



**UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA
EXPERIMENTAL YACHAY**

Escuela de Ciencias Químicas e Ingeniería

**TÍTULO: Predictive Modeling of the Primary Settling Tank in the Ibarra
Wastewater Treatment Plant based on Artificial Neural Networks**

Trabajo de integración curricular presentado como requisito para
la obtención del título de Petroquímico

Autor

Carlos Eduardo Veloz Marmolejo

Tutor

Marvin Ricaurte, Ph.D

Urququí, Noviembre 2020

SECRETARÍA GENERAL
(Vicerrectorado Académico/Cancillería)
ESCUELA DE CIENCIAS QUÍMICAS E INGENIERÍA
CARRERA DE PETROQUÍMICA
ACTA DE DEFENSA No. UITEY-CHE-2020-00062-AD

A los 5 días del mes de noviembre de 2020, a las 17:00 horas, de manera virtual mediante videoconferencia, y ante el Tribunal Calificador, integrado por los docentes:

Presidente Tribunal de Defensa	Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.
Miembro No Tutor	Dra. HIDALGO BONILLA, SANDRA PATRICIA , Ph.D.
Tutor	Dr. RICAURTE FERNÁNDEZ, MARVIN JOSÉ , Ph.D.

El(la) señor(ita) estudiante **VELOZ MARMOLEJO, CARLOS EDUARDO**, con cédula de identidad No. **0503748766**, de la **ESCUELA DE CIENCIAS QUÍMICAS E INGENIERÍA**, de la Carrera de **PETROQUÍMICA**, aprobada por el Consejo de Educación Superior (CES), mediante Resolución **RPC-SO-39-No.456-2014**, realiza a través de videoconferencia, la sustentación de su trabajo de titulación denominado: **PREDICTIVE MODELING OF THE PRIMARY SETTLING TANK IN THE IBARRA WASTEWATER TREATMENT PLANT BASED ON ARTIFICIAL NEURAL NETWORKS**, previa a la obtención del título de **PETROQUÍMICO/A**.

El citado trabajo de titulación, fue debidamente aprobado por el(los) docente(s):

Tutor	Dr. RICAURTE FERNÁNDEZ, MARVIN JOSÉ , Ph.D.
--------------	---

Y recibió las observaciones de los otros miembros del Tribunal Calificador, las mismas que han sido incorporadas por el(la) estudiante.

Previamente cumplidos los requisitos legales y reglamentarios, el trabajo de titulación fue sustentado por el(la) estudiante y examinado por los miembros del Tribunal Calificador. Escuchada la sustentación del trabajo de titulación a través de videoconferencia, que integró la exposición de el(la) estudiante sobre el contenido de la misma y las preguntas formuladas por los miembros del Tribunal, se califica la sustentación del trabajo de titulación con las siguientes calificaciones:

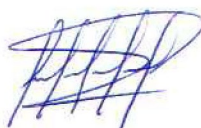
Tipo	Docente	Calificación
Tutor	Dr. RICAURTE FERNÁNDEZ, MARVIN JOSÉ , Ph.D.	10,0
Presidente Tribunal De Defensa	Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.	10,0
Miembro Tribunal De Defensa	Dra. HIDALGO BONILLA, SANDRA PATRICIA , Ph.D.	10,0

Lo que da un promedio de: **10 (Diez punto Cero)**, sobre 10 (diez), equivalente a: **APROBADO**

Para constancia de lo actuado, firman los miembros del Tribunal Calificador, el/la estudiante y el/la secretario ad-hoc.

Certifico que en cumplimiento del Decreto Ejecutivo 1017 de 16 de marzo de 2020, la defensa de trabajo de titulación (o examen de grado modalidad teórico práctica) se realizó vía virtual, por lo que las firmas de los miembros del Tribunal de Defensa de Grado, constan en forma digital.

VELOZ MARMOLEJO, CARLOS EDUARDO
Estudiante



Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.
Presidente Tribunal de Defensa

ALICIA ESTELA SOMMER MARQUEZ
 Firmado digitalmente por ALICIA ESTELA SOMMER MARQUEZ
 Fecha: 2020.11.05 18:57:08 -05'00'

Dr. RICAURTE FERNÁNDEZ, MARVIN JOSÉ , Ph.D.
Tutor

MARVIN JOSE
RICAURTE
FERNANDEZ

Firmado digitalmente por
MARVIN JOSE RICAURTE
FERNANDEZ
Fecha: 2020.11.05 20:45:05
-05'00'

Dra. HIDALGO BONILLA, SANDRA PATRICIA , Ph.D.
Miembro No Tutor

SANDRA PATRICIA
HIDALGO BONILLA

Firmado digitalmente por
SANDRA PATRICIA HIDALGO
BONILLA
Fecha: 2020.11.05 20:53:10
-05'00'

CIFUENTES TAFUR, EVELYN CAROLINA
Secretario Ad-hoc



Firmado electrónicamente por:
**EVELYN CAROLINA
CIFUENTES TAFUR**

AUTORÍA

Yo, **CARLOS EDUARDO VELOZ MARMOLEJO**, con cédula de identidad 0503748766, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así como, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autor(a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, Noviembre 2020



Carlos Eduardo Veloz Marmolejo

CI: 0503748766

AUTORIZACIÓN DE PUBLICACIÓN

Yo, **CARLOS EDUARDO VELOZ MARMOLEJO**, con cédula de identidad 0503748766, cedo a la Universidad de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, Noviembre 2020



Carlos Eduardo Veloz Marmolejo

CI: 0503748766

ACKNOWLEDGEMENTS

First, I express my gratitude to Yachay Tech University for bringing me the opportunity to obtain a high-quality education to improve my personal and professional formation. In the same way, I thank the School of Chemical Science and Engineering for receiving me in the best way and provide me an integral and multidisciplinary education.

I express all my gratitude to my family, especially to my mother and father, for the intense and constant effort that they did to support me during my professional formation. Thank you for all the sacrifices that you have made for me and your teachings; you are the most important part of my life.

Also, I express my gratitude to my professor Marvin Ricaurte for being an important guide in my professional formation such as in the development of this work. Thank you for the dedicated time for training my classmates and me with your knowledge and experience.

I thank professor Carlos Pazmiño for assisting in the development of this work. Thank you for always being available to discuss, share ideas, and help me solve all the problems and questions that I had during this period.

I would like to thank professor José Ángel Rivera for motivating me to develop my internship experience in the Ibarra Wastewater Treatment Plant and supporting me in every moment of my academic experience.

I want to express my acknowledgments to all my professors for dedicating their time and effort to my personal and professional development. Thank you for all your teachings and hard work that you daily made to give us the best education.

I also like to express my gratefulness to the Rivadeneira Aguirre family for receiving me in their home during the internship experience period. Thank you for all your kindness and hospitality.

I would like to thank all my friends and classmates for being a fundamental part of my entire academic life. Daniela, who was always present to help and motivate me in the

development of this work. Michael for sharing part of his laboratory data analysis during the internship experience to help me with this work. Also, Luis for supporting me in many important moments.

Finally, I thank the authorities and staff of EMAPA-Ibarra for giving me the opportunity to perform my internship experience in the Ibarra Wastewater Treatment Plant and providing me the necessary information for the development of the present work.

DEDICATION

I dedicate my dissertation thesis to my beloved family, who have supported me all my life to achieve my goals. Especially to my parents, Mariela and Tobías, whose words of encouragement and thrust are a source of inspiration to me. My brother, David, who has always impulse me to become a better person. My grandmothers, Rosa and María Beatriz, who are my motivation for doing my best effort in every aspect of my life. My grandfather, Tobías, who always supports me unconditionally in my entire academic life.

Finally, I would also want to dedicate this work to all my professors and friends who have become part of my academic formation at Yachay Tech University.

RESUMEN

Las aguas residuales son la combinación de desechos líquidos y desechos transportados en el agua debido a la contaminación producida por la actividad antropogénica. Por lo tanto, es necesario el respectivo proceso de tratamiento para reducir los niveles de contaminación del agua residual previo a su descarga al medio ambiente. La planta de tratamiento de aguas residuales de la ciudad de Ibarra (PTAR-I) procesa un caudal promedio de 43,200 m³/día proveniente de la comunidad, y el efluente tratado es descargado en el río Tahuando. El proceso de tratamiento consiste en un proceso biológico en donde la etapa de sedimentación durante el tratamiento primario es parte fundamental del proceso. El decantador primario es usado para la remoción de sólidos suspendidos y así reducir la carga orgánica presente en el agua residual. Por lo tanto, el modelaje y monitoreo del funcionamiento del decantador primario es necesario para obtener un control efectivo del proceso. Los modelos usados son basados principalmente en relaciones empíricas derivadas de la observación diaria. Sin embargo, estos modelos empíricos frecuentemente suelen presentar varias limitaciones para su aplicación. Por lo tanto, en los últimos años, se han desarrollado diferentes herramientas computacionales orientadas a obtener mejores modelos. Las redes neuronales artificiales (RNA) son una popular herramienta computacional inspiradas en el funcionamiento del cerebro. Este estudio tiene como objetivo desarrollar un modelo predictivo en base a redes neuronales artificiales para modelar el funcionamiento del decantador primario. Dos redes neuronales artificiales separadas fueron elaboradas utilizando información sobre las características del agua residual y condiciones de operación del decantador primario. Una de las redes fue destinada a la predicción de la concentración de sólidos suspendidos totales (SST) en el efluente del sedimentador, y la segunda para la predicción de la demanda química de oxígeno (DQO). La metodología consistió primero en realizar un análisis del proceso de sedimentación para identificar las variables más representativas. Posteriormente, el funcionamiento del modelo propuesto fue comparado con los modelos empíricos tradicionales reportados en la literatura. Finalmente, el modelo propuesto fue validado utilizando información actual proporcionada por la PTAR-I.

Palabras Claves: aguas residuales, decantador primario, análisis de procesos, modelaje de procesos, redes neuronales.

ABSTRACT

Wastewater is the combination of liquid waste and water-carried wastes due to the pollution produced by anthropogenic activity. Therefore, the adequate treatment of wastewater is required to reduce the level of pollution before discharge to the environment. The Ibarra wastewater treatment plant (WWTP) processes an average flow of 43,200 m³/day from the community, and the treated effluent is discharged to the Tahuando river. The treatment process consists of a biological process where the sedimentation stage during the primary treatment is fundamental in the entire process. The primary settling tank is used to remove the suspended solids, and, as a result, to reduce the organic load of wastewater. Hence, the modeling and monitoring of primary settling tank performance are necessary to obtain effective process control. The models are mainly based on empirical relationships that are derived from daily observation. However, these empirical models frequently present some limitations to the application. Therefore, in recent years, several computational tools have been developed in order to obtain better models. Artificial neural networks (ANN) are popular computational tools inspired by the performance of the brain. The objective of this study is to develop a predictive model based on ANN for modeling the primary settling tank (PTS) of the Ibarra WWTP. Two separated artificial neural networks were built using the information about wastewater characteristics and the operational conditions of the primary settling tank. The first of the networks are used to predict the total suspended solids (TSS) concentration of the effluent from the settling tank, and the second one makes predictions about the chemical oxygen demand (COD). The methodology consisted of first realize a process analysis of the sedimentation process to identify the most representative variables. Posteriorly, the performance of the proposed model was compared to traditional empirical models reported in the literature. Finally, the model was validated using current information provided by the Ibarra WWTP.

Keywords: wastewater, primary settling tank, process analysis, process modeling, neural networks.

TABLE OF CONTENT

LIST OF TABLES.....	x
LIST OF FIGURES	xi
ABBREVIATIONS	xiii
CHAPTER I.....	1
1. INTRODUCTION	1
1.1. Problem Approach.....	3
1.2. Objectives.....	4
1.2.1. General Objective	4
1.2.2. Specific Objectives	4
CHAPTER II	5
2. BACKGROUND AND LITERATURE REVIEW	5
2.1. Sources of Wastewater	5
2.2. Wastewater Composition	5
2.2.1. Physical Parameters.....	7
2.2.1.1. Solids.....	7
2.2.1.2. Turbidity.....	9
2.2.1.3. Temperature	10
2.2.1.4. Inlet Flow	10
2.2.2. Chemical Parameters	11
2.2.2.1. Biochemical Oxygen Demand (BOD)	11
2.2.2.2. Chemical Oxygen Demand (COD)	11
2.2.2.3. Electrical Conductivity.....	12
2.2.2.4. Nitrogen Content.....	12
2.2.3. Biological Parameters	13
2.2.3.1. Indicator Organisms	13
2.3. Wastewater Treatment Fundamentals	14
2.3.1. Preliminary Treatment	15
2.3.1.1. Screening.....	15
2.3.1.2. Grit Removal.....	15
2.3.2. Primary Treatment	16
2.3.3. Secondary Treatment	19

2.4.	Dynamic Modeling.....	20
2.5.	Fundamentals of Artificial Neural Networks.....	22
2.5.1.	Components of Artificial Neural Networks.....	23
2.5.2.	Artificial Neural Networks Topology.....	25
2.5.3.	Training of Artificial Neural Networks.....	26
2.5.3.1.	Overfitting.....	28
2.5.4.	Evaluation of Artificial Neural Networks.....	30
CHAPTER III.....		32
3.	IBARRA WASTEWATER TREATMENT PLANT.....	32
3.1.	Process Description.....	32
3.1.1.	Inlet Structure.....	34
3.1.2.	Pretreatment.....	35
3.1.2.1.	Screening.....	35
3.1.2.2.	Grit Removal-Degreasing.....	36
3.1.3.	Primary Treatment.....	37
3.1.3.1.	Primary Settling Tank (PST).....	38
3.1.4.	Biological Treatment.....	40
3.1.5.	Secondary Sedimentation.....	41
3.1.6.	Sludge Treatment.....	41
3.1.6.1.	Sludge thickening.....	41
3.1.6.2.	Anaerobic digestion.....	42
3.1.6.3.	Dewatering and drying.....	42
CHAPTER IV.....		44
4.	METHODOLOGY.....	44
4.1.	Process Analysis.....	46
4.2.	Data Processing for Artificial Neural Network Models.....	48
4.3.	Structure of Artificial Neural Network Model.....	49
4.4.	Artificial Neural Network Model Training.....	50
4.5.	Sensitivity Analysis.....	51
4.6.	Results and Discussion Section: Outline.....	53
CHAPTER V.....		56
5.	RESULTS AND DISCUSSION.....	56
5.1.	Data Processing for Artificial Neural Network Models.....	56
5.2.	Artificial Neural Network Models Training.....	59

5.3. Performance of Selected Models.....	64
5.4. Sensitivity Analysis.....	73
5.4.1. Local Sensitivity Analysis	73
5.4.2. Global Sensitivity Analysis.....	78
5.5. Model Validation.....	81
5.6. Applications	82
CONCLUSIONS AND RECOMMENDATIONS	84
REFERENCES	86
APPENDIX A:	94
APPENDIX B:.....	97
APPENDIX C:.....	101

LIST OF TABLES

Table 1. Typical characteristics of untreated municipal wastewater.	6
Table 2. Design parameters for primary settling tanks.	17
Table 3. Design flows of the Ibarra WWTP.	34
Table 4. Design wastewater parameters of the Ibarra WWTP.....	35
Table 5. Purge equipment to sludge extraction.....	40
Table 6. Wastewater properties that are measured during primary treatment.	46
Table 7. Results and Discussion Section: Outline.	55
Table 8. Descriptive statistics of experimental data used for the TSS model.....	58
Table 9. Descriptive statistics of experimental data used for the COD model.	58
Table 10. R ² values for the TSS model with different number of hidden nodes.	60
Table 11. R ² values for the COD model with different number of hidden nodes.....	61
Table 12. MSE value for model performance with different hidden nodes.....	63
Table 13. Distribution of error for the TSS model versus the empirical TSS model. ...	68
Table 14. Distribution of error for the COD model versus the empirical COD model. 70	
Table 15. Statistical description of error distribution produced by TSS models.	71
Table 16. Statistical description of error distribution produced by COD models.	72
Table 17. Influence of input parameters to neural network models.	78
Table 18. Weights and bias of the input layer of the TSS model	99
Table 19. Weights and bias of the output layer of the TSS model	99
Table 20. Weights and bias of the input layer of the COD model.....	100
Table 21. Weights and bias of the output layer of the COD model.....	100
Table 22. Coefficients of empirical models.	102

LIST OF FIGURES

Figure 1. Interrelationships of solids found in wastewater.....	7
Figure 2. A node that receives three inputs and bias.	23
Figure 3. Feed-forward multilayer neural network.....	26
Figure 4. Schematic representation of supervised learning algorithm.	27
Figure 5. Representation of early stopping method.....	30
Figure 6. Block Process Diagram of Ibarra WWTP.....	33
Figure 7. Primary settling tank (PST) cross-section view.....	38
Figure 8. Methodology diagram.	45
Figure 9. Multilayer-neural network architecture used for the TSS model and COD model.....	50
Figure 10. Correlation functions between wastewater properties.	56
Figure 11. R^2 values for the TSS model with different number of hidden nodes.....	60
Figure 12. R^2 values for the COD model with different number of hidden nodes.	62
Figure 13. Variation of MSE value for model performance with different hidden nodes.	63
Figure 14. Performance obtained by the COD model for the total data set.....	65
Figure 15. Performance obtained by the TSS model for the total data set.	66
Figure 16. Correlation between experimental data and model predictions.	67
Figure 17. Average error of neural network models against empirical models.....	67
Figure 18. Distribution of error for the TSS model versus the empirical TSS model...	68
Figure 19. Comparison of predicted values of effluent TSS versus experimental data for different models.....	69
Figure 20. Distribution of error for the COD model versus the empirical COD model.	70
Figure 21. Comparison of predicted values of effluent COD versus experimental data for different models.....	71
Figure 22. Normal distribution of TSS model errors and empirical TSS model errors.	72
Figure 23. Normal distribution of COD model errors and empirical COD model errors.	73
Figure 24. Local SA analysis of the TSS model.....	74
Figure 25. Local SA analysis of the COD model.	76
Figure 26. Global SA of the TSS model.....	79
Figure 27. Global SA of the COD model.	80

Figure 28. Correlation between model predictions and new experimental data.....	81
Figure 29. Schematic representation of notation used in ANN models.	98
Figure 30. Correlation between experimental data and empirical model predictions .	102

ABBREVIATIONS

ANN	Artificial Neural Networks
BC	Base Case
CFD	Computational Fluid Dynamics
COD	Chemical Oxygen Demand
FC	Fecal Coliform
FS	Fixed Solids
LMA	Levenberg-Marquardt Algorithm
MLP	Multilayer Perceptron
MSE	Mean Square Error
MPN	Most Probable Number
NTU	Nephelometric Turbidity Units
PST	Primary Settling Tank
R^2	Coefficient of Determination
SA	Sensible Analysis
TC	Total Coliform
TDS	Total Dissolved Solids
TKN	Total Kjeldahl Nitrogen
TS	Total Solids
TSS	Total Suspended Solids
VS	Volatile Solid
WWTP	Wastewater Treatment Plant

CHAPTER I

1. INTRODUCTION

Wastewater is the combination of the liquid wastes and wastes transported in water from residences, commercial establishments, and industries, as well as stormwater and other surface runoff ^[1]. Wastewater contains high concentrations of organic and inorganic pollutants, suspended solids, pathogenic microorganisms, as well as toxic compounds. Therefore, the disposal of wastewater without treatment in the environment has an important negative consequence on human health and sustainable development ^[2].

In the last few years, population growth and industrialization have increased the degradation of the environment. According to the fourth World Water Development Report (2012), only 20% of globally produced wastewater receives proper treatment ^[3]. Therefore, the engineering of wastewater treatment has been focused in recent years to design and construct adequate treatment processes. As a general idea, every wastewater treatment process has the objective to reduce the levels of pollutant concentrations in the wastewater to acceptable levels before it discharges in the final disposal.

The wastewater treatment consists of a combination of unit operations and unit processes designed to reduce the contaminants in the water to an acceptable level ^[4]. The term unit operation refers to physical treatment operations, while the unit process corresponds to biological or chemical treatment methods. Hence, the organization of these elements produces several levels of treatments classified into: (i) preliminary treatment, (ii) primary treatment, (iii) secondary treatment, and (iv) tertiary treatment.

In Ecuador, there are approximately 421 wastewater treatment plants (WWTP), and around 62% of the municipalities provide proper treatment to wastewater before be discharged to rivers ^[5]. In particular, the Ibarra WWTP processes an average wastewater flow of 43,200 m³/day from the community and industrial activities, and the treated effluent is finally discharged to the Tahuando river. The treatment process consists of a

biological treatment method as a secondary treatment to reduce pollutants concentration in the wastewater.

The reduction of suspended solids is a fundamental part of the entire biological process of wastewater treatment. Primary treatment is focused on the physical removal of suspended solids fraction from the wastewater composition ^[6]. Hence, sedimentation is one of the principal physical operations used in the primary treatment of wastewater. Primary settling tanks are widely operated to remove suspended solids and reduce the organic load of wastewater and are considered a fundamental part of the biological process ^[7].

Then, modeling of primary settling tanks is an important task in WWTP design and process control. The modeling of these operational units is mainly based on empirical relationships that are derived from daily observation ^[8]. However, these empiric models usually fail to model full-scale settling tanks due to the complexity of the sedimentation mechanism. Furthermore, the laboratory test also does not provide satisfactory results under operating conditions. Consequently, in recent years, many attempts for modeling the settling tank performance have been developed using different computational tools. Artificial neural networks (ANN) are a popular machine learning technique that is based on how the brain works. They are powerful non-linear regression models that can establish complicated relationships between variables through the examination of a data set ^[9].

Therefore, the objective of this study is to develop a predictive model based on ANN for modeling the primary settling tank (PST) of the Ibarra WWTP. The model consists of two separate artificial neural networks. The first one uses influent wastewater characteristics and operational conditions as input data to predict effluent total suspended solids (TSS) concentration, and the other one makes predictions of the effluent chemical oxygen demand (COD) using the same input data. Ibarra WWTP supplied the available data of the operational conditions and wastewater characteristics related to the primary treatment for the development of the model.

The methodology of this study consisted first of developing an integral process analysis to establish the critical variables associated with the PST operation. Second, the ANN models were built using the available data. Then, the performance of ANN models was

compared to the traditional empiric models reported in the literature. Moreover, a sensitivity analysis was considered to identify the relationship between input variables and model response. Finally, a new data set was provided by Ibarra WWTP to evaluate the performance and accuracy of ANN models developed in this study.

1.1. Problem Approach

Wastewater treatment plants frequently have a satisfactory performance under steady-state conditions because these conditions are similar to design parameters. However, load variations are usually during the routine operation of these plants. The inflow to the treatment processes is a random variable that necessarily influences the integral performance of the treatment process ^[3]. Hence, the development and availability of process models for the design and control of treatment operation is a challenging task to improve the process control system.

Primary settling tanks are one of the principals controlling equipment in the performance and removal efficiency of the treatment process, especially in biological treatments ^[1]. Therefore, monitoring the dynamic response of primary settling tanks is an important aspect to improve the process control system of the entire treatment process. However, empiric models of the sedimentation process and laboratory settling tests have failed to predict the behavior of settling tanks under actual operating conditions due to the difficulties in simulating the effects of flocculation and density currents ^[2]. Consequently, the design and modeling of settling tanks are still based on empirical relationships that are derived from the daily observation and accumulated operational experience ^[8]. However, these empiric models have several limitations to modeling full-scale settling tanks during real operation. Hence, the development of models to describe the dynamic response of settling tanks is an important and difficult task.

In recent years, many attempts for modeling the settling tank performance have been developed using computational tools that are focused on describing the fluid dynamics of the settling tanks. Models based on computational fluid dynamics (CFD) have been used to predict flow patterns and suspended solids distribution within the settling tanks ^[7]. However, these models are not commonly used due to the complexity of Navier-Stokes equations for turbulent flow and the cost associated with the specialized hardware and software required ^[10].

On the other hand, other types of computational modeling methodologies have been developed. Artificial neural networks are a popular machine learning technique that simulates some important features of the brain. They are able to establish complicated relationships through an examination of only the data set without assuming theoretical considerations or pre-specified formula ^[9]. Therefore, many models based on artificial neural networks have been successfully applied in many areas of engineering. In this context, this study has the main objective to develop a predictive model based on artificial neural networks for modeling the dynamic response of the primary settling tank of Ibarra WWTP.

1.2. Objectives

1.2.1. General Objective

- To develop a predictive model based on artificial neural networks for modeling the primary settling tank behavior, in order to determine the chemical oxygen demand and total suspended solids concentration in the clarified effluent.

1.2.2. Specific Objectives

- To carry out a process analysis to identify the critical variables associated with the primary settling tank operation.
- To establish the optimum characteristics of artificial neural network models that provide the best performance for predicting chemical oxygen demand and total suspended solids concentration of the effluent from the primary settling tank.
- To compare the performance of artificial neural network models against the traditional empiric regression model based on the error prediction of each model.
- To analyze the effect of input variables in the artificial neural network model response, employing a sensitivity analysis.

CHAPTER II

2. BACKGROUND AND LITERATURE REVIEW

2.1. Sources of Wastewater

Wastewater is the water supply of the community after it has been used in a variety of applications. Thus, wastewater is defined as a combination of liquid wastes and waste transported in water from households, commercial establishments, and industries, as well as stormwater and other surface runoff ^[1]. The common sources or types of wastewater are:

- *Domestic or municipal wastewater*: this includes wastewater produced in residences, institutions as schools, restaurants, etc.
- *Industrial wastewater*: wastewater discharged from industrial processes.
- *Stormwater*: rainfall-runoff and snowmelt.

2.2. Wastewater Composition

Wastewater composition is approximately 99% water and 1% suspended, colloidal, and dissolved solids ^[3]. Municipal sewage also contains organic matter, nutrients (N, P, and K), inorganic matter, dissolved compounds, toxic substances, and pathogen microorganisms. Therefore, these pollutants in water need to be treated before the wastewater is discharged into nature (water bodies).

The composition of sewage depends on several factors as the socio-economic characteristics of residential communities, number, and types of industrial activities ^[3]. Consequently, the design of a wastewater treatment process partially depends on the composition of wastewater. Usually, in the design of WWTP, there is no interest in

determining all the compounds present in the raw sewage. It is due to the difficulty of the laboratory test, and the fact that some results themselves cannot be directly applied as elements in design and operation ^[11].

Hence, for the WWTP design and operation, many indirect parameters are used, which represent the pollution potential of raw sewage. Wastewater is characterized in terms of physical, chemical, and biological parameters in order to define its composition and the nature of contaminants. The typical wastewater composition is detailed in Table 1.

Table 1. Typical characteristics of untreated municipal wastewater.

Component	Concentration Range
Biochemical oxygen demand (BOD ₅)	100 – 360 mg/L
Chemical oxygen demand (COD)	250 – 100 mg/L
Total organic carbon (TOC)	80 – 300 mg/L
Total Kjeldahl nitrogen (TKN)	20 – 85 mg/L as N
Total phosphorus	5 – 15 mg/L as P
Oil and grease	50 – 120 mg/L
Total solids (TS)	400 – 1200 mg/L
Total dissolved solids (TDS)	250 – 850 mg/L
Total suspended solids (TSS)	110 – 400 mg/L
Volatile suspended solids (VSS)	90 – 320 mg/L
Fixed suspended solids (FSS)	20 – 80 mg/L
Settleable solids	5 – 20 ml/L
Total coliforms (TC)	10 ⁶ – 10 ¹⁰ MPN/100ml
Fecal coliforms (FC)	10 ³ – 10 ⁸ MPN/100ml

Source: Adapted from Riffat ^[4].

2.2.1. Physical Parameters

The most significant physical property of wastewater is total solids (TS) content, which is composed of settleable matter, floating material, colloidal matter, and matter dissolved in solution ^[1]. However, other important physical characteristics include temperature, turbidity, conductivity, density, color, etc.

2.2.1.1. Solids

All pollutants of raw sewage, except dissolved gases, contribute to solid concentration ^[12]. Wastewater contains several solids materials varying from coarse solids as rags to colloidal particles. In order to characterize the raw sewage, coarse solids usually are removed before the sample is analyzed for solids. The concentration of solids in wastewater is commonly measured in mg/L, but in the case of the settleability test is measured in ml/L. The relationships between the several types of solid fractions in wastewater are illustrated in Figure 1.

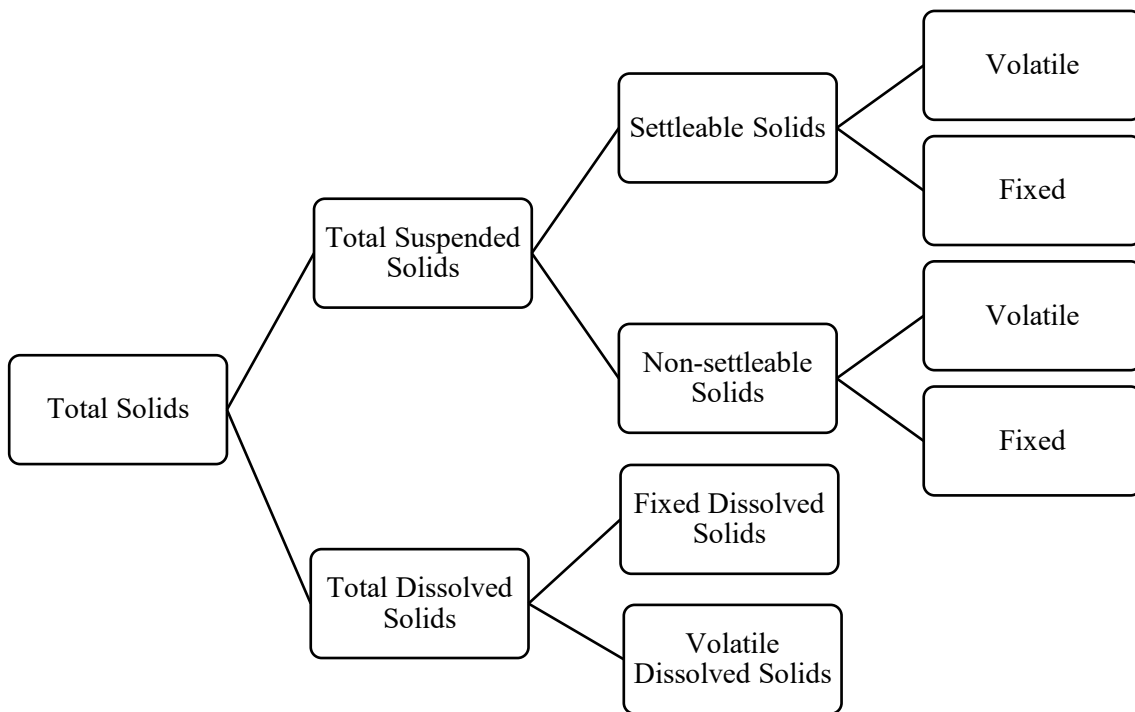


Figure 1. Interrelationships of solids found in wastewater.

Solids particles can be classified by their size and state, their chemical characteristics, and their settleability, as following ^[11]:

i. Solid Classification by Size and State

- *Total Solids (TS)*: substances in an aqueous solution that can exist in either the dissolved or undissolved state. Total solids are the residues remaining from the water sample dried at 103-105 °C for a specific time (one hour or overnight)^[13]. Thus, TS correspond to everything that was in the sample that is not water. However, any organic or inorganic substance that volatilized at 103 °C, or less, will not be considered in the residue.
- *Total Suspended Solids (TSS)*: fraction of the TS retained on a filter with a specified pore size (0.45-2.0 µm), measured after being dried in an oven for at least one hour at a specified temperature (103-105 °C)^[14]. Thus, the result of the TSS test depends partially on the pore size of the filter paper used for the test. The TSS test result is an important parameter used to evaluate the performance of the treatment process and measure the quality of the treated effluent. For this reason, TSS is one of the two universally used effluent parameters by which the performance of the treatment process is judged for regulatory control purposes^[1].
- *Total Dissolved Solids (TDS)*: fraction of TS in a water sample that passes through a filter with a nominal pore size of 2.0 µm (or smaller)^[15]. The filtrate is evaporated for at least one hour in an oven at a specified temperature (180 ± 2 °C)^[14]. Wastewater contains a high fraction of colloidal solids, which contribute to the TDS concentration. The size of colloidal particles is less than 1.0 µm, so they will not settle even if allowed to sit quietly for days or weeks^[12]. The number of colloidal particles in untreated sewage and after primary sedimentation is typically within the range 10⁸-10¹²/ml^[1]. Consequently, colloidal material and truly dissolved material are considered in the TDS test result.

ii. Solid Classification by Chemical Characteristics

- *Volatile Solids (VS)*: matter that can be volatilized when ignited at a specific temperature (500 ± 50 °C) is classified as volatile^[1]. In general, VS is associated with organic matter present in the wastewater sample, which can be oxidized to CO₂ at high temperatures^[16].

- *Fixed Solids (FS)*: solid fraction that remains after a sample has been ignited is referred to as fixed solids, and it is considered the inorganic portion of the wastewater sample^[11]. Thus, this inorganic fraction will not be oxidized and volatilized during the ignition process.

Hence, VS concentration represents an estimation of organic matter in the solids, while FS represents the inorganic fraction of solids^[16]. As a result, TS, TSS, and TDS include both fixed and volatile solids. Nevertheless, the determination of VS and FS concentrations do not distinguish precisely between organic and inorganic matter. Some inorganic mineral salts will be lost during the ignition, and few organic substances will not burn. The ratio of VS to FS is useful in controlling wastewater treatment plant operations because it provides a rough estimation of the amount of organic and inorganic matter present in sewage^[1].

iii. Solid Classification by Settleability

- *Settleable Solids*: they are non-dissolved solids that can settle to the bottom of a container under the influence of gravity in 1 hour^[16]. The standard test for settleable solids consists of filling a 1-liter Imhoff cone to the full mark with a well-mixed sewage sample. After 45 minutes, the sample is gently agitated near the sides of the Imhoff cone with a rod, continue to settle for an additional 15 minutes and measure the volume of solids that settled at the bottom of the cone as ml/L^[14]. In general, about 60% of the suspended solids in municipal sewage are settleable solids^[1].
- *Non-settleable Solids*: fraction that not settle during the specified time, and it is usually not expressed in the result of the analysis^[11].

The equations used to determine the several fractions of solids in wastewater are detailed in Appendix A.

2.2.1.2. Turbidity

As a general idea, turbidity refers to the clarity of the water, and this is one of the first characteristics that people notice. Turbidity is a physical measure of the extent to which

light is scattered or absorbed by suspended material in water ^[12]. The measurement is based on a comparison of the intensity of light scattered by a sample to the light scattered by a reference suspension at the same conditions ^[1]. Colloidal material in the sample will absorb or scatter light and block its transmission. Turbidity measurements usually are reported as nephelometric turbidity units (NTU). However, the presence of bubble air in sewage will cause errors in turbidity readings. A turbidimeter is an equipment used to measure the turbidity of wastewater and provide useful data for the process control. Turbidimeter is an electronic device in which a beam of light is directed through a specific path length of the sample. The photometer placed at right angles to the direction travel of the light beam detects the amount of light diverted, which is proportional to the turbidity ^[16].

The larger amount of TSS concentration in water, the murkier it appears, and higher measured turbidity ^[17]. However, dissolved solids do not cause turbidity in the sample, so the water of low turbidity is not necessarily without dissolved solids. High turbidity will cause operational problems as the reduction of the effectiveness of the disinfection process using UV radiation ^[18]. Furthermore, the components which cause turbidity will generate taste and odor problems.

2.2.1.3. Temperature

It is an important parameter in the wastewater treatment process because it affects the chemical reactions, the viscosity of the liquid, and the microbial activity ^[19]. The temperature of wastewater is commonly higher than that of the local water supply due to the addition of warm water from households and industrial activities ^[1]. Additionally, the temperature can vary depending on the geographical location and season of the year.

2.2.1.4. Inlet Flow

The actual volume of wastewater is used as a physical characterization of wastewater, and it is usually expressed in terms of cubic meters per day (m^3/day) ^[12]. It is an essential parameter in the design of treatment plants. Flow rates will vary throughout the day. This variation can be a wide difference from the average daily flow.

2.2.2. Chemical Parameters

2.2.2.1. Biochemical Oxygen Demand (BOD)

Amount of oxygen utilized by a mixed population of microorganisms to oxidize organic matter present in sewage, through aerobic biochemical processes, at a controlled temperature of 20 °C for a specified time ^[4]. Therefore, this is an indirect measurement of biodegradable organic carbon. Three classes of materials exert the oxygen demand of sewage: (1) carbonaceous organic materials used as a source of food by aerobic organisms; (2) oxidizable nitrogen derived from ammonia, nitrites, and organic nitrogen compounds, which serve as nutriment for specific bacteria (*Nitrosomonas* and *Nitrobacter*); (3) chemical reducing compounds like sulfites, ferrous ions, and sulfides ^[6].

The BOD value is time-dependent, and it would take an infinitely long time for the microorganism to degrade all the organic matter in the sample. The oxidation of about 95% of the carbonaceous organic matter is achieved at 20 days. However, in the wastewater industry, the BOD₅ is used as a standard value that is obtained from a BOD test conducted for five days. About 60% to 70% of carbonaceous matter is oxidized after five days ^[4].

The main ecological effect of organic pollution in a water body is the reduction in the level of dissolved oxygen. Hence, the adequate quantity of oxygen is fundamental in the aerobic treatment process. This parameter is a basic idea about the "strength" of the pollution potential of sewage by the measurement of the oxygen consumption that it could cause ^[11]. Therefore, this is an indirect quantification of the potential impact that the pollutant concentration will produce.

2.2.2.2. Chemical Oxygen Demand (COD)

Amount of oxygen required to oxidize the fraction of the sample, which is susceptible to dichromate or permanganate oxidation in an acid solution ^[6]. Hence, this value is an indirect measurement of the level of organic matter in the sewage. The main difference between the COD test and BOD₅ test is that the BOD₅ test relates to the biochemical oxidation of the organic matter, undertaken entirely by microorganisms.

The main advantages of the COD test are:

1. The time that the test required. The BOD test needs at least five days to provide a result. On the other hand, the COD test takes only two to three hours. Thus, the COD test is commonly used for operational control in treatment processes.
2. The test results provide an indication of the oxygen required to stabilize the organic matter. However, both biodegradable and the inert fractions of organic matter are oxidized during the test. Therefore, the test may overestimate the oxygen required to oxidized organic matter.
3. The test is not affected by nitrification.

For raw domestic sewage, the ratio COD/BOD₅ varies between 1.7 and 2.4 ^[11]. However, for industrial wastewater, this ratio can vary widely. Depending on the value of the ratio, it can infer the biodegradability of the wastewater and select the appropriate treatment process. If COD is much greater than BOD₅ in raw sewage (COD/BOD₅ greater than 3.5 or 4.0), the pollutants are not readily biodegradable, and it may be toxic to the microorganisms ^[22]. If the COD/BOD₅ ratio is lower than 3.0, the waste in raw sewage is readily biodegradable and is a good indication for biological treatment.

2.2.2.3. Electrical Conductivity

In the wastewater treatment process, conductivity is a measurement of the ability of an aqueous solution to carry an electric current ^[15]. This ability highly depends on the number of dissolved salts (ions) in the aqueous solution because ions transport the electrical current ^[20]. Thus, conductivity can be used to calculate the amount of TDS in wastewater and evaluate its variation. In order to determine the TDS value, the conductivity value is multiplied by an empirical factor that can range from 0.55 to 0.95, depending on the temperature and the soluble substances of the solution ^[21]. The electrical conductivity is expressed as millisiemens per meter (mS/m) or in micromhos per centimeter (μmho/cm).

2.2.2.4. Nitrogen Content

The type and amount of nitrogen compounds present will vary from the raw wastewater to the treated effluent ^[12]. Total nitrogen includes organic nitrogen, ammonia, nitrites, and

nitrates. It is an essential nutrient for microorganism growth in biological wastewater treatment. Most of the nitrogen content in untreated wastewater will be in the forms of organic nitrogen and ammonia ^[11]. The sum of these two forms of nitrogen compounds is measured and is known as total Kjeldahl nitrogen (TKN).

2.2.3. Biological Parameters

2.2.3.1. Indicator Organisms

An important aspect in terms of the biological quality of wastewater is related to disease transmission by pathogenic organisms. Water-borne diseases include typhoid, cholera, paratyphoid, diarrhea, fever, and dysentery. Pathogenic organisms found in wastewater may be excreted by human beings and animals who are infected with these diseases or carry a particular infectious disease ^[1].

The pathogenic organism found in wastewater can be classified into four broad categories: bacterias, protozoa, helminths, and viruses ^[11]. However, the detection of pathogenic microorganisms in a wastewater sample is a hard task due to their low concentration and because they are difficult to isolate and identify ^[14]. Therefore, some microorganisms which are more numerous and more easily analyzed are commonly used as an indicator organism for the target pathogen(s) ^[1]. These organisms are predominantly non-pathogenic, but they give a satisfactory indication of whether the water is contaminated by human or animal feces, and, therefore, of its potential to transmit diseases ^[11]. Bacterias of the coliform group are the organisms most commonly used as an indicator. Frequently, total coliform and fecal coliform are used as indicator organisms ^[14].

- *Total coliform*: coliform organisms are rod-shaped bacteria, which are present in the human intestinal tract ^[6]. Each person discharges from 100 to 400 billion coliform bacteria per day, in addition to other kinds of bacteria ^[1]. Total coliform has long been used as an indication that pathogen organisms may also be present. All members of the total coliform groups can occur in human feces, however, some coliform organisms can also be present in animal manure, soil, and submerged wood, as well as in other places outside the human body ^[12]. Therefore,

some coliform organisms are related to other free-living organisms such as bacterias related to the genus *Klebsiella*, and not only to the intestinal or pathogenic ones ^[11].

- *Fecal coliform*: it is a subset of total coliform bacteria. Fecal coliforms are more fecal-specific. This group of coliform bacteria is predominantly originated from the intestinal tract of humans and other animals ^[11]. Therefore, fecal coliform has been used as a standard indicator organism of pathogenic contamination in water ^[14]. Nevertheless, even this group contains some organisms that are not necessarily fecal in origin.

The determination of indicator organisms is a fundamental task for assessing the quality of natural water, drinking water, and wastewater. The multiple-tube fermentation technique for total and fecal coliform testing is useful in determining the coliform density in wastewater ^[12]. The technique is based on the most probable number (MPN) of bacterias present in the sample that produces gas in a series of fermentation tubes with various volumes of diluted samples ^[6]. The MPN is determined by the application of the Poisson distribution for extreme values to the analysis of the number of positive and negative results obtained from the multiple-tube fermentation test ^[1]. The MPN is obtained from charts based on statistical studies of known concentrations of bacteria. The concentrations of coliform bacteria are usually reported in MPN/100mL. It is also important to mention that the MPN is not the absolute concentration of organisms that are present in the water, but it is only a statistical estimation of the concentration.

2.3. Wastewater Treatment Fundamentals

The main objective of wastewater treatment is to reduce the pollutant concentration to acceptable levels before discharge to streams and rivers ^[11]. Therefore, wastewater treatment processes are designed to reduce the amount of solids, the level of biodegradable organic matter, the level of pathogens and nutrients, and the presence of toxic compounds in wastewater ^[4]. Usually, the wastewater treatment process is divided into preliminary, primary, secondary, and tertiary treatment. Nevertheless, in developing countries, tertiary treatment is unusual in wastewater treatment plants because it is mainly

used for the removal of specific pollutants (usually toxic or non-biodegradable compounds), which cannot be eliminated during secondary treatment ^[19].

2.3.1. Preliminary Treatment

Preliminary treatment involves the removal of larger suspended solids and inert materials from the sewage using physical unit operations commonly. The objectives of pretreatment are: (1) remove coarse material from the flow stream that may cause damage to subsequent process equipment, and (2) removal inert material before secondary treatment ^[1]. The unit operations used in this part are screen, grit chamber, and comminutors.

2.3.1.1. Screening

Sewage contains a significant amount of floating and suspended materials, which includes rags, bags, organic matter, and several types of solids. Large solids can damage the mechanical equipment and interfere with the flow in pipes or channels ^[23]. Therefore, the first unit operation usually encountered in this part of the process is screening. A screen is a device with openings, generally, of uniform size, that is used to retain solids found in the influent wastewater. Different types are available depending on wastewater characteristics and site requirements ^[6]. Two general types of screens are commonly used in the wastewater treatment process: coarse screen and fine screen. Both of these screens are mostly used in the preliminary treatment of sewage. Coarse screens have a size of the opening between 25 and 75 mm (1 to 3 in) and, fine screens correspond to openings less than 6 mm (0.25 in) ^[4].

2.3.1.2. Grit Removal

Grit is defined as sand, gravel, or other mineral material that has a nominal diameter of 0.15-0.20 mm or larger ^[24]. Grit chambers are settling tanks that are placed after screens and before primary clarifiers. The purpose of a grit chamber is to remove materials that may produce heavy deposits in pipelines, protect pumps and other mechanical equipment from abrasion, and separate heavier inert solids from lighter biodegradable organic solids that are sent to secondary biological treatment ^[4]. The sand removal mechanism is simply by sedimentation: the sand grains go to the bottom of the tank due to their larger particle size and density, while the organic matter, which settles much slower, remains in

suspension and goes on to the downstream units ^[11]. Grit removed is usually washed to remove organic matter and then transported to a sanitary landfill for disposal. The amount of grit material removed depends on the wastewater characteristics and the design of the grit chamber established in the treatment process. The design of grit chambers is usually based on the removal of grit particles having a specific gravity of 2.65 and a wastewater temperature of 15.5 °C (60°F) ^[1]. However, the values of the specific gravity of grit particles can oscillate in the range from 1.3 to 2.7.

2.3.2. Primary Treatment

The objective of primary treatment is to remove a significant fraction of settleable suspended solids and floating material from the wastewater by sedimentation process (primary sedimentation or clarification) ^[4]. A significant part of TSS is composed of organic matter in suspension. Usually, the primary treatment step can be expected to remove 90 to 95% of settleable solids, 40 to 60% of total suspended solids, and 25 to 35% of BOD ^[12]. Primary treatment mainly consists of sedimentation or settling by gravity. Sedimentation is a physical operation that consists of separate solid particles with density higher than that of the surrounding liquid ^[25]. Consequently, the supernatant liquid becomes clarified, while the particles at the bottom form a sludge layer, which is subsequently removed. This process is a fundamental unit operation in several phases in many wastewater treatment systems.

Settling tanks, also called sedimentation tanks or clarifiers, are commonly used in primary treatment; they reduce the flow velocity of the wastewater sufficiently to allow suspended solids to settle, in order to remove settleable material ^[14]. Many configurations of primary clarifiers are established, but the clarifier selection depends on the site conditions, size of the plant, local regulations, and engineering judgment ^[4]. All treatment plants use mechanically cleaned settling tanks of standardized circular or rectangular design ^[1]. The mass of solids accumulated in the bottom is called raw primary sludge, and mechanical scrapers collect the settled solids into a hopper where they are pumped to a sludge-processing area ^[12]. Floating matter, such as oils and greases, tends to have a lower density than surrounding liquid and rise to the surface of the settling tank. Therefore, floating material is removed (skimmed) from the tank surface for subsequent treatment.

Settling tanks are designed to operate on a continuous flow basis. Frequently, the most important design parameters for primary clarifiers are detention time, overflow rate, and weir loading rate [4]. These design parameters are based on the average flow rate conditions. Design and operation are still based on empirical relationships, consequently flow patterns, homogeneous turbulence, and the direct way of describing the movement of suspended particles and their interaction mechanism with settling and removal are rarely considered [26]. Hence, the optimal design is difficult to achieve in practice due that settling tanks do not show an ideal behavior during operation. Several factors and conditions can affect the performance of the sedimentation unit, such as several particle sizes and densities (different settling velocity), liquid density, variation in liquid temperature (viscosity), rate flow, and tank design [22]. Typical design information for primary settling tanks is provided in Table 2.

Table 2. Design parameters for primary settling tanks.

Parameter	Range	Typical Value
Detention time (h)	1.5 - 2.5	2.0
Overflow rate (m³/m².d)	-	-
-At average flow	32 - 50	40
-At peak hourly flow	78 - 120	100
Weir loading rate (m ³ /m.d)	125 - 500	260
Rectangular Tank (m)	-	-
-Length	15 - 90	25 - 40
-Width	3 - 24	5 - 10
-Depth	3 - 5	4.5
Circular Tank (m)	-	-
-Diameter	3 - 60	12 - 40
-Depth	3 - 5	4.5

Source: Adapted from Riffat [4].

The primary role of sedimentation units in a conventional treatment plant is the removal of suspended matter, however, in terms of the effect of primary sedimentation on the downstream biological processes, the removal of organic matter in the primary stage is of significant importance [27]. A significant part of TSS concentration consists of organic matter in suspension, which contributes to the BOD and COD of the wastewater. Thus, the reduction of TSS during primary sedimentation minimizes operational problems in the biological treatment because it reduces the energy necessary for the oxidation of particles. Hence, these effects improve the elimination of substrate during biological treatment and reduce the volume of activated sludge generated [28].

On the other hand, a poor solids removal during this step of the process causes organic overloading of the biological treatment. Frequently, the performance of settling tanks is measured by the removal efficiency, which is defined in Equation 1:

$$E = \frac{S_i - S_e}{S_i} \quad (1)$$

where:

E : removal efficiency (-)

S_i : influent contaminant concentration (mg/L)

S_e : effluent contaminant concentration (mg/L)

Therefore, the performance evaluation of primary settling tanks consists of comparing the removal efficiency during operation with typical ranges mentioned previously. In some cases, the sedimentation is enhanced by the addition of coagulants agents. The process is called enhanced clarification, or chemically enhanced primary treatment [4]. Coagulants may be aluminum sulfate ($\text{Al}_2(\text{SO}_4)_3$), ferric chloride (FeCl_3), or other, aided or not by a polymer [11]. Thus, more sludge is produced, resulting from a higher amount of settleable solids removed from the liquid.

2.3.3. Secondary Treatment

The main objective of secondary treatment is the removal of organic matter in the effluent from primary treatment. Secondary treatment mainly consists of biological treatment of primary effluent for reducing the BOD and suspended solids of the effluent to acceptable levels [4]. However, the overall objectives of the biological treatment of wastewater are: (1) transform (i.e., oxidize) dissolved and particulate biodegradable constituents, (2) transform or remove nutrients, such as nitrogen and phosphorus, (3) capture and incorporate suspended and non-settleable colloidal solids into a biological floc or biofilm, (4) in some cases, remove specific trace organic compounds and constituents [1]. The organic matter is present in the following forms:

- *Dissolved organic matter*: this organic material cannot be removed by physical operations as primary sedimentation.
- *Organic matter in suspension*: it can be removed mainly using primary sedimentation, but solids with slower settling velocity remain in the liquid phase.

Biological treatment occurs entirely by biological mechanisms, which reproduces the natural processes that take place in a water body after a wastewater discharge. The secondary treatment is considered in such a way to accelerate the degradation mechanism that naturally occurs in the receiving water bodies [11]. Therefore, a wastewater treatment plant uses technology for making the natural purification process of sewage developed under controlled conditions (operational control) and at high rates [25].

The decomposition of organic pollutants in the second stage is carried out through biochemical reactions undertaken by organisms [29]. Several microorganisms take part in the process as bacteria, protozoa, fungi, and others. The microorganisms convert the organic matter principally into CO₂, water, and cellular material (growth and reproduction of the microorganism) [1]. The microorganisms are able to decompose the organic matter by two different biological processes: biological oxidation and biosynthesis [30].

The biological oxidation produces some end-products that remain in the solution, and biosynthesis transforms the colloidal particles and organic matter into new cells that form in turn the dense biomass that can be removed by sedimentation [29]. Thus, the secondary treatment is composed of a biological reactor that produces the bioconversion of organic matter and a settling tank, which separates the microbial biomass from the treated

effluent. The secondary sedimentation process is similar to primary sedimentation except that the sludge contains microorganisms rather than fecal solids. The biological decomposition of organic matter requires the presence of oxygen as a fundamental component of the aerobic processes, besides the maintenance of other environmental conditions, such as temperature, contact time, and pH ^[11].

The tertiary treatment is not considered in this study because it is unusual in wastewater treatment plants, particularly in developing countries ^[2, 19].

2.4. Dynamic Modeling

Even though sedimentation is one of the most widely used unit processes in wastewater treatment, no satisfactory mathematical models have been developed as yet, mainly due to the complexity of sedimentation mechanisms ^[8]. Several sedimentation mechanisms take place during primary sedimentation, depending on the suspended particle concentration and the degree of particle interactions ^[31]. Wastewater contains flocculent particles, which can coalesce while settling and increase in size, and, as a result, they do not have a constant settling velocity ^[11]. Flocculation and settling are influenced by several factors such as the TSS concentration, particle size and density, specific velocity field, the density and viscosity of the fluid ^[8]. Many mathematical models of spatially one-dimensional solid-liquid separation processes of flocculated suspensions have been developed to unify studies of discrete settling and flocculating suspensions ^[32].

In practice, modeling the dynamic performance of full-scale primary settling tanks has been frequently done using regression-based models, which are empirical relationships derived strictly from observed daily average influent and effluent data ^[27]. These empirical models have generally been restricted to related the TSS removal efficiency to overflow rate or influent TSS concentration ^[8]. However, these empirical models are also used to determine the COD removal efficiency. Smith ^[33] has proposed an empirical relationship based on data collected at a large number of several treatment plants and different flow rates. The model proposed by Smith relates the TSS removal to the overflow rate in a non-linear function suggesting that the performance of the settling tank will always deteriorate at higher flow rates. The empirical relationship developed by

Smith is defined in Equation 2. This empirical model was proposed by Jover-Smet *et al.* [7] for modeling a semi-technical primary settling tank with successful results.

$$E = A \cdot e^{-(Bq)} \quad (2)$$

where:

E : removal efficiency of TSS (-)

q : surface overflow rate ($\text{m}^3/\text{m}^2 \cdot \text{h}$)

A and B : adjusted constants (-)

However, a pilot-scale study over a wide range of hydraulic loadings shows a lack of correlation between the overflow rate and the TSS removal efficiency, and it was attributed to the flocculent nature of wastewater solids [27]. Tebbutt and Christoulas [34] suggested an empirical relationship that relates the influent TSS concentration and overflow rate to the TSS removal efficiency (Equation 3).

$$E = D \cdot e^{-(F/S+Gq)} \quad (3)$$

where:

q : surface overflow rate ($\text{m}^3/\text{m}^2 \cdot \text{h}$)

S : influent TSS concentration (mg/L)

D, F and G : adjusted constants (-), see Appendix C

These empirical models regularly have a satisfactory performance under steady-state conditions because these conditions are similar to design parameters (ideal behavior). However, load variations represent an important part during the operation life of treatment plants, and most of the observed problems in complying with permit requirements are due to these transient conditions [27]. Thereby, to describe the performance of primary settling units under transient conditions is necessary to be considered to define the effect of variation loads on the removal efficiency. Hydraulic efficiency models are also used to describe the dynamic behavior of settling tanks. These models identify that real flows are turbulent and encounter a certain degree of mixing or eddy diffusion [27]. However, implementing these models to predict the dynamic response of a full-scale settling tank is very difficult because these models have been developed using controlled studies of model tanks.

In recent years, models based on computational fluid dynamics (CFD) have been used to predict flow patterns and suspended solids distributions within sedimentation [7]. However, in practice, CFD-based models are not commonly used due to the complexity of the Navier-Stokes equations in turbulent flows and the cost associated with specialized instrumentation and software required [10].

Finally, there exists another class of models that do not try to describe the internal function of the system. However, they can establish complex relationships between measurable inputs and outputs data of the system. These models are called "black box" class models. Based on this idea, artificial neural networks have been successfully applied in many engineering areas. The main advantage of these models is that neural networks can generate complicated relationships through the inspection of a data set without assuming theoretical and mathematical formulation.

2.5. Fundamentals of Artificial Neural Network

Artificial neural networks (ANN) are popular machine learning techniques that simulate the mechanism of learning in biological organisms [35]. They are powerful non-linear regression models inspired by how the brain works [36]. For example, the human nervous system contains cells that can process and transmit information to each other; these cells are called neurons. These neurons are the information-processing cells in the brain [37]. Neurons are connected between structures referred to as axons and dendrites, and this connecting region between one cell to another is named as synapses. Thus, neuron connections form complex information processing networks that are responsible for our sensations, feelings, and actions. The strength of synapsis connections changes depending on the external stimulations, and this change is how learning takes place in living organisms [35]. Biological organisms learn how to solve problems using information gained from past experiences, and this mechanism is replicated in an ANN. As a general idea, the ANN modeling approach is a computer methodology that attempts to simulate some important features of the human nervous system, in order to develop the ability to solve problems by applying information gained from the past to new problems or case scenarios [27]. Neural networks are widely used in a variety of applications pattern classification, complex system modeling, control, optimization, and prediction [38].

2.5.1. Components of Artificial Neural Networks

ANN consists of a group of simple processing units which communicate by sending signals to each other over a large number of weighted connections [39]. A neuron or node is a processing unit in a neural network, and these units are connected by connection weight, which determines the strength of the relationship between two connected neurons [40]. Hence, a neuron receives input data from neighbors or external sources and uses it to compute an output signal which is propagated to other units.

In Figure 2, the circle and arrows represent the node and signal flow, respectively. The input data is denoted by x_1, x_2, x_3 . The connections between input data and the node are represented by w_1, w_2, w_3 , which are the weights for the corresponding input signals. Lastly, b is the bias, which is a factor associated with the storage of information in neural networks. The information on neural networks is stored in the form of weights and bias [41].

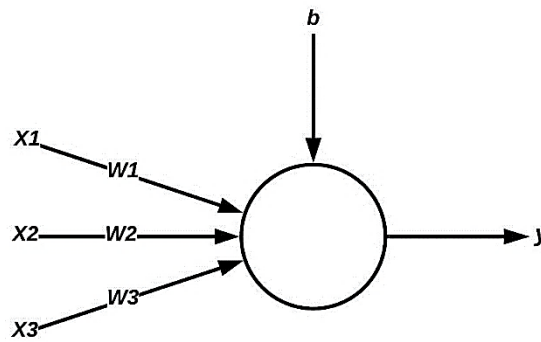


Figure 2. A node that receives three inputs and bias.

The input data is multiplied by the weight before it reaches the node. Thus, the total input (*net*) to the neuron is the weighted sum plus the bias. Hence, the total input for a neuron which receives n input signals is defined as the following in Equation 4:

$$net = \sum_{i=0}^n w_i \cdot x_i + b \quad (4)$$

Nodes computing the output signal (y) according to a transfer function called activation function (ϕ). The activation function determines the behavior of the node; it represents a

linear or non-linear mapping from the input data to the output signal ^[42]. The mathematical expression of the node output signal is presented in Equation 5.

$$y = \phi(\text{net}) \quad (5)$$

The activation function is the core of a neural network structure, and it is used to increase the modeling ability of the neural network model, which can make the neural network has the meaning of artificial intelligence ^[43]. Some of the most common functions used as activation functions are to solve non-linear problems. The use of non-linear activation functions is fundamental to improve the modeling power of a neural model ^[35]. An important characteristic of activation functions is that they limit the amplitude range of the output signal of a neuron to some finite values ^[44]. The activation function is usually some continuous or discontinuous function mapping the real numbers into the interval (-1,1) or (0,1) ^[42]. We list below some functions that are frequently used as activation functions.

- *Identity or Linear Function*: it is the most basic activation function, and it provides no nonlinearity (Equation 6). This function is commonly in output nodes when the target is a real value.

$$\phi(x) = x \quad (6)$$

- *Logistic or Sigmoid Function*: it is a widely used non-linear activation function in machine learning (Equation 7). The sigmoid activation outputs a value in (0,1), which is useful in performing a computation that should be interpreted as probabilities ^[35].

$$\phi(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

- *Hyperbolic Tangent Function (tanh)*: it is similar to the sigmoid function, but the output values range is (-1,1). The tanh function (Equation 8) is preferable to the sigmoid function when the output of the computations are desired to be both negative and positive ^[35].

$$\phi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

2.5.2. Artificial Neural Networks Topology

Despite each neuron that can only perform simple computations operations, the hierarchical organization of interconnected neurons makes an ANN capable of performing complex tasks as pattern classification and prediction ^[27]. Depending on how the nodes are connected, several neural network structures will be developed. The architecture of neural networks consists of a network of nodes that are normally arranged in layers and executed in parallel, and the layered arrangement for the network is referred to as the topology of a neural network ^[38].

Depending on the pattern of connections between nodes and the propagation of data, several types of neural network architectures of complex structures can be developed ^[39]. The most simple architecture of neural networks is the perceptron, which consists of an input layer and an output layer ^[41]. However, when hidden layers are added to the perceptron structure, more complex structures are produced, which are called multilayer perceptron (MLP) neural networks. Therefore, MLP neural networks contain multiple functional computational layers; it consists of an input layer, a hidden layer(s), and an output layer ^[42]. A specific architecture of a multilayer neural network is referred to as feed-forward networks, where the data flow from input to output units is strictly feed-forward ^[40]. Consequently, no feedback connections are present in the network structure. Figure 3 shows a schematic diagram of the feed-forward multilayer neural network.

For understanding neural network systems, it is useful to distinguish three types of units:

1. Input units that receive data from outside the neural network.
2. Output units that send data out of the neural network.
3. Hidden units whose input and output signals remain within the neural network.

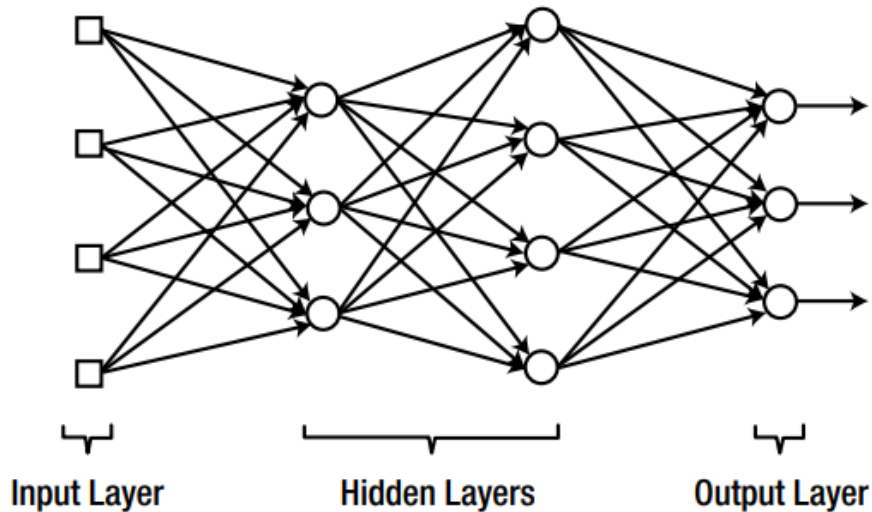


Figure 3. *Feed-forward multilayer neural network.*

(Source: adapted from Kim ^[41])

The group of input nodes is defined as the input layer. The nodes of the input layer have no functionality; they only receive input signals and transmit the input information to the next nodes ^[45]. In contrast, the output nodes provide the final result of the neural network. Finally, the layers in between input and output layers are called hidden layers. Hidden nodes connect the input nodes to the output nodes and provide nonlinearity to the neural network model ^[27]. Hence, the output signals of hidden layers are subsequently processed by the output layer.

2.5.3. Training of Artificial Neural Networks

The main characteristic of neural models is the ability to solve problems by using information gained from past experiences. Therefore, neural networks have the capacity to generalize. In machine learning, the generalization ability is developed during the training phase. There are several learning algorithms that are used to train neural networks; however, the most used in practice is the class of algorithms that are based on supervised training ^[38]. The supervised training consists of providing to the neural network a training set comprising input data and desired outputs (target output) ^[41]. Therefore, the supervised training of neural networks for a specific task is done by adjusting the weights of the network to minimize the discrepancy between the network

output and the target output on a training data set ^[46]. As a general idea, the algorithms used for supervised training consist of the following steps (Figure 4):

1. All weights are assigned initial random values.
2. Input data are propagated in a feed-forward way through the network.
3. Network output data is produced according to the weight and the activation function.
4. Outputs produced by the neural network are compared with the target outputs.
5. The error generated is propagated back through the network.
6. The weights are adjusted according to the errors propagated back.

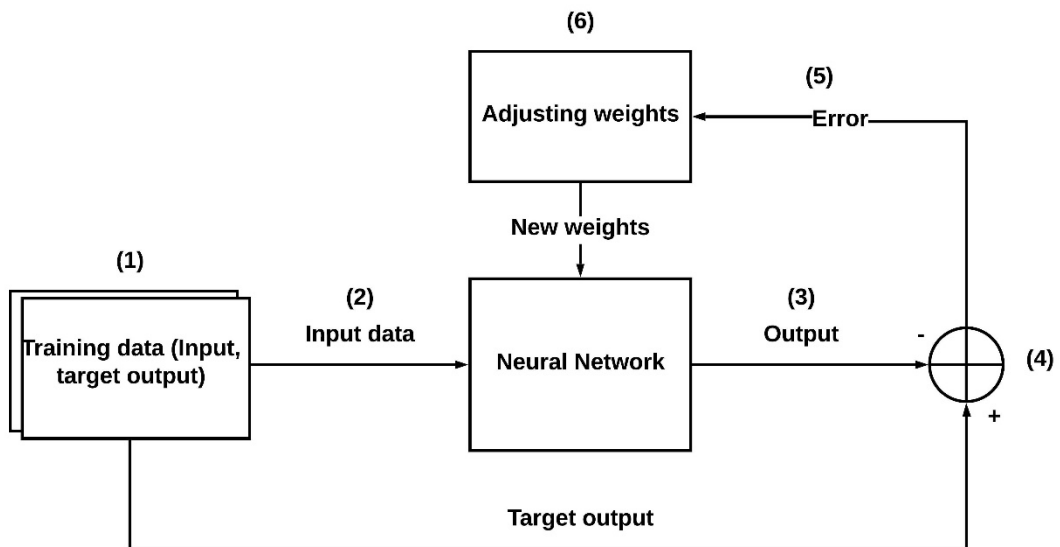


Figure 4. Schematic representation of supervised learning algorithm.

This procedure is repeated until the errors are minimized; it means that the training data set is processed many times as the weights between nodes are refined during the training of the network. One pass over the whole training data set represents a cycle of training, and it is defined as an epoch. Multiple epochs are required until achieving a satisfactory data mapping. The backpropagation algorithm has been widely used for training multilayer neural networks. The algorithm uses a gradient search technique to minimize

the cost function equal to the mean square error (MSE) between the target output and the actual network output [27]. The backpropagation algorithm is often a training method too slow for practical problems. Subsequently, numerical optimization techniques have been implemented to develop algorithms that can converge faster. Levenberg-Marquardt algorithm (LMA) has become a standard technique for non-linear least-square problems, and it is ranked as one of the most efficient training algorithms for multilayer neural networks [47]. A combination of gradient descent and Gauss-Newton methods is used in LMA, and it guarantees problem-solving through its adaptive behavior [48]. A detailed mathematical description of these training algorithms is provided elsewhere [35,41].

Hence, the main advantage of neural models is that neural networks can generate complicated relationships through examinations of only the data points in the training set without assuming a pre-specified functional form [27]. As a result, complicated mathematical models are replaced by a training period using experimental data. The processing ability of neural networks is stored in the strengths of the connections between nodes (weights), which are obtained by the process of learning from a set of training patterns [45].

The generalization capability of a neural network is determined by the size of the training data set, the complexity of the problem, and the architecture of the network [42]. Also, the learning algorithm and the number of epochs used for the training process will influence the performance of the neural network model. The training data set should be sufficiently large and diverse in order to represent the problem well [42]. Typically, supervised training starts with a neural network comprising an arbitrary number of hidden nodes, a fixed architecture of connections, and randomly selected values for the weight [50]. The modeling capacity of the model is proportional to the number of hidden nodes used in the neural network. However, using an excessive number of hidden nodes will inhibit the network's ability to generalize [39]. Hence, determining the optimum number of hidden nodes is a fundamental task to guarantee the convergence of neural networks during the training process, without inhibiting the generalization capacity of the model.

2.5.3.1. Overfitting

In statistics, overfitting corresponds to the situation wherein a model possesses too many parameters and fits the noise in the data rather than the underlying function [42]. Overfitting occurs when the model has been overly customized to the training data that it

yields a poor performance for new input data, but its performance for the training data is excellent ^[41]. Consequently, the model has memorized the training patterns, but it has not learned to generalize to new scenarios. Therefore, overfitting is one of the principal factors that reduce the ability of generalization in neural models.

There are several methods to confront the overfitting problems. However, the two typical methods used are regularization and validation. Regularization is a numerical method that attempts to construct a model structure as simple as possible, modifying the cost function ^[51]. This numerical method is commonly implemented in many training algorithms; a detailed mathematical description is shown elsewhere ^[52]. On the other hand, validation is a method that reserves a part of the training data set and uses it to monitor the performance ^[41]. Thus, the available data for training a neural network is divided into three subsets: training, validation, and test. The first data set is used for computing the learning algorithm and modifying the network weights and biases. The second subset is used to evaluate the performance of modeling, and it is an indicator of the generalization capability of the model during the training. The third subset does not affect the learning process, it is used as an independent measurement of neural network performance after the training phase. It is important to mention that the validation data set is not used for the training process; it does not affect the adjusting of weights and biases in the neural network. The errors of both training and validation data set are monitored during the training process. The error of the training data fails to indicate overfitting, but the error on validation data provides information about when the model begins overfitting the training data. During the initial phase of training, the training and validation error will normally decrease. However, when the model begins overfitting the data training, the error on the validation set begins to increase. Hence, when validation error increases for a specified number of iterations, the training process is stopped, and the weights and biases of the network at the minimum of the validation error are returned ^[51]. Therefore, overfitting can be avoided by stopping the network training before the absolute minimum of training error is reached. This method to avoid the overfitting problem is known as early stopping. Figure 5 illustrates a schematic representation of the early stopping method.

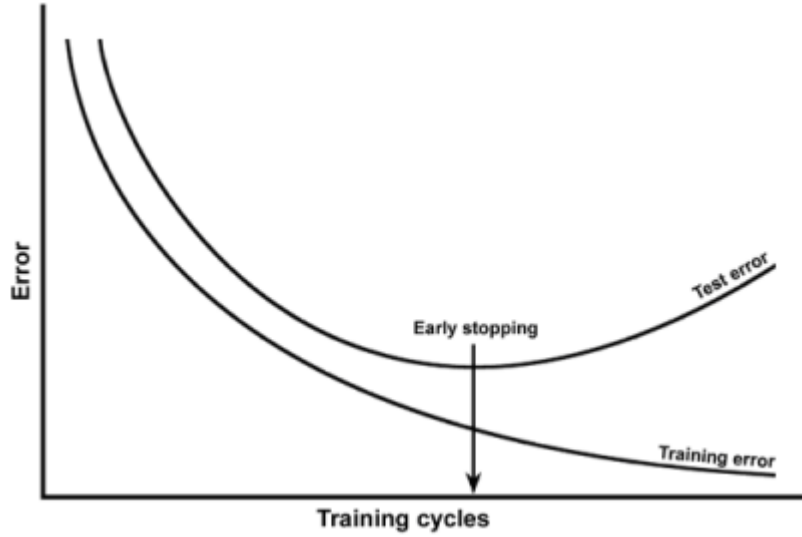


Figure 5. Representation of early stopping method.

2.5.4. Evaluation of Artificial Neural Network

The objective of evaluating the performance of neural models is to find a model that is as simple as possible that fits a given training data with sufficient accuracy, and provide a good generalization capability to new scenarios ^[42]. A test data set is used to evaluate the performance of the neural network after the training phase. Several statistics indicators are used to estimate the performance of neural network models. However, for this study were selected the following:

- **Mean Square Error (MSE):** for supervised learning, an error measure is used to control and evaluate the training process. The error measure is usually defined by the mean square error (MSE) ^[42]. The MSE has been a useful measure widely used in model evaluations. In statistics, MSE is an estimator to quantify the difference between the values estimated by a model and the population parameter ^[53]. Specifically, MSE is the average squared difference between the model predictions and target outputs. Its mathematical expression is defined in Equation 9:

$$MSE = \frac{1}{N} \sum_i^N (y_i - y_i^p)^2 \quad (9)$$

Where N is the number of patterns in the sample data set, y corresponds to the value target output, and y^p is the value of model prediction. The best performance of the model is usually achieved when the MSE function is minimized.

- **Coefficient of Determination (R^2):** the coefficient of determination (R^2) is a statistical indicator that measures the proportion of variance in measured data explained by the model ^[54]. For linear models, it measures the goodness of fit and precision in predictions for the general linear model ^[55]. The coefficient of determination is mathematically described as the following in Equation 10:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i^p)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (10)$$

Where N is the total number of data records, y is the observed value of a specific parameter, y^p is the model prediction and \bar{y} is the mean value of observed data for the constituent being evaluated. The coefficient of determination ranges from 0 to 1. Higher values of R^2 indicates a better data fit of the model; a perfect fit would achieve in an R^2 value of 1. Typically, values of R^2 greater than 0.5 are considered acceptable ^[54].

CHAPTER III

3. IBARRA WASTEWATER TREATMENT PLANT

The objective of Ibarra WWTP is to treat the wastewater from the mixed sewerage of Ibarra. This plant has been operating since September 2018, and it processes an average sewage flow of 43,200 m³/day from the community. The treatment plant was built on a terrain situated on the left bank of the Tahuando River with a total surface area of around 4.5 hectares. The plant is located between Carchi Avenue and the river. The treated effluent is discharged on the Tahuando river, where the natural purification process continues to reduce the levels of pollutant concentration.

3.1. Process Description

The Ibarra WWTP uses a biological treatment system to process mainly municipal effluents, but industrial effluents are also present in the wastewater. The activated sludge system is used as a biological treatment, which is a widely used process for wastewater treatment. The treatment process is divided principally into four steps: Pretreatment, primary treatment, secondary treatment, and sludge treatment. Figure 6 corresponds to the Block Process Diagram (BPD), which provides a schematic representation of the process.

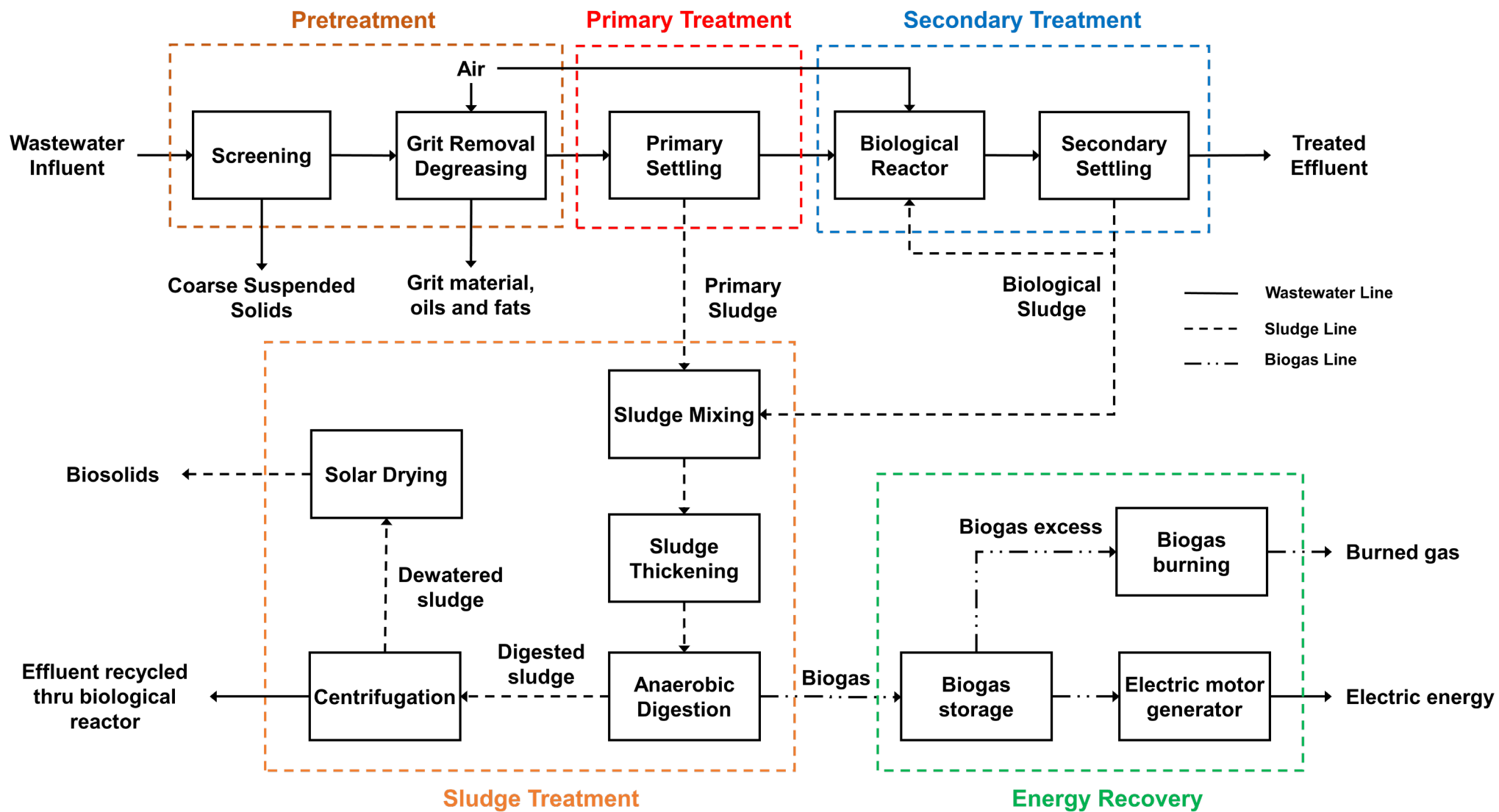


Figure 6. Block Process Diagram of Ibarra WWTP.

The wastewater is collected in a connection chamber located outside the plant under the Avenida Carchi. From this structure, the water is sent employing a PVC pipeline of approximately 250 m in length to the WWTP [56]. The inlet structure and preliminary treatment process are located in a single building, and an accumulative electromagnetic flowmeter is installed in the WWTP inlet line to record the inlet flow of wastewater (influent).

3.1.1. Inlet Structure

The sewage pipeline discharges into large-particle wells, which enables calming and homogenize the inflow prior to treatment. The large-particle well is fitted with a relief spillway, which acts as a general bypass line. The bypass line can carry the maximum wastewater inflow to the plant in case of an emergency. Detailed information on design flows and wastewater parameters considered to the design of the Ibarra WWTP are shown in Table 3 and Table 4, respectively.

Table 3. Design flows of the Ibarra WWTP.

Design flow parameters	Design Value
Average flow (m ³ /h)	1,656
Average Daily Flow (m ³ /day)	39,744
Peak Flow (Primary Treatment) (m ³ /h)	3,312
Maxim Flow (Pretreatment) (m ³ /h)	6,624

Source: Acciona and BTD Proyectos [28].

Table 4. Design wastewater parameters of the Ibarra WWTP.

Parameters	Inlet Ibarra WWTP	Outlet Ibarra WWTP
BOD ₅ (mg/L)	210	100
COD (mg/L)	500	250
TSS (mg/L)	210	100
TKN (mg/L)	40	25

Source: Acciona and BTD Proyectos ^[28].

A large-particle well was built to remove bulky solids present in the raw water inflow. This well is designed to handle the maximum inflow to the plant, using a retention time of 1 minute ^[28]. Thus, it has a working capacity of 113.03 m³. The lower structure of the well has an inverted pyramidal shape to allow the concentration of bulky solids at the bottom. A clamshell grab with a capacity of 100 L is installed for the removal of these solids. An electric hoist with a capacity of 4000 Kg is installed to handle the grab and carry solid wastes to the storage containers.

3.1.2. Pretreatment

The inlet wastewater to the treatment plant principally has domestic characteristics. It is from the community and the municipal sewerage. The raw water contains several bulky solids and non-biodegradable material, which could damage the equipment and pipelines during the treatment process. Therefore, a pretreatment process is necessary to remove these solids. This part of the treatment consists of the following processes:

3.1.2.1. Screening

Two bar screens with a space between bars of 80 mm are installed at the outlet of the large-particle well. The screening is the first step of the treatment process, and its objective is filtering coarse and fine particles using bar screens. The screening takes place in 4 channels of 2 m in height and 1 m in width. The maximum flow rate in these channels is 0.34 m/s, and each channel is equipped with a bar screen with a size opening of 10 mm

^[56]. Screened material is removed using a screw-press conveyor fitted with hoppers to collect the discharged material from the bar screens. The screw conveyor transports the collected material to a container with 1100 L of capacity, and it is used mainly for landfill. The screening system is controlled through the combinations of programmed timing and level controller.

Motorized gates are located at the inlet and outlet of each channel, so filtration lines go into operation automatically in accordance with the inflow of sewage. Each channel can operate with the average inflow of the total capacity of the plant. Fine screens are installed subsequent to the coarse screens into the channel. Fine screens have a size opening of 3 mm. A bypass channel with the capacity to carry the maximum inflow to the plant is fixed in parallel to the screening channels. A manual bar screen with a 16 mm passage size is fitted in the bypass channel to screened large solids ^[28]. The bypass channel will be used exclusively when maintenance work is being carried out on the screening channels.

3.1.2.2. Grit Removal-Degreasing

A mixed structure is used to remove oils, greases, floating solids, and grit particles from the wastewater. The structure consists of two channels with 25 m large and 4.5 m width, and it has the capacity to treat the maximum flow from the pretreatment. However, the inflow of these channels is regulated with motorized gates. Hydraulic equilibrium in this structure is reached using a weir located at the outlet of the channel. Thus a constant level of water is ensured to facilitate the removal of grease and floating solids ^[56]. Each channel is swept by a mobile bridge, which runs forward and backward in both channels. The following equipment is suspended on the bridge:

1. Surface scrapers with a counterweight elevation system to collect floating material. These scrapers have a 1.5 m length as the width of the grease collection channel.
2. Bottom scrapers are used to reduce the formation of deposits at the lowest part of channels.

Grit particles are collected on the floor of the channels and extracted using pumps, which move it with the trajectory of the bridge. These pumps are specially designed for this work, and they feature manually adjustable suction heights. Each channel is equipped with one pump with a capacity of 20 m³/h for grit extraction, and these pumps feed a grit/water mixture collection channel, which separates the mixture using a screw grit classifier [28]. The water removed from the mixture is sent to a collection well and then returned to the WWTP inlet.

On the other hand, the grease has to be emulsified before removal from the sewage. Therefore, this emulsion is realized by an air distribution system, which supplies air through diffusers. The aeration system consists of blowers and fine bubble diffusers, and this system is used in both channels of the mixed structure. Each channel has three blowers (two operating and one standby) with a unitary flow of 950 m³/h, with 200 diffusers per channel [56]. The air injected by blowers maintains a cross-circulation speed, which, as a result of the turbulence effect, enables the separation of the organic matter attached to the grit particles and prevents the mass accumulation of grit material into the system [28]. Superficial scrapers send the removed grease and floating substance to the collection zone, where they are extracted and sent to the grease concentrator. Subsequent to thickening in the concentrator, the greases and supernatants generated are sent to a waste container using a scraper system. The degreasing equipment and containers are arranged alongside the channels of mixed structure.

An electromagnetic flowmeter is installed at the outlet of the grit removal-degreasing process to record the inflow rate into primary treatment. Regulation of an excess flow is carried out through a weir installed in the collection channel of the grit removal-degreasing structure.

3.1.3. Primary Treatment

Primary settling is used for the removal of suspended solids that are not removed in the preliminary treatment. Primary treatment consists of removing suspended solids by gravitational sedimentation [8]. Two circular settling tanks are operating in the plant, and they can function as primary or secondary setting tanks in the process. This is an essential characteristic of flexible design in terms of operating conditions. At average flow, a

complete process consisting of primary sedimentation, biological treatment, and secondary sedimentation, is operating. At peak flow, the concentration of pollutants in sewage increases, and a high loading treatment is used. Thereby, at these conditions, both circular sedimentation units are used for secondary sedimentation after biological treatment.

Pretreated wastewater is sent to a distribution chamber, which directs the flow to the primary settling tanks through spillways. Sluice gates installed at the outlet of the distribution chamber allow each settling tank to operate isolated. Thus, the inflow to each settler is distributed through separate pipelines. Consequently, the clarified water returns separately to the distribution chamber from each settling tank, and then it is sent to the biological treatment.

3.1.3.1. Primary Settling Tank (PST)

A circular settling tank fitted with a scarper system is used for primary sedimentation. Circular sedimentation tanks are center-feed with the purpose to produce a radial flow pattern [31]. The PST has a diameter of 34 m and a cylindrical height of 4 m, but 2.8 m cylindrical useful height. The volume of the unit is 2,902.32 m³, and the total surface is 907.92 m². A schematic representation of the primary settling unit is shown in Figure 7.

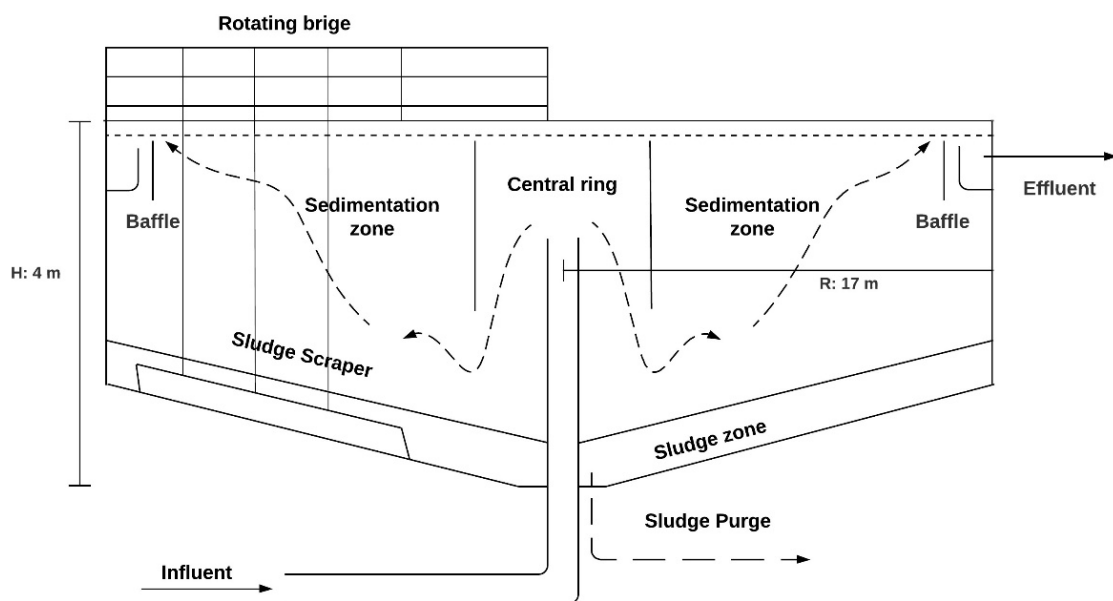


Figure 7. Primary settling tank (PST) cross-section view.

The influent flows through a vertical pipe and enters the settling tank at the bottom of the structure. The inlet flow is collected in a central structure called the central ring, which generates a homogenous radial flow through a head loss. Typically, the central ring has a diameter between 15 and 20 % of the total tank diameter ^[1]. The circular ring structure consists of narrow vertical windows. The windows cause the head loss of water, and it ensures a homogenous distribution on a horizontal plane ^[28]. Therefore, the flow rate through the windows needs to be sufficiently high to produce a uniform distribution, but low enough to avoid turbulent flows. This reduction in the flow velocity allows that suspended particles can settle along the sedimentation zone. The flocculation settling mechanism usually occurs in primary sedimentation ^[4]. The layer of accumulated settled solids at the bottom of the tank is denominated sludge.

Additionally, the sedimentation unit is fitted with a rotating bridge that provides a constant rotatory movement to a sludge scraper. The sludge scraper moves the settled sludge to a hopper structure in the middle of the settling tank. The accumulated sludge is posteriorly removed using centrifugal pumps. Also, the rotating bridge is fitted with surface skimmers to remove floating material from the surface of the settling tank. Baffles are the structure that prevents the remaining floating solids and scum leaving the settling tank with the clarified effluent ^[12]. The clarified water is collected using a perimeter spillway in the clarified water collection reservoir and returns through the outlet pipeline to the distribution chamber. However, the clarified water can be sent to biological treatment or for final discharge; it depends on the operational mode and wastewater characteristics. Chemical flocculants are not used during this process.

Sludge extraction or sludge purge from the settling tank is performed through pipelines that transport the accumulated sludge to the pumping well. The sludge purge from PTS is carried out every 8 hours for 3 hours, with a sludge extraction flow of 40 m³/h. The sludge production corresponds to 5,425.06 kg/day, and the concentration of this sludge on extraction from the settler is 10 kg/m³ ^[56]. The sludge is sent to a screening process and after a sludge treatment process. The equipment used for sludge extraction is detailed in Table 5.

Table 5. Purge equipment to sludge extraction.

Description equipment	
Type of pump	Centrifugal
Operating Units	2
Reserve Units	1
Unit pumping flow	50 m ³ /h
Pressure	14.22 psi

Source: Acciona and BTD Proyectos ^[28].

3.1.4. Biological Treatment

The clarified effluent is sent to biological treatment through a pipeline that feeds a distribution channel to biological treatment at a constant flow. The maximum inflow to biological treatment is 3,312 m³/h. Biological treatment is performed by two plug flow activated sludge reactors, with aeration by diffusers. The volume of each biological reactor is 5,190 m³, and the aeration is produced by a membrane diffuser at a depth of 5 m ^[28]. Air is supplied by uniformly using diffuser grids.

The treated effluent leaves the reactor by the spillway, and it is sent through pipelines to the secondary settling unit. However, in the activated sludge biological treatment with a high loading rate adopted at the plant, it is necessary for sludge recirculation. After separating the biomass from treated effluent, the biomass is reintroduced to biological reactors frequently in order to keep the microbial mass active and avoid the loss of microorganisms. Sludge recirculation is carried out with submersible pumps from the outlet of reactors (internal recirculation) or the secondary settling tank (external recirculation).

3.1.5. Secondary Sedimentation

The effluent from biological treatment is sent to the same distribution chamber that feeds the primary settling unit. However, in this case, it acts as a distribution chamber to secondary settling. The performance and physical characteristics of the secondary settling tank is similar to the primary settling unit. The purge of the secondary settling tank is performed in continuous mode using a centrifugal pump. A fraction of sludge is returned to biological reactors, but it depends on the concentration of pollutants in wastewater. The rest of the sludge is sent to the sludge treatment process. The average production of sludge is 3,386.25 kg/d, and the concentration of this sludge on extraction from the secondary settling tank is 6 kg/m³ [56].

The decanted effluent is collected by a double perimeter spillway in the decanted water collection channel and sent to the decanted water collection chamber. The treated water flow is measured by an accumulative electromagnetic flowmeter installed in the main outlet pipeline of the plant. After the point at which is installed the flowmeter, the pipe is connected to the bypass line by a control well. Finally, treated water is discharged to the Tahuando river.

3.1.6. Sludge Treatment

The sludge settled in sedimentation units need to be treated to produce stable and non-contaminant products. The sludge treatment implanted in the Ibarra WWTP involves three general steps: Sludge thickening, anaerobic digestion, dewatering and drying.

3.1.6.1. Sludge thickening

The installed system for the reception of sludge excess consists of an inlet structure, sieving, compaction zone, and discharge zone. The sludge excess from the primary sedimentation unit is screened in a rotatory drum sieve with an open size of 3 mm, in order to remove coarse solids and fibers that will affect the equipment and treatment process. The filtered sludge is dewatered at the required conditions in the compaction zone, and it is sent to a mixing chamber with the sludge of secondary sedimentation. The mixing chamber has a retention time of 2 hours, and it is fitted with submersible mixers for homogenizing the sludge before the thickening [28].

The mixed sludge is pumped from the mixing chamber to the thickener. The sludge thickening is carried out in a gravity thickener with a diameter of 15 m. The thickened mixed sludge is stored in a tank with 34.02 m³ of capacity before digestion. The storage tank is fitted with a submersible agitator to homogenize the thickened sludge.

3.1.6.2. Anaerobic digestion

A narrow digester with a pyramid-shaped is used for the digestion process. This reactor requires equipment to enable feed-in, heating, homogenization, extraction, etc. All these operational variables need to be controlled in order to achieve the optimal performance of the digestion unit. The final products produced by digestion are principally digested sludge, CO₂ and CH₄, but secondary gases as hydrogen sulfide (SH₂) and saturated vapor water are also produced in the process. Hydrogen sulfide is a highly oxidizing gas that affects metallic structures (premature aging), and also it forms sulfuric acid (H₂SO₄) in contact with water. Therefore, a dosing system of ferric chloride (FeCl₃) is installed to reduce the formation of SH₂ during the digestion process. A fraction of the biogas generated in the digester is sent to boilers to be used as fuel. On the other hand, the rest of the biogas is stored in a spherical double-membrane gas holder with a capacity of 550 m². The gas holder has a service pressure of 20 mbar, which is adequate to feed the gas engine to produce electric energy, and a gas flare with a capacity to burn off 200 Nm³/h is installed to eliminate the gas excess ^[56].

The digested sludge is stored in a sludge buffer tank with a capacity of 542.84 m³ prior to the dewatering process. The tank is covered by a plastic cover connected to the odor control system. The stored sludge is pumped to the dewatering process gradually.

3.1.6.3. Dewatering and drying

With the purpose of reducing the volume of excess sludge, the sludge is dewatered in a centrifuge to obtain a dry matter content of approximately 20% ^[28]. The dewatering system consists of two centrifuges with each corresponding pump system. However, one complete line corresponds to a complete standby line.

Before the sludge enters the dewatering system, it undergoes chemical conditioning by polyelectrolyte (commercial product: FLOPAM AKA 8400) dosing. The addition of polyelectrolyte is used to improve the flocculation of sludge in order to get easier the

separation liquid-solid during the centrifugation. The dried sludge is transported by two screw conveyors and stored in a storage hopper with a capacity of 60 m³. Finally, solar thermal drying is used as an eco-friendly sludge drying alternative. A tractor-drawn wheel loader is used to transport the sludge from the storage hopper to the solar drying system. The solar drying transforms the sludge from the centrifuge, which has a large volume due to its high-water content, into a smaller granulated mass. This final product has a heating power similar to coal so that it can be used as fuel or fertilizer ^[56]. The solar thermal drying involves two drying surfaces, with a unitary volume of 960 m². Additionally, each drying surface is fitted by a scarper system to achieve a homogenous distribution of sludge.

CHAPTER IV

4. METHODOLOGY

In order to develop a predictive model for the PST based on neural networks, first, a process analysis was developed to identify the critical variables associated with the performance of PST. The Ibarra WWTP supplied the available data about wastewater characteristics and inlet flow rates corresponding to the primary treatment unit. The determination of chemical, biological, and physical properties of wastewater composition was done according to the Standard Methods ^[15]. Operational variables and wastewater characteristics measured were classified to obtain the most representative information about the behavior of the sedimentation unit. Data analysis was realized to guarantee the quality of the data set used for neural network models.

The objective of developing neural network models was to predict the COD and TSS concentration in the effluent wastewater from PST. Thus, two neural networks were trained for this task using the available experimental data. One neural network was trained to predict effluent COD (COD model), and the other to predict effluent TSS concentration (TSS model). The performance of neural network models was compared to the traditional empirical models. Additionally, a complete sensitivity analysis (SA) was developed to study the effect of each input variable on the model response. Hence, local and global SA was developed to determine the model response under several conditions. Once the models were developed and trained, new experimental data were provided by the laboratory of Ibarra WWTP to validate the accuracy and precision of both model predictions. Figure 8 shows a schematic representation of the methodology developed in this study.

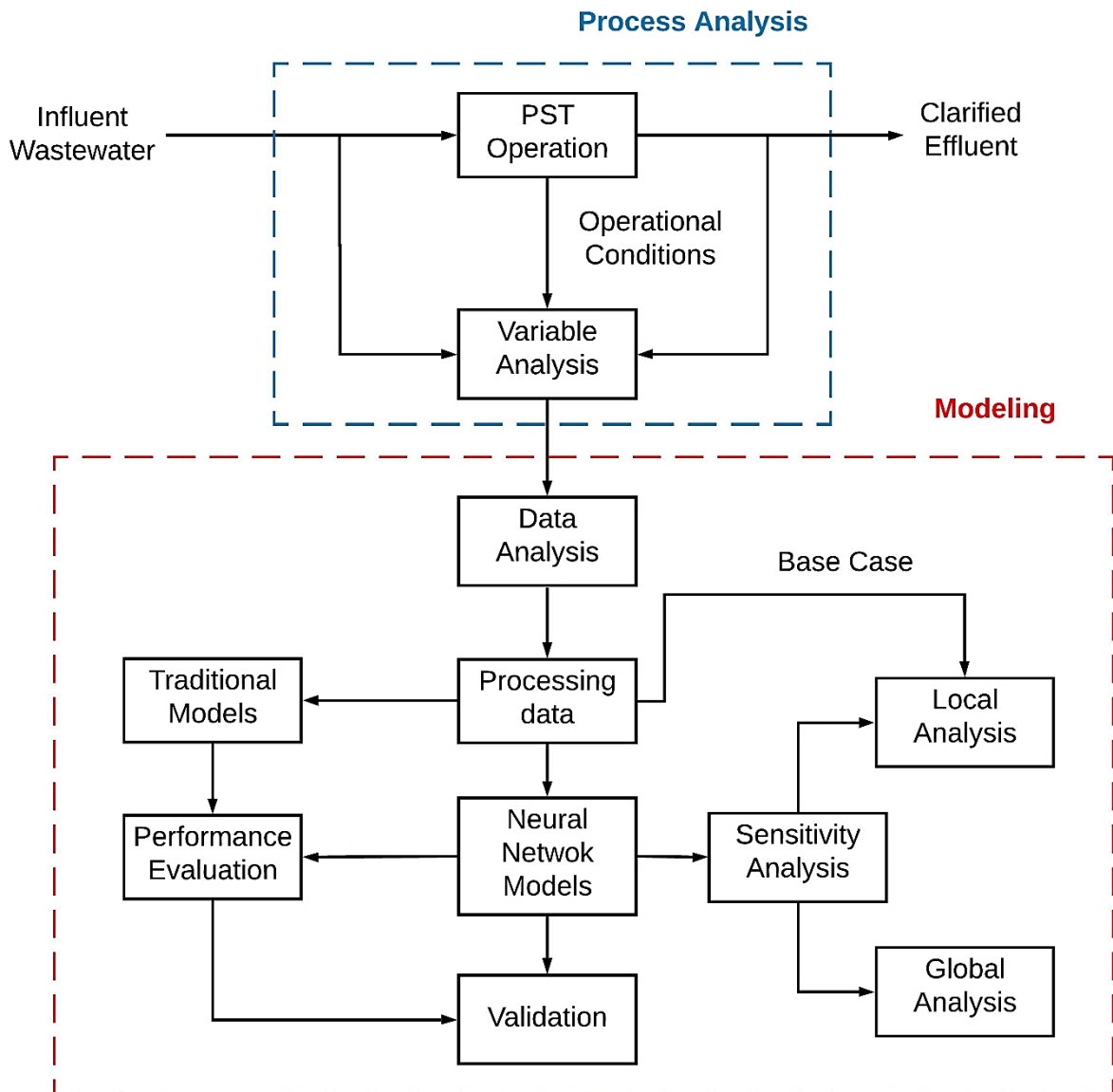


Figure 8. Methodology diagram.

4.1. Process Analysis

An integral process analysis was considered to identify the variables associated with operational conditions and wastewater characteristics of Ibarra WWTP. In the treatment process, 14 parameters are measured to evaluate the performance of the treatment process. The chemical, physical and biological properties measured in Ibarra WWTP are presented in Table 6. In primary treatment, these parameters are measured at the settling tank inlet and outlet. According to the Ibarra WWTP laboratory data, the delay time between inlet measurement and outlet measurement was approximately one hour.

Table 6. Wastewater properties that are measured during primary treatment.

Properties	Parameters	Units
Physical	Inlet Flow rate	m ³ /h
	Temperature	°C
	Turbidity	NTU
	Conductivity	μS/cm
	TSS	mg/L
	VSS	mg/L
	FSS	mg/L
Chemical	BOD ₅	mg/L
	COD	mg/L
	DO	mg/L
	pH	-
	TKN	mg/L
Biological	TC	MPN/100mL
	FC	MPN/100mL

For modeling, it was necessary to select the most influential and critical variables associated with the sedimentation process. Many parameters will complicate the learning process of neural networks, and more hidden nodes will be necessary. Therefore, the selection of the most representative parameters to the data set for neural networks is an important step in the modeling. Hence, an expert on the field of the wastewater treatment process was consulted to select the most important parameters associated with primary sedimentation.

Biological properties were not considered because sedimentation is essentially a physical process. The microbiological activity is not significant during primary sedimentation; it does not have an important influence on the effluent COD and TSS concentration. Nevertheless, some properties as the concentration of total and fecal coliforms, BOD₅, and nitrogen content change during primary sedimentation due to the removal of suspended solids and organic matter. Similarly, pH is an important factor in the biological treatment process, but it does not have a significant effect on the primary sedimentation. As suggested in Equation 3, Gamal and Smith ^[27] proposed to use influent TSS concentration and inlet flow as input data to the TSS neural network. In the same way, the COD model used influent COD and inlet flow as input parameters to estimate effluent COD. Both neural network models reported acceptable results. The inlet flow rate is a critical variable in primary sedimentation because the removal efficiency of the settling process partially depends on the flow distribution and retention time of wastewater in the settling tank. Also, the performance of the clarifier tank is a function of the constituent influent concentration because it is fundamentally required to estimate the effluent parameters ^[1].

From a modeling standpoint, COD cannot differentiate between biodegradable and inert organic matter ^[57]. Additionally, both fractions are composed of particulate and soluble compounds (particulate and soluble COD, respectively) ^[1]. The particulate COD of sewage generally ranges from 30 to 70% of its total COD, but in domestic wastewater, particulate COD in some cases will represent 85% of total COD ^[58]. Consequently, a fraction of particulate COD will be removed, but the soluble and colloidal fraction of COD remains after the primary sedimentation process. Hence, additional information data was considered to provide more input information to neural models.

The addition of influent TSS concentration and influent turbidity as input data for the COD model was considered to supply an approximation of particulate and colloidal matter in sewage composition. Analogous to the COD model, influent COD and influent turbidity give complementary information about the sewage chemical composition, soluble compounds, and suspended matter. Therefore, influent COD and influent turbidity were considered as input data for the TSS model.

Although the temperature is an important factor in the wastewater treatment process, it has not a significant effect on primary settling tanks performance [1]. Therefore, the influent temperature was not considered as input data for modeling. Additionally, it presented a relatively constant value in the measurements of the entire data set. Conductivity was not considered because there exists a large number of cases where this parameter has been not registered. Finally, the data set for neural models included inlet flow, influent turbidity, influent COD, and influent TSS concentration as input parameters. For the output data, only was considered effluent COD and TSS concentration.

4.2. Data Processing for Artificial Neural Network Models

In ANN modeling, the efficiency of process data and their ability to accurately predict the target output is largely dependent on the relationship between the input data and the output desired [59]. Some values in the available data set were inconsistent, specifically, the measurements associated with influent and effluent TSS concentration. The removal efficiency in these particular cases presented negative values. These inconsistencies could be produced by the wrong performance of the sedimentation unit or mistakes during the laboratory analysis. Hence, these erratic values were discarded only in the data set used for the TSS model because they could affect the training process of the neural network model negatively. However, in the case of the COD model, these irregularities are not significant because effluent TSS concentration was not considered.

In contrast, some cases of incomplete data were found in the data set. The measurement of COD and TSS concentration mainly were not registered in many instances. Some studies available in the literature suggest the idea of using specific auxiliary parameters to estimate other process parameters that are not temporarily available. Hack and Kohne

found a very strong correlation between turbidity and the influent COD ^[60]. Therefore, turbidity data were used for estimating COD values on incomplete data. As COD estimation, there exists a strong linear relationship between turbidity and TSS concentration, which can be used for estimating TSS concentration using turbidity measurements ^[61]. However, in this study, it was considered a logarithm relationship because it showed a better mathematical adjustment for experimental data than a simple linear correlation. All these considerations were realized to maintain a larger amount of representative data samples as possible for the training phase. Hence, data sets of 50 and 53 samples were used to develop the TSS neural network and COD neural network, respectively. Additionally, a statistical description of data sets used for the training phase was developed to obtain a detailed insight into the general data used for the study.

4.3. Structure of Artificial Neural Network Model

A prototype was developed to find a priori which characteristics provide the best results in both neural network models. A preliminary test was developed in Excel to establish which parameters and activation functions get better results. From the initial model testing, several activation functions were considered. Five different activation functions by the hidden layer were used (sigmoidal, sine, cosine, hyperbolic tangent, and gaussian). Hyperbolic tangent (tanh) in the hidden layer and linear function in the output layer showed the best results during the initial test for both neural networks. Additionally, the same configuration of layers and activation functions produced satisfactory results reported by Gamal and Smith in the literature ^[27].

The Neural Network Toolbox of MATLAB R2019a software was used to design the architecture and to train the ANN models. A multi-layer-feed forward neural network was selected as the architecture for both neural models. A feed-forward network with only one hidden intermediate layer is widely used in the automation and control of processes ^[60]. Many researchers in the field of artificial intelligence suggest that it is usually unnecessary to use more than one hidden layer in multi-layer feed-forward architecture, and varying the number of nodes in the hidden layer is sufficient for delivering different results ^[27]. Consequently, in both cases, only one hidden layer and one output node were used in an MLP neural network architecture (Figure 9), and the number of hidden nodes

ranged from 5 to 10 in order to determine the optimum number of neurons in the ANN architecture.

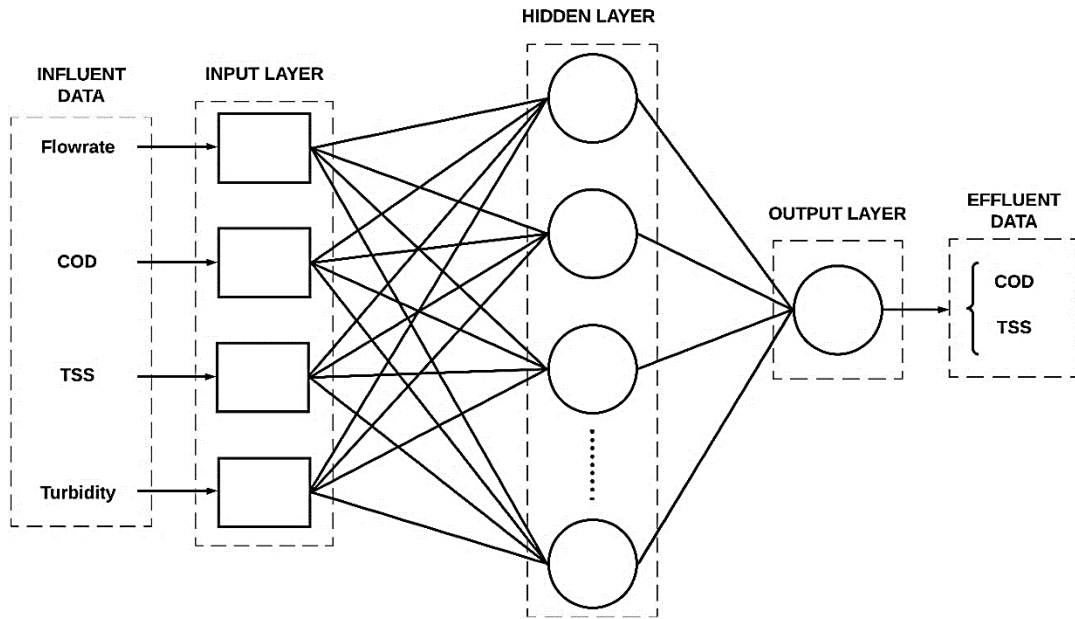


Figure 9. Multilayer-neural network architecture used for the TSS model and COD model.

4.4. Artificial Neural Network Model Training

The data set that is fed to the network has to be scaled into a numeric range that is efficient for the calculation by the neural network^[50]. If the data used do not scale to an appropriate range, the neural network will not converge on the training phase, or it will not produce accurate results^[62]. Therefore, the data set were scaled linearly (normalized) into the range [-1,1] before being presented to the network. Thus, network outputs also have to be scaled to the initial range to obtain the real value. The mathematical expression used for the normalization method is defined in Equation 11.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \cdot (r_{max} - r_{min}) + r_{min} \quad (11)$$

where x is the original value, x_{norm} is the normalized value, x_{max} and x_{min} are the maximum and minimum values of the concerned variable, respectively. Lastly, r_{max} and r_{min} are the desired values of the normalized variable range. The data set were divided into training, validation, and test set. A random division of available data is a satisfactory

method to generate the training, validation, and test data set ^[50]. Hence, 70% of the total data set was used for the training phase, 15% were used for the validation process, and the remaining 15% were used for model testing. This division has been used successfully and reported by Shahin *et al.* ^[63] in the literature.

The method for training MLP is based on the minimization of a suitable cost function ^[62]. The mean square error was selected as a cost function for this study. For prediction problems, supervised learning is frequently utilized for teaching the neural network how to relate input data patterns to output data ^[27]. Therefore, a supervised learning algorithm was used for the training phase of both neural network models.

LMA was used as a training algorithm; it is often the fastest backpropagation algorithm and is highly recommended as a first-choice supervised algorithm ^[51]. Moreover, batch-mode for adjusting weights (weights updates were done after each epoch and not after each training pattern) was used during the training phase. To compare the performance of different MPL architectures and select the best option, both statistical criteria MSE and R^2 were used. Additionally, the performance of selected models was compared to the traditional empiric models. The correlation between empirical model predictions and experimental data was determined. Additionally, the error associated with the model prediction produced by each model was calculated.

4.5. Sensitivity Analysis

One of the most important disadvantages of neural network models is the comprehension of internal relationships generated by the network ^[64]. In contrast to classical mathematical and statistical models, it is not evident the importance and influence that each input variable has over the model response. Consequently, the exact relationship between input and output data are not well understood. Different attempts have been developed to interpret the weights and internal operations in neural network models, but the sensitivity analysis (SA) is the most widely used ^[64]. Therefore, a sensitivity analysis was developed in this study to obtain information about the internal relationships established by the neural network model.

The sensitivity analysis studies how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input [65]. Hence, as a general idea, the sensitivity analysis is a method of determining the rate of change in model output concerning the change in model inputs [54]. It is often employed to quantify the importance of each of the model's parameters on the behavior of the system [66]. In the literature, many methods to estimate the sensitivity measurements have been proposed. In general, we can distinguish between local and global SA. The local SA addresses sensitivity relative to a change of a specific parameter value; it means that local SA focuses on a single input's behavior while other parameters remain constant [67]. The most common method to carry out a local SA is to set the value of all input variables to their mean value (Base Case) and vary the value of one of them throughout its range, with the object to observe the effect it has on the model response [64]. Thus, a local SA was considered in this study to understand the effect that input parameters generate on the ANN model response. A Base Case (BC) of both the TSS model and the COD model was established. The input value parameters were varied in the range from a reduction of 40% of the initial value to an increment of 40% of the initial value.

On the other hand, a global SA focuses on the variance of model outputs and determines how input parameters influence the output parameters [66]. Based on this idea, in a neural model study, Irigoyen *et al.* [68] defined the sensitivity measurement as the relationship between the error estimation with the missing variable and the original error estimation of the model. The error estimation can be measured as the MSE associated with the model. The formula used to determine the sensitivity of the input parameter using the MSE as error measurement is defined in Equation 12.

$$S_i = \frac{MSE_i}{MSE} \quad (12)$$

Where S_i is the sensitivity value of the missing input variable i , MSE_i correspond to the mean square error associated with the model performance with the missing input variable i , and MSE is the mean square error of the original model using all input parameters. Thus, more sensitive is the model to a specific input variable as the ratio between error estimations increases. In contrast with the previous method, the analysis proposed by Irigoyen *et al.* [68] provides a measurement of the robustness of the model. Hence, a global SA was considered to identify which are the most influential variable inputs for the ANN

model and evaluates the model performance under a specific combination of input variables. In order to obtain a complete sensitivity analysis, local and global SA was developed in this study applying the methodologies described previously.

4.6. Results and Discussion Section: Outline

A modular methodology was considered for the development of this study. Thus, a segmental division of the entire work was applied to establish a systematic method for the organized resolution of specific tasks. Therefore, to achieve a better understanding, the results obtained in this study, which are reported in Chapter V, were classified into sections. Hence, Chapter V is divided into the following six sections:

- *Section 5.1*: it corresponds to the adequation and processing of available data used for the ANN modeling. In this section, some correlations between wastewater parameters were established to complete the incomplete information using complementary data. Additionally, a descriptive statistic of the used data sets was incorporated.
- *Section 5.2*: it presents the results obtained during the training phase using different number of hidden nodes in ANN models. This section involves the R^2 and MSE values obtained with different options of ANN architectures. Hence, Section 5.2 is focused on determining the optimum number of hidden nodes in the ANN models architectures considering the R^2 and MSE values in each case.
- *Section 5.3*: it consists of the evaluation of the performance of selected models in the previous section. Therefore, the ANN model predictions were compared with the experimental data. Furthermore, correlation functions between experimental data and model predictions were established to evaluate the performance of the ANN model. In the same way, a comparison based on the error of model predictions between the ANN model proposed in this work and the empirical models reported in the literature is presented in this section.
- *Section 5.4*: it presents the sensitivity analysis of ANN models. This section describes the relationship between the input variables and the model response. In

the same way, this section shows the effect that each input variable has on the accuracy of the model output value.

- *Section 5.5*: it corresponds to the validation of ANN models. Hence, this section shows correlation functions between new experimental data supplied by Ibarra WWTP and ANN model predictions used to validate the accuracy of ANN model predictions.
- *Section 5.6*: it is destined to discuss some potential applications of this study in the wastewater treatment process.

Table 7 summarizes the organization of results obtained in this study, which will be presented in detail and commented in Chapter V.

Table 7. Results and Discussion Section: Outline.

Sections	Titles	Content
5.1	Data Processing for Artificial Neural Network Models	Correlation functions between wastewater properties Descriptive statistics of experimental data
5.2	Artificial Neural Network Models Training	R ² values with different hidden nodes. Variation of MSE value with different hidden nodes.
5.3	Performance of Selected Models	Correlation between experimental data and model predictions Average error of ANN model against empirical models Error distribution of ANN model against empirical models.
5.4	Sensitivity Analysis (SA)	Local SA Global SA
5.5	Model Validation	Correlation between model predictions and new experimental data.
5.6	Applications	Possible potential uses of the ANN model in the wastewater treatment process.

CHAPTER V

5. RESULTS AND DISCUSSION

5.1. Data Processing for Artificial Neural Network Models

In this study, a strong correlation between TSS concentration and turbidity measurements was found. The result obtained with the linear regression of the logarithmic relationship between turbidity and TSS concentration is shown in Figure 10 (b). It suggests a positive correlation between turbidity and TSS. This result is expected because turbidity is a measure of water clarity and a measurement of the suspended matter in wastewater. As the volume of suspended solids increases, the light is absorbed or scattered and results in higher values of turbidity [69]. However, the presence of dissolved, color-causing substances that absorb light may cause interference in the measurement [15]. These interferences could be affecting the relationship between TSS concentration and turbidity data.

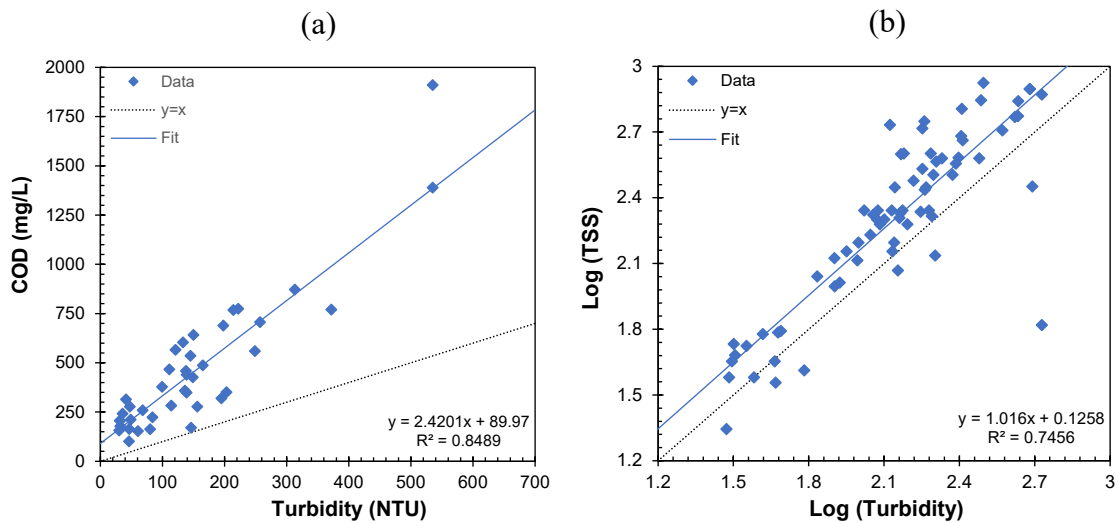


Figure 10. Correlation functions between wastewater properties.

The relationship between COD and turbidity is exposed in Figure 10 (a). Similar to TSS, the COD linear regression model suggests a rise in the COD as the turbidity increases. As was exposed previously, higher values of turbidity indicate a higher concentration of suspended solids. These suspended particles contain much organic matter, which is responsible for the COD of the wastewater. As a result, measuring the turbidity gives an estimation of COD. As Figure 10 shows, the turbidity/COD linear relationship achieved a better R^2 value compared to the turbidity/TSS logarithmic relationship. This result is produced due that turbidity provides an integrated measurement of suspended and dissolved particles^[69]. Dissolved substances are also included in the COD measurements. Thus, the interferences associated with turbidity measurements do not negatively affect the relationship between COD and turbidity. These interferences provide more general and integral information of pollutant concentrations, which is useful to estimate more precisely the COD in wastewater composition.

In summary, positive linear correlations between turbidity/TSS and turbidity/COD were found to be consistent. A slightly lower correlation coefficient is observed in this study for turbidity/COD relationship, and turbidity/TSS relationship compared to the values reported in the literature^[66,67]. However, the wastewater characteristics and conditions were not the same. These results will be used to develop other models or empirical relationships to estimate wastewater parameters. From an operational and management point of view, COD is an efficient monitoring parameter compared to BOD_5 analysis, because of its quick determination between 2-3 hours. However, it requires specific laboratory equipment and hazardous chemicals substances. Thus, a significant reduction of time and risk can be reached in estimating COD monitoring using the regression model based on turbidity data.

In the same way, the determination of TSS concentration required long periods due to the gravimetric method that it involves. Therefore, these correlations could be used as a complementary tool to improve the treatment process control. These results might be implemented and used as operational tools to obtain rapid estimations and detection of monitoring parameters.

These correlations between turbidity/TSS and turbidity/COD were used to complete a few cases where COD or TSS experimental measurements were not available in the data set. Then a statistical analysis of the data set was developed to obtain a general statistical

description. The descriptive statistical information about the data set used for the TSS model and the COD model are summarized in Table 8 and Table 9, respectively.

Table 8. Descriptive statistics of experimental data used for the TSS model.

Parameter	Mean Value	Maximum Value	Minimum Value	Standard deviation	Coefficient of variation
Flow rate (m ³ /h)	934	2748	193	495	0.53
Influent COD (mg/L)	589	1389	213	298	0.51
Influent TSS (mg/L)	336	800	48	216	0.64
Influent Turb (NTU)	210	535	51	123	0.59
Effluent TSS (mg/L)	71	228	20	40	0.56

Table 9. Descriptive statistics of experimental data used for the COD model.

Parameter	Mean Value	Maximum Value	Minimum Value	Standard deviation	Coefficient of variation
Flow rate (m ³ /h)	921	2748	193	466	0.51
Influent COD (mg/L)	550	1389	213	276	0.50
Influent TSS (mg/L)	330	960	22	241	0.73
Influent Turb (NTU)	197	535	50	113	0.58
Effluent COD (mg/L)	209	389	101	56	0.27

5.2. Artificial Neural Network Models Training

The training phase was carried on for 1000 epochs, but the training process was stopped when the validation data MSE value started to increase. This procedure avoids the overfitting problem. To find the best MLP architecture, different numbers of nodes in the hidden layer were used varying from 5 to 10. Every neural network model completed the training phase before 20 epochs. The R^2 values corresponding to the training phase of each candidate TSS model and each candidate COD model are shown in Table 10 and Table 11, respectively. Additionally, the corresponding notation used to identify each specific neural network architecture is detailed in the tables.

As one can observe in Table 10 and Table 11, the R^2 value of each neural network architecture used in the training phase was calculated for every subdivision of the entire data set (training, validation, and test). Subsequently, the same statistical parameter was determined for the whole data set (combined). In this study, a value of $R^2 \geq 0.5$ was considered a satisfactory result. In order to guarantee the generalizability of the model, the model must achieve at least a satisfactory R^2 value in all subdivisions of the data set.

In the case of the TSS model, all the options used in the training phase show an acceptable R^2 value in terms of training data. In the same way, R^2 values above 0.5 are obtained in all cases concerned with the validation data set (Figure 11). However, in the test data set, only the TSS-05 architecture was not able to achieve a satisfactory R^2 value. Therefore, the TSS-05 option was discarded. The R^2 value of each subdivision provided an idea of the performance of a particular neural network architecture about specific group data. However, there is not enough information to determine which architecture presents the best performance. For example, Table 10 shows that TSS-08 architecture reached the highest R^2 value in the training data. However, it did not reach the highest R^2 value corresponding to the validation and test data set.

On the other hand, despite the TSS-07 option not obtaining the highest value in the training data set, it achieved the highest value in terms of validation and test data set. Although specific information of each subdivision of the data set is useful to evaluate the performance of the model initially, it is necessary to determine the performance of each model about the entire data set (combined).

Therefore, the R^2 value of the entire data set was determined to establish which number of nodes in the hidden layer provided the best performance in general terms. The highest R^2 value concerned with the total data set was achieved with the TSS-07 architecture. In contrast, the TSS-06 was discarded because it did not reach a satisfactory R^2 value in terms of the entire data set.

Table 10. R^2 values for the TSS model with different number of hidden nodes.

Notation	N° of Nodes	Training	Validation	Test	Combined
TSS-05	5	0.63	0.63	0.49	0.58
TSS-06	6	0.56	0.60	0.84	0.44
TSS-07	7	0.57	0.78	0.86	0.60
TSS-08	8	0.66	0.67	0.62	0.53
TSS-09	9	0.58	0.70	0.62	0.57
TSS-10	10	0.54	0.56	0.66	0.51

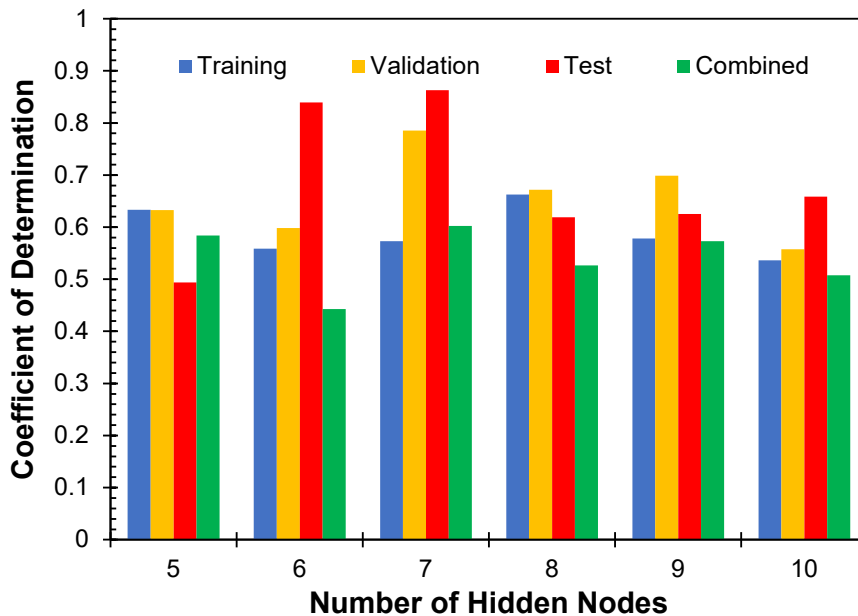


Figure 11. R^2 values for the TSS model with different number of hidden nodes.

In the case of the COD model, all the proposed architectures reached a satisfactory R^2 value in terms of training, validation, and test data set. The COD-09 model showed the highest R^2 value corresponding to the training and test data set. However, COD-09 architecture did not reach the highest value of R^2 in general terms. This fact is associated with the R^2 value obtained in the validation data set, which is the second-lowest. On the other hand, as it is observed in Figure 12, the COD-10 architecture obtained similar values of R^2 in all subdivisions of the data set.

Moreover, the COD-10 architecture achieved the second-highest values in the training, validation, and test data set. Consequently, this option obtained the highest R^2 value in terms of the entire data set. Nevertheless, COD-09 and COD-10 achieved similar results of R^2 value corresponding to the total data set. Otherwise, the COD-07 option was discarded because it did not provide an acceptable result about the combined data set. Despite the COD-07 model reached the highest R^2 value in the validation data set, it obtained the lowest R^2 value in terms of the training and test data set. Therefore, this has a significant effect on the R^2 value obtained in the entire data set.

Table 11. R^2 values for the COD model with different number of hidden nodes.

Notation	N° of Nodes	Training	Validation	Test	Combined
COD-05	5	0.61	0.77	0.62	0.57
COD-06	6	0.56	0.56	0.79	0.50
COD-07	7	0.54	0.89	0.60	0.47
COD-08	8	0.55	0.64	0.76	0.55
COD-09	9	0.65	0.56	0.86	0.58
COD-10	10	0.61	0.62	0.69	0.60

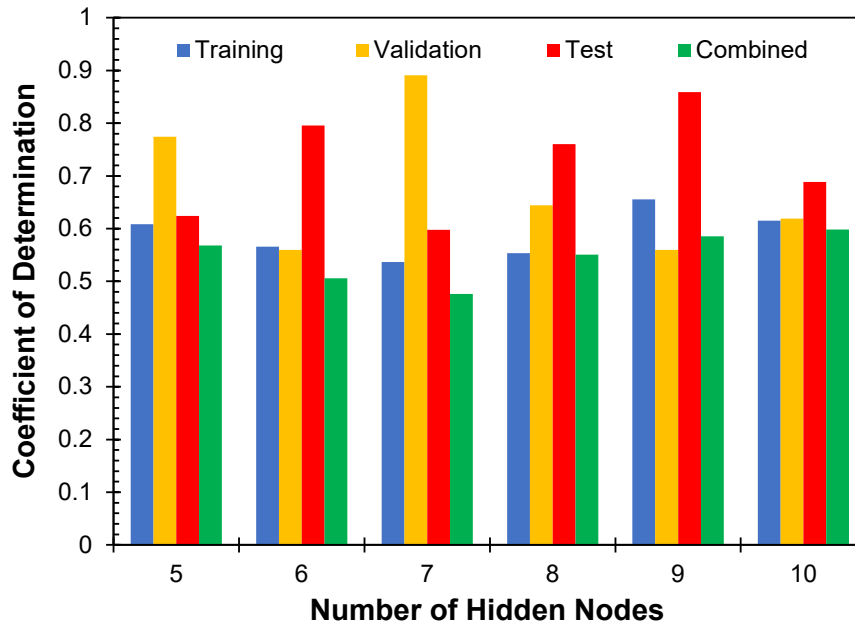


Figure 12. R^2 values for the COD model with different number of hidden nodes.

Although R^2 has been widely used for model evaluation, this statistical indicator is oversensitive to high extreme values and insensitive to additive and the proportional difference between model predictions and measured data [54]. Das and Sivakugan [71] explain that use only R^2 can be confusing because that higher values of R^2 might not necessarily indicate better model performance due to the tendency of the model to deviate towards higher or lower values in a wide range of data set. Therefore, if R^2 is used for model evaluation, it is necessary to take into account additional information that can cope with that problem [72]. Hence, the MSE value was also considered to evaluate model performance. The MSE values of the normalized total data set corresponding to the training phase are shown in Table 12.

Table 12. MSE value for model performance with different hidden nodes.

N° of Nodes	TSS Model	COD Model
5	0.060	0.063
6	0.082	0.084
7	0.057	0.085
8	0.070	0.066
9	0.062	0.061
10	0.071	0.060

Note: The MSE values correspond to the normalized data set.

The lower MSE values mean better performance of the network (zero means no error). Figure 13 shows that the minimum MSE value of the TSS model was obtained using seven hidden nodes. In the case of the COD model, the minimum MSE value was reached using ten nodes in the hidden layer.

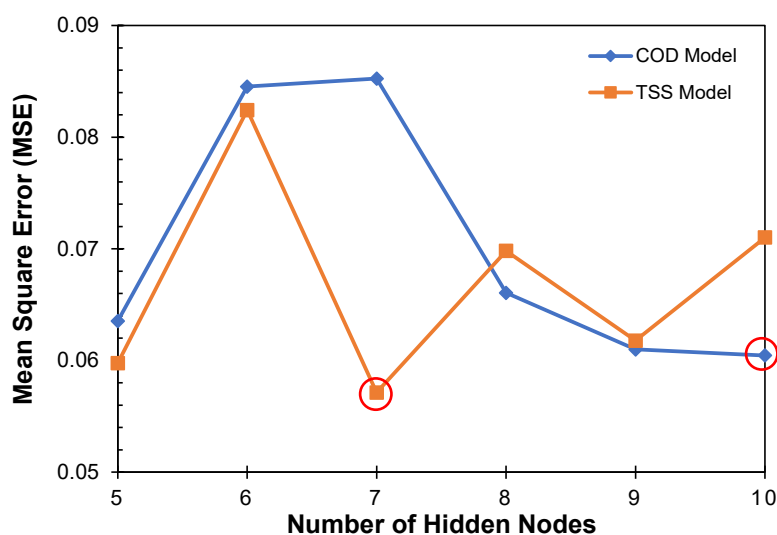


Figure 13. Variation of MSE value for model performance with different hidden nodes.

Shahin ^[73] and Eidgahee *et al.* ^[74] established that the best measure for the performance of neural network models should be based on the highest values of R^2 and the lowest MSE values. According to these considerations, the TSS-07 architecture and COD-10 achieved the best performance. Therefore, the TSS-07 model was selected as the best option to

predict TSS concentration effluent. However, the simplicity of the model is an essential factor that should be considered in model selection. Kolay *et al.* [75] concluded that using more than ten nodes in the hidden layer will cause the saturation of the neural network, which results in lesser quality simulated results due to undesirable feedback to the network. Particularly in the COD model, despite the COD-10 architecture achieving the highest R^2 value and the lowest MSE value, the COD-09 architecture obtained similar results with less nodes. Therefore, in order to avoid saturation problems, the COD model of nine nodes in the hidden layer was selected as the best option to predict effluent COD. The values of weights and bias obtained during the training phase of ANN models are detailed in Appendix B.

In a similar study about primary settling tank modeling, Gamal and Smith [27] reported that the neural network structure that used nine hidden nodes yielded the best performance to predict effluent TSS concentration and effluent COD. In contrast, in this study, the architecture of seven hidden nodes showed the best performance for the TSS model. This is a significant result because a simpler model will be more efficient and easier to implement in process control. In the case of the COD model, the structure of 10 hidden nodes achieved the best performance. However, despite the additional hidden node, the COD-10 architecture did not represent a significant improvement of MSE value and R^2 compared to the COD-09 architecture. It should be mentioned that the neural network models developed by Gamal and Smith [27] used a more extended data set, another algorithm for the training phase, and the wastewater characteristics and conditions could not be similar.

5.3. Performance of Selected Models

Figure 14 corresponds to the comparison between COD model predictions and the experimental data used as the data set. The COD model predictions fit very well the experimental data in general terms. In many cases, the model provides very accurate predictions or at least a good approximation. However, there are some particular cases which are not fitted by the model. Particularly, the COD model fails to predict the extreme values of effluent COD. A sudden change in the wastewater composition will produce considerable incorrect predictions [60]. Effluent discharged to the primary settling tank is often quite variable due to stormwater and intermittent discharges from an industrial plant

[31]. Due to the mixed sewage system in Ibarra, industrial effluents are also treated in the Ibarra WWTP. Consequently, these extreme cases will be associated with a higher concentration of pollutants in industrial discharges, which considerably affect the chemical composition of untreated wastewater. As a result, the COD fractions are also altered significantly.

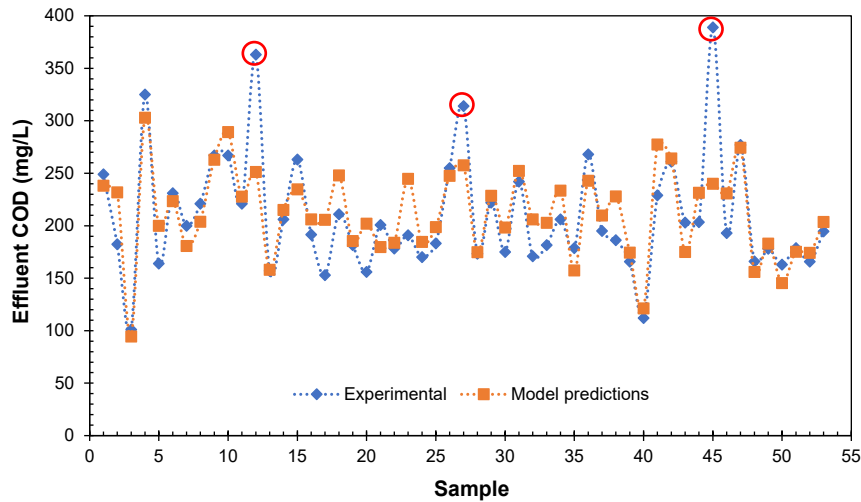


Figure 14. Performance obtained by the COD model for the total data set.

The comparison between TSS model prediction and experimental data set is presented in Figure 15. Similar to the COD model, the TSS model fits very well with the experimental data. However, the TSS model fails to predict the extreme values of effluent TSS concentration. These extreme values could affect the learning process of both neural network models significantly. There are few cases of extreme values about effluent COD and TSS concentration in the available data set. Therefore, due to the limited data set used in the training phase, both neural network models were not trained with enough examples to generalize these extreme cases. However, the COD model and the TSS model were able to generalize and fit acceptably the rest of the samples of the data set used in the training phase. Therefore, it should be considered the application range of ANN models based on the range of data set used for the training process.

It also should be considered that the quality of data is an important factor that can considerably affect model performance. The errors associated with experimental measurements also could cause important differences between experimental data and model predictions. Especially in the case of TSS analysis, the gravimetric method used to determine solids can be associated with several sources of errors as systematic, random, and human error.

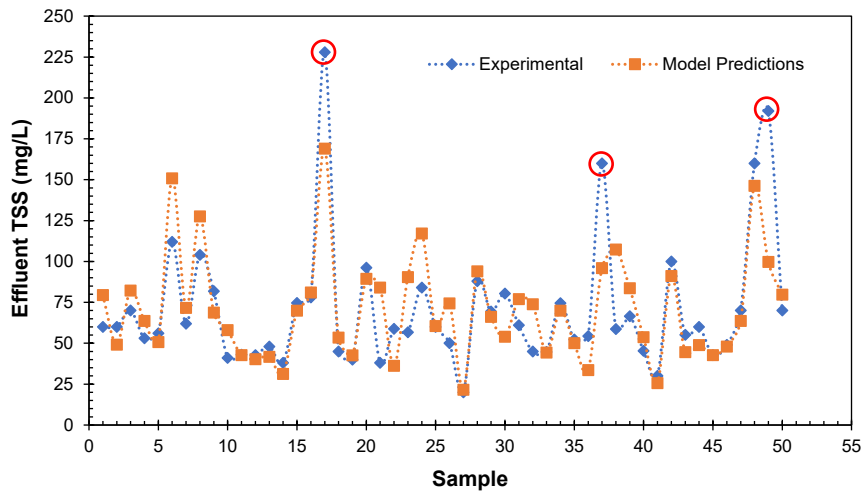


Figure 15. Performance obtained by the TSS model for the total data set.

The prediction ability of neural network models developed was tested against the empirical regression model proposed by Tebbutt and Christoulas (Equation 3). A nonlinear regression method was used to estimate the values of parameters D, F, and G. The overflow rate and removal efficiency were calculated to adapt the available data to the empirical equation. The R^2 value obtained by the empirical regression model for the effluent TSS concentration was 0.0014, which is very low compared with the value obtained by the TSS neural network model ($R^2= 0.60$). The empirical model for the effluent COD showed an R^2 value of 0.15, which is also much lower than that of the COD neural network model ($R^2=0.58$). A more detailed description of empirical regression models is shown in Appendix C. The relationship between TSS model predictions and experimental data is observed in Figure 16 (a), and the COD model and experimental data is shown in Figure 16 (b). If there is a perfect agreement between the model and experimental results, all the points will lie along the 45° line ($y=x$). In both Figure 16 (a) and Figure 16 (b), it can be seen that the values simulated by models spread around the 45° line, which implies neither overestimation nor underestimation [76].

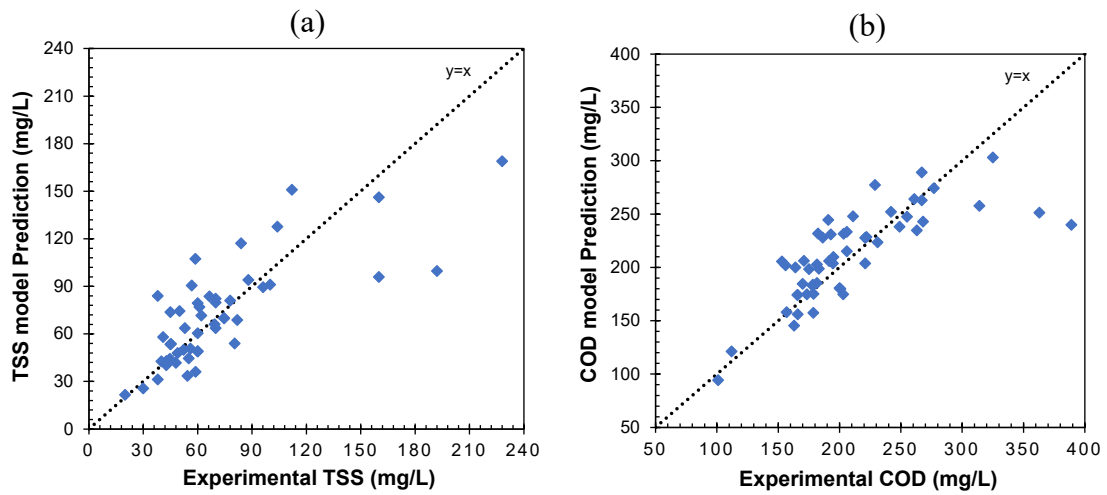


Figure 16. Correlation between experimental data and model predictions.

The error distribution of models, in terms of the percentage difference between model prediction and experimental measurement, was determined (Figure 17). The average error of the TSS model and the COD model for predicting experimental measurements is equal to 22.87% and 11.12%, respectively. In comparison, the average error for the empirical TSS model and the empirical COD model is 42.42% and 18.24%, respectively.

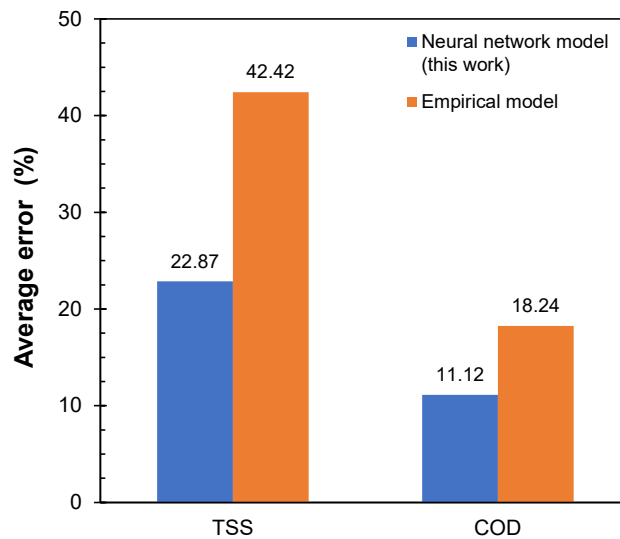


Figure 17. Average error of neural network models against empirical models.

The error distribution of TSS models and COD models are summarized in Table 13 and Table 14, respectively. From Table 13, it is observed a reasonably acceptable performance of the TSS neural network model. More than 60% of TSS model predictions are within the error range from 0 to 20% of experimental values, while less than 40% of the empirical

TSS model predictions are within the same range. In Figure 18, it is observed a significant difference between the accuracy of the TSS model and the empirical TSS model. A considerable difference in the error range from 0 to 5% is evident in Figure 18. In the same way, in the error range from 5 to 10%, the TSS model gives better results than the empirical TSS model. Additionally, the predictions of the TSS model are principally accumulated in the error range from 5 to 20%, while the empirical TSS model predictions are concentrated in the range from 20 to 30%.

Table 13. Distribution of error for the TSS model versus the empirical TSS model.

Range of error (%)	Number of data in the range for models		Percentage of whole data (%)	
	TSS model	Empirical TSS model	TSS model	Empirical TSS model
$X \leq 5$	8	2	16	4
$5 < X \leq 10$	11	4	22	8
$10 < X \leq 20$	12	11	24	22
$20 < X \leq 30$	5	8	10	16
$30 < X \leq 40$	6	5	12	10
$40 < X \leq 50$	4	7	8	14

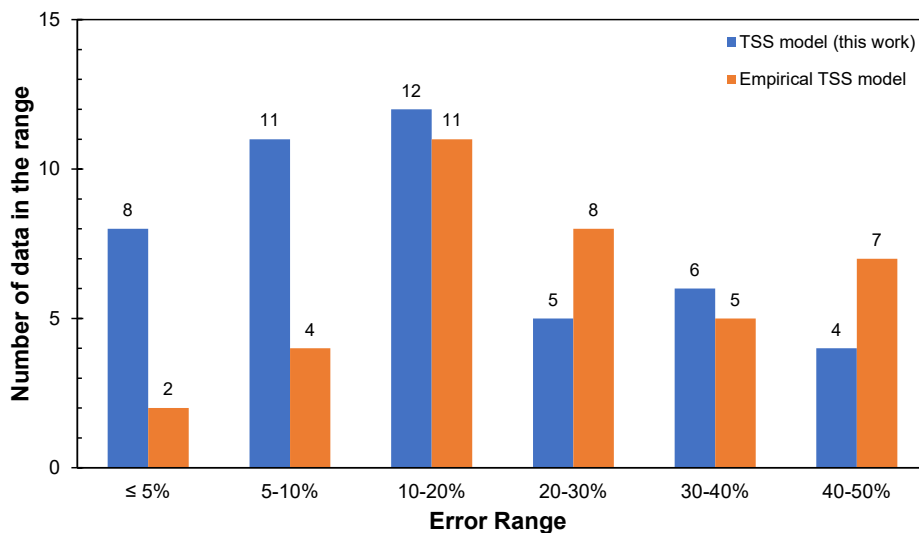


Figure 18. Distribution of error for the TSS model versus the empirical TSS model.

The comparison of predicted values obtained by the TSS model and the empirical TSS model is shown in Figure 19. It exposes a high difference in the accuracy of predictions between both models. The TSS model predictions are mainly concentrated within the $\pm 20\%$ error range. While the predictions of the empirical TSS model show a higher error dispersion than the TSS model. This is a significant indication that the TSS neural network model has learned to generalize the information better than the traditional TSS empiric model.

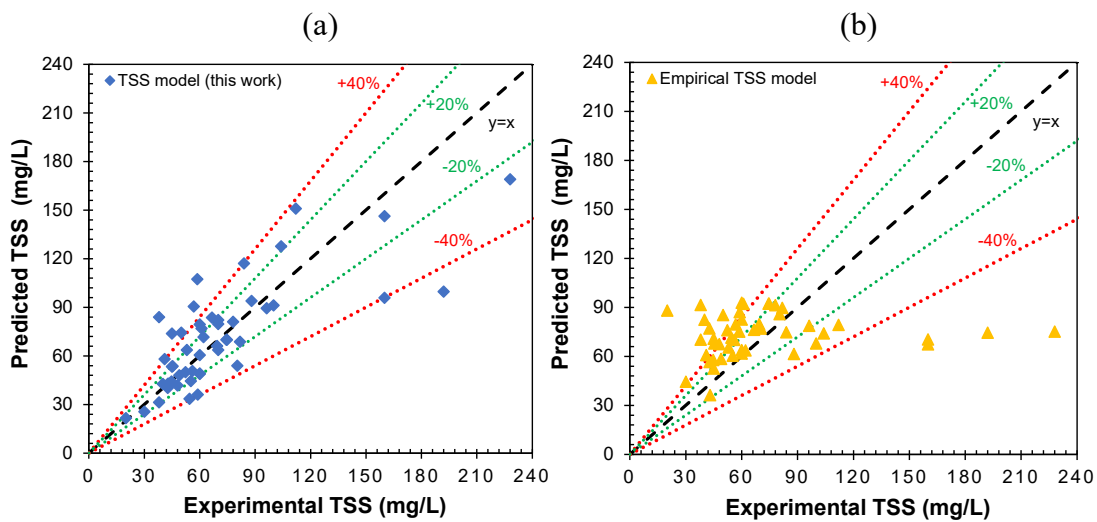


Figure 19. Comparison of predicted values of effluent TSS versus experimental data for different models.

For the COD model, from Table 14 is observed that more than 80% of COD model predictions are within the error range from 0 to 20%. In contrast, approximately 62% of empirical COD model predictions are inside in the same range. It indicates a significant difference in predicting ability between both models. Figure 20 exposes better results of the COD model in the range from 0 to 10% in comparison to the empirical COD model. It means that more values with higher accuracy are obtained by the COD model.

Table 14. Distribution of error for the COD model versus the empirical COD model.

Range of error (%)	Number of data in the range for models		Percentage of whole data (%)	
	COD model	Empirical COD model	COD model	Empirical COD model
$X \leq 5$	17	8	32	15
$5 < X \leq 10$	14	7	26	13
$10 < X \leq 20$	12	18	23	34
$20 < X \leq 30$	7	11	13	21
$30 < X \leq 40$	3	4	6	7
$40 < X \leq 50$	0	4	0	7

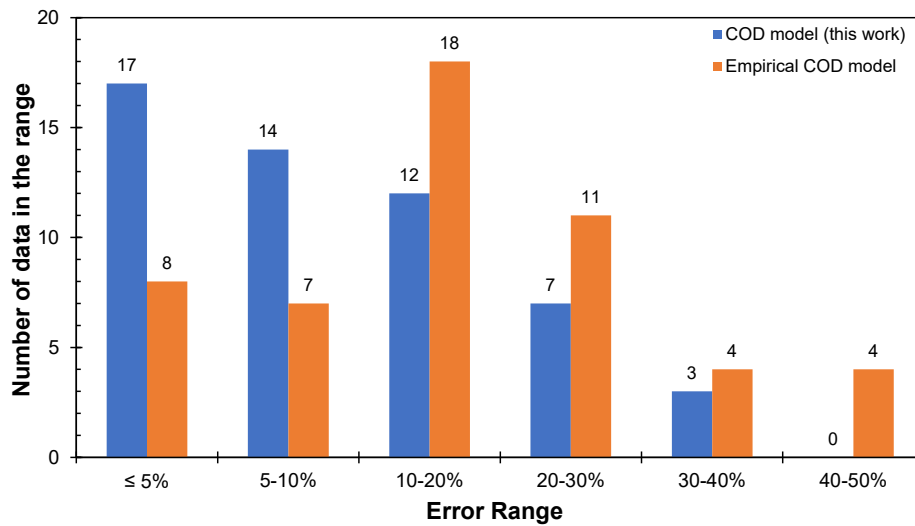


Figure 20. Distribution of error for the COD model versus the empirical COD model.

Figure 21 corresponds to the comparison of predicted values obtained by the COD model and the empirical COD model. It exposes that the COD model reached a better performance in predicting experimental values in comparison to the traditional COD empiric model. The COD model predictions are mainly accumulated within the $\pm 20\%$ error range. In contrast, the predicted values of the empirical COD model expose a higher error dispersion than the COD model. Therefore, the COD model provided more accurate predictions than the empirical COD model.

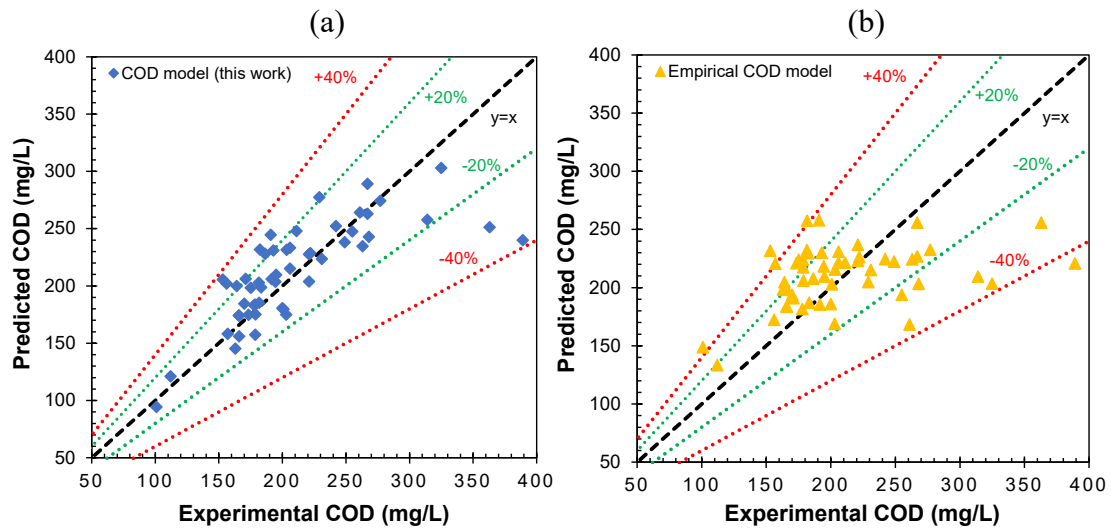


Figure 21. Comparison of predicted values of effluent COD versus experimental data for different models.

Actually, these results suggest an acceptable performance of both neural network models in predicting experimental values. Additionally, neural network models give better results than traditional empirical models.

The previous statistical analysis provides essential information about the error distribution of models. In Figure 18 and Figure 20 is observed that the error distribution of neural network models and empirical models could be described by a normal distribution. Therefore, a normal distribution was used to describe and compare the error distribution of each model. The statistical parameters used in each case of TSS models and COD models are detailed in Table 15 and Table 16, respectively.

Table 15. Statistical description of error distribution produced by TSS models.

Parameter	TSS model (%)	Empirical TSS model (%)
Mean value (μ)	22.87	42.42
Standard deviation (σ)	22.98	51.01
Max value	121.00	340.04
Min value	0.04	2.66

Table 16. Statistical description of error distribution produced by COD models.

Parameter	COD model (%)	Empirical COD model (%)
Mean value (μ)	11.12	18.24
Standard deviation (σ)	9.43	13.03
Max value	38.30	51.48
Min value	0.75	0.88

The normal distribution of TSS model errors and empirical TSS model errors is shown in Figure 22. It is observed an important reduction of average error using the TSS model in comparison to the traditional empiric TSS model. Additionally, the TSS model shows a significant reduction of error dispersion. The errors produced by the empirical TSS model predictions are widely dispersed. As a result, the variation range of prediction error will be greater for the empirical TSS model. In contrast, the TSS model errors are principally concentrated around the mean value (22.87%). These results suggest that the TSS model provides predictions with less variability and more accuracy than the empirical TSS model.

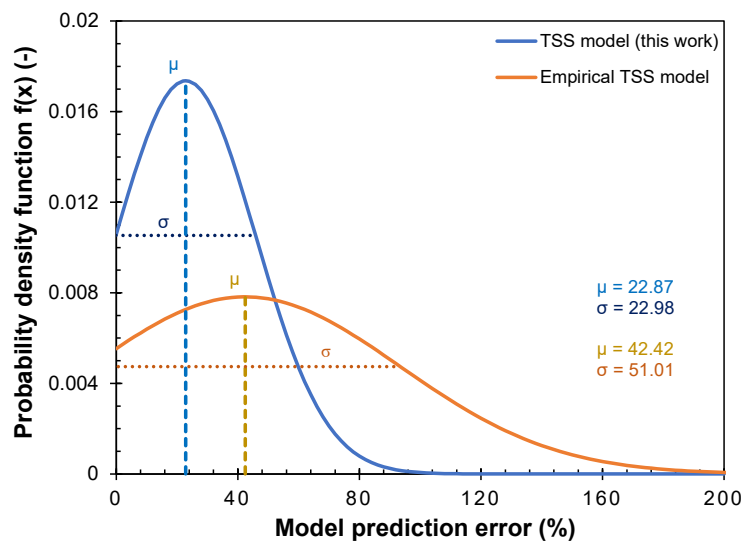


Figure 22. Normal distribution of TSS model errors and empirical TSS model errors.

Figure 23 corresponds to the comparison between the COD model errors and empirical COD model errors. It shows that the COD model reached a significant reduction of average error in comparison to the empirical COD model. In the same way as the TSS model, the COD model also obtained a considerable reduction of error dispersion compared to the empirical model. The COD model predictions are mainly concentrated around the mean value (11.12%). These results indicate that the COD model developed a better prediction ability than the empirical COD model. In both cases, neural network models achieved better performance than traditional empirical models. A significant reduction of average error and error dispersion related to model predictions were obtained using ANN models.

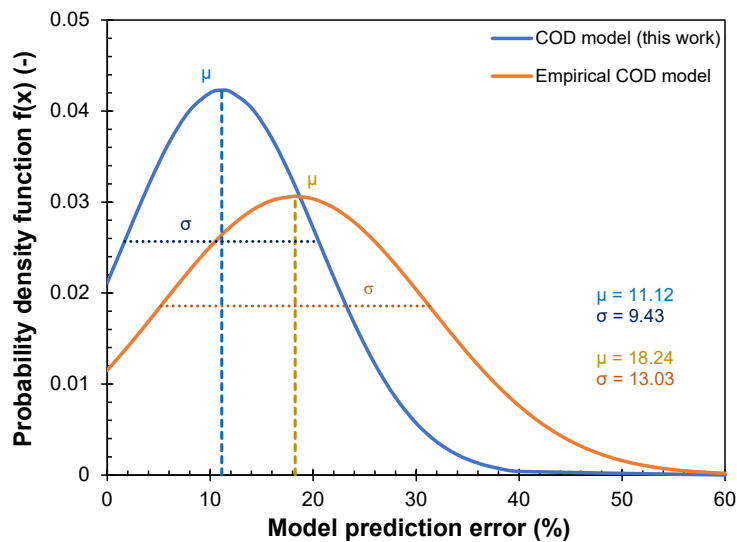


Figure 23. Normal distribution of COD model errors and empirical COD model errors.

5.4. Sensitivity Analysis

5.4.1. Local Sensitivity Analysis

A representative Base Case for each ANN model was selected to develop the local SA. Although the literature suggests establishing a hypothetical BC using the mean values of each input parameter, in this study was considered a real representative case of the data set as the BC. The reason is that using a real case of the data set, the base case is composed of coherent values of wastewater composition. In contrast, the BC designed by mean values could present some inconsistencies between input data due to the different

statistical deviations associated with each input variable. The input parameters were varied in a range from a reduction of 40% of the initial value to an increment of 40% of the initial value. Additionally, the average error bars related to the model prediction were included in both the TSS model and the COD model. An average error of 20% and 10% were used in the TSS model and the COD model, respectively.

Figure 24 corresponds to a schematic representation of the relationship between input parameters and the TSS model response. In the case of the TSS model, Figure 24 shows a tendency that indicates a proportional relationship between input parameters and TSS model predictions. It means that an increment of any input variable produces an increment of the TSS model prediction value. In Figure 24 is illustrated that the TSS model suggests that an increment of inlet flow rate produces higher values of TSS concentration in the effluent. The tendency shows that the TSS model response reaches a plateau or constant value as inlet flow increases. The same behavior is observed when the inlet flow decreases.

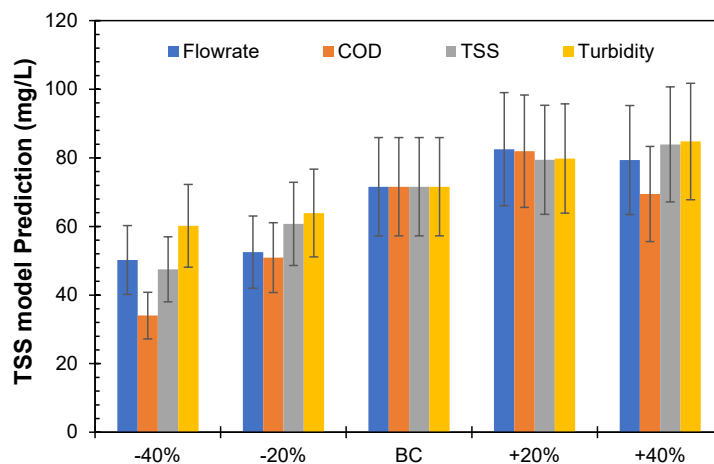


Figure 24. Local SA analysis of the TSS model.

Inlet flow rates greater than the design overflow rate, resulting in an increase in the effluent solids in practice^[31]. Several factors produce this fact. First, high inlet flow rates reduce the detention time of wastewater in the PST. Consequently, the suspended particles have less time to settle. On the other hand, a flocculent settling mechanism commonly occurs in primary sedimentation, and the removal efficiency of flocculant particle removal is dependent on both overflow rate and detention time^[11]. The removal efficiency improves as detention time increases because the suspended particles have a greater opportunity to coalesce and form flocs, which increases the settling velocity of

the particles, and they settle at the tank bottom.

Second, the performance of the settling tank is strongly influenced by the effectiveness of energy dissipation at the inlet to avoid turbulent flow^[77]. Thus, the inlet flow is uniform across the cross-sectional area of the tank as it enters the settling zone, and it generates quiescent conditions to allow the sedimentation of suspended solids. Hence, high inlet flow rates reduce the efficiency of energy dissipation and do not allow to achieve a uniform flow in the settling tank. Furthermore, short-circuiting or density currents occur when the flow through the tank is not uniform, and a current carries the particulate material to the effluent launders before the particles can settle^[77]. Hence, the flow velocity through the tank should be kept sufficiently low to avoid the resuspension of settled particles^[1]. Low inlet flow rates avoid these problems in settling tanks. Consequently, the TSS concentration in the effluent decreases as the inlet flow rate decreases.

The COD input data has a direct relationship with the effluent TSS concentration according to the model predictions. Figure 24 shows that TSS model prediction values increase as the COD of the influent increases. However, it is also observed a slight reduction of the output value in the case of a 40% increment of influent COD. It suggests that with a higher value than an increment of 40% of input COD value, the COD model prediction value will decrease. Hence, the COD model response reached a maximum value at an increment of 20%. This fact could be related to the nature of suspended particles and pollutants in wastewater. Despite this particular case, the tendency shows a proportional relationship of influent COD and TSS model prediction, which is the result expected due to the proportional relationship between COD and TSS concentration. In the same way, a constant reduction of the TSS model prediction value is obtained as the COD input value decreases.

Due to the direct relation between TSS concentration and turbidity, it is expected that both variables produce a similar effect on TSS model performance. The tendency demonstrates in Figure 24 that the TSS model prediction value increases as influent turbidity and initial TSS concentration increase. It shows a constant increment of the TSS model response in all the input variation range. This result is consistent because high values of TSS or turbidity corresponds to high concentrations of suspended particles in the wastewater. In a determined operational condition of the settling tank, specific

removal efficiency of TSS will be obtained. Thus, an increment of suspended particles will produce an increment of TSS concentration in the effluent.

In general, the TSS model is consistent with the settling tank performance in practice. It is an important result because models should be according to the properties of the modeled system. Additionally, Figure 24 exposes similar values of TSS model predictions produced by the variations of input parameters. This result indicates that every input variable has a similar level of influence on the TSS model performance.

Figure 25 corresponds to the relationship between COD model response and variation of input parameters. It is observed in Figure 25 that the inlet flow rate and TSS concentration of the influent not have a considerable effect on the COD model prediction. The COD model prediction value remains relatively constant through the entire variation range of flow rate and TSS concentration. Particularly in the case of the inlet flow rate, it is observed no change in the COD model response. Therefore, the inlet flow rate parameter does not have a relevant influence on COD model performance. This result is not consistent with the normal performance of the settling tank in practice. As was described before, the inlet flow rate is an important factor in the settling tank performance. The TSS removal efficiency highly depends on the inlet flow rate, and it will increase as the inlet flow rate decrease. Due to the proportional relationship between TSS concentration and COD, a reduction of TSS concentration generates a reduction of COD in the wastewater. Therefore, the behavior described in Figure 25 about the inlet flow rate is not the result expected based on the performance of the settling tank.

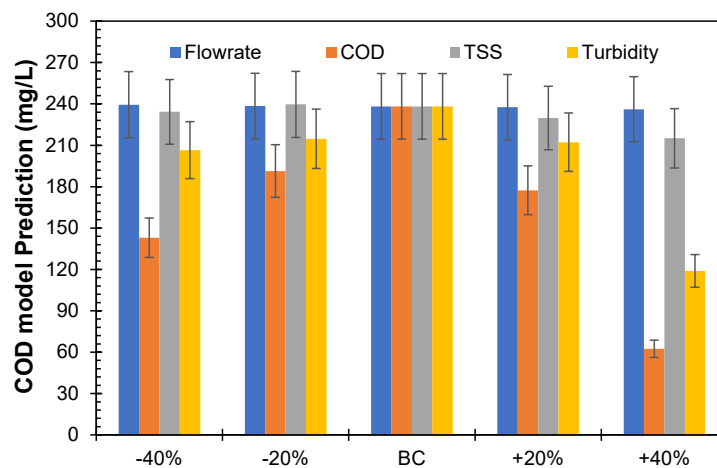


Figure 25. Local SA analysis of the COD model.

On the other hand, a slight reduction of the effluent COD produced by an increase of TSS concentration of the influent is observed. Thus, the trend indicates that an increment of the TSS concentration value of input data generates a reduction of the COD model prediction value. This particular case will be associated with a case of high particulate COD fraction in the influent composition. Thus, according to the COD model behavior, the removal of COD in the wastewater increases as TSS concentration increases. However, this particular result is not always right in practice. Despite high initial TSS concentration promotes flocculent sedimentation in the settling tank, several factors also determine the sedimentation mechanism. These several factors determine different scenarios of COD removal efficiency. In contrast, a reduction of the TSS input data does not produce a significant change in the COD model response.

Figure 25 also illustrates a not proportional tendency of the effect produced by influent COD and turbidity on the COD model predictions. In both cases, any variation in the COD or turbidity input parameter causes a reduction in the COD model prediction value. Both variables produced the same behavior in the COD model response. This outcome is associated with the direct correlation between both parameters. Instead, the variation of both input variables generates the same behavior on the COD model; the variation of the COD input parameter showed a more pronounced effect. Furthermore, this result suggests that COD model prediction principally depends on input COD value. In contrast, the tendency in Figure 25 shows that a gradual reduction of turbidity will produce that COD model prediction reaches a constant value. Input parameters produce a different effect on the model response. In some cases, ANN model responses are not coherent with the expected behavior observed in practice. This fact is related to the internal relationship developed during the training phase. In order to reach the best fit for experimental data and reduce the error function, ANN models establish complicated relationships between variables following a specific algorithm without physical considerations. As a result, the performance of ANN models usually differs from the physical behavior of the modeled system. The effect of each input parameter on the TSS model and the COD model response is summarized in Table 17.

Table 17. Influence of input parameters to neural network models.

Increases in the input variable	TSS model response	COD model response
Flow rate	+ (Reaches a plateau)	= (No change)
COD	+ (Reaches a maximum)	- (Sudden reduction)
TSS	+ (Constant increment)	- (Slight reduction)
Turbidity	+ (Constant increment)	- (Constant reduction)
Decreases in the input variable	TSS model response	COD model response
Flow rate	- (Reaches a plateau)	= (No change)
COD	- (Constant reduction)	- (Sudden reduction)
TSS	- (Constant reduction)	= (No change)
Turbidity	- (Constant reduction)	- (Reaches a plateau)

Note: (+) increasing effect, (-) decreasing effect, (=) no effect.

5.4.2. Global Sensitivity Analysis

The global sensitivity analysis of the TSS model and the COD model is exposed in Figure 26 and Figure 27, respectively. Figure 26 relates the sensitivity (S) value obtained to each input parameter according to the MSE associated with the TSS model performance with the corresponding missing input variable. According to Figure 26, the maximum S value is generated by the COD input parameter. Therefore, the TSS model is most sensitive to changes in the COD input variable. Consequently, it indicates that the absence of COD input data will generate the case of the highest estimation error in the TSS model. Thus, COD is the most influential input variable for TSS model predictions. Whereas, influent TSS concentration and turbidity produce a similar effect to the model response. Both input parameters obtained similar S values, which indicate that the TSS model is relatively equally sensitive to both input parameters. On the other hand, the minimum S value is linked to the inlet flow rate. It means that the TSS model is less sensitive to

variations of inlet flow rate data compared to the rest of the input data parameters. This is not an expected result due to the importance of the inlet flow rate in the TSS removal efficiency. Despite the difference of S values between input parameters, any input variable produces a similar magnitude increment of the estimation error of the TSS model.

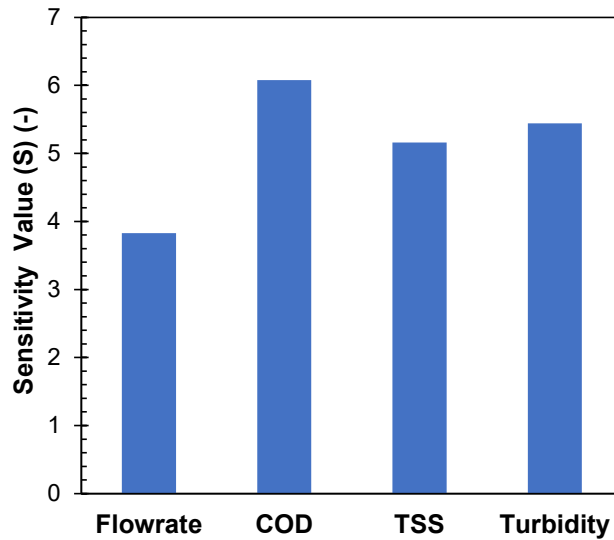


Figure 26. Global SA of the TSS model.

In the case of the COD model, there is a large difference between the S values of each input parameter. Figure 27 shows that the most influential input variables are turbidity and COD. The highest S value is associated with the influent COD, so the COD model is most sensitive to variations of the COD input parameter. The influent COD is the most important variable in the COD model because it exists a strong correlation between the influent and effluent COD [27]. Hence, the COD model indicates that effluent COD mainly depends on the influent COD. However, the COD and turbidity input parameters obtained similar S values. This result will be associated with the strong correlation between both variables. Therefore, these input parameters are essential to the correct performance of the COD model. In contrast, the TSS concentration and inlet flow rate do not have a significant effect on the model response in comparison to the COD and turbidity input parameters. Thus, the absence or variation of inlet flow rate or TSS concentration data will not generate a significant change in the model prediction value.

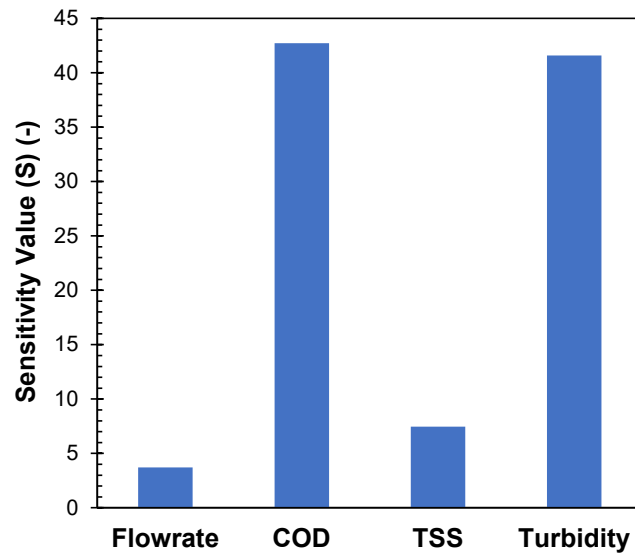


Figure 27. Global SA of the COD model.

In both cases, the COD model and TSS model show less sensitivity to variations of the flow rate input parameter. It suggests that model predictions mainly depend on the wastewater composition in both cases. Furthermore, the estimation error associated with the TSS model and the COD model was less when the whole input parameters about influent composition were used. However, the TSS model exposes higher stability to changes in COD and turbidity in comparison to the COD model.

5.5. Model Validation

Ibarra WWTP provided new experimental data that was used to evaluate the performance and accuracy of model predictions. A data set of 10 patterns was conditioned to simulate the performance of PST and predict the effluent characteristics. The range of application of both models is limited by the max and min value of the training data set. It should be mentioned that experimental values outside this range will not be predicted correctly by the models. Similarly, due to the error distribution obtained from the training data, an acceptable error range of $\pm 20\%$ was considered as the error expected for model predictions in both cases. Figure 28 corresponds to the relationship between the ANN model predictions and the new experimental data.

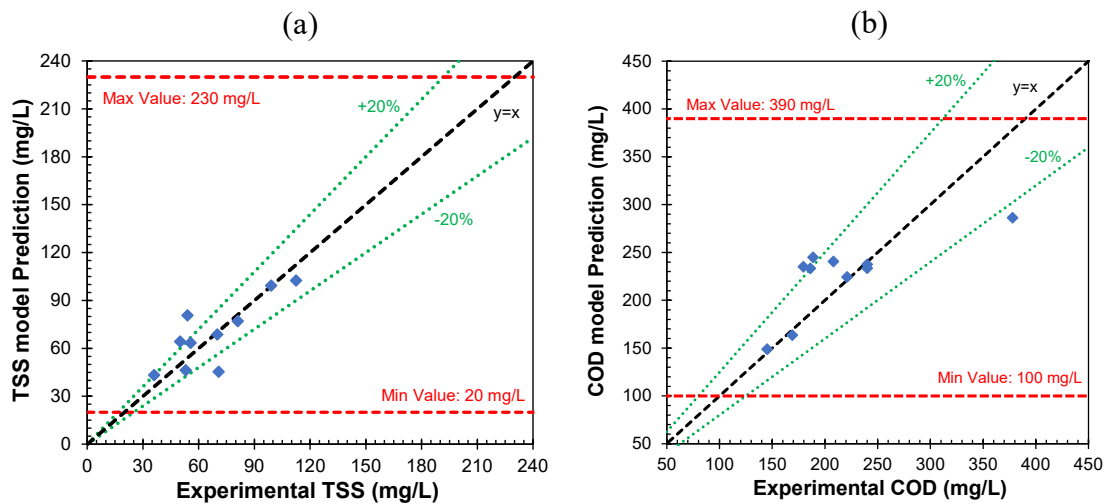


Figure 28. Correlation between model predictions and new experimental data.

The performance of the TSS model is exposed in Figure 28 (a). The TSS model obtained satisfactory results in general terms. As Figure 28 (a) shows, only two samples of the data set are outside the $\pm 20\%$ error range. Moreover, some of the samples lie along the 45° line, which implies a perfect agreement between the TSS model predictions and experimental data. Hence, in general, the TSS model provides good approximations of TSS concentration effluent.

Figure 28 (b) corresponds to the relationship between COD model predictions and experimental data. Only one pattern of the data set is completely outside the $\pm 20\%$ error range. In contrast, two samples are within the limit of the range, so they are inside the

expected error range. As it is observed in Figure 28 (b), seven patterns of the data set are totally within the ± 20 error range. Furthermore, five of these patterns were predicted with high accuracy by the COD model.

In general, both models achieved satisfactory performance. Furthermore, in both cases, the R^2 values obtained with this new data set are similar to the R^2 values registered in the training phase. For the TSS model, it obtained an R^2 value equal to 0.64, while the COD model reached an R^2 value equal to 0.56 for the new data set. It should be mentioned that the neural network models have been trained on a very limited data set; however, they still were able to provide good predictions.

5.6. Applications

These models could be applied to two main potential uses. The first, simulating the response of PST to different conditions in order to develop different strategies of the primary treatment operation. Furthermore, the design of settling tanks also can be improved using the information provided by simulations. Additionally, the response of PST to different conditions can then be used as input data to the development of control systems or operational strategies for the downstream biological processes [27].

Second, these models could be adapted to the online control processes of PST. Thus, these models will be useful operational tools for monitoring the PST operation in real-time. In order to implement an online control process, the input data of the models should be measured online. The inlet flow rate can be measured online and incorporated into the control system. Nevertheless, TSS concentration cannot be measured online, and online COD analyzers are expensive. However, a turbidity analyzer for monitoring the quality of water through continuous measurements of turbidity is available. Due to the strong correlation between sewage composition and turbidity, TSS concentration and COD can be estimated using turbidity data. Hence, turbidity can supply the required information to implement neural network models to online processes control. As well, these estimations using turbidity data could be used to compensate for a high extent of the problems caused by breakdowns of process analyzers. Despite a sudden change of the wastewater composition after the training period could produce significant incorrect predictions, these adaptative models could be re-trained using the current data. This flexibility of

neural models is an important advantage because the models will be readapted to new conditions. Therefore, these models can be used to model sedimentation units in other wastewater treatment plants.

CONCLUSIONS AND RECOMMENDATIONS

- The predictive model based on ANN obtained satisfactory results for predicting COD and TSS concentration in the clarified effluent. The proposed model obtained successful results during the training phase and validation.
- The process analysis determined that inlet flowrate, TSS, COD, and turbidity of the influent wastewater are the most representative input variables that must be considered for the PST modeling. Additionally, useful correlations between COD, TSS, and turbidity were found. They can be applied to obtain a rapid estimation of COD or TSS concentration using turbidity measurements.
- A multi-layer feed-forward architecture with only one hidden layer was used in both ANN models. In the case of the TSS model, a hidden layer of seven nodes provided the best results ($R^2=0.60$). For the COD model, the hidden layer with ten hidden nodes gave the best performance ($R^2=0.60$). However, the COD model using nine nodes provide similar results with a simpler architecture ($R^2=0.58$).
- The ANN models reached better results and more accurate predictions in comparison to empiric models reported in the literature. The ANN models obtained a significant reduction of average error associated with the model predictions. The TSS model obtained an average error of 22.87% relative to the experimental data, while the average error for the empirical TSS model was 42.42%. In the same way, the COD model obtained an average error of 11.12%, whereas the average error of the empirical COD model was 18.24%. Furthermore, ANN models reduce the variability of estimation error compared to empiric models.
- The local SA evidenced a proportional relationship between all input variables and the TSS model response. The behavior of the TSS model under different conditions is consistent with the normal PST operation. On the other hand, the COD model showed different responses under the variation of input variables. Due to the complicated internal relationships that neural networks developed during the training

phase without physical considerations, the COD model response is not coherent with the physical behavior of the PST. However, it provides good predictions of effluent characteristics. Moreover, the SA exposed a strong correlation between the COD model response and influent COD.

- The global SA showed that the predictions of the COD model and TSS model mainly depends on the influent COD. The MSE value of the TSS model is six times greater when the input COD value is not considered. In the case of the COD model, the MSE value increases more than forty times. This indicates a higher sensitivity to input COD data of the COD model in comparison to the TSS model. In contrast, the inlet flowrate has the lowest influence in the response of both the COD model and the TSS model. Therefore, the predictions of ANN models mostly depend on the composition of influent wastewater.
- As a recommendation, it should be considered a stricter and detailed sampling program at different delay times of the PST operation in the Ibarra WWTP. Thus, the ANN models will be trained with a complete description of the PST response at different delay times. Moreover, a more extensive data set should be used in the training phase to improve the predicting ability of the ANN models.
- Finally, this study should be considered to implement a process control system for monitoring the PST operation in the Ibarra WWTP. The proposed model will be retrained using the current operation data of the PST, and an online turbidity analyzer should be considered for the implementation of the continuous monitoring of the PST operation.

REFERENCES

- [1] Metcalf & Eddy Inc, G. Tchobanoglous, F. L. Burton and H. D. Stensel, *Wastewater Engineering: Treatment and Reuse*, 4th ed. New York, USA: McGraw-Hill, 2003, pp. 1-411. ISBN: 0071122508.
- [2] Programa Mundial de Evaluación de los Recursos Hídricos de las Naciones Unidas, *Informe Mundial de las Naciones Unidas sobre el Desarrollo de los Recursos Hídricos 2017. Aguas residuales: El recurso no explotado*. París, Francia: UNESCO, 2017, pp. 39-43. ISBN:9789233000582.
- [3] World Health Organization (WHO) and United Nations Environment Programme (UNEP), *Wastewater management: A UN-Water Analytical Brief*. USA: UN-Water, 2015, sec. 2, pp. 9-10. Available at: <https://www.unenvironment.org/resources/report/wastewater-management-un-water-analytical-brief>.
- [4] R. Riffat, *Fundamentals of Wastewater Treatment and Engineering*. Florida, USA: CRC Press, 2012, pp. 75- 119. ISBN: 978020385717.
- [5] Instituto Nacional de Estadística y Censo (INEC), “*Estadística de Información Ambiental Económica en Gobiernos Autónomos Descentralizados Municipales 2015 (Agua Y Alcantarillado)*”, Dirección de Estadísticas Agropecuarias y Ambientales, Ecuador, Oct 2016, pp. 22-25. Available at: https://www.ecuadorencifras.gob.ec/documentos/web-inec/Encuestas_Ambientales/Municipios_2015/Documento_Tecnico-Gestion_de_Agua_y_Alcantarillado_2015.pdf.
- [6] R. S. Ramalho, *Introduction to Wastewater Treatment Processes*, 2nd ed. New York, USA: Academic Press, 1983, pp. 63-108. ISBN: 0125765606.
- [7] M. Jover-Smet, J. Martín-Pascual and A. Trapote, “Model of suspended solids removal in the primary sedimentation tanks for the treatment of urban wastewater”, *Water*, vol. 9, no. 6, pp. 448, Jun. 2017. doi: 10.3390/w9060448.
- [8] D. G. Christoulas, P. H. Yannakopoulos and A. D. Andreadakis, “An empirical model for primary sedimentation of sewage”, *Environ. Int.*, vol. 24, no. 6, pp. 925-934, Nov. 1998. doi: 10.1016/S0160-4120(98)00076-2.
- [9] Y. Djebbar and P. T. Kadota, “Estimating sanitary flows using neural networks”, *Water Sci. and Technol.*, vol. 38, no. 10, pp. 215-222, Jun. 1998. doi: 10.1016/S0273-1223(98)00752-5.
- [10] G. Martínez-González, H. Loría-Molina, D. Taboada-López, F. Ramírez-Rodríguez, J. L. Navarrete-Bolaños and H. Jiménez-Islas, “Approximate method for designing a primary settling tank for wastewater treatment,” *Ind. & Eng. Chem.*

- Res., vol. 48, no. 16, pp. 7842-7846, Jul. 2009. doi: 10.1021/ie801869b.
- [11] M. Von Sperling and C. A. De Lemos Chernicharo, “Biological Wastewater Treatment in Warm Climate Regions”, in *Biological Wastewater Treatment Series*, London, UK: IWA Publishing, 2005, vol. 1, pp. 30-183. ISBN: 1843390027.
- [12] F. R. Spellman, *Handbook of Water and Wastewater Treatment Plant Operations*. Florida, USA: CRC Press, 2013, pp. 583-701. ISBN: 9781466553385.
- [13] R. H. G. Perry, D. W. Green and J. O. Maloney, *Perry’s Chemical Engineers’ Handbook*, 7th ed. New York, USA: McGraw-Hill, 1997. sec. 25, pp. 62-64, ISBN: 0070498415.
- [14] C. C. Lee and S. D. Lin, *Handbook of Environmental Engineering Calculations*, 2nd ed. New York, USA: McGraw-Hill, 2007, pp. 156-1453. ISBN: 0071475834.
- [15] R. B. Baird, A. D. Eaton and E. W. Rice, *Standard Methods for the Examination of Water and Wastewater*, 23th ed. Washington D.C., USA: American Water Works Association, 2017, sec. 2, pp 56-73. ISBN: 9780875532875.
- [16] Woodard & Curran Inc, *Industrial Waste Treatment Handbook*, 2nd ed. Boston, USA: Butterworth-Heinemann, 2006, pp. 97-109. ISBN: 9780750679633.
- [17] H. Patel and R. T. Vashi, “Use of Naturally Prepared Coagulants for the Treatment of Wastewater from Dyeing Mills,” in *Characterization and Treatment of Textile Wastewater*, London, UK: Elsevier, 2015, ch. 6, pp. 147-158. doi: 10.1016/C2014-0-02395-7.
- [18] N. P. Cheremisnoff, *Handbook of water and wastewater Treatment Technologies*. Boston, USA: Butterworth-Heinemann, 2002, pp. 4-5. ISBN: 0750674989.
- [19] M. von Sperling and C. de Lemos, “Biological Wastewater Treatment in Warm Climate Regions”, in *Biological Wastewater Treatment Series*, London, UK: IWA Publishing, 2005, vol. 1, pp. 30-183. ISBN: 1843390027.
- [20] E. Levlin, “Conductivity Measurements for Controlling Municipal Waste-Water Treatment,” in *Research and application of new technologies in wastewater treatment and municipal solid waste disposal in Ukraine, Sweden and Poland: Proceedings of a Polish-Swedish-Ukrainian seminar*, 2010, pp. 51-62. Available at: <http://www.energiomiljo.org/kth/Polishproject/rep15/ConductV15.pdf>.
- [21] F. P. García, E. B. Esteban, M. Vega and L. Debá, “A rapid estimation of metal contents in wastewater treatment for conductivity measurements”, *J. Chil. Chem. Soc.*, vol. 50, no. 3, pp. 547-551, Sep. 2005. doi: 10.4067/S0717-97072005000300004.
- [22] M. Scholz, *Wetlands for Water Pollution Control*, 2nd ed. Oxford, UK: Elsevier, 2015, pp. 27-35. ISBN: 9780444636072.
- [23] United States Environmental Protection Agency (EPA), “Wastewater Technology Fact Sheet: Screening and Grit removal”, United States Environ. Prot. Agency (EPA), Washington, D.C, USA, Tech. Rep. EPA 832-F-03-011, 2003. Available

at: <https://www.semanticscholar.org/paper/Wastewater-Technology-Fact-Sheet-Screening-and-Grit/9c4b0d67e7330b7d8c248ffe123b98c42f6bf326>.

- [24] R. Droste and R. Gehr, *Theory and Practice of Water and Wastewater Treatment*, 2nd ed. New Jersey, USA: Wiley, 2019, pp. 754-757. ISBN: 0471124443.
- [25] M. von Sperling, *Basic Principles of Wastewater Treatment*, in *Biological Wastewater Treatment Series*, vol. 2, London, UK IWA Publishing, 2007, pp. 125-140. ISBN: 9781843391616.
- [26] M. Patziger and K. Kiss, "Towards a hydrodynamically enhanced design and operation of primary settling tanks - Results of a long term in situ measurement investigation program", *Water Environ. J.*, vol. 29, no. 3, pp. 338-345, May. 2015. doi: 10.1111/wej.12125.
- [27] A. Gamal El-Din and D. W. Smith, "Modeling a full-scale primary sedimentation tank using artificial neural networks," *Environ. Technol.*, vol. 23, no. 5, pp. 479-496, Oct. 2002. doi: 10.1080/09593332308618384.
- [28] Acciona Agua and BTD Proyectos, "Planta de Tratamiento de Aguas Residuales de Ibarra: Manual de Operación y Mantenimiento," Acciona Agua, Madrid, España, Tech. Rep. 144201-000-JRF-0001, Aug. 1, 2017.
- [29] M. Samer, "Biological and Chemical Wastewater Treatment Processes", in *Wastewater Treatment Engineering*, 1st ed. London, UK: IntechOpen, 2015, ch.1. pp. 1-5. doi: 10.5772/61250.
- [30] N. F. Gray, *Water Technology: an Introduction for Environmental Scientists and Engineers*, 2nd ed. Amsterdam, Netherlands: Elsevier, 2005, pp. 400-402. ISBN: 97807506-6633-6.
- [31] Nalco Company, *The Nalco Water Handbook*, 3rd ed. New York, USA: McGraw-Hill, 2009, sec. 2.4, ch. 22, pp. 3-18. ISBN: 9780071548830
- [32] F. Concha and R. Bürger, "A century of research in sedimentation and thickening", *KONA Powder Part. J*, vol. 20, pp. 38-70. 2002. doi: 10.14356/kona.2002009.
- [33] R. Smith, "Preliminary design of wastewater treatment systems," *J. Sanit. Eng. Div.*, vol. 95, no. 1, pp. 117-148, 1969. Available at: <https://cedb.asce.org/CEDBsearch/record.jsp?dockkey=0016256>.
- [34] T. H. Y. Tebbutt and D. G. Christoulas, "Performance relationships for primary sedimentation", *Water Res.*, vol. 9, no. 4, pp. 347-356, 1975. doi: 10.1016/0043-1354(75)90180-3.
- [35] C. C. Aggarwal, *Neural Networks and Deep Learning*. Cham, Switzerland: Springer 2018, pp. 1-42. ISBN: 9783319944623.
- [36] M. Kuhn and K. Johnson, *Applied predictive modeling*. New York, USA: Springer, 2013, pp. 141-151. ISBN: 9781461468486.
- [37] J. Stiles and T. L. Jernigan, "The basics of brain development", *Neuropsychology*

- Review*, vol. 30, pp. 327-348, Nov. 2010. doi: 10.1007/s11065-010-9148-4.
- [38] Y. P. Chen, E. P. Ivanova, F. Wang and P. Carloni, “Bioinformatics”, in *Comprehensive Natural Products II: Chemistry and Biology*, vol. 9, pp. 569-593, 2010, Mar. 2010. ISBN: 9780080453828.
- [39] B. Krose and P. V. Smagt, *An introduction to Neural Networks*, 8th ed. Amsterdam, Netherlands: University of Amsterdam, 1996, pp. 13-19. Available at: <https://www.infor.uva.es/~teodoro/neurointro.pdf>.
- [40] D. Kriesel, *A Brief Introduction to Neural Networks*. Bonn, Germany: University of Bonn, 2005, pp. 33-44. Available at: http://www.dkriesel.com/en/science/neural_networks.
- [41] P. Kim, *MATLAB Deep Learning*. New York, USA: Apress, 2017, pp. 1-50. ISBN: 9781484228456.
- [42] K. L. Du and M. N. S. Swamy, *Neural networks in a Softcomputing Framework*. London, UK: Springer, 2006, pp. 1-40. ISBN: 9781846283024.
- [43] Y. Wang, Y. Li, Y. Song and X. Rong, “The influence of the activation function in a convolution neural network model of facial expression recognition”, *Appl. Sci.*, vol. 10, no. 5, pp. 1897, Mar. 2020. doi: 10.3390/app10051897.
- [44] B. Karlik and A. Olgac, “Performance analysis of various activation functions in generalized MLP architectures of neural networks”, *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111-122, Dec. 2010. Available at: https://www.researchgate.net/publication/228813985_Performance_Analysis_of_Various_Activation_Functions_in_Generalized_MLP_Architectures_of_Neural_Networks.
- [45] K. Gurney, “Neural networks for perceptual processing: From simulation tools to theories,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 362, no. 1479, pp. 339-353, Mar. 2007. doi: 10.1098/rstb.2006.1962.
- [46] J. Larsen, *Introduction to Artificial Neural Networks*, 1st ed. Kongens Lyngby, Denmark: University of Denmark, 1999, pp. 14-15. Available at: https://d1wqtxts1xzle7.cloudfront.net/49157436/imm2443.pdf?1474990280=&response-content-disposition=inline%3B+filename%3DIntroduction_to_artificial_neural_network.pdf&Expires=1598856112&Signature=Cv0VUvBiqevoKyavS0oBAR5BBmirMpTMC1juS-TaLRWwp3QZluEBs~3GZn-c15XVwnICZOk-k4uXkNe6Ch0L5A00IoLB2C4zS8BWH9sGQrvvGU~wLW1n9-Xlis2Nro912n8syKOQwaqx98enjfx8T26~VRSCFnc-1h5uKTh2vdpIbCfi56JrVt6vf-mEiYEQJpHAb~IeFkuatv85Fy-0vnxfwyfmb~piq00c8QL4aw7RqiZPqEUIVyxji5xF55t4fxDoCYb~eJyzeRJnne p2l-BvpMn2UHQHINYQ7DpuXk0GBj51rn~DeFNmtUT2ifZcPzLsZLM1lq8qCo6HsHrA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA.

- [47] N. M. Nawi, A. Khan and M. Z. Rehman, “CSLM: Levenberg Marquardt based backpropagation algorithm optimized with Cuckoo search”, *J. ICT Res. Appl.*, vol. 7, no. 2, pp. 105-119, Nov. 2013. doi: 10.5614/itbj.ict.res.appl.2013.7.2.1.
- [48] Y. C. Du and A. Stephanus, “Levenberg-Marquardt neural network algorithm for degree of arteriovenous fistula stenosis classification using a dual optical photoplethysmography sensor”, *Sensors*, vol. 18, no. 7, pp. 2322, Jul. 2018. doi: 10.3390/s18072322.
- [49] M. I. a Lourakis, “A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar”, in *Foundation of Research and Technology - Hellas*, Greece, Feb. 2005, pp. 1-6. Available at: https://www.researchgate.net/publication/239328019_A_Brief_Description_of_the_Levenberg-Marquardt_Algorithm_Implemented_by_levmar.
- [50] I. Flood and N. Kartam, “Neural networks in civil engineering. I: Principles and understanding”, *J. Comput. Civ. Eng.*, vol. 8, no. 2, pp.131-148, Apr. 1994. doi: 10.1061/(ASCE)0887-3801(1994)8:2(131).
- [51] H. Demuth and M. Beale, *Neural Network Toolbox - For Use with MATLAB*. Massachusetts, USA: MathWorks, 2002, Section 5, pp. 2-57. Available at: http://128.174.199.77/matlab_pdf/nnet.pdf.
- [52] S. Skansi, *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*, 1st ed. Cham, Switzerland: Springer International Publishing AG, 2018, pp. 107-111. [ISBN: 9783319730035].
- [53] A. Michalos, *Encyclopedia of Quality of Life and Well-Being Research*, 1st ed. Heidelberg, Germany: Springer, 2014. ISBN: 9789400707528.
- [54] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel and T. L. Veith, “Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations,” *Trans. ASABE*, vol. 50, no. 3, pp. 885-900, May 2007. doi: 10.13031/2013.23153.
- [55] J. P. Barrett, “The coefficient of determination-some limitations,” *Am. Stat.*, vol. 28, no. 1, pp. 19-20, Mar. 1974. doi: 10.1080/00031305.1974.10479056.
- [56] A. Gonzáles, U. Pérez, X. López, A. Garcia and L. Lopez, “EDAR de Ibarra, Ecuador”, *FuturENViRO*, no. 43, pp. 51-69, Sep. 2017. Available at: http://www.futurenviro.com/pdf/reportajes-especiales/09-2017/EDAR_Ibarra.pdf.
- [57] D. Orhon and E. U. Çokgör, “COD Fractionation in Wastewater Characterization—The State of the Art”, *J. Chem. Technol. Biotechnol.*, vol. 68, no. 3, pp. 283-293, Mar. 1997. doi: 10.1002/(SICI)1097-4660(199703)68:3<283::AID-JCTB633>3.0.CO;2-X.
- [58] S. Prashanth, P. Kumar and I. Mehrotra, “Anaerobic degradability: Effect of particulate COD”, *J. Environ. Eng.*, vol. 132, no. 4, pp. 488-496, Apr. 2006. doi: 10.1061/(ASCE)0733-9372(2006)132:4(488).

- [59] A. B. Salahudeen, T. Ijimdiya, A. Eberemu and K. Osinubi, “Artificial Neural Networks Prediction of Compaction Characteristics of Black Cotton Soil Stabilized with Cement Kiln Dust”, *J. Soft Comput. Civ. Eng.*, vol. 4, no. 3, pp. 50-71, May. 2018. doi: 10.22115/SCCE.2018.128634.1059.
- [60] M. Häck and M. Köhne, “Estimation of wastewater process parameters using neural networks”, *Water Sci. & Technol.*, vol. 33, no. 1, pp. 101-115, 1996. doi: 10.1016/0273-1223(96)00163-1.
- [61] A. Hannouche, G. Chebbo, G. Ruban, B. Tassin, B. J. Lemaire and C. Joannis, “Relationship between turbidity and total suspended solids concentration within a combined sewer system”, *Water Sci. Technol.*, vol. 64, no. 12, pp. 2445–2452, Dec. 2011. doi: 10.2166/wst.2011.779.
- [62] V. Rostampour, A. Modares, M. Hasan, M. Sadeghi, I. Bernousi and T. Ghanbari, “Using Artificial Neural Network (ANN) technique for prediction of apple bruise damage”, *Aust. J. Crop Sci.*, vol. 7, no. 10, pp. 1442-1448, Jan. 2013. Available at: https://www.researchgate.net/publication/286204203_Using_artificial_neural_network_ann_technique_for_prediction_of_apple_bruise_damage.
- [63] M. A. Shahin, H. R. Maier and M. B. Jaksa, “Predicting settlement of shallow foundations using neural networks”, *J. Geotech. Geoenviron. Eng.*, vol. 128, no. 9, pp. 758-793, Sep. 2002. doi: 10.1061/(ASCE)1090-0241(2002)128:9(785).
- [64] A. Palmer Pol, J. J. Montano Moreno and A. Calafat Far, “Predicción del consumo de extasis a partir de redes neuronales artificiales”, *Adicciones*, vol. 12, no. 1, pp. 29-41, Mar. 2000. doi: 10.20882/adicciones.623.
- [65] A. Saltelli, “Sensitivity Analysis for Importance Assessment”, *Risk Anal.*, vol. 22, no. 3, pp. 579-590, Jul. 2002. doi: 10.1111/0272-4332.00040.
- [66] S. Hoops *et al.*, Ordinary Differential Equations (ODEs) Based Modeling, in *Computational Immunology: Models and Tools*, New York, USA: Springer, 2016, ch. 5, pp 63-78. doi: 10.1016/B978-0-12-803697-6.00005-9.
- [67] V. Abedi, R. Hontecillas, A. Carbo, C. Philipson, S. Hoops and J. Bassaganya-Riera, “Multiscale Modeling: Concepts, Technologies, and Use Cases in Immunology”, in *Computational Immunology: Models and Tools*, New York, USA: Academic Press, 2016, ch. 8, pp 145-173. doi: 10.1016/B978-0-12-803697-6.00008-4.
- [68] A. Irigoyen, A. Maggiora and L. Angelocci, “Análisis de Sensibilidad en Redes Neuronales Artificiales Entrenadas para Estimar la Lámina Total de Agua en el Suelo,” presented *XVII Congresso Brasileiro de Meteorologia*, Gramado, Brazil, 2012. Available at: https://www.researchgate.net/publication/313837942_ANALISIS_DE_SENSIBILIDAD_EN_REDES_NEURONALES_ARTIFICIALES_ENTRENADAS_PARA_ESTIMAR_LA_LAMINA_TOTAL_DE_AGUA_EN_EL_SUELO.

- [69] L. A. T. Nguyen, A. J. Ward and D. Lewis, "Utilisation of turbidity as an indicator for biochemical and chemical oxygen demand", *J. Water Process Eng.*, vol. 4, pp.137-142, Dec. 2014. doi: 10.1016/j.jwpe.2014.09.009.
- [70] T. Bersinger, I. Le Hécho, G. Bareille and T. Pigot, "Assessment of erosion and sedimentation dynamic in a combined sewer network using online turbidity monitoring", *Water Sci. Technol.*, vol. 72, no. 8, pp. 1375-1382, Oct. 2015. doi: 10.2166/wst.2015.350.
- [71] S. K. Das and N. Sivakugan, "Discussion of Intelligent computing for modeling axial capacity of pile foundations", *Can. Geotech. J.*, vol. 47, no. 2, pp. 230-243, Aug. 2010. doi: 10.1139/t10-048.
- [72] P. Krause, D. P. Boyle and F. Bäse, "Comparison of different efficiency criteria for hydrological model assessment", *Adv. Geosci.*, vol. 5, pp. 89-97, Dec. 2005. doi: 10.5194/adgeo-5-89-2005.
- [73] M. Shahin, "Artificial Intelligence in Geotechnical Engineering: Applications, Modeling Aspects, and Future Directions", in *Metaheuristics in Water, Geotechnical and Transport Engineering*, London, UK: Elsevier Science, 2012, pp. 169-194. ISBN: 9780123982964.
- [74] D. Rezazadeh Eidgahee, A. Haddad and H. Naderpour, "Evaluation of shear strength parameters of granulated waste rubber using artificial neural networks and group method of data handling", *Sci. Iran.*, vol. 26, no. 6, pp. 3233-3244, Dec. 2019. doi: 10.24200/SCI.2018.5663.1408.
- [75] P. K. Kolay, A. B. Rosmina and N. W. Ling, "Settlement prediction of tropical soft soil by Artificial Neural Network (ANN)", in *12th International Conference on Computer Methods and Advances in Geomechanics (12th IACMAG)*, Goa, India, Oct. 1-6, 2008, pp. 1843-1849. ISBN: 9781622761760.
- [76] H. Naderpour, A. Kheyroddin and G. G. Amiri, "Prediction of FRP-confined compressive strength of concrete using artificial neural networks", *Compos. Struct.*, vol. 92, no. 12, pp. 2817-2829, Nov. 2010. doi: 10.1016/j.compstruct.2010.04.008.
- [77] M. L. Davis, *Water and wastewater engineering: Design Principles and Practice*. Chicago, USA: McGraw-Hill, 2010, Section 10, pp. 19-20. ISBN: 9780071713856.

APPENDIX

APPENDIX A:
Calculations for Solids Determination

The equations used to determine the several fractions of solids are detailed below.

$$TS (mg/L) = \frac{(A - B) \cdot 1000}{\text{sample volume (ml)}}$$

A = Weight of dried residue + crucible, mg

B = Weight of crucible, mg

$$TSS (mg/L) = \frac{(C - D) \cdot 1000}{\text{sample volume (ml)}}$$

C = Weight of dried residue + crucible+ paper filter, mg

D = Weight of crucible + paper filter, mg

$$TDS (mg/L) = \frac{(E - F) \cdot 1000}{\text{sample volume (ml)}}$$

E = Weight of dried residue + crucible, mg

F = Weight of crucible, mg

$$VS (mg/L) = \frac{(G - H) \cdot 1000}{\text{sample volume (ml)}}$$

$$FS (mg/L) = \frac{(H - I) \cdot 1000}{\text{sample volume (ml)}}$$

G = Weight of residue + crucible or filter before ignition, mg

H = Weight of residue + crucible or filter after ignition, mg

I = Weight of crucible or filter, mg

APPENDIX B:
**Weights and Bias of Artificial Neural
Network Models**

The weights values and bias values of selected neural network models are detailed in the following tables below. Table 18 and Table 20 correspond to the weights and bias of the input layer of the TSS model and the input layer of the COD model, respectively. Whereas Table 19 and Table 21 correspond to the weights values and bias values of the output layer of the TSS model and the output layer of the COD model, respectively. The notation used in the tables is described in Figure 29.

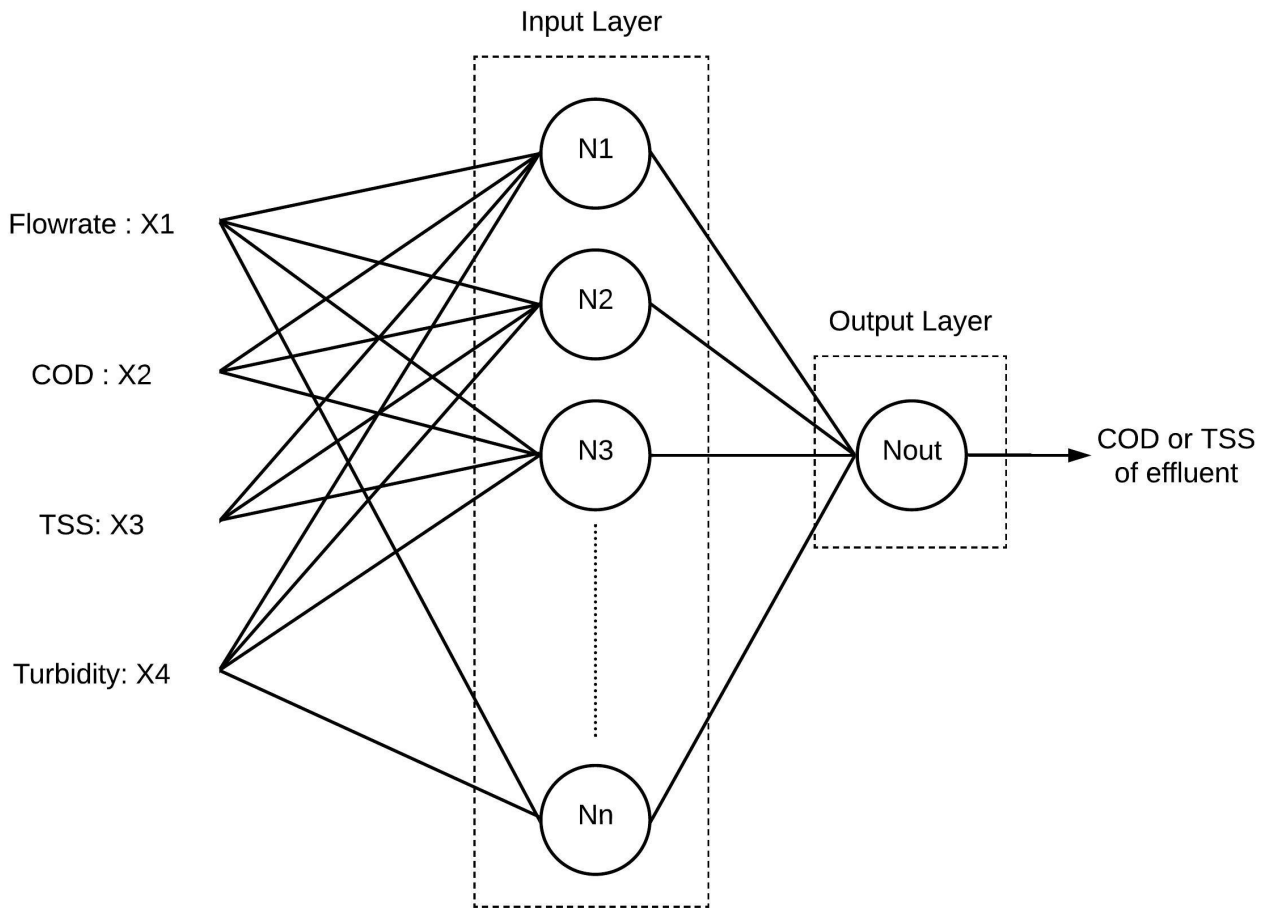


Figure 29. Schematic representation of notation used in ANN models.

Table 18. Weights and bias of the input layer of the TSS model

	N1	N2	N3	N4	N5	N6	N7
X1	-0.5738	-0.7025	-2.0459	-2.6945	-0.9615	-1.1116	-4.0269
X2	1.5529	2.0842	2.9136	0.9065	0.1095	-0.1213	0.8986
X3	-0.7681	0.2069	1.7006	-1.7156	-1.1087	-0.8787	-1.0047
X4	1.6573	2.1172	0.5508	0.6969	-0.1818	-1.1972	0.5602
b	2.6277	2.3316	1.2419	-1.9281	-0.8488	-1.9235	-3.4963

Table 19. Weights and bias of the output layer of the TSS model

	N1	N2	N3	N4	N5	N6	N7	b
Nout	-0.3049	1.673	-0.9114	-1.6247	-0.491	1.4756	1.7368	0.2423

Table 20. *Weights and bias of the input layer of the COD model*

	N1	N2	N3	N4	N5	N6	N7	N8	N9
X1	0.8933	1.1764	0.0456	0.3398	1.2355	-1.0854	-0.9091	-0.8425	-1.9629
X2	-1.3764	2.1995	1.0877	4.5112	-2.4768	-2.6379	-0.6889	0.0839	1.6374
X3	2.0882	2.6052	0.9321	1.2998	1.4563	-0.0667	-0.5774	0.1386	-0.6133
X4	-0.7427	2.8721	-3.7563	-4.215	-2.0396	-2.1571	-1.1806	0.3832	-0.5741
b	-2.0707	-1.4866	-0.8859	1.5561	-0.7479	0.2673	-2.0426	-3.005	-2.7054

Table 21. *Weights and bias of the output layer of the COD model*

	N1	N2	N3	N4	N5	N6	N7	N8	N9	b
Nout	-0.4209	1.576	-2.4578	1.8245	1.1141	2.2126	1.639	0.441	1.0603	-0.3192

APPENDIX C:
Results of Empirical Models

The coefficients of empirical relationships were determined using the Excel tool SOLVER. A non-linear regression method was applied to the calculations. The results of each case are summarized in Table 22. The correlation between experimental data and the empirical model is shown in Figure 30.

The empirical model proposed by Tebbutt and Christoulas ^[34]:

$$E = D \cdot e^{-(F/S+Gq)}$$

Table 22. Coefficients of empirical models.

Parameters	Empirical TSS model	Empirical COD model
D	0.98	1.20
F	86.13	474.75
G	14.81	105.38

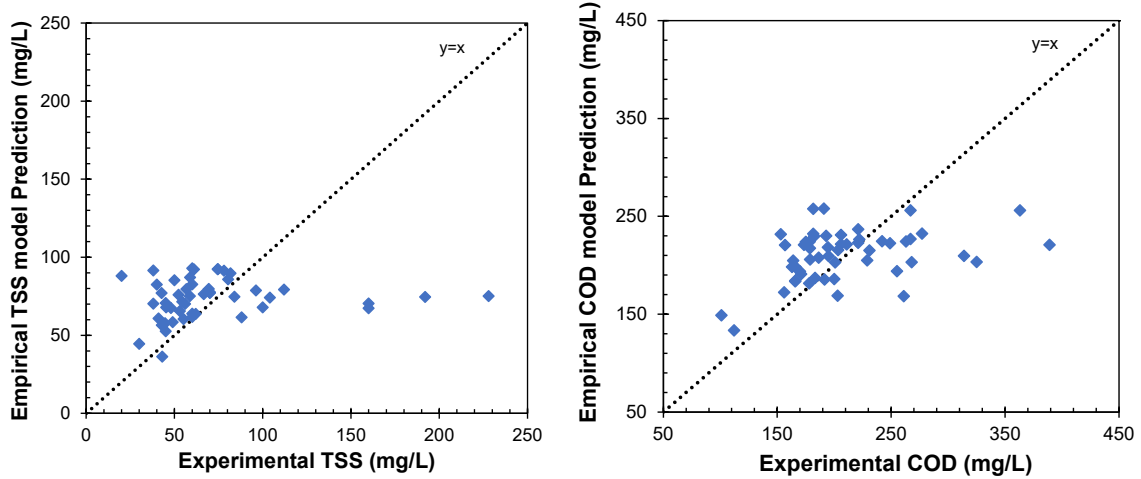


Figure 30. Correlation between experimental data and empirical model predictions