentration of pollutants in the effluent stream and comply with the strict effluent regulations.

Keywords: Wastewater treatment, Modeling, Multi-objective optimization

Summary For Lay Audience

The purpose of wastewater treatment is to remove the impurities before discharging them back into the environment. Untreated wastewater is harmful to both humankind and the environment. Improper operation of WWTP can cause environmental and various health issues like cholera and dysentery. The optimal operation of WWTP can improve efficiency and reduce the costs associated with various processes. In this research work, a multi-objective optimization approach has been used to minimize the concentration of pollutants in the effluent stream instead of a single-optimization approach. In the real world, multi-objective problems with conflicting objectives are frequently encountered. In this case, a set of equally good solutions is generated, also known as the Pareto set. Though sometimes it becomes difficult for the decision maker to choose a single optimal solution from a set of optimal solutions.

Wastewater treatment is a complex system, and it is difficult to explore various design ideas on a pilot plant. Modeling helps in understanding how a system would behave in various conditions without experimentation. A WWTP model is a representation of physical and chemical processes involved in the purification of wastewater. In my research work, a black-box modeling approach has been employed to model WWTP. This type of modeling is based on the input-output behavior of the process in contrast to physical modeling which is time-consuming. A model based on ANN was developed to predict the quality of effluent stream.

To minimize the concentration of the pollutants in the effluent stream multi-objective optimization is suggested. A genetic algorithm is used to solve multi-objective optimization of the treatment plant. An equalization tank or buffer system is suggested to counterbalance the fluctuating flow and composition of influent to the treatment plant. The decision variables associated with this process are the temperature of the influent stream, total sewage flow, biochemical oxygen demand, suspended solids, pH, total phosphorus, and ammonia of the influent stream. The optimizer was able to minimize the concentration of pollutants in the effluent stream and comply with the strict effluent regulations.

Acknowledgement

I would like to thank my supervisor Dr. Ajay K. Ray, whose expertise was invaluable in formulating the research questions and methodology. Your insightful feedback pushed me to sharpen my thinking and brought my work to higher level.

I would like to thank my colleagues Mrinmoyee Mandal, Malini Ghosh, and Stanislav Ivanov for their patient support and guidance.

In addition, I would like to thank my parents for their wise counsel and sympathetic ear. Finally, I would like to thank my friends for their support and motivation.

Table of Contents

List of Figures

List of Tables

List of abbreviations

ACO	Ant colony optimization
ANN	Artificial neural network
ARIMA	Autoregressive Integrated moving average
ARMA	Autoregressive moving average
ARX	Autoregressive with exogenous inputs
ASM1	Activated sludge model 1
ASM2	Activated sludge model 2
ASM3	Activated sludge model 3
ATAD	Autothermal thermophilic aerobic digestion
BOD	Biochemical oxygen demand
COD	Chemical oxygen demand
DE	Diatomaceous earth
DNN	Dynamic neural network
DO	Dissolved oxygen
EBPR	Enhanced biological phosphorus removal
GA	Genetic algorithm
k-NN	k-nearest neighbor

LPP	Linear programming problem
MOGA	Multi-objective genetic algorithm
MOO	Multi-objective optimization
MSE	Mean square error
NLPP	Non-linear programming problem
NN	Neural network
NSGA	Non-dominated sorting genetic algorithm
PAOs	Phosphorus accumulating organisms
PSO	Particle swarm optimization
RF	Random forests
SVM	Support vector machine
SQP	Sequential quadratic programming
TN	Total nitrogen
TP	Total phosphorus
TSS	Total suspended solids
UV	Ultraviolet
WWTP	Wastewater treatment plant
WWTPs	Wastewater treatment plants

Nomenclature

BOD_e Effluent biochemical oxygen demand

BOD_{eff} Effluent biochemical oxygen demand

BOD_i Influent biochemical oxygen demand

BOD_{inf} Influent biochemical oxygen demand

$Flow_i$ Influent flow

K_A ammonification rate

K_h hydrolysis rate constant

K_{NH4} saturation coefficient of nitrogen

K_{OH} saturation coefficient of oxygen

K_{SS} saturation coefficient of readily biodegradable substrate

K_x saturation coefficient for particulate COD

NH_{3e} Effluent ammonia

NH_{3eff} Effluent ammonia

NH_{3i} Influent ammonia

NH_{3inf} Influent ammonia

P_c Crossover probability

$p^t_{(gb)}$ Global best

$p^t_{(i,lb)}$ Local best of ith particle

SS	Suspended solids
SS_e	Effluent suspended solids
SS_{eff}	Effluent suspended solids
SS_i	Influent suspended solids
SS_{inf}	Influent suspended solids
$temp_i$	Influent temperature
TP_e	Effluent total phosphorus
TP_{eff}	Effluent total phosphorus
TP_i	Influent total phosphorus
TP_{inf}	Influent total phosphorus
Vs	Volume of sample
W_f	Weight of filter
W_{fss}	Weight of filter with suspended solids
$\mu_{max, H}$	heterotrophic specific growth rate

Chapter 1

1 Introduction

Water is essential to all forms of life and makes up about 70% of the human body. The availability and quality of water play an important role in determining the quality of life. Effluents from industries are disposed into streams without proper treatment, which severely affects the quality of receiving streams. Wastewater treatment is the process of removing impurities from wastewater before they reach natural water bodies such as lakes, rivers, estuaries, and oceans. In this process, the impurities are removed and converted to effluent. This effluent is returned to the water cycle with minimum impact on the environment. Freshwater is a limited environmental resource because freshwater sources have become polluted due to human activity: wastewater produced from agriculture and industrial processes has been discarded into freshwater resources. By treating wastewater, water-borne diseases can be controlled. The conventional treatment plants consist of three stages(US EPA, 1998): primary, secondary, and tertiary treatment. During primary treatment, large particles and coarse material are removed. Secondary treatment (also called biological treatment) degrades organic matter using microorganisms. The tertiary treatment removes pathogens and microorganisms, to meet stringent effluent quality standards.

Table 1-1: Various constituents of wastewater

Microorganisms	Virus, bacteria
Biodegradable organic matter	
Other organic material	Oil and grease, fat, detergents, Pesticides
Nutrients	Nitrogen, Phosphorus, Ammonium
Metals	Hg, Pb, Cd, Cr, Cu
Inorganic material	Acids, bases
Odor	Hydrogen sulfide

Radioactivity	

Wastewater treatment plants are major consumers of energy. They consume 7% of electrical energy worldwide (Plappally & Lienhard V, 2012). Generally, wastewater treatment plants are operated based on experience and small-scale experiments. Therefore, plants are not operated optimally. The growing population and restricted effluent emission rights mean that the existing process needs to be looked at in terms of optimization. Optimization is the selection of the best element from a set of available alternatives.

Figure 1.1 Study Area

https://london.ca/projects/adelaide-wastewater-treatment-plant-climate-change-resiliency-environmental-assessment

1.1 Introduction to modeling and simulation

Simulation is the process of designing a model of a real system and conducting experiments with this model. The main objective of the simulation is to gain a thorough understanding of the behaviour of the system. It also involves the evaluation of various alternative strategies. Broadly a model is a representation of reality. For instance, a blueprint is a model of a building and is a two-dimensional model of a three-dimensional reality.

Various types of models:

1. Physical models resemble the system being studied e.g., flight simulator

2. Scaled models also represent the system under study, but at a different size e.g., a scaled-up model of an atom.

3. Mathematical models have relationships represented by mathematical functions. These are generalized and oversimplified models.

4. A heuristic model is a collection of decision rules, usually computer-based and is not limited by physical or mathematical boundaries.

Every model employs abstraction which is an important characteristic of modeling. A model alters reality to some degree as a model can be larger, slower, or faster. Therefore, a model is a simpler version of a real-world system. Since physical modeling is a time-consuming process, the black box model is suggested to overcome this problem. In the black box model, a set of inputs is mapped to a set of outputs through a transformation.

Figure 1.2 Black box model

1.1.1 Types of simulation models

1. The static model represents a system at a particular point in time. A static model is trained offline.

2. The dynamic model represents systems as they change over time. A dynamic model is trained online. e.g., simulation of a banking system

3. The stochastic model has one or more random variables as inputs. Random inputs lead to random outputs. e.g., the simulation of a bank involves random interarrival and service time.

4. The deterministic model has defined results for a known set of inputs and has no randomness associated with it e.g., the arrival of patients to the dentist at the scheduled appointment time.

5. Discrete-event simulation involves changes to the system at a discrete set of points in time e.g., a manufacturing system with parts entering and leaving at specific times.

6. A continuous model involves changes to the system continuously over time e.g., water flowing in and out of a reservoir.

7. The mixed model contains both discrete and continuous elements e.g., a refinery with continuously changing pressure inside the vessel and discreetly occurring shutdowns.

Modeling and simulation increase the understanding of very complex systems. This reduces the risks and costs of experimentation. The likely outcomes of all the alternatives can be estimated before building the actual system. A simulation is a powerful approach for making evidence-based decisions and improving efficiency and profitability. For example, in a transportation system how a new bus or rail line can affect people's lives in terms of travelling times can be estimated. Therefore, with the help of a simulation model performance and improvements can be easily evaluated before constructing a new line in a transportation system. Every model is limited by the assumptions that created it, even the best models don't predict the future accurately (Maria, 1997). Applications of modeling and simulation can be found in various fields like economics, financial industry, engineering, biological transportation, and epidemiology. Epidemiology is the study of epidemics and in the situation of the COVID-19 outbreak, mathematical modeling has played a vital role in understanding the dynamics and predicting the curves and behavior of virus spread.

1.1.2 Steps in the simulation study

1. Define the system intended to simulate.

2. Identify and collect data: collect data on system specifications, input variables, output variables, and performance of the system.

3. Formulate and develop model: A computer program is designed where a set of inputs give a set of outputs. To understand the performance of the computer program verification is done. Verification techniques include varying the input parameters and comparing the outputs and tracing intermediate results with actual outcomes.

4. Validate the simulation: Comparing the model's performance under known conditions with the performance of the actual system.

5. Analyze the results of the simulation and recommend further courses of action.

6. Document model for future use (Maria, 1997).

1.2 Introduction to Optimization

Optimization is the process of making the best use of a situation or resource. It is choosing the best element from a set of alternatives according to some criterion. A mathematical programming problem is one that typically maximizes or minimizes a function (called an objective function) for a set of criteria, called the constraints. For example, minimizing the cost of transportation of goods, maximizing the profit of a company, etc.

$$Max/Min \quad f(x) \;\; \leftarrow Objective\ function \tag{1.1}$$

$$Subject\ to$$

$$g_j \leq 0\ , \;\; j \in 1,2,3 \ldots m$$

There are broadly two types of programming problems:

Linear programming problem(LPP)-where the objective function and constraints are both linear functions of decision variables. For example, Min x+y, s/t x$\geq 2, y \geq 4$ is an LPP as both the objectives and the constraints are linear functions of x and y.

Non-linear programming problem(NLPP)-where either the objective function or constraints or both are non-linear functions of the decision variables. For example: Max x^2+y^2 s/t x+y ≤ 10, x, y≥ 0 is an NLP as the objective function is non-linear function of x and y.

1.2.1 Optimization techniques

Mathematical programming techniques based on geometric properties of the problem

1. Simplex algorithm 2. interior point method

Simplex Algorithm

The simplex method is a commonly used algorithm to solve LPP (Dantzig, 1990). This method uses slack variables and pivot variables to find the optimal solution. The steps involved in simplex method are as follows:

1. Standard form-In order to perform the simplex algorithm problem must be formulated appropriately and consistently.

2. Slack variables-They are introduced to replace inequality constraints with an equality constraint and a non-negative constraint.

3. Find the initial feasible solution-This can be done by setting all non-basic variables to 0.

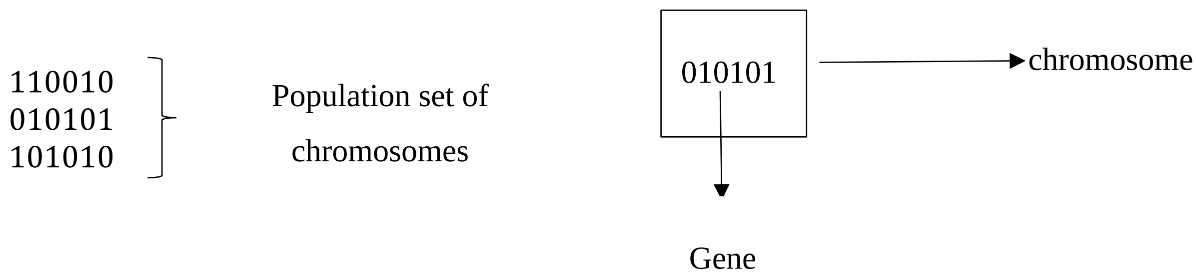
4. Perform iteration of the simplex algorithm.

Metaheuristic techniques

Meta means a high-level methodology, while heuristics means art in finding new strategies for solving a problem (Gunantara & Nurweda Putra, 2019). Examples of metaheuristic methods based on population are:-

1. Genetic algorithm
2. Ant colony optimization
3. Particle Swarm optimization

A genetic algorithm(GA) is a search-based optimization technique based on the principles of natural selection. It is frequently used to find the optimal solution to difficult problems which otherwise would take a lifetime to solve. This algorithm is based on Darwin’s principle of natural selection and uses operations such as crossover and mutation (Haldurai et al., 2016). GA is a population-based random search technique and every individual in the population corresponds to a possible solution.



In GA every solution is characterized by a set of parameters known as genes. Genes are joined together to form a chromosome(solution). The first step of GA is to generate an initial population and calculate the fitness of each solution. Fitness is the value objective function in the optimization problem. The fitness function determines the ability of an individual to compete with other individuals. Based on the value of the objective function fitness score is assigned to everyone. The individuals with higher fitness scores will be selected for recombination and mutation to form a new generation. Each solution is represented by a bit string as shown above.

Crossover

In a crossover, a mating pair is selected randomly from a population. The crossover operator is used to create new solutions from the existing solutions. This operator exchanges the gene information between two parents and as a result, two offspring solutions are produced. The crossover point is chosen using a random number generator, generating an integer between 1 and L. Where L is the length of the chromosome. Various crossover operations are single-point crossover, two-point crossover, uniform crossover, and multi-point crossover. Single point crossover is shown below and genes after the crossover point are swapped between parents.

$$\begin{array}{l|lcl} 0 & 001 & & 0010 \\ 1 & 010 & \rightarrow & 1001 \end{array}$$

Only good strings get propagated and less good ones slowly die during the copying process. Not all good strings in the mating pool undergo crossover, if crossover probability is P_c then 100(1-P_c)% of strings continue unchanged to the next generation.

Mutation

The purpose of the mutation operator is to introduce new features into the solution string of population to maintain diversity. This is analogous to biological mutation. The mutation operator is used to avoid convergence to local minima by preventing the population of chromosomes from becoming identical to each other. For example, the first position of the strings shown below can never become 1 by crossover.

$$\begin{array}{c} 0110\ldots \\ 0011\ldots \\ 0001\ldots \end{array}$$

Since GA is an iterative process and the algorithm stops when the population is no longer able to generate better individuals.

Ant colony optimization(ACO) is inspired by the behavior of a colony of ants. Communication between ants is based on the use of a chemical compound called pheromone (Gunantara & Nurweda Putra, 2019). The main motive behind ACO is to minimize the path and power consumption in finding the optimal solution analogous to how ants search for food. Ants lay down pheromone while searching for food and way back to their colony. If other ants find this path, they are likely to follow the same path and eventually reach their objective of finding food. However, the pheromone trail starts to evaporate with time. Ants are more likely to choose a path with a higher pheromone and therefore the probability of choosing the shortest path is more. When one ant finds a shorter path, other ants are likely to follow that path. This kind of positive feedback eventually leads to all ants following the shortest path.

Particle swarm optimization(PSO) was proposed by Kennedy and Eberhart in 1995. The basis of this algorithm is that when a flock of birds move in a group, they share their discovery and help the entire flock get the best hunt. PSO mimics the phenomenon of the flocking of birds or a school of fish. It is a direct search method and does not require any gradient information. Each member in the swarm learns from their experience and other members for changing the search pattern to locate the food.

PSO algorithm

1. PSO is initialized with a group of random particles or solutions and searches for an optimal solution by updating generations

2. Particles move through the solution space and are evaluated according to some fitness criteria after each step. In every iteration, each particle is updated by two best values.

3. The first one is the best solution a particle has achieved so far. This value is referred to as the pbest.

4. Another best value is the best value obtained so far by any particle in the population. This value is called the global best or gbest. The particles work independently and keep track of the pattern of other members to find the optimal solution to a problem.

5. Each particle tries to update its current position and velocity according to the distance between its current position and pbest and the distance between its current position and gbest.

The position of the particle is updated according to the following equation

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \tag{1.2}$$

$$v_i^{(t+1)} = v_i^{(t)} + c_1 r_1 \left(p_{(i,lb)}^{(t)} - x_i^{(t)}\right) + c_2 r_2 \left(p_{gb}^{(t)} - x_i^{(t)}\right) \tag{1.3}$$

Where:-

i is the ith particle

t is generation counter

c_1 and c_2 are the acceleration coefficients

r_1 and r_2 are random numbers

$p^t_{(i,lb)}$ is the local best of ith particle

$p^t_{(gb)}$ is the global best

1.2.2 Introduction to Multi-objective optimization

Multi-objective optimization(MOO) problems involve more than one objective function to be maximized or minimized simultaneously. This results in trade-off decision-making as there is no single best solution for problems with more than one objective function (Deb, 2011). For example, while buying any commodity price and quality are constantly being compared. As it is hard to minimize price and maximize the quality of a commodity. A single objective function has a single optimal solution but in the real world, it is hard to find problems that involve only one objective function.

There are three components of any objective function

1. Objectives

   minimize or maximize $f_i(x_1, x_2, \dots \dots . x_n), i = 1,2 \dots . . m$

2. Constraints

   Subject to $g_j(x_1, x_2, \dots \dots . x_n), j = 1,2, \dots \dots . . l$

3. Design variables

$$x_k\ , k = 1,2, \dots \dots . m$$

Pareto optimality

A state is said to be Pareto optimal if there is no other state dominating the state for a set of objective functions. A state X dominates a state Y if X is better than Y in at least one objective function and not worse with respect to all other objective functions. This concept gives a set of solutions and not a single solution. This idea was first proposed by Goldberg in 1989 (Ngatchou et al., 2005).

1.2.2.1 Classic MOO methods

Weighted sum method

In the weighted sum method, a set of objectives is combined into a single objective function by adding each objective pre-multiplied by a user-defined weight. The weight of an objective is chosen in proportion to the relative importance of the objective (Marler & Arora, 2010). A disadvantage of the weighted sum method is that it is difficult to set the weight vector to obtain a Pareto-optimal solution in the desired region.

$$\text{Minimize } F(x) = \sum_{m=1}^{M} w_m f_m(x)$$

$$\text{Subject to } g_j(x) \geq 0, \quad j = 1,2, \dots \dots \dots \dots J \qquad (1.4)$$

$$h_k(x) = 0, \quad k = 1,2, \dots \dots \dots K$$

$$x_i^{(L)} \leq x_i \leq x_i^U, i = 1,2, \dots \dots \dots . n$$

ϵ-Constraint Method

the ϵ-constraint method is a classical method for handling multi-objective optimization problems by converting them to a single objective optimization problem. In this method, only one objective is kept, and the rest of the objectives are restricted by user-specified values (Ngatchou et al., 2005).

$$\text{Minimize} f_\mu(x)$$

$$\text{Subject to } f_m(x) \leq \epsilon_m, m = 1,2, \dots \dots \dots . . M \text{ } and \text{ } m \neq \mu$$

$$g_j(x) \geq 0, \quad j = 1,2, \dots \dots \dots \dots J \qquad (1.5)$$

$$h_k(x) = 0, \quad k = 1,2, \dots \dots \dots K$$

$$x_i^{(L)} \leq x_i \leq x_i^U, i = 1,2, \dots \dots \dots . n$$

1.2.2.2 Multi-objective Genetic Algorithm

This approach consists of a scheme in which the rank of an individual corresponds to the number of individuals in the current population by which it is dominated. It maintains diversity in the non-dominated solutions. The first step of a multi-objective genetic algorithm(MOGA) is to sort the population according to rank. Rank 1 is given to the individuals with the best fitness. All non-dominated individuals are assigned rank 1. Fitness is assigned to individuals based on a linear function. Non-dominated sorting genetic algorithm(NSGA) is a modified version of MOGA proposed by Srinivas and Deb in 1994 (Coello et al., 2007). In this algorithm, before selection is performed, the population is ranked based on domination. All non-dominated individuals are classified into one category. Then this group of classified individuals is removed from the population and another layer of non-dominated individuals is considered. This process continues until all the individuals in the population are classified. Since the individuals in the first front have the maximum fitness value, they will get more copies than the rest of the population. This allows to search for non-dominated regions and eventually results in convergence.

NSGA-II is a commonly used MOO technique and in this algorithm, individuals are selected based on rank and crowding distance (Deb et al., 2002)

The key process steps of NSGA-II are: -

1) Start with a random population of solutions(P) encoded in binary form(chromosomes)
2) Create a child population (Q) through crossover and mutation.
3) Combine P&Q and score for all objectives
4) Identify the first Pareto front (F1)
5) If F1 is larger than the maximum permitted solution, then reduce the size of F1 by crowding distance.
6) If F1 is smaller than the required population size, then repeat the Pareto selection. This new set of solutions is F2.
7) Repeat Pareto selection until the required population is reached.

8) Repeat from step 2 for the required number of generations or until stop criterion is reached.

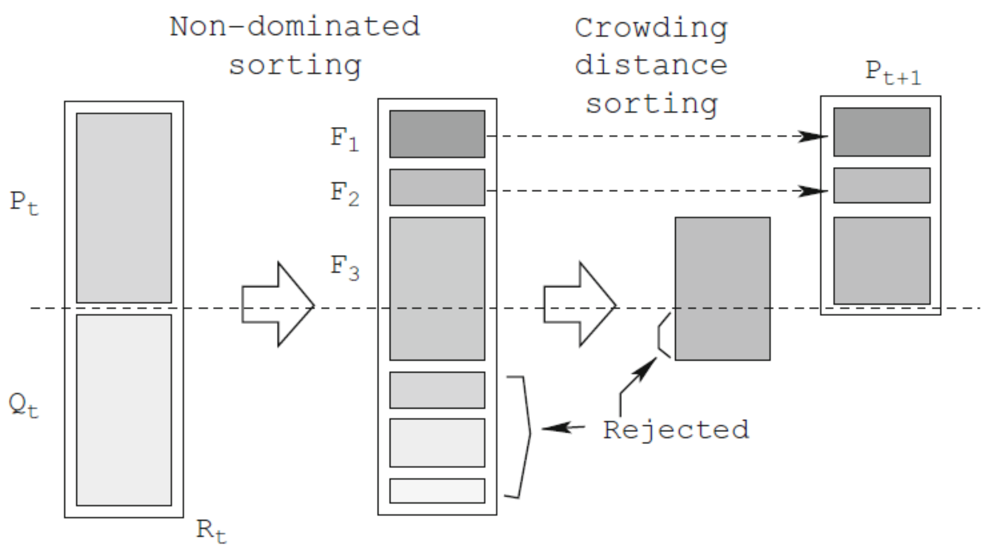


Figure 1.3 Schematic of NSGA-II procedure

1.3 Summary

The motivation behind wastewater treatment and the various stages involved in wastewater treatment is discussed in this chapter. Wastewater treatment plants are major consumers of energy. Therefore, it is important to operate them optimally. Various optimization techniques such as the simplex method, GA, ACO, and PSO have been briefly discussed in this chapter. In this study, multi-objective optimization is performed using GA to minimize the concentration of pollutants in the effluent stream. Pareto optimality and MOGA are explained in this chapter.

Chapter 2

2 Literature Review

2.1 Background

Modeling helps in understanding how a system would behave in various conditions without experimentation. For complex systems like wastewater treatment, it is difficult to explore various design ideas on a pilot plant. A wastewater treatment plant model is a representation of the physical and biochemical processes involved in the purification of wastewater. The biochemical process involves the conversion of organic material and nutrients into carbon dioxide, nitrogen, and particulate fraction. This particulate fraction is further removed from water through physical separation. The “state-of-the-art models” for activated sludge processes are activated sludge model 1(ASM1)-Activate sludge model 3(ASM3) models developed by the international water association (IWA) task group. These models incorporate oxidation of organic matter, nitrification, and denitrification. ASM2d also describes biological and chemical phosphorus removal. The ASM models have been updated several times. ASM1 has been considered as the reference model. ASM3 was developed to include additional processes that were missing in the ASM1 model.

ASM1 and ASM 3 were developed for domestic water and hence cannot be used for industrial wastewater. These models are developed for a temperature range of 8-23°C. A significant error can occur if these models are applied beyond the suggested temperature range (Szilveszter et al., 2010). To overcome these limitations black-box modeling is suggested. In this type of modeling input enters a system and output comes out of the system, but the process by which that input is considered to generate the output is not fully understood. An artificial neural network black-box modeling approach was used to predict the performance of the Doha West wastewater treatment plant (Mjalli et al., 2007). The model provided accurate predictions of the effluent stream in terms of biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total suspended solids (TSS).

The purpose of wastewater treatment is to protect clean water. Approximately 330 km^3 of municipal wastewater is generated globally per year. Wastewater treatment plants(WWTP) are energy-intensive (Zhang et al., 2016). Major electrical energy consumers in the wastewater treatment process are pumps and aeration systems. Electrical energy consumption depends on the size and design of the plant. Moreover, wastewater treatment plants are highly dynamic, which leads to fluctuations in the influent characteristics. Optimization can reduce both capital and operating costs. The activated sludge process is one of the main wastewater treatment processes. It converts organic waste to stable inorganic forms or cellular mass. In this process soluble and colloidal organic material left after primary sedimentation is metabolized by a group of microorganisms to carbon dioxide and water. The activated sludge process is a part of biological wastewater treatment systems. In this process, the air is added to a liquid or substance. Insufficient aeration can lead to violation of discharge permits. Oxygen is required by microorganisms to biodegrade organic materials. Dissolved oxygen (DO) should be maintained at a proper level to keep the microorganisms alive. Dissolved oxygen concentration is an important parameter for controlling the activated sludge process. Several research papers have been published on how to control and model WWTP. Oxygen is also required for the removal of nitrogen. Nitrogen is removed in two steps. Firstly, ammonium is oxidized to nitrate in the presence of oxygen. Then under anoxic conditions nitrate is converted to nitrogen gas. The nitrification and denitrification processes depend on the concentration of oxygen. Efficient operation of wastewater treatment improves the performance of the plant while meeting strict effluent norms set by the pollution control board. Optimization of a wastewater treatment plant involves adjustment of various process variables such as chemical dosing rate, air supply, solid retention time and hydraulic loading rate. In the past several optimization studies have been carried out to efficiently operate wastewater treatment plants. Aerobic and anaerobic conditions are alternately carried out in an aeration tank. Optimization of aeration profile using sequential quadratic programming (SQP) technique can reduce electrical energy consumption by 30% without compromising the effluent quality (Chachuat et al., 2000). To maintain effluent quality higher aeration rate is required. Various studies on minimization of aeration in activated sludge have been reported in the past. To avoid the effect of upstream disturbances due to fluctuations in the flow and variable characteristics of municipal wastewater, aeration rate is kept higher than the required. Electrical energy consumption was reduced by more than 60% while maintaining effluent quality using genetic algorithm (Ozturk et al., 2016).

Oxygen transfer dynamics and growth rate function are both non-linear and there is a huge potential for saving electrical energy by optimal use of DO profile to avoid zones with unnecessary high aeration rates. Each aeration zone can have a separate DO probe instead of having only two probes in the first and last tank. This control strategy was able to reduce total airflow by 18% (Thunberg et al., 2009). A larger reduction of the airflows in the first zone was replaced by a smaller increase in the last zones.

Nowadays, the focus is not only on improving the performance of WWTPs but also to reduce the impact on the environment. Greenhouse gases cause the greenhouse effect. Greenhouse gases are the compounds that are responsible for keeping the earth's surface warmer by trapping heat in the atmosphere. Greenhouse gases, carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) can all be produced in wastewater treatment operations (Gupta & Singh, 2012). N_2O gas is an intermediate of biological processes such as heterotrophic denitrification and nitrification (Campos et al., 2016). Carbon dioxide is generated from the combustion of fossil fuels and the oxidation of organic matter. Methane gas is produced during the decomposition of organic matter under anaerobic conditions. Mathematical modelling helps in evaluating various scenarios. There are three ways to minimize greenhouse gases emission: (1) change of operational conditions (2) treatment of gaseous streams (3) prevention by applying the new configuration (Campos et al., 2016). Treatment of gaseous streams is not a feasible option as it involves huge capital costs. Change of configuration by using microalgae or partial nitration process instead of conventional nitrification-denitrification can reduce greenhouse gas emissions but currently not much information is available about this process. Therefore, the most economical way to minimize greenhouse gases is through modification of the operational conditions of WWTPs units.

2.2 Optimization

Sewage sludge is the solid, semisolid or slurry residual material that is produced as a by-product of the wastewater treatment process. This residue is classified as primary and secondary sludge. Primary sludge is generated from sedimentation, chemical precipitation, and various primary processes. Secondary sludge is a waste resulting from biological treatments. Treatment of sewage sludge includes thickening, digestion, and dewatering. Aerobic digestion is a process that occurs in the presence of oxygen. The objective of autothermal thermophilic aerobic digestion (ATAD) is to stabilize and pasteurize the sludge (Rojas & Zhelev, 2012). Conventional systems make use of invariable air supply regardless of the strong variations of bacterial activity during the reaction. Continuous and invariable rate of aeration leads to higher consumption of energy. A dynamic optimization approach can minimize the energy requirement of ATAD (Rojas & Zhelev, 2012).

2.2.1 Multi-variable optimization

Single variable optimization has been widely studied especially for dissolved oxygen concentration, to improve the optimal performance multi-variable optimization strategies are proposed (Qiao & Zhang, 2018; Egea & Gracia, 2013). Most studies have focused on the optimization of specific units that comprise a wastewater treatment plant. The results fail to represent true optimal conditions because of a lack of interactions with other unit operations (Asce et al., 1983). Moreover, optimal design and operation of a wastewater treatment plant requires several conflicting objectives to be optimized simultaneously.

Effluent quality and energy consumption, two contradicting objective functions were studied by (Béraud et al., 2007). The goal was to find a global solution because wastewater systems are highly complex and non-linear. Optimization for the reduction of operational greenhouse gas emissions from the wastewater treatment plant in a cost-effective manner can be done using control strategy (Sweetapple et al., 2014) . Multi-objective optimization can facilitate a significant reduction in greenhouse gas emissions without the need for plant redesign. Another study focused on simultaneous minimizing operating cost, greenhouse gas emissions and maximizing effluent quality (Kim et al., 2015). Multi-objective optimization improved the effluent quality by 2%, and reduced greenhouse gas emission and operating costs by 31% and 11% respectively. Optimization done using the simplex method was quite time-consuming and efforts should be made to reduce the computational time (D. Kim et al., 2015). Data-mining approach can minimize energy consumption and maximize the pumped wastewater flow rate (Zhang et al. 2016). The optimization model was solved by an artificial immune network algorithm. The results indicated that the efficiency of the pumping system can be increased by using optimal pump speeds.

2.2.2 Dynamic multi-objective optimization

The WWTP is a non-linear dynamic system with many fluctuations, such as influent flow, pollutant concentration and weather variations. Multi-objective optimization provides a group of equally excellent Pareto optimal solutions. One of the challenges modelling and optimization of wastewater treatment faces is that sometimes data is noisy, uncertain, and incomplete. Another challenge is the optimization of the process model. Models are usually non-linear and dynamic therefore evolutionary computational algorithm is proposed for this research work. A dynamic multi-objective optimization, to reflect the dynamic characteristic of WWTP MOO was constructed by a neural network (Qiao & Zhang, 2018).

Mechanistic models have been used for modelling biological wastewater treatment systems. However, in these models' many empirical parameters are difficult to estimate and many chemical and biological species are lumped into one model component. To overcome this problem machine learning and computational intelligence can be used. Wastewater treatment plants are highly complex therefore single simulation can take hours to complete. To reduce computation time surrogate models are used, as they can replicate the behavior of a simulation model. Surrogate modelling has been used in various fields of engineering due to its use in computationally expensive such as Monte Carlo based global sensitivity analysis and process design optimization (Al et al., 2018). The potential of surrogate modelling in multi-objective optimization was explored to evaluate the efficiency of this method, results were compared with NSGA-II (Fu et al., 2009). The objective functions chosen for study were maximization of dissolved oxygen and minimization of ammonium concentration. Comparing the results, it was clear that ParEGO a surrogate-based method had lesser objective evaluations.

2.2.3 Prediction of influent quality

Estimation of influent flow rate is important for selecting pump configurations and their speed settings. A few main sewers are equipped with sensors. Therefore, it is difficult to determine the influent flow rate to WWTP based on data from these sensors. Extreme weather events pose one of the major threats to drinking water treatment plants (Whitehead et al., 2009). When precipitation falls onto impervious surfaces it drains as stormwater runoff. Extreme rainfall and wet weather events can generate large quantities of stormwater. High flow rates can impact the performance of the treatment plant if they exceed the capacity of the treatment facility. Sewage treatment plants are designed to treat 2-6 times the average dry weather flow. The total potential inflow into the combined sewer system is higher than its design flow during wet weather. The maximum capacity of the plant can be reached even at an early stage of a storm. Improper operation of wastewater treatment plants may bring about serious environmental and public health problems (Hamed et al., 2004). The high cost involved in the operation of wastewater treatment facilities has led to the use of simulation models to optimize the performance. Accurate prediction of water quality will provide knowledge for intelligent decision-making regarding ecological conservation. Moreover, the early prediction will ensure the smooth operation of the treatment plant.

Global growth of population and industrial development has led to an increase in daily water consumption. Wastewater treatment plants are energy intensive. For instance, higher biochemical oxygen demand (BOD) concentration requires longer aeration. Therefore, for managing wastewater treatment plants and maintaining effluent quality prediction of influent flow is suggested. Forecasting of influent flow is useful in the optimum operation of wastewater devices and pumps. The influent flow rate estimated by plant operators based on experience and weather forecast is not accurate. Knowing the amount of influent flow a few hours or even days ahead can reduce the impact of diurnal flow.

Influent flow rate is a combination of the contribution of households, industry, rainfall, and infiltration. Rainfall contributes to the total flow rate in two ways: 1. Major portion of rainfall is directly transported to the sewer 2. Rainfall on permeable surfaces will influence groundwater level. The cold and warm season modifies the amount of infiltration therefore a seasonal correction factor will be combined with the rainfall falling on permeable surfaces. The net contribution of infiltration will be combined with the overall flow rate resulting from households, industry and flow contribution from rainfall on impermeable surfaces (Flores-Alsina et al., 2011). During the cold season, the groundwater level is high resulting in high infiltration into the sewer system.

Various approaches for generating influent data are phenomenological models, models based on harmonic functions, and data-driven methods based on creating databases with monitoring and experimental data. Phenomenological models are detailed influent models, that give a phenomenological representation of dynamics of WWTP influent, including weekend, seasonal and holiday variations. Phenomenological approach was applied for the generation of dynamic influent flow rate (Gernaey et al., 2005). Despite being a promising model cannot reproduce the dynamics in WWTP during wet weather due to poor representation of the buildup and wash-off of pollutants (Martin & Vanrolleghem, 2014). Models based on harmonic functions are well suited for dry weather conditions and less for wet weather flow(WWF) situations (Martin & Vanrolleghem, 2014).

2.3 Data mining

Data mining is a promising approach for building prediction models. As this method does not require a physical understanding of the system to be modelled. In the past several techniques have been applied to predict influent flow which includes the time series model. Time series models are categorized into two types: statistical methods, e.g., Autoregression, Moving Average, Autoregressive moving average (ARMA), and Autoregressive integrated moving average. The second type includes AI methods such as support vector machine (SVM), artificial neural networks, etc. Autoregressive Integrated moving average (ARIMA) is a statistical analysis model that uses time series data to better understand the data set and predict future trends. Influent flow was forecasted using AIRMA (Boyd et al., 2019). RMSE, MAPE and R-squared were used to analyze the results. The results from this study were found to be acceptable in predicting influent flow, though it was suggested that hybrid models can be used to improve the accuracy of results. Four data mining methods namely Random forests (RF), support vector machines (SVM), Kernel regression(K) and k-nearest neighbor(k-NN) were applied to predict the inflow of wastewater into the Rzeszow city plant (Szelag et al., 2017). A k-NN algorithm was used to predict influent flow rate, chemical oxygen demand (COD), suspended solid (SS), total nitrogen(TN) and total phosphorus(TP) (Kim et al., 2016). In another study ARIMA, NAR, and SVM were used to predict the inflow of sewage treatment plants (Ansari et al., 2018).

2.3.1 Applications of ANN

Traditional methods of modelling require rate constants of various physical, chemical and biological processes, which depend largely on space and time (Emamgholizadeh et al., 2014). In recent years several pieces of research have been conducted on forecasting water quality using non-linear models such as artificial neural network, and adaptive neuro-fuzzy inference systems.

The ANN model was used to compute COD and BOD levels in the Gomti river, India (Singh et al., 2009). It was successfully demonstrated in a research that ANN can predict influent water quality parameters with a correlation coefficient between observed and predicted output values reaching up to 0.93 (Aminabad et al., 2014). Statistical-based water quality models assume a linear relationship between input and output. ANN works well when the diversity of data is large and the relationship between variables is not clearly understood or is difficult to describe with conventional approaches (Singh et al., 2009). A NN based model was developed for predicting WWTP performance (Hamed et al., 2004). In one study a MLP based model was selected to build a prediction model of influent flow rate (Wei et al., 2013). The model was able to predict influent flow rate up to 180 min ahead, however, MAE and MSE increased with a longer time horizon. To overcome this problem one study proposed a dynamic neural network(DNN) with an online corrector (Wei & Kusiak, 2015). This method was able to provide good prediction up to 300 min ahead with 85% accuracy.

Sewers are prone to corrosion due to the production of hydrogen sulfide generated in sewage under anaerobic conditions (Jiang et al., 2009). To minimize the formation of hydrogen sulfide, chemicals such as oxygen, nitrate, magnesium hydroxide and iron salts are added. The main dosing strategies commonly used are (1) constant dosing: In this, the dosing rate is maintained at a constant value without considering variations of the wastewater characteristics. (2) flow-passed dosing: the chemical dosing rate is proportional to the sewage flow rate. (3) profiled dosing: dosing rate is according to a predefined profile. Real-time prediction of sewage flow is necessary for optimal addition of chemicals (Li et al., 2019)

An autoregressive moving average was developed for real-time prediction of future flow in sewers (Chen et al., 2014). The model was able to predict future flow rates with good accuracy under different weather conditions. However, the prediction accuracy for wet weather conditions was lower as compared to dry weather conditions. The autoregressive with exogenous inputs (ARX) model had better prediction accuracy as compared to the autoregressive model for real-time prediction of sewage flow (Li et al., 2019). Generally, models do not consider the additional flow to sanitary systems during rainfall. Due to this additional flow error between the predicted and actual flow occurs. Models based on hydrology require detailed knowledge of the system and rely on a large number of parameters, which are difficult to determine (El-Din & Smith, 2002). In this research, the ANN model will be used to make a long-term prediction of the wastewater inflow rate to the plant.

2.4 Summary

Wastewater treatment is a complex system, and it is difficult to explore various design ideas on a pilot plant. Modeling helps in understanding how a system would behave in various conditions without experimentation. A WWTP model is a representation of physical and chemical processes involved in the purification of wastewater. This chapter presents the work of researchers in the area of modeling and simulation of wastewater treatment. Most of the work done in the past employs first principle methods to model the treatment plant and not much work has been done which utilizes machine learning and black-box modeling. WWTPs exhibit non-linear behaviour and therefore becomes difficult to describe with first principle methods. In my research work, a black-box modeling approach has been employed to model WWTP and predict the performance of the plant. This type of modeling is based on the input-output behaviour of the process in contrast to physical modeling which is time-consuming.

Chapter 3

3 Description of plant and statistical analysis

3.1 Introduction

The operators of wastewater treatment plants are constantly facing challenges due to stricter regulations and aging infrastructure. To address the increasing environmental issues WWTPs should be operated properly, and effluent quality should be continuously monitored. Data generated from wastewater treatment plants(WWTPs) can be utilized to improve the quality of effluent discharge. The risk involved in proper management of the system can be reduced by data analysis and constant monitoring of the plant. Analysing historical and real-time data can help in better management and decision-making of the plant.

3.2 Description of plant

Adelaide WWTP is located at 1157 Adelaide Street North and treats about 15% of the total wastewater produced in London. The plant capacity of Adelaide WWTP is 36,650 cubic metres per day and the approved peak flow rate is 59,000 m^3/d. The wastewater enters the treatment plant through screens where grit removal takes place in a vortex-type grit removal chamber. After preliminary treatment wastewater enters primary treatment tanks (clarifiers). Solids are removed in the primary clarifier this process is also called primary settling. The treated water flows to aeration tanks for secondary treatment. Air is pumped into the aeration tanks to provide oxygen for the bacteria to grow. Secondary treatment involves the conversion of organic compounds into carbon dioxide and water. The treated wastewater is then pumped to secondary settling tanks to allow solids to settle down. These solids are called activated sludge and mostly consist of active bacteria. A portion of this activated sludge is returned to the aeration tanks, and this is termed as returned activated sludge.

The purpose of tertiary treatment is to improve the quality of wastewater before discharging it into the environment. As water passes through the UV unit, an ultraviolet range of light inactivates harmful bacteria, particularly E.coli. After proper treatment of wastewater, it is important to test the quality of effluent in terms of dissolved oxygen, pH level, suspended solids, total phosphorous and ammonia.

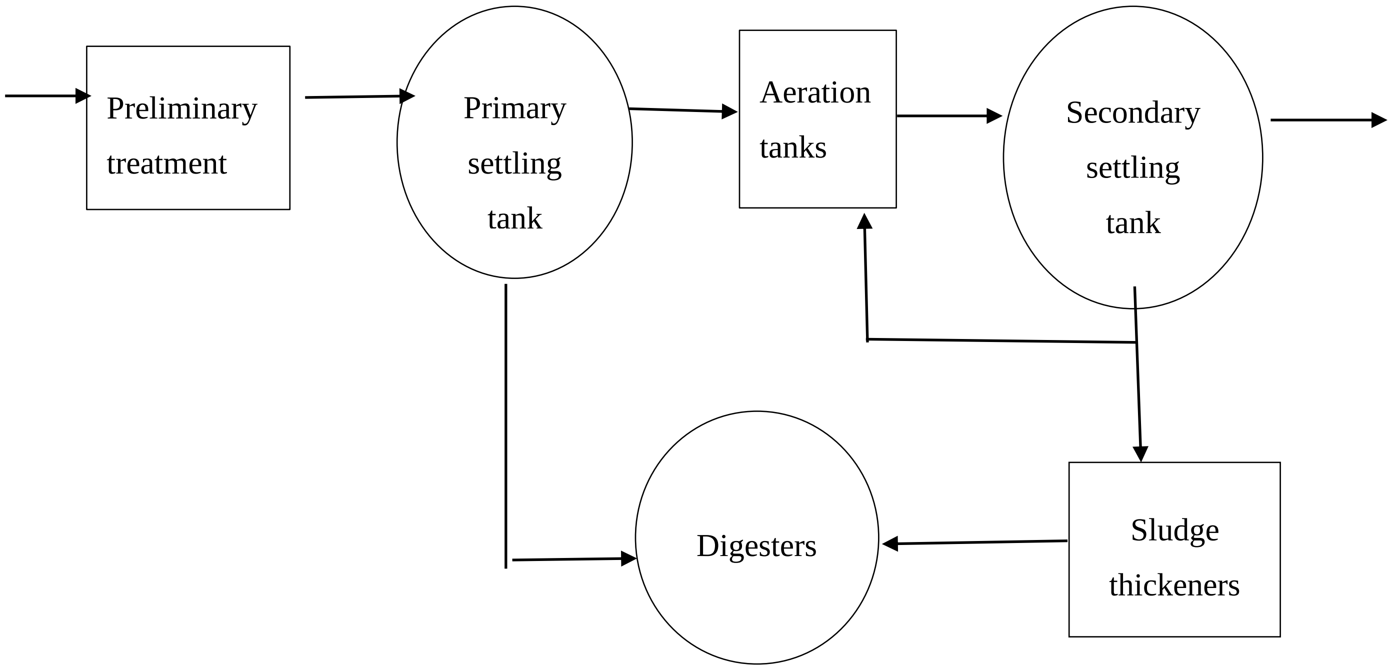


Figure 3.1 Schematic diagram of WWTP

Contamination of freshwater sources by nutrients induces excessive growth of algae and reduces oxygen levels in the water. This lack of oxygen also known as eutrophication suffocates plants and animals and can create dead zones. Major nutrients that contribute to algal bloom are carbon, nitrogen, and phosphorus. Excess nutrient concentration in water bodies can lead to depletion of oxygen thereby making it inhospitable for aquatic life.

The main source of nitrogen in water are fertilizers that contain ammonia, ammonium, urea, nitrate, and amines. Phosphorous is a natural component of biological tissue. Detergents and personal care products are also major contributors to phosphorus in wastewater. Ammonia is commonly used in various cleaning solutions, and fertilizers and is also a component of human and animal waste. A high concentration of ammonia in freshwater bodies can lead to algal bloom. Conversion of ammonia to nitrate in the presence of bacteria causes depletion of dissolved oxygen levels in the water.

Measurement of water quality Indicator

Biochemical oxygen demand(BOD)

BOD indirectly measures the amount of organic matter contained in a water sample. It is a measure of the amount of oxygen consumed by heterotrophic bacteria for the oxidation of organic matter.

$$organic\ matter\ + microorg. + O_2\ + nutrients\ \rightarrow\ CO_2 \uparrow\ + H_2O\ + microorg.$$

Hence, BOD is expressed in mg O_2/L and the higher the amount of organic matter in wastewater the higher the BOD value. BOD measures the aerobic degradation of organic matter and chemical oxidation of inorganic matter. It is measured by comparing the dissolved oxygen concentration before and after a 5-day incubation period. The test method uses 5 days incubation period because after 5 days most of the organic matter degrades. The test sample is continuously agitated in the absence of light and a sensor measures the decline in pressure caused by the consumption of oxygen.

Table 3-1:BOD values

BOD level in mg/L	**Water quality**
1-2	Very Good: Not much organic matter is present
3-5	Fair: Moderately Clean
6-9	Poor: Somewhat Polluted
100 or more	Very Poor: Polluted

Suspended Solids(SS)

Suspended solids in water are due to fine particles of soil or organic material such as algae. Suspended solids are visible if they are present in substantial quantities. When the concentration of suspended solids exceeds a certain limit, can be harmful to the aquatic organisms. A high concentration of suspended solids can block light from reaching submerged vegetation. As the amount of light passing through the water is reduced, photosynthesis slows down and causes less dissolved oxygen to be released into the water by plants. A higher concentration of suspended solids can also cause an increase in surface temperature because suspended solids absorb heat from the sunlight. This can cause dissolved oxygen levels to fall even further because warmer waters hold less dissolved oxygen.

The concentration of suspended solids is measured in laboratories by filtering a known volume of sample, the filter will capture the suspended solids and let the water pass through. Then, the filter is dried in an oven to remove the moisture. The weight of the suspended solids is evaluated by calculating the difference in the weight of the filter with suspended solids and the weight of the filter.

$$SS\left(\frac{mg}{L}\right) = (Wfss - Wf)/Vs$$

Where:

Wfss:-weight of the filter with suspended solids after drying in mg

Wf:- the weight of filter in mg

Vs:- Volume of sample in L

Suspended solids in a sample affect the light scattering properties of the sample. The light scattering is proportional to the concentration of suspended solids in the sample. Installing a sensor that measures the concentration of suspended solids can give instantaneous information about the sample at that location, unlike the above method which takes around 2 hours to complete. Online measurement is a continuous method of monitoring the quality of the sample. Suspended solids sensors are calibrated in unit’s mg/L and diatomaceous earth(DE), primarily composed of silicon dioxide is used as a standard for turbidity calibrations.

Ammonia

Ammonia is a critical nutrient in wastewater that causes water pollution. This is an essential nutrient that is consumed by bacteria to break down organic compounds. Insufficient availability of ammonia will cause production of excessive filaments and hence will interfere with the compaction of sludge. Presence of nutrients is important for removal of BOD and deficiency of nutrients will make it difficult for bacteria to grow. Online monitoring of ammonia ensures that overloading of ammonia is avoided. However, for laboratory testing ammonia-sensitive electrode is used which uses hydrophobic gas-permeable membrane. This membrane separates the sample solution from internal solution of ammonium chloride. Ammonia diffuses through the membrane and changes the pH of internal solution. This change in pH is sensed by a pH electrode and concentration of ammonia is proportional to change in pH. Salicylate method is useful for determination of small quantities of ammonia in the sample. In this method the ammonia present is converted to intense blue indophenol, which is then quantified by UV-visible spectrometry (Giner-Sanz et al., 2020).

Total Phosphorus

Phosphorus is a nutrient that is essential for plant, animal, and human growth. However, an increase in the concentration of phosphorus in water bodies can accelerate plant growth, algal bloom, and low dissolved oxygen thus affecting aquatic life. Phosphorus entering the treatment plant is removed by primary and secondary treatment. In a treatment plant phosphorus comes in either soluble form or particulate form and the goal is to convert the phosphorus into particulate form. There are two techniques of removal of phosphorus one is chemical treatment and the second is biological treatment. Chemical treatment for phosphorus removal involves the addition of metal salts to react with soluble phosphate to form precipitates. These precipitates are then removed by physical separation processes such as filtration and clarification. The commonly used metal salts are alum(aluminium sulphate), ferric chloride and ferrous chloride. The basic reaction involved in chemical phosphorus removal is a reaction between metal ions and phosphate that is dissolved in water leading to metal phosphate that is insoluble in water and acid as a by-product.

$$Metal\ ions + Phosphate \rightarrow Metal\ Phosphate + Acid$$

Enhanced biological phosphorus removal(EBPR) is accomplished by microorganisms and these organisms are called phosphorus accumulating organisms(PAOs).

3.3 Materials and Methods

The methodology of data analysis for wastewater treatment process is shown in figure 3.2.

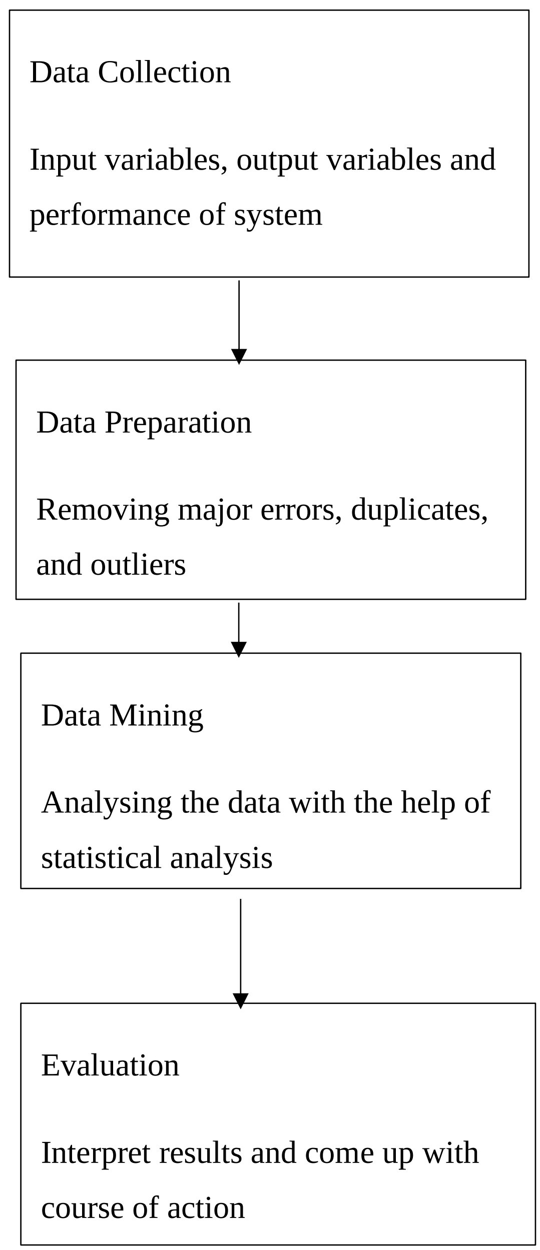


Figure 3.2 Flow Chart of Data Analytics

3.3.1 Data collection and preprocessing

The data of Adelaide WWTP was collected and carefully analysed for the years 2015, 2016, 2017 and 2018. Data collection is one of the most important steps in building machine learning models. The quality of data can influence the usability and relevance of machine learning models. Data preprocessing is the process of converting raw data into a suitable form. The purpose of data preprocessing is to clean the raw data set and remove missing data, and other inconsistencies. Every day data of Adelaide WWTP for a period of four years was analysed and if for a particular day information about a particular variable (e.g. BOD_5, SS, TP, etc.) was missing that data entry was discarded. After analysis 1460(365x4) data points, some of the variables were missing and was left with dataset of 280 data points. A dataset(10 data points) of Adelaide WWTP is shown in Table 3.2.

Table 3-2: A dataset of Adelaide WWTP

Influent Temp (°C)	Effluent Temp (°C)	Total sewage flow ML/D	Influent BOD_5 mg/L	Effluent BOD_5 mg/L	Influent SS mg/L	Effluent SS mg/L	Influent TP mg/L	Effluent TP mg/L	Influent NH3 mg/L	Effluent NH3 mg/L
17	15	23.62	196	2	343	4	10.3	0.26	18.1	0.1
15	14.7	21.63	147	4	123	6	5.4	0.42	24.6	0.62
20	21.1	22.75	166	1	205	1	8.1	0.56	34.4	0.1
21	21.5	21.93	210	2	273	1	6.3	0.72	24.7	0.1
20	19.9	18.77	286	3	349	6	7.7	0.54	26.1	1.37
19	20.6	29.87	169	1	289	2	5.1	0.22	27.3	0.1

20	20.3	26.18	154	1	333	2	4.3	0.28	18.5	0.12
21	21.4	27.79	211	2	417	3	7.2	0.38	22.3	0.14
18	17.8	38.47	301	1	325	4	11.2	0.69	32.9	0.96
17	17.7	31.27	282	1	256	2	6.3	0.26	20.8	0.65

Table 3-2: A dataset of Adelaide WWTP (contd.)

Data point	Influent pH	Effluent pH	Effluent DO mg/L
1.	7.4	7.3	7.9
2.	7.4	6.9	7.3
3.	7.5	7.2	5.5
4.	7.2	6.9	7.5
5.	7.4	6.9	7
6.	7.6	7.2	7
7.	7.7	7.4	7.9
8.	7.5	7.3	7.5
9.	7.7	7.5	5.8

10.	7.6	7.4	7

3.3.2 Data mining

Data mining is the process of extracting useful information from raw data by identifying patterns and relationships between variables. In this research work total of fourteen variables were analysed. The influent stream is characterized by seven parameters (Temperature, BOD, SS, pH, TP and NH_3) and the effluent stream is characterized by seven parameters (Temperature, BOD, SS, pH, TP, NH_3 and DO).

3.4 Results

3.4.1 Data visualization

Data visualization is the graphical representation of data. The scatter plot of parameters temperature, biological oxygen demand(BOD), suspended solids(SS), total phosphorous(TP), and ammonia(NH_3) in the influent and effluent stream of Adelaide WWTP are shown in figures 3.3-3.7.

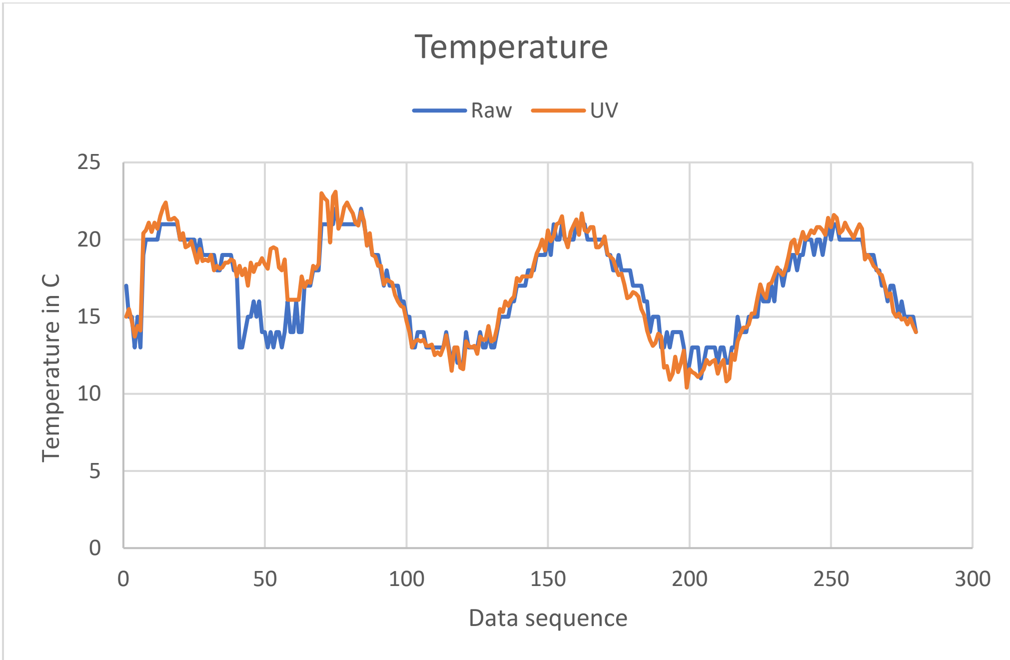


Figure 3.3 Scatter plot of temperature in influent and effluent stream

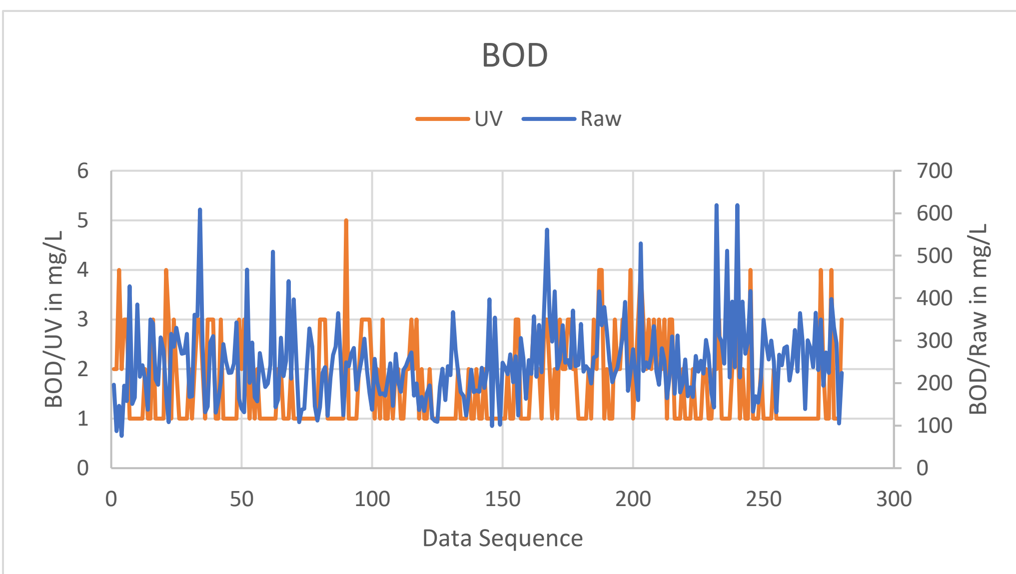


Figure 3.4 Scatter plot of BOD in the influent and effluent stream

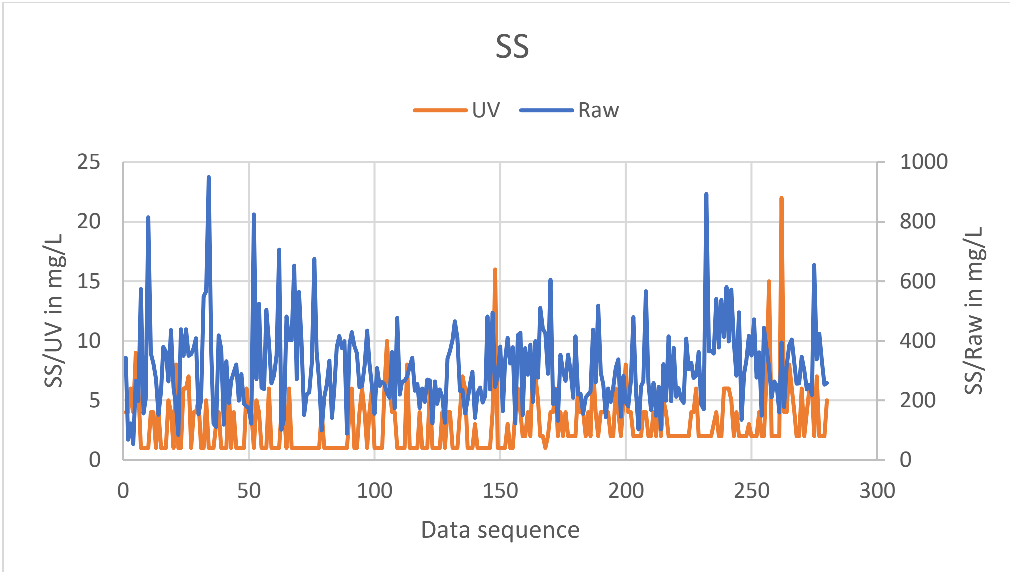


Figure 3.5 Scatter plot of suspended solids in the influent and effluent stream

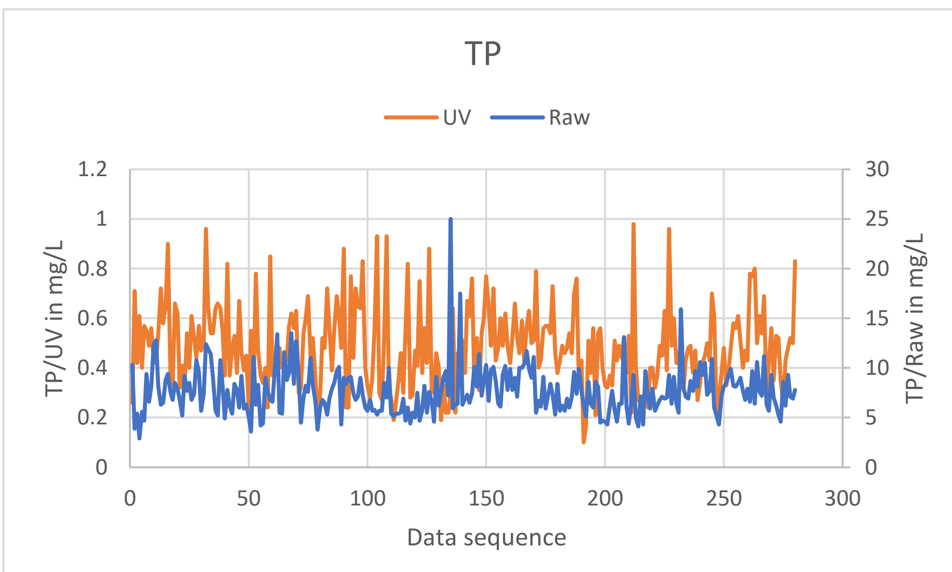


Figure 3.6 Scatter plot of total phosphorous in the influent and effluent stream

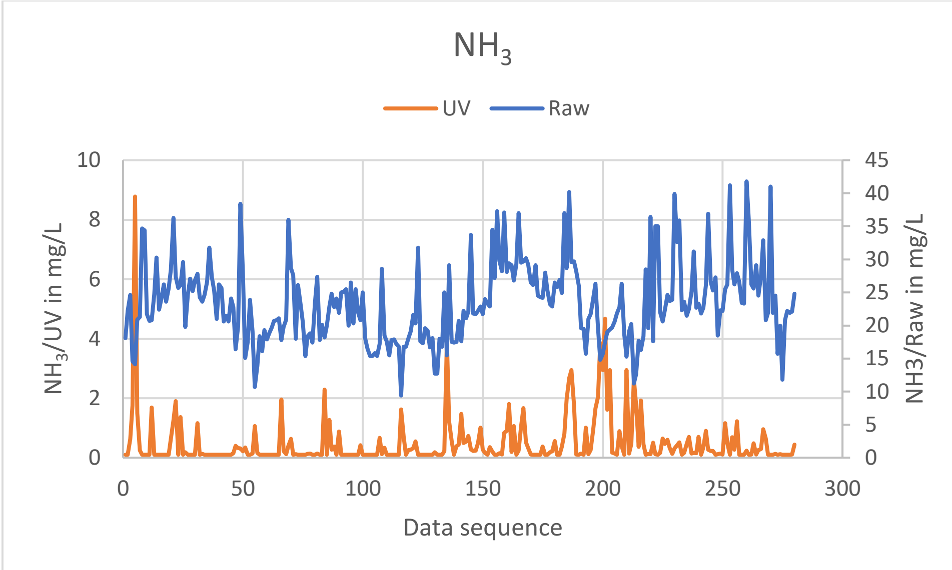


Figure 3.7 Scatter plot of ammonia in the influent and effluent stream

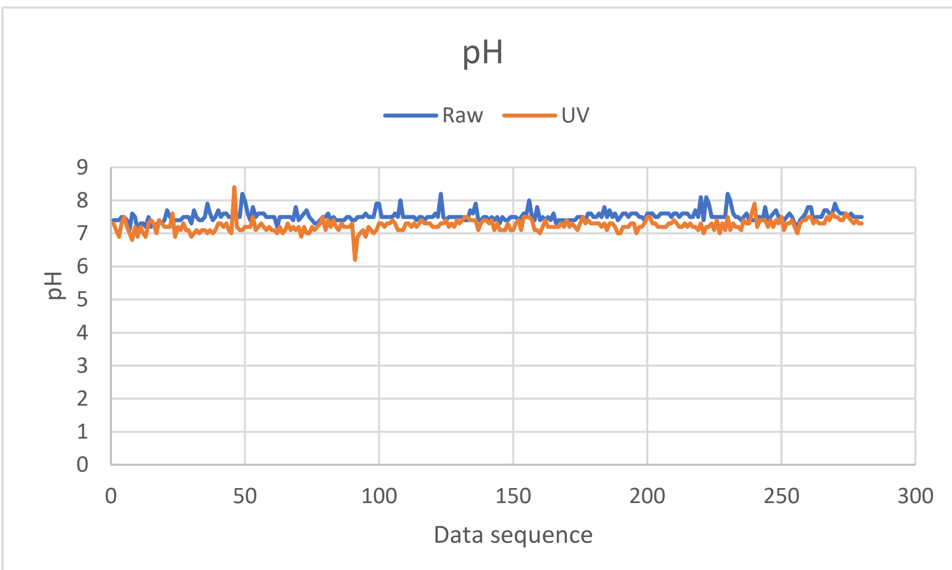


Figure 3.8 Scatter plot of pH in the influent and effluent stream

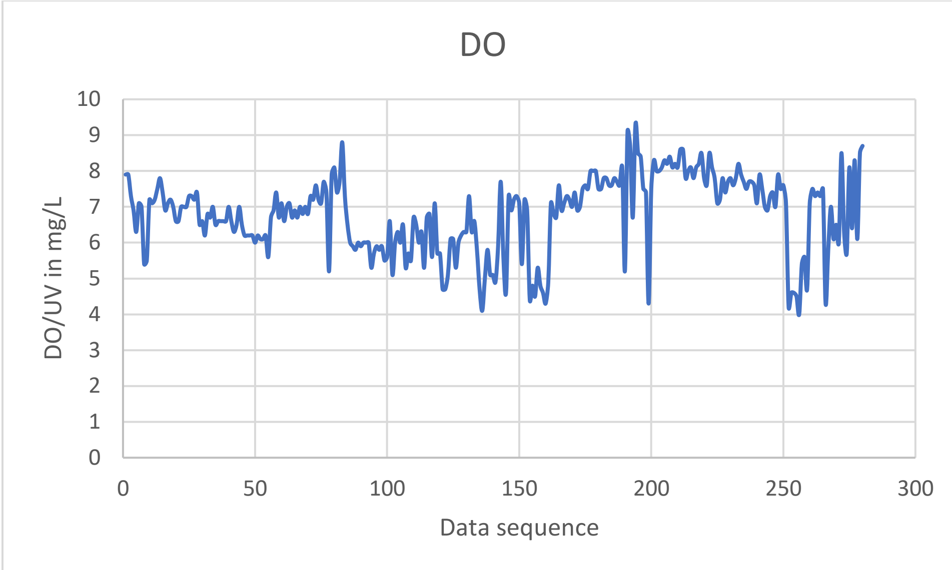


Figure 3.9 Scatter plot of DO in the effluent stream

3.5 Discussion

Data mining involves the use of statistical methods to identify patterns in data and help predict future trends and behaviour. It focuses on finding relevant information and data sets which can be used for analytics and predictive modeling. Statistical analysis is the collection and interpretation of data. Identifying patterns from historical data will help in better management and decision-making of wastewater treatment plants. The efficiency of the plant can be increased by applying data analytics and optimizing plant operations.

Table 3.2 -Statistical Analysis of Adelaide WWTP data

	unit	**Minimum**	**Maximum**	**standard deviation**
Temperature sewage raw	°C	11	22	2.93
Total sewage flow	ML/D	16.4	69.5	5.6

BOD 5 day Raw	mg/L	76	619	90.8
Suspended solids Raw	mg/L	53	950	137.8
pH Raw	-	7.1	8.2	0.16
Total Phosphorus raw	mg/L	2.9	25	2.4
NH3 raw	mg/L	9.4	41.8	5.9
Temp UV	°C	10.4	23.1	3.2
BOD5 UV channel	mg/L	1	5	0.89
Suspended solids UV channel	mg/L	1	22	2.5
pH UV channel	-	6.2	8.4	0.18
Total Phosphorus UV	mg/L	0.1	0.98	0.16
NH3 UV	mg/L	0.1	8.78	0.85
DO plant effluent	mg/L	4	9.3	1.1

3.5.1 Correlation coefficient

The correlation coefficient is used to understand the relationship between two variables. The most used correlation coefficient is Pearson product-moment correlation. It is calculated by dividing the sample covariance by the product of their sample standard deviation. It measures the strength and direction of the linear relationship between two variables. The value of the correlation coefficient ranges from -1 to +1. A positive correlation coefficient close to 1 indicates a strong linear relationship between the variables, zero correlation indicates no relation and negative correlation indicates linear correlation in opposite direction. However, the Pearson correlation coefficient does not indicate any non-linearity between the variables. Table 3.2 shows the correlation between seven parameters in the influent stream and seven parameters in the effluent stream. The maximum correlation exists between BOD in the influent stream and SS in the influent stream. There is a strong correlation between temperature in the influent stream and temperature in the effluent stream. Suspended solids in the influent stream have a strong correlation with BOD and TP in the influent stream.

Table 3-3: Correlation coefficients matrix

	$Temp_i$	$Flow_i$	BOD_i	SS_i	pH_i	TP_i	NH_{3i}	$Temp_e$	BOD_e	SS_e	pH_e	TP_e	NH_{3e}	DO_e
$Temp_i$	1													
$Flow_i$	-0.38	1												
BOD_i	0.16	-0.16	1											
SS_i	0.215	-0.097	**0.752**	1										
pH_i	-0.265	0.061	-0.328	-0.403	1									
TP_i	0.313	-0.162	0.541	**0.667**	-0.257	1								
NH_{3i}	0.436	-0.384	0.137	0.000	0.392	0.195	1							
$Temp_e$	**0.906**	-0.292	0.117	0.220	-0.212	0.291	0.402	1						
BOD_e	-0.159	-0.003	0.181	0.015	-0.008	-0.072	0.000	-0.225	1					
SS_e	-0.041	-0.069	0.020	-0.050	0.023	-0.009	0.048	-0.071	0.147	1				
pH_e	-0.075	0.123	0.044	-0.099	0.178	-0.098	-0.047	-0.067	-0.017	0.06	1			
TP_e	0.21	-0.202	0.037	-0.019	-0.086	0.111	0.210	0.197	0.138	0.03	-0.046	1		

NH_{3e}	-0.177	0.143	-0.011	-0.104	-0.003	-0.029	-0.136	-0.243	0.250	0.20	0.103	-0.033	1	
DO_e	-0.060	-0.220	0.150	0.041	-0.016	-0.084	-0.049	-0.159	0.134	0.00	0.006	-0.075	-0.005	1

3.6 Summary and conclusions

In this chapter Adelaide wastewater treatment plant is briefly discussed and data obtained from plant is analysed. The purpose of this research is to model the wastewater treatment plant and predict its performance in terms of effluent quality. The data set used to build a model was collected and analysed. The performance of a model is dependent on the preparation of data. Various steps involved in the preparation of data are:-gathering the data, handling missing data, deciding which key factors are important, and splitting the data into training and validation set. The measurements of influent flow rate, temperature, BOD, SS, pH, TP,NH_3 in the influent stream and effluent stream were collected and analysed over four-year period.

Chapter 4

4 Prediction of effluent quality using Artificial Neural Network(ANN)

4.1 Introduction

Modeling helps in understanding how a system would behave in various conditions. For complex systems like wastewater treatment, it is difficult to explore various design ideas on a pilot plant. A wastewater treatment plant model describes the physical and biochemical processes involved in the purification of wastewater. The application of mechanistic models relies on material and energy balances as well as empirical correlations. The wastewater treatment system involves several such equations and correlations. As the complexity of the model increases, the accuracy of mechanistic modeling decreases. To overcome these difficulties black box modeling is suggested.

4.2 Methodology

4.2.1 Structure of ANN

The first step toward a neural network took place in 1943 when Warren McCulloch and Walter Pitts modeled a neural network with electrical circuits. A neural network derives its origin from the human brain. The human brain is a complex, non-linear and parallel computer that can do tasks in a very small amount of time. This inspired researchers to mimic the working of a brain. A neural network can learn and produce outputs for inputs that were not encountered during the training phase. Compared to mechanistic models neuron-based modeling has better accuracy(Mjalli et al., 2007).

ANN has several neurons arranged in a series of layers. Neurons are the core processing unit of the algorithm. These artificial neurons model the neurons present in the human brain. Like a biological brain, each artificial neuron receives a signal process it and transmits the signal to other neurons connected to it. The signal flows from the input layer to the output layer. Neurons of one layer are connected to neurons of the next layer through channels. Each of these channels is assigned a numerical value known as weights. The inputs are multiplied by the corresponding weights and their sum is sent as input to the neurons of the next layer. Each neuron is associated with a specific value called bias. Bias is then added to the input sum, this value is then passed through a threshold function called the activation function.

The result of the activation function determines if the neuron will get activated or not. The activated neuron will transmit data to the neurons of the next layer. This transmission of data through layers is known as forward propagation. A commonly used activation function is sigmoid. If the input value to a sigmoid function is negative, then it will transform the input value close to 0. The sigmoid function will transform a positive number into a value close to 1. Input value close to 0 will be transformed into a value between 0 and 1.

Artificial neural networks are trained by processing input and output data. The predicted output is compared with the actual output to determine the error in prediction. This information is propagated back to the network and based on this information weights are adjusted. The cycle of forward propagation and backward propagation is iteratively performed with multiple input and output data. This process continues until weights are assigned such that the predicted output is close to the target output.

For training and testing purposes the data should be divided into two data sets. The two data sets are called the training set and validation set. The training set is used to train the model, during each epoch the model will be trained to learn about the features of this dataset. A validation set is a set of data separate from the training set and used to validate the model. The performance of the trained network is evaluated from the error function of the validation set.

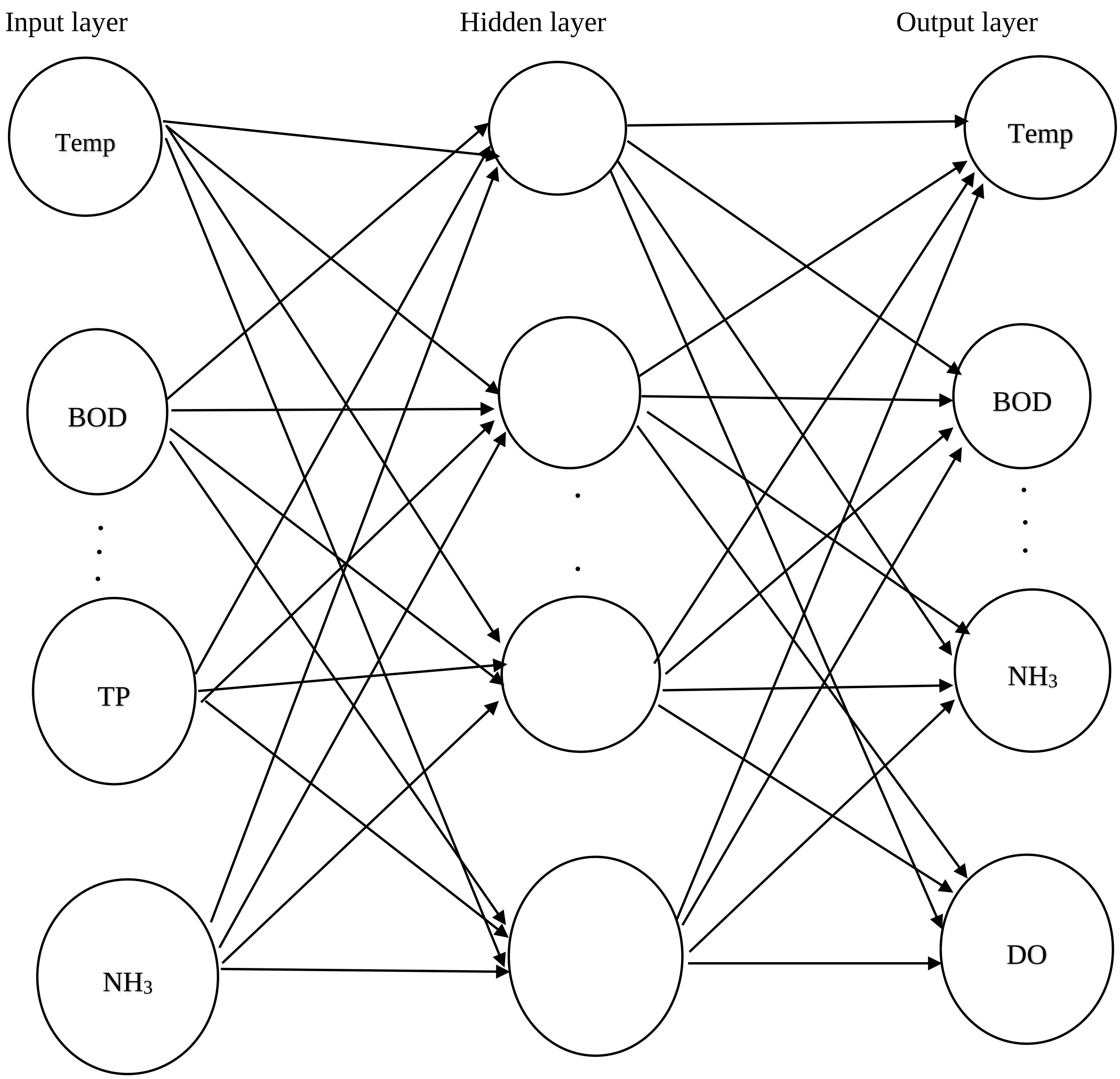


Figure 4.1 Schematic of ANN

To select the architecture of a neural network it is important to decide the number of hidden layers and the number of neurons in each hidden layer.

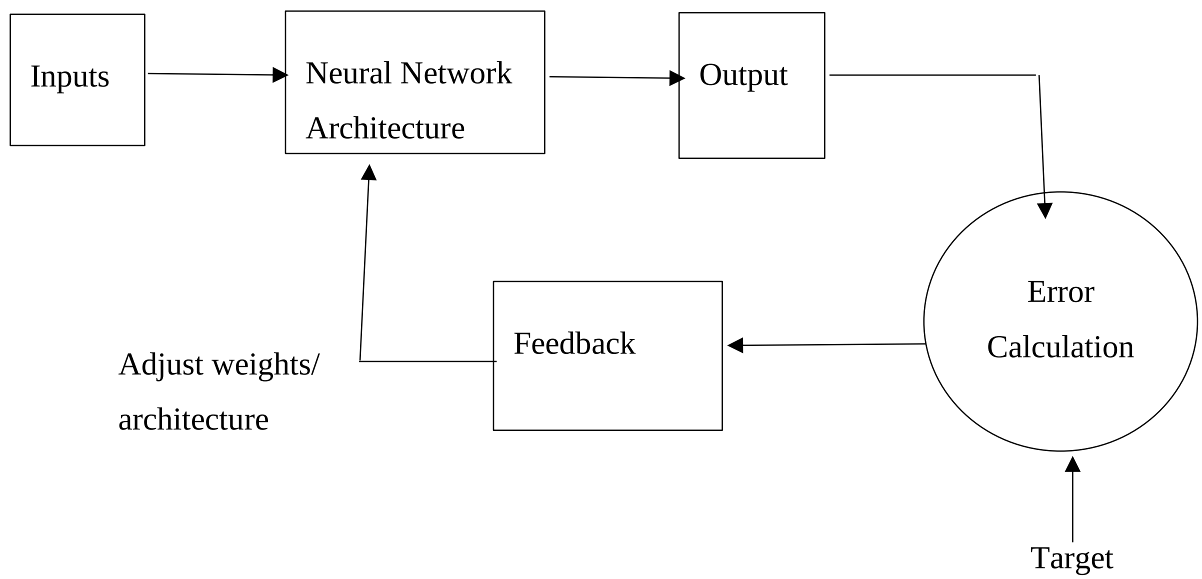


Figure 4.2 Framework of feedback neural network

4.3 Modeling Results

The artificial neural network is used to model the Adelaide WWTP. In this study, MATLAB 2019a is used to model and optimize the WWTP. The goal of this research work is to predict the quality of effluent in terms of four major pollutants namely BOD, SS, TP, and NH3. To predict the pollutants in the effluent stream, the model is trained using seven parameters in the influent stream. The preprocessed data is divided into two set one is the training set and the second is the validation set. The complete data set consists of 280 data points. The data set is divided in the ratio of 7:3 for training and validation.

The network structure consists of three layers input, hidden and output layers. There are 7, 10, and 7 neurons in the input, hidden, and output layers respectively. The number of neurons in the input layer is equivalent to the number of variables in the influent stream. The seven variables in the influent stream are temperature, flow rate, BOD, SS, pH, TP, and ammonia. Similarly, the number of neurons in the output layer is equivalent to the number of variables in the effluent stream. The seven variables in the effluent stream are temperature, BOD, SS, pH, TP, ammonia, and DO. The number of neurons in the hidden layer is decided based on the mean square error(MSE) between the predicted and actual values.

To reach suitable network architecture, several trials were conducted to attain a suitable learning rate, number of hidden layers and number of neurons per each hidden layer. TRAINLM training function is used for optimization. This updates weights and bias values according to Levenberg-Marquardt optimization (Gavin, 2020). This is the fastest backpropagation algorithm and is highly recommended as a first-choice supervised algorithm. The two activation functions used for this architecture are the tan-sigmoidal and linear functions.

4.3.1 Training

The 70% of the data is used for training purposes. The performance of the model is evaluated from regression analysis. The overall R-value for the training dataset is close to 1. Comparisons between modeled values and actual values are shown in figure 4.3-4.9.

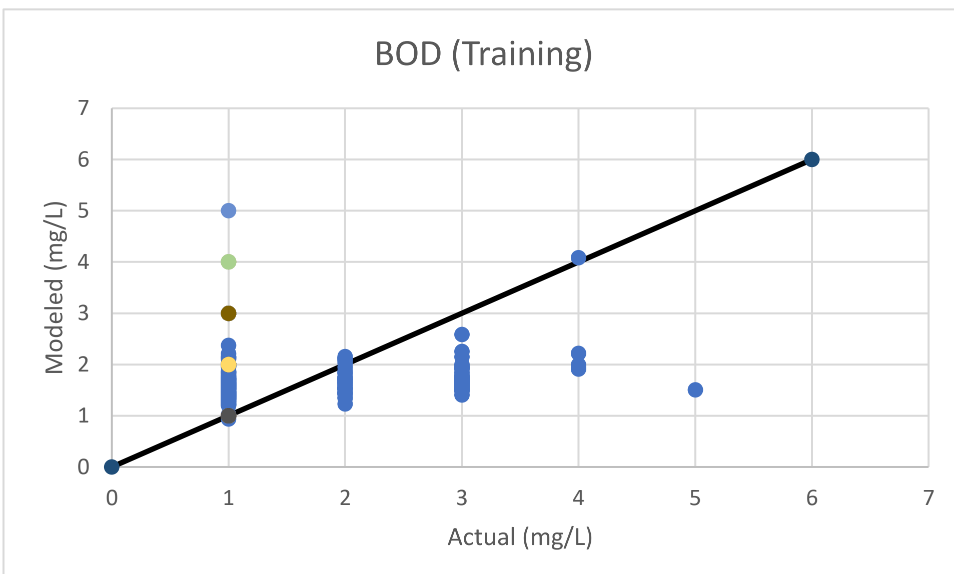


Figure 4.3 Regression analysis of actual versus ANN-modeled BOD concentration in effluent stream

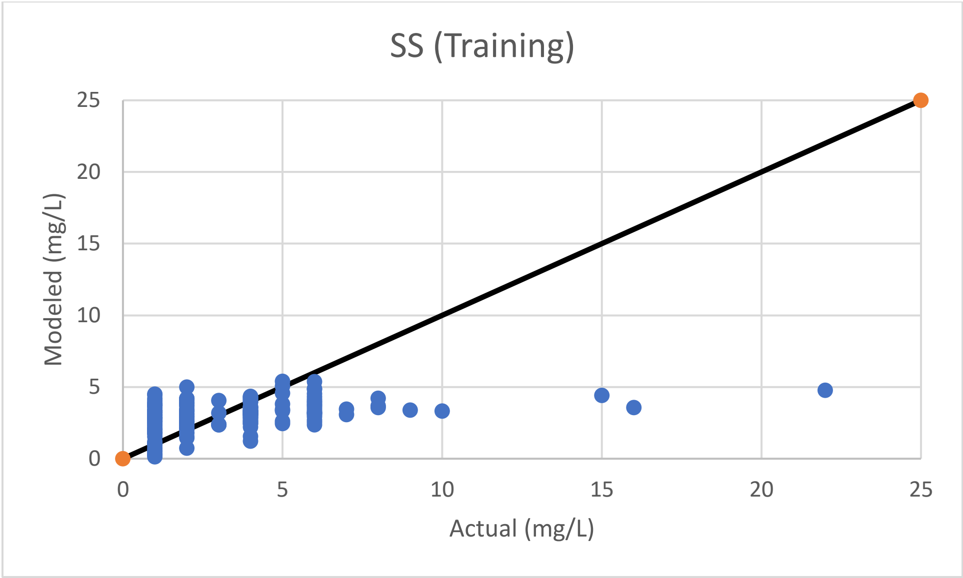


Figure 4.4 Regression analysis of actual versus ANN-modeled SS concentration in effluent stream

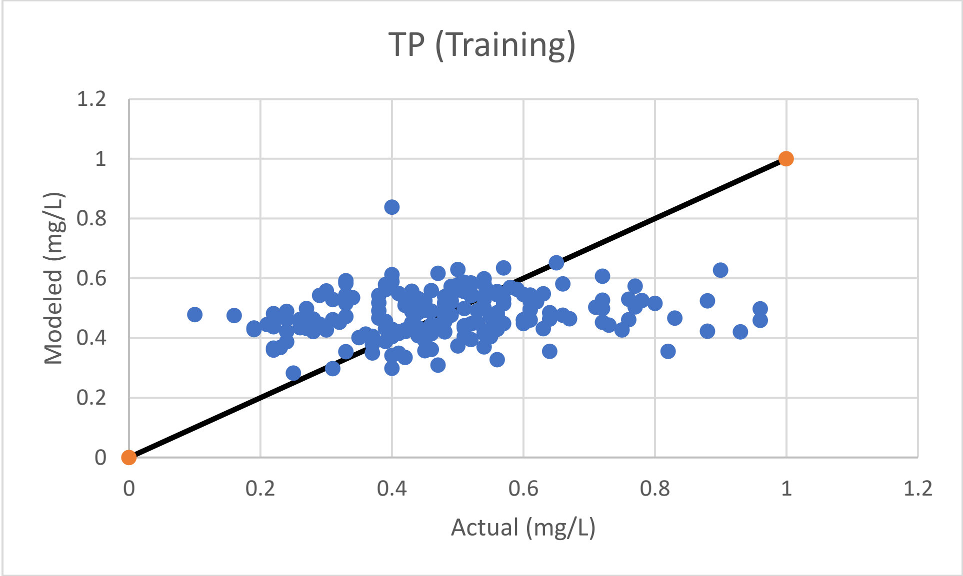


Figure 4.5 Regression analysis of actual versus ANN-modeled TP concentration in effluent stream

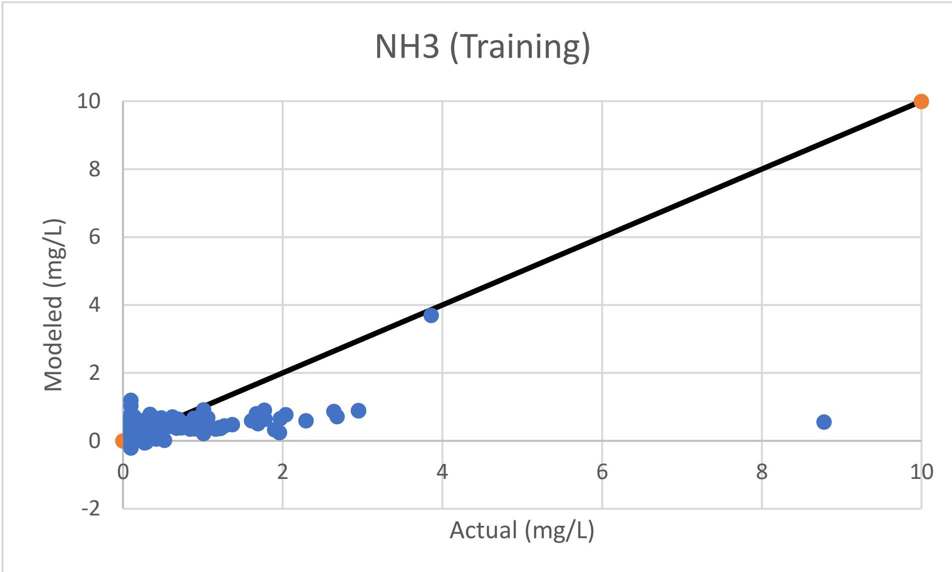


Figure 4.6 Regression analysis of actual versus ANN-modeled NH3 concentration in effluent stream

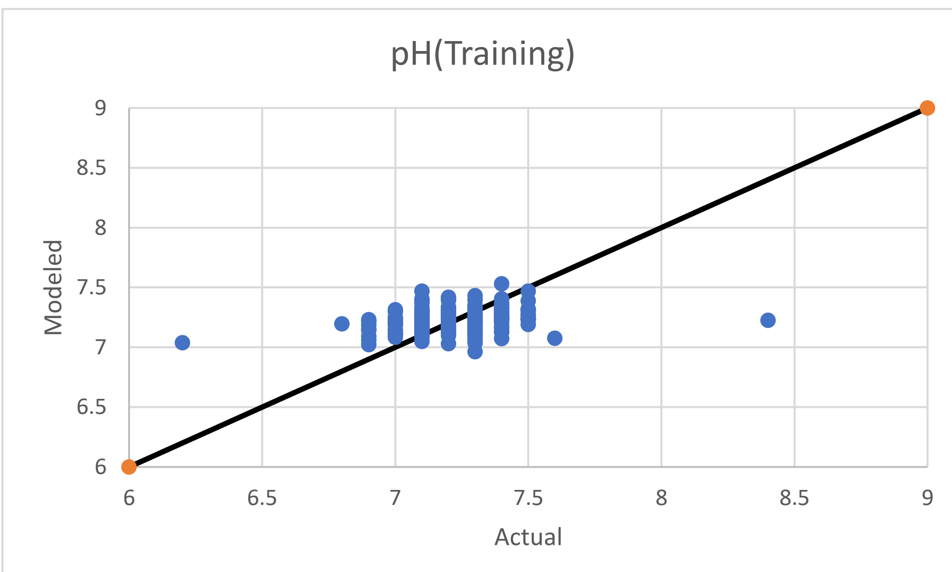


Figure 4.7 Regression analysis of actual versus ANN-modeled pH values in effluent stream

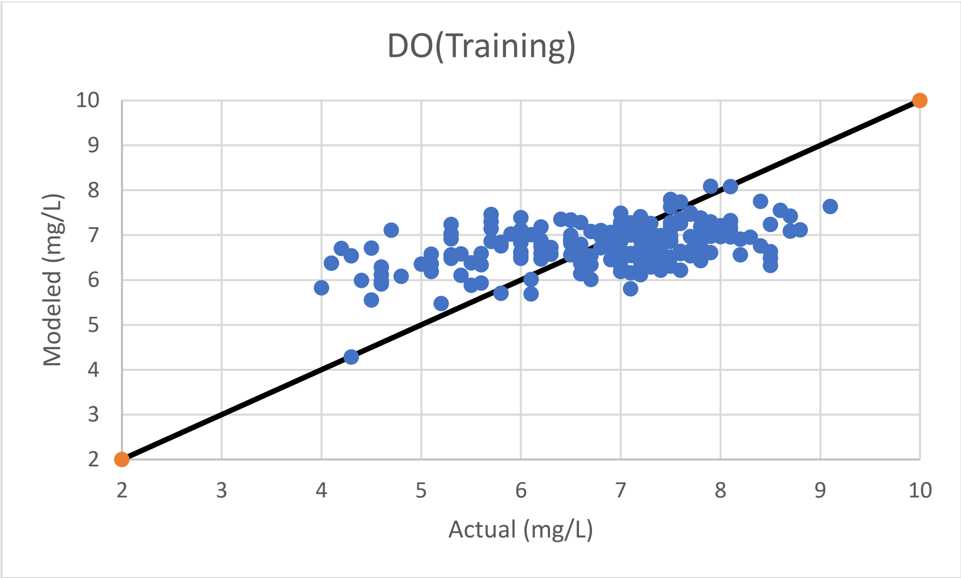


Figure 4.8 Regression analysis of actual versus ANN-modeled DO concentration in effluent stream

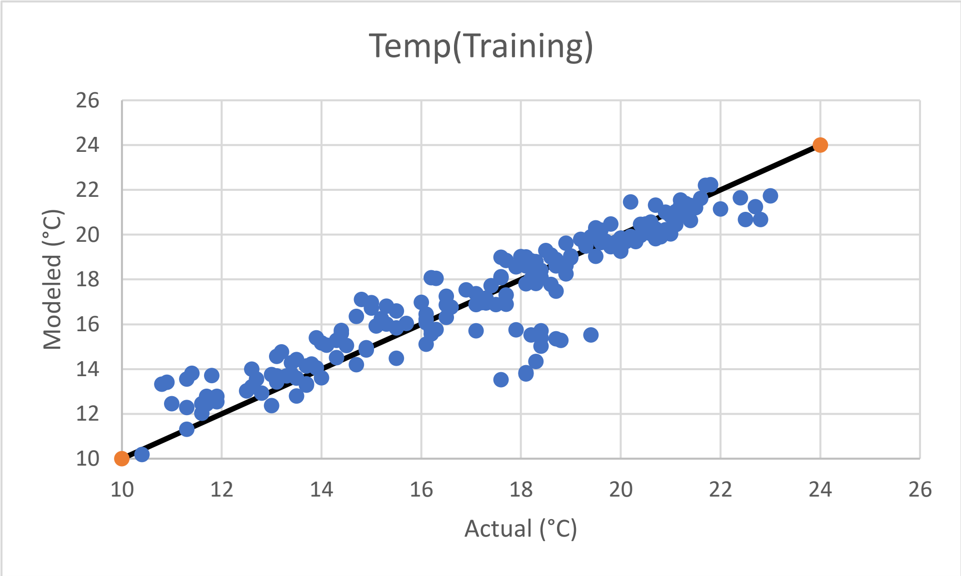


Figure 4.9 Regression analysis of actual versus ANN-modeled temperature values in effluent stream

Validation (30% of data used for validation)

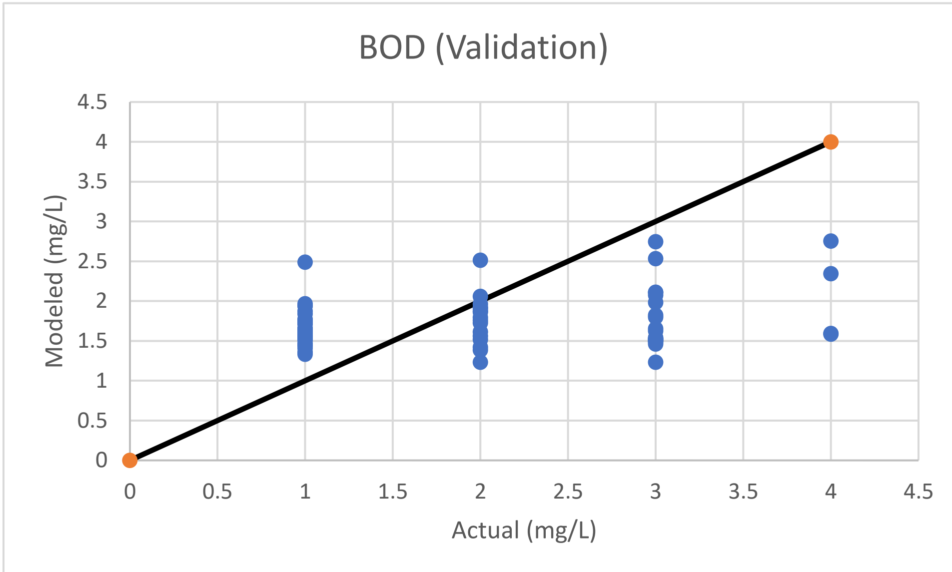


Figure 4.10 Regression analysis of actual versus ANN-modeled BOD concentration in the effluent stream

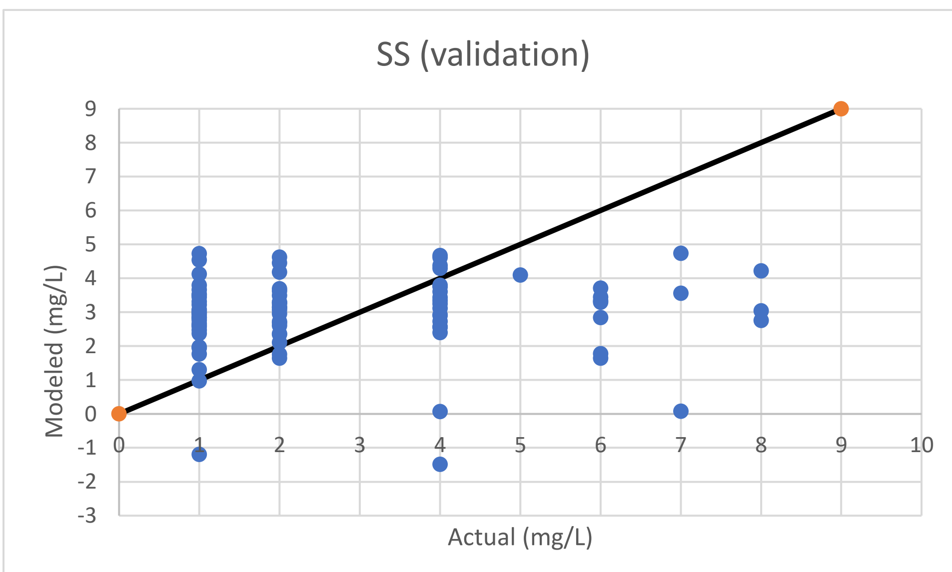


Figure 4.11 Regression analysis of actual versus ANN-modeled SS concentration in the effluent stream

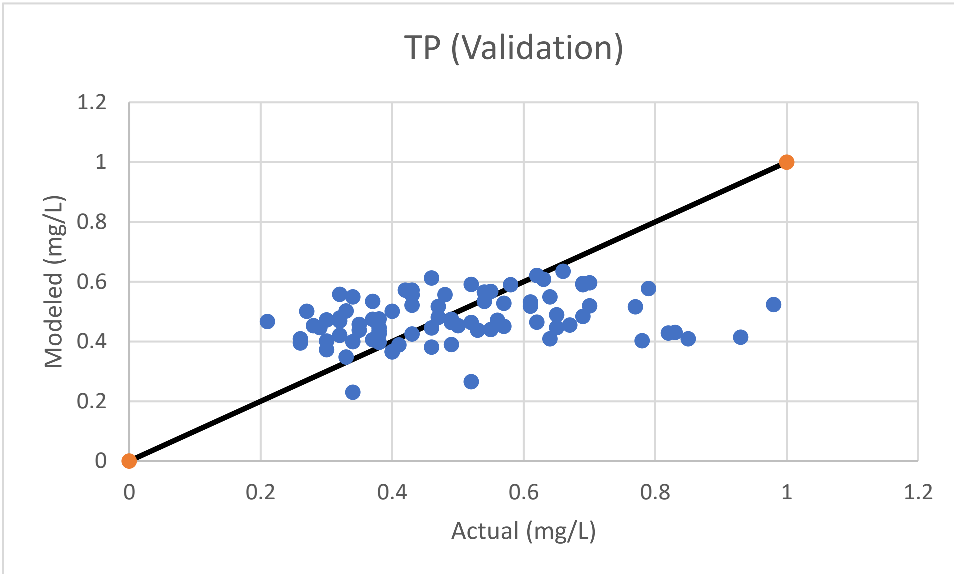


Figure 4.12 Regression analysis of actual versus ANN-modeled TP concentration in the effluent stream

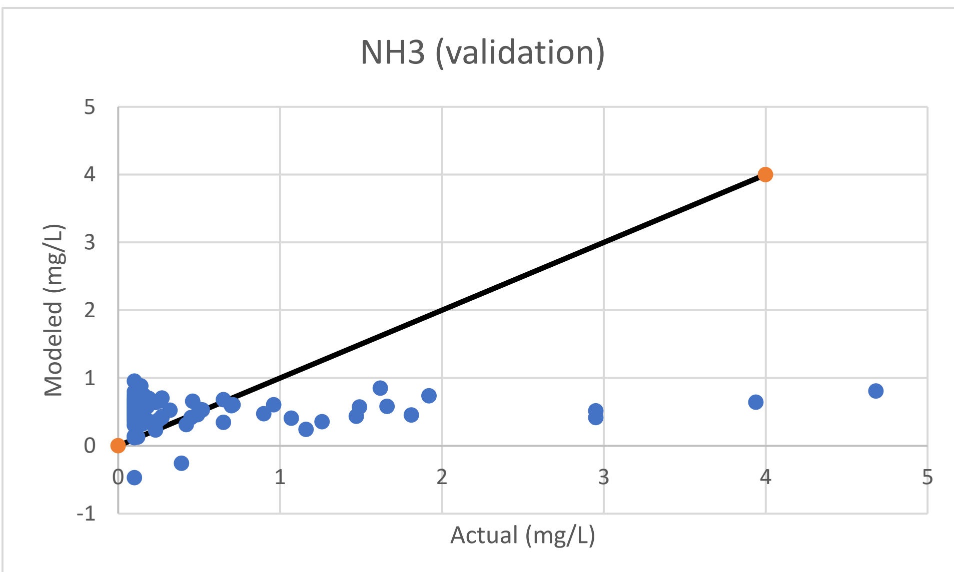


Figure 4.13 Regression analysis of actual versus ANN-modeled NH3 concentration in the effluent stream

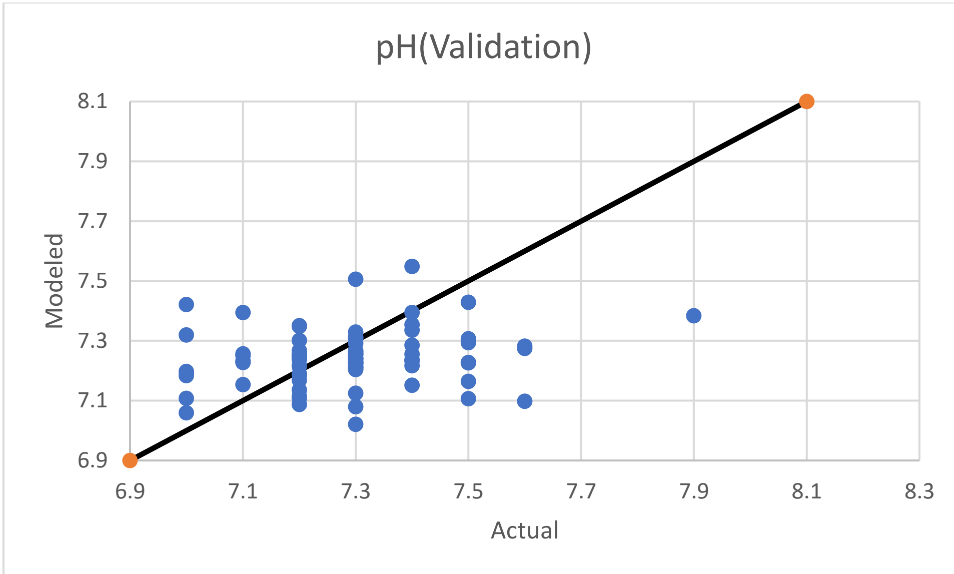


Figure 4.14 Regression analysis of actual versus ANN-modeled pH values in effluent stream

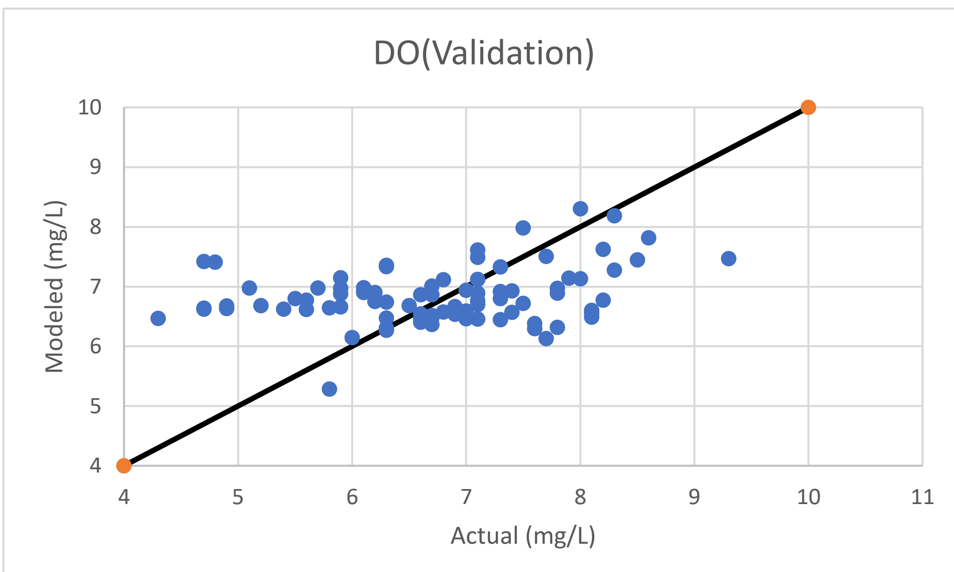


Figure 4.15 Regression analysis of actual versus ANN-modeled DO concentration in effluent stream

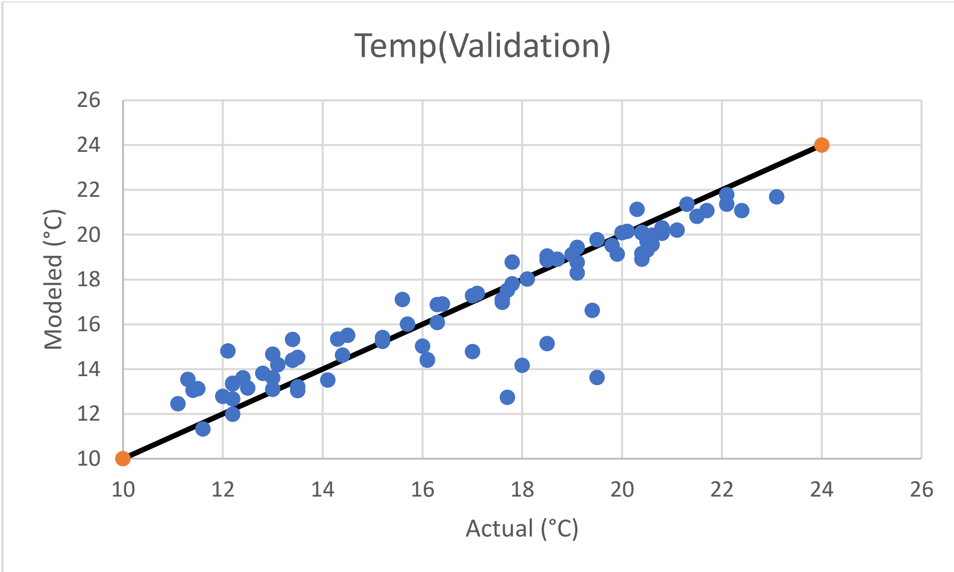


Figure 4.16 Regression analysis of actual versus ANN-modeled temperature values in effluent stream

4.3.2 Discussion of results and limitations

The graphs for model predictions are shown in figure 4.3-4.9 for the training set. It is observed that the correlation coefficient for the training set of variable BOD is 0.36. A negative slope on the parity plot can be observed in figure 4.3. The SS concentration in the effluent stream cannot be predicted by the model beyond 5 mg/L although actual values vary till 20 mg/L as seen in figure 4.4. The modeled TP concentration in the effluent stream varies between 0.2-0.6 mg/L whereas the actual TP concentration varies between 0-1 mg/L. The range of output values is narrow as compared to the input values. The model is not able to capture the variation of NH_3 concentration in the effluent stream and is underestimating its values as seen in figure 4.6.

From table 4.1 it is observed that a maximum correlation exists between the predicted temperature of the effluent stream for both the training and validation set. The least correlation coefficient for the training set exists for TP concentration in the effluent stream as observed in figure 4.5. The least correlation coefficient for the validation set exists for SS and NH_3 concentration in the effluent stream as observed in Figures 4.11 and 4.13 respectively. Since the correlation coefficient for the training set varies between 0.2 to 0.9 therefore the variables with low correlation can further impact the optimization of the system. The model can be further improved by incorporating more information about the operating conditions(e.g., chemical dosage). The incorporation of more data can improve the predictive capability of the network.

Table 4-1 Correlation coefficient of training and validation set

	Training set	Validation set
Temp UV	0.92	0.91
BOD UV	0.36	0.30
SS UV	0.38	0.06
pH UV	0.24	0.20
TP UV	0.21	0.26
NH_3 UV	0.38	0.16
DO UV	0.48	0.29

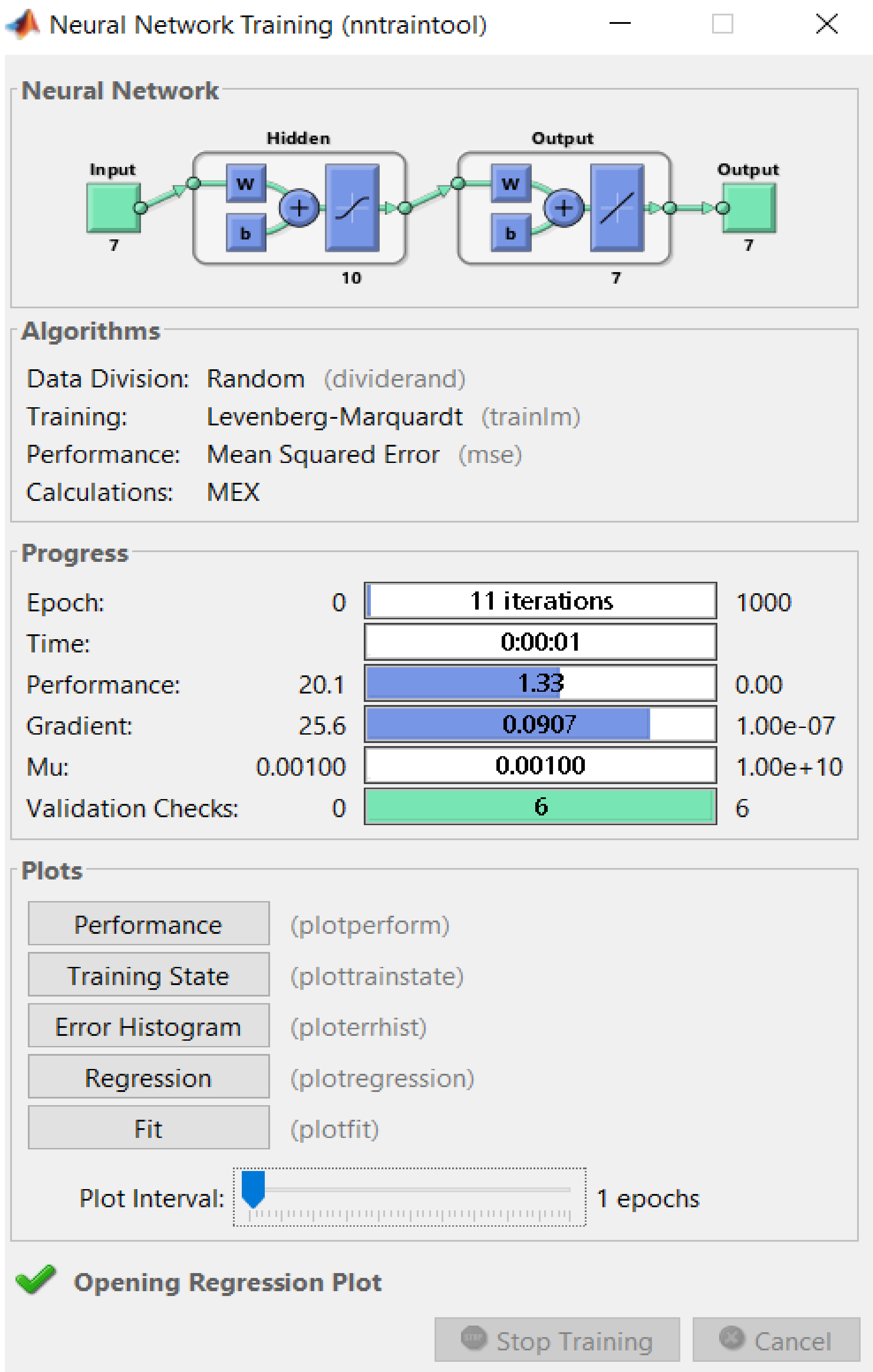


Figure 4.17 Architecture of ANN

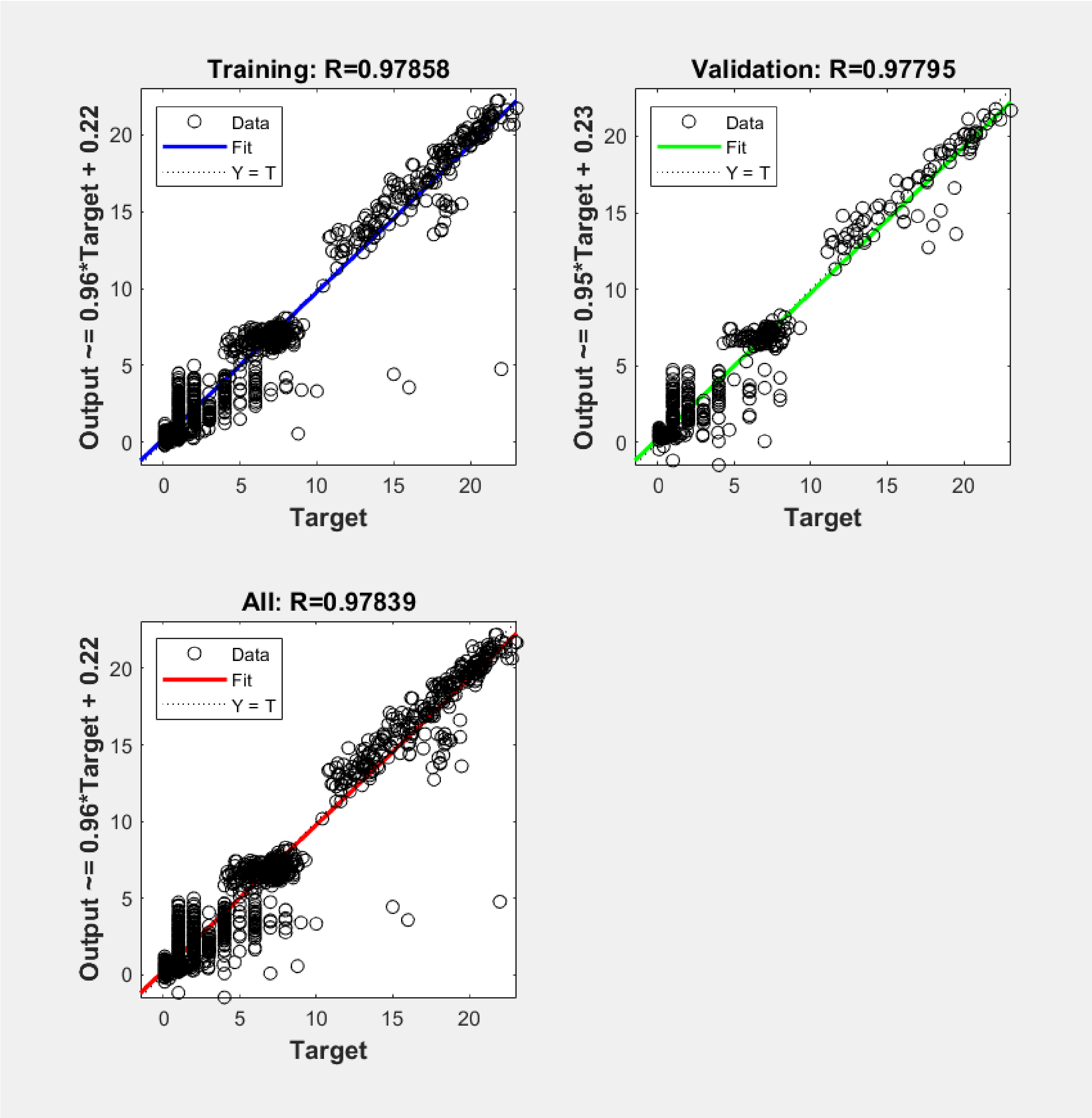


Figure 4.18 Regression analysis

4.4 Sensitivity analysis

Sensitivity analysis is a tool to analyze and assess how multiple independent variables affect dependent variables. Sensitivity analysis is also known as 'what-if analysis' or simulation analysis.

The biggest advantage of sensitivity analysis is to understand which variables have a lot of say in the result. It identifies the most significant variable and prepares for a less favorable scenario.

There are two methods for performing sensitivity analysis: local sensitivity analysis and global sensitivity analysis. The term local refers to the fact that the derivatives are calculated at a particular location. This method works well for simple cost functions but is not practical for complex models. Global sensitivity analysis explores the design space using Monte Carlo methods.

Sensitivity analysis is done to gain insight into the interaction between plant input and output. In this case, one variable is changed by 10% while keeping other variables at their baseline values. From table 4.1 it is evident that suspended solids and ammonia in the effluent are most sensitive to temperature and pH.

Table 4-2:Sensitivity analysis

			BOD_{eff} mg/L	SS_{eff} mg/L	TP_{eff} mg/L	NH_{3eff} mg/L
	unit	Reference values	**1.9704**	**1.7685**	**0.4656**	**0.1316**
Temperature	°C	**18**	-3%	-55%	11%	119%
Flow rate	ML/D	**23.8**	-90%	9%	1%	41%
BOD_{inf}	mg/L	**440**	6%	1%	2%	62%
SS_{inf}	mg/L	**652**	-5%	-21%	-6%	-116%
pH_{inf}	-	**7.4**	35%	-254%	-38%	-431%
TP_{inf}	mg/L	**13.5**	2%	-17%	-1%	31%

NH_{3inf}	mg/L	**20.9**	-4%	23%	2%	14%

4.5 Summary and conclusions

In this chapter model based on ANN was developed to predict the quality of effluent stream. The purpose of ANN model was to identify the pattern between various parameters in the influent and effluent stream. In this research work the seven variables that define the influent stream are temperature, flow rate, BOD, SS, pH, TP, and NH_3 and seven variables that define the effluent stream are temperature, BOD, SS, pH, TP, NH_3 and DO. The effluent quality is measured in terms of four major pollutants namely BOD, SS, TP, and NH_3. The complete data set is divided in 7:3 ratio for training and validation. The results indicate that ANN can predict the quality of effluent stream as the correlation coefficient between the actual values and the predicted values is close to 0.98. In conclusion ANN is an effective tool for predicting the performance of non-linear and complex WWTP.

Chapter 5

5 Multi-objective optimization in WWTP

One approach to reducing nutrient loading is the use of enhanced treatment systems. But this approach can be expensive so to reduce the expenses optimisation approach is suggested in this research work. Optimization in WWTP is a challenging issue because of the complexity of the process. In the past several studies have examined energy optimization in WWTP to increase the treatment efficiency and comply with the discharge limits.

5.1 Introduction

In single objective function optimization, one attempts to find the best solution, which is usually the global minimum (or maximum). However, most real-world problems involve the simultaneous optimization of multiple objective functions (a vector). Such problems are conceptually different from single objective function problems. In multiple objective function optimization, there may not exist a solution that is the best (global optimum) with respect to all objectives. Instead, there could exist a complete set of optimal solutions that are equally good. These solutions are known as Pareto-optimal (or non-dominated) solutions. A Pareto set, for example, for a two-objective function problem is described by a set of points such that when one moves from one point to any other, one objective function improves, while the other worsens. Thus, one cannot say that any one of these points is superior (or dominant) to any other. Since none of the non-dominated solutions in the Pareto set is superior to any other, any one of them is an acceptable solution. The choice of one solution over the other requires additional knowledge of the problem, and often, this knowledge is intuitive and non-quantifiable. The Pareto set, however, is extremely useful since it narrows down the choices and helps to guide a decision-maker in selecting a desired operating point (called the preferred solution) from among the (restricted) set of Pareto-optimal points, rather than from a much larger number of possibilities. One real-life example of multi-objective optimisation is minimizing fuel consumption and maximizing the performance of a vehicle.

The Pareto front is a set of non-dominated solutions that are equally optimal. If there are two objectives f_1(to be maximized) and f_2(to be minimized) as shown in figure 5.1. Solution 1 dominates solution 2 because for point 1 the value of the first objective is higher as compared to the value of point 2 and for objective 2 also point 1 has a better value as compared to point 2. Similarly, from the figure, it can be observed that point 3 is better than point 1. Point 3 and point 5 are non-dominated because for objective 1 point 5 has a better value as compared to point 3 but for objective 2 points 3 has a better value as compared to point 5, hence point 3 and point 5 can not be compared. Points 3, 5 and 6 are non-dominated points and visualised as the non-domination front. Furthermore, choosing one solution from a set of optimal solutions requires higher-level information (Deb, 2011).

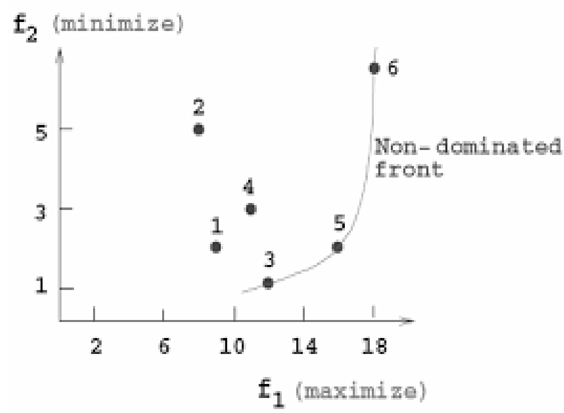


Figure 5.1: Pareto Front

In earlier years, multi-objective optimization problems were usually solved using a single scalar objective function, which was a weighted average of the several objectives (‘scalarization’ of the vector objective function). This process allows a simpler algorithm to be used, but unfortunately, the solution obtained depends largely on the values assigned to the weighting factors used, which is done quite arbitrarily. An even more important disadvantage of the scalarization of the several objectives is that the algorithm may miss some optimal solutions, which can never be found, regardless of the weighting factors chosen. Several methods are available to solve multi-objective optimization problems, e.g., the ε-constraint method, goal attainment method and the non-dominated sorting genetic algorithm (NSGA). In this study, we use NSGA to obtain the Pareto set. This technique offers several advantages, as for example:

1. its efficiency is relatively insensitive to the shape of the Pareto optimal front.

2. problems with uncertainties, stochasticities, and with discrete search spaces can be handled efficiently.

3. the ‘spread’ of the Pareto set obtained is excellent (in contrast, the efficiency of other optimization methods decides the spread of the solutions obtained).

4. it involves a single application to obtain the entire Pareto set (in contrast to other methods, e.g., the ε-constraint method, which needs to be applied several times over).

5.2 Materials and Methods

5.2.1 Genetic Algorithm

A genetic algorithm is an optimization technique used to solve the nonlinear and nondifferentiable optimization problems. The genetic algorithm is inspired by Charles Darwin’s theory of natural evolution to search for a global optimum value.

Firstly, GA is initialized, and a random population is generated, this randomly generated population is also known as the solutions. Each solution is represented as a string of 0s and 1s. For every solution, the value of the objective function is evaluated. This value is called the fitness value and represents the quality of the solution. Individuals are selected based on the quality of the solution.

Crossover results in combining good components of a string to yield an even better string. This new population is called the offspring population and the previous population is called the parent population. The size of the parent population and the offspring population is the same. Therefore, in every iteration, the size of the population remains the same. The purpose of mutation and crossover is to create a better set of the population as compared to the previous population and move towards the optimal solution. After the creation of a new population, the objective function is again evaluated. If the termination criterion specified by the user has been met, then the algorithm can be terminated. But if the termination criterion is not satisfied then step 2 is repeated.

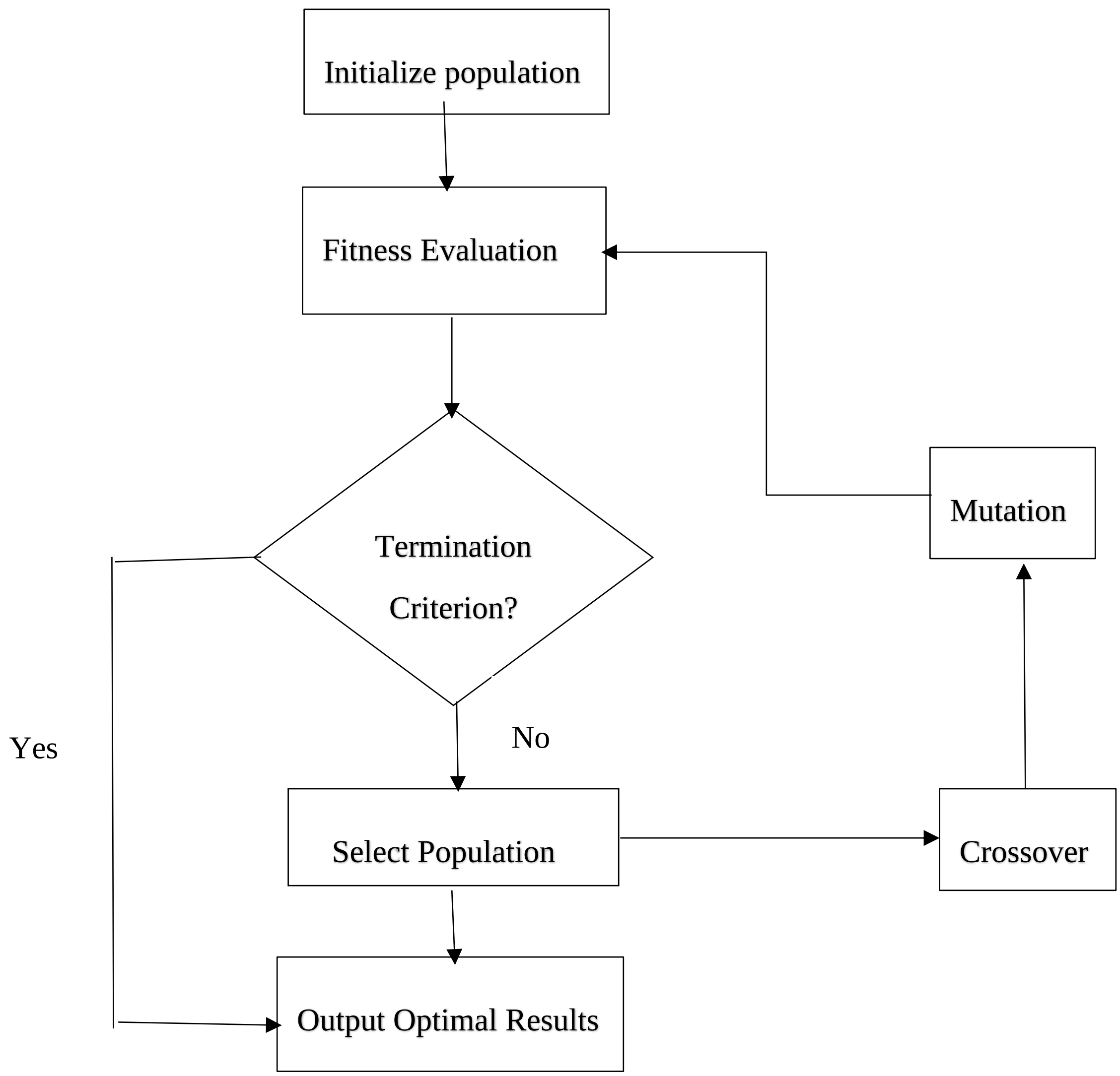


Figure 5.2: Structure of Genetic Algorithm

5.2.2 Effluent Regulations

Wastewater contains pollutants, chemicals and pathogens which are harmful to human health and environment. In Canada to reduce the harmful effects on humans and the environment effluent regulations are imposed under the Fisheries Act. Wastewater management is a collaborative responsibility of federal, provincial, and municipal governments. The federal government specifies limits on wastewater effluents and the provincial government issues permits to operate under those regulations. The municipal government is responsible for the management and operation of wastewater systems.

Table 5-1:Wastewater effluent quality regulations

Effluent parameter	**Regulation**
BOD_5	10 mg/L
SS	10 mg/L
NH_3	3 mg/L
TP	0.5 mg/L

5.2.3 Problem formulation

A genetic algorithm is used to solve the multi-objective optimization problems. An equalization tank or buffer system is suggested to counterbalance the fluctuating flow and composition of influent to the treatment plant. The purpose of equalization is to pretreat the wastewater and equalization tanks are located after the primary clarification and before the aeration tank.

5.2.3.1 Decision Variables

For optimal operation of a wastewater treatment plant number of conflicting objectives needs to be dealt with. The purpose of this study is to minimize the concentration of four major pollutants in the effluent stream and satisfy the restrictions on the effluent stream. For proper operation of the wastewater treatment, four cases of multi-objective optimization are formulated. The system includes an equalization tank therefore the decision variables associated with this process are the temperature of the influent stream, total sewage flow, BOD, SS, pH, TP, and NH_3 of the influent stream.

5.3 Multi-Objective Optimization using Genetic Algorithm

5.3.1 Optimization Problem

For case 1 two objectives are minimized, BOD and TP concentration in the effluent stream (BOD_{eff} and TP_{eff}). It is important to minimize the concentration of the pollutants (BOD and TP) in the effluent stream and meet environmental regulations. The constraints to be satisfied in this problem are the limits imposed on the effluent quality by the regulatory bodies. To comply with the regulatory requirements, the concentration of NH_3 in the effluent stream (NH_{3eff}) should be below 3 mg/L and the concentration of SS in the effluent stream (SS_{eff}) should be below 10 mg/L. Since the concentration of the pollutants cannot be a negative number therefore the concentration of all four pollutants should be greater than zero. The decision variables involved in this process are the temperature of the influent stream, total sewage flow, BOD_{inf}, SS_{inf}, pH_{inf}, TP_{inf}, and NH_{3inf}. The upper and lower bounds of the decision variables are chosen based on the industrial values as shown in Table 5.2.

In this case of minimizing BOD_{eff} and TP_{eff} concentration, the solution is violating the fourth constraint. As the value of concentration cannot be negative so to overcome this problem penalty term is added to the objective function. The purpose of the penalty function is to penalize the objective function in case of constraint violation by adding a penalty to it. In the case of a minimization problem, a penalty is added to the objective function. So that the infeasible solution will be penalized. On the other hand, if the maximization problem is being solved penalty is subtracted from the objective function. Various types of penalty functions are the death penalty, static penalty, dynamic penalty and adaptive penalty (Yeniay, 2005).

Case1

Objective 1: Min BOD_{eff}

Objective 2: Min TP_{eff} +0.1 (where 0.1 is the penalty function)

Constraint 1 $0 < NH_{3\ eff} < 3$ (mg/L)

Constraint 2 $0 < SS_{eff} < 10$ (mg/L)

Constraint 3 $0 < BOD_{eff}$

Constraint 4 $0 < TP_{eff}$

To find the optimal values of (7 Parameters)

1. Temperature of the influent stream (°C) (x1)
2. Total sewage flow (ML/D) (x2)
3. BOD_{inf} (mg/L) (x3)
4. SS_{inf} (mg/L) (x4)
5. pH_{inf} (x5)
6. TP_{inf} (mg/L) (x6)
7. NH_{3inf} (mg/L) (x7)

Where Y=f(X) Y = [BOD_{eff} TPeff]

Table 5-2: Bounds on decision variables(X) for case 1

Decision Variable	**Lower Bound**	**Upper Bound**
Temperature(°C)	11	22
Total sewage flow (ML/D)	16.3	59
BOD_{inf} (mg/L)	76	619
SS_{inf} (mg/L)	53	950
pH_{inf}	7.1	8.2
TP_{inf} (mg/L)	2.9	25
NH_{3inf} (mg/L)	9.0	41.8

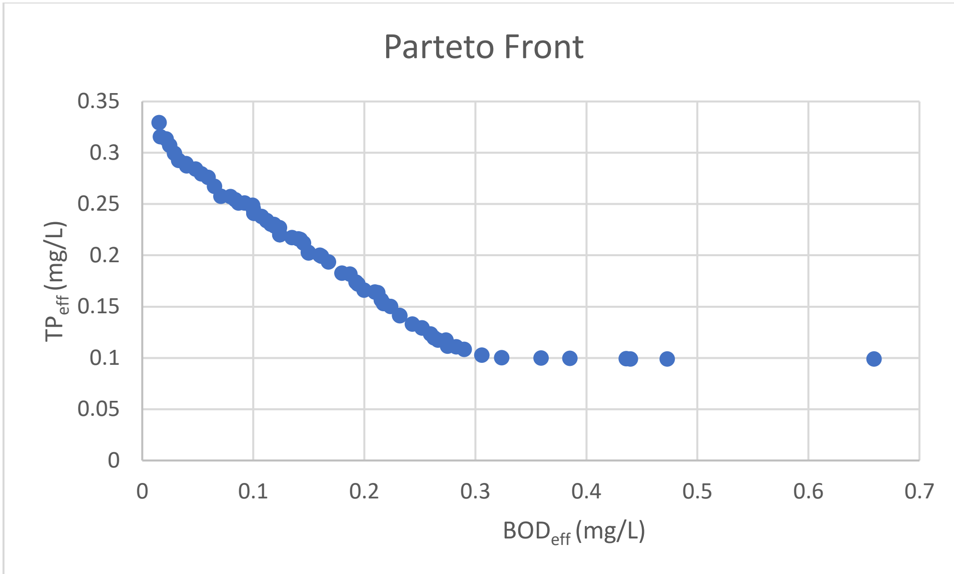
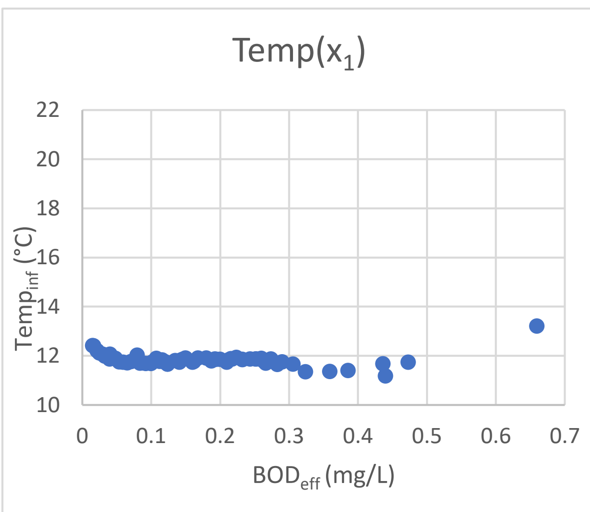


Figure 5.3: Pareto front for case 1



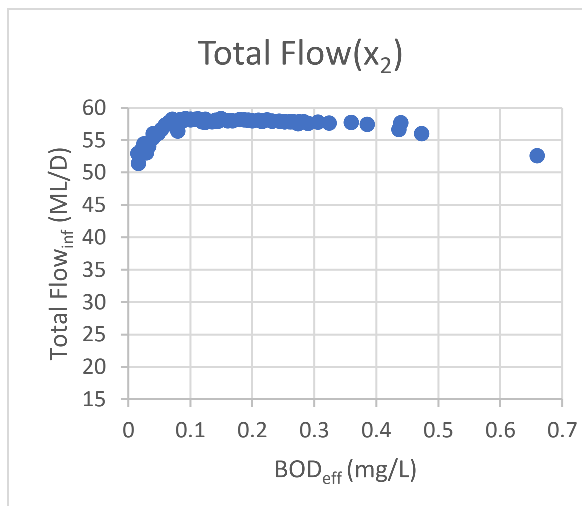


Figure 5.4 Optimal variation of temp and total flow rate with BOD_{eff} for case 1

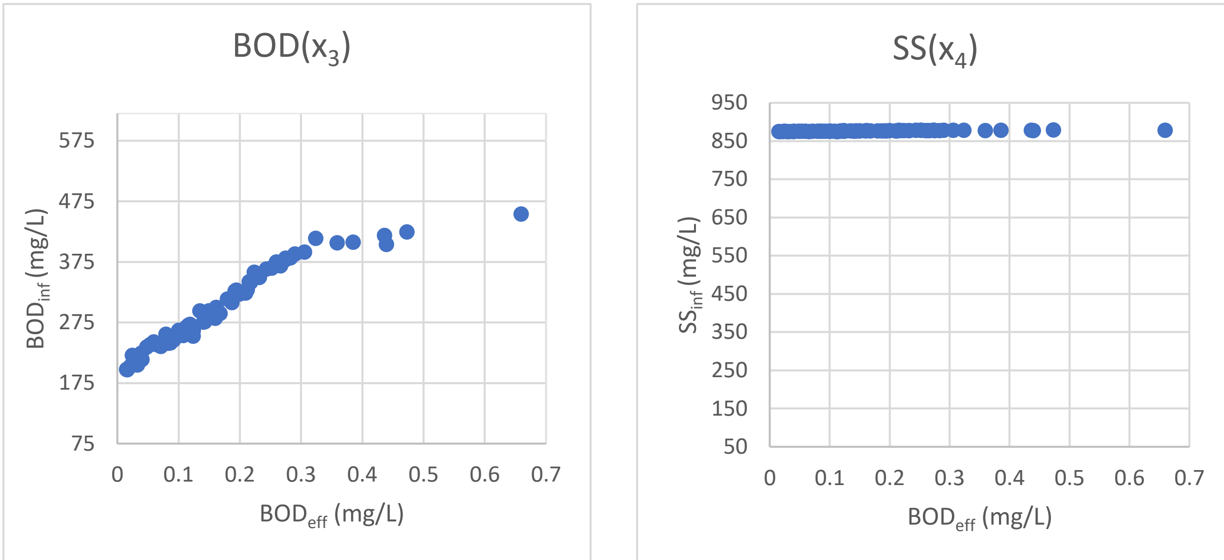


Figure 5.5 Optimal variation of BOD_{inf} and SS_{inf} with BOD_{eff} for case 1

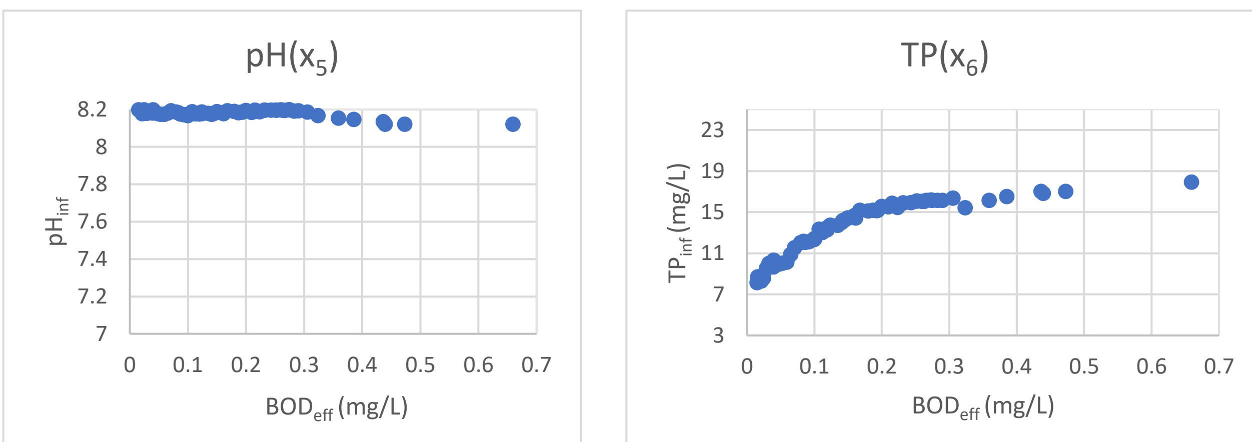


Figure 5.6 Optimal variation of pH_{inf} and TP_{inf} with BOD_{eff} for case 1

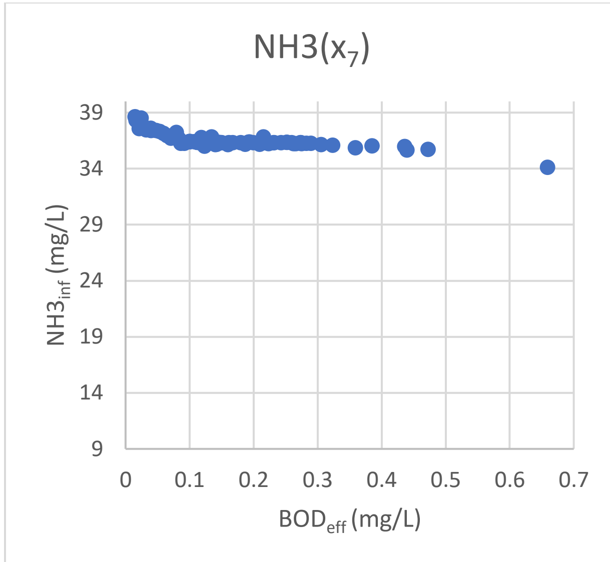


Figure 5.7 Optimal variation of $NH3_{inf}$ with BOD_{eff} for case 1

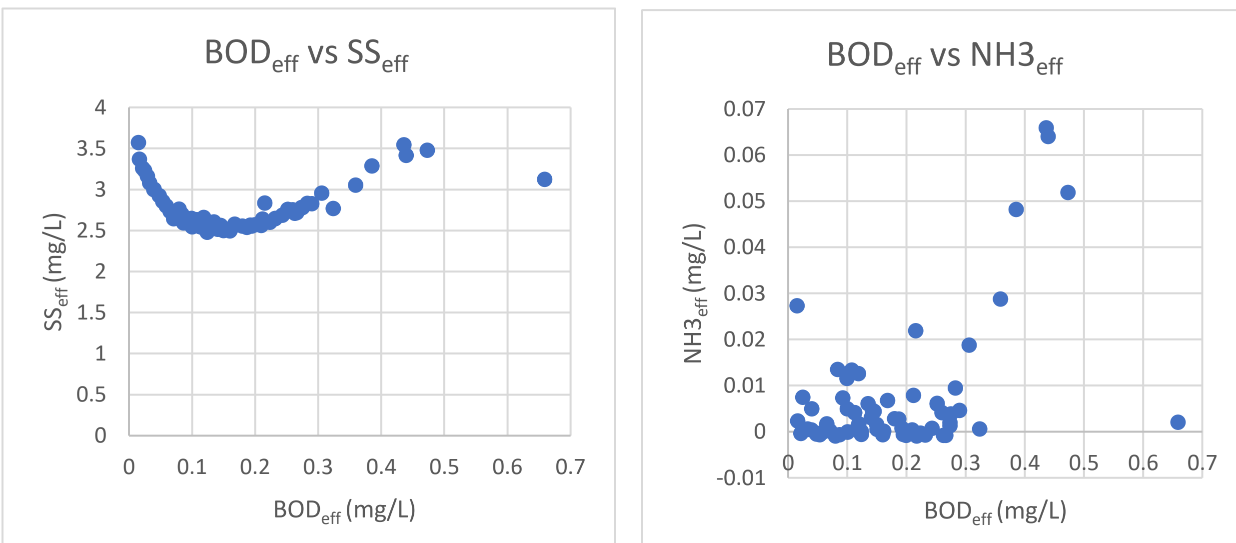


Figure 5.8 Variation of SS_{eff} and $NH3_{eff}$ with BOD_{eff} for case 1

5.4 Results and discussions

A genetic algorithm(GA) is used to solve the multi-objective optimization problem. The two objectives that are simultaneously minimized in case 1 are BOD and TP concentration in the effluent stream. The optimal values of the decision variables are hence evaluated. For evaluation of the optimum values of the decision variables, every data was collected and analysed for a period of four years. The upper bound and lower bounds of the decision variable are decided based on the system capacity and historical data. The computation time taken for optimizing seven decision variables was less than 10 minutes on 8GB RAM and a 1.8 GHz processor. The population size specifies the number of individuals in each generation and the number of generations specifies the maximum number of iterations the genetic algorithm performs. The maximum number of generations specified for case 1 is 1400(200 x 7). It was observed that after the 122th generation there was no improvement in the Pareto optimal front. The Pareto optimal front of BOD and TP concentration in the effluent stream is shown in figure 5.3. The function tolerance used in case 1 is 1×10^{-4} and the significance of the tolerance function is that if the weighted average change in the spread of Pareto solutions over the previous generations is less than function tolerance and the spread is smaller than the average spread over the last generations then the algorithm stops. The constraint tolerance used for case 1 is 1×10^{-3} and is the tolerance for linear constraint violations.

Table 5-3: GA parameters for Case 1

Population size	200 x 7
Maximum number of generations	200 x 7
Crossover fraction	0.8
Pareto fraction	0.35
Constraint tolerance	1×10^{-3}
Function tolerance	1×10^{-4}

The optimal variation of the decision variables corresponding to BOD_{eff} is shown in figures 5.4-5.7. The optimal values of temperature of the influent stream lie towards the lower bound. This suggests that to minimize the concentration of BOD_{eff} and TP_{eff}, the temperature should be kept around 12°C. The total flow rate of influent wastewater lies towards the upper bound value of 60 ML/D. It is observed from figure 5.4 that the optimal values of total flow of influent wastewater increase from 50 to 60 ML/D as BOD_{eff} increases from 0 to 0.1 mg/L. From figure 5.5 it is observed that BOD_{eff} increases with an increase in BOD_{inf}. Therefore, the optimal range of BOD_{inf} is from 175 to 475 mg/L. The optimal value of SS_{inf} is around 850 mg/L and not much variation in the optimal values can be observed in figure 5.5. Similarly, the optimal values of pH_{inf} remain constant around the upper bound value of 8.2 as observed in figure 5.6. As the optimal value of TP_{inf} increases from 7 to 19 mg/L as the concentration of BOD_{eff} also increases from 0 to 0.7 mg/L as shown in figure 5.6. The optimal range of NH_{3inf} lies towards the upper bound value of 39 mg/L and ranges between 34-39 mg/L as observed in figure 5.7. Since in case 1 the objective is to minimize the concentration of BOD and TP in the effluent stream it is also important to evaluate the concentration of SS and NH_3 in the effluent stream. From figure 5.8 it is observed that the value of SS_{eff} and NH_{3eff} is less than 4 mg/L and 0.1 mg/L respectively, thereby meeting the regulatory requirements.

Case 2 is a multi-objective optimization problem which involves seven decision variables and two objectives. The two objectives that are simultaneously being minimized are BOD and SS concentration in the effluent stream. This problem is solved using a Genetic Algorithm.

Case2

Objective 1: Min BOD_{eff}

Objective 2: Min SS_{eff}

Constraint 1	$0 < NH_{3\,eff} < 3$ (mg/L)
Constraint 2	$0 < TP_{eff} < 0.5$ (mg/L)
Constraint 3	$0 < BOD_{eff}$
Constraint 4	$0 < SS_{eff}$

Table 5-4: Bounds on decision variables(X) for case 2

Decision Variable	Lower Bound	Upper Bound
Temperature(°C)	11	22
Total sewage flow (ML/D)	16.3	59
BOD_{inf} (mg/L)	76	619
SS_{inf} (mg/L)	53	950
pH_{inf}	7.1	8.2
TP_{inf} (mg/L)	2.9	25
NH_{3inf} (mg/L)	9.	41.8

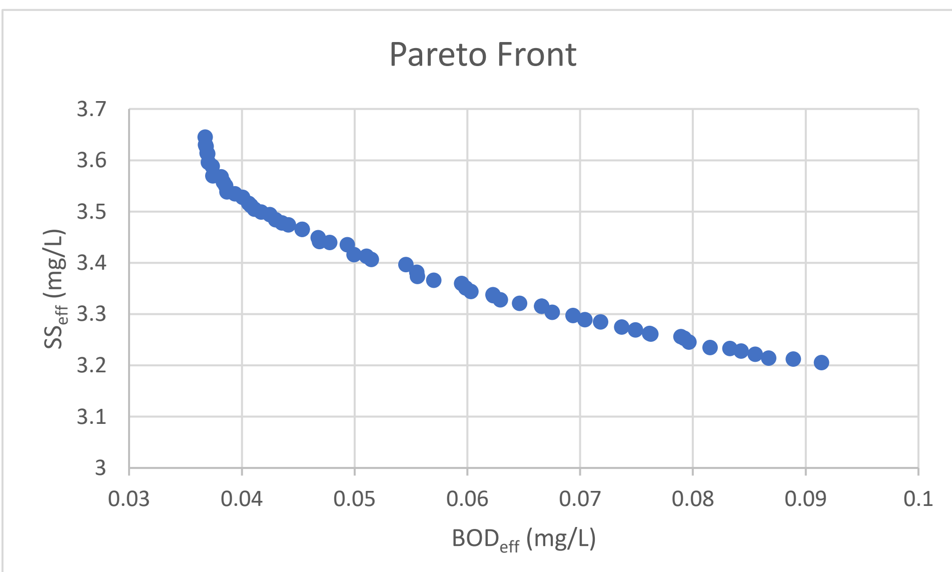


Figure 5.9 Pareto front of BOD_{eff} and SS_{eff} for case 2

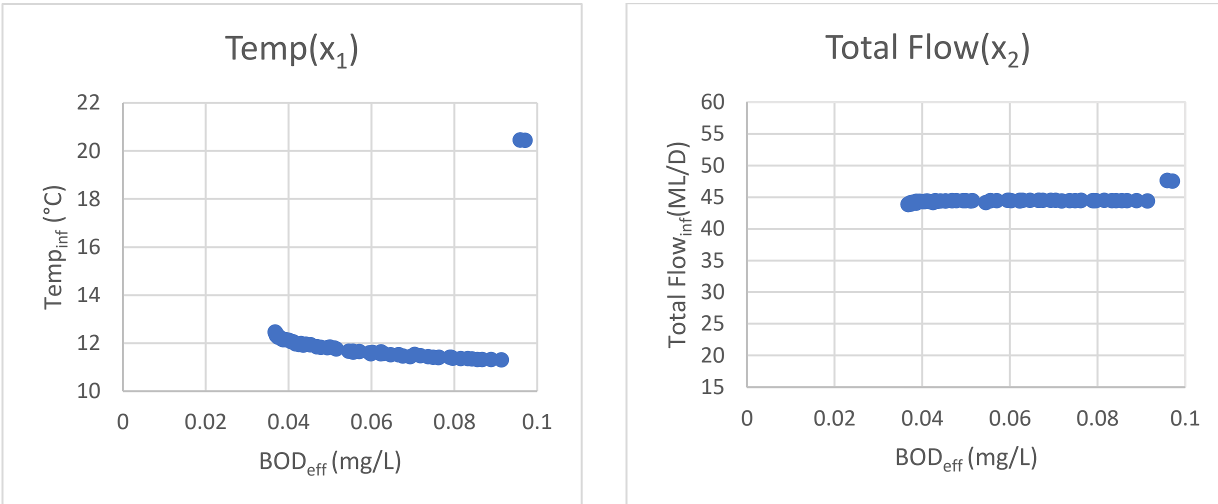


Figure 5.10 Optimal variation of temp and total flow rate with BOD_{eff} for case 2

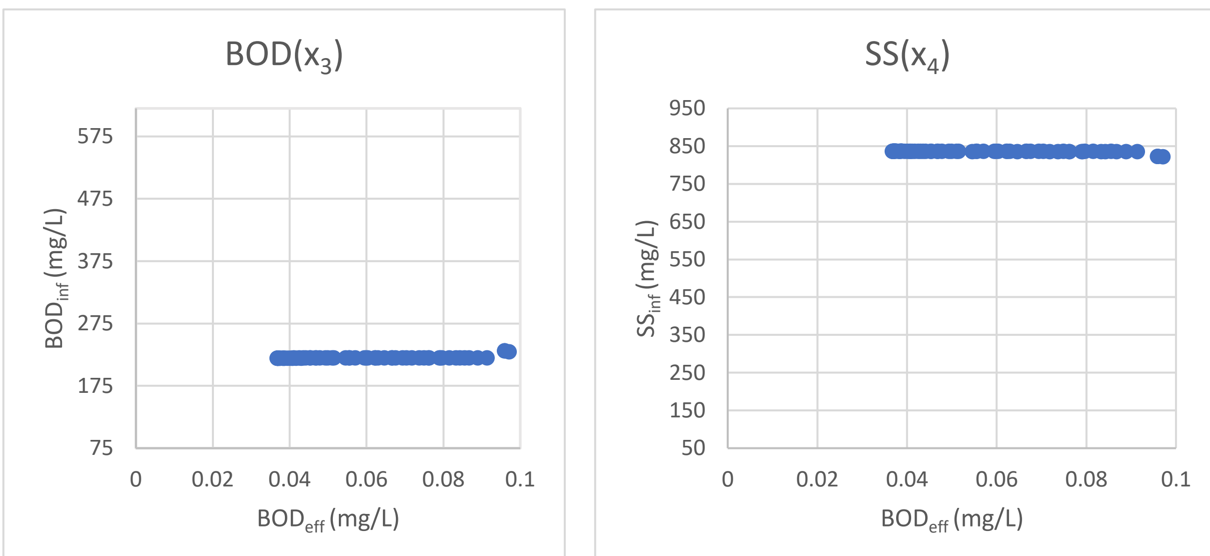


Figure 5.11 Optimal variation of BOD_{inf} and SS_{inf} with BOD_{eff} for case 2

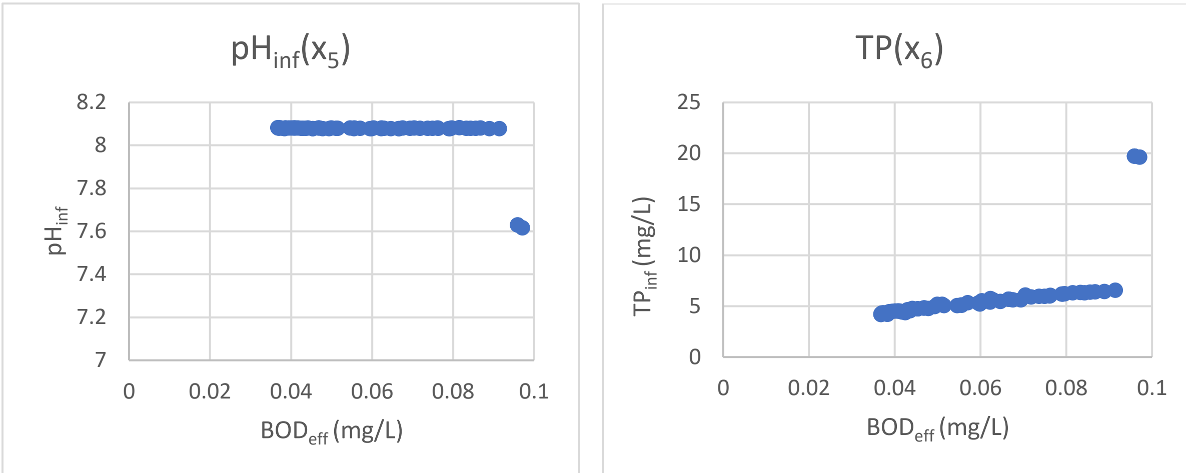


Figure 5.12 Optimal variation of pH_{inf} and TP_{inf} with BOD_{eff} for case 2

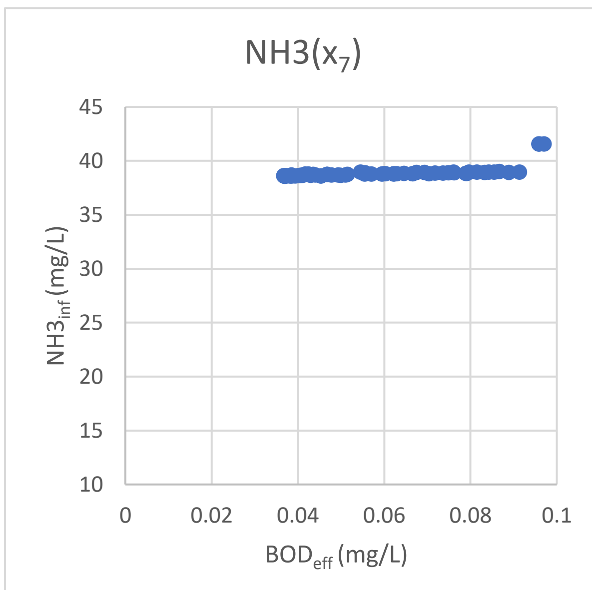


Figure 5.13 Optimal variation of $NH3_{inf}$ with BOD_{eff} for case 2

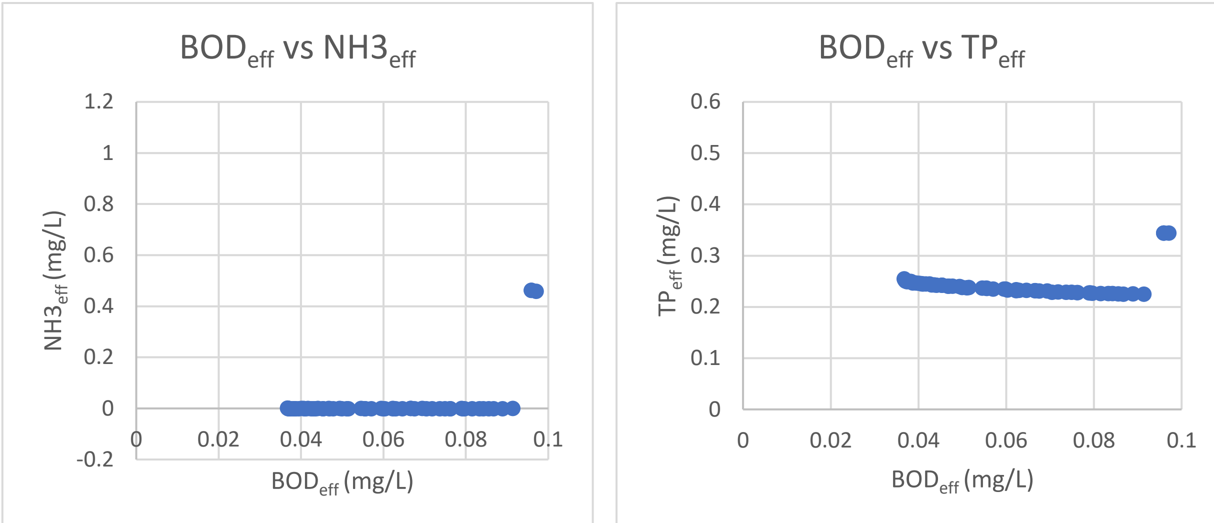


Figure 5.14 Variation of $NH3_{eff}$ and TP_{eff} with BOD_{eff} for case 2

Table 5-5: GA parameters for case 2

Population size	200 x 7
Maximum number of generations	200 x 7
Crossover fraction	0.8
Pareto fraction	0.35
Constraint tolerance	1 x 10^{-3}
Function tolerance	1 x 10^{-4}

The computational time taken for optimizing seven decision variables of case 2 was 12 minutes on an 8GB RAM and 1.8 GHz processor. It was observed that after the 173rd generation there was no improvement in the Pareto optimal front. The Pareto optimal front of BOD_{eff} and SS_{eff} is shown in figure 5.9. The population size and the maximum number of generations for case 2 are the same as in case 1. The crossover fraction used for case 2 is 0.8 and this value specifies the fraction of each population, other than elite solutions, that are made up of crossover solutions. A crossover fraction of 1 means that all the solutions other than the best individuals are crossover solutions. While the crossover fraction of 0 represents that all the solutions are generated after mutation. The Pareto fraction chosen for case 1 and case 2 is 0.35. This specifies that fraction of the population on the best Pareto frontier is to be kept on the optimal Pareto front. In this case, population size of each generation is 200, therefore the number of solutions kept on the optimal Pareto front is 200 x 0.35=70.

The optimal variation of decision variables corresponding to BOD_{eff} for case 2 is shown in figures 5.10-5.13. The optimal values of temperature of the influent stream lie towards the lower bound value of 12°C. From figure 5.10 it is observed that the optimal value of temperature decreases with an increase in BOD_{eff} concentration. There is not much variation in the optimal values of total $flow_{inf}$ with change in BOD_{eff}. The optimal value of total $flow_{inf}$ is close to 45 ML/D as shown in figure 5.10. The optimal value of BOD_{inf} is around 225 mg/L and there is no variation in the optimal values of BOD_{inf} as observed in figure 5.11. The optimal values of SS_{inf} lie close to its upper bound value of 850 mg/L. Like the BOD_{inf} values, the optimal values of SS_{inf} also remain constant corresponding to BOD_{eff} values. The optimal values of pH remain close to upper bound values of 8.2 as shown in figure 5.12. The optimal values of TP_{inf} increase with an increase in BOD_{eff} as observed in figure 5.12. The optimal value of NH_{3inf} lies towards the upper bound value of 45 mg/L and there is not much variation in NH_{3inf} values as observed in figure 5.13.

In case 3 two objectives that are simultaneously being minimized are TP_{eff} and SS_{eff}. The limits imposed on the concentration of pollutants in the effluent stream is per the regulatory requirements. Therefore, NH_{3eff} should be less than 3 mg/L and the concentration of BOD_{eff} should be less than 10 mg/L. Since the concentration of the pollutants cannot be a negative number therefore the concentration of all four pollutants should be greater than zero. The decision variables involved in this process are $temp_{inf}$, total sewage flow, BOD_{inf}, SS_{inf}, pH_{inf}, TP_{inf}, and NH_{3inf}. The upper and lower bounds of the decision variables are chosen based on the industrial values as shown in Table 5.6.

Case3

Objective 1: Min $TP_{eff} + 0.1$

Objective 2: Min $SS_{eff} + 0.1$ (where 0.1 is penalty function)

Constraint 1 $0 < NH_{3\ eff} < 3$

Constraint 2 $0 < BOD_{eff} < 10$

Constraint 3 $0 < TP_{eff}$

Constraint 4 $0 < SS_{eff}$

Table 5-6: Bounds on decision variables(X) for case 3

Decision Variable	**Lower Bound**	**Upper Bound**
Temperature	11	22
Total sewage flow (ML/D)	16.3	59
BOD_{inf} (mg/L)	76	619
SS_{inf} (mg/L)	53	950
pH_{inf}	7.1	8.2
TP_{inf} (mg/L)	2.9	25
NH_{3inf} (mg/L)	9.	41.8

This multi-objective optimization problem is solved using GA. Initially, the MATLAB code was run without a penalty function, but it was observed that some solutions on the Pareto front were negative. Therefore, to overcome this issue static penalty function of 0.1 was added to both objectives. The computation time taken for optimizing seven decision variables was 12 minutes on 8GB RAM and a 1.8 GHz processor. After the 202nd iteration there was no improvement in the Pareto optimal front. The Pareto optimal front of TP_{eff} and SS_{eff} is shown in figure 5.15. From figure 5.15 it is observed that there is no variation in SS_{eff} concentration with respect to the TP_{eff} concentration.

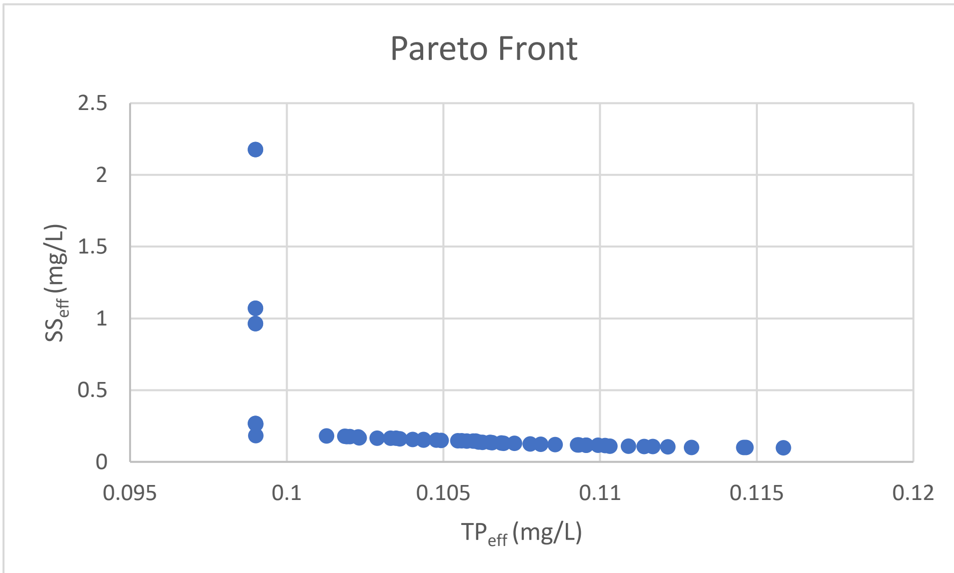


Figure 5.15 Pareto front of case 3

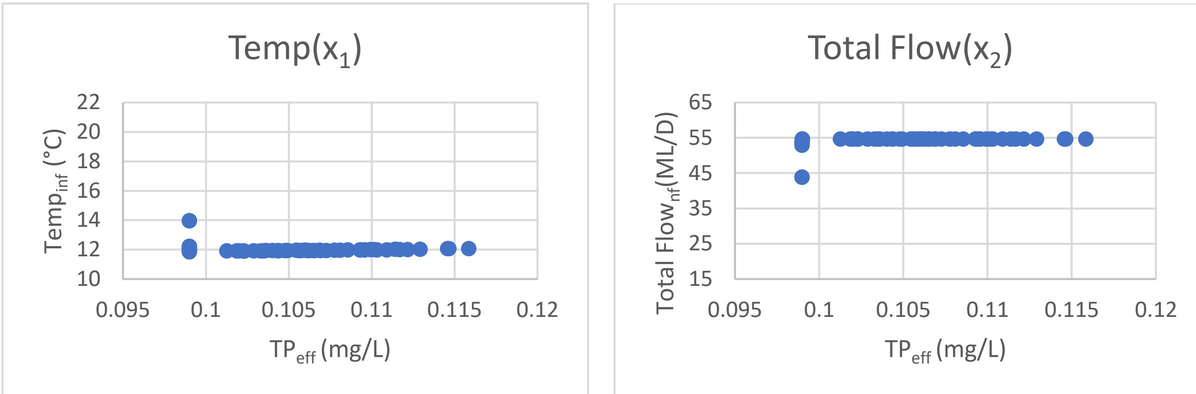


Figure 5.16 Optimal variation of temp and total flow with TP_{eff} for case 3

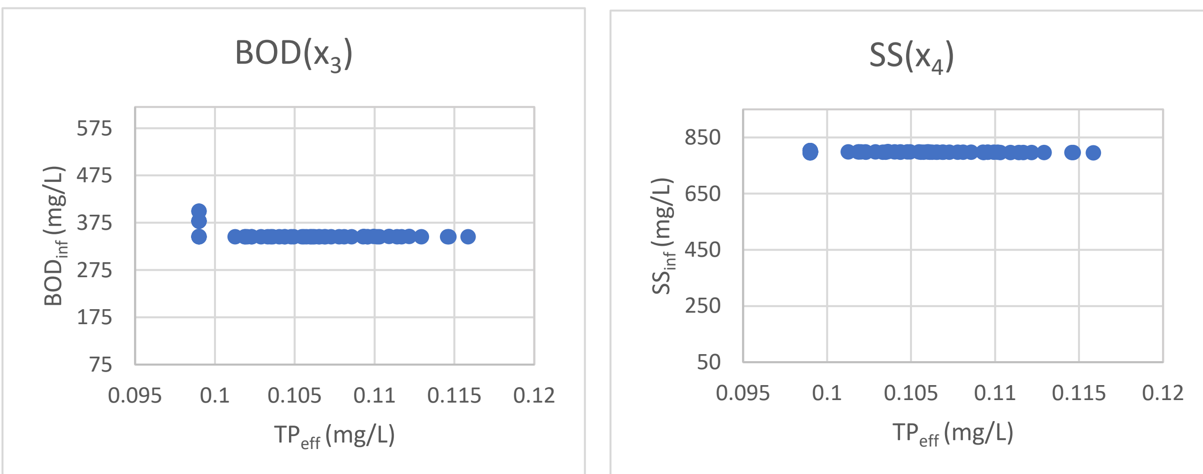


Figure 5.17 Optimal variation of BOD_{inf} and SS_{inf} with TP_{eff} for case 3

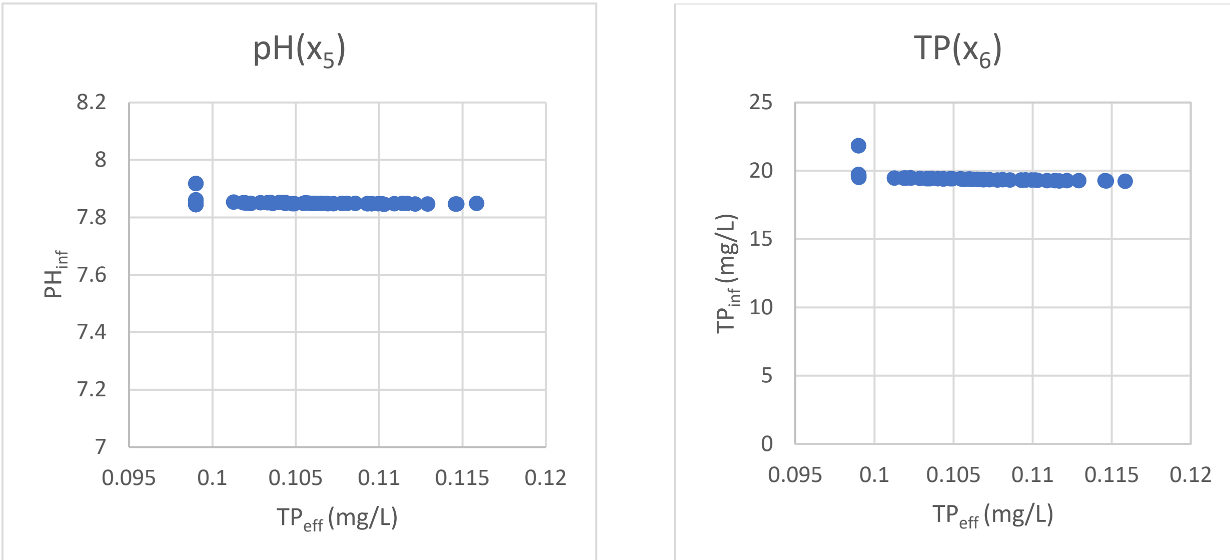


Figure 5.18 Optimal variation of pHi$_{nf}$ and TP$_{inf}$ with TP$_{eff}$ for case 3

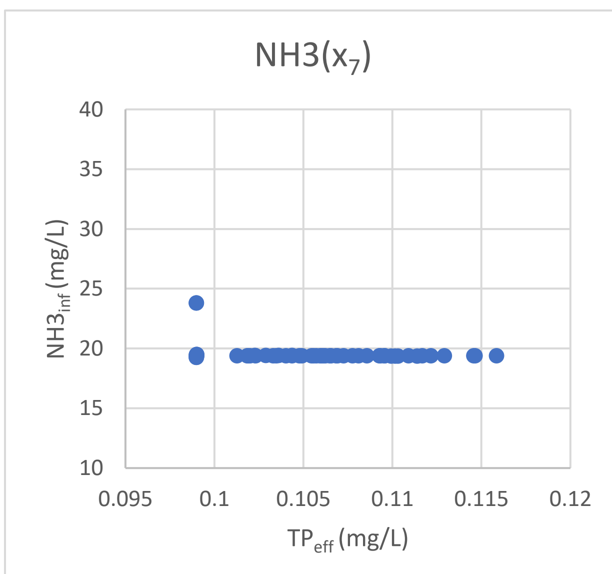
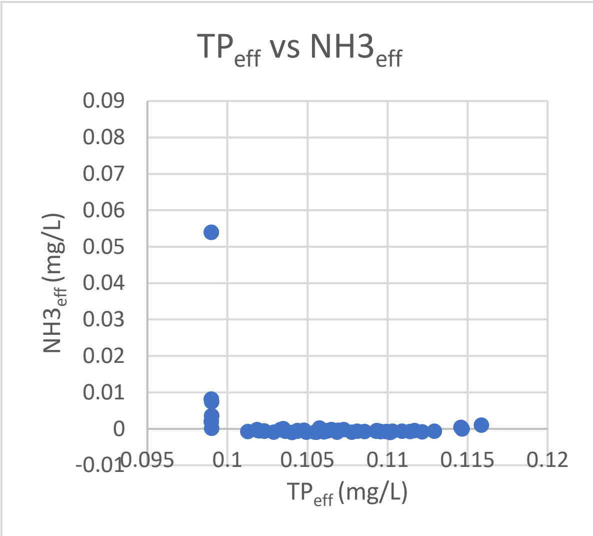


Figure 5.19 Optimal variation of NH3$_{inf}$ with TP$_{eff}$ for case 3



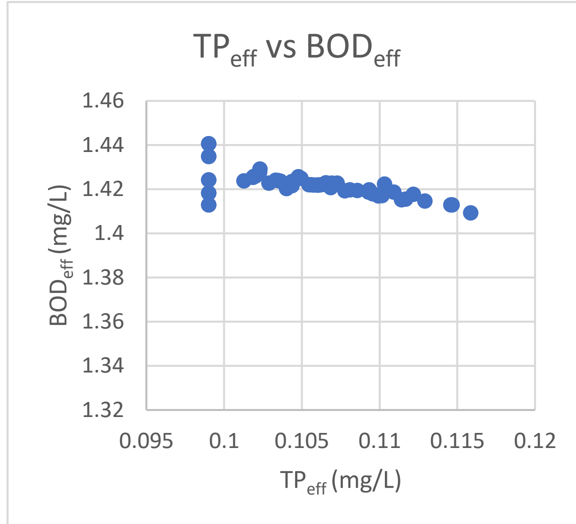


Figure 5.20 Variation of $NH3_{eff}$ and BOD_{eff} with TPeff for case 3

Table 5-7: GA parameters for case 3

Population size	200 x 7
Maximum number of generations	200 x 7
Crossover fraction	0.8
Pareto fraction	0.35
Constraint tolerance	1 x 10^{-3}
Function tolerance	1 x 10^{-4}

The optimal variation of the seven decision variables corresponding to TP_{eff} is shown in Figures 5.16-5.19. As compared to case 1 and case 2 not much variation in the optimized values of the decision variables can be observed. The optimal values of $temp_{inf}$ lie close to lower bound values similar to case 1 and case 2. Therefore, to minimize the concentration of TP_{eff} and SS_{eff} temperature should be kept around 12°C as shown in figure 5.16. The optimized values of the total flow of wastewater lie towards the upper bound value similar to case 1. From figure 5.16 it is observed that the optimal value of the total flow are around 55 ML/D. The reason for constant values of the decision variables corresponding to the optimal Pareto front could be the narrow range of TP_{eff} on the Pareto front. The results are satisfactory because TP_{eff} concentration can't exceed the limit of 0.5 mg/L as per effluent regulations mentioned in table 5.1. The optimal value of BOD_{eff} are around 375 mg/L as shown in figure 5.17. The optimal values of SS_{eff} lie towards the upper value of 850 mg/L similar to case 1 and case 2 as shown in figure 5.17.

The optimal value of pH_{inf} are around 7.8 and this value is lower as compared to case 1 and case 2. From figure 5.18 it is observed that the optimal value of TP_{inf} is around 20 mg/L and is close to the upper bound value of 25 mg/L. The optimal values of TP_{inf} for case 3 are higher as compared to case 2. The optimal values of NH_{3inf} for case 3 remain constant at around 20 mg/L as TP_{eff} concentration varies from 0.1-0.12 mg/L as shown in figure 5.19. As case 3 involves minimization of TP_{eff} and SS_{eff} to understand the variation of the other two pollutants(BOD_{eff} and NH_{3eff}) is also important. Variation of NH_{3eff} and BOD_{eff} corresponding to the optimal Pareto front is shown in figure 5.20. It is observed that there is no violation of the regulatory norms.

In case 4 three objectives are simultaneously minimized, which include: (i) TP_{eff} (ii) BOD_{eff} (iii) SS_{eff}. The constraints to be satisfied in this problem are the limits imposed on the effluent quality by the regulatory bodies. To comply with the regulatory requirements, NH_{3eff} should be less than 3 mg/L. Since the concentration of the pollutants cannot be a negative number therefore the concentration of all four pollutants should be greater than zero. The decision variables involved in this process are the temperature of the influent stream, total sewage flow, BOD_{inf}, SS_{inf}, pH_{inf}, TP_{inf}, and NH_{3inf}. The upper and lower bounds of the decision variables are chosen based on the industrial values as shown in Table 5.8.

Case 4

Objective 1: Min TP_{eff} + 0.1 (where 0.1 is the penalty function)

Objective 2: Min BOD_{eff}

Objective 3: Min SS_{eff} + 0.1 (where 0.1 is the penalty function)

Constraint 1 $0 < NH_{3\,eff} < 3$ (mg/L)

Constraint 2 $0 < BOD_{eff}$

Constraint 3 $0 < TP_{eff}$

Constraint 4 $0 < SS_{eff}$

Table 5-8: Bounds on decision variables(X) for case 4

Decision Variable	Lower Bound	Upper Bound
Temperature	11	22
Total sewage flow (ML/D)	16.3	59
BOD_{inf} (mg/L)	76	619
SS_{inf} (mg/L)	53	950
pH_{inf}	7.1	8.2
TP_{inf} (mg/L)	2.9	25
NH_{3inf} (mg/L)	9.	41.8

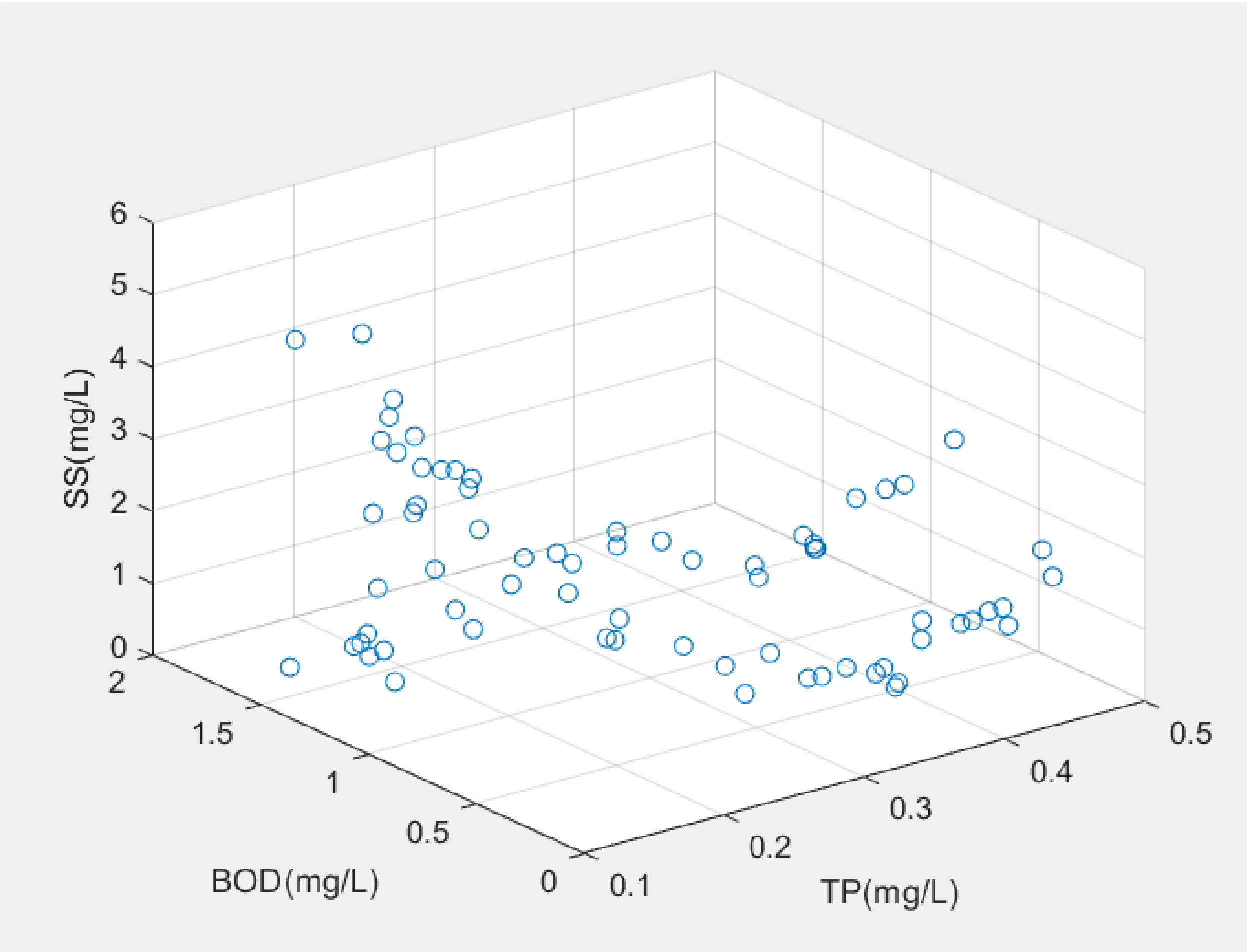


Figure 5.21 3-D Pareto front of case 4

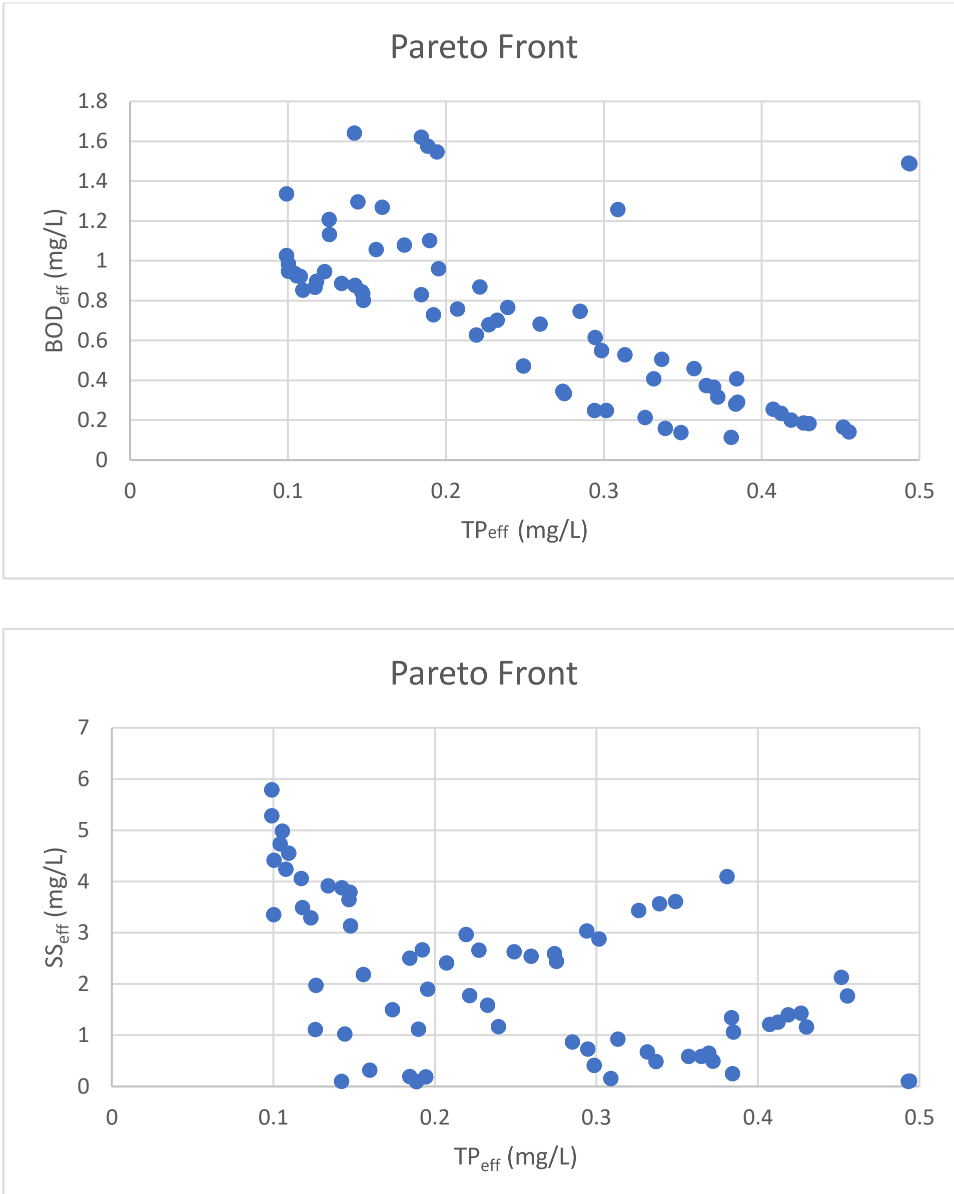


Figure 5.22 2-D plot of Pareto front of case 4

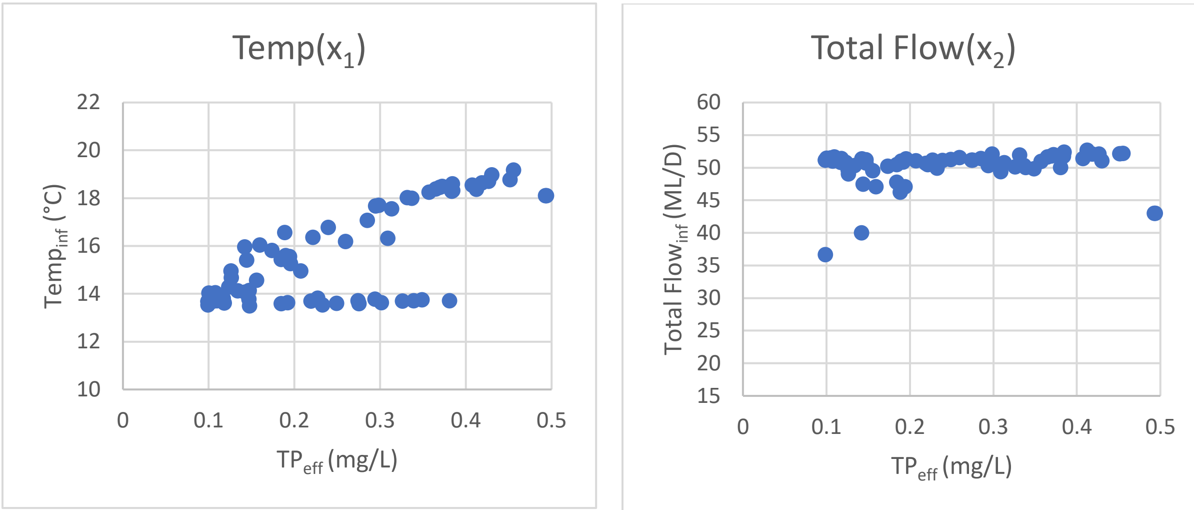


Figure 5.23 Optimal variation of temp and total flow with TPeff for case 4

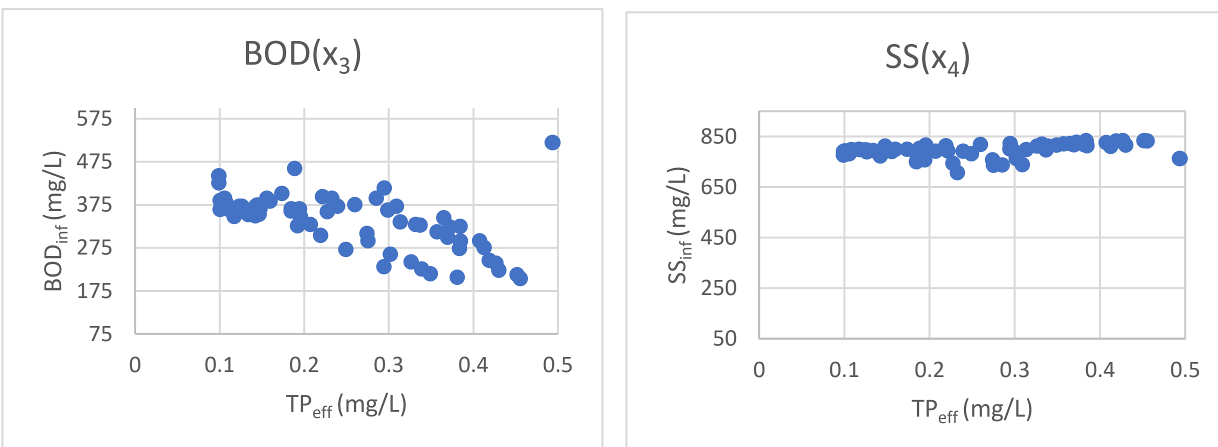


Figure 5.24 Optimal variation of BOD_{inf} and SS_{inf} with TP_{eff} for case 4

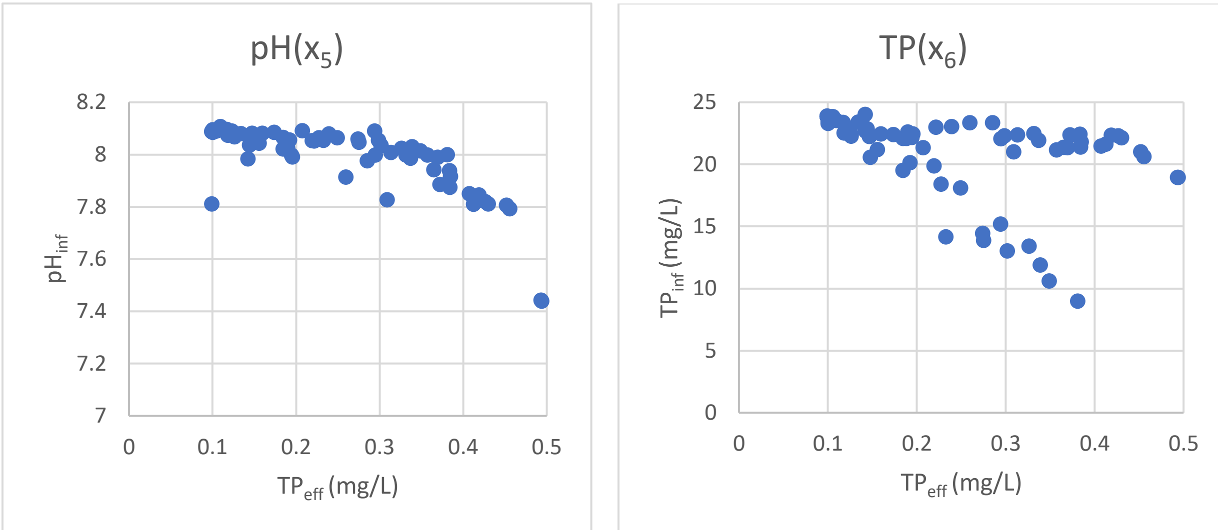


Figure 5.25 Optimal variation of pH_{inf} and TP_{inf} with TP_{eff} for case 4

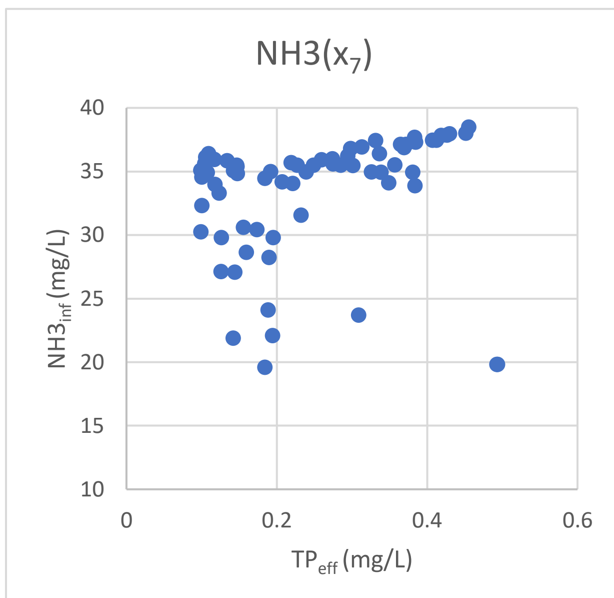


Figure 5.26 Optimal variation of $NH3_{inf}$ with TP_{eff} for case 4

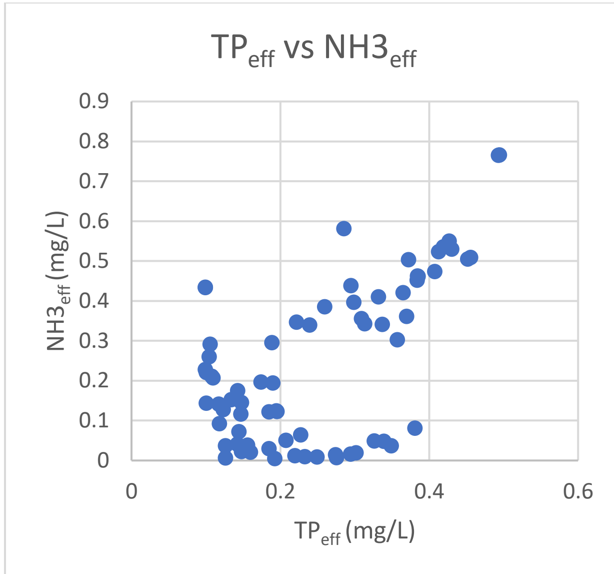


Figure 5.27 Variation of $NH3_{eff}$ with TP_{eff} for case 4

To visualize the variation of seven decision variables the optimal values of the decision variables are plotted against TP_{eff} (one of the objectives that are being minimized in case 4). From figures 5.23-5.26 it is observed that the graphs are more scattered as compared to cases 1-3. This is due to the addition of a third objective function. The optimal values of temperature range between 14-20°C as shown in figure 5.23. Like case 1 the optimal values of total flow lie close to the upper bound value and range between 35-55 ML/D as TP_{eff} concentration varies between 0-0.6 mg/L as observed in figure 5.23. From figure 5.24 it is observed that the optimal values of BOD_{inf} vary between 175-475 mg/L like in case 1. The optimal values of SS_{inf} lie close to upper bound value of 850 mg/L like cases 1, 2, and 3. For case 4 the optimal variation of SS_{inf} is least compared to six other decision variables. The optimal values of pH_{inf} range between 7.8-8.1 as shown in figure 5.25. From figure 5.26 it is observed that the optimal values of NH_{3inf} vary between 20-40 mg/L. Figure 5.27 represents the variation of NH_{3eff} with respect to TP_{eff} and there is no violation of the effluent regulations is terms of NH_{3eff} as the effluent concentration is less than 3 mg/L.

The computation time taken for optimizing seven decision variables was 8 minutes on 8GB RAM and a 1.8 GHz processor. After the 109th iteration, there was no improvement in the Pareto optimal front.

Table 5-9: GA parameters for case 4

Population size	200 x 7
Maximum number of generations	200 x 7
Crossover fraction	0.8
Pareto fraction	0.35
Constraint tolerance	1 x 10^{-3}
Function tolerance	1 x 10^{-4}

5.5 Summary and conclusions

In this chapter multi-objective optimization in WWTP is proposed to minimize the concentration of pollutants in the effluent stream. Multi-objective optimization problems are commonly encountered in real word as compared to single-objective optimization. When the goal is to improve the effluent quality, minimizing BOD in the effluent stream only might deteriorate SS and TP removal from the wastewater. In case of multi-objective optimization there is no single best solution but a set of solutions, also known as Pareto set. In this section of research GA is employed to minimize the concentration of BOD, SS, TP, and NH_3 in the effluent stream. Three cases of multi-objective optimization are formulated with two objectives each for better visualization of the Pareto curve. In fourth case three objectives have been minimized simultaneously which includes BOD, SS, and TP.

The goal of this research work is to find the optimum values of the decision variables and satisfy the objectives and constraints. The decision variables involved in this process are the temperature of the influent stream, total sewage flow, BOD_{inf}, SS_{inf}, pH_{inf}, TP_{inf}, and NH_{3inf}. The constraints imposed are in accordance with the regulatory requirements of effluent quality of treated wastewater.

Chapter 6

6 Modeling of activated sludge process

6.1 Introduction

An activated sludge process in an integral process of wastewater treatment. It is conventionally used across the world in sewage treatment plants. In an aeration tank, the organic compounds present in the wastewater are stabilized by microorganisms such as bacteria. In the presence of oxygen, microorganisms convert organic matter to carbon dioxide, ammonia, and new microorganisms. The oxygen keeps the contents of the aeration tank in a mixed state and allows bacteria to grow in suspension hence called a suspended growth system.

The activated sludge process consists of two separate chambers, an aeration tank and secondary sedimentation tank. In aeration tank microbes feed on organic, forming flocs which settle down easily. In a secondary sedimentation tank, which is also known as a secondary settling tank or secondary clarifier, where biological cell mass is separated from the effluent. The activated sludge which is also known as mixed liquor volatile suspended solids settles at the bottom of the tank while the effluent finds its way from the top. The recirculation of activated sludge is essential to maintain the concentration of microorganisms at a certain level. The sludge from the bottom of the secondary clarifier is recirculated back to the aeration tank.

Aeration conditions in the aeration tank are achieved using diffused or mechanical aerators. The diffused aeration takes place from the bottom of the tank whereas mechanical aeration occurs at the bottom. A diffused aeration system normally operates vertically. Compressed air is pumped through pipes and filters into the water through the diffusers, which creates small bubbles. As the bubbles rise, they create a spiral flow pattern. Thus, transferring oxygen into the water help bacteria in degrading organic material. The two main types of diffused aeration systems are fine bubble and coarse bubble aerators. Fine bubble aerators improve aeration and efficiency, whereas coarse bubble aerators have enhanced mixing and increased the level of dissolved oxygen.

A conventional mechanical aeration system operates horizontally. It consists of a pump and a motor that turns a propeller. This system churns up water and creates a current which in turn provides a mixing of wastewater. The equipment mixes aerated water with the rest of the water in the tank and brings in more air for aeration and mixing. Mechanical aerators provide stronger localized mixing, whereas diffused aerators provide complete mixing throughout the tank. In diffused aeration systems the bubbles originate from the bottom and rise upwards, hence mixing the whole tank. This avoids the chances of dead zones, in contrast, mechanical aerators cannot always reach the bottom of deep tanks. Therefore, diffused aerators provide more air to the wastewater system per unit of power.

6.2 Influent characterization

Influent characterization is important for development of good model as influent drives the rest of the model. Organic matter in wastewater can be broadly divided into two categories biodegradable and non-biodegradable(inert) material. Non-biodegradable organics are not degraded under any conditions.

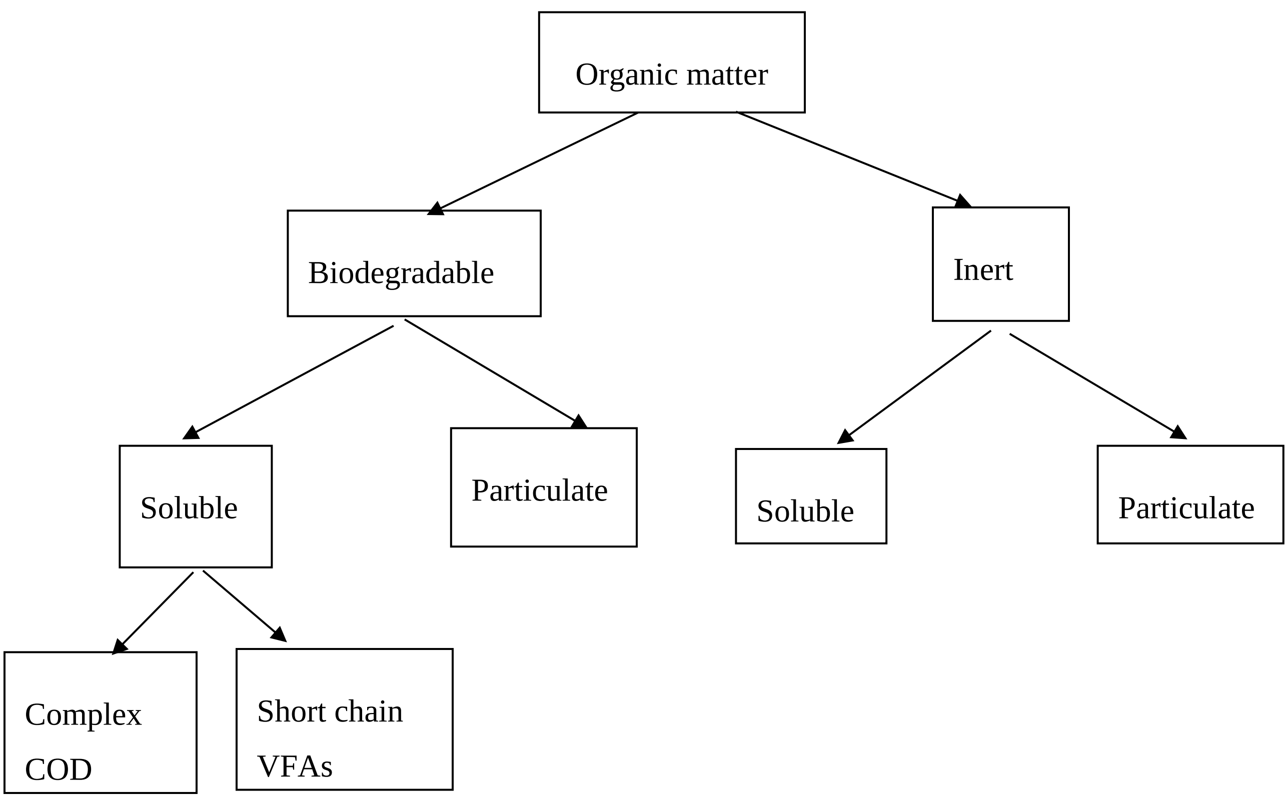


Figure 6.1 Classification of organic matter

Non-biodegradable organics are further sub-divided into soluble and particulate compounds. Particulate compounds form a significant part of primary sludge and impacts plant sludge production. Readily biodegradable organic material consists of small molecules and microorganisms that rapidly take up and consume. Slowly biodegradable organic material consists of larger molecules that require extracellular breakdown before the uptake of organic matter.

6.3 Activated sludge models

Activated sludge models are used to study the biological reactions in aeration tanks which consists of wastewater and microbes. In 1983 International water association (IWA) was formed to coordinate the modeling of the activated sludge process. The goal of IWA was to develop a simple model that could accurately predict biological processes. In 1987 ASM1 was developed to describe carbon oxidation, nitrification, and denitrification (Henze et al., 1987). In 1995 ASM2 was introduced that incorporated the removal of phosphorus by phosphorus-accumulating organisms (PAOs). ASM2 was further extended to ASM2d which included simultaneous phosphorus removal and denitrification by PAOs.

Various reactions involved in ASM1 are: -

1. Aerobic growth of heterotrophs

$$\rho_1 = \mu_H \left(\frac{S_s}{K_S + S_s}\right)\left(\frac{S_o}{K_{o,H} + S_o}\right) X_{B,H}$$

2. Anoxic growth of heterotrophs

$$\rho_2 = \mu_H \left(\frac{S_s}{K_S + S_s}\right)\left(\frac{K_{o,H}}{K_{o,H} + S_o}\right)\left(\frac{S_{NO}}{K_{NO} + S_{NO}}\right) X_{B,H}\eta_g$$

3. Aerobic growth of autotrophs

$$\rho_3 = \mu_A \left(\frac{S_{NH}}{K_{NH} + S_{NH}}\right)\left(\frac{S_o}{K_{o,A} + S_o}\right) X_{B,H}$$

4. Decay of heterotrophs

$$\rho_4 = b_H X_{B,H}$$

5. Decay of autotrophs

$$\rho_5 = b_A X_{B,A}$$

6. Ammonification of soluble organic nitrogen

$$\rho_6 = K_a S_{ND} X_{B,H}$$

7. Hydrolysis of entrapped organics

$$\rho_7 = K_h \frac{X_s / X_{B,H}}{K_X + X_s / X_{B,H}} \left[\left(\frac{S_o}{K_{o,H} + S_o} \right) + \eta_g \left(\left(\frac{K_{o,H}}{K_{o,H} + S_o} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \right) \right] X_{B,H}$$

8. Hydrolysis of entrapped organic nitrogen

$$\rho_8 X_{ND} / X_S$$

Table 6-1 Parameters and characteristics

Symbol	Name
S_{NO}	Soluble nitrate nitrogen concentration in water
S_{NH}	Soluble ammonia nitrogen concentration in water
S_S	Concentration of readily biodegradable COD in water
μ_A	Maximum specific growth rate of autotrophic biomass

K_{NH}	Ammonia half saturation coefficient of autotrophic biomass
b_H	Decay coefficient for heterotrophic biomass
X_s	Slowly biodegradable organic matter concentration
η_g	Correction factor for μ_H under anoxic conditions
η_H	Correction factor for hydrolysis under anoxic conditions
μ_H	Maximum specific growth rate for heterotrophic biomass
K_S	Half saturation coefficient for heterotrophic biomass
k_h	Maximum specific hydrolysis rate
K_X	Half-saturation coefficient for hydrolysis of slowly biodegradable
K_A	Ammonification rate
$K_{O,H}$	Oxygen half saturation coefficient for heterotrophic biomass

6.4 Methodology

In this study, GPS-X software by the company Hydromantis is used for the modeling and simulation of the wastewater treatment plant. Mantis2 a comprehensive model including biological, physical, and chemical processes in WWTP was developed by GPS-X. This model includes carbon, nitrogen, and phosphorus removal with an integrated anaerobic digestion process. The model considers a two-step nitrification and two-step denitrification process.

Various processes included in Mantis2 model are: -

1. Adsorption of colloidal COD: Adsorbed COD is considered slowly biodegradable COD and requires hydrolysis before its uptake.

2. Aerobic hydrolysis: Heterotrophic microorganisms hydrolysis slowly biodegradable substrate X_S to soluble substrate (SS).

$$\rho_2 = K_h.\frac{S_{O_2}}{K_{O_2} + S_{O_2}}.\frac{X_s/X_H}{K_X + X_s/X_H}.X_H$$

3. Anoxic hydrolysis: This process occurs under anoxic conditions. The oxygen saturation term in aerobic hydrolysis rate expression is replaced by an oxygen inhibition term. The specific hydrolysis rate is reduced by anoxic hydrolysis reduction factor(η_{NOx})

$$\rho_3 = K_h \eta_{NO3}\frac{K_{O_2}}{K_{O_2} + S_{O_2}}.\frac{S_{NO_3}}{K_{NO_3} + S_{NO_3}}.\frac{X_s/X_H}{K_X + X_s/X_H}.X_H$$

4. Anaerobic hydrolysis: This process occurs under anaerobic conditions. Anaerobic conditions mean that the process occurs in the absence of oxygen and electron acceptors such as nitrate. The specific hydrolysis rate is reduced by anaerobic hydrolysis reduction factor(η_{anaer})

$$\rho_4 = K_h \eta_{anaer}\frac{K_{O_2}}{K_{O_2} + S_{O_2}}.\frac{K_{NO_3}}{K_{NO_3} + S_{NO_3}}.\frac{X_s/X_H}{K_X + X_s/X_H}.X_H$$

5. Ammonification: In this process, soluble organic nitrogen is converted to ammonia nitrogen.

$$\rho_5 = K_A S_{ND} X_{BH}$$

6. Decay of heterotrophs: This process is described by the following equation

$$\rho_6 = b_H X_{BH}$$

6.4.1 Model Calibration

Calibration is the process of adjusting model parameters to improve the fit between predicted and actual data. First step towards model calibration is to assess the data. In this research influent and effluent data is collected from Adelaide wastewater treatment plant. The influent data is collected before primary treatment and is characterized in terms of temperature, flowrate, BOD, TSS, NH_3, and TP. The effluent data is collected after UV disinfection and is characterized in terms of temperature, BOD, TSS, NH_3, and TP. Table 6.1 shows average monthly influent and effluent quality parameters used for this study.

Table 6-2 Influent and effluent parameters of WWTP for study

case	1		2		3		4	
Parameter	Influent	Effluent	Influent	Effluent	Influent	Effluent	Influent	Effluent
Flow($10^3 m^3$/L)	26.24		36.32		27.04		20.4	
BOD (mg/L)	199	2	171	1	238	1	342	2
TSS (mg/L)	264	4	225	2	306	4	374	4
NH_3 (mg/L)	25	0.16	31	0.10	34	0.41	37	0.47
TP (mg/L)	10	0.46	11	0.47	13	0.46	15	0.54

6.4.2 Modeling and Simulation in GPS-X

Various steps involved in modeling and simulation are: -

1. Collection of data required for modeling
2. Construction of WWTP layout in GPS-X environment.
3. Characterization of influent wastewater quality parameters (such as BOD, TSS, NH_3, and TP)
4. Adjusting influent fractionation of organic and nitrogen compounds which are difficult to measure directly, using GPS-X influent advisor.
5. Running the model and calibration via adjusting kinetic, stoichiometric, and other parameters to fit the model.
6. Validate the calibrated data using a different set of data.

6.5 Results and discussions

In this section of the study, steady-state simulations were performed for model calibration and validation. Four sets of data were used to calibrate and validate the predicted results. Case 1 was used for calibration and Cases 2-4 were used for the validation of results. The kinetic parameters of activated sludge model were adjusted to match actual data. The performance of the model was measured in terms of effluent quality. In case 1 GPS-X default values were used for model calibration. The default values of GPS-X were not able to predict TSS and TP concentrations in the effluent stream as shown in figure 6.2. The kinetic parameters ($\mu_{max, H}$ K_{SS}, $K_{OH,}$ K_{NH4}, K_h, K_x, and K_A) were changed from 3.2, 5.0, 0.2, 0.05, 3, 0.1, and 0.08 to 7, 0.3, 0.15, 0.2, 12, 0.3, and 0.4 as shown in Table 6.3. From figure 6.2 it is observed that TSS concentration in the effluent stream was predicted best for case 3. However, this model fails to predict the TP concentration in the effluent stream.

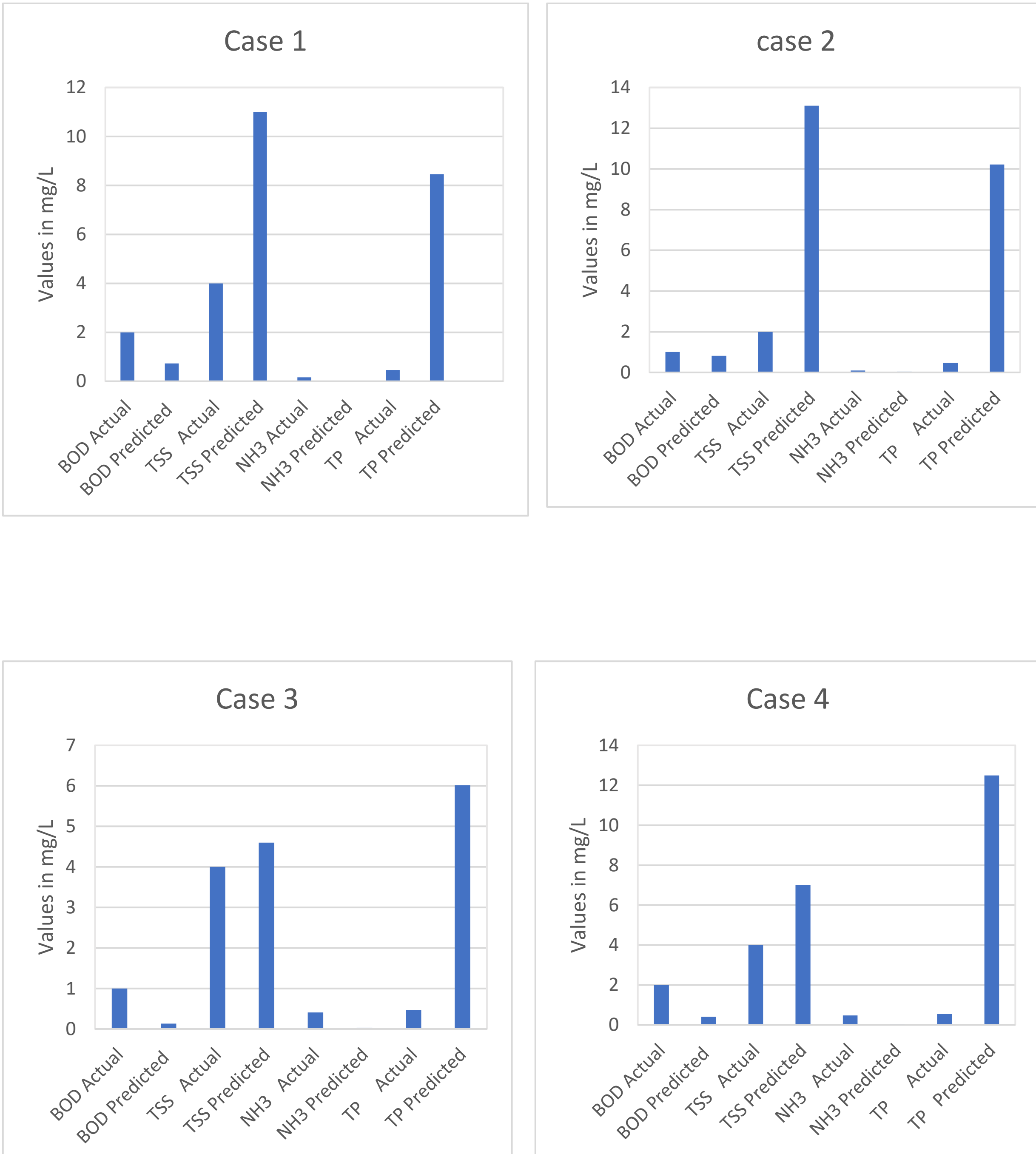


Figure 6.2 Calibration and validation data

Table 6-3 Influent parameters based on GPS-X influent advisor

Parameter	Unit	Case1	Case 2	Case 3	Case 4

VSS/TSS ratio	gVSS/gTSS	0.53	0.48	0.48	0.55
Soluble inert material	$gCOD/m^3$	17.5	16.7	23.3	33.5
Readily degradable soluble substrate	$gCOD/m^3$	59.9	65	102.8	160.8
Particulate inert material	$gCOD/m^3$	6.07	5.8	8.07	11.6
Heterotrophic biomass	$gCOD/m^3$	0	0.2	0.2	0.2

Table 6-4 Kinetic parameters for calibration and validation data

Parameter	Unit	Case 1	Case 2	Case 3	Case 4
$\mu_{max,\,H}$	1/d	3.2	7	7	7
K_{SS}	mgCOD/L	5	0.3	0.3	0.3
K_{OH}	mgO2/L	0.2	0.15	0.15	0.15
K_{NH4}	mgN/L	0.05	0.2	0.2	0.2
K_h	1/d	3	12	12	12

K_x	gCOD/gCOD	0.1	0.3	0.3	0.3
K_A	m^3/gCOD/d	0.08	0.4	0.4	0.4

6.6 Summary and conclusions

In this chapter modeling and simulation of WWTP using GPS-X are discussed. In this study, four sets of data for influent and effluent quality parameters were used for the calibration and validation of data sets. The modeling results showed that BOD, TSS, and NH_3 concentrations in the effluent stream were well predicted by the model but failed to capture TP concentration in the effluent stream. The model can be further improved to predict the TP concentration in the effluent stream. The volume of the aeration tank can be minimized using the optimization tool in GPS-X.

Chapter 7

7 Conclusion and Future Work

The purpose of wastewater treatment is to remove the impurities before discharging them back into the environment. Untreated wastewater is harmful to both humankind and the environment. Improper operation of WWTP can cause environmental and various health issues like cholera and dysentery. The optimal operation of WWTP can improve efficiency and reduce the costs associated with various processes. In this research work, a multi-objective optimization approach has been used to minimize the concentration of pollutants in the effluent stream instead of a single-optimization approach. In the real world, multi-objective problems with conflicting objectives are frequently encountered. In this case, a set of equally good solutions is generated, also known as the Pareto set. Though sometimes it becomes difficult for the decision maker to choose a single optimal solution from a set of optimal solutions.

The motivation behind wastewater treatment and the various stages involved in wastewater treatment is discussed in chapter 1. Wastewater treatment plants are major consumers of energy. Therefore, it is important to operate them optimally. Various optimization techniques such as the simplex method, GA, ACO, and PSO have been briefly discussed in this chapter. In this study, multi-objective optimization is performed using GA to minimize the concentration of pollutants in the effluent stream. Pareto optimality and MOGA are explained in Chapter 1.

Wastewater treatment is a complex system, and it is difficult to explore various design ideas on a pilot plant. Modeling helps in understanding how a system would behave in various conditions without experimentation. A WWTP model is a representation of physical and chemical processes involved in the purification of wastewater. Chapter 2 presents the work of researchers in the area of modeling and simulation of wastewater treatment. In my research work, a black-box modeling approach has been employed to model WWTP. This type of modeling is based on the input-output behaviour of the process in contrast to physical modeling which is time-consuming.

In chapter 3 Adelaide WWTP is briefly discussed and data obtained from plant is analysed. The purpose of this research is to model the WWTP and predict its performance in terms of effluent quality. In first section of this research ANN has been used to model the WWTP. The data set used to build a model was (temp, flowrate, BOD, SS, pH, TP, and NH_3 of influent and effluent stream) collected and analysed. The performance of a model is dependent on the preparation of data. Various steps involved in the preparation of data are:-gathering the data, handling missing data, deciding which key factors are important, and splitting the data into training and validation set. The measurements of influent flow rate, temperature, BOD, SS, pH, TP,NH_3 in the influent stream and effluent stream were collected and analysed over four-year period.

In chapter 4 model based on ANN was developed to predict the quality of effluent stream. The purpose of ANN model was to identify the pattern between various parameters in the influent and effluent stream. In this research work the seven variables that define the influent stream are temperature, flow rate, BOD, SS, pH, TP, and NH_3 and seven variables that define the effluent stream are temperature, BOD, SS, pH, TP, NH_3 and DO. The effluent quality is measured in terms of four major pollutants namely BOD, SS, TP, and NH_3. The complete data set is divided in 7:3 ratio for training and validation. The results indicate that ANN can predict the quality of effluent stream as the correlation coefficient between the actual values and the predicted values is close to 0.98. In conclusion ANN is an effective tool for predicting the performance of non-linear and complex WWTP.

In chapter 5 GA is employed to minimize the concentration of BOD, SS, TP, and NH_3 in the effluent stream. Three cases of multi-objective optimization are formulated with two objectives each for better visualization of the Pareto curve. In fourth case three objectives have been minimized simultaneously which includes BOD, SS, and TP.

Chapter 6 focuses on modeling and simulation of WWTP using GPS-X are discussed. In this study, four sets of data for influent and effluent quality parameters were used for the calibration and validation of data sets. The modeling results showed that BOD, TSS, and NH_3 concentrations in the effluent stream were well predicted by the model but failed to capture TP concentration in the effluent stream. The model can be further improved to predict the TP concentration in the effluent stream. The volume of the aeration tank can be minimized using the optimization tool in GPS-X.

Future work

The results reported in this research work were accurate in prediction of the effluent quality. Implementation of optimization techniques as described in chapter 5 can increase the performance of the plant and reduce the operating cost by running the plant at optimum conditions. Future research work should focus on prediction of influent flow and the quality of influent stream. This information can provide useful insight to the operators and help them in better management of the plant. The research can further be extended to dynamic modeling of WWTP. Dynamic modeling will help to simulate the plant operation precisely and this will take into the account of hourly, daily, and monthly variation of influent flow and quality of influent stream.

Aeration tanks are important components of wastewater treatment and significant amount of energy is consumed in aeration of the tanks. Future work should focus on optimization of aeration energy consumption. This will improve the efficiency of the treatment plant. The optimum design of aeration tanks can reduce the fixed capital investment.

8 Appendices

Table 1 - Dataset of Adelaide WWTP (Influent stream)

Data point	Temperature sewage Raw (°C)	Total sewage flow (ML/D)	BOD_5 Raw (mg/L)	Suspended solids Raw (mg/L)	pH Raw	Total Phosphorus Raw (mg/L)	Ammonia (mg/L)
1.	17	23.62	196	343	7.4	10.3	18.1
2.	15	21.15	87	68	7.4	3.9	22.2
3.	15	21.63	147	123	7.4	5.4	24.6
4.	13	29.23	76	53	7.5	2.9	14.6
5.	15	27.8	194	265	7.5	5.6	14.1
6.	13	33.36	158	198	7.4	4.7	20.9
7.	19	23.61	428	574	7.1	9.4	21.3
8.	20	23.41	151	156	7.6	6.6	34.7
9.	20	22.75	166	205	7.5	8.1	34.4
10.	20	24.57	385	815	7.1	12.4	21.7
11.	20	22.62	215	358	7.3	12.8	20.7
12.	20	22.45	242	323	7.3	8.1	20.8

13.	21	21.93	210	273	7.2	6.3	24.7
14.	21	24.53	138	152	7.5	6.5	30.3
15.	21	24.85	350	232	7.2	8.8	22.4
16.	21	23.83	341	379	7.3	9.4	24
17.	21	17.61	209	363	7.2	7.6	26.2
18.	21	18.6	196	263	7.4	6.8	23.6
19.	21	22.87	308	436	7.3	8.5	25.6
20.	20	24.33	276	242	7.4	7.9	28.5
21.	20	20.64	174	178	7.7	6.8	36.3
22.	20	21.31	108	84	7.5	5.2	27.4
23.	20	18.67	316	439	7.4	9.2	25.7
24.	20	18.77	286	349	7.4	7.7	26.1
25.	20	18.52	330	439	7.4	8.5	29.6
26.	19	23.85	294	349	7.4	6.8	19.8
27.	20	24.18	269	355	7.5	7.4	24.7
28.	19	25.42	271	376	7.5	10.8	27.1

29.	19	25.39	316	408	7.5	9.6	25.2
30.	19	26.41	168	153	7.3	5.7	26.8
31.	19	24.76	169	200	7.7	7.4	27.8
32.	19	24.91	361	549	7.5	12.4	24.3
33.	18	25.93	358	566	7.4	11.9	23.6
34.	18	24.59	609	950	7.4	11.4	24.6
35.	19	23.22	288	357	7.5	9	26.6
36.	19	23.15	132	121	7.9	5.7	31.8
37.	19	22.58	145	111	7.6	5.2	27.3
38.	19	23.40	296	418	7.4	10.8	25.3
39.	18	23.14	311	372	7.5	8.2	21
40.	18	30.02	131	118	7.7	4.9	26.2
41.	13	24.75	158	331	7.5	7.8	25.7
42.	13	25.60	202	191	7.6	6.2	20.6
43.	14	25.07	292	266	7.6	5.4	21.5
44.	15	23.76	252	293	7.5	8.4	20.6

45.	15	23.37	224	320	7.5	7.8	24.1
46.	16	23.07	225	199	7.5	6	22.8
47.	15	26.61	246	288	7.5	9.2	16.4
48.	16	24.63	343	188	7.5	5.9	19.9
49.	14	24.33	163	180	8.2	6.3	38.4
50.	14	28.66	140	174	8	5	27.8
51.	13	33.73	132	121	7.6	3.6	15.1
52.	14	29.23	467	825	7.4	11.1	17.5
53.	13	27.35	201	271	7.8	6.3	23.9
54.	14	28.62	296	525	7.5	8.3	19.1
55.	14	37.78	167	242	7.6	4.2	10.7
56.	13	32.78	157	238	7.6	4.4	13.9
57.	14	30.44	271	504	7.6	9	18.4
58.	16	28.72	230	396	7.5	7.7	16.1
59.	14	30.11	191	257	7.5	6.8	19.3
60.	14	27.42	199	284	7.5	6.6	17.9

61.	16	26.22	243	352	7.5	9.7	18.8
62.	14	25.86	509	706	7.2	13.4	19.7
63.	14	26.14	144	101	7.5	5.5	20.7
64.	17	29.27	161	137	7.5	5.4	20.8
65.	17	21.84	307	482	7.5	11.6	21.1
66.	17	23.21	217	403	7.5	8.8	17.8
67.	18	23.13	253	402	7.5	9.5	19.8
68.	18	23.8	440	652	7.4	13.5	20.9
69.	18	24.44	211	271	7.8	8.4	36
70.	21	23.94	397	564	7.4	12.7	28.7
71.	21	24.65	242	393	7.5	9.2	27.6
72.	21	27.44	108	150	7.6	4.5	18
73.	21	24.54	138	220	7.7	7	26.1
74.	21	20.99	140	227	7.5	8.2	23.2
75.	22	21.77	237	325	7.4	7.6	20.5
76.	21	25.85	329	675	7.3	11	15.4

77.	21	28.2	286	364	7.3	8.3	18.4
78.	21	33.39	146	265	7.4	6.7	18.8
79.	21	29.82	112	99	7.5	3.8	17.4
80.	21	26.38	144	211	7.5	6	23.5
81.	21	38.39	224	254	7.6	6.8	27.4
82.	21	26.95	238	333	7.4	6.5	17.8
83.	21	26.97	123	141	7.5	5.3	20.1
84.	22	27.46	214	232	7.4	6.9	18.2
85.	21	26.11	266	378	7.4	7.9	20
86.	20	26.85	286	416	7.4	8.5	22.9
87.	20	25.17	365	375	7.4	9.7	24.8
88.	19	27.18	251	399	7.5	10.1	22.8
89.	19	27.59	125	88	7.5	4.3	24.1
90.	19	21.55	249	388	7.4	9	21.9
91.	18	21.24	240	429	7.4	8.4	25
92.	17	20.10	269	383	7.5	9	25.1

93.	18	21.19	283	358	7.5	9.1	25.5
94.	17	20.69	184	244	7.5	7.4	20
95.	17	20.94	223	248	7.6	6.8	26.5
96.	17	21.25	259	316	7.5	7.1	20.3
97.	17	21.17	305	434	7.5	9	25.5
98.	16	20.34	231	326	7.5	7.3	21.9
99.	16	27.32	177	247	7.9	6	20.8
100.	15	31.65	138	155	7.9	5.7	25
101.	15	26.76	257	309	7.5	6.8	18
102.	13	38.03	202	248	7.5	5.7	16.5
103.	13	31.04	174	261	7.5	5.8	15.4
104.	14	25.91	176	255	7.5	5.3	15.4
105.	14	24.59	171	225	7.5	5.7	15.8
106.	14	25.13	208	208	7.6	5.7	15.4
107.	13	24.1	247	362	7.5	8.5	17.2
108.	13	23.12	147	173	8	7	28.6

109.	13	32.72	269	477	7.5	10	18.5
110.	13	29.56	201	219	7.5	5.4	17.6
111.	13	25.97	180	262	7.5	5.1	15.5
112.	13	24.63	231	266	7.5	5.4	17.8
113.	13	25.61	241	281	7.5	5.4	17.9
114.	14	27.21	257	314	7.4	5.5	17.2
115.	13	37.52	272	343	7.5	6.9	16.8
116.	12	34.41	171	232	7.5	4.7	9.4
117.	13	35.38	199	255	7.4	6	16.8
118.	12	26.8	137	174	7.5	4.4	16.8
119.	12	24.41	168	240	7.5	5.5	18.1
120.	12	25.64	135	195	7.5	5	19.1
121.	14	26.42	176	270	7.6	7.5	21.6
122.	13	26.36	195	267	7.5	4.7	20.3
123.	13	27.1	118	123	8.2	5.7	31.8
124.	13	38.57	111	263	7.5	8.2	17.6

125.	13	43.25	109	188	7.4	5.5	17.3
126.	14	30.15	192	236	7.5	7.6	19.6
127.	13	46.94	234	211	7.5	6.6	19.2
128.	13	31.7	160	125	7.5	4.6	16.7
129.	14	34	240	340	7.5	9.1	18.1
130.	13	35.28	219	366	7.5	6.2	12.7
131.	13	34.71	367	404	7.5	7.6	12.7
132.	14	30.2	278	465	7.4	9.1	18
133.	15	28.46	226	409	7.4	9.7	16.8
134.	15	37.96	179	231	7.7	7.2	25
135.	15	29.73	170	234	7.6	25	15.5
136.	15	28.52	124	156	7.9	5.9	29.1
137.	16	29.04	169	201	7.4	6.4	17.5
138.	16	27.24	232	258	7.4	6.4	17.4
139.	17	26.25	179	296	7.5	17.5	17.5
140.	17	24.91	183	141	7.5	6.3	20.7

141.	17	24.85	180	225	7.4	6.9	17.6
142.	17	25.18	236	243	7.5	7.3	22.2
143.	18	24.84	189	192	7.4	6.5	21.1
144.	18	26.65	230	215	7.5	7.4	22.1
145.	18	25.47	397	482	7.3	10.1	33.7
146.	19	26.81	99	237	7.5	8.1	21.8
147.	19	26.81	354	494	7.4	11.4	21.7
148.	19	28.55	202	245	7.4	7.2	22.2
149.	19	26.77	102	283	7.5	8.8	22.9
150.	20	26.08	248	380	7.5	10.3	21.7
151.	19	25.97	236	163	7.5	6.7	24
152.	21	25.92	221	333	7.4	9.8	23.3
153.	20	26.5	268	409	7.5	10.1	22.9
154.	20	26.05	202	324	7.6	8.7	34.5
155.	21	25.62	263	378	7.6	6.4	27.2
156.	20	24.87	124	122	8	6.1	37.3

157.	20	26.17	306	419	7.5	9.3	29.8
158.	20	25.86	259	427	7.4	10.2	28.2
159.	20	26.04	163	151	7.8	7.8	37.1
160.	21	22.88	254	375	7.4	9.9	28.1
161.	21	23.15	222	286	7.5	7.8	29.4
162.	21	21.72	357	387	7.4	9	29.1
163.	21	19.88	215	195	7.5	7.1	26.8
164.	20	18.53	337	399	7.4	10	29.3
165.	20	19.61	226	276	7.6	10	37
166.	20	20.72	396	511	7.3	10.4	29.5
167.	20	19.80	561	441	7.4	11.7	29.7
168.	20	19.28	358	425	7.4	9.9	30.2
169.	20	19.63	298	289	7.4	9	29.2
170.	20	18.98	416	605	7.4	11.1	26.5
171.	19	19.12	234	187	7.4	5.5	26.1
172.	19	18.06	256	236	7.4	6.7	29.1

173.	18	25.73	336	131	7.4	6.1	24.5
174.	18	21.17	248	353	7.5	9.1	24.3
175.	19	20.20	254	299	7.5	5.9	24.2
176.	18	19.91	236	266	7.5	7.3	28
177.	18	20.45	371	354	7.4	8.4	25.3
178.	18	24.76	240	298	7.6	7.3	23.2
179.	18	20.39	243	209	7.6	5.3	22.9
180.	17	20.69	339	415	7.5	8.4	26.5
181.	17	21.08	227	219	7.5	5.7	25.8
182.	17	20.46	240	221	7.6	6.2	26.9
183.	17	19.43	232	155	7.5	5.7	24.9
184.	16	19.73	199	210	7.8	6.9	37
185.	16	18.31	260	220	7.5	6	28.7
186.	14	16.38	264	229	7.7	7.1	40.2
187.	15	18.83	416	437	7.5	9.6	29.6
188.	15	18.78	336	260	7.6	7.4	29.7

189.	15	18.38	379	518	7.4	9.9	28.2
190.	13	56.48	322	293	7.5	7.7	26
191.	13	22.05	252	253	7.6	6.2	19.6
192.	14	20.42	202	143	7.6	5.1	19.5
193.	13	44.22	222	240	7.5	8.6	15.7
194.	14	21.31	237	195	7.6	6.4	21
195.	14	22.49	270	243	7.6	6	21.7
196.	14	32.82	307	310	7.6	8.6	24.1
197.	14	21.62	391	338	7.5	8.1	26.3
198.	13	32.69	182	145	7.5	4.5	19.6
199.	11	69.51	226	282	7.4	4.7	14.8
200.	12	29.98	280	193	7.6	4.6	15.7
201.	13	26.62	218	179	7.6	4.3	17.5
202.	13	24.38	160	263	7.6	4.3	19
203.	13	25.14	529	479	7.5	7.7	19.4
204.	11	22.99	229	232	7.5	5.7	19.7

205.	12	22.24	249	102	7.6	4.6	20.6
206.	13	22.19	236	247	7.6	6.4	21.8
207.	13	22.03	265	263	7.6	6.4	22.8
208.	13	24.81	334	566	7.6	13.1	26.3
209.	13	33.68	228	259	7.5	6.1	18.6
210.	12	29.18	196	168	7.6	4.4	15.3
211.	13	25.66	282	258	7.6	6.5	19.1
212.	13	25.22	253	148	7.5	9.3	20.2
213.	12	35.4	165	245	7.6	5	11.2
214.	12	33.41	215	102	7.6	4.1	12.6
215.	13	30.73	310	321	7.6	7.1	17.7
216.	13	29.66	175	199	7.5	4.3	16.3
217.	15	31.17	312	415	7.5	9.7	18.4
218.	14	29.37	178	197	7.7	6.9	28.5
219.	14	27.88	198	376	7.5	5.9	19.6
220.	14	25.32	256	212	8.1	7.9	36.4

221.	15	29.91	169	241	7.4	5.7	17.6
222.	15	27.44	190	205	8.1	6.2	35
223.	15	25.19	168	191	7.9	6.8	35
224.	15	24.65	264	407	7.5	7.1	22
225.	17	28.76	227	304	7.5	6.9	20.6
226.	16	28.64	253	326	7.5	7	22.4
227.	16	24.01	223	275	7.5	9.3	24.6
228.	16	25.21	302	283	7.5	6.4	23.7
229.	17	25.41	266	362	7.5	9.1	23.9
230.	16	25.26	177	186	8.2	6.4	39.9
231.	18	28.74	143	170	8	5.5	32.6
232.	18	25.42	619	893	7.6	15.9	35.9
233.	17	25.18	311	364	7.5	8	22.3
234.	18	25.26	297	365	7.5	7.1	23.6
235.	18	26.37	246	357	7.4	6.9	21.5
236.	19	35.59	512	541	7.5	8.7	22.5

237.	19	26.28	214	377	7.6	7.7	25.1
238.	18	25.63	392	537	7.4	9.7	31.2
239.	19	29.52	238	413	7.4	8.1	22.7
240.	19	27.68	619	580	7.4	10.6	23.3
241.	20	23.72	214	398	7.5	9.3	21.8
242.	20	22.14	392	572	7.5	10.5	22.7
243.	20	24	269	394	7.4	7.3	26.1
244.	19	23.46	298	283	7.8	7.7	36.9
245.	20	24.59	417	495	7.4	10.9	26.6
246.	20	25.63	133	134	7.5	6	25.3
247.	19	29.87	169	289	7.6	5.1	27.3
248.	20	26.18	154	333	7.7	4.3	18.5
249.	21	27.79	211	417	7.5	7.2	22.3
250.	20	26.39	349	350	7.5	8.1	22.2
251.	21	27.72	295	472	7.4	8.5	25.5
252.	21	26.4	256	276	7.5	9.3	26.2

253.	20	23.83	300	361	7.6	9.9	41.2
254.	20	26.01	251	148	7.5	8.2	28.5
255.	20	24.45	132	444	7.3	8.1	26.2
256.	20	23	267	374	7.2	8.5	27.9
257.	20	21.75	242	312	7.4	9	26.6
258.	20	26.98	283	210	7.5	7.7	23.4
259.	20	23.81	287	263	7.6	6.8	23.3
260.	20	27.74	206	249	7.8	7.9	41.8
261.	20	25.95	244	159	7.8	6.6	35.9
262.	19	23.58	325	395	7.4	9.7	26.2
263.	19	22.67	227	177	7.5	6.4	25.6
264.	19	22.58	365	315	7.5	10.6	29.1
265.	19	22.22	297	385	7.5	7.6	24.5
266.	18	26.57	139	404	7.7	7.2	27.5
267.	18	38.46	301	325	7.7	11.2	32.9
268.	17	31.26	282	256	7.6	6.3	20.8

269.	17	23.81	238	257	7.6	5.7	21.9
270.	16	23.03	366	346	7.9	9.3	41
271.	17	22.40	231	301	7.7	7	21.9
272.	17	21.65	350	236	7.6	6.2	24.5
273.	16	27.00	194	253	7.6	5.3	15.7
274.	15	26.17	272	218	7.6	4.6	20
275.	16	26.72	224	655	7.5	8.6	11.8
276.	15	24.97	398	338	7.6	6.2	20.8
277.	15	24.57	332	423	7.5	9.3	22.2
278.	15	24.87	294	332	7.5	7.1	21.9
279.	15	23.19	105	253	7.5	6.9	22.1
280.	14	23.92	224	258	7.5	7.8	24.8

Table 2 - Dataset of Adelaide WWTP (Effluent stream)

Data	Temp UV	BOD5 UV (	Suspended s	pH UV cha	Total Phospsphorus (	NH3 UV	DO plant
point			solids UV		UV		effluent
	°C	mg/L	mg/L		mg/L	mg/L	mg/L
1	15	2	4	7.3	0.26	0.1	7.9
2	15.5	2	4	7.1	0.71	0.1	7.9
3	14.7	4	6	6.9	0.42	0.62	7.3
4	13.7	2	4	7.3	0.61	1.75	6.9
5	14.4	3	9	7.5	0.4	8.78	6.3
6	14.1	3	6	7.2	0.57	1.49	7.1
7	20.4	1	1	7	0.55	0.25	7
8	20.6	1	1	6.8	0.49	0.1	5.4
9	21.1	1	1	7.2	0.56	0.1	5.5
10	20.5	1	1	6.9	0.42	0.1	7.2
11	21.1	1	4	7.2	0.43	0.1	7.1
12	20.7	1	4	7	0.54	1.69	7.2
13	21.5	2	1	6.9	0.72	0.1	7.5
14	22.1	1	4	7.3	0.58	0.1	7.8
15	22.4	1	1	7.4	0.65	0.1	7.4
16	21.3	3	1	7.3	0.9	0.1	6.9
17	21.3	1	1	7	0.38	0.1	7.1
18	21.4	1	5	7.4	0.33	0.1	7.2
19	21.2	1	4	7.3	0.66	0.1	7
20	20	1	1	7.2	0.62	0.65	6.6
21	20.4	4	8	7.2	0.32	1.26	6.6

22	19.5	3	1	7.2	0.41	1.9	7
23	19.6	1	1	7.6	0.33	0.1	7
24	19.9	3	6	6.9	0.54	1.37	7
25	19.2	2	6	7.2	0.4	0.1	7.3
26	18.5	1	7	7.1	0.61	0.18	7.3
27	19.4	1	1	7.3	0.52	0.1	7.2
28	18.6	1	4	7.1	0.44	0.1	7.4
29	18.7	1	4	7.1	0.57	0.1	6.5
30	18.6	2	5	6.9	0.47	0.1	6.6
31	19	1	1	7	0.5	1.16	6.2
32	18	2	1	7.1	0.96	0.1	6.8
33	18.3	3	5	7	0.64	0.13	6.7
34	18.1	3	1	7.1	0.54	0.1	7
35	18.2	2	1	7.1	0.54	0.1	6.5
36	18.5	1	1	7	0.64	0.1	6.6
37	18.5	3	3	7.1	0.66	0.1	6.6
38	18.7	3	4	7	0.64	0.1	6.6
39	18.6	3	1	7.1	0.55	0.1	6.6
40	17.6	1	1	7.3	0.37	0.1	7
41	18.3	1	1	7.3	0.82	0.1	6.6
42	17.7	3	6	7.2	0.37	0.1	6.3
43	18.1	1	1	7.3	0.48	0.1	6.5

44	17	1	4	7.1	0.53	0.1	7
45	18.5	1	1	7	0.38	0.1	6.5
46	17.9	1	1	8.4	0.67	0.15	6.2
47	18.4	1	1	7.2	0.46	0.4	6.2
48	18.4	1	1	7.1	0.39	0.32	6.2
49	18.8	3	6	7.1	0.45	0.31	6.2
50	18.4	2	4	7.2	0.23	0.22	6
51	18.1	3	4	7.2	0.55	0.34	6.2
52	19.4	2	1	7.2	0.34	0.1	6.1
53	19.5	1	5	7.5	0.78	0.1	6.1
54	19.4	2	4	7.1	0.47	0.14	6.2
55	18.2	1	1	7.2	0.37	1.06	5.6
56	18	2	1	7.3	0.34	0.16	6.7
57	18.7	1	1	7.2	0.4	0.1	6.9
58	16.1	1	6	7.1	0.24	0.1	7.4
59	16.1	1	1	7.2	0.85	0.1	6.7
60	16.1	1	1	7.1	0.37	0.1	7.1
61	16.1	1	1	7.1	0.4	0.1	6.6
62	16.1	1	1	7	0.4	0.1	7
63	17.6	1	4	7.2	0.48	0.1	7.1
64	16.9	2	4	7	0.4	0.1	6.7
65	17.3	2	1	7.1	0.41	0.1	6.9
66	17.2	1	6	7.3	0.44	1.96	6.7

67	18.3	1	1	7.1	0.57	0.23	7
68	18.1	1	1	7.2	0.62	0.12	6.8
69	18.4	2	1	7.1	0.56	0.4	7
70	23	1	1	7.2	0.63	0.63	6.8
71	22.7	1	1	6.9	0.3	0.1	7.3
72	22.5	1	1	7.2	0.32	0.12	7.2
73	19.8	1	1	7	0.54	0.1	7.6
74	22.8	1	1	7	0.6	0.1	7.2
75	23.1	1	1	7.2	0.69	0.1	7.1
76	20.7	1	1	7.1	0.45	0.1	7.7
77	21.2	1	1	7.2	0.52	0.13	7.4
78	22.1	1	1	7.3	0.37	0.14	5.2
79	22.4	1	4	7.5	0.21	0.1	7.9
80	22	3	1	7.1	0.29	0.1	8.1
81	21.7	3	1	7.4	0.52	0.14	7.4
82	21.1	3	1	7.2	0.48	0.1	7.7
83	20.9	1	1	7.4	0.72	0.1	8.8
84	21.8	1	1	7.2	0.56	2.29	7.3
85	21.2	1	1	7.1	0.39	0.1	6.5
86	19.6	1	1	7.3	0.55	1.27	6
87	20.4	1	1	7.2	0.69	0.27	5.9
88	19	1	1	7.2	0.61	0.38	5.8
89	18.9	1	1	7.2	0.48	0.1	6

90	18.3	5	4	7.3	0.88	0.88	5.9
91	18.3	1	6	6.2	0.24	0.1	6
92	17.1	1	1	6.9	0.24	0.1	6
93	17.4	1	1	7	0.77	0.1	6
94	17.3	1	4	7.1	0.44	0.1	5.3
95	17.2	2	6	6.9	0.72	0.1	5.7
96	16.4	3	4	7.2	0.65	0.1	5.9
97	16	3	1	7.1	0.64	0.1	5.8
98	15.7	3	4	7	0.83	0.1	5.9
99	15.6	3	6	7.1	0.4	0.42	5.5
100	14.7	2	1	7.3	0.33	0.1	5.6
101	14	1	1	7.3	0.28	0.1	6.6
102	13	2	1	7.2	0.36	0.1	5.1
103	13.4	1	1	7.3	0.55	0.1	6
104	13.5	3	6	7.3	0.93	0.1	6.3
105	13.4	1	10	7.4	0.3	0.1	6
106	13.5	1	6	7.3	0.27	0.1	6.5
107	13.1	2	4	7.1	0.51	0.67	5.3
108	13.1	1	4	7.1	0.93	0.12	5.7
109	13.2	2	1	7.1	0.31	0.33	5.5
110	12.5	2	1	7.3	0.28	0.1	6.7
111	12.7	1	1	7.3	0.19	0.1	6.5
112	12.5	1	1	7.2	0.26	0.1	6

113	13	2	8	7.3	0.35	0.1	6.3
114	13.8	2	1	7.2	0.46	0.1	5.3
115	12.8	3	1	7.3	0.3	0.1	6.7
116	11.5	2	1	7.4	0.55	1.62	6.8
117	13	3	1	7.3	0.82	0.71	5.6
118	13	1	4	7.3	0.28	0.1	7.1
119	11.7	2	1	7.3	0.29	0.26	5.7
120	11.6	1	1	7.2	0.47	0.27	5.7
121	13.4	1	1	7.2	0.41	0.32	4.7
122	13.1	2	6	7.2	0.75	0.55	4.7
123	13	1	1	7.3	0.38	0.1	5.1
124	13.1	1	1	7.3	0.56	0.1	6.1
125	12.6	1	1	7.4	0.42	0.1	6.1
126	13.7	1	1	7.2	0.88	0.1	5.3
127	13.5	1	4	7.3	0.3	0.1	6
128	13.5	1	1	7.2	0.32	0.1	6.2
129	14.4	1	4	7.4	0.46	0.1	6.3
130	13.4	1	4	7.3	0.4	0.18	6.3
131	13.5	1	1	7.4	0.19	0.1	7.3
132	14.3	1	1	7.5	0.33	0.1	6.3
133	15.5	2	1	7.5	0.22	0.1	6.6
134	15.3	1	4	7.4	0.22	0.22	5.8
135	16	1	7	7.4	0.3	3.94	4.7

136	15.7	1	6	7.4	0.64	1.21	4.1
137	16.1	2	1	7.1	0.22	0.56	5
138	16.3	1	1	7.3	0.46	0.1	5.8
139	17.5	1	1	7.4	0.39	0.39	5.1
140	17.3	2	3	7.4	0.51	0.44	5.1
141	17.6	1	1	7.3	0.38	1.47	4.9
142	17.6	1	1	7.4	0.67	0.49	5.9
143	17.6	2	1	7.1	0.61	0.54	7.7
144	17.6	1	1	7.3	0.76	0.73	5.6
145	18.4	1	1	7.1	0.34	0.28	4.6
146	19.1	1	1	7.1	0.52	0.23	7.3
147	19.5	1	4	7.1	0.31	0.24	6.9
148	20	1	16	7.3	0.54	0.48	7.2
149	19.3	1	1	7.1	0.6	1.01	7.3
150	20.6	1	1	7.1	0.77	0.27	7.1
151	19.9	1	1	7.3	0.65	0.16	5.4
152	20.2	2	1	7.4	0.49	0.1	7.2
153	21	1	3	7.1	0.72	0.35	6.9
154	21.1	1	1	7.5	0.43	0.2	4.4
155	21.5	3	1	7.5	0.48	0.1	4.8
156	20.1	3	4	7.5	0.6	0.1	4.5
157	19.5	1	6	7.4	0.49	0.15	5.3
158	20.5	1	4	7.1	0.62	0.11	4.8

159	20.9	1	2	7.1	0.49	0.84	4.6
160	21.3	1	2	7	0.42	0.9	4.3
161	20.3	2	4	7.2	0.52	1.81	4.9
162	21.7	3	2	7.4	0.66	0.19	7.1
163	20.6	2	8	7.2	0.52	1.06	6.8
164	20.5	2	7	7.2	0.46	0.1	6.7
165	20.8	1	5	7.2	0.59	0.24	7.6
166	20.8	3	2	7.2	0.43	1.07	6.9
167	19.5	3	2	7.2	0.57	1.67	7.1
168	19.5	2	1	7.3	0.63	0.52	7.3
169	19.7	1	2	7.2	0.5	0.32	7.2
170	20.2	2	4	7.4	0.51	0.11	7
171	19.1	1	4	7.2	0.79	0.1	7.4
172	18.9	3	6	7.3	0.4	0.1	6.9
173	18.7	2	2	7.2	0.43	0.1	7
174	18.1	2	4	7.1	0.56	0.1	7.5
175	17.7	3	2	7.3	0.57	0.38	7.6
176	17.8	3	4	7.5	0.57	0.1	7.5
177	17.1	2	2	7.3	0.54	0.1	8
178	16.2	2	2	7.4	0.73	0.18	8
179	16.3	1	2	7.3	0.46	0.22	8
180	16.6	1	2	7.3	0.38	0.56	7.5
181	16.5	1	6	7.3	0.43	0.1	7.5

182	16.3	1	4	7.3	0.49	0.1	7.8
183	15.5	2	4	7.2	0.46	0.36	7.8
184	15.1	1	2	7.3	0.48	0.81	7.6
185	14.1	3	4	7.1	0.54	1.97	7.6
186	13.5	2	2	7.3	0.46	2.68	7.8
187	13.1	4	8	7.3	0.7	2.95	7.7
188	13.3	4	4	7.2	0.76	1.78	7.6
189	13.9	1	2	7	0.33	0.11	8.1
190	13.7	2	4	7	0.43	0.1	5.2
191	11.7	1	4	7.2	0.1	0.14	9.1
192	11.8	1	4	7.2	0.16	0.1	8.7
193	10.9	3	4	7.2	0.51	1.01	6.7
194	11.3	2	2	7.3	0.38	0.12	9.3
195	12.4	2	2	7.3	0.56	0.26	8.5
196	11.4	3	6	7	0.21	0.95	8.4
197	12	3	4	7.2	0.54	1.66	7.5
198	12.8	2	2	7.2	0.56	2.04	7.4
199	10.4	4	5	7.3	0.4	3.86	4.3
200	11.6	1	8	7.5	0.33	2.95	7.5
201	11.4	2	4	7.5	0.32	4.68	8.3
202	11.3	3	4	7.3	0.37	1.61	8
203	11.1	4	2	7.3	0.33	2.95	8
204	11.3	3	2	7.2	0.51	0.17	8.1

205	11.6	2	2	7.2	0.43	0.14	8.3
206	12.2	3	2	7.2	0.49	0.1	8.2
207	11.9	2	4	7.2	0.44	0.89	8.4
208	12.1	3	4	7.3	0.52	0.39	8.1
209	12.2	2	2	7.3	0.38	0.1	8.2
210	11.3	3	2	7.4	0.53	2.95	8.1
211	11.9	2	4	7.3	0.22	0.14	8.6
212	12.2	3	2	7.2	0.98	0.46	8.6
213	10.8	1	6	7.2	0.41	2.64	7.8
214	11	3	2	7.3	0.27	1.77	8
215	12.6	3	5	7.2	0.26	0.37	8.1
216	12.2	1	4	7.3	0.29	1.92	7.8
217	13.4	1	2	7.2	0.26	0.45	8.1
218	13.9	1	2	7.2	0.24	0.1	8.2
219	14.3	2	2	7.1	0.4	0.13	8.5
220	14.3	1	2	7.3	0.4	0.12	7.8
221	14.5	1	2	7	0.32	0.5	7.6
222	15.2	1	2	7.2	0.37	0.12	8.5
223	15.2	2	2	7.2	0.49	0.1	8.1
224	16.2	1	2	7.3	0.45	0.16	7.8
225	17.1	1	2	7.1	0.63	0.65	7.1
226	16.5	1	4	7.4	0.39	0.46	7.2
227	16.2	2	4	7	0.96	0.6	7.8

228	17.1	2	6	7.3	0.49	0.32	7.4
229	17.2	1	2	7.1	0.6	0.13	7.7
230	17.7	1	2	7.5	0.42	0.31	7.8
231	18.2	2	2	7.1	0.46	0.42	7.6
232	17.9	2	2	7.3	0.41	0.52	7.8
233	17.7	3	2	7.2	0.31	0.1	8.2
234	18.2	1	2	7.2	0.43	0.16	7.9
235	18.9	1	3	7.1	0.48	0.46	7.7
236	19.8	1	4	7.4	0.49	0.7	7.5
237	20	1	2	7.3	0.38	0.14	7.7
238	19.1	2	2	7.3	0.47	0.16	7.7
239	19.8	3	6	7.6	0.27	0.15	7.6
240	20.5	3	6	7.9	0.34	0.7	7.1
241	20	1	6	7.2	0.42	0.1	7.9
242	20.2	2	5	7.4	0.45	0.34	7.5
243	20.6	1	2	7.4	0.5	0.9	7
244	20.4	1	4	7.4	0.43	0.27	6.9
245	20.8	4	2	7.2	0.7	0.23	7.3
246	20.8	1	2	7.4	0.61	0.23	7.4
247	20.6	1	2	7.2	0.22	0.1	7
248	20.3	1	2	7.4	0.28	0.12	7.9
249	21.4	2	3	7.3	0.38	0.14	7.5
250	20.8	1	2	7.5	0.48	0.11	7.6

251	21.6	1	2	7.1	0.33	1.16	7.1
252	21.4	1	2	7.3	0.39	0.48	4.2
253	20.4	1	4	7.3	0.48	0.1	4.6
254	20.6	2	2	7.4	0.58	0.69	4.6
255	21.1	1	2	7.2	0.56	0.27	4.5
256	20.7	1	7	7	0.61	1.22	4
257	20.4	1	15	7.3	0.51	0.1	5.4
258	20.1	1	2	7.4	0.4	0.1	5.6
259	20.6	1	2	7.4	0.47	0.1	4.7
260	21	1	2	7.5	0.43	0.24	7.1
261	20.7	1	2	7.5	0.78	0.1	7.5
262	18.7	1	22	7.3	0.77	0.11	7.3
263	19	1	4	7.4	0.8	0.48	7.4
264	18.7	1	4	7.3	0.5	0.11	7.3
265	18.3	1	8	7.3	0.61	0.26	7.5
266	18.1	1	6	7.3	0.54	0.29	4.3
267	17.8	1	4	7.5	0.69	0.96	5.8
268	17.7	1	2	7.4	0.26	0.65	7
269	17	1	2	7.6	0.43	0.1	6.1
270	16.3	1	6	7.5	0.56	0.1	6.5
271	16.5	1	2	7.5	0.35	0.1	6
272	15.3	4	4	7.4	0.53	0.13	8.5
273	15	2	6	7.4	0.52	0.1	6.3

274	15.2	1	6	7.6	0.35	0.12	5.7
275	14.8	1	2	7.5	0.25	0.1	8.1
276	14.9	4	7	7.4	0.43	0.1	6.4
277	14.5	1	2	7.3	0.48	0.1	8.3
278	14.9	1	2	7.4	0.52	0.1	6.1
279	14.4	1	2	7.3	0.5	0.11	8.5
280	14	3	5	7.3	0.83	0.44	8.7

MATLAB Code for modeling WWTP using ANN

```
i=Inputandoutputdata(:,1:7);
o=Inputandoutputdata(:,8:14);
hiddenlayersize=10;
net=fitnet(hiddenlayersize);
net.divideParam.trainRatio=70/100;
net.divideParam.valRatio=30/100;
net.divideParam.testRatio=0/100;
[net,tr]=train(net,it,ot);

iTrain=it(:,tr.trainInd);
iVal=it(:,tr.valInd);
oTrain=net(it(:,tr.trainInd));
oTrainTrue=ot(:,tr.trainInd);
oValTrue=ot(:,tr.valInd);
oVal=net(it(:,tr.valInd));
oValTruet=oValTrue';
oValt=oVal';
sum(sqrt(mean((oValt-oValTruet).^2)))
%omin=oValt(:,2:5);
%Extracting weights
w1=net.IW(1,1);
w2=net.Lw(2,1);
b1=net.b(1);
b2=net.b(2);
save('Trained-Network.mat')
%end
```

W1 (Weights)

10 – neurons in hidden layer 7 – neurons in input layer

2.253922	1.012001	0.776434	-0.76715	-0.30355	1.518476	-0.25827
0.108927	2.51738	-0.51042	-0.57277	-0.27406	0.574018	-0.82383
-0.62377	-0.1403	0.557544	-0.98586	0.870841	0.538559	-0.43542
1.355239	-0.08942	0.392527	-0.63427	0.996416	-0.29188	-0.45111
0.1643	1.310789	-0.36146	-0.78076	-1.38619	-0.48319	0.432218
0.480692	0.644446	-0.59926	0.778919	-1.59881	0.326442	1.639477
-0.77806	-1.31416	0.641651	0.191947	-0.16566	1.39329	-0.47019
0.190104	-0.29839	0.949314	-1.48205	0.837722	0.086483	-0.97314
1.810092	0.115601	-0.58561	1.348571	1.10291	-0.16102	-0.80061
0.947126	0.391462	-0.94071	-0.59502	-0.29847	-1.55539	0.086428

W2 (Weights)

10 – neurons in hidden layer 7 – neurons in output layer

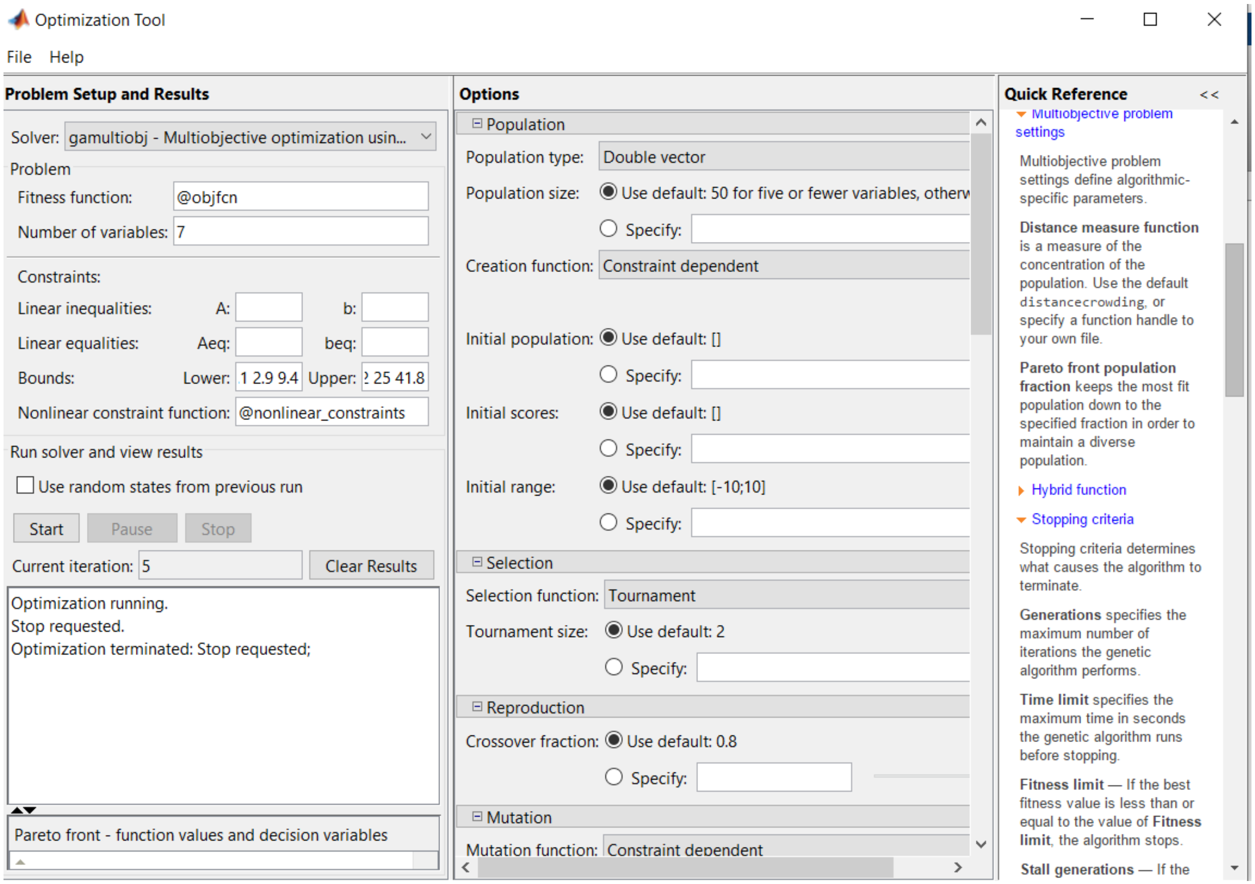
0.23957	-0.31176	-0.08818	0.498041	0.184522	-0.004	0.123773	-0.34742	0.319021	0.211791
-0.13407	0.71778	0.016002	0.341452	0.140406	0.010841	0.476699	0.113107	-0.3365	0.22181
-0.24538	0.277369	-0.30762	-0.47931	-0.34014	0.460775	-0.41579	0.884604	0.043217	-0.58364
0.217608	-0.21141	0.436445	0.364821	0.221385	-0.23608	0.2643	-0.55732	0.029908	0.683168
-0.08001	0.62341	-0.45555	-0.15021	-0.17874	0.481181	-0.08369	0.89001	-0.27152	0.277857
0.072898	0.446614	-0.05111	-0.18507	-0.0756	0.147128	-0.13035	0.32896	-0.02287	-0.12603
-0.06786	-0.39208	-0.85468	0.15253	-0.39986	-0.27746	0.046166	0.473178	-0.59217	-0.73152

b1 (bias)

-0.93089
-2.2919
0.329838
-0.43255
-0.24852
0.068391
-0.5646
0.147647
1.426659
2.260947

b2

-0.15493
0.176487
-0.70683
-0.45933
0.028225
-0.4663
0.520238


Optimization Tool
File Help
Problem Setup and Results
Solver: gamultiobj - Multiobjective optimization usin...
Problem
Fitness function: @objfcn
Number of variables: 7
Constraints:
Linear inequalities: A: b:
Linear equalities: Aeq: beq:
Bounds: Lower: 1 2.9 9.4 Upper: 25 41.8
Nonlinear constraint function: @nonlinear_constraints
Run solver and view results
Use random states from previous run
Start Pause Stop
Current iteration: 5 Clear Results
Optimization running.
Stop requested.
Optimization terminated: Stop requested;
Pareto front - function values and decision variables
Options
Population
Population type: Double vector
Population size: Use default: 50 for five or fewer variables, otherw
Specify:
Creation function: Constraint dependent
Initial population: Use default: []
Specify:
Initial scores: Use default: []
Specify:
Initial range: Use default: [-10;10]
Specify:
Selection
Selection function: Tournament
Tournament size: Use default: 2
Specify:
Reproduction
Crossover fraction: Use default: 0.8
Specify:
Mutation
Mutation function: Constraint dependent
Quick Reference
Multiobjective problem settings
Multiobjective problem settings define algorithmic-specific parameters.
Distance measure function is a measure of the concentration of the population. Use the default distancecrowding, or specify a function handle to your own file.
Pareto front population fraction keeps the most fit population down to the specified fraction in order to maintain a diverse population.
Hybrid function
Stopping criteria
Stopping criteria determines what causes the algorithm to terminate.
Generations specifies the maximum number of iterations the genetic algorithm performs.
Time limit specifies the maximum time in seconds the genetic algorithm runs before stopping.
Fitness limit — If the best fitness value is less than or equal to the value of Fitness limit, the algorithm stops.
Stall generations — If the

Case 1
Objective function for minimization of 2 objectives

```
function effconc=objfcn(x)
load('Trained-Network.mat');
y=sim(net,x');
effconc(1)=y(2)';
if y(5)<0
effconc(2)=y(5)'+0.1;
end
if y(5)>0
effconc(2)=y(5)';
end

function [C Ceq]=nonlinear_constraints(x)
load('Trained-Network.mat');
y=sim(net,x');
C(1)=-y(2);
C(2)=y(6)-3;
C(3)=-y(6);
C(4)=-y(5);
C(5)=-y(3);
C(6)=y(3)-10;
Ceq=[];
```

Case 2
Objective function for minimization of 2 objectives

```
function effconc=objfcn(x)
load('Trained-Network.mat');
y=sim(net,x');
effconc(1)=y(2)';
effconc(2)=y(3)';

function [C Ceq]=nonlinear_constraints(x)
load('Trained-Network.mat');
y=sim(net,x');
C(1)=-y(2);
C(2)=y(6)-3;
C(3)=-y(6);
C(4)=-y(5);
C(5)=y(5)-0.5;
C(6)=-y(3);
Ceq=[];
```

Case 3
Objective function for minimization of 2 objectives

```
function effconc=objfcn(x)
load('Trained-Network.mat');
y=sim(net,x');
if y(3)>0
effconc(1)=y(3)';
end
if y(3)<0
effconc(1)=y(3)'+0.1;
end
if y(5)>0
effconc(2)=y(5)';
end
if y(5)<0
effconc(2)=y(5)'+0.1;
end


function [C Ceq]=nonlinear_constraints(x)
load('Trained-Network.mat');
y=sim(net,x');
C(1)=-y(2);
C(2)=y(2)-10;
C(3)=y(6)-3;
C(4)=-y(6);
C(5)=-y(5);
C(6)=-y(3);
Ceq=[];
```

```
Case 4
Objective function for minimization of 3 objectives

function effconc=objfcn(x)
load('Trained-Network.mat');
y=sim(net,x');
if y(5)<0
effconc(1)=y(5)'+0.1;
end
if y(5)>0
effconc(1)=y(5)';
end
effconc(2)=y(2)';
if y(3)<0
effconc(3)=y(3)'+0.1;
end
if y(3)>0
effconc(3)=y(3)';
end


function [C Ceq]=nonlinear_constraints(x)
load('Trained-Network.mat');
y=sim(net,x');
C(1)=-y(2);
C(2)=y(6)-3;
C(3)=-y(6);
C(4)=-y(5);
C(5)=-y(3);
Ceq=[];
```

MATLAB code for plotting three objectives

```
x=paretofrontplot(:,1);
y=paretofrontplot(:,2);
z=paretofrontplot(:,3);
%plot3(x,y,z)
figure
%[X,Y,Z]=meshgrid(x,y,z);
scatter3(x,y,z)
xlabel('TP(mg/L)');
ylabel('BOD(mg/L)');
zlabel('SS(mg/L)');
```

Kinetics and stoichiometry for ASM1(Henze et al., 1987)

j	Component → i / Process ↓	1 S_I	2 S_S	3 X_I	4 X_S	5 $X_{B,H}$	6 $X_{B,A}$	7 X_P	8 S_O	9 S_{NO}	10 S_{NH}	11 S_{ND}	12 X_{ND}	13 S_{ALK}	Process Rate, ρ_j $[ML^{-3}T^{-1}]$
1	Aerobic growth of heterotrophs		$-\frac{1}{Y_H}$			1			$-\frac{1-Y_H}{Y_H}$		$-i_{XB}$			$-\frac{i_{XB}}{14}$	$\hat{\mu}_H\left(\frac{S_S}{K_S+S_S}\right)\left(\frac{S_O}{K_{O,H}+S_O}\right)X_{B,H}$
2	Anoxic growth of heterotrophs		$-\frac{1}{Y_H}$			1				$-\frac{1-Y_H}{2.86Y_H}$	$-i_{XB}$			$\frac{1-Y_H}{14\cdot 2.86Y_H}$ $-i_{XB}/14$	$\hat{\mu}_H\left(\frac{S_S}{K_S+S_S}\right)\left(\frac{K_{O,H}}{K_{O,H}+S_O}\right)$ $\times\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right)\eta_g X_{B,H}$
3	Aerobic growth of autotrophs						1		$-\frac{4.57-Y_A}{Y_A}$	$\frac{1}{Y_A}$	$-i_{XB}-\frac{1}{Y_A}$			$-\frac{i_{XB}}{14}-\frac{1}{7Y_A}$	$\hat{\mu}_A\left(\frac{S_{NH}}{K_{NH}+S_{NH}}\right)\left(\frac{S_O}{K_{O,A}+S_O}\right)X_{B,A}$
4	'Decay' of heterotrophs				$1-f_P$	-1		f_P					$i_{XB}-f_P i_{XP}$		$b_H X_{B,H}$
5	'Decay' of autotrophs				$1-f_P$		-1	f_P					$i_{XB}-f_P i_{XP}$		$b_A X_{B,A}$
6	Ammonification of soluble organic nitrogen										1	-1		$\frac{1}{14}$	$k_a S_{ND} X_{B,H}$
7	'Hydrolysis' of entrapped organics		1		-1										$k_h \frac{X_S/X_{B,H}}{K_X+(X_S/X_{B,H})}\left[\left(\frac{S_O}{K_{O,H}+S_O}\right)+\eta_h\left(\frac{K_{O,H}}{K_{O,H}+S_O}\right)\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right)\right]X_{B,H}$
8	'Hydrolysis' of entrapped organic nitrogen											1	-1		$\rho_7(X_{ND}/X_S)$
	Observed Conversion Rates $[ML^{-3}T^{-1}]$	$r_i=\sum_j \nu_{ij}\rho_j$													
	Stoichiometric Parameters: Heterotrophic yield: Y_H Autotrophic yield: Y_A Fraction of biomass yielding particulate products: f_P Mass N/Mass COD in biomass: i_{XB} Mass N/Mass COD in products from biomass: i_{XP}	Soluble inert organic matter [M(COD)L^{-3}]	Readily biodegradable substrate [M(COD)L^{-3}]	Particulate inert organic matter [M(COD)L^{-3}]	Slowly biodegradable substrate [M(COD)L^{-3}]	Active heterotrophic biomass [M(COD)L^{-3}]	Active autotrophic biomass [M(COD)L^{-3}]	Particulate products arising from biomass decay [M(COD)L^{-3}]	Oxygen (negative COD) [M(−COD)L^{-3}]	Nitrate and nitrite nitrogen [M(N)L^{-3}]	$NH_4^+ + NH_3$ nitrogen [M(N)L^{-3}]	Soluble biodegradable organic nitrogen [M(N)L^{-3}]	Particulate biodegradable organic nitrogen [M(N)L^{-3}]	Alkalinity – Molar units	Kinetic Parameters: Heterotrophic growth and decay: $\hat{\mu}_H$, K_S, $K_{O,H}$, K_{NO}, b_H Autotrophic growth and decay: $\hat{\mu}_A$, K_{NH}, $K_{O,A}$, b_A Correction factor for anoxic growth of heterotrophs: η_g Ammonification: k_a Hydrolysis: k_h, K_X Correction factor for anoxic hydrolysis: η_h

References

Al, R., Behera, C. R., Zubov, A., & Sin, G. (2018). *Systematic framework development for the construction of surrogate models for wastewater treatment plants*. *Computer Aided Chemical Engineering* (Vol. 44). Elsevier Masson SAS. https://doi.org/10.1016/B978-0-444-64241-7.50313-X

Aminabad, M. S., Malek, A., & Hadi, M. (2014). Application of artificial neural network (ANN) for the prediction of water treatment plant influent characteristics. *J Adv Environ Health Res*, *1*(2). https://doi.org/http://jaehr.muk.ac.ir

Ansari, M., Othman, F., Abunama, T., & El-Shafie, A. (2018). Analysing the accuracy of machine learning techniques to develop an integrated influent time series model: case study of a sewage treatment plant, Malaysia. *Environmental Science and Pollution Research*, *25*(12), 12139–12149. https://doi.org/10.1007/s11356-018-1438-z

Asce, M., Godfrey, C. S., & Stewart, H. T. (1983). Wastewater treatment: sensitivity analysis, *109*(1), 120–138.

Béraud, B., Steyer, J. P., Lemoine, C., Latrille, E., Manic, G., & Printemps-Vacquier, C. (2007). Towards a global multi objective optimization of wastewater treatment plant based on modeling and genetic algorithms. *Water Science and Technology*, *56*(9), 109–116. https://doi.org/10.2166/wst.2007.592

Boyd, G., Na, D., Li, Z., Snowling, S., Zhang, Q., & Zhou, P. (2019). Influent Forecasting for Wastewater Treatment Plants in North America. *Sustainability*, *11*(6), 1764. https://doi.org/10.3390/su11061764

Campos, J. L., Valenzuela-Heredia, D., Pedrouso, A., Val Del Río, A., Belmonte, M., & Mosquera-Corral, A. (2016). Greenhouse Gases Emissions from Wastewater Treatment Plants: Minimization, Treatment, and Prevention. *Journal of Chemistry*, *2016*. https://doi.org/10.1155/2016/3796352

Chachuat, B., Roche, N., & Latifi, M. A. (2000). Dynamic optimisation of small size wastewater treatment plants including nitrification and denitrification processes. *Computer Aided*

Chemical Engineering, *8*(C), 853–858. https://doi.org/10.1016/S1570-7946(00)80144-3

Chen, J., Ganigué, R., Liu, Y., & Yuan, Z. (2014). Real-time multistep prediction of sewer flow for online chemical dosing control. *Journal of Environmental Engineering (United States)*, *140*(11), 1–9. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000860

Coello, C. A. C., Lamont, G. B., & Veldhuizen, D. A. Van. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. *Evolutionary Algorithms for Solving Multi-Objective Problems*. https://doi.org/10.1007/978-0-387-36797-2

Dantzig, G. B. (1990). Origins of the simplex method. In *A history of scientific computing* (pp. 141–151). https://doi.org/10.1145/87252.88081

Deb, K. (2011). *Multi-Objective Optimization Using Evolutionary Algorithms: An Introduction*. *Kanpur Genetic Algorithm Laoratory*. Retrieved from http://www.iitk.ac.in/kangal/deb.htm

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197. https://doi.org/10.1109/4235.996017

Egea, J. A., & Gracia, I. (2013). *Dynamic Multiobjective Global Optimization of a Waste Water Treatment Plant for Nitrogen Removal*. *IFAC Proceedings Volumes* (Vol. 45). IFAC. https://doi.org/10.3182/20120215-3-at-3016.00066

El-Din, A. G., & Smith, D. W. (2002). A neural network model to predict the wastewater inflow incorporating rainfall events. *Water Research*, *36*(5), 1115–1126. https://doi.org/10.1016/S0043-1354(01)00287-1

Emamgholizadeh, S., Kashi, H., Marofpoor, I., & Zalaghi, E. (2014). Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models. *International Journal of Environmental Science and Technology*, *11*(3), 645–656. https://doi.org/10.1007/s13762-013-0378-x

Flores-Alsina, X., Gernaey, K. V., & Jeppsson, U. (2011). Global sensitivity analysis of the BSM2 dynamic influent disturbance scenario generator. In *8th IWA Symposium on Systems*

Analysis and Integrated Assessment (pp. 722–729). https://doi.org/10.2166/wst.2012.089

Fu, G., Khu, S. T., & Butler, D. (2009). Use of surrogate modelling for multiobjective optimisation of urban wastewater systems. *Water Science and Technology*, *60*(6), 1641–1647. https://doi.org/10.2166/wst.2009.508

Gavin, H. P. (2020). The Levenberg-Marquardt Algorithm For Nonlinear Least Squares Curve-Fitting Problems. *Duke University*, 1–19. Retrieved from http://people.duke.edu/~hpgavin/ce281/lm.pdf

Gernaey, K. V, Rosen, C., Benedetti, L., & Jeppsson, U. (2005). Phenomenological Modeling of Wastewater Treatment Plant Influent Disturbance Scenarios. *10th International Conference on Urban Drainage, Copenhagen/Denmark*, *1*(1), 21–26.

Giner-Sanz, J. J., Leverick, G. M., Pérez-Herranz, V., & Shao-Horn, Y. (2020). Salicylate Method for Ammonia Quantification in Nitrogen Electroreduction Experiments: The Correction of Iron III Interference. *Journal of The Electrochemical Society*, *167*(13), 134519. https://doi.org/10.1149/1945-7111/abbdd6

Gunantara, N., & Nurweda Putra, I. D. N. (2019). The Characteristics of Metaheuristic Method in Selection of Path Pairs on Multicriteria Ad Hoc Networks. *Journal of Computer Networks and Communications*, *2019*. https://doi.org/10.1155/2019/7983583

Gupta, D., & Singh, S. K. (2012). Greenhouse Gas Emissions From Wastewater Treatment Plants : A Case Study of Greenhouse Gas Emissions from Wastewater Treatment Plants : A Case Study of Noida. *Journal of Water Sustainability*, *2*(2), 131–139.

Hamed, M. M., Khalafallah, M. G., & Hassanien, E. A. (2004). Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling and Software*, *19*(10), 919–928. https://doi.org/10.1016/j.envsoft.2003.10.005

Henze, M., Gujer, W., Mino, T., & van Loosedrecht, M. (1987). *Activated Sludge Model No.1*. *International Association on Water Pollution Research and Control*. https://doi.org/10.2166/9781780402369

Jiang, G., Sharma, K. R., Guisasola, A., Keller, J., & Yuan, Z. (2009). Sulfur transformation in rising main sewers receiving nitrate dosage. *Water Research*, *43*(17), 4430–4440. https://doi.org/10.1016/j.watres.2009.07.001

Kim, D., Bowen, J. D., & Ozelkan, E. C. (2015). Optimization of wastewater treatment plant operation for greenhouse gas mitigation. *Journal of Environmental Management*, *163*, 39–48. https://doi.org/10.1016/j.jenvman.2015.07.005

Kim, M., Kim, Y., Kim, H., Piao, W., & Kim, C. (2016). Evaluation of the k-nearest neighbor method for forecasting the influent characteristics of wastewater treatment plant. *Frontiers of Environmental Science and Engineering*, *10*(2), 299–310. https://doi.org/10.1007/s11783-015-0825-7

L. Haldurai, Madhubala, T., & Rajalakshmi, R. (2016). A Study on Genetic Algorithm and its Applications. *International Journal of Computer Sciences and Engineering*, *6*(10), 139–143.

Li, J., Sharma, K., Liu, Y., Jiang, G., & Yuan, Z. (2019). Real-time prediction of rain-impacted sewage flow for on-line control of chemical dosing in sewers. *Water Research*, *149*, 311–321. https://doi.org/10.1016/j.watres.2018.11.021

Maria, A. (1997). Introduction to modeling and simulation. *Winter Simulation Conference Proceedings*, 7–13. https://doi.org/10.1145/268437.268440

Marler, R. T., & Arora, J. S. (2010). The weighted sum method for multi-objective optimization: New insights. *Structural and Multidisciplinary Optimization*, *41*(6), 853–862. https://doi.org/10.1007/s00158-009-0460-7

Martin, C., & Vanrolleghem, P. A. (2014a). Analysing, completing, and generating influent data for WWTP modelling: A critical review. *Environmental Modelling and Software, 60,* 188–201. https://doi.org/10.1016/j.envsoft.2014.05.008

Martin, C., & Vanrolleghem, P. A. (2014b). Analysing, completing, and generating influent data for WWTP modelling: A critical review. *Environmental Modelling and Software*, *60*(December 2015), 188–201. https://doi.org/10.1016/j.envsoft.2014.05.008

Mjalli, F. S., Al-Asheh, S., & Alfadala, H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, *83*(3), 329–338. https://doi.org/10.1016/j.jenvman.2006.03.004

Ngatchou, P., Zarei, A., & El-Sharkawi, M. A. (2005). Pareto multi objective optimization. In *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems, ISAP'05* (Vol. 2005, pp. 84–91). IEEE. https://doi.org/10.1109/ISAP.2005.1599245

Ozturk, M. C., Martin Serrat, F., & Teymour, F. (2016). Optimization of aeration profiles in the activated sludge process. *Chemical Engineering Science*, *139*, 1–14. https://doi.org/10.1016/j.ces.2015.09.007

Plappally, A. K., & Lienhard V, J. H. (2012). Energy requirements for water production, treatment, end use, reclamation, and disposal. *Renewable and Sustainable Energy Reviews*, *16*(7), 4818–4848. https://doi.org/10.1016/j.rser.2012.05.022

Qiao, J., & Zhang, W. (2018). Dynamic multi-objective optimization control for wastewater treatment process. *Neural Computing and Applications*, *29*(11), 1261–1271. https://doi.org/10.1007/s00521-016-2642-8

Rojas, J., & Zhelev, T. (2012). Energy efficiency optimisation of wastewater treatment: Study of ATAD. *Computers and Chemical Engineering*, *38*, 52–63. https://doi.org/10.1016/j.compchemeng.2011.11.016

Singh, K. P., Basant, A., Malik, A., & Jain, G. (2009). Artificial neural network modeling of the river water quality-A case study. *Ecological Modelling*, *220*(6), 888–895. https://doi.org/10.1016/j.ecolmodel.2009.01.004

Sweetapple, C., Fu, G., & Butler, D. (2014). Multi-objective optimisation of wastewater treatment plant control to reduce greenhouse gas emissions. *Water Research*, *55*(0), 52–62. https://doi.org/10.1016/j.watres.2014.02.018

Szelag, B., Bartkiewicz, L., Studziński, J., & Barbusiński, K. (2017). Evaluation of the impact of explanatory variables on the accuracy of prediction of daily inflow to the sewage treatment plant by selected models nonlinear. *Archives of Environmental Protection*, *43*(3), 74–81. https://doi.org/10.1515/aep-2017-0030

Szilveszter, S., Ráduly, B., Ábrahám, B., Lányi, S., & Niculae, D. R. (2010). Mathematical models for domestic biological wastewater treatment process. *Environmental Engineering and Management Journal*, *9*(5), 629–636. https://doi.org/10.30638/eemj.2010.086

Thunberg, a., Sundin, A.-M., & Carlsson, B. (2009). Energy optimization of the aeration process at Käppala wastewater treatment plant. *10th IWA Conference on Instrumentation, Control and Automation*.

US EPA. (1998). Wastewater Treatment Works... The Basics, (May).

Wei, X., & Kusiak, A. (2015). Short-term prediction of influent flow in wastewater treatment plant. *Stochastic Environmental Research and Risk Assessment*, *29*(1), 241–249. https://doi.org/10.1007/s00477-014-0889-0

Wei, X., Kusiak, A., & Sadat, H. R. (2013). Prediction of Influent Flow Rate: Data-Mining Approach. *Journal of Energy Engineering*, *139*, 118–123. https://doi.org/10.1061/(ASCE)EY.1943-7897

Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., & Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, *54*(1), 101–121. https://doi.org/10.1623/hysj.54.1.101

Yeniay, Ö. (2005). Penalty function methods for constrained optimization with genetic algorithms. *Mathematical and Computational Applications*, *10*(1), 45–56. https://doi.org/10.3390/mca10010045

Zhang, Z., Kusiak, A., Zeng, Y., & Wei, X. (2016). Modeling and optimization of a wastewater pumping system with data-mining methods. *Applied Energy*, *164*, 303–311. https://doi.org/10.1016/j.apenergy.2015.11.061

Curriculum Vitae

Name:	Navneet Kaur
Post-Secondary Education And Degrees:	Thapar University Patiala, Punjab, India 2010-2014 B.E.
	University of Petroleum and Energy Studies Dehradun, Uttarakhand, India 2015-2017 M.Tech.
	The University of Western Ontario London, Ontario, Canada 2017-2022 Ph.D.
Honours and Awards:	Western Graduate Research Scholarship 2017-2022
Related Work Experience:	Teaching Assistant The University of Western Ontario 2017-2021
	Research Intern Indian Oil Corp Ltd. (IOCL) R&D Faridabad, Haryana, India 2016-2017

Publications:

Navneet Kaur and Ajay K. Ray Modeling of wastewater treatment process using Artificial Neural Network(ANN), In progress

Navneet Kaur and Ajay K. Ray Multi-objective optimization of wastewater treatment process using Genetic Algorithm(GA), In progress