

UCUENCA

Facultad de Ciencias Químicas

Ingeniería Ambiental

Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020

Trabajo de titulación previo a la obtención del título de Ingeniero Ambiental

Autores:

María Paula Picón Mosquera

CI: 0105818868

Correo electrónico: paupiconm1999@hotmail.com

Pedro Andrés Vázquez Flores

CI: 0105519995

Correo electrónico: pedrovf99@outlook.com

Tutor:

Ing. Fausto Fernando García Ávila PhD.

CI: 0301454617

Cuenca, Ecuador

25-julio-2022

RESUMEN

Los eventos naturales de precipitación significativos ocasionan el incremento de la turbidez del agua cruda usada para potabilización; desencadenando problemas, tales como el incremento en el uso de productos químicos y la dificultad en la capacidad del tratamiento en la planta que finalmente afecta la calidad del agua tratada. Esta tesis tuvo como objetivo analizar el impacto de la precipitación pluvial en la turbidez del agua cruda que ingresa a la Planta de Tratamiento de Agua Potable Mahuarca (PTAP-M) de la ciudad de Azogues. El estudio contempló varias etapas; en la primera etapa, se trabajó con la data histórica de precipitación, turbidez y cantidad de coagulante utilizado en el período comprendido entre enero 2014 – mayo 2020, datos proporcionados por la empresa administradora de la planta, EMAPAL-EP; posteriormente se evaluó la representatividad de los datos y su distribución de probabilidad. En la segunda etapa, se realizó un análisis de relación entre la turbidez del agua cruda que ingresa a la PTAP-M y la precipitación pluvial medida directamente de las estaciones meteorológicas próximas a las captaciones del agua cruda. La relación entre estas dos variables se determinó en base a la regresión lineal, redes neuronales artificiales (RNA) y la metodología de superficie de respuesta; para la calibración del modelo se utilizó el 80 % del total de datos disponibles y para la validación el 20 % restante. Los modelos matemáticos obtenidos predicen de manera poco satisfactoria el valor de turbidez del agua cruda conociendo el dato de precipitación. En la tercera y última etapa, se evaluó el esfuerzo del tratamiento químico en la PTAP-M, para lo cual, se desarrolló un modelo matemático que describe la relación entre el esfuerzo de tratamiento químico (cantidad de coagulante), calidad del agua cruda (nivel de turbidez) y precipitación pluvial; la modelación se la realizó mediante las metodologías de análisis de regresión múltiple, redes neuronales artificiales (RNA) y metodología de superficie de respuesta. Finalmente, los resultados de este trabajo demostraron que no fue posible la obtención de modelos que sean capaces de representar de manera óptima el comportamiento real, es decir, se evidenció que los modelos calculados no presentan los mejores resultados. Por lo tanto, estos no permiten predecir de manera efectiva el valor de turbidez del agua cruda que ingresa a la PTAP-M tras un evento de precipitación; de igual forma, no es posible estimar con mayor certeza la cantidad de coagulante a utilizar en el proceso.

Palabras claves: Agua cruda. Precipitación. Turbidez. Esfuerzo de tratamiento químico. Regresión lineal. Regresión múltiple. Redes neuronales artificiales (RNA). Superficie de respuesta. Elasticidad.

ABSTRACT

Significant natural precipitation events cause an increase in the turbidity of raw water used for drinking water treatment, triggering problems such as increased use of chemicals and difficulty in plant treatment capacity that ultimately affect the quality of treated water. The objective of this thesis was to analyze the impact of rainfall on the turbidity of the raw water entering the Mahuarcay Drinking Water Treatment Plant (PTAP-M) in the city of Azogues. The study included several stages; in the first stage, we worked with the historical data of precipitation, turbidity and amount of coagulant used for the period January 2014 - May 2020, data provided by the plant's management company, EMAPAL-EP; subsequently, we evaluated the representativeness of the data and its probability distribution. In the second stage, an analysis was made of the relationship between the turbidity of the raw water entering the PTAP-M and the rainfall measured directly from the meteorological stations near the raw water catchments. The relationship between these two variables was determined based on linear regression, artificial neural networks (ANN) and response surface methodology; 80% of the total available data was used for model calibration and the remaining 20% for validation. The mathematical models obtained predict unsatisfactorily the turbidity value of the raw water knowing the precipitation data. In the third and final stage, the chemical treatment effort in the PTAP-M was evaluated, for which a mathematical model was developed to describe the relationship between the chemical treatment effort (amount of coagulant), raw water quality (turbidity level) and rainfall; the modeling was carried out using multiple regression analysis methodologies, artificial neural networks (ANN) and response surface methodology. Finally, the results of this work showed that it was not possible to obtain models that are capable of optimally representing the real behavior, i.e., it was evident that the calculated models do not present the best results. Therefore, they do not allow to effectively predicting the turbidity value of the raw water entering the PTAP-M after a precipitation event; likewise, it is not possible to estimate with greater certainty the amount of coagulant to be used in the process.

Keywords: Raw water. Raw water quality. Precipitation. Turbidity. Chemical treatment effort. Linear regression. Multiple regression. Artificial neural networks (ANN). Response surface. Elasticity.

ÍNDICE

1. INTRODUCCIÓN Y OBJETIVOS	17
1.1. INTRODUCCIÓN	17
1.2. OBJETIVOS	20
1.2.1. <i>Objetivo general</i>	20
1.2.2. <i>Objetivos específicos</i>	20
2. MARCO TEÓRICO	21
2.1. GENERALIDADES DEL AGUA.....	21
2.1.1. <i>Importancia del agua</i>	21
2.1.2. <i>Tipos de agua</i>	22
2.1.2.1. Agua cruda.....	22
2.1.2.2. Agua potable	22
2.2. ABASTECIMIENTO DE AGUA PARA POTABILIZACIÓN.....	22
2.2.1. <i>Captación de agua cruda</i>	22
2.2.2. <i>Línea de conducción</i>	22
2.2.3. <i>Planta de potabilización</i>	23
2.3. CARACTERÍSTICAS QUE DETERMINAN LA CALIDAD DEL AGUA POTABLE PARA CONSUMO HUMANO.....	23
2.3.1. <i>Características fisicoquímicas del agua potable</i>	23
2.3.1.1. Turbiedad	23
2.3.1.2. Color	23
2.3.1.3. pH.....	23
2.3.1.4. Alcalinidad.....	24
2.3.1.5. Dureza.....	24
2.3.1.6. Nitritos y Nitratos.....	24
2.3.1.7. Sulfatos	25
2.3.1.8. Cloruros	25
2.3.2. <i>Características microbiológicas del agua</i>	25
2.3.2.1. Coliformes: <i>Escherichia coli</i>	25
2.3.3. <i>Turbidez como parámetro de control de la calidad de agua</i>	25
2.4. PROCESOS DE POTABILIZACIÓN DEL AGUA.....	27
2.4.1. <i>Coagulación o Mezcla Rápida</i>	27
2.4.2. <i>Floculación</i>	27
2.4.3. <i>Sedimentación</i>	28
2.4.4. <i>Filtración rápida</i>	28
2.4.5. <i>Desinfección</i>	28
2.5. PRECIPITACIONES.....	29
2.5.1. <i>Tipos de precipitación</i>	29
2.5.2. <i>Pluviometría</i>	30
2.5.3. <i>Estación Meteorológica</i>	30
2.5.4. <i>Influencia de las precipitaciones en la calidad del agua</i>	30
2.6. ESFUERZO DE TRATAMIENTO EN LA PTAP.....	31
2.6.1. <i>Impacto de la calidad del agua cruda en el esfuerzo químico del tratamiento del agua</i>	31
2.6.2. <i>Esfuerzo del tratamiento químico</i>	32
2.6.2.1. Uso de Coagulantes.....	32
2.6.2.1.1. Sulfato de aluminio, alumbre	33
2.6.2.1.2. Policloruro de aluminio, PAC	33
2.6.2.2. Parámetros adicionales.....	34
2.6.3. <i>Elasticidad del agua cruda</i>	35
2.7. ANÁLISIS ESTADÍSTICO	36
2.7.1. <i>Pruebas paramétricos y no paramétricos</i>	36
2.7.1.1. Pruebas paramétricas.....	36
2.7.1.2. Pruebas no paramétricas.....	37
2.7.2. <i>Análisis de regresión lineal simple y múltiple</i>	39

2.7.2.1.	Regresión lineal simple.....	39
2.7.2.2.	Regresión múltiple.....	40
2.7.3.	<i>Redes Neuronales Artificiales, ANN</i>	43
2.7.3.1.	Estructura y elementos básicos que componen la Red Neuronal Artificial.....	44
2.7.3.2.	Aprendizaje y Validación.....	46
2.7.3.2.1.	Niveles o capas de una Red Neuronal Artificial.....	46
2.7.3.2.2.	Tipos de Redes Neuronales Artificiales.....	46
2.7.3.2.3.	Mecanismos de Aprendizaje.....	46
2.7.3.2.4.	Validación de la Red Neuronal Artificial.....	47
2.7.4.	<i>Metodología de superficie de respuesta</i>	47
2.7.5.	<i>Calibración y validación del modelo</i>	48
2.7.5.1.	Error cuadrático medio, RMSE.....	48
2.7.5.2.	Índice de Nash-Sutcliffe, EF.....	49
2.7.6.	<i>Análisis de sensibilidad del modelo</i>	49
3.	MATERIALES Y MÉTODOS	51
3.1.	DESCRIPCIÓN DEL ÁREA DE ESTUDIO Y LA PTAP-M.....	51
3.2.	RECOPIACIÓN DE DATOS.....	55
3.3.	EVALUACIÓN DE LOS DATOS DE TURBIDEZ DEL AGUA CRUDA, PRECIPITACIÓN PLUVIAL Y ESFUERZO DE TRATAMIENTO QUÍMICO DE LA PTAP-M.....	56
3.3.1.	<i>Verificación de la representatividad de los datos</i>	56
3.3.2.	<i>Representatividad de los datos</i>	56
3.3.3.	<i>Comprobación de la normalidad</i>	57
3.4.	ANÁLISIS DE CORRELACIÓN.....	57
3.5.	MODELACIÓN POR REGRESIÓN.....	57
3.5.1.	<i>Modelo de Regresión simple y múltiple</i>	57
3.5.2.	<i>Naturaleza de modelos</i>	59
3.6.	REDES NEURONALES ARTIFICIALES.....	60
3.6.1.	<i>Análisis de Sensibilidad</i>	61
3.7.	METODOLOGÍA DE SUPERFICIE DE RESPUESTA.....	62
3.8.	CALIBRACIÓN Y VALIDACIÓN.....	62
3.9.	ELASTICIDADES DEL AGUA CRUDA DE LA PTAP-M.....	63
3.10.	INCERTIDUMBRES DEL CASO DE ESTUDIO.....	65
4.	RESULTADOS	67
4.1.	EVALUACIÓN DE LOS PARÁMETROS DE LA PTAP-M.....	67
4.1.1.	<i>Verificación de la representatividad de los datos</i>	67
4.1.2.	<i>Representación gráfica de los datos</i>	70
4.1.3.	<i>Comprobación de la normalidad de los datos</i>	72
4.2.	RELACIÓN MATEMÁTICA DE LA PRECIPITACIÓN PLUVIAL Y LA TURBIDEZ DEL AGUA CRUDA DE LA PTAP-M.....	76
4.2.1.	<i>Análisis de correlación y regresión entre turbidez y precipitación</i>	76
4.2.2.	<i>Modelo entre turbidez y precipitación con Redes Neuronales Artificiales</i>	80
4.2.2.1.	Análisis de sensibilidad de parámetros para el modelo entre turbidez y precipitación.....	80
4.2.2.2.	Ensayo de Redes Neuronales Artificiales para el modelo entre turbidez y precipitación.....	82
4.2.2.2.1.	Índice de sensibilidad del número de redes neuronales para la relación turbidez y precipitación.....	88
4.2.3.	<i>Modelo entre turbidez y precipitación con la metodología de superficie de respuesta</i>	89
4.3.	RESPUESTA DEL ESFUERZO DEL TRATAMIENTO QUÍMICO SEGÚN TURBIDEZ DEL AGUA CRUDA Y PRECIPITACIÓN PLUVIAL EN LA PTAP-M.....	90
4.3.1.	<i>Análisis del esfuerzo del tratamiento químico en la PTAP-M</i>	90
4.3.2.	<i>Enfoque de modelización de regresión simple y múltiple entre el esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial</i>	92
4.3.3.	<i>Modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial con Redes Neuronales Artificiales</i>	97

4.3.3.1. Análisis de sensibilidad de parámetros para el modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial	97
4.3.3.2. Ensayo de Redes Neuronales Artificiales para el modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial	99
4.3.3.2.1. Índice de sensibilidad del número de redes neuronales para la relación esfuerzo de tratamiento químico, turbidez y precipitación	105
4.3.4. <i>Modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial con la metodología de superficie de respuesta</i>	<i>106</i>
4.3.5. <i>Elasticidades de la calidad del agua cruda en la PTAP-M.....</i>	<i>107</i>
4.4. COMPARACIÓN DE LOS RESULTADOS CON LA LITERATURA	107
4.4.1. <i>Comparación del modelo entre la relación turbidez y precipitación</i>	<i>107</i>
4.4.2. <i>Comparación del modelo entre la relación de esfuerzo de tratamiento químico, turbidez y precipitación</i>	<i>109</i>
5. CONCLUSIONES	113
6. RECOMENDACIONES.....	117
7. BIBLIOGRAFÍA	118
ANEXOS	124
ANEXO A: RECORRIDO IN SITU DE LOS PUNTOS DE CAPTACIÓN, LÍNEAS DE CONDUCCIÓN Y RESERVORIO DENTRO DEL ÁREA DE ESTUDIO DE LA MICROCUENCA DEL RÍO TABACAY	125
ANEXO B: CURVA DE DOSIFICACIÓN DEL SULFATO DE ALUMINIO Y POLICLORURO DE ALUMINIO DE LA PTAP-M	131

LISTA DE ILUSTRACIONES

ILUSTRACIÓN 1 SISTEMA GLOBAL DE PROCESO DE UNA RED NEURONAL. FUENTE: LARRAÑAGA ET AL., 2012.....	44
ILUSTRACIÓN 2 RED NEURONAL ARTIFICIAL TOTALMENTE CONECTADA. FUENTE: MATICH, 2001.....	45
ILUSTRACIÓN 3 UBICACIÓN DE LA ZONA DE ESTUDIO. A) UBICACIÓN DE LA PROVINCIA DE CAÑAR, REPÚBLICA DEL ECUADOR, B) MICROCUENCA DEL RÍO TABACAY EN LA PROVINCIA DE CAÑAR, C) MICROCUENCA DEL RÍO TABACAY, RÍO TABACAY Y RED HIDROGRÁFICA DE QUEBRADAS	51
ILUSTRACIÓN 4 QUEBRADAS PRINCIPALES DE LA MICROCUENCA DEL RÍO TABACAY	52
ILUSTRACIÓN 5 MAPA ÁREA DE ESTUDIO DE LA MICROCUENCA DEL RÍO TABACAY	52
ILUSTRACIÓN 6 MAPA DE PENDIENTES DEL ÁREA DE ESTUDIO DE LA MICROCUENCA DEL RÍO TABACAY	53
ILUSTRACIÓN 7 REPRESENTATIVIDAD DE LOS DATOS	68
ILUSTRACIÓN 8 VARIACIÓN DIARIA DE LA TURBIEDAD	68
ILUSTRACIÓN 9 VARIACIÓN DIARIA DE LA PRECIPITACIÓN EN LLAUCAY.....	69
ILUSTRACIÓN 10 VARIACIÓN DIARIA DE LA PRECIPITACIÓN EN NUDPUD	69
ILUSTRACIÓN 11 VARIACIÓN DIARIA DE SULFATO DE ALUMINIO.....	70
ILUSTRACIÓN 12 DENSIDAD DE LA TURBIEDAD	71
<i>ILUSTRACIÓN 13 DENSIDAD DE LA PRECIPITACIÓN DE LA ESTACIÓN DE LLAUCAY.....</i>	<i>71</i>
ILUSTRACIÓN 14 DENSIDAD DE LA PRECIPITACIÓN DE LA ESTACIÓN DE NUDPUD	72
ILUSTRACIÓN 15 DENSIDAD DEL ESFUERZO DEL TRATAMIENTO QUÍMICO	72
ILUSTRACIÓN 16 BOXPLOT DATA TURBIEDAD VS BOXPLOT NORMAL	74
ILUSTRACIÓN 17 BOXPLOT DATOS DE TURBIEDAD	74
ILUSTRACIÓN 18 Q-QPLOT TURBIEDAD.....	74
ILUSTRACIÓN 19 BOXPLOT DATA PRECIPITACIÓN LLAUCAY VS BOXPLOT NORMAL	74
ILUSTRACIÓN 20 BOXPLOT DATOS DE PRECIPITACIÓN ESTACIÓN LLAUCAY	74
ILUSTRACIÓN 21 Q-QPLOT PRECIPITACIÓN ESTACIÓN LLAUCAY	75
ILUSTRACIÓN 22 BOXPLOT DATA PRECIPITACIÓN NUDPUD VS BOXPLOT NORMAL	75
ILUSTRACIÓN 23 BOXPLOT DATOS DE PRECIPITACIÓN ESTACIÓN NUDPUD.....	75
<i>ILUSTRACIÓN 24 Q-QPLOT PRECIPITACIÓN ESTACIÓN NUDPUD</i>	<i>76</i>
ILUSTRACIÓN 25 BOXPLOT DATA SULFATO DE ALUMINIO VS BOXPLOT NORMAL	76
ILUSTRACIÓN 26 BOXPLOT DATOS SULFATO DE ALUMINIO	76
ILUSTRACIÓN 27 Q-QPLOT SULFATO DE ALUMINIO	76
ILUSTRACIÓN 28 MODELO LINEAL SIN INTERCEPTO TURBIEDAD – LLAUCAY	77
ILUSTRACIÓN 29 MODELO LINEAL TURBIEDAD – NUDPUD	77
ILUSTRACIÓN 30 MODELO CUADRÁTICO TURBIEDAD - LLAUCAY – NUDPUD	77

LISTA DE TABLAS

TABLA 1 RECLASIFICACIÓN DE LA PENDIENTE EN PORCENTAJE PARA LA ZONA DE ESTUDIO	53
TABLA 2 ESCENARIOS CONSIDERADOS PARA EL ANÁLISIS DE SENSIBILIDAD PARA LAS RNA	61
TABLA 3 REPRESENTATIVIDAD DE LOS DATOS	67
TABLA 4 RESUMEN DE LA EVALUACIÓN DESCRIPTIVA DE LOS DATOS ANALIZADOS EN LA PTAP-M.....	68
TABLA 5 RESUMEN DE NORMALIDAD DE LAS VARIABLES.....	73
TABLA 6 COEFICIENTES DE CORRELACIÓN DE LOS PARÁMETROS ANALIZADOS	76
TABLA 7 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE CALIBRACIÓN DE LOS MODELOS DE REGRESIÓN	78
TABLA 8 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE VALIDACIÓN DE LOS MODELOS DE REGRESIÓN	79
TABLA 9 CAMBIOS EN LA PRECISIÓN DE LA PREDICCIÓN BAJO DIFERENTES PARÁMETROS RNA PARA LA RELACIÓN DE TURBIDEZ Y PRECIPITACIÓN	81
TABLA 10 ÍNDICE DE SENSIBILIDAD (TIEMPO DE ENTRENAMIENTO Y VELOCIDAD DE APRENDIZAJE) PARA LA RELACIÓN ENTRE TURBIDEZ Y PRECIPITACIÓN	82
TABLA 11 RESUMEN DE LOS MODELOS DE RNA ENTRENADOS PARA LA RELACIÓN TURBIDEZ Y PRECIPITACIÓN ..	84
TABLA 12 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE CALIBRACIÓN Y VALIDACIÓN DE LOS MODELOS DE RNA PARA LA RELACIÓN TURBIDEZ Y PRECIPITACIÓN.....	85
TABLA 13 ÍNDICE DE SENSIBILIDAD (NÚMERO DE NEURONAS) PARA LA RELACIÓN ENTRE TURBIDEZ Y PRECIPITACIÓN	89
TABLA 14 ERRORES DEL MODELO TURBIDEZ Y PRECIPITACIÓN CON LA METODOLOGÍA DE SUPERFICIE DE RESPUESTA	89
TABLA 15 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE CALIBRACIÓN DE LOS MODELOS DE REGRESIÓN PARA EL ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	93
TABLA 16 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE VALIDACIÓN DE LOS MODELOS DE REGRESIÓN PARA EL ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	96
TABLA 17 CAMBIOS EN LA PRECISIÓN DE LA PREDICCIÓN BAJO DIFERENTES PARÁMETROS RNA PARA LA RELACIÓN ESFUERZO TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	98
TABLA 18 ÍNDICE DE SENSIBILIDAD (TIEMPO DE ENTRENAMIENTO Y VELOCIDAD DE APRENDIZAJE) PARA LA RELACIÓN ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	99
TABLA 19 RESUMEN DE LOS MODELOS DE RNA ENTRENADOS PARA LA RELACIÓN ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	101
TABLA 20 RESULTADOS DE LOS ERRORES DE LAS ETAPAS DE CALIBRACIÓN Y VALIDACIÓN DE LOS MODELOS DE RNA PARA LA RELACIÓN ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	103
TABLA 21 ÍNDICE DE SENSIBILIDAD (NÚMERO DE NEURONAS) PARA LA RELACIÓN ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN	106
TABLA 22 ERRORES DEL MODELO ESFUERZO DE TRATAMIENTO QUÍMICO, TURBIDEZ Y PRECIPITACIÓN CON LA METODOLOGÍA DE SUPERFICIE DE RESPUESTA	107

LISTA DE ECUACIONES

ECUACIÓN 1 TASA DE DOSIS DE COAGULANTE	34
ECUACIÓN 2 RELACIÓN ENTRE DOSIS DE COAGULANTE Y CALIDAD DEL AGUA	34
ECUACIÓN 3 ELASTICIDAD DE LA TURBIDEZ	36
ECUACIÓN 4 ESTRUCTURA BASE DE LA ECUACIÓN DE REGRESIÓN MÚLTIPLE	40
ECUACIÓN 5 FORMA DE LA SUPERFICIE DE RESPUESTA	47
ECUACIÓN 6 ERROR CUADRÁTICO MEDIO, RMSE.....	49
ECUACIÓN 7 ÍNDICE DE NASH-SUTCLIFFE, EF.....	49
ECUACIÓN 8 ÍNDICE DE SENSIBILIDAD "S"	50
ECUACIÓN 9 CANTIDAD DE COAGULANTE TOTAL.....	64
ECUACIÓN 10 RNA MODELO 2	86
ECUACIÓN 11 RNA MODELO 10-15-5.....	88
ECUACIÓN 12 CÁLCULO DE LA TURBIDEZ CON LA METODOLOGÍA DE SUPERFICIE DE RESPUESTA DEL PROCESO ...	89
ECUACIÓN 13 RNA MODELO 3	103
ECUACIÓN 14 RNA MODELO 8-6-4-2	105
ECUACIÓN 15 CANTIDAD DE COAGULANTE MODELADA CON LA METODOLOGÍA DE SUPERFICIE DE RESPUESTA ..	106

LISTA DE ANEXOS

ANEXO 1 QUEBRADA NUDPUD, MICROCUENCA DEL RÍO TABACAY	125
ANEXO 2 SISTEMA DE CAPTACIÓN DE AGUA CRUDA DE LA QUEBRADA NUDPUD, MICROCUENCA DEL RÍO TABACAY	125
ANEXO 3 ZONA ALEDAÑA AL SISTEMA DE CAPTACIÓN DE AGUA CRUDA DE LA QUEBRADA NUDPUD, MICROCUENCA DEL RÍO TABACAY	126
ANEXO 4 QUEBRADA LLAUCAY, MICROCUENCA DEL RÍO TABACAY	126
ANEXO 5 SISTEMA DE CAPTACIÓN DE AGUA CRUDA DE LA QUEBRADA LLAUCAY, MICROCUENCA DEL RÍO TABACAY	126
ANEXO 6 ZONA ALEDAÑA AL SISTEMA DE CAPTACIÓN DE AGUA CRUDA DE LA QUEBRADA LLAUCAY, MICROCUENCA DEL RÍO TABACAY	127
ANEXO 7 TRAMOS DE LA LÍNEA DE CONDUCCIÓN A CIELO ABIERTO DEL AGUA CRUDA PROVENIENTE DE LA QUEBRADA NUDPUD, MICROCUENCA DEL RÍO TABACAY	127
ANEXO 8 TRAMO DE LA LÍNEA DE CONDUCCIÓN CUBIERTA DEL AGUA CRUDA PROVENIENTE DE LAS QUEBRADAS NUDPUD Y LLAUCAY, MICROCUENCA DEL RÍO TABACAY	128
ANEXO 9 TRAMO DE LA LÍNEA DE CONDUCCIÓN CUBIERTA CON MALEZA, HOJARASCA Y RESTOS DE FLORA DEFORESTADA, MICROCUENCA DEL RÍO TABACAY	128
ANEXO 10 RESERVORIO DE AGUA CRUDA EN LA PTAP-M (DESEMBOCADURA DE LA LÍNEA DE CONDUCCIÓN DE AGUA CRUDA DE NUDPUD Y LLAUCAY). INGRESO DEL AGUA CRUDA DE LA LÍNEA DE CONDUCCIÓN AL RESERVORIO (CÍRCULO AMARILLO)	129
ANEXO 11 CONTINUACIÓN DEL RESERVORIO DE AGUA CRUDA EN LA PTAP-M	129
ANEXO 12 SECCIÓN FINAL DEL RESERVORIO DE AGUA CRUDA EN LA PTAP-M. INGRESO DEL AGUA CRUDA DEL RESERVORIO A LA TUBERÍA (1 KM) PARA INGRESAR A LA PTAP-M	129
ANEXO 13 INGRESO DEL AGUA CRUDA DE LA TUBERÍA (1 KM) A LA PTAP-M. PUNTO DE MEDICIÓN DEL NIVEL DE TURBIDEZ DEL AGUA CRUDA	130
ANEXO 14 INICIO DEL PROCESO DE POTABILIZACIÓN DEL AGUA CRUDA EN LA PTAP-M	130
ANEXO 15 GRÁFICA DE SULFATO DE ALUMINIO Y PAC, CURVA DE DOSIFICACIÓN DE LA PTAP-M	131

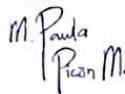
Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Cláusula de licencia y autorización para publicación en el Repositorio Institucional

María Paula Picón Mosquera en calidad de autora y titular de los derechos morales y patrimoniales del trabajo de titulación "Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 25 de julio de 2022



María Paula Picón Mosquera

C.I: 0105818868

Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Cláusula de licencia y autorización para publicación en el Repositorio Institucional

Pedro Andrés Vázquez Flores en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación "Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020", de conformidad con el Art. 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN reconozco a favor de la Universidad de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Asimismo, autorizo a la Universidad de Cuenca para que realice la publicación de este trabajo de titulación en el repositorio institucional, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Cuenca, 25 de julio de 2022



Pedro Andrés Vázquez Flores

C.I: 0105519995

Cláusula de Propiedad Intelectual

Cláusula de Propiedad Intelectual

María Paula Picón Mosquera, autora del trabajo de titulación "Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autora.

Cuenca, 25 de julio de 2022



María Paula Picón Mosquera

C.I: 0105818868

Cláusula de Propiedad Intelectual

Cláusula de Propiedad Intelectual

Pedro Andrés Vázquez Flores, autor del trabajo de titulación "Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 25 de julio de 2022



Pedro Andrés Vázquez Flores

C.I: 0105519995

DEDICATORIA

A mis apreciados padres, Patricio y Raquel, por su amor y apoyo incondicional quienes se merecen esta investigación, que al finalizarla me genera mucha felicidad y gratitud hacia ellos. A mis hermanos, Christian Patricio y Juan José, que son mi ejemplo a seguir, y, de forma general a todo mi círculo familiar que son uno de mis pilares fundamentales y mi fuerza.

A cada uno de mis Favos que son mis mejores amigos, la familia que elegí, en especial énfasis para aquellos que siempre se mantuvieron a mi lado, brindándome su amistad, tiempo, lealtad y apoyo constante.

A todas estas personas, que son relevantes en mi vida y han aportado para forjar mi esencia y calidad como persona y ahora como profesional.

Muchas Gracias.

María Paula Picón Mosquera

DEDICATORIA

Lleno de orgullo y sentimiento dedico mi tesis a mis seres queridos, con especial énfasis a papi Rober, quien desde pequeño supo guiarme y aconsejarme de la mejor manera, aunque hoy no este conmigo fue la persona que más espero este día.

A mis amados padres, a mi mami que ha sabido formarme con buenos sentimientos, hábitos y valores lo cual me han ayudado a seguir en los momentos difíciles, por todo su amor y consejos a lo largo de la carrera; a mi papi quién ha sido un ejemplo de dedicación y superación, por todo su apoyo y cariño.

A mis hermanos, Sol y Sebas, siempre serán mi mayor orgullo. A mi tía Marcela y mi tío Bayron, a mis primos que son como mis otros hermanos: Ari, David y Xavi, que con sus palabras de aliento y consejos no me dejaron decaer para seguir siempre adelante y así cumplir con todos mis ideales y propósitos.

A mi amiga de toda la vida Micaela, y a mis amigas de toda la carrera Marilyn y Dome.

Gracias a todos.

Pedro Andrés Vázquez Flores

AGRACEDIMIENTOS

En conjunto, agradecemos a nuestros familiares y amigos, por su comprensión, estímulo constante y el apoyo incondicional a lo largo de la carrera.

A todos nuestros docentes que nos acompañaron en este recorrido, que han compartido el conociendo y experiencia de la vida con nosotros; forjándonos como personas. A nuestro tutor de tesis Ingeniero Fernando García por habernos brindado la oportunidad de recurrir a su capacidad y conocimiento científico, así como también su paciencia y tiempo durante el desarrollo de esta tesis.

A la Planta de Tratamiento de Agua Potable Mahuarcay, por brindarnos toda la información necesaria para la realización de la tesis.

CAPÍTULO I

1. INTRODUCCIÓN Y OBJETIVOS

1.1. INTRODUCCIÓN

El acceso al agua potable a nivel mundial es un eje primordial para la salud pública y el desarrollo social, siendo el líquido vital para la humanidad, además, de ser uno de los derechos humanos básicos más trascendentales para las sociedades. Es por ello, la importancia de garantizar la calidad del agua potable y su abastecimiento satisfactorio; teniendo presente, que este criterio satisfactorio involucra suficiencia, seguridad y accesibilidad. La calidad del agua cruda se ha visto afectada por el cambio climático, el crecimiento de la población y el aumento de las actividades antropogénicas, que genera grandes desafíos a las instituciones encargadas del suministro del agua potable alrededor del mundo, con especial énfasis en los países en desarrollo.

Según el Informe Mundial de las Naciones Unidas sobre el Desarrollo de los Recursos Hídricos 2020, explica que las masas de agua superficial del mundo están experimentando un deterioro de la calidad del agua por agentes externos como son: los sedimentos, patógenos, sales, nutrientes, fertilizantes, pesticidas, metales pesados, residuos orgánicos o contaminantes emergentes (UNESCO & ONU-Agua, 2020). Conjuntamente, el Informe sobre el Desarrollo de los Recursos Hídricos en el Mundo 2021, informa que la calidad del agua en casi todos los ríos principales de África, América Latina y Asia se han deteriorado notablemente (Naciones Unidas, 2021).

Las fuentes del deterioro de la calidad del agua cruda son variadas; fuentes puntuales, no puntuales o de origen natural como son las precipitaciones. Es así como los episodios de lluvias abundantes aumentan notablemente la turbidez del agua cruda, pudiendo ocasionar problemas en los diferentes procesos de potabilización si el operador no toma los correctivos necesarios a su debido tiempo, ocasionando cambios evidentes en la calidad del agua tratada y repercutiendo directamente en el funcionamiento de cualquier planta de tratamiento de agua potable (PTAP). Los cambios de la turbidez del agua cruda originada por fuertes precipitaciones pluviales suelen caracterizarse por ser rápidos y esporádicos, ocasionando un potencial problema para los diferentes procesos en la planta de tratamiento, llegando a saturar los filtros o sobrecargarse los sedimentadores; desencadenando una contaminación del agua tratada y del sistema de distribución (Gorchev & Ozolins, 1984). Del mismo modo, la presencia de sedimentos y partículas asociadas a la turbidez aumentan los esfuerzos de tratamiento químico en los

procesos; siendo la cantidad de productos químicos uno de los factores que más varía en el proceso según la calidad del agua cruda, incrementando consecuentemente el costo del tratamiento.

Con lo mencionado, a nivel integral, se concibe el objetivo de trabajo de toda planta de tratamiento de agua potable; que es mantener y brindar agua potable de calidad constante. Este último argumento, aporta y suma de forma general para cumplir con los Objetivos de Desarrollo Sostenible, los globalmente conocidos ODS, enfocándose de forma directa al Objetivo 6 *Agua limpia y Saneamiento*. El objetivo 6 busca garantizar la disponibilidad de agua y saneamiento para todos; por lo que abarca la higiene, la gestión de cuencas fluviales y la gestión integrada de sus recursos hídricos y las preocupaciones ambientales.

A nivel nacional, no se han realizado investigaciones profundas sobre la influencia de la precipitación en la turbidez del agua cruda que es usada para la potabilización. Por ello, la respectiva investigación resulta de importancia para poder controlar y tomar las medidas necesarias en el proceso de potabilización ante un evento de un incremento súbito de turbidez del agua cruda. El conocer con anticipación la turbidez del agua cruda que ingresa a una PTAP, evitará desperdicios de productos químicos, una deficiente eliminación de patógenos del agua, formación de compuestos tóxicos, cortes del servicio de abastecimiento, pérdidas económicas, etc.; todo esto aporta para la toma de decisiones institucionales, mejorando el sistema y la gestión de las PTAPs, e, inmediatamente, mejorar el servicio hacia la comunidad.

El presente estudio se llevó a cabo en la Planta de Tratamiento de Agua Potable Mahuarcay (PTAP-M), implementada en el cantón Azogues, la cual inició sus actividades de operación el 14 de febrero de 2005. La PTAP-M, es administrada y operada por la empresa pública EMAPAL EP. Trata las aguas provenientes de las quebradas Nudpud y Llaucay, en promedio se potabiliza alrededor de 100 l/s de agua superficial, es una planta convencional con operación a gravedad y, consta de las siguientes operaciones: coagulación, floculación, sedimentación, filtración y desinfección. El agua tratada es suministrada a aproximadamente 50 000 habitantes del cantón.

Haciendo alusión a la zona de estudio, referente a la calidad del agua, las quebradas se han visto afectadas en los últimos años por la contaminación provocada por eventos ambientales o actividades antropogénicas en la zona (GAD de Azogues, 2018). La contaminación natural está asociada a la influencia de la precipitación, debido a que provoca el lavado de los suelos, el

arrastre de sedimentos y demás sucesos; causando el aumento considerable de la concentración de ciertos parámetros fisicoquímicos y microbiológicos en el agua cruda.

El parámetro primordial considerado en el estudio es la turbidez, ya que es uno de los principales criterios de la calidad del agua cruda y del agua tratada en la PTAP-M. En cuanto a la precipitación, en el cantón Azogues se ha cuantificado que la precipitación promedio anual oscila entre 750 – 1000 mm (Z. Romero, 2015). Según el PDOT del cantón, la zona nororiental es la de mayor precipitación, en junio de 2018 alcanzó un promedio máximo de 530 mm/mes, mientras que la zona sur occidental es la de menor precipitación mínima con 18 mm/mes en agosto del mismo año (GAD de Azogues, 2018).

Por lo antes mencionado, este estudio analizó las relaciones entre las precipitaciones y su impacto en la calidad del agua cruda, que se utiliza como fuente para el proceso de potabilización en la PTAP-M. Se analizó y comparó las precipitaciones en la zona de captación del agua cruda, la calidad del agua cruda de entrada y el esfuerzo de tratamiento químico que fluctúa en reacción al cambio de la calidad del agua cruda que ingresa a la PTAP-M. Estas relaciones se enfocan en la característica física del agua como es la turbidez; es decir que la calidad del agua cruda de entrada será regida y medida en función de la turbidez, que puede o no estar influenciada por las precipitaciones. Del mismo modo, este cambio de turbidez y de calidad del agua cruda de entrada se interpreta en un cambio en el uso de la cantidad de productos químicos coagulantes. Todas estas interacciones fueron interpretadas mediante modelos matemáticos que explican la relación entre calidad del agua superficial (turbidez), precipitación y esfuerzo del tratamiento químico en la PTAP-M. Con este fin se buscó estimar el grado de influencia que tiene los cambios en la calidad del agua superficial en los esfuerzos de tratamiento y, además, estimar el nivel de turbidez del agua cruda ante un evento de lluvia específico. Para ello, se investigó de forma enfocada en la microcuenca del río Tabacay los datos de las precipitaciones, la calidad del agua cruda y el esfuerzo del tratamiento químico durante el período 2014 – 2020. Además, se correlacionó estas variables y se encontró un modelo que explica el impacto de la precipitación en la calidad del agua cruda y en el esfuerzo de tratamiento mediante el uso de distintas metodologías y herramientas como son la regresión simple y múltiple, redes neuronales artificiales (RNA) y metodología de superficie de respuesta mediante el software Minitab. Por último, se analizó las elasticidades de la turbidez o el cambio porcentual en el esfuerzo de tratamiento resultante de un cambio del 1% en la calidad del agua cruda.

Esta investigación procura desarrollar un modelo matemático que permita predecir la turbidez del agua cruda que ingresará a la PTAP-M, conociendo únicamente el dato de la precipitación pluvial de la zona de estudio, obtenido de las estaciones meteorológicas ubicadas en la zona de captación del agua cruda. De igual forma se utilizó esta información para conocer y modelar la cantidad de coagulante a ser utilizado en función de la turbidez del agua cruda y precipitación del momento, para el futuro aprovechamiento en el proceso y tratamiento del agua cruda.

1.2. OBJETIVOS

1.2.1. Objetivo general

- Evaluar el impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en la ciudad de Azogues durante el período 2014 – 2020.

1.2.2. Objetivos específicos

- Evaluar la evolución de la calidad del agua cruda que ingresa a la planta de tratamiento Mahuarca y de la ciudad de Azogues durante los años 2014 – 2020.
- Relacionar matemáticamente la precipitación pluvial con la turbidez del agua cruda usada como indicador de la calidad en la planta de tratamiento de Mahuarca y.
- Modelar la respuesta matemática en el esfuerzo del tratamiento químico según la variación de la turbidez del agua cruda y precipitación pluvial.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Generalidades del agua

2.1.1. Importancia del agua

El agua juega un papel crucial en nuestro planeta, al ser uno de los principales constituyentes del medio natural y de la materia viva, cubre más del 70 % de la superficie de la Tierra, se la encuentra en ríos, océanos, lagos, entre otros. De igual forma, los organismos vivos se componen en mayor o menor medida, de agua, siendo entre 65 – 70 % el porcentaje medio en agua de la masa total de un organismo (Fernández, 2012).

El agua está implicada en el proceso de modelado del paisaje, al modificar la morfología, estructura y composición del medio. Es un recurso renovable pero finito y sus fuentes de agua, aunque disponibles en mayor o menor cantidad, han sido contaminadas gradualmente con el paso de los años (Fernández, 2012).

El movimiento masivo del agua en el planeta, causado por la energía solar se lo conoce como ciclo hidrológico, el cual incluye a la precipitación, escurrimiento, evapotranspiración y la infiltración. En el futuro no habrá sustentabilidad si no se conocen y tienen en cuenta todas las fases del ciclo del agua. Para ello, es necesario no solo velar por la utilización y la distribución eficiente del agua dulce cruda, sino también salvaguardar el estado de la cuenca de captación y las aguas subterráneas, así como el tratamiento y la eliminación adecuada de las aguas de desecho o residuales, (Martos, 2015).

El agua utilizada para el consumo humano brinda varios beneficios para la salud de las personas, ya que ayuda a mantener la temperatura corporal, transporte de nutrientes, así como la eliminación de desechos del cuerpo. Por ello, es recomendable beber periódicamente este líquido vital. El agua que consumen las personas debe ser absolutamente segura, no debe suponer un riesgo de ninguna clase para las personas y debe ser apta para el uso al cual se lo encomiende de manera general, sin que pueda ser perjudicial para el consumidor final de este recurso, principalmente para los lactantes y los ancianos que son el foco principal de afectados. Resulta indispensable un efectivo y correcto tratamiento del agua cruda a agua potable, antes de su distribución a la colectividad y posterior utilización, logrando evitar posibles enfermedades de carácter hídrico; principalmente en zonas de escasa pobreza donde no se

cuenta con un sistema de tratamiento de agua potable convencional, o con sistemas que no cubren con la mínima calidad de agua de consumo (Carreño et al., 2019).

2.1.2. Tipos de agua

2.1.2.1. Agua cruda

Es el agua que se encuentra en la naturaleza y que no ha recibido ningún tratamiento ni modificación en su estado natural para modificar sus características: físicas, químicas o microbiológicas (INEN 1108, 2020).

La calidad del agua cruda depende de la posición geográfica, origen (superficial y subterráneo) y hábitos de las poblaciones aledañas (Sierra, 2011).

2.1.2.2. Agua potable

El agua potable es aquella que ha recibido un tratamiento destinado a cumplir ciertos parámetros físicos, químicos y microbiológicos con el fin de suministrar a una población sin generar problemas de salud (INEN 1108, 2020). Para dar cumplimiento a lo anterior, se requiere de un conjunto de operaciones y procesos unitarios que se realizan sobre el agua cruda, con el fin de mejorar sus características organolépticas, físicas, químicas y microbiológicas. Esto incluye detectar cualquier posible contaminante microbiológico o químico, y aplicar las metodologías para que no continúe la contaminación. La calidad del agua potable depende del uso que se le vaya asignar o a dar (Cárdenas & Patiño, 2010).

2.2. Abastecimiento de agua para potabilización

2.2.1. Captación de agua cruda

La captación del agua cruda se realiza por medio de tomas de agua que se hacen en los ríos o diques, preferiblemente en la cuenca alta del río (Idrovo, 2010).

2.2.2. Línea de conducción

La línea de conducción es la parte del sistema constituido por el conjunto de conductos, obras de arte y accesorios destinados a transportar el agua cruda procedente de la fuente de abastecimiento, desde el lugar de la captación hasta la planta potabilizadora y/o tanque de almacenamiento (Vargas, 2011).

2.2.3. Planta de potabilización

Consiste de manera general en una estructura destinada al tratamiento del agua cruda con el fin de hacerla apta para el consumo humano, mediante un proceso físico-químico convencional el cual se enfoca en la remoción de sólidos suspendidos y algunos coloides presentes en el agua cruda, y una desinfección que garantice la eliminación del material patógeno; se encuentra integrada por las siguientes fases: coagulación, floculación, sedimentación, filtración y desinfección (Idrovo, 2010).

2.3. Características que determinan la calidad del agua potable para consumo humano

2.3.1. Características fisicoquímicas del agua potable

2.3.1.1. Turbiedad

La turbiedad o turbidez se origina por las partículas en suspensión, es decir, aquellas que por su tamaño se encuentran suspendidas y reducen la transparencia del agua en mayor o menor grado (Barrenechea, 2015). La turbiedad tiene un origen inorgánico (arenas, arcillas aportadas por la erosión) y un alto grado de material orgánico (microorganismos, compuestos orgánicos y otros) (Sierra, 2011).

2.3.1.2. Color

El color, aunque está íntimamente relacionado con la turbidez del agua, se lo puede considerar como una característica independiente. La turbidez es ocasionada por partículas de gran tamaño (diámetros $> 10^{-3}$ mm), mientras que el color es generado por partículas disueltas y coloides. El color es producido por la descomposición natural de la materia vegetal de las plantas y por la disolución de ciertos minerales, principalmente hierro y manganeso (Sierra, 2011).

Existen dos tipos de color en el agua: el color verdadero, es decir, el color del agua removida la turbidez (luego de que el agua ha sido filtrada), y el color aparente, aquel que presenta el agua. El color juega un papel importante en el resultado final del tratamiento del agua, debido a que su presencia causa un rechazo inmediato en los consumidores, así el agua potable presente una buena calidad. El color se expresa en unidades de color (UC) (Idrovo, 2010).

2.3.1.3. pH

El pH es el término utilizado para expresar la intensidad de condiciones ácidas o básicas del agua. El pH influye en algunos fenómenos como: la corrosión y la incrustación en las redes de

distribución; se podría decir que no tiene efectos directos en la salud de las personas, pero si puede afectar en el proceso de tratamiento del agua cruda, como la coagulación y la desinfección (Barrenechea, 2015). El pH se puede medir en campo o en el laboratorio con un pH metro, por lo general el agua cruda tiene un pH entre 5 y 9 (Idrovo, 2010).

2.3.1.4. Alcalinidad

La alcalinidad del agua es la suma de concentraciones de hidróxidos, carbonatos y bicarbonatos, expresadas en términos de carbonato de calcio, es decir, mide la capacidad del agua para neutralizar ácidos. La alcalinidad está influenciada por el pH, la composición general de agua, la temperatura y la fuerza iónica. Es un parámetro importante en el tratamiento del agua porque reacciona con coagulantes hidrolizables (sales de hierro y aluminio) durante el proceso de coagulación (Ministerio de Salud & Control de Calidad del Agua, 2013).

2.3.1.5. Dureza

La dureza es la suma de las concentraciones de iones de calcio y magnesio presente en el agua, expresados como carbonato de calcio. La dureza del agua puede ser temporal o permanente, la dureza temporal o dureza de carbonatos, es producida por presencia de bicarbonatos de calcio y magnesio, y la dureza permanente o dureza de no carbonatos se debe a la presencia de sulfatos, cloruros y nitratos de calcio y magnesio (Ministerio de Salud & Control de Calidad del Agua, 2013).

Las aguas duras como las aguas blandas no afectan la salud de las personas, sin embargo, un agua dura requiere demasiado jabón para la formación de espuma y genera problemas de lavado, además, producen incrustaciones en las líneas de conducción y mucha más contaminación hídrica (J. Romero, 2009).

2.3.1.6. Nitritos y Nitratos

El nitrógeno es un nutriente importante para el desarrollo de la flora y fauna. Por lo general, en el agua se concentra en forma de amoníaco, nitrato y nitrito. Si una fuente de agua recibe descargas nitrogenadas, el nitrógeno estará presente como nitrógeno orgánico amoniacal y en contacto con el oxígeno disuelto (OD), se transformará por oxidación en nitritos y nitratos. El proceso de nitrificación depende de la temperatura, OD y el pH del agua (Barrenechea, 2015).

Los nitratos (sales de ácido nítrico, HNO_3) y los nitritos (sales de ácido nitroso HNO_2) son muy solubles en agua. Los nitritos se transforman naturalmente a partir de los nitratos, ya sea por

oxidación bacteriana incompleta del N o, por reducción bacteriana en sistemas acuáticos y terrestres (Barrenechea, 2015).

2.3.1.7. Sulfatos

Los sulfatos son un componente principal en aguas superficiales, provienen de la oxidación de los sulfuros existentes en el agua, pueden tener su origen por las aguas que atraviesan terrenos ricos en yesos o por la contaminación de aguas residuales. El sulfato no suele ser un problema en el proceso de potabilización del agua, pero un alto contenido de sulfatos puede proporcionar sabor al agua y podría tener un efecto laxante provocando trastornos gastrointestinales principalmente en los niños, sobre todo cuando se encuentra presente el magnesio (Jiménez et al., 2010).

2.3.1.8. Cloruros

El ion cloruro es uno de los aniones inorgánicos principales en el agua natural, las aguas superficiales no contienen cloruros en altas concentraciones como para afectar el sabor. A partir de ciertas concentraciones, los cloruros pueden ejercer una acción disolvente sobre ciertas sales en el agua (Secretaría de Economía, 2001).

2.3.2. Características microbiológicas del agua

2.3.2.1. Coliformes: Escherichia coli

Escherichia coli son bacterias gram-negativo y son un tipo de bacterias coliformes fecales que se encuentran comúnmente en los intestinos de los animales y los seres humanos. La mayoría de las bacterias *E. coli* no causan enfermedad, pero si una persona ingiere agua contaminada con una alta carga bacteriana de este tipo de bacteria, el punto primario de infección es el tracto gastrointestinal y los síntomas pueden incluir náuseas, vómito, diarrea y fiebre. Estos microorganismos se eliminan en el material fecal, o las heces, y la ruta de transmisión es fecal – oral. Los alimentos y agua contaminada son las formas más comunes de ser expuestos a *E. coli* (Rock & Rivera, 2014).

La presencia de *E. coli* en el agua es un indicador de una reciente contaminación con aguas residuales o contaminación de residuos fecales de animales (Rock & Rivera, 2014).

2.3.3. Turbidez como parámetro de control de la calidad de agua

La turbidez es una medida del grado en la cual el agua pierde su transparencia, causada por partículas en suspensión, precipitados químicos, partículas orgánicas y organismos. Se muestra

que una turbidez elevada alberga y protege a los microorganismos patógenos ya que estos tienen la posibilidad de adherirse a las partículas en suspensión, estimulando así la proliferación de bacterias y afectando significativamente la eficiencia de la desinfección e, igualmente, aumentando la demanda de cloro requerida para el óptimo tratamiento (WHO, 2017). Generalmente está expresada como unidades nefelométricas de turbidez NTU y se puede medir con el turbidímetro o nefelómetro (OMS, 2019).

Los valores de turbiedad sirven para establecer el grado de tratamiento requerido para el agua cruda, la tasa de filtración más adecuada, así, como la efectividad de sus procesos de coagulación, sedimentación y filtración, de igual forma es un parámetro para determinar la potabilidad del agua (Cardona, 2011).

La turbiedad puede incidir en la eficiencia de determinados procesos de tratamiento del agua, por ejemplo, la filtración y la desinfección. La filtración del agua se vuelve más difícil y costosa cuando la turbiedad es alta; la turbiedad obstruye los filtros, genera problemas de operación en el tratamiento del agua y ocasiona un aumento en la frecuencia de lavado de las unidades. Cuando la turbiedad es alta se necesitan dosis mayores de cloro para desinfectar el agua. Los microorganismos se ocultan entre las partículas de sólidos suspendidos que ocasionan turbiedad por lo que se requieren mayores cantidades de cloro para eliminarlos. Si el agua cruda presenta valores superiores a 1000 NTU se recomienda realizar una pre-sedimentación y se establece 3000 NTU como valor máximo para un tratamiento convencional (Montoya et al., 2011).

En el Texto Unificado de Legislación Secundaria de Medio Ambiente (TULSMA), Libro VI Anexo 1, la Tabla 1 establece los límites permisibles para aguas de consumo humano y uso doméstico, que únicamente requieren tratamiento convencional, en la cual se especifica que el límite máximo permisible de la turbiedad es 100 NTU (Ministerio del Ambiente del Ecuador, 2011).

La OMS indica que la turbidez debe mantenerse por debajo de 1 NTU (OMS, 2019), para propiciar una desinfección eficaz y, en la Norma NTE INEN 1108:2020 se estipula el límite máximo permisible de turbiedad de agua potable, 5 NTU (INEN 1108, 2020).

2.4. Procesos de potabilización del agua

Las operaciones unitarias que se realizan para potabilizar el agua dependen de la calidad físicoquímica y microbiológica de la misma, el proceso de un tratamiento completo se describe a continuación:

2.4.1. Coagulación o Mezcla Rápida

Proceso a través del cual los coagulantes son adicionados al agua reduciendo las fuerzas que tienden a mantener separadas las partículas en suspensión. La coagulación comienza en el mismo instante en que se agregan los coagulantes al agua cruda. Básicamente es el resultado de dos fenómenos (Salazar, 2020):

- Químico: consiste en las reacciones del coagulante con el agua y la formación de especies hidrolizadas con carga positiva. Esto dependerá de la concentración del coagulante y el pH del agua.
- Físico: consiste en el transporte de especies hidrolizadas para que hagan contacto con las impurezas del agua.

Este proceso es muy rápido ya que tarda décimas de segundo, de acuerdo con características del agua como: pH, temperatura, cantidad de partículas, etc. Se realiza el proceso en una unidad de tratamiento denominada mezclador rápido (Salazar, 2020).

2.4.2. Flocculación

La flocculación se define con el lento movimiento del agua para permitir el crecimiento del floc. Este crecimiento es inducido por el contacto de partículas de diámetro mayor a una micra, el contacto de partículas se crea por el gradiente de velocidad de la masa líquida (Arboleda, 1992).

La flocculación persigue dos objetivos básicos (Arboleda, 1992):

- Reunir micro flóculos para formar partículas mayores con peso específico superior al agua.
- Compactar el floc (disminuyendo su grado de hidratación) para producir una baja concentración volumétrica, que permite una alta eficiencia en la etapa de sedimentación y filtración.

Este proceso se lo realiza en una unidad llamada floculador, distinto al mezclador rápido, los floculadores pueden ser de acción mecánica y otros de acción hidráulica (SAMSA, 2008).

2.4.3. Sedimentación

Se entiende por sedimentación la remoción por efecto gravitacional de las partículas en suspensión presentes en el agua, las partículas deben presentar un peso específico mayor que el fluido; en esencia es un proceso netamente físico (SAMSA, 2008).

Las partículas en suspensión sedimentan de diferente forma dependiendo de sus características y de su concentración. Por ello, se puede referir a la sedimentación de partículas discretas y sedimentación de partículas floculentas (Arboleda, 1992).

- Partículas discretas: aquellas partículas que no cambian de características (forma, tamaño, densidad) durante la caída.
- Partículas floculentas: aquellas producidas por la aglomeración de partículas coloidales desestabilizadas a consecuencia de la aplicación de agentes químicos, las características de este tipo de partículas si cambian durante la caída. El proceso de sedimentación de estas partículas también se le conoce como decantación y es el proceso intermedio entre la coagulación-floculación y la filtración rápida.

Los decantadores o sedimentadores en su tramo final poseen un vertedero en los cuales se capta la capa superior del agua que contiene mejor turbiedad para pasar al siguiente proceso (Arboleda, 1992).

2.4.4. Filtración rápida

Después de terminada la decantación el agua pasa a estanques donde se realiza la clarificación final o filtración. La filtración es un proceso físico, químico y en algunos casos biológico para separar las impurezas suspendidas mediante el paso del agua a través de un medio poroso. La filtración remueve aquellas partículas que contienen una densidad muy cercana a la del agua o que han sido re - suspendidas y, por lo tanto, no pudieron ser removidas en el proceso anterior (Okun & Schulz, 1990).

Cuando el agua ingresa al filtro, atraviesa capas porosas principalmente de arena y grava, el agua es recolectada y conducida mediante tuberías a la siguiente etapa (SAMSA, 2008).

2.4.5. Desinfección

El último proceso del tratamiento de la potabilización del agua es la desinfección. Esta etapa consiste en la inyección de cloro que permite eliminar los microorganismos que aún pueden estar presentes en el agua. Con la desinfección se busca prevenir contaminaciones en las redes

de distribución, además es un indicador de la calidad del agua. Es por ello, que la desinfección permite asegurar la calidad sanitaria del agua (SAMSA, 2008).

El cloro se inyecta a través de dosificadores automáticos en una cantidad de entre 0.6 y 0.8 miligramos de cloro por litro de agua. Esto permite la eliminación de los microorganismos que pueden quedar en el agua luego del tratamiento. El agua potable debe contener un nivel de cloro residual con el fin de evitar cualquier contaminación en el transporte del agua potable desde la planta de tratamiento hasta los hogares de los consumidores (Orellana, 2016).

2.5. Precipitaciones

Se considera precipitación a cualquier producto de condensación del agua atmosférica que llega hasta la superficie terrestre, incluye; lluvia, nieve y granizo (Segerer & Villodas, 2013). Para su formación es necesario que una masa de aire ascienda en la atmósfera, se enfríe hasta llegar a la saturación y para que pueda condensar en gotitas de agua es preciso que se cumplan dos condiciones: que la masa de aire se haya enfriado lo suficiente y que existan en el aire núcleos de condensación sobre los que puedan formarse gotitas de agua (Tapia, 2016). Los instrumentos que se utilizan para cuantificar las precipitaciones caídas sobre un punto es el pluviómetro y pluviógrafo, la unidad de medida generalmente es mm/día (Segerer & Villodas, 2013).

2.5.1. Tipos de precipitación

En relación a su origen pueden distinguirse los siguientes tipos de precipitación:

- Precipitaciones convectivas: la elevación convectiva del aire debido al calentamiento de éste por radiación solar. Se produce la condensación de vapor al alcanzar zonas más frías, este tipo de precipitación acostumbra a ir acompañado de descargas eléctricas y suele darse generalmente en zonas ecuatoriales y templadas. Dando origen a las precipitaciones en forma de chubascos o tormentas (Segerer & Villodas, 2013).
- Precipitaciones orográficas: se denomina así a la precipitación que tiene origen en el ascenso de la masa de aire forzado por una barrera montañosa. Se presentan en forma de lluvia o nieve, siendo muy irregulares en importancia y localización. A veces, en casos de masas inestables, el efecto orográfico no supone más que el mecanismo de disparo de la inestabilidad convectiva (Tapia, 2016).
- Precipitaciones ciclónicas: en este caso la elevación de masas de aire se produce sobre una masa de aire frío por un pasaje frontal o superficie de contacto. Las masas más

húmedas y cálidas se elevan hasta zonas frías, donde se condensan y originan las precipitaciones (Segerer & Villodas, 2013).

2.5.2. Pluviometría

Se denomina pluviometría al estudio y tratamiento de los datos de precipitación que se obtienen en los pluviómetros que se localizan a lo largo y ancho del territorio, obteniendo así unos datos de alto interés para la toma de decisiones en las cuencas fluviales (Duarte et al., 1991).

2.5.3. Estación Meteorológica

Una estación meteorológica realiza labores de medición y registro de los parámetros meteorológicos, con el fin de conocer el comportamiento de la atmósfera y del medio ambiente, son muy utilizadas para realizar pronósticos y la ocurrencia de fenómenos severos, además, se utiliza para pronósticos hidrológicos y meteorológicos, así como para varias investigaciones (Novoa & Guillén, 2018).

Una estación meteorológica posee diversas e innovadoras funciones que hacen que se constituya como un equipo de gran aporte en el estudio del comportamiento de las variables climáticas, las características generales de una estación meteorológica son: estructura de alto grado de protección para ambientes hostiles; registro de estadística de tiempo incluyendo datos mínimos, máximos, promedios, etc.; unidad de transmisión y recepción de datos a larga distancia; operación mediante la provisión de energía por paneles solares y por batería; almacenamiento continuo de información en un datalogger; captación de variaciones climáticas como: temperatura, humedad, presión, velocidad y dirección del viento, precipitación, entre otras (Novoa & Guillén, 2018).

2.5.4. Influencia de las precipitaciones en la calidad del agua

El clima es un factor determinante en las cuencas hidrográficas. Las variaciones climáticas pueden influir en la calidad y el acceso del agua, lo cual puede provocar el uso de otras fuentes de abastecimiento y aumentar el riesgo de la presencia de microorganismos patógenos y contaminantes en el agua que será utilizada para el consumo humano (Morales et al., 2019).

Según estudios previos se conoce que la calidad microbiológica y química del agua se ve afectada por eventos climatológicos extremos, causando un incremento significativo en la concentración de patógenos, turbidez y otros parámetros del agua; por el aumento de la escorrentía (Morales et al., 2019).

La precipitación favorece el lavado de los suelos, arrastre de sedimentos y resuspensión de los lechos de los ríos principalmente en zonas con escasa cobertura vegetal, provocando un cambio importante en la calidad del agua cruda que ingresa a las plantas de tratamiento de agua potable, en el caso de que el nivel de turbidez sea muy alto la PTAP debe impedir el ingreso del agua cruda, generando una interrupción momentánea en el proceso de potabilización del agua cruda, lo que puede repercutir en la falta de agua potable para la población y problemas económicos para la empresa potabilizadora (Morales et al., 2019).

2.6. Esfuerzo de tratamiento en la PTAP

2.6.1. Impacto de la calidad del agua cruda en el esfuerzo químico del tratamiento del agua

Indudablemente se ratifica con evidencia y experiencia profesional que la calidad del agua cruda al ingreso de la planta de tratamiento es un factor importante a considerar para la planificación, funcionamiento y toma de decisiones en toda PTAP. Diversos estudios concuerdan en su totalidad en que la fuente de agua cruda de una PTAP, es el factor más crítico para determinar los procesos de tratamiento que necesitan para la potabilización. De forma más concisa la calidad del agua cruda impacta directamente en el esfuerzo del tratamiento químico; dicha calidad va a fijar el uso de productos químicos, es decir, desde la cantidad de estos hasta el tipo de producto químico usado; lo que conjuntamente engloba e interviene en la variación de costes económicos del tratamiento. Es por ello que, en dichos casos en donde el agua cruda se clasifica como muy buena es posible prescindir de etapas específicas del tratamiento, usando menos productos químicos y economizando el proceso como tal, una circunstancia idónea señalada por los autores (Mondorf, 2020).

Como se ha indicado la calidad del agua superficial, utilizada como agua cruda en una PTAP, está comprometida por sedimentos y partículas transportadas por las masas de aguas superficiales de escorrentía y por las aportaciones de las precipitaciones. Por lo tanto, al utilizarla como fuente de agua para el tratamiento de potabilización se debe trabajar y enfrentar con este tipo de sustancias. Según dichas condiciones la calidad del agua mejora o se ve empobrecida y, del mismo modo, el costo del tratamiento se ve afectado de forma proporcional por el uso imprescindible de productos químicos versus el nivel de partículas que deben ser tratadas. El impacto y la relación que se mantiene entre la calidad del agua cruda y el esfuerzo químico se la conoce en el ámbito teórico y cualitativo o, por simple experiencia, pero, no existe estudios científicos aceptados universalmente que ofrezcan una metodología específica

para cuantificar de forma matemática el impacto de la calidad del agua cruda en el esfuerzo químico; en contraste sólo se indican relaciones o correlaciones entre estas dos variables que encaminan su comportamiento (Mondorf, 2020).

2.6.2. Esfuerzo del tratamiento químico

Esta sección se enfoca en relacionar la calidad del agua cruda según la turbidez (medida de calidad) y el uso de productos químicos coagulantes para describir el esfuerzo del tratamiento químico de una PTAP. El esfuerzo químico se ve reflejado en la cantidad de productos químicos que deben ser utilizados para potabilizar el agua en toda PTAP, considerando que esta cantidad está directamente relacionada por la calidad del agua cruda o su nivel de turbidez. Se conoce que las fluctuaciones de la turbidez del agua cruda son un contratiempo y dificultan el tratamiento y la dosificación de coagulantes, lo que deriva en un agua potable con calidad no constante (Mondorf, 2020).

2.6.2.1. Uso de Coagulantes

El uso de productos químicos coagulantes es una parte fundamental en el sistema de una PTAP, y el tipo de producto químico que se aplique en el tratamiento considera un conjunto de factores que son específicos según las condiciones de la planta, la calidad del agua cruda o la intensidad de las precipitaciones. Especificando más a detalle, la selección del coagulante depende de la concentración y características de las partículas, características física-químicas del agua cruda y, las características del coagulante, a más de consideraciones externas de costo.

De forma general existen muchos tipos de productos químicos coagulantes que se clasifican en coagulantes metálicos inorgánicos y polímeros; en este caso se hace énfasis en los coagulantes con base en el aluminio.

El uso de coagulantes es un tema de precisión porque si se trata de un coagulante metálico inorgánico, este se disociará e hidrolizará rápidamente para asociarse con las partículas del agua; sin embargo, para que este proceso sea eficiente en relación a la cantidad usada se debe considerar la dosis exacta del coagulante, pH del agua y la eficiencia de la mezcla rápida. En el caso de agregar una dosis de coagulante inapropiada se corre el riesgo de formar precipitados químicos lo que conlleva un agua de salida no apta para consumo humano y una pérdida económica referente al producto (Mondorf, 2020). Los coagulantes metálicos inorgánicos más utilizados a nivel mundial son las sales de iones de aluminio y hierro, siendo estos el sulfato de

aluminio más conocido como alumbre, el cloruro férrico y las respectivas sales prehidrolizadas de estos dos metales.

2.6.2.1.1. Sulfato de aluminio, alumbre

El sulfato de aluminio se obtiene en la reacción de un mineral alumínico con ácido sulfúrico a temperaturas elevadas; obteniéndose $\text{Al}_2(\text{SO}_4)_3$ en presentación sólida y líquida con dos especificaciones: estándar y libre de hierro. Es uno de los coagulantes más utilizados mundialmente porque reúne un conjunto de características que lo hace idóneo para el tratamiento de agua potable; debido a su eficacia, fácil disponibilidad en el mercado y su bajo costo (Silicatos y Derivados S.A, 2011).

2.6.2.1.2. Policloruro de aluminio, PAC

Para el caso del policloruro de aluminio, se trata de un coagulante inorgánico a base de sales de aluminio polimerizadas; este es un tipo de coagulante de sales metálicas prehidrolizadas que puede reemplazar al sulfato de aluminio. La eficacia y rendimiento del PAC es igual o superior al ya mencionado alumbre, en un estudio realizado en la Universidad de Córdoba (Rodríguez et al., 2019) se demostró que este coagulante disminuyó la turbidez y los sólidos suspendidos totales en un 92 % en aguas de tipo industrial. Además, a este se le suman algunos beneficios que cita Crittenden (2012) sobre las sales metálicas prehidrolizadas:

- Potencialmente con mayor eficacia; esto se resume en la necesidad de una menor dosis del producto químico.
- Los flóculos formados tienen mejores características, con respecto a su dureza y mayor densidad.
- Este tipo de coagulantes tienen menos sensibilidad al pH, pudiendo hacerse uso dentro de rangos de 4.5 – 9.5.
- No tienen tanta sensibilidad al tiempo de mezcla inicial por su naturaleza de prehidrolizados.
- Su uso presenta menores niveles de aluminio residual en el agua potable tratada.
- A comparación con los coagulantes metálicos inorgánicos, estas sales metálicas prehidrolizadas producen menos lodos residuales a dosis equivalentes.

Todas estas ventajas aceleran el proceso, reducen la dosificación y, en general, hacen más eficiente el proceso. Sin embargo, en el mercado y, a comparación del alumbre común, el

policloruro de aluminio tiene un coste mucho más elevado, por lo que las mayorías de PTAP siguen apostando por el sulfato de aluminio como principal coagulante para sus tratamientos de agua cruda.

2.6.2.2. Parámetros adicionales

Para establecer el vínculo entre la calidad del agua cruda (nivel de turbidez) y el esfuerzo de tratamiento químico (uso de coagulante) de una PTAP, es necesario trabajar con la dosificación de sulfato de aluminio y policloruro de aluminio diaria; así se conoce la relación de coagulante con la turbidez. Mediante la dosis de coagulante se puede relacionar las cantidades de consumo químico de coagulante que es necesario para un volumen determinado de agua cruda; así se logra evaluar la intensidad con la que se agregan los químicos coagulantes. Para este fin, se maneja la *Ecuación 1* para el cálculo de la dosis del coagulante (Mondorf, 2020).

$$\text{Tasa de dosis de coagulante} \left(\frac{mg}{l} \right) = \frac{\text{coagulante usado} \left(\frac{kg}{d} \right)}{\text{agua cruda tratada} \left(\frac{m^3}{d} \right)} * 1000 \left(\frac{\frac{mg}{kg}}{\frac{l}{m^3}} \right)$$

Ecuación 1 Tasa de dosis de coagulante

Adicionalmente, se determina la relación entre la dosis de coagulante y la calidad de agua definida por el parámetro de turbidez; se evalúa y deriva conclusiones sobre la coherencia y precisión de la dosificación de los coagulantes usados en la PTAP. Para establecer dicha relación se aplica la *Ecuación 2* (Mondorf, 2020).

$$R_{WT} \left(\frac{mg}{l * NTU} \right) = \frac{W \left(\frac{kg}{d} \right)}{Q \left(\frac{m^3}{d} \right) * T(NTU)} * 1000 \left(\frac{\frac{mg}{kg}}{\frac{l}{m^3}} \right)$$

Ecuación 2 Relación entre dosis de coagulante y calidad del agua

En donde,

R_{WT} = relación entre coagulante y turbidez [mg/l*NTU]

W = cantidad de coagulante usado diariamente [kg/d]

Q = caudal de agua tratada diariamente [m³/d]

T = turbidez del agua cruda [NTU]

Lo que denota la relación R_{WT} es la cantidad de coagulante (mg) que se utiliza por litro de agua cruda y por unidad de turbidez (NTU). Básicamente la *Ecuación 2* se resume en dividir la *Ecuación 1* o la tasa de dosis de coagulante para el valor de turbidez. El comportamiento de la relación R_{WT} se explica de la siguiente manera: para un valor alto de la tasa de dosis de coagulante y un valor más bajo de turbidez se obtiene un valor de R_{WT} igual o mayor a la unidad, mientras más extrema es esta diferencia mayor será el valor de R_{WT} . En el caso contrario, al valor de la tasa de dosis de coagulante menor que el valor de la turbidez se obtiene un valor de R_{WT} entre 0 – 1 y de la misma forma, mientras mayor sea esta diferencia el valor de la relación R_{WT} será más cercano a 1. Con esta relación se podrá deducir la precisión del esfuerzo del tratamiento químico o cantidad de coagulante considerando la cantidad de coagulante, caudal de agua cruda y el nivel de la turbidez; fundamentando que los valores resultantes no tienen calificación definida como un proceso óptimo u pésimo, ya que el resultado depende del peso de las variables en mención y de cómo se operó in situ en la PTAP-M por las situaciones específicas del momento.

2.6.3. Elasticidad del agua cruda

El impacto de la calidad del agua cruda en el esfuerzo químico del tratamiento de potabilización es relativamente importante y compleja; siendo un hecho que todos los investigadores concuerdan. Mediante las elasticidades del agua cruda se puede determinar el esfuerzo del tratamiento; en distintas bibliografías esta relación está basada con varios criterios como, por ejemplo: los parámetros de calidad del agua cruda como la turbidez, la carga de la cuenca con respecto a sedimentos o la clasificación del uso del suelo. Entonces, se entiende por elasticidad del agua cruda como el cambio porcentual en el esfuerzo del tratamiento químico del agua resultante al cambio del 1% en la calidad del agua cruda; con estos pesos se puede representar de forma empírica la capacidad de respuesta del esfuerzo del tratamiento químico (teniendo en cuenta que dicho esfuerzo es la cantidad de coagulante usado en un momento determinado en función al nivel de turbidez) según los cambios de la turbidez o según el cambio de la calidad del agua cruda. En la *Ecuación 3* se muestra la fórmula utilizada por Mondorf (2020) para obtener las elasticidades de la turbidez, además se sugiere adicionar el cálculo del error estándar (SE) correspondiente a cada elasticidad.

$$\text{Elasticidad de la turbidez} = \frac{\Delta \text{esfuerzo de tratamiento químico (\%)}}{\Delta \text{turbidez del agua cruda (\%)}}$$

Ecuación 3 Elasticidad de la turbidez

En la literatura se indica que ciertos estudios han determinado que la mayoría de elasticidades de turbidez son estadísticamente significativas, variando desde un rango de -0.11 a 0.30; esto hace referencia a que, en promedio, la disminución del 1 % en el nivel de turbidez del agua cruda reduce 0.14 % los costes de tratamiento por la disminución de productos químicos (Price & Heberling, 2018).

2.7. Análisis estadístico

2.7.1. Pruebas paramétricos y no paramétricos

Para lograr los objetivos de la presente investigación fue necesario analizar la data recolectada desde el 2014 hasta el 2020. Mediante las pruebas paramétricas y no paramétricas se conoce el modelo de distribución de la población, si mantienen alguna tendencia, si son datos similares entre sí, son cercanos o existen datos atípicos.

2.7.1.1. Pruebas paramétricas

Por un lado, las pruebas paramétricas son aplicadas a variables continuas y se logra identificar la distribución de los datos o de la variable aleatoria con la que se está trabajando. En estas pruebas se asume que la población de datos sigue una distribución normal; por lo tanto, si se cumple la hipótesis de la prueba se afirma que los datos tienden a la normalidad, de lo contrario, no siguen una distribución normal (Quispe, 2015). Lo que se busca preferiblemente es contar con una serie de datos que tiendan a la normalidad. Algunas características de los procedimientos para los test paramétricos son (Aragón, 2015):

- Los datos deben ser continuos y cuantitativos; es decir, que no aplica para datos tipo cualitativo.
- Se deben cumplir supuestos de forma estricta, como son: los datos de la muestra deben ser extraídos de forma aleatoria de la población global con distribución normal, y las observaciones deben ser independientes entre sí.

Dependiendo de lo que se quiera lograr con las pruebas estadísticas y del tipo de data, existen varias pruebas paramétricas que pueden ser aplicadas, entre las más empleadas están las siguientes detalladas brevemente (Bautista-Díaz et al., 2020):

- **Prueba del valor Z de la distribución normal**

Se trabaja con la curva de probabilidad estándar o la campana; en donde se ubican tres medidas de tendencia central: media, mediana y moda y define la desviación estándar. Para esta prueba se obtiene un valor Z, el que significa el número de desviaciones estándar que se desvían según la media aritmética.

- **Prueba T de Student para datos no relacionados** (muestras independientes)

Para aplicar este tipo de prueba se necesita que las observaciones sean independientes, los datos vengan de una población con distribución normal y las varianzas de los grupos deben ser homogéneas.

- **Prueba F** (análisis de varianza o ANOVA)

Esta prueba se basa en el estudio de la variación total entre los datos y su interacción entre los diversos factores; usando el estadístico F que se obtiene de la tabla ANOVA. Aquí se busca diferencias significativas entre las medias de las poblaciones o si aquellas diferencias son resultado de limitaciones del muestreo; por eso, la hipótesis nula dice que las medias de dos o más poblaciones son iguales y, su hipótesis alternativa dice que al menos una es diferente.

Como se indica en el nombre, el procedimiento de esta prueba se basa en comparar la varianza entre las medias de los grupos y dentro de los grupos; con eso se determina si los grupos son todos parte de una población o por el contrario, cada grupo es único y con características únicas.

Otra prueba paramétrica para hacer mención por su amplio uso y confiabilidad es la correlación de Pearson que actúa como un índice que mide el grado de covarianza o relación lineal entre las distintas variables; expuesta más adelante.

2.7.1.2. Pruebas no paramétricas

Por el lado contrario, se aplica las pruebas no paramétricas si no se puede determinar la distribución original de los datos y, comúnmente, se aplica cuando no se cumplen ciertas

condiciones. A continuación, se describen algunas de las características de las pruebas no paramétricas de forma general (Aragón, 2015):

- No tengan relación con los parámetros de una población.
- Que la prueba estadística que se seleccione no dependa de la distribución de la población inicial.

A comparación de las pruebas paramétricas, estas tienen menos consideraciones y son mucho más fáciles de aplicar; se las utiliza especialmente cuando las observaciones son susceptibles a poder ordenarse. Dentro de las pruebas no paramétricas encontramos algunos tipos como los siguientes en mención (Aragón, 2015):

- **Prueba del signo**

Esta prueba se la aplica para datos pareados; de forma simple, esta prueba hace uso de los signos, positivo y negativo, para reemplazar a los valores como tal y obtener un resultado final que acepte o rechace las hipótesis planteadas de normalidad.

- **Prueba del rango con signo, Wilcoxon**

En este caso también se trabaja con los signos, pero se considera la diferencia entre los pares de valores de la muestra. Wilcoxon se utiliza cuando no se satisfacen las suposiciones reglamentadas de una prueba paramétrica para dos muestras.

- **Pruebas de suma de rangos**

Este tipo de prueba se cita a la prueba de Kruskal-Wallis, se basa en el análisis de varianza ya mencionado anteriormente suponiendo que las poblaciones están distribuidas con normalidad.

- **Prueba de Kolmogorov-Smirnov**

Kolmogorov-Smirnov determina si existe una diferencia significativa para afectar la población; usando la distribución de frecuencias observadas y una distribución teórica de frecuencia.

En definitiva, la aplicación de las pruebas paramétricas y no paramétricas son un análisis estadístico que, con su uso correcto, permite conocer las características de la data para poder

trabajarla y aprovecharla de la mejor manera; las pruebas arrojan un resultado teórico que debe ser interpretado según el argumento y criterio del investigador.

2.7.2. Análisis de regresión lineal simple y múltiple

Mediante el análisis de regresión se genera una ecuación que describe la relación estadística entre uno o más predictores y la variable de respuesta; obteniendo así una ecuación específica para cada caso que tiene la capacidad de predecir las nuevas observaciones con una variable independiente de entrada. Es así como, gracias a esta herramienta, se pueden identificar y consolidar relaciones de dos o más variables y llegar a obtener esta información futura para aprovecharla en distintas áreas científicas y de la vida cotidiana (Pardo Merino & Ruiz Díaz, 2005).

2.7.2.1. Regresión lineal simple

La regresión lineal simple es la técnica estadística más elemental; con esta se encuentra la relación lineal entre la variable dependiente y una variable independiente. El propósito de esta es: en primera instancia conocer en qué grado la variable dependiente está siendo explicada por su variable independiente y, en segundo instante, se va a predecir la respuesta de la variable dependiente según el cambio de la variable independiente (Pardo Merino & Ruiz Díaz, 2005).

Entonces, con el modelo de regresión lineal simple se obtiene la línea que mejor se ajusta a los datos. De forma global este análisis simple se trabaja, primero, con la ecuación base de mínimo-cuadrática que va a describir la relación idónea entre las variables actuantes y, a este se le suma el cálculo del coeficiente de determinación R para evaluar la calidad de la ecuación o modelo obtenido (Pardo Merino & Ruiz Díaz, 2005).

El coeficiente de determinación o de correlación, conocido comúnmente como r de Pearson, expresa lo siguiente (Morales, 2011):

- R = indica el grado en que las variables x y y se ordenan o relacionan. Mientras el valor de R se acerque más a la unidad significa que la relación entre las variables es mejor.
- R^2 = el coeficiente de determinación cuadrado expresa la proporción de varianza de la variable dependiente que está explicada por la variable independiente.

2.7.2.2. Regresión múltiple

El análisis de regresión múltiple se deriva del análisis de regresión simple; la diferencia recae en que en este caso se va a establecer la relación entre una variable dependiente y una o más variables independientes, esto implica que la regresión múltiple sea una mejora de la regresión simple. De igual forma, el análisis múltiple es mucho más aplicable a la realidad porque se ajusta de mejor manera a las situaciones reales como fenómenos naturales, situaciones y procesos sociales; es decir, como se sabe estas situaciones tienen un alto nivel de complejidad que van en conjunto con un sinnúmero de interacciones entre varios factores y, en consecuencia, deben ser explicados por todas estas variables que actúan e intervienen de forma directa o indirecta (Rodríguez Jaume & Mora Catalá, 2001). Dicho esto, se debe tener en consideración la elección del modelo con el menor número de variables para explicar el comportamiento de la variable dependiente de manera más eficiente.

La ecuación matemática del modelo de regresión lineal múltiple es la que se muestra a continuación, *Ecuación 4* (Rodríguez Jaume & Mora Catalá, 2001):

$$Y = a + b_{1x1} + b_{2x2} + \dots + b_{n \times n} + e$$

Ecuación 4 Estructura base de la ecuación de regresión múltiple

En donde,

Y = variable a predecir

a, b_{1x1}, b_{2x2} = parámetros desconocidos a estimar

e = error de la predicción de los parámetros

A medida que se construye el modelo de regresión múltiple se debe incluir los estadísticos de bondad de ajuste, el análisis de varianza, etc.; con esto se logra comparar los resultados obtenidos con cada modelo para seleccionar el mejor en la etapa final. Seguidamente, se presenta estos seis elementos que hay que considerar y aplicar al análisis de regresión múltiple (Rodríguez Jaume & Mora Catalá, 2001):

a. *Coefficiente de correlación múltiple, R*

Con el coeficiente de correlación múltiple se logra medir la intensidad de la relación entre la variable dependiente y el conjunto de variables independientes con las que se esté

trabajando; a medida que se va introduciendo variables se debe calcular el R, teniendo en cuenta que las variables independientes que se van incluyendo debe ser en orden de nivel de correlación con la variable dependiente. En este punto dependiendo del valor de R se visualizará las relaciones entre las variables, buscando siempre la más alta o el valor más cercano a la unidad; en otras palabras, todas las variables independientes van a lograr explicar a la variable dependiente, pero algunas deben hacerlo con una mayor significancia y estas son las que deben ser consideradas para el modelo futuro (Rojo, 2008).

Para mayor detalle sobre el valor del coeficiente de correlación múltiple; se explica que su valor puede ir desde +1 a -1; pasando por 0 que hace alusión que entre las variables no existe asociación lineal y, por lo tanto, no se puede definir un modelo para esa situación. Por el contrario, a medida que el valor se aproxima más a +1 se demuestra que existe una asociación lineal positiva o que mientras el valor de una variable aumenta, de igual forma, aumentará el valor de la otra variable respectivamente. Y de manera viceversa, ocurre lo mismo con el valor negativo de la unidad; indicando que existe una asociación lineal negativa o que mientras aumenta el valor de una variable, en respuesta, el valor de la otra variable disminuye.

b. *Coficiente de correlación múltiple al cuadrado o coeficiente de determinación, R^2*

El coeficiente de correlación múltiple al cuadrado indica la proporción o el porcentaje de la variabilidad de la variable dependiente que es explicada por el modelo; con las variables independientes seleccionadas. Mientras se trabaja con el modelo, incluyendo o reemplazando variables independientes en este, se espera que el porcentaje de R^2 sea alto o lo más cercano a 100. Para adicionar, es importante juzgar y valorar si los incrementos del R^2 son realmente significativos al momento de sumar o restar más variables independientes (Rojo, 2008).

c. *Coficiente de determinación ajustado, R^2 ajustado*

Se trabaja con el R^2 ajustado específicamente para la regresión múltiple. Este cumple el mismo papel que el R^2 con el contraste que su valor no se ve influenciado por el número de variables que se está introduciendo al modelo. En otras palabras, el R^2 puede aumentar su valor cuando más variables independientes se está considerando en el modelo, pero este incremento puede deberse únicamente al número de variables más no a que si estas son

realmente significativas; por lo que, complementar el análisis con el R^2 ajustado brinda más confiabilidad para validar el modelo (Rojo, 2008).

d. Error típico de predicción, ETB

El error típico de predicción es aquel error que se denota por la parte de la variable dependiente que no se está explicando en el modelo. Este está ligado con el coeficiente de determinación; mientras este último aumenta el ETB disminuye (Rojo, 2008).

e. Análisis de varianza

Durante el análisis de regresión múltiple es apropiado generar la tabla de análisis de varianza para conocer si el modelo que se genera es adecuado para estimar los valores de la variable dependiente en estudio.

Para el análisis de varianza se considera la variabilidad total de la muestra descompuesta entre la variabilidad explicada por el modelo y la variabilidad residual. La tabla muestra los valores de p-value que están asociados al estadístico F (ANOVA) para cada modelo que fue generado; se busca que el valor de p-value sea menor que el nivel de significancia seleccionado, por lo general de 0.05 (Rojo, 2008).

f. Análisis de residuales

Previamente se señaló el error de la predicción de los parámetros e en la *Ecuación 4*; en este análisis se los considera para analizar más a fondo la relación y cómo actúan las variables. Los residuos de los modelos deben seguir una distribución normal y ser independientes. Para comprobar estos hechos se puede utilizar distintos métodos; desde representación gráfica con el gráfico de residuales tipificados para mostrar la distribución de estos o el gráfico de probabilidad normal. E incluso recurrir a estadísticos que arrojan un único valor; como el caso del estadístico de Durbin-Watson que mide el grado de autocorrelación entre el residuo, la observación y el anterior residuo; mientras este valor es próximo a 2 los residuos no están relacionados, si se acerca a 4 están incorrelacionados negativamente y, con un valor próximo a 0 están incorrelacionados positivamente (Rojo, 2008).

2.7.3. Redes Neuronales Artificiales, ANN

Las Redes Neuronales Artificiales, abreviada como *ANN* por sus siglas en inglés *Artificial Neural Networks* o RNA por su nombre en español, mantienen su estructura y fundamentos de las redes neuronales biológicas del cerebro humano; por lo que, tienen elementos inspirados en las funciones de la neurona biológica y se encuentran organizados de forma similar (Basogain, 2015).

Con lo ya mencionado se puede definir a una red neuronal artificial como un modelo matemático compuesto inspirado en modelos biológicos y formado por elementos procesales organizados en niveles o jerarquías que procesan información por medio de su estado dinámico como respuesta a entradas externas (Matich, 2001).

Básicamente al funcionamiento de las ANN se les asemeja con un modelo de caja negra, este busca modelar matemáticamente el comportamiento de un sistema real con funciones de transferencia conocidas y, sin poder acceder de forma directa a las dinámicas internas (Chang & Liao, 2012). Entonces, una ANN está compuesta por elementos de procesamientos simples llamados neuronas, que reciben y modifican las señales recibidas y las transmiten de forma sucesiva a través de conexiones con pesos asociados; representando la fuerza de la conexión (ASCE, 2000).

Una característica de las ANN es su capacidad de aprender de una experiencia. De esta forma las ANN pueden: primero, aprender de los conocimientos que son adquiridos por estudios, ejemplos o experiencias y cambiar su comportamiento según la situación del entorno, con esto solo se indican entradas y la ANN arroja salidas según toda la información aprendida con anterioridad. Segundo, generalizar una situación; las ANN generalizan de forma automática ofreciendo varias respuestas correctas con diferencias mínimas debido a efectos del ruido. Y, tercero, las ANN pueden abstraer, aislar o considerar por separado cualidades específicas; con esto se facilita encontrar similitudes o aspectos comunes de una serie de entradas que pudieran a simple vista carecer de ellos (Basogain, 2015).

De lo mencionado se desglosan brevemente algunas ventajas de las ANN que resalta (Matich, 2001):

- **Aprendizaje Adaptativo:** gozan de la capacidad de aprender a realizar tareas basadas en un entrenamiento previo.

- **Auto organización:** puede organizarse u ordenar la información recibida en la etapa de aprendizaje de forma autónoma.
- **Tolerancia a fallos:** en el caso de la destrucción parcial de una red, la ANN mantiene sus capacidades a pesar de los daños.
- **Operación en tiempo real:** con los cómputos neuronales realizados en paralelo es posible realizar procesos con datos de forma rápida.
- **Fácil inserción dentro de la tecnología existente:** se puede mejorar el rendimiento de la red neuronal con chips especializados.

2.7.3.1. Estructura y elementos básicos que componen la Red Neuronal Artificial

Las Redes Neuronales Artificiales tienen un elemento procesador (PE, process element); este cuenta con entradas que va combinando con una operación básica como la suma. Posteriormente, la suma de las entradas es modificada por una función de transferencia y el valor resultante es la salida del elemento procesador. La salida del PE tiene la posibilidad de conectarse con otras ANN (Basogain, 2015).

Entonces la ANN se compone de un conjunto de unidades elementales PE conectadas entre sí de formas singulares; estos a su vez, están organizados en grupos llamados niveles o capas, *Ilustración 1.*

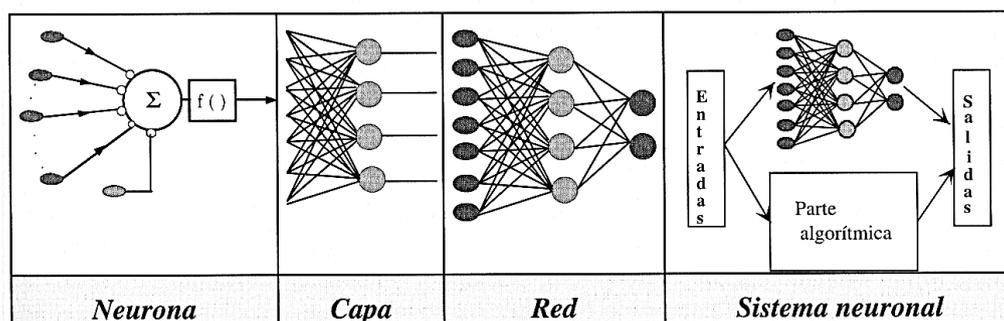


Ilustración 1 Sistema global de proceso de una red neuronal. Fuente: Larrañaga et al., 2012

Como se muestra en la *Ilustración 2*, una ANN cuenta con tres tipos de capas; una capa de entrada o input buffer, una capa de salida u output buffer y el resto de capas que son llamadas capas ocultas o hidden layer (Basogain, 2015).

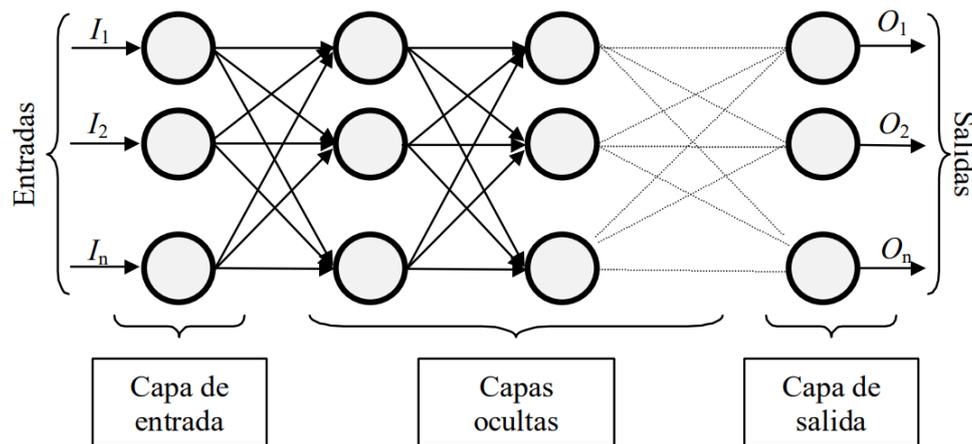


Ilustración 2 Red Neuronal Artificial totalmente conectada. Fuente: Matich, 2001

A continuación, se describe con un poco más a detalle las funciones que las ANN.

- ***Función de entrada, input function***

La ANN tiene una entrada global o varias entradas que reciben todos los datos; en esta se lleva a cabo algunas funciones de entrada como las siguientes: “sumatoria de las entradas pesadas”, en donde suma todos los valores de entrada a la neurona y lo multiplica por sus correspondientes pesos; “productora de las entradas pesadas”, aquí se lleva a cabo el producto de todos los valores de entrada y se los multiplica por sus correspondientes pesos y; por último, “máximo de las entradas pesadas” que consiste en tomar en consideración el valor de entrada más fuerte ya multiplicado por su peso (Basogain, 2015).

- ***Función de activación, activation function***

En este punto se calcula el estado de actividad de una neurona, transformando la entrada global en un valor de activación; 0 o -1 para indicar la inactividad y 1 para la actividad. En este caso las funciones de activación utilizadas comúnmente son la función lineal, función sigmoidea y la función tangente hiperbólica (Basogain, 2015).

- ***Función de salida, output function***

El valor que resulta en este punto es la salida de la neurona; y este es el valor que se va a transferir a las otras neuronas vinculadas. Este valor puede estar comprendido entre 0 – 1 o -1 – 1 o, ser de tipo binario (Basogain, 2015).

2.7.3.2. Aprendizaje y Validación

2.7.3.2.1. Niveles o capas de una Red Neuronal Artificial

Como se indicó la distribución de las ANN, *Ilustración 2*, cuenta de tres tipos de capas que se van organizando en niveles o capas como tal.

Las capas de entrada que, como su nombre lo explica, es la capa que recibe de forma directa la información externa. Las capas ocultas son las internas y las que carecen de contacto con el exterior; el número de capas ocultas puede ser desde 0 a un número superior. Por último, las capas de salida que cumplen la función de transferir la información final a otras ANN (Isasi & Galván, 2004).

2.7.3.2.2. Tipos de Redes Neuronales Artificiales

Las Redes Neuronales Artificiales se clasifican según los valores que pueden tomar; entre los principales están las Neuronas binarias y las Neuronas reales. De estas se distingue lo siguiente: por un lado, las primeras solamente trabajan y toman valores dentro del intervalo $\{0, 1\}$ o $\{-1, 1\}$; mientras que las segundas ocupan los valores entre $[0, 1]$ o $[-1, 1]$ (Isasi & Galván, 2004).

2.7.3.2.3. Mecanismos de Aprendizaje

Para el funcionamiento idóneo de una ANN es fundamental el aprendizaje o el proceso en el cual la red neuronal modifica sus pesos en respuesta a una información de entrada nueva; durante el proceso en mención ocurren cambios de destrucción, modificación y creación de conexiones neuronales nuevas. Cuando los pesos de las conexiones de la red dejan de modificarse y se mantienen estables es cuando se sabe que este proceso ha finalizado (Isasi & Galván, 2004).

En toda literatura se mencionan dos métodos de aprendizaje:

- **Aprendizaje supervisado:** en este tipo de aprendizaje se da un entrenamiento controlado por un supervisor externo que indica la respuesta que debería generar la red neuronal según una cierta entrada; es así como el supervisor controla la salida y la modifica en caso de que no se obtenga lo deseado (Basogain, 2015).

El aprendizaje supervisado se puede dar de tres formas; siendo estas: el aprendizaje por corrección de error, aprendizaje por refuerzo y aprendizaje estocástico.

- **Aprendizaje no supervisado:** este tipo de aprendizaje es auto supervisado ya que no requiere de un supervisor externo para ajustar los pesos de las conexiones entre las neuronas. La ANN no tiene información para generar la salida o saber si está correcta; por lo tanto, esta red debe encontrar las características, regularidades, correlaciones o categorías que puedan establecerse de las entradas para arrojar una salida potencialmente correcta (Basogain, 2015).

Los algoritmos que se utilizan en este tipo de aprendizaje auto supervisado son el aprendizaje hebbiano y el aprendizaje competitivo y comparativo.

2.7.3.2.4. *Validación de la Red Neuronal Artificial*

Con la etapa del entrenamiento de la ANN concluido es hora de comprobar la eficacia de la red neuronal y su capacidad para resolver nuevos problemas. Para este fin se requiere de un conjunto de datos que ejercen como el conjunto de validación o testeo del modelo (Basogain, 2015). Esta data cuenta con su entrada y respectiva salida, sin embargo, lo que se proporciona al modelo de ANN es únicamente la información de entrada para verificar y comparar la salida obtenida por el modelo con la salida original.

Para un buen desempeño de la ANN es importante la calidad de la base de datos histórica y, a este se suma la calibración y entrenamiento. De tal forma, se ajustarán constantemente los parámetros según su sensibilidad; tales como el tiempo de entrenamiento, la velocidad de aprendizaje o el número de neuronas; para alcanzar el objetivo; que los valores estimados por la ANN sean similares o muy cercanos a los valores observados (Chang & Liao, 2012).

2.7.4. **Metodología de superficie de respuesta**

La metodología superficie de respuesta (RSM por sus siglas en inglés) es un conjunto de técnicas estadísticas y matemáticas para el desarrollo de una relación funcional adecuada entre una respuesta de interés y un número de variables de control o variables independientes denotadas por x_1, x_2, \dots, x_n . Dicha relación es desconocida, pero puede predecirse mediante un modelo polinómico de bajo grado de la forma (Khuri & Mukhopadhyay, 2010):

$$y = f'(x)\beta + \varepsilon$$

Ecuación 5 Forma de la superficie de respuesta

En donde,

x = variables independientes o de control x_1, x_2, \dots, x_n .

f' = función vectorial de p elementos que consiste en potencias y productos cruzados de potencias x_1, x_2, \dots, x_n .

β = vector de p coeficientes constantes desconocidos denominados parámetros

\mathcal{E} = error experimental aleatorio

El propósito de considerar un modelo con la metodología de respuesta es triple (Myers et al., 1989):

1. Establecer una relación, aunque sea aproximada, entre y y x_1, x_2, \dots, x_n que pueda utilizarse para predecir los valores de la respuesta.
2. Determinar, mediante pruebas de hipótesis, la significación de los factores cuyos niveles están representados por x_1, x_2, \dots, x_n .
3. Determinar los ajustes óptimos de x_1, x_2, \dots, x_n que dan lugar a la máxima (o mínima) en una determinada región de interés.

2.7.5. Calibración y validación del modelo

Tanto la calibración como la validación de cualquier modelo construido es requerida para medir la calidad de las respuestas del modelo obtenidas para diferentes valores de los parámetros. De forma general se recomienda utilizar la propia data para la parte de la calibración y validación.

Por un lado, se utiliza el 80% del total de los datos para la parte de calibración y, el 20% restante es destinado para la validación del modelo. Con la validación se verifica si el modelo es capaz de predecir el valor futuro de forma correcta; comparando mi dato modelado versus el observado.

En la próxima sección se detalla los estadísticos potenciales para cumplir con la calibración y validación del modelo y así, lograr evaluar el desempeño del modelo.

2.7.5.1. Error cuadrático medio, RMSE

El error cuadrático medio mide la cantidad de error que existe entre el conjunto de dato observado y el modelado; compara cada valor observado con el valor obtenido por el modelo. La *Ecuación 6* muestra la fórmula del RMSE; su resultado puede variar desde el 0 hasta el infinito, pero mientras más cercano a 0 es mejor el resultado (Montgomery & Runger, 1994).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P - O)^2}{n}}$$

Ecuación 6 Error Cuadrático medio, RMSE

En dónde,

P = valor modelado

O = valor observado

n = número de simulaciones

2.7.5.2. Índice de Nash-Sutcliffe, EF

El índice de Nash-Sutcliffe es comúnmente utilizado para evaluar el comportamiento del modelo. El EF es muy útil porque se basa en verificar que el modelo construido tiene que ser parecido en variabilidad o R^2 y en el valor en sí; se hace énfasis en que tener como resultado en valor alto de R^2 no significa que el modelo está simulando de la forma correcta ya que puede estar subestimando o sobreestimando el valor. Según el valor que arroja el EF se tiene: entre un rango de 0.5 – 1 se considera una buena simulación, entre el rango de 0 – 0.5 es una simulación mala y un rango negativo indica una pésima simulación; donde es preferible considerar el valor del promedio de los datos observados al valor simulado por el modelo. La *Ecuación 7* indica el índice de Nash (Montgomery & Runger, 1994):

$$EF = \frac{\sum_{i=1}^n (O - \bar{O})^2 - \sum_{i=1}^n (P - O)^2}{\sum_{i=1}^n (O - \bar{O})^2}$$

Ecuación 7 Índice de Nash-Sutcliffe, EF

En dónde,

P = valor modelado

O = valor observado

2.7.6. Análisis de sensibilidad del modelo

Mediante el análisis de sensibilidad se busca estudiar los cambios que se producen en la variable respuesta cuando se modifican variaciones de los parámetros, básicamente se buscará conocer cómo el cambio en un parámetro de entrada influye o afecta en la variabilidad de la salida del modelo. Para este análisis se modificará aquellos parámetros que sean considerados

importantes e ir verificando la calidad de la variable de respuesta; con esto se encontrará el rango óptimo en el que estos fluctúan. Adicionalmente, se complementará el análisis de sensibilidad con el índice de sensibilidad “S” propuesto en el artículo (Chang & Liao, 2012). Este índice se calcula mediante la *Ecuación 8*.

$$S = \left(\frac{O_2 - O_1}{I_2 - I_1} \right) * (\bar{I} - \bar{O})$$

Ecuación 8 Índice de sensibilidad "S"

En dónde,

S = índice de sensibilidad “S”

I_1 = valor de entrada del parámetro mínimo

I_2 = valor de entrada del parámetro máximo

O_1 = valor de salida del modelo correspondiente a I_1

O_2 = valor de salida del modelo correspondiente a I_2

I_{ave} = media entre I_1 e I_2 [\bar{I}]

O_{ave} = media entre O_1 e O_2 [\bar{O}]

Se aclara que la fórmula del índice “S” trabaja con los valores mínimos y máximos del rango de iteración establecido para cada parámetro y, el valor de salida se considera al R^2 obtenido para cada iteración de los parámetros.

CAPÍTULO III

3. MATERIALES Y MÉTODOS

3.1. Descripción del área de estudio y la PTAP-M

La microcuenca del río Tabacay, zona de enfoque de la presente tesis, pertenece a la subcuenca del río Burgay, dentro de la cuenca del río Paute; está ubicada en la provincia del Cañar (*Ilustración 3*). Esta microcuenca varía su altitud desde los 2490 hasta los 3730 m.s.n.m. y se abastece de sus principales afluentes que mantienen el orden de quebradas: Llaucay, Nudpud, Cóndor Yacu, Rosario, Mapayacu y Rubís; dentro de las principales quebradas: “Llaucay”, “Nudpud”, “Concor Yacu y Rosario”, “Mapayacu” y “Tabacay” (*Ilustración 4*). En general, la microcuenca abarca una superficie total de 66.5 Km² y está administrada por el GAD del cantón Azogues (Municipio de Azogues et al., 2003). De toda el área en mención, el presente estudio se enfocó específicamente en 23.68 Km² pertenecientes a la microcuenca del río Tabacay; abarcando únicamente a las quebradas Llaucay y Nudpud (*Ilustración 5*).

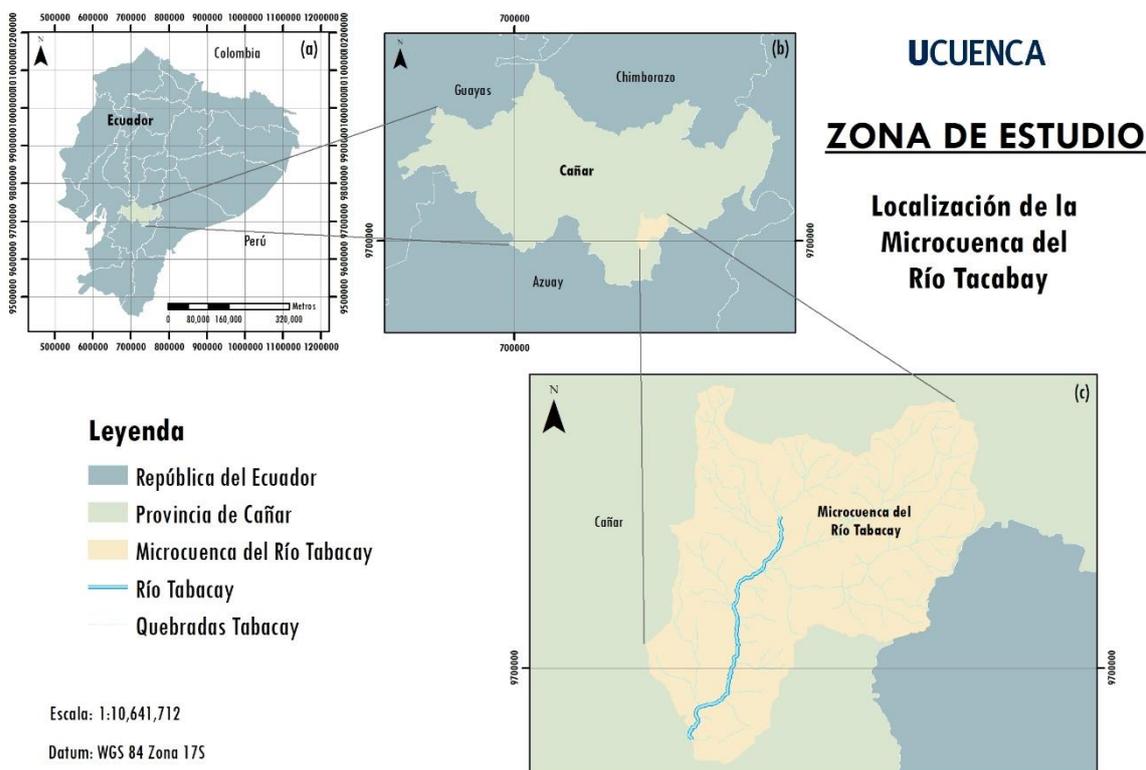


Ilustración 3 Ubicación de la zona de estudio. a) Ubicación de la provincia de Cañar, República del Ecuador, b) Microcuenca del río Tabacay en la provincia de Cañar, c) Microcuenca del río Tabacay, río Tabacay y red hidrográfica de quebradas

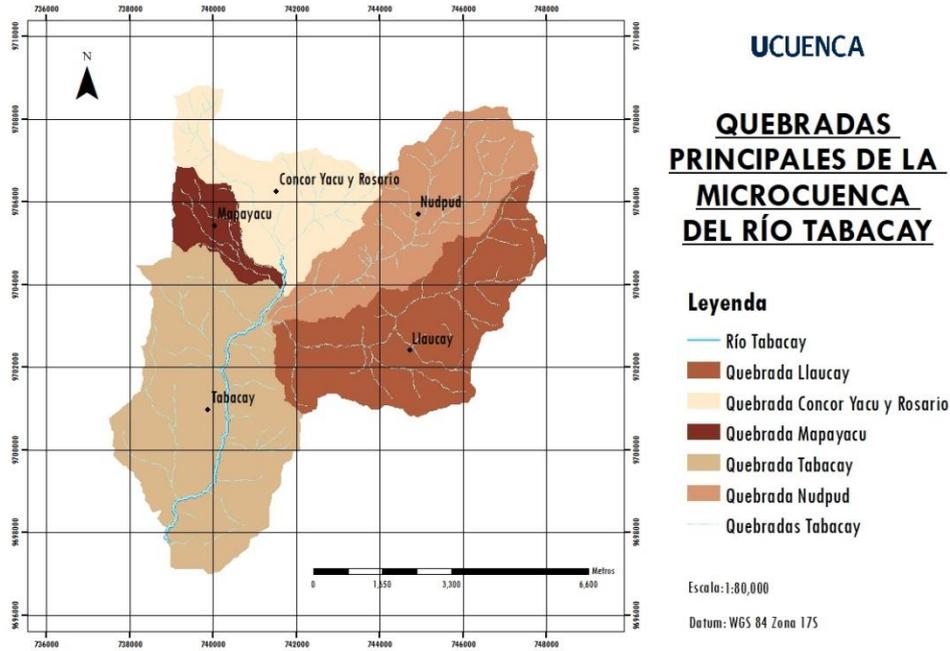


Ilustración 4 Quebradas principales de la Microcuenca del río Tabacay

Según Vásconez (2019) se clasifica a esta microcuenca como “cuenca muy pequeña” por abarcar un área menor a los 25 Km², e incluso, haciendo énfasis en la cuenca del río Tabacay de 66.5 Km² se la clasifica como “cuenca pequeña” por estar dentro del rango de 25 a 250 Km².

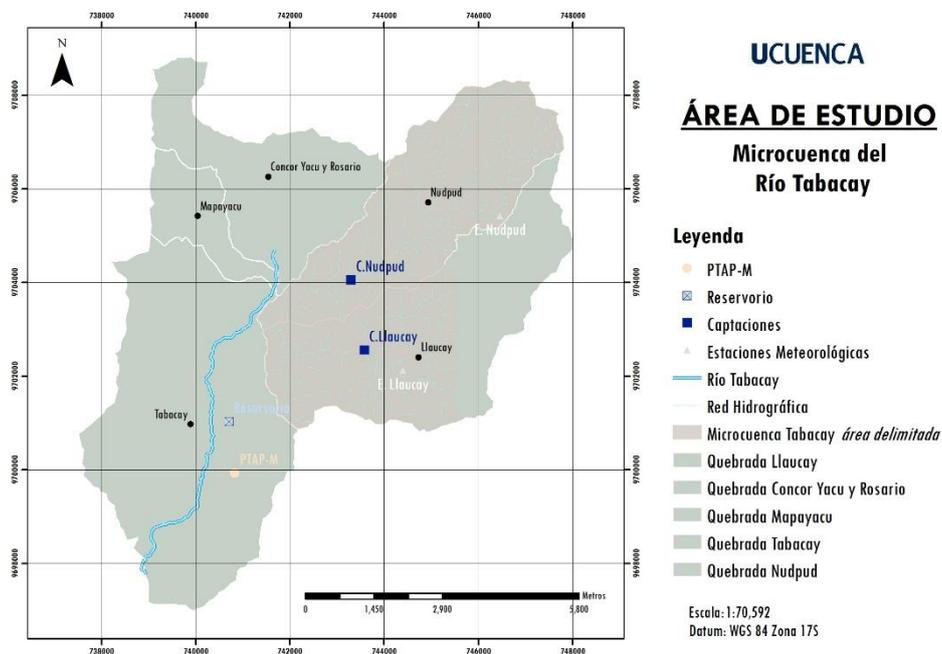


Ilustración 5 Mapa área de estudio de la Microcuenca del río Tabacay

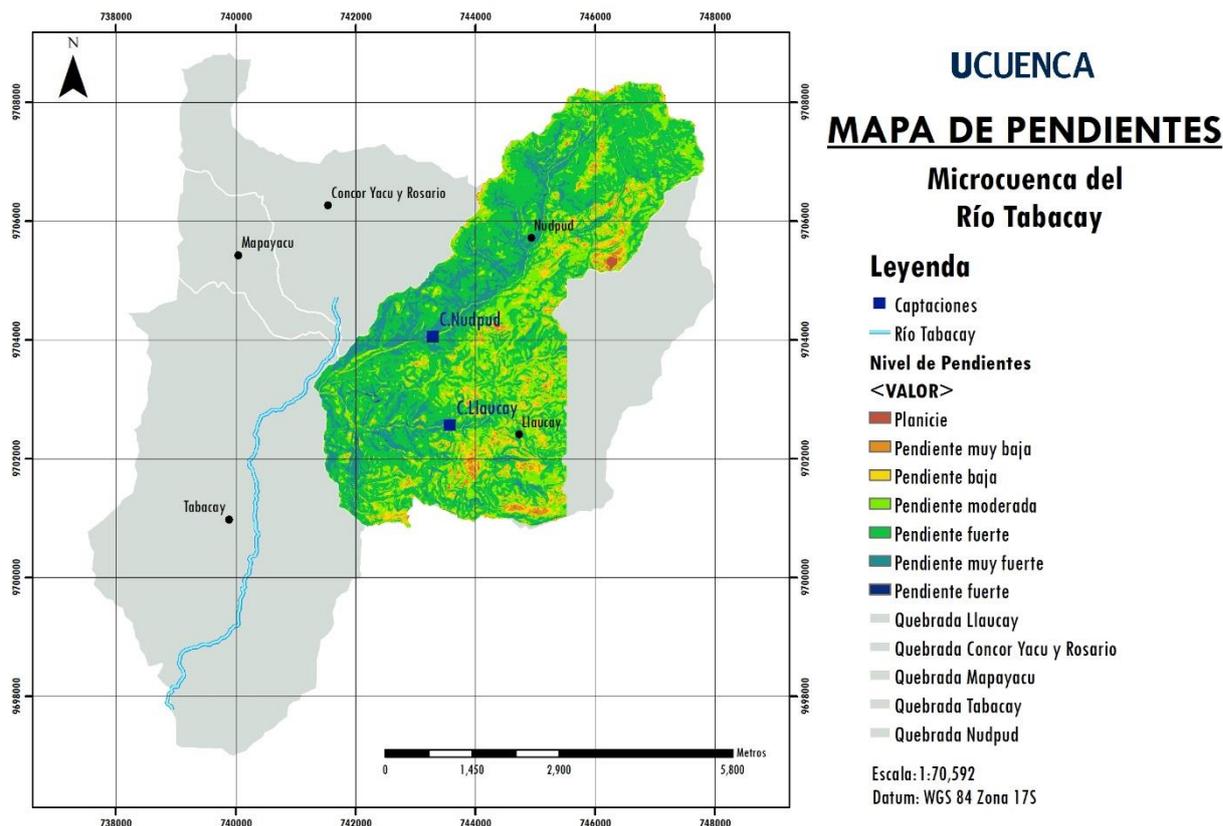


Ilustración 6 Mapa de pendientes del área de estudio de la Microcuenca del río Tabacay

Tabla 1 Reclasificación de la pendiente en porcentaje para la zona de estudio

PENDIENTE - PORCENTAJE	CONDICIONES DEL TERRENO
0 – 2 %	Planicie, sin denudación apreciable.
2 – 7 %	Pendiente muy baja, peligro de erosión.
7 – 15 %	Pendiente baja, peligro severo de erosión.
15 – 30 %	Pendiente moderada, deslizamientos ocasionales, peligro severo de erosión.
30 – 70 %	Pendiente fuerte, procesos denudacionales intensos (deslizamientos), peligro extremo de erosión de suelos.
70 – 140 %	Pendiente muy fuerte, afloramientos rocosos, procesos denudacionales intensos, reforestación posible.
> 140 %	Extremadamente fuerte, afloramientos rocosos, procesos denudacionales severos (caída de rocas), cobertura vegetal limitada.

* Denudacionales: acción y efecto de denudar o denudarse. Desprendimiento o desaparición de la parte más externa de la corteza terrestre a causa de erosión.

Se presenta en la *Ilustración 6* las pendientes de la zona de estudio, clasificándolas en base a porcentajes según la clasificación Van Zuidam como se indica en la *Tabla 1* (Rolando Mora et al., 2006). Casi la totalidad del área a trabajar se clasifica con pendiente fuerte que corresponde a inclinaciones del 30 al 70 %, le sigue las pendientes de tipo moderadas con inclinaciones del 15 al 30 %. Denotando, un potencial peligro de erosión de suelos de la zona. El grado de la

pendiente está vinculado principalmente con el grado de erosión del suelo y, de igual forma la cantidad de sedimentos que llegan a la cuenca y se transportan vía acuática por las quebradas en mención. Además, es importante mencionar que la cantidad de sedimentos está sujeta al tipo y densidad de la vegetación, cantidad e intensidad de la lluvia, a esto se le suma también la época del año.

Como se describió anteriormente, la tesis se centra en las dos captaciones de Nudpud y Llaucay (*Anexo A: Anexo 2 y Anexo 5*); siendo estas la fuente de abastecimiento para la PTAP-M. Las quebradas están localizadas a 4.8 Km de la PTAP-M en el caso de Nudpud (*Anexo A: Anexo 1*) y a 3.8 Km para el caso de Llaucay (*Anexo A: Anexo 4*). Estos cursos de agua se caracterizan por tener poco caudal y nacen de forma natural gracias a las aguas lluvia; por tal motivo, su caudal disminuye en época de verano y aumenta en época de invierno; a pesar de aquello, el caudal de ambas quebradas es suficiente para abastecer y mantener del servicio a la población de Azogues. Como se ilustra en la *Ilustración 5* el sistema de abastecimiento de agua cruda para la PTAP-M inicia en las dos captaciones que llevan el nombre de las mismas quebradas; se transporta el agua cruda por separado para cada captación por medio de dos líneas de conducción a canal cerrado hasta un punto de convergencia en donde se unen las dos líneas de conducción (*Anexo A: Anexo 8*) para formar una única línea que desagua en un reservorio revestido de concreto y descubierto a la intemperie; en total el agua recorre un aproximado de 5 Km desde el norte de la PTAP-M hasta el reservorio (*Anexo A: Anexo 10, Anexo 11 y Anexo 12*). Este reservorio mantiene una forma irregular; con 100 m de largo, 15 m de ancho y 2.3 – 3 m de profundidad. Desde el reservorio el agua cruda se dirige por 1 Km de tubería cerrada hasta la PTAP-M (*Anexo A: Anexo 13*), esta planta opera a gravedad e integra las fases del proceso de potabilización: coagulación, sedimentación, filtración rápida y desinfección. La PTAP-M usa productos químicos como el sulfato de aluminio, policloruro de aluminio, polielectrolito y cloro gas.

La PTAP-M es administrada por EMAPAL EP y está ubicada al noreste de la ciudad de Azogues, a una altitud de 2810 m.s.n.m.; específicamente en la parroquia de Bayas y en la comunidad rural de Mahuarca, a 5 Km del centro urbano de la ciudad. Esta planta se abastece del agua de las quebradas Nudpud y Llaucay. En el *Anexo A* se presenta una selección de fotografías del recorrido por las líneas de conducción Nudpud y Llaucay, el reservorio y el ingreso del agua cruda en la PTAP-M, de la misma forma una fotografía de la curva de dosificación de los coagulantes (alumbre y PAC) que utiliza la PTAP-M.

Para documentar diariamente la calidad del agua cruda y los parámetros operativos que exige el protocolo de la PTAP-M, el personal de la planta monitorea horariamente los parámetros del agua durante todo el proceso de potabilización, es decir, desde el ingreso del agua cruda hasta la salida del agua potable. Entre los parámetros medidos se enlistan: pH, color, turbidez, alcalinidad, dureza total, cloro residual, entre otros; de estos el parámetro a considerar en la presente tesis es la turbidez.

Cabe recalcar que la PTAP-M trabaja principalmente con el sulfato de aluminio, $Al_2(SO_4)_3$, o conocido comercialmente como alumbre; a este se le suma el policloruro de aluminio que es usado en menor cantidad por su elevado coste en el mercado, sin embargo, se usa de manera exclusiva en casos de niveles de turbidez superiores a 200 NTU.

3.2. Recopilación de datos

La totalidad de los datos utilizados para la presente tesis fueron obtenidos y recopilados de forma directa por los técnicos de la EMAPAL EP, quienes, a su vez, proporcionaron dichos datos. Se mantuvo contacto y se realizó visitas de campo a la PTAP-M con el fin de familiarizarse con el área de estudio, establecer contacto con los profesionales y recibir de forma directa información verbal por experiencia y la data necesaria para el desarrollo de la investigación. Se obtuvo la información registrada desde el 2005, año de inicio de funcionamiento de la PTAP-M; pero para el presente caso se trabajó exclusivamente con la información que comprende el período del 1 de enero de 2014 hasta el 31 de mayo de 2020. Desde la fecha en mención se cuenta con información de precipitación por la implementación de estaciones meteorológicas cercanas a las captaciones de la PTAP-M. En la data proporcionada por EMAPAL EP se cuenta con información completa de las operaciones dentro de la planta; registros como: control de operación, control de la solución de coagulante, control de la solución de polímero, control del lavado de los filtros, consumo de coagulante, consumo de polímero, control del caudal de ingreso y salida de agua, indicadores de operación y representación gráfica, etc. El último registro resume toda la información pertinente para la tesis; desglosando el consumo de coagulantes, ya sea el sulfato de aluminio o el policloruro de aluminio, turbidez e ingreso de agua cruda.

De forma colateral la data de precipitación fue proporcionada por personal de EMAPAL EP. Se trabajó con la estación meteorológica de Nudpud, ubicada a 3.45 Km aguas arriba del punto de la captación y la estación Llaucay a 0.93 Km al sureste de la captación del mismo nombre

(Ilustración 5). Esta colección de datos brinda la precipitación horaria, diaria y mensual; extrayendo la información desde el 1 de enero de 2014 hasta el 31 de mayo de 2020.

3.3. Evaluación de los datos de turbidez del agua cruda, precipitación pluvial y esfuerzo de tratamiento químico de la PTAP-M

3.3.1. Verificación de la representatividad de los datos

Se analizó los valores diarios de turbidez medida en el laboratorio de calidad de agua de la PTAP.M, así como la precipitación tomada de las estaciones meteorológicas ubicadas cerca de las captaciones del agua cruda y el esfuerzo del tratamiento químico medido por el personal técnico de la planta. Posteriormente, se evaluó la representatividad de los datos teniendo en cuenta el porcentaje faltante de datos y el período del muestreo.

Para el desarrollo de la presente investigación, se trabajó con un período de estudio de 6 años, obteniendo un total de 2341 días evaluados. A partir de lo anterior, se procesó los datos y se estableció el porcentaje de datos faltantes. Teniendo en cuenta las variaciones de los datos a lo largo del tiempo, se descartaron aquellos días que presentaron precipitación de 0 mm en las dos estaciones meteorológicas, ya que, no aportarían con ningún tipo de información para el análisis y de manera contraria producirían inconvenientes en el modelo para la predicción de las variables.

Seguido a la verificación de la representatividad de los datos, se realizó una evaluación descriptiva de la turbidez, precipitación (en la estación Llaucay y en la estación Nudpud) y cantidad de coagulante utilizado, teniendo en cuenta medidas estadísticas descriptivas mediante el software R Studio: en primer lugar, se determinó medidas de tendencia central tales como la media y mediana para cada una de las variables analizadas, con el fin de abreviar en un único valor el conjunto de valores estudiados. Además, se calculó medidas de dispersión como la varianza, desviación estándar, rango; siendo necesario además el cálculo del valor máximo y mínimo de cada una de las variables y, la asimetría; esta información es necesaria para conocer sobre el grado de variabilidad que presentan las distintas variables.

3.3.2. Representatividad de los datos

Por medio de la representación gráfica de las variables turbidez, precipitación en la estación Llaucay, precipitación en la estación Nudpud y cantidad de coagulante utilizado con el uso del software R Studio, se identificó si los datos presentan patrones o tendencias que no pueden ser percibidos al realizar el análisis estadístico mencionado en el apartado anterior. Es decir, se

elaboró las gráficas de densidad de cada una de las variables, lo que permitió establecer los rangos con la mayor y menor cantidad de valores.

3.3.3. Comprobación de la normalidad

El análisis de la normalidad de los datos tiene como objetivo conocer la diferencia de la distribución de los datos observados, con los que se pretende determinar si estos presentan una distribución normal con el mismo valor de media y desviación estándar.

Es necesario realizar pruebas de normalidad antes de un análisis más complejo. En este caso se trabajó tanto test de normalidad gráficos como numéricos. Cabe mencionar que, a pesar de realizar los test estadísticos, se considera el teorema del límite central que estipula, si el tamaño de una muestra es mayor a 30 datos, la distribución de las medias muestrales sigue una distribución normal; para la comprobación de este teorema se calculó la normalidad de los datos. En este caso específico se utilizó la prueba de Lilliefors, este test se aplica para datos mayores a 50, además, presenta correcciones al test de Kolmogorov Smirnov, que asume la media y desviación estándar poblacional como conocida y, en la mayoría de los casos es imposible. Lo que se pretende al realizar esta prueba es no rechazar la hipótesis nula (datos siguen una distribución normal), si el valor p es mayor a 0.05. La prueba de Lilliefors se calculó mediante el software R Studio, con el comando `lillie.test`.

3.4. Análisis de correlación

Se calculó la correlación que existe entre las variables turbidez, precipitación de Llaucay y precipitación de Nudpud mediante el coeficiente de correlación de Pearson, el cual mide la relación estadística entre dos variables. El coeficiente de correlación puede variar desde -1 a +1. Si toma un valor de 0 indica que no existe relación entre las variables, un valor cercano a +1 indica una relación positiva y conforme aumenta el valor de una variable aumenta la otra y, un valor negativo indica que si aumenta el valor de una variable la otra variable disminuye.

Este análisis se realizó con el fin de conocer la presencia de relaciones directas (positivas) o indirectas (negativas) entre las variables.

3.5. Modelación por regresión

3.5.1. Modelo de Regresión simple y múltiple

Mediante el software R Studio se realizó la regresión simple con los datos de calibración entre la turbidez (variable dependiente) y la precipitación (variable independiente) de las dos

estaciones meteorológicas, para conocer esta relación de manera cuantitativa se calculó el coeficiente de determinación. Para ello fue necesario formar conjuntos de datos para el análisis: el primer conjunto correspondió a la turbiedad del agua cruda y la precipitación de la estación de Llaucay; el segundo conjunto fue entre la turbiedad del agua cruda y la precipitación de la estación de Nudpud; por último, el tercer conjunto fue entre la turbiedad del agua cruda, la precipitación de Llaucay y precipitación de Nudpud, para ello se realizó la suma de las precipitaciones entre ambas estaciones meteorológicas quedando así un único valor de análisis para la variable precipitación.

De igual forma, se utilizó la regresión simple para el análisis entre el esfuerzo de tratamiento químico y la turbidez del agua cruda, así como para la relación entre esfuerzo de tratamiento químico y precipitación; para ello, se utilizó como se mencionó anteriormente únicamente el valor resultante de la suma de las precipitaciones de las dos estaciones meteorológicas. Se determinó que la variable dependiente fue el esfuerzo de tratamiento químico o cantidad de coagulante utilizado y la variable independiente fue la turbidez del agua cruda y la precipitación en el primer y segundo suceso respectivamente.

En el caso de la regresión múltiple, para estimar la respuesta del esfuerzo de tratamiento químico de la PTAP-M en base a la calidad del agua y factores ambientales externos de forma incorporada se optó por un modelo enfocado en la regresión múltiple. Es decir, se estimó el uso de la cantidad del coagulante con los datos de calibración según la turbidez y la precipitación con un modelo. Para el modelo se determinó variables dependientes (esfuerzo del tratamiento químico) e independientes (turbidez del agua cruda), así como variables de control (precipitación), según el tipo de influencia que tienen en el sistema.

Es necesario incluir variables de control dentro del modelo de regresión múltiple para mantener los factores que puedan sesgar los parámetros de interés. Este tipo de variables pueden ser variadas como volúmenes de agua procesada, condiciones climáticas, tipo de agua de origen, propiedades del sistema y las características de la población. Para el presente estudio se consideró una variable de control descrita brevemente a continuación:

- La precipitación fue una de las principales variables para este modelo que afectó a los resultados. Esta a su vez, condiciona la movilidad de sedimentos y partículas de la cuenca por la influencia de su intensidad y el volumen de lluvia; estas partículas se transportan y llegan al punto de captación afectando a la turbidez del agua cruda y consecuentemente el tratamiento de potabilización como tal. Además, en base a la

experiencia profesional en la PTAP-M y por las pruebas de jarras, se ajusta y se anticipa la cantidad de coagulante en función de la precipitación del momento. Se ha estimado de forma experimental que tras un evento de precipitación esta se refleja después de 2 a 4 horas en la turbidez de la planta.

La literatura argumenta que la precipitación es una variable de control significativa para reducir el riesgo de una estimación sesgada; porque esta aumenta la escorrentía de contaminantes hacia los puntos de captación.

Para el análisis de la regresión simple y múltiple se consideró el valor del R^2 ajustado para conocer la correlación de los datos, porque el R^2 ordinario aumenta si incrementa el número de variables independientes, por lo que el modelo es mejor por incluir más variables; se usa este criterio para proteger al modelo de un sobreajuste.

Además del cálculo del coeficiente de determinación, se calculó los estimadores del error como: el error cuadrático medio (RMSE) y el índice de Nash (EF) mediante las *Ecuación 6* y *Ecuación 7* respectivamente, la determinación de estos estimadores se realizó tanto para la data de calibración como para la data de validación. Con el desarrollo de todo lo mencionado anteriormente se culminó con ecuaciones de regresión simple y múltiple que pueden ser aplicada en la PTAP-M.

3.5.2. Naturaleza de modelos

Las regresiones lineales y múltiples se las realizó con distintos modelos para la etapa de calibración, en cambio, que para la etapa de validación se utilizó la ecuación ya generada en la etapa de calibración para únicamente calcular los estimadores del error y conocer que tan bueno es el modelo para usos futuros.

En este caso, cada conjunto de regresión mencionado anteriormente, trabajó con siete modelos: modelo lineal, modelo lineal sin intercepto, modelo cuadrático, modelo cuadrático sin intercepto, modelo exponencial, modelo cúbico y modelo cúbico sin intercepto. No se calculó el modelo logarítmico ya que, este se lo puede aplicar únicamente si la variable independiente o de control del modelo no incluye valores de cero y en el presente estudio existe precipitaciones de 0 mm.

Adicionalmente, para cada conjunto de datos analizados, del modelo que se consideró como el mejor, se graficó su diagrama de dispersión para visualizar la curva de regresión y el R^2 .

3.6. Redes Neuronales Artificiales

El uso de redes neuronales artificiales, abreviadas como RNA, resulta prometedor para la estimación de variables ambientales como la turbidez en relación con la precipitación; llegando a predecir un valor de turbidez del agua cruda en la PTAP-M en función de los datos pluviométricos de las estaciones meteorológicas aguas arriba; así como el esfuerzo de tratamiento químico en función de la turbidez del agua cruda y la precipitación. Debido a que la RNA tiene la capacidad de aprender de los ejemplos sin necesidad de una física explícita. Básicamente, el funcionamiento de las RNA se asemeja con un modelo de caja negra, este busca modelar matemáticamente el comportamiento de un sistema real con funciones de transferencia conocidas y sin poder acceder de forma directa a las dinámicas internas (Chang & Liao, 2012). Entonces, una RNA está compuesta por elementos de procesamientos simples llamados neuronas que reciben y modifican las señales recibidas; para posteriormente transmitirla de forma sucesiva a través de conexiones con pesos asociados, representando la fuerza de la conexión (ASCE, 2000).

Para un buen desempeño de las RNAs es importante la calidad de la base de datos histórica y a este se suma la calibración y entrenamiento. De tal forma, se ajustó constantemente los parámetros según su sensibilidad; tales como el tiempo de entrenamiento, la velocidad de aprendizaje o el número de neuronas que permiten alcanzar el objetivo y para que los valores estimados por la RNA sean similares o muy cercanos a los valores observados (Chang & Liao, 2012).

Este estudio utilizó el software R Studio, específicamente el paquete neuralnet, para desarrollar una red neuronal de entrenamiento de tres capas, es decir, una capa de entrada, una capa interna y una capa de salida. En el primer caso se utilizó únicamente a la precipitación como variable de entrada y la turbidez del agua cruda como variable de salida; en el segundo caso las variables de entrada fueron la turbidez del agua cruda y la precipitación y el esfuerzo del tratamiento químico como variable de salida.

Una vez definido los valores óptimos de los parámetros examinados dentro del análisis de sensibilidad, se procedió a correr el código en el software modificando el número de neuronas dentro de la capa interna, es decir, se obtuvo redes neuronales con distinta arquitectura. Para el caso de la predicción de la turbidez, se utilizó 6 modelos con distinto número de neuronas y, para la predicción del esfuerzo del tratamiento químico se trabajó con 7 modelos con distinto número de neuronas cada uno. Adicionalmente, se graficó con el mismo software, la estructura

base del RNA, el sistema global final de la RNA y una comparación de los datos observados y obtenidos con el modelo calculado.

Para la determinación del sesgo de los modelos obtenidos, se calculó los estimadores de error para cada uno de ellos, tanto para la etapa de calibración como para la etapa de validación. En base a este resultado, se conoció cuál fue el modelo que mejor se ajustó a la realidad de los datos y se presentó la ecuación de los mismos.

3.6.1. Análisis de Sensibilidad

El análisis de sensibilidad de los modelos permite identificar los parámetros con mayor sensibilidad a cualquier cambio para obtener el mejor conjunto de combinación de los diferentes parámetros. Para un correcto análisis de sensibilidad (Chang & Liao, 2012) menciona que cada parámetro a ser considerado debe ser independiente.

Para el presente estudio se trabajó el análisis de sensibilidad en dos partes, teniendo en cuenta que la arquitectura base de la RNA con la que se trabajó en la primera parte es de tres capas: una capa de entrada, una capa oculta y una capa de salida, y en la capa oculta se trabajó únicamente con tres neuronas constantes para todas las pruebas de sensibilidad de los demás parámetros. En la primera parte se estudió y analizó tres parámetros: tiempo de entrenamiento, velocidad de aprendizaje y función de transferencia. En la *Tabla 2* se aprecia que existen diferentes escenarios que fueron evaluados para seleccionar la mejor opción de parámetro. Para el tiempo de entrenamiento se trabajó con cinco escenarios que corrió desde 500, 5000, 10000, 15000 hasta 20000; para la velocidad de aprendizaje se trabajó con cuatro escenarios: 0.001, 0.01, 0.1 y 0.5; por último, la función de transferencia se trabajó exclusivamente con un escenario (*Tabla 2*).

Tabla 2 Escenarios considerados para el análisis de sensibilidad para las RNA

<i>Escenarios para el análisis de sensibilidad</i>		
<i>Parámetros de la RNA</i>	Número de escenarios	Escenarios
<i>Tiempo de entrenamiento</i>	5	500, 5000, 10000, 15000, 20000
<i>Velocidad de aprendizaje</i>	4	0.001, 0.01, 0.1, 0.5
<i>Función de Transferencia</i>	2	Log Sigmoide (S), Tangente hiperbólica (T)

Elaboración propia

Posteriormente, para la segunda parte, se enfocó en el parámetro de número de neuronas que debe ser considerado en las capas ocultas de los modelos de RNA para optimizar los resultados. En este punto ya se consideró los resultados de la sensibilidad de los parámetros y se trabajó

con el conjunto de los mejores escenarios de cada parámetro; siendo estos valores constantes y, ahora el número de neuronas fue variando hasta encontrar el mejor resultado; se hace hincapié en que las variaciones de todos los parámetros van en función de la literatura pertinente y según las observaciones del momento.

Aclarado el análisis de sensibilidad, se complementa este análisis de permutar y variar con un índice de sensibilidad utilizado por Chang & Liao (2012). El índice de sensibilidad “S” (mencionado en el Capítulo II “*Marco Teórico*”) presenta un solo valor que permite evaluar el grado de efecto de los parámetros en la RNA. Se empleó la *Ecuación 8* para obtener este índice para cada parámetro estudiado; tiempo de entrenamiento, velocidad de aprendizaje, función de transferencia y número de neuronas; e interpretando al valor absoluto del índice “S” a modo que mayor este valor, mayor es el efecto del parámetro de entrada sobre la salida. El ingreso de los datos para la ecuación fue de la siguiente manera: para las variables entrada *I* se consideró el rango (valor de los escenarios) de cada parámetro y para las variables de salida *O* se consideró como valor de salida al estadístico R^2 de cada parámetro.

3.7. Metodología de superficie de respuesta

El fundamento teórico en el que se asienta la metodología de superficie de respuesta está detallado en el Capítulo II “*Marco Teórico*”; sección “2.7.4. *Metodología de superficie de respuesta*”. Se optó por utilizar una herramienta muy funcional, así como eficaz en tiempo, versatilidad y complejidad; el software Minitab examina datos actuales y pasados para descubrir tendencias, encontrar y predecir patrones, descubriendo relaciones entre variables. Minitab permite realizar gráficas, estadísticas básicas, regresiones, análisis de varianza, análisis de sistemas de medición, herramientas de calidad, diseño de experimentos, fiabilidad y demás posibilidades. Entre todo lo señalado, se realizó la metodología de superficie de respuesta mediante las características y módulos del programa; necesitando únicamente la data de entrada no codificada (turbidez, precipitación y esfuerzo de tratamiento químico), trabajando con las mejores opciones que brinda el programa. Finalmente, se obtuvo de forma instantánea una ecuación para los modelos que representan las relaciones entre turbidez – precipitación y esfuerzo de tratamiento químico – turbidez – precipitación.

3.8. Calibración y validación

Como se ha comentado en el Capítulo II “*Marco Teórico*” la calibración y validación es necesaria para verificar la calidad del modelo; en otras palabras, es necesario entrenarlo y

calentarlo con un porcentaje mayoritario de la base de datos para posteriormente validar su funcionamiento con el porcentaje restante, verificando si las respuestas de salida son comparables con las respuestas originales.

Para la presente tesis se utilizó como herramienta el software R Studio para realizar las dos etapas en mención para todos los modelos que fueron trabajados continuamente. En total se trabajó con una base de datos válidos de 2026 de los 2341 datos originales; de forma aleatoria (porque el tipo de data permite trabajar de forma aleatoria sin alterar las relaciones y los futuros resultados) el programa fraccionó un 80 % de los datos que es igual a 1620 para calibración y el 20 % igual a 406 valores para la parte de validación.

Adicionalmente, se complementa el análisis de los modelos en las dos secciones de calibración y validación, mediante la aplicación de los estadísticos detallados el Capítulo II “*Marco Teórico*”, secciones “2.7.5.1. *Error cuadrático medio, RMSE*” y “2.7.5.2 *Índice de Nash-Sutcliffe, EF*”. Proporcionando así una respuesta cuantitativa que demuestra la calidad del modelo en cuanto a su ajuste, precisión y predicción.

3.9. Elasticidades del agua cruda de la PTAP-M

Primeramente, se analizó la cantidad de coagulante utilizado en la PTAP-M durante el período estudiado conjuntamente con los datos válidos de precipitación y turbidez para el correspondiente análisis inicial del esfuerzo de tratamiento químico. Como ya se mencionó en los apartados anteriores, no se trabajó con los datos correspondientes a los días con precipitaciones igual a cero en ambas estaciones meteorológicas para evitar problemas con el valor de la turbidez cuando existe una precipitación nula; de la misma forma, con los datos de aquellas fechas que no tenían datos de entrada o eran datos negativos de la cantidad de coagulante utilizado en el momento.

Seguidamente, se transformó el valor de la cantidad de coagulante utilizado por día; es decir, se unificó los pesos de los dos tipos de coagulantes usados en la planta: sulfato de aluminio y policloruro de aluminio (PAC). Esta transformación fue realizada para modificar al policloruro de aluminio e igualarlo o convertirlo en cantidades equivalentes de sulfato de aluminio y poder mantener una comparación que sea posible y fidedigna entre estos dos coagulantes. Para este fin se puede basar en los criterios descritos por Sari et al. (2017), entonces, es posible dividir los valores de PAC para un factor de conversión igual a 0.33 (valor seleccionado por la literatura en mención); se recalca que este valor fue estimado en base a la bibliografía existente

para este caso en particular; otro caso es el de Gebbie (2001) que señala un factor de conversión igual a 0.2. Por último, según el conocimiento empírico de los operadores y la curva de dosificación del sulfato y PAC de la PTAP-M se determina que la relación entre policloruro de aluminio a sulfato de aluminio se equipara un valor promedio entre dosis PAC/dosis de sulfato de aluminio mediante un factor de conversión de 0.39. Comprendiendo que cualquiera sea el método seleccionado para los posteriores cálculos, este tiende a traer consigo un mínimo de error (sobrestimando la cantidad de coagulante) que no resulta significativo según la literatura; pero para la presente tesis, y considerando las particularidades de la PTAP-M, se seleccionó como factor de conversión el valor de 0.39.

Entonces, para aquellas fechas que tenían el uso únicamente de sulfato de aluminio (sabiendo que este es el coagulante principal en la PTAP-M) se tomó directamente este valor, para las fechas restantes que a más de haber usado el sulfato de aluminio se complementó con el PAC; se procedió a dividir el valor de PAC para 0.39 (factor equivalente entre dosis PAC/dosis sulfato de aluminio) y sumarle el valor original del sulfato de aluminio utilizado en el día, esto se expone en la *Ecuación 9*.

$$\text{Coagulante total} = \frac{\text{PAC}}{0.39} + \text{Sulfato de aluminio}$$

Ecuación 9 Cantidad de coagulante total

En dónde,

Coagulante total = valor final de la cantidad de coagulante usado por día (sulfato de aluminio + PAC) [kg/d]

PAC = cantidad de policloruro de aluminio utilizado [kg/d]

Sulfato de Aluminio = cantidad de sulfato de aluminio utilizado [kg/d]

Ya con esto se procedió a realizar los respectivos cálculos para el análisis del esfuerzo de tratamiento químico en la PTAP-M. Se calculó la tasa de dosis de coagulante en mg/l para verificar la cantidad de coagulante usada por unidad de agua tratada o agua cruda, utilizando la *Ecuación 1* expuesta en el Capítulo II “*Marco Teórico*”. Mediante el cálculo de la dosis de coagulante requerida por día en la PTAP-M se logró evaluar la intensidad con la que se está usando los coagulantes. Con el fin de corroborar lo anterior, se calculó la relación R_{WT} entre la cantidad usada de coagulantes, la cantidad de agua cruda ingresada y la turbidez, *Ecuación 2*, en otras palabras, se relacionó la dosis de coagulante recién calculada con la calidad del agua cruda. Con esto se logró un primer vistazo para verificar la proporción entre estas variables y

mantener una idea de su relación, además, sobre la precisión de la dosificación de los químicos coagulantes.

Referente a las elasticidades de la calidad del agua cruda o elasticidades de la turbidez, se aplicó la *Ecuación 3* teniendo claro que la elasticidad es igual al cambio porcentual en la cantidad de coagulante usado (esfuerzo de tratamiento químico) sobre el cambio porcentual de la turbidez del agua cruda. Acotando lo mencionado y lo que expresa la *Ecuación 3*; se entiende que se trabaja con los cambios porcentuales para evitar problemas de unidad, así se divide las dos variables para una misma base y se obtiene un resultado adimensional con el que se puede trabajar y se presta para una clara interpretación. También, es importante recalcar el significado del *cambio porcentual*, este cambio expresa el cambio relativo de una variable con respecto al tiempo; es decir entre un valor antiguo y el nuevo. Para el presente caso este cambio se lo consideró en el sentido arriba hacia abajo o del día anterior al día siguiente; aunque cualquiera sea el sentido seleccionado (viceversa) la respuesta no varía debido a que el cambio permanece igual. Entonces, la relación de la elasticidad de la turbidez implica que a cualquier cambio de la turbidez se modificará la cantidad demandada de coagulante.

El valor de la elasticidad de la calidad del agua cruda puede variar desde cero en adelante y su interpretación más simple puede fundamentarse en ser una elasticidad o muy elástica o muy inelástica. Cuando se trata de una elasticidad muy elástica se entiende que un cierto cambio porcentual de la turbidez causa un alto cambio porcentual en la cantidad de coagulante requerida, mientras que, por el contrario, una elasticidad muy inelástica se entiende cuando al mismo cierto cambio porcentual de la turbidez causa un bajo cambio porcentual en la cantidad de coagulante requerida.

3.10. Incertidumbres del caso de estudio

Se considera que la data con la que se cuenta para el presente estudio (abarcando data de turbidez, precipitación y toda la concerniente al esfuerzo de tratamiento químico) incluye errores de tipo sistemático y aleatorio como comúnmente sucede en todos los casos de estudio científicos. Específicamente, para el presente caso se percibió en primera instancia que no era posible trabajar con la serie de tiempo completa (1/enero/2014 – 31/mayo/2020) por la ausencia de datos e incoherencias en la misma. Estos errores siempre implican que los resultados estén propensos a cierto grado de error y desviación, también, a subestimaciones o sobreestimaciones. Dicho esto, en esta sección, se pretende identificar y exponer posibles fuentes de error en el análisis de la base de datos cuantitativos; evidentemente no se mantuvo

un control en las etapas de recopilación y almacenamiento de datos porque la información es de tipo recopilatorio de años posteriores. De igual forma, se conoce empíricamente que existen fuentes de error de otra índole que siempre influyen en los resultados. A continuación, se muestran las fuentes de errores potenciales:

- La base de datos pertenece a una recopilación de siete años (2014 – 2020) de funcionamiento de la PTAP-M. Durante este período de tiempo se sabe que la generación, registro y almacenamiento o compilación de la información fue por medio de diferentes miembros del personal de la planta. Esto abre la posibilidad de generar un error de operador; no asegurando que, durante el muestreo, las mediciones del nivel de calidad del agua, la determinación de los parámetros de calidad, el registro de valores y la recopilación en los registros; haya sido correcta la generación de información. Esto no certifica la veracidad de los datos al cien por ciento. Sumado a esto último, también pudieron existir errores de medición por muestreos incorrectos, resultados imprecisos, falta de calibración del equipo, perturbaciones externas, errores de registro, entre otros errores; con todo este conjunto no se puede evaluar la exactitud de los datos.
- La existencia de valores individuales o series de valores que fueron considerados poco intuitivos e irreales. Esto se defiende porque existían valores, por ejemplo; una serie de días consecutivos que no existía uso de coagulante o, por el contrario, una cantidad relativamente exagerada para el primer día y valores nulos para los siguientes. En estos casos se optó por no considerar los valores inciertos o repetidos; aunque el eliminar dichos valores se considera una incertidumbre dentro del estudio.
- Referente a los datos de precipitación de las estaciones meteorológicas en las dos captaciones, también pueden estar sujetos a errores aleatorios o sistemáticos, pero muestran una buena concordancia entre sí.

Con esto se exhibe brevemente las posibles fuentes de error para el estudio y la posible limitación informativa que puede caracterizar al estudio actual. Estos tipos de errores no son evaluados ni apreciados y los resultados futuros no presentan una exactitud de la realidad de la PTAP-M y de sus condiciones meteorológicas, la calidad de agua cruda o del esfuerzo de tratamiento químico; pero se pueden hacer afirmaciones relativamente aceptables a base de los resultados por mantener y seguir correctamente la metodología. E incluso, este estudio puede ofrecer información para mejorar condiciones dentro de la PTAP-M con relación a la obtención de datos o simplemente ofrecer la información para generar futuras observaciones, toma de decisiones y demás.

CAPÍTULO IV

4. RESULTADOS

A continuación, se presentarán los resultados obtenidos para cada uno de los objetivos planteados con su respectivo análisis.

4.1. Evaluación de los parámetros de la PTAP-M

4.1.1. Verificación de la representatividad de los datos

Esta sección se conformó por dos etapas, en la primera de ellas se recopiló la información de las variables de estudio, en este caso: precipitación (estación Llaucay y Nudpud), turbiedad, y cantidad de coagulante expresado como sulfato de aluminio de la PTAP -M; seguidamente se analizó la representatividad de los datos y evaluación descriptiva de las variables.

Para el período de estudio comprendido entre enero de 2014 y mayo de 2020, se recopiló 2341 datos diarios de precipitación, turbidez y esfuerzo del tratamiento químico. Cabe mencionar que se eliminaron datos que presentaban precipitación de 0 mm en las dos estaciones meteorológicas, en base a las consideraciones mencionadas por Mondorf (2020), en total se eliminó 315 datos del análisis, debido a que esos valores podrían influenciar de manera negativa en el resultado del modelo. El 87 % de los datos analizados se los consideró como válidos, es decir, 2026 datos en total. En la *Tabla 3* e *Ilustración 7* se especifica lo mencionado anteriormente.

Tabla 3 Representatividad de los datos

Período de muestreo	Inicio	01/01/2014
	Fin	31/05/2020
Datos	Válidos	2026
	%	87
	Faltantes	315
	%	13
	Total	2341

Elaboración propia

Posterior a la verificación de la representatividad de los datos, se realizó una evaluación descriptiva de las distintas variables mediante medidas estadísticas. A continuación, se presentan los resultados de turbidez en la *Tabla 4*, precipitación medida en las dos estaciones meteorológicas y esfuerzo de tratamiento químico expresado como sulfato de aluminio.

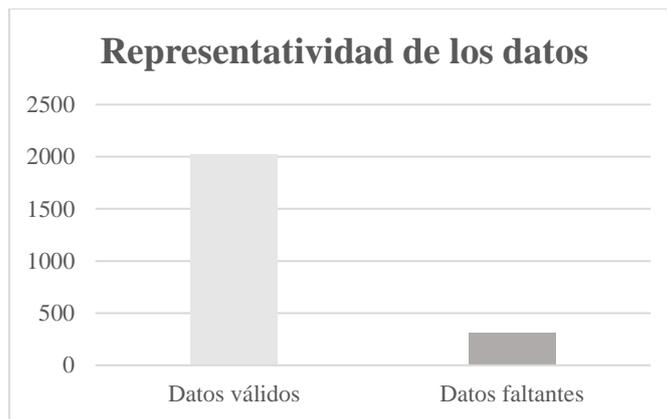


Ilustración 7 Representatividad de los datos
Elaboración propia

Tabla 4 Resumen de la evaluación descriptiva de los datos analizados en la PTAP-M

Parámetro	Unidad	Media	Mediana	Varianza	Desviación estándar	Mínimo	Máximo	Rango	Asimetría
<i>Turbiedad</i>	NTU	42.34	22.83	3620.49	60.17	3.05	804.96	801.9	4.73
<i>Precipitación Llaucay</i>	mm	3.19	1.27	21.75	4.66	0.00	39.88	39.88	2.57
<i>Precipitación Nudpu</i>	mm	2.33	0.60	14.00	3.74	0.00	26.80	26.80	2.64
<i>Sulfato de aluminio</i>	kg/día	241.27	215	10072	100.36	78.04	1069.3	991.3	2.69

Elaboración propia

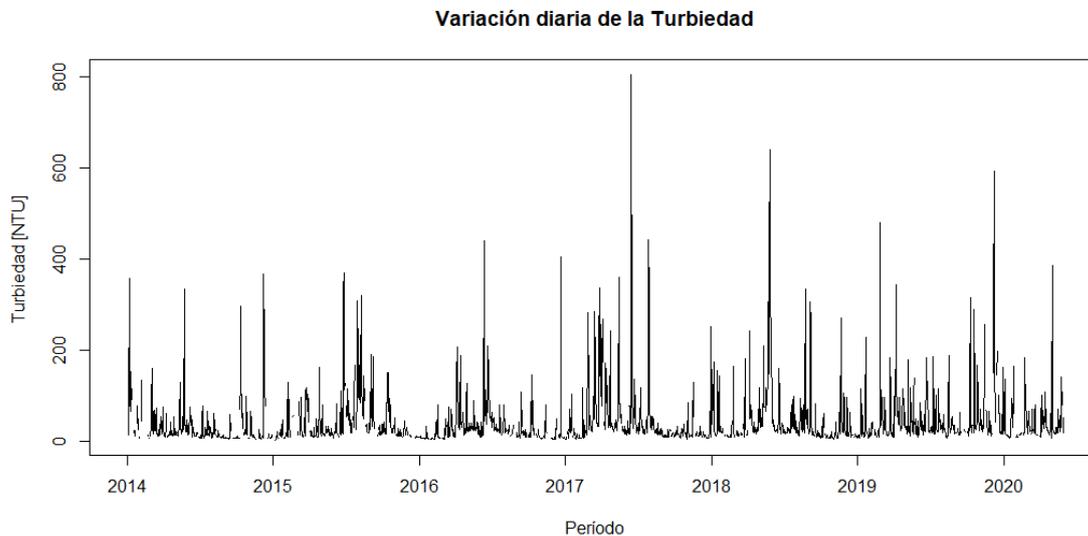


Ilustración 8 Variación diaria de la turbiedad
Elaboración propia

En primer lugar, la turbiedad presentó como valor máximo 804.96 y como mínimo de 3.05 NTU, con una media y desviación estándar de 42.34 y 60.17 NTU respectivamente, es decir en base a la desviación calculada los datos se encontraron muy dispersos con respecto a su media. La tendencia de los datos se puede observar en la *Ilustración 8*, se identificó que el valor más

alto de la turbiedad en la planta se presentó a mediados del año 2017; además de ello, en los años 2018 y 2019 existieron datos muy cercanos a los 600 NTU.

Los resultados obtenidos para el parámetro de la precipitación diaria con la estación de Llaucay, mostraron que la media es de alrededor 3.19 mm, con un valor mínimo de 0 y un máximo de 39.88 mm a inicios del año 2017, tal como se muestra en la *Ilustración 9*.

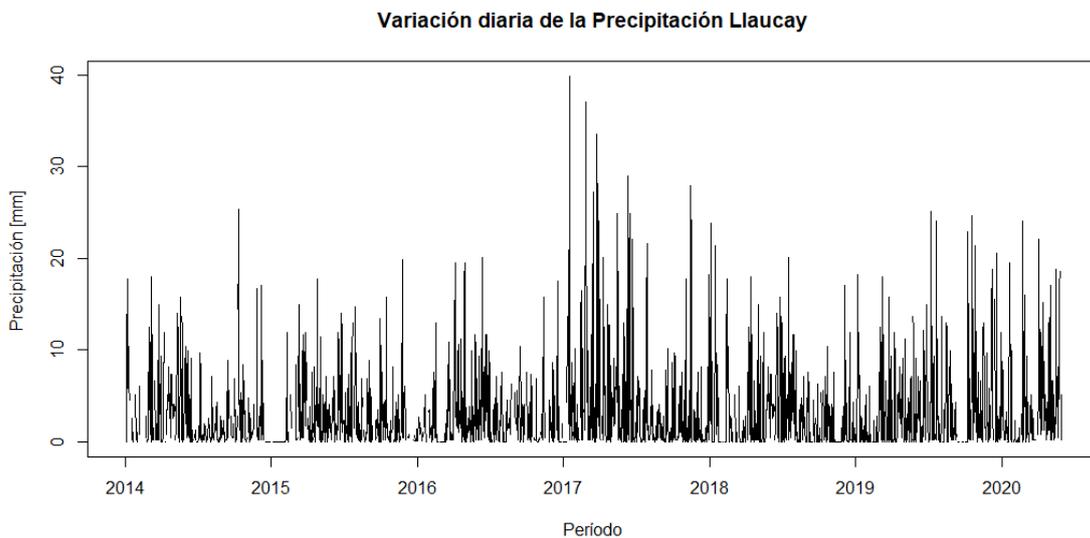


Ilustración 9 Variación diaria de la precipitación en Llaucay

Elaboración propia

Además, para la variable de precipitación diaria en la estación Nudpud se observó una media de 2.33 mm, un valor mínimo y máximo de 0 y 26.80 mm respectivamente. El dato máximo de precipitación diaria se presentó a finales del año 2014, aunque a finales de 2015 existió de igual forma un valor alto de precipitación como se muestra en la *Ilustración 10*.

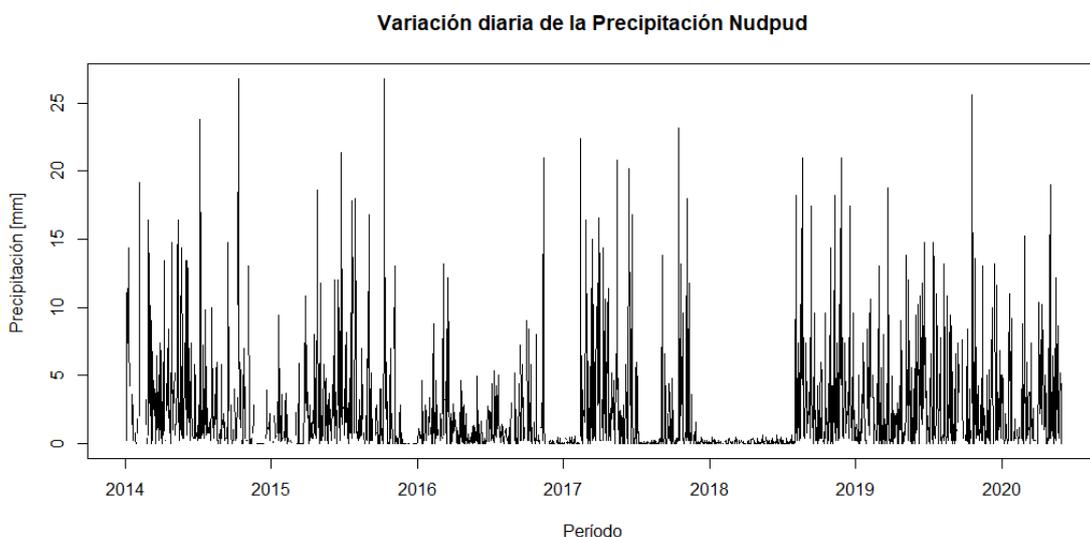


Ilustración 10 Variación diaria de la precipitación en Nudpud

Elaboración propia

Por último, en cuanto a la dosis de coagulante utilizado en el proceso de potabilización del agua cruda, se obtuvo que se consume en promedio 241.27 Kg de sulfato de aluminio por día, con un consumo máximo de 1069.3 Kg a mediados de 2015. Se puede observar en la *Ilustración 11* que el uso de sulfato de aluminio sigue una tendencia casi similar a lo largo de los años de análisis.

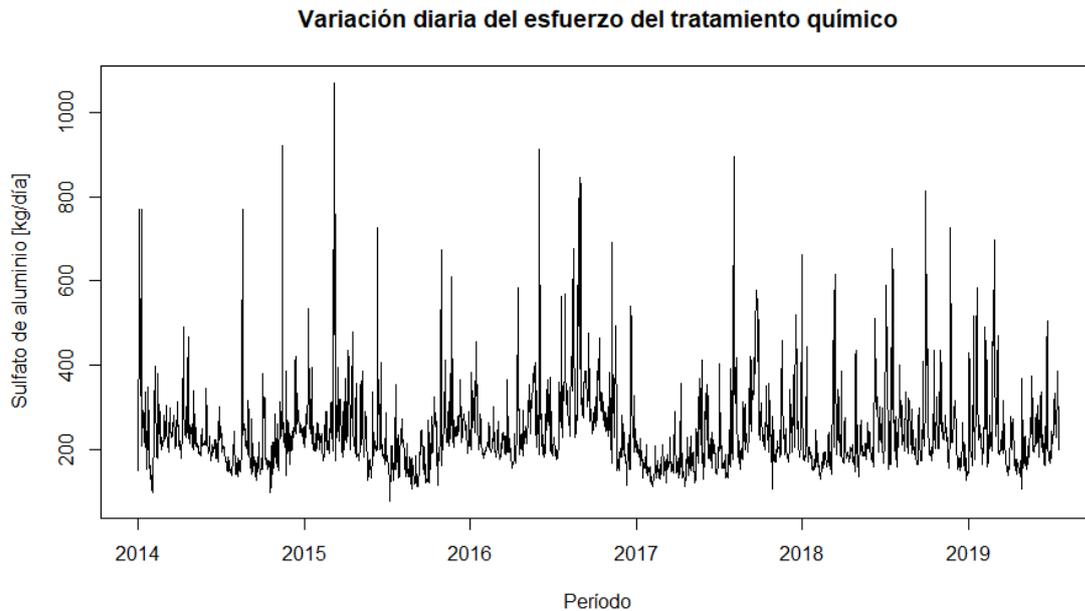


Ilustración 11 Variación diaria de sulfato de aluminio

Elaboración propia

4.1.2. Representación gráfica de los datos

Como se mencionó en el Capítulo II “*Marco Teórico*”, la representación gráfica de los datos ayudó a identificar ciertos patrones, tendencias que presentan los datos y pueden ser inadvertidos al momento de realizar el análisis estadístico básico.

Para ello, con ayuda del software libre R Studio se graficó la densidad de todas las variables utilizadas, tal como se observa en la *Ilustración 12*, el mayor número de datos para la variable turbidez se encuentra en el rango de 0 a 200 NTU, con 1975 datos en total.

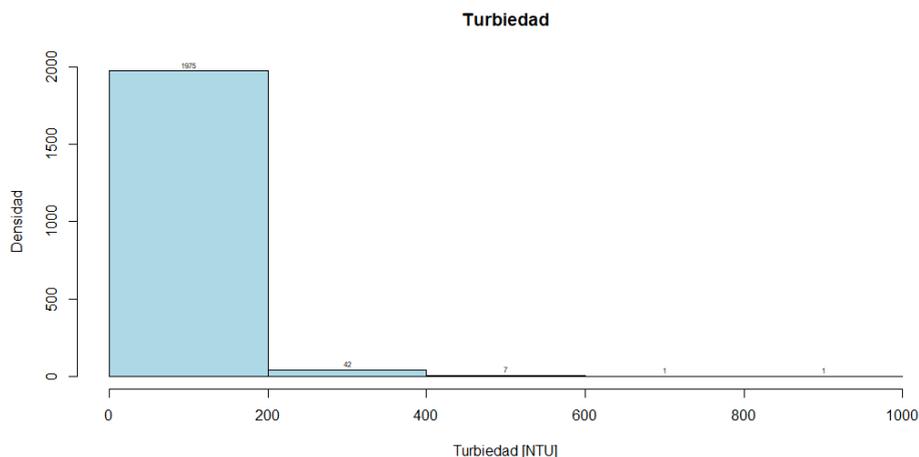


Ilustración 12 Densidad de la turbiedad
Elaboración propia

Con base en la *Ilustración 13*, el rango de precipitación de mayor densidad para la estación meteorológica de Llaucay es de 0 a 5 mm/ día con 1584 datos, seguido por los datos del rango de 5 a 10 mm/día con un total de 275 días, con 1 dato dentro del rango de 30 a 35 mm/día y con 2 datos en el mayor rango considerado para la estación de 35 a 40 mm/día.

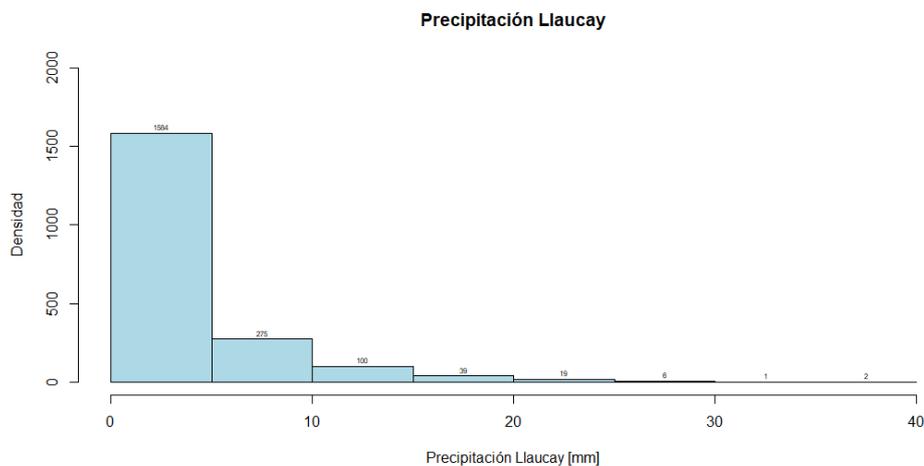
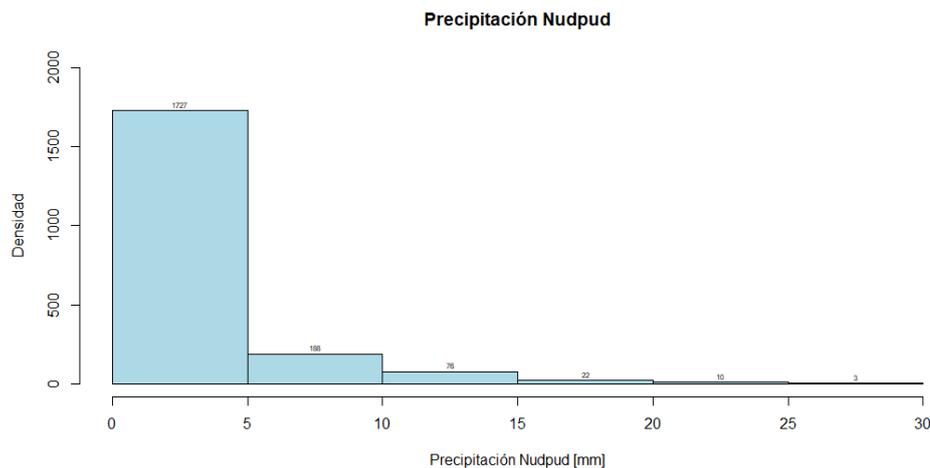


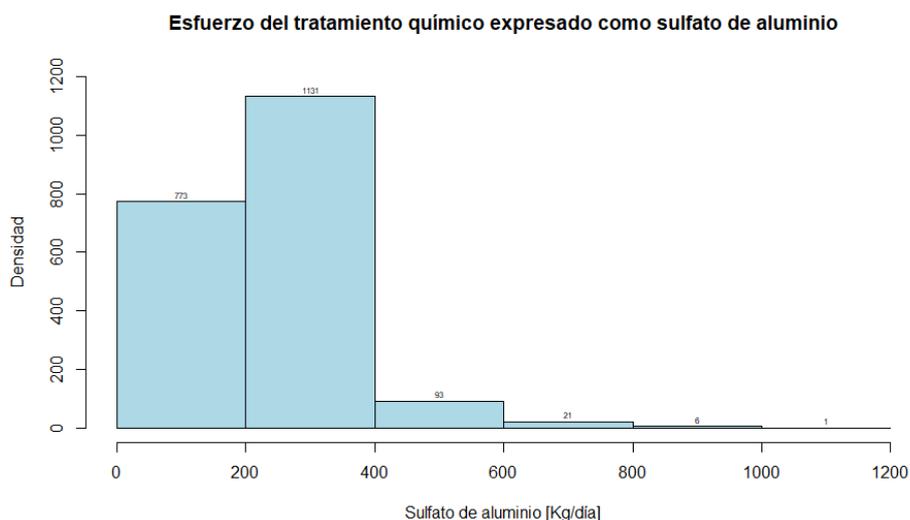
Ilustración 13 Densidad de la precipitación de la estación de Llaucay
Elaboración propia

En la *Ilustración 14* se presenta la densidad de los datos de la estación meteorológica de Nudpud de igual forma que la estación anterior, el mayor número de datos se encuentra dentro del rango de 0 a 5 mm/día con 1727 datos, el rango mayor para la estación está dentro del rango de 25 a 30 mm/día con 3 días.



*Ilustración 14 Densidad de la precipitación de la estación de Nudpud
Elaboración propia*

La *Ilustración 15* muestra la densidad de la cantidad de coagulante utilizado en el proceso de coagulación en la PTAP-M, en este caso expresado como Kg/día de sulfato de aluminio, el rango con mayor reiteración es de 200 a 400 Kg/día. El rango que presenta el mayor consumo de productos químicos es entre los 1000 a 1200 Kg/día, ha ocurrido únicamente una vez, episodio de alta turbiedad del agua cruda.



*Ilustración 15 Densidad del esfuerzo del tratamiento químico
Elaboración propia*

4.1.3. Comprobación de la normalidad de los datos

En cuanto a la normalidad de los datos analizados mediante el software R Studio, con las pruebas de Kolmogorov-Smirnov, Shapiro Wilk, Liliefors, se conoció que todas las variables analizadas: turbiedad, precipitación (estación Llaucay y Nudpud), turbiedad y coagulante para

el proceso de potabilización de la PTAP -M; los valores de p_value fueron inferiores a $2.2e^{-16}$, lo que nos indica que los datos no siguen una distribución normal, ya que, el valor es inferior a 0.05, como se observa en la *Tabla 5*.

Tabla 5 Resumen de normalidad de las variables

Turbiedad	
Kolmogorov- Smirnov	p-value < $2.2e^{-16}$
Shapiro Wilk	p-value < $2.2e^{-16}$
Lilliefors	p-value < $2.2e^{-16}$
Precipitación Llaucay	
Kolmogorov- Smirnov	p-value < $2.2e^{-16}$
Shapiro Wilk	p-value < $2.2e^{-16}$
Lilliefors	p-value < $2.2e^{-16}$
Precipitación Nudpud	
Kolmogorov- Smirnov	p-value < $2.2e^{-16}$
Shapiro Wilk	p-value < $2.2e^{-16}$
Lilliefors	p-value < $2.2e^{-16}$
Sulfato de aluminio	
Kolmogorov- Smirnov	p-value = $2.669e^{-12}$
Shapiro Wilk	p-value < $2.2e^{-16}$
Lilliefors	p-value < $2.2e^{-16}$

Elaboración propia

Además, mediante un análisis gráfico se pudo observar en las *Ilustraciones 12, 13, 14 y 15* que las densidades de las variables no siguen una distribución gaussiana (forma de campana), sino que en la mayoría de los casos tienen un sesgo de la curva hacia la izquierda.

De igual forma, se realizó los boxplots y los q-qplots de cada una de las variables. Como se puede observar en la *Ilustración 16 y 17*, el boxplot de la turbiedad nos indica que los datos no siguen una distribución normal. Al lado derecho de la *Ilustración 16* se encuentra un boxplot normal con media y distribución de los datos y, al lado izquierdo se encuentran los datos sin ningún tipo de transformación, si los datos fueran normales las gráficas serían similares. Dentro del boxplot la línea punteada (media) y la línea sólida (mediana) no se empalman, es decir, no siguen una distribución normal. Además, la sección de los bigotes en una distribución normal es del mismo tamaño tanto en la parte superior como inferior; lo que no aplica al presente estudio. Por último, la presencia de datos atípicos es poco común en serie de datos con distribución normal.

El q-qplot es un gráfico de probabilidad para comparar distribuciones de probabilidad al trazar sus cuantiles uno contra el otro, lo ideal es que los puntos se acerquen o sigan la recta diagonal, pero como se aprecia en la *Ilustración 18* los datos analizados de la turbiedad solo siguen el

sentido de la diagonal en el rango de -1 a 1; mientras que entre el rango mayor a 1 en adelante no se entrelazan en ningún punto, considerando que la variable no presenta una distribución normal.

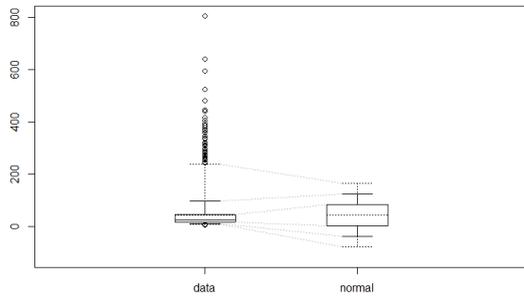


Ilustración 16 Boxplot data turbiedad vs boxplot normal

Elaboración propia

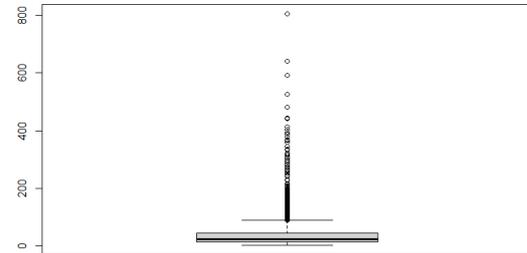


Ilustración 17 Boxplot datos de turbiedad

Elaboración propia

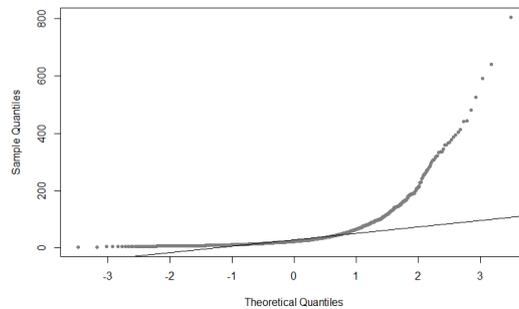


Ilustración 18 Q-Qplot turbiedad

Elaboración propia

De la misma manera, con lo ya mencionado anteriormente sobre el boxplot, en este caso para la precipitación en la estación de Llaucay, se puede identificar en la *Ilustración 19* y *20* que los datos no siguen una distribución normal, ya que, la media y la mediana de los datos no se empalman, la dimensión de los bigotes superior e inferior son distintas y, la existencia considerable de datos atípicos.

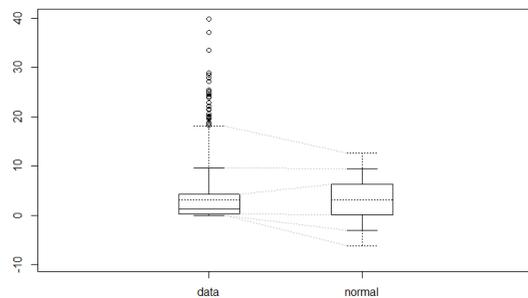


Ilustración 19 Boxplot data precipitación Llaucay vs

boxplot normal

Elaboración propia

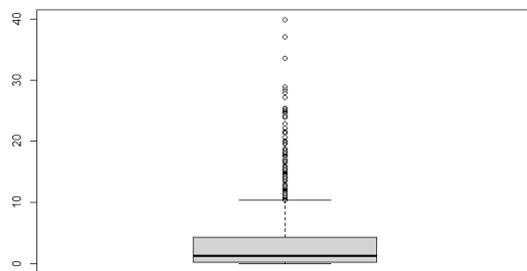
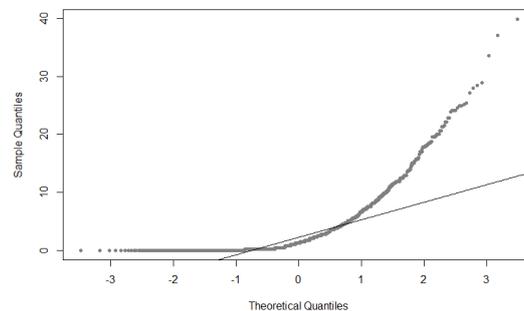


Ilustración 20 Boxplot datos de precipitación estación

Llaucay

Elaboración propia

En la *Ilustración 21* se presenta el q-qplot de la precipitación en Llaucay y se observa que los datos no se asemejan o no siguen la misma distribución que la línea diagonal, es decir, los datos de esta variable no siguen una distribución normal.

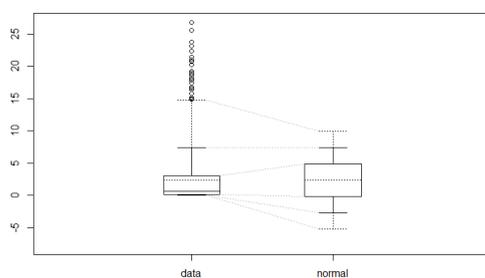


*Ilustración 21 Q-Qplot precipitación estación Llaucay
Elaboración propia*

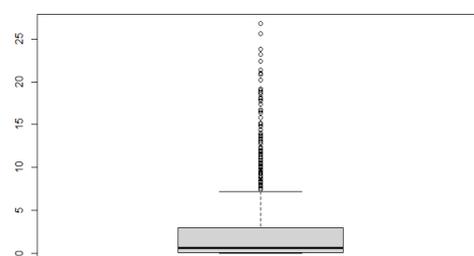
De igual forma, se presenta el mismo análisis para el resto de variables analizadas: precipitación de la estación meteorológica de Nudpud y coagulante utilizado en el proceso de coagulación expresado como sulfato de aluminio; en estos casos el resultado fue similar al presentado anteriormente. El boxplot de cada uno de ellos demuestra que los datos no siguen una distribución normal, considerando que los valores de media y mediana no son los mismos, que los bigotes tanto superior como inferior son de dimensiones distintas y existe la presencia significativa de datos atípicos.

En cuanto, al análisis de q-qplot de las variables mencionadas anteriormente, el análisis demostró que los datos de ninguno de ellos siguen una distribución normal ya que, en ningún caso se asemeja la línea diagonal con la dispersión de los datos.

En las *Ilustraciones 22, 23 y 24* se observa el análisis de la precipitación de la estación de Nudpud, las *Ilustraciones 25, 26 y 27* de sulfato de aluminio.



*Ilustración 22 Boxplot data precipitación Nudpud vs
boxplot normal
Elaboración propia*



*Ilustración 23 Boxplot datos de precipitación estación
Nudpud
Elaboración propia*

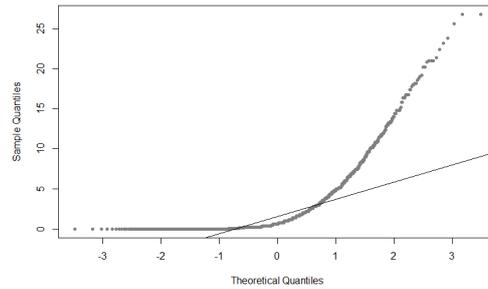


Ilustración 24 Q-Qplot precipitación estación Nudpud

Elaboración propia

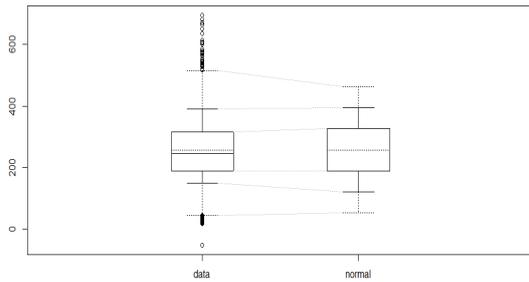


Ilustración 25 Boxplot data sulfato de aluminio vs boxplot normal

Elaboración propia

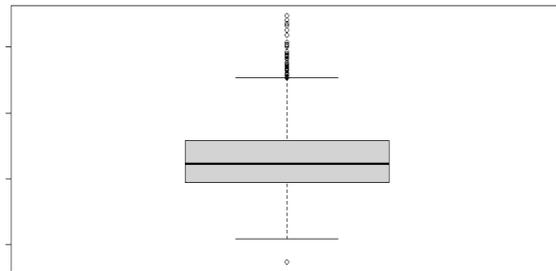


Ilustración 26 Boxplot datos sulfato de aluminio

Elaboración propia

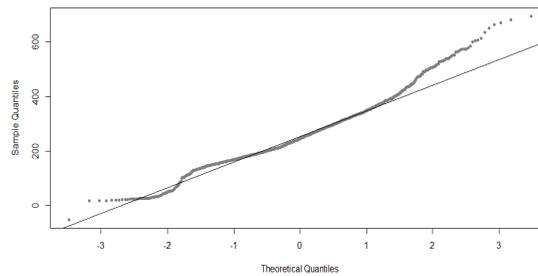


Ilustración 27 Q-Qplot sulfato de aluminio

Elaboración propia

4.2. Relación matemática de la precipitación pluvial y la turbidez del agua cruda de la PTAP-M

4.2.1. Análisis de correlación y regresión entre turbidez y precipitación

Los resultados mostraron que para el caso de la turbiedad y la precipitación de Llaucay se obtuvo un r de 0.44. En tanto que, al realizar el análisis entre la turbiedad y la precipitación de Nudpud el valor de r fue de 0.31 y entre ambas precipitaciones se obtuvo un coeficiente de correlación de 0.35, como se puede observar en la *Tabla 6*.

Tabla 6 Coeficientes de correlación de los parámetros analizados

Parámetro	Turbiedad	Precipitación Llaucay	Precipitación Nudpud
Turbiedad	1	-	-
Precipitación Llaucay	0.44	1	-
Precipitación Nudpud	0.31	0.35	1

Elaboración propia

Con el fin de generar una ecuación para la turbidez y precipitación, se utilizó la regresión lineal, cabe mencionar que los conjuntos para la regresión fueron: Turbiedad – Precipitación de Llaucay, Turbiedad – Precipitación de Nudpud y para el caso de Turbiedad – Precipitación Llaucay – Precipitación Nudpud se procedió a sumar las precipitaciones de las dos estaciones y con ese único valor se realizó la regresión.

Al realizar la regresión lineal simple con los datos de calibración los análisis mostraron que la mejor relación para el caso de Turbiedad – Precipitación Llaucay fue el modelo lineal con un R^2 de 0.20 (ver *Ilustración 28*) con un RMSE de 0.69 y un EF de 0.58, seguido del modelo cuadrático con el mismo R^2 , con un RMSE de 166.33 y un EF negativo. En cuanto, al segundo conjunto de Turbiedad – Precipitación Nudpud el mejor modelo fue el modelo lineal con un R^2 de 0.10 (*Ilustración 29*), seguido del modelo lineal sin intercepto, cuadrático, cuadrático sin intercepto y exponencial todos con el mismo R^2 , de este conjunto el mejor modelo presentó un RMSE de 0.23 y un EF de 0.54.

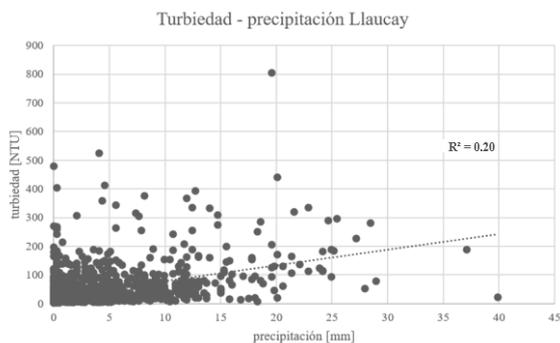


Ilustración 28 Modelo lineal sin intercepto turbiedad – Llaucay
Elaboración propia

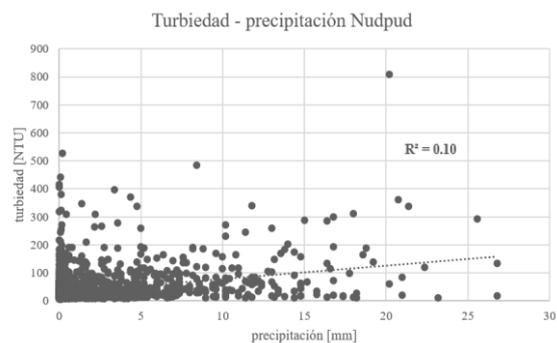


Ilustración 29 Modelo lineal turbiedad – Nudpud
Elaboración propia

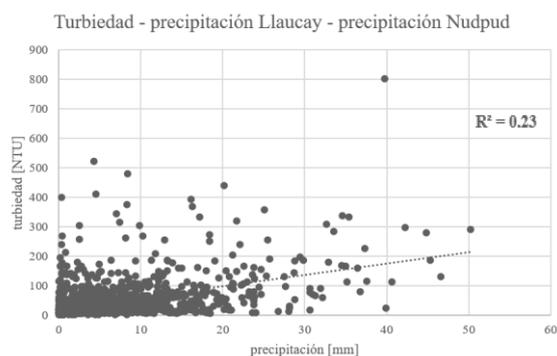


Ilustración 30 Modelo cuadrático turbiedad - Llaucay – Nudpud
Elaboración propia

Por último, el conjunto Turbiedad – Precipitación Llaucay – Precipitación Nudpud, obtuvo la correlación más significativa del estudio, de igual forma el mejor modelo fue el cuadrático y presentó un R^2 de 0.23 (ver *Ilustración 30*), seguido del modelo lineal, lineal sin intercepto y cuadrático sin

intercepto con 0.22. Este conjunto de regresión presentó los mejores resultados porque asume la precipitación sumada de las dos estaciones meteorológicas y el valor de la turbidez es medido en la entrada de la PTAP-M donde ya se mezcló el agua procedente de ambas captaciones Llaucay y Nudpud.

Tabla 7 Resultados de los errores de las etapas de calibración de los modelos de regresión

		Calibración			
Conjunto	Modelo	Ecuación	R²	RMSE	EF
Turbiedad – Llaucay	Lineal	$y = 5.4485x + 25.0397$	0.20	0.69	0.58
	Lineal sin intercepto	$y = 7.8992x$	0.19	0.89	0.51
	Cuadrático	$y = -7.453x^2 + 1034.912x + 42.178$	0.20	166.33	-7731.12
	Cuadrático sin intercepto	$y = -7.453x^2 + 1034.912x$	0.19	165.29	-7665.38
	Exponencial	$y = 20.006e^{0.0872x}$	0.12	1.34	0.13
	Cúbico	$y = -286.7x^3 - 7.453x^2 + 1034.912x + 42.178$	0.08	2129.79	-7843628
	Cúbico sin intercepto	$y = -286.7x^3 - 7.453x^2 + 1034.912x$	0.08	2130.83	-7845736
Turbiedad – Nudpud	Lineal	$y = 4.7326x + 31.0730$	0.10	0.23	0.54
	Lineal sin intercepto	$y = 8.3543x$	0.10	0.75	0.60
	Cuadrático	$y = 129.474x^2 + 728.918x + 42.178$	0.10	926.06	-32228.83
	Cuadrático sin intercepto	$y = 129.5x^2 + 728.9x$	0.10	925.01	-32093.88
	Exponencial	$y = 22.202e^{0.0725x}$	0.10	0.67	0.06
	Cúbico	$y = -90.619x^3 + 129.474x^2 + 728.918x + 42.178$	0.07	5792.58	-395474.23
	Cúbico sin intercepto	$y = -90.619x^3 + 129.474x^2 + 728.918x$	0.07	5793.63	-395947.61
Turbiedad – Llaucay – Nudpud	Lineal	$y = 3.8520x + 21.0230$	0.22	0.32	0.76
	Lineal sin intercepto	$y = 5.3026x$	0.22	0.56	0.83
	Cuadrático	$y = 218.57x^2 + 1090.720x + 42.178$	0.23	3027.03	-1020467.15
	Cuadrático sin intercepto	$y = 218.57x^2 + 1090.720x$	0.22	3025.99	-1019707.11
	Exponencial	$y = 18.859e^{0.0607x}$	0.21	0.54	0.37
	Cúbico	$y = 21.224x^3 + 218.57x^2 + 1090.72x + 42.178$	0.19	8089.58	-6869046
	Cúbico sin intercepto	$y = 21.224x^3 + 218.57x^2 + 1090.72x$	0.19	8088.53	-6867074.21

Se puede evidenciar que las diferencias no son muy significativas entre los distintos modelos para cada uno de los conjuntos analizados teniendo como el mejor coeficiente el ya mencionado de 0.23, y el peor de los coeficientes dentro del análisis de 0.07 en el conjunto Turbiedad – Precipitación Nudpud, para el caso del modelo cúbico y cúbico sin intercepto. Pero se

selecciona el mejor modelo al lineal para la relación Turbiedad – Nudpud – Llaucay por presentar los mejores estadísticos en conjunto: 0.22, 0.32 y 0.76 para el R^2 , RMSE y EF respectivamente. Finalmente, la *Tabla 7* muestra los resultados obtenidos con cada uno de los modelos de los diferentes conjuntos.

En la *Tabla 8* se puede observar los resultados obtenidos para los datos en la validación. Se busca el valor más alto de R^2 igual o más cercano a la unidad, un valor mínimo de RMSE y por último un valor de EF más cercano a 1. Para el primer conjunto de variables, Turbiedad – Precipitación Llaucay, se atribuye el mejor modelo según los estadísticos al modelo lineal porque presenta un R^2 de 0.20, un RMSE de 0.15 y un EF (Índice de Nash) de -6.89. En el caso del segundo conjunto de variables, Turbiedad – Precipitación Nudpud, se atribuye el mejor modelo en base a los estadísticos al modelo lineal con un R^2 de 0.07, un RMSE de 1.35 y un EF de 0.72. Finalmente, en el tercer caso donde interactúan el conjunto de variables, Turbiedad – Precipitación Llaucay – Precipitación Nudpud, el mejor modelo es el exponencial con valores de R^2 de 0.21, un RMSE de 1.65 y un EF de 0.96. Interpretando estos últimos valores se comprende: el 0.21 del R^2 significa que el modelo propuesto no brinda una predicción adecuada porque no está cerca del valor de la unidad; el 1.65 del RMSE demuestra que el modelo puede ser preciso con la respuesta que se espera recibir y el valor de 0.96 de EF está indicando que el modelo tiene una buena simulación de la realidad.

Tabla 8 Resultados de los errores de las etapas de validación de los modelos de regresión

Conjunto	Modelo	Ecuación	Validación		
			R^2	RMSE	EF
Turbiedad – Llaucay	Lineal	$y = 5.4485x + 25.0397$	0.20	0.15	-6.89
	Lineal sin intercepto	$y = 7.8992x$	0.20	0.43	-12.82
	Cuadrático	$y = -7.453x^2 + 1034.912x + 42.178$	0.19	581.02	-5941.81
	Cuadrático sin intercepto	$y = -7.453x^2 + 1034.912x$	0.19	578.93	-5883.52
	Exponencial	$y = 20.006e^{0.0872x}$	0.15	1.52	-1.35
	Cúbico	$y = -286.7x^3 - 7.453x^2 + 1034.912x + 42.178$	0.10	2698.15	-62358.89
	Cúbico sin intercepto	$y = -286.7x^3 - 7.453x^2 + 1034.912x$	0.10	2692.68	-62364.58

Continuación

Tabla 8 Resultados de los errores de las etapas de validación de los modelos de regresión

	Lineal	$y = 4.7326x + 31.0730$	0.07	0.38	0.79
Turbiedad – Nudpud	Lineal sin intercepto	$y = 8.3543x$	0.07	1.35	0.72
	Cuadrático	$y = 129.474x^2 + 728.918x + 42.178$	0.06	1767.89	-651826.71
	Cuadrático sin intercepto	$y = 129.5x^2 + 728.9x$	0.06	1765.79	-648601.16
	Exponencial	$y = 22.202e^{0.0725x}$	0.06	1.42	-0.03
	Cúbico	$y = -90.619x^3 + 129.474x^2 + 728.918x + 42.178$	0.02	10603.33	-5406788.9
	Cúbico sin intercepto	$y = -90.619x^3 + 129.474x^2 + 728.918x$	0.02	10605.43	-5416094.3
		Lineal	$y = 3.8520x + 21.0230$	0.18	3.04
Turbiedad – Nudpud – Llaucay	Lineal sin intercepto	$y = 5.3026x$	0.18	3.94	0.90
	Cuadrático	$y = 218.57x^2 + 1090.720x + 42.178$	0.21	9401.61	-4378.02
	Cuadrático sin intercepto	$y = 218.57x^2 + 1090.720x$	0.21	9399.51	-4171.20
	Exponencial	$y = 18.859e^{0.0607x}$	0.21	1.65	0.96
	Cúbico	$y = 21.224x^3 + 218.57x^2 + 1090.72x + 42.178$	0.21	30238.39	-4596.43
	Cúbico sin intercepto	$y = 21.224x^3 + 218.57x^2 + 1090.72x$	0.21	30236.30	-4384.45

* x=precipitación

Elaboración propia

4.2.2. Modelo entre turbidez y precipitación con Redes Neuronales Artificiales

4.2.2.1. Análisis de sensibilidad de parámetros para el modelo entre turbidez y precipitación

Los resultados obtenidos en el análisis de sensibilidad de los parámetros utilizados en la Red Neuronal Artificial para el cálculo del modelo de la turbidez con la precipitación se muestran en la *Tabla 9*, este proceso sirve para conocer qué parámetro presenta mayor sensibilidad en comparación a los otros y de esta forma obtener los mejores resultados.

Según Chang & Liao (2012), en su estudio al realizar el análisis de la sensibilidad para la predicción de la turbidez del agua cruda, establecieron que el parámetro más sensible fue el número de neuronas de las capas ocultas. Para el caso puntual los mejores resultados fueron: tiempo de entrenamiento 500; velocidad de aprendizaje 0.1; el número de neuronas en las capas ocultas 3 y la función de transferencia de la capa oculta y la capa de salida "función de

transferencia log sigmoide". Considerando este antecedente, para el presente estudio se realizó un análisis más detallado del número de neuronas como se muestra en el siguiente apartado, por ende, para esta sección se consideró el parámetro número de neuronas en las capas ocultas constante (3) y los distintos parámetros fueron modificados.

Tabla 9 Cambios en la precisión de la predicción bajo diferentes parámetros RNA para la relación de turbidez y precipitación

a. Escenarios con los tiempos de entrenamiento					
<i>Escenario</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Número de neuronas	3	3	3	3	3
Tiempo de entrenamiento	500	5000	10000	15000	20000
Velocidad de aprendizaje	0.1	0.1	0.1	0.1	0.1
Función de transferencia	S	S	S	S	S
R ² calibración	0.377	0.231	0.225	0.235	0.223
R ² validación	0.131	0.134	0.135	0.134	0.141
b. Escenarios con las velocidades de aprendizaje					
<i>Escenario</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	
Número de neuronas	3	3	3	3	
Tiempo de entrenamiento	20000	20000	20000	20000	
Velocidad de aprendizaje	0.001	0.01	0.1	0.5	
Función de transferencia	S	S	S	S	
R ² calibración	0.243	0.243	0.223	0.241	
R ² validación	0.125	0.127	0.141	0.130	
c. Escenario con distinta función de transferencia					
<i>Escenario</i>	<i>1</i>				
Número de neuronas	3				
Tiempo de entrenamiento	20000				
Velocidad de aprendizaje	0.1				
Función de transferencia	T				
R ² calibración	0.23				
R ² validación	0.134				

Elaboración propia

En la *Tabla 9* se puede observar los distintos resultados de la simulación del modelo al modificar los parámetros en la RNA. El resultado muestra que, en el caso de los escenarios con

varios tiempos de entrenamiento, el mejor R^2 se obtuvo con el escenario 5, es decir, con un tiempo de entrenamiento de 20000, se puede observar que conforme aumenta este parámetro el error disminuye igual a lo reportado por Chang & Liao (2012), a diferencia del escenario 4, en donde existe un aumento del error.

Para los escenarios con varias velocidades de aprendizaje, el que presentó mejores resultados fue el escenario 3 (Tabla 9), con una velocidad de 0.1; con un R^2 de 0.141 para la etapa de validación del modelo; de igual forma al tiempo de entrenamiento, conforme aumenta la velocidad de aprendizaje el coeficiente R^2 aumenta. Por último, para el escenario con distinta función de transferencia, en donde se utilizó la función tangente hiperbólica, se observó un R^2 de 0.134.

Es por ello, que en base al análisis realizado se optó que la mejor combinación de parámetros para el modelo entre turbidez del agua cruda y precipitación pluvial fue: tiempo de entrenamiento 20000; velocidad de aprendizaje 0.1; en tanto, la función de transferencia de la capa oculta y la capa de salida fue "función de transferencia log sigmoide" (Tabla 9).

En cuanto al cálculo del índice de sensibilidad "S", este se indica en la Tabla 10, observándose que el parámetro que tiene mayor sensibilidad al comparar los dos parámetros es el tiempo de entrenamiento con un índice de 0.000715, es decir, una mayor variación de este parámetro tendrá un efecto más significativo en el modelo realizado. Es por ello, que se puede decir que la velocidad de aprendizaje presentó una menor sensibilidad.

Tabla 10 Índice de Sensibilidad (tiempo de entrenamiento y velocidad de aprendizaje) para la relación entre turbidez y precipitación

Análisis de Sensibilidad utilizando el Índice de Sensibilidad "S"	
Parámetro	Índice "S"
Tiempo de entrenamiento	0.000715
Velocidad de aprendizaje	0.00032

Elaboración propia

4.2.2.2. Ensayo de Redes Neuronales Artificiales para el modelo entre turbidez y precipitación

Una vez conocido los valores óptimos de los parámetros ensayados en el análisis de sensibilidad, se procedió a utilizar los parámetros tiempo de entrenamiento y velocidad de aprendizaje como constantes y se modificó el número de neuronas en la capa interna. Para ello, se entrenó seis modelos con distintas neuronas en la capa interna, el primer modelo que se trabajó fue con 30 neuronas, 20 neuronas en la primera capa y 10 neuronas en la segunda capa

interna. El segundo modelo contó con 6 neuronas en la primera capa, 4 neuronas en la segunda capa y 2 neuronas en la última capa. El tercer modelo también presenta 30 neuronas, 10 neuronas en la primera capa, 15 y 5 en la segunda y tercera capa respectivamente. El cuarto modelo contó con una única capa de 30 neuronas. El quinto modelo con apenas 2 neuronas en la primera capa. En tanto, el sexto modelo fue calculado con 18 neuronas, con 9, 6 y 3 neuronas en las distintas capas. En la capa de entrada los modelos utilizan el valor de la precipitación y la capa de salida es el valor de la turbiedad.

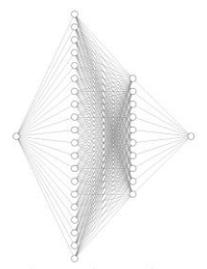
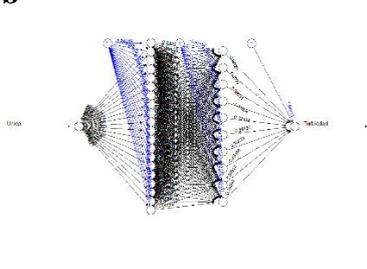
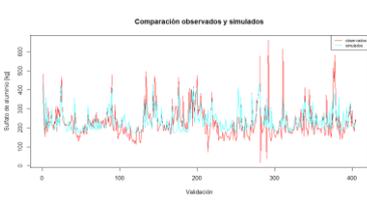
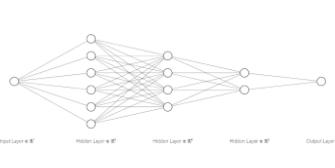
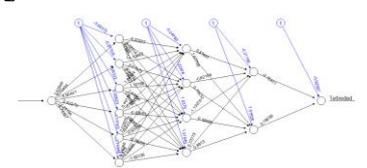
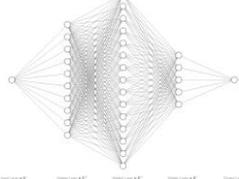
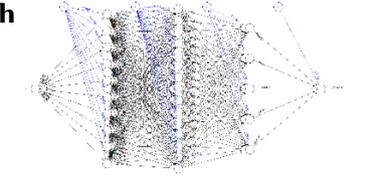
Los seis modelos analizados se presentan en la *Tabla 11*, se consideró la estructura del modelo (secciones a, d, g, j, m, p), además, el sistema global de la red neuronal (secciones b, e, h, k, n, q) obtenido en el software R Studio con los coeficientes de cada una de las redes y neuronas para el cálculo de la variable respuesta. En la tercera columna se presenta la comparación de los datos de la turbidez observados y calculados con el modelo generado con los datos de calibración (secciones c, f, i, l, o, r). Se puede percibir que, para todos los casos la línea de color cian que representa a los datos predichos mantiene una distribución similar y que el modelo generado no es capaz de predecir los valores máximos y mínimos de turbiedad medidos en PTAP-M. Para ratificar estas diferencias se calculó los errores de los datos observados y predichos para cada uno de los modelos, exhibidos en la *Tabla 12*. En el proceso de calibración la mayoría de los modelos presentan un valor de R^2 de 0.22, a excepción del Modelo 6-4-2 que muestra un coeficiente de 0.23. Es decir, la variable independiente es capaz de predecir únicamente el 22 o 23 % el comportamiento de la variable independiente, considerándose un valor bajo y el modelo no sería el más adecuado para futuras predicciones de turbidez. En cuanto a la validación del modelo, existió una disminución en la predicción de los datos, empeorando el desempeño del modelo obteniendo un R^2 de 0.17 para tres modelos, 0.18 para el Modelo 6-4-2, el peor Modelo fue el 9-6-3 que obtuvo un coeficiente de correlación de 0.0017.

Al considerar los errores calculados, se decidió que los modelos que presentan los mejores resultados son el Modelo 2 y el Modelo 10-15-5. El primer modelo escogido (Modelo 2) cuenta con una única capa interna de 2 neuronas, obteniendo un R^2 de calibración de 0.22, para el caso de la validación el coeficiente disminuyó a 0.17; es decir, el modelo fue capaz de explicar únicamente el 17 % de los datos, considerándolo como un valor muy bajo. Adicionalmente, para un análisis más completó de la validación, se consideró el RMSE y el índice de Nash, se obtuvo un valor de 0.84 y 0.96 respectivamente. Es importante mencionar que, para el RMSE se busca el valor más cercano a 0, considerando el valor de este modelo el más bajo de todo el

análisis, en cuanto al EF lo que se espera es un valor entre 0.5 y 1, para considerar que el modelo tiene una buena simulación, lo que muestra que el modelo tiene un rendimiento hasta cierto punto aceptable.

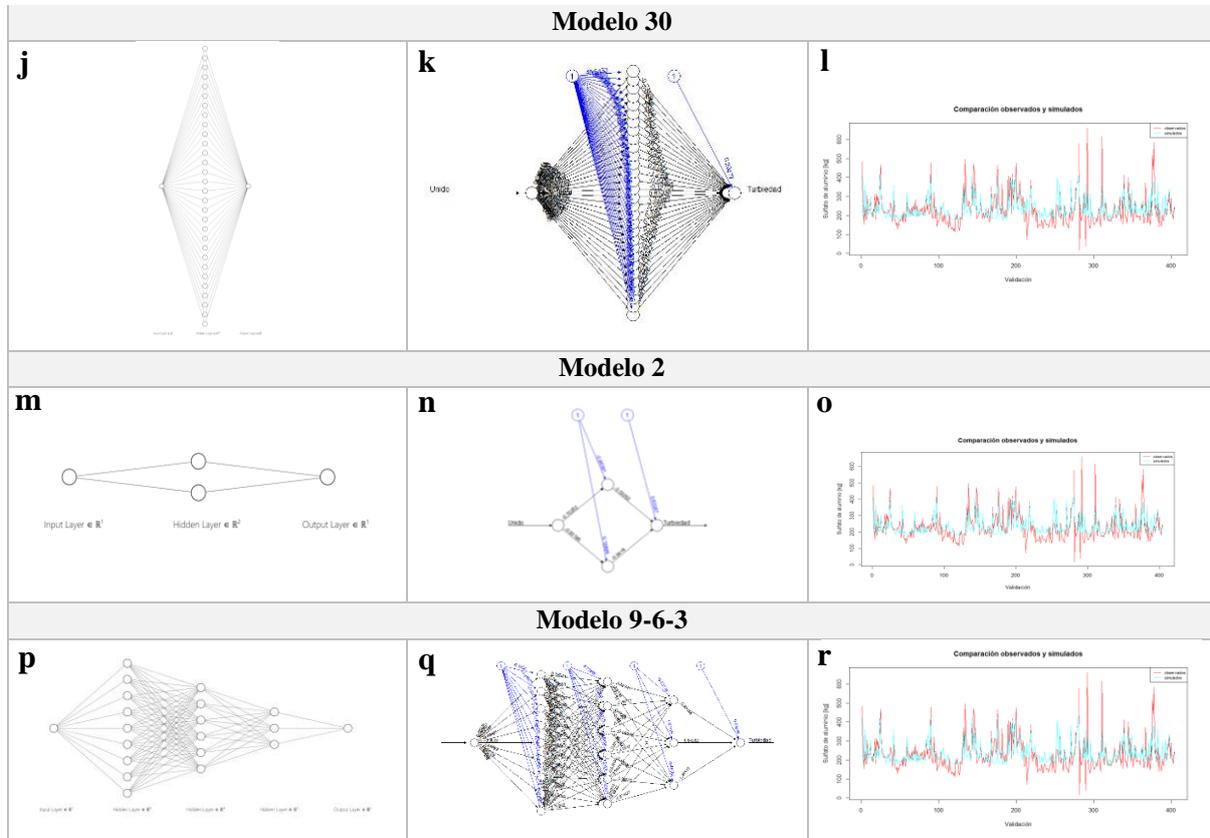
En cuanto al segundo modelo escogido, Modelo 10-15-5, este cuenta con una estructura muy diferente a la mencionada anteriormente, aquí se encuentra 10 neuronas en la primera capa oculta, 15 neuronas en la segunda capa y en la última capa existen 5 neuronas; en cuanto al desempeño observado este es muy similar al Modelo 2. Este modelo presentó un R^2 para la calibración de 0.22 y para la validación disminuyó este valor a 0.19, considerado el valor más alto en este parámetro de todos los modelos. El valor de RMSE fue de 0.90 y el EF de 0.89, de igual forma se muestra que el modelo tiene un rendimiento aceptable hasta cierto punto en base a los dos últimos parámetros analizados.

Tabla 11 Resumen de los modelos de RNA entrenados para la relación turbidez y precipitación

Estructura base de la RNA	Sistema global final de la RNA	Calidad del agua observada y predicha
Modelo 20-10		
a 	b 	c 
Modelo 6-4-2		
d 	e 	f 
Modelo 10-15-5		
g 	h 	i 

Continuación

Tabla 11 Resumen de los modelos de RNA entrenados para la relación turbidez y precipitación



Elaboración propia

Tabla 12 Resultados de los errores de las etapas de calibración y validación de los modelos de RNA para la relación turbidez y precipitación

Relación entre variable turbidez y variable de precipitación (Llaucay + Nudpud)

Modelo	Calibración			Validación		
	R ²	RMSE	EF	R ²	RMSE	EF
Modelo 20-10	0.22	2.03	0.98	0.19	0.92	0.90
Modelo 6-4-2	0.23	1.88	0.97	0.18	0.90	0.88
Modelo 10-15-5	0.22	2.06	1.00	0.19	0.90	0.89
Modelo 30	0.22	2.05	0.99	0.19	0.90	0.88
Modelo 2	0.22	2.02	0.99	0.17	0.84	0.96
Modelo 9-6-3	0.22	2.06	1.00	0.0017	1.46	-15.41

Elaboración propia

De esta forma, se demuestra que los dos modelos considerados como los mejores, presentan desempeños muy similares entre ellos, pero, es desalentador el uso de los mismos para futuras aplicaciones en la PTAP-M con el fin de predecir la turbiedad del agua cruda que ingresa a la planta, debido a sus valores de predicción tan bajos.

La ecuación para el Modelo 2 se presenta a continuación, la Ecuación 10 de salida:

$$H1 = \sigma((-0.7035*x) - 0.8938))$$

$$H2 = \sigma((-0.8319*x) + 0.1889))$$

$$\text{Salida} = \sigma((-0.3325 * H1) + (-0.9416 * H2) - 0.6339))$$

Ecuación 10 RNA Modelo 2

En donde,

$H1$ = ecuación de la primera neurona de la RNA Modelo 2

$H2$ = ecuación de la segunda neurona de la RNA Modelo 2

x = precipitación pluvial [mm/día]

$Salida$ = turbidez [NTU]

σ = función log sigmoide $\left[\frac{1}{1+e^{-z}} \right]$

De igual forma, se obtuvo el sistema de ecuaciones para el Modelo 10-15-5 como se muestra:

Primera capa interna (10 neuronas)

$$H1 = \sigma((-1.7978*x) + 1.5929))$$

$$H2 = \sigma((-1.4863*x) - 1.0817))$$

$$H3 = \sigma((-0.7591*x) - 0.339))$$

$$H4 = \sigma((1.5522*x) + 0.0927))$$

$$H5 = \sigma((0.1228*x) - 1.5369))$$

$$H6 = \sigma((0.758*x) + 0.4689))$$

$$H7 = \sigma((-0.1889*x) + 0.3023))$$

$$H8 = \sigma((-2.1299*x) + 0.9522))$$

$$H9 = \sigma((-1.0499*x) - 1.4009))$$

$$H10 = \sigma((-0.1973x) - 0.1757))$$

Segunda capa interna (15 neuronas)

$$H11 = \sigma((-1.055*H1) + (-0.6623*H2) + (-1.8082*H3) + (-0.5748*H4) + (-1.7048*H5) + (-2.2746*H6) + (-1.301*H7) + (-1.3001*H8) + (1.1352*H9) + (-0.8827*H10) - 2.0684))$$

$$H12 = \sigma ((-0.3951*H1) + (-0.5807*H2) + (-1.0971*H3) + (-1.3979*H4) + (0.5295*H5) + (1.5808*H6) + (0.9996*H7) + (1.6683*H8) + (1.6587*H9) + (0.9582*H10) + 0.4215)$$

$$H13 = \sigma ((-1.698*H1) + (-1.8368*H2) + (-0.1872*H3) + (-1.0555*H4) + (0.5289*H5) + (-0.4357*H6) + (0.3186*H7) + (0.4192*H8) + (-0.018*H9) + (-0.1935*H10) + 0.4268)$$

$$H14 = \sigma ((-0.1348*H1) + (-1.3499*H2) + (1.2566*H3) + (-1.8778*H4) + (-1.2672*H5) + (0.5847*H6) + (0.3469*H7) + (0.8623*H8) + (-0.9386*H9) + (-1.6057*H10) - 0.8012)$$

$$H15 = \sigma ((0.4790*H1) + (2.2438*H2) + (-1.0855*H3) + (-1.9399*H4) + (0.5316*H5) + (1.5628*H6) + (-0.1799*H7) + (0.7593*H8) + (-0.4383*H9) + (1.0776*H10) + 0.9012)$$

$$H16 = \sigma ((0.4968*H1) + (-1.3047*H2) + (-0.7002*H3) + (-0.6709*H4) + (-0.3155*H5) + (-1.4036*H6) + (0.6297*H7) + (-0.9173*H8) + (0.2991*H9) + (-1.6104*H10) + 0.3654)$$

$$H17 = \sigma ((2.0805*H1) + (0.5686*H2) + (0.0603*H3) + (-1.2363*H4) + (-0.7512*H5) + (-0.9852*H6) + (0.4560*H7) + (1.9799*H8) + (1.5735*H9) + (0.1702*H10) - 0.9206)$$

$$H18 = \sigma ((1.4187*H1) + (1.2502*H2) + (-0.1889*H3) + (1.7824*H4) + (0.4072*H5) + (-0.7816*H6) + (-0.9842*H7) + (-0.7162*H8) + (0.4502*H9) + (-0.4289*H10) - 0.4068)$$

$$H19 = \sigma ((-1.1513*H1) + (1.4232*H2) + (0.0081*H3) + (1.0856*H4) + (-0.6389*H5) + (-1.0856*H6) + (-1.3174*H7) + (0.4832*H8) + (-0.2698*H9) + (-0.7434*H10) - 0.4884)$$

$$H20 = \sigma ((0.0323*H1) + (-0.8933*H2) + (-1.7688*H3) + (0.9440*H4) + (-2.1087*H5) + (-1.3580*H6) + (0.9749*H7) + (1.0532*H8) + (-0.4210*H9) + (0.3531*H10) + 0.7629)$$

$$H21 = \sigma ((0.7095*H1) + (0.4819*H2) + (1.3177*H3) + (0.7522*H4) + (0.1181*H5) + (1.5808*H6) + (0.1032*H7) + (-0.4236*H8) + (-0.2000*H9) + (0.3182*H10) - 0.4882)$$

$$H22 = \sigma ((-1.7134*H1) + (-0.5809*H2) + (-0.2642*H3) + (-1.3979*H4) + (0.5295*H5) + (-0.8175*H6) + (0.3186*H7) + (0.5663*H8) + (0.4160*H9) + (0.1568*H10) + 0.2670)$$

$$H23 = \sigma ((0.4481*H1) + (0.2963*H2) + (0.6105*H3) + (0.7408*H4) + (0.54476*H5) + (1.7767*H6) + (1.3785*H7) + (0.6669*H8) + (1.038*H9) + (-2.2205*H10) - 0.0232)$$

$$H24 = \sigma ((-1.4927*H1) + (1.0113*H2) + (-0.3531*H3) + (-1.0913*H4) + (-0.0109*H5) + (1.4288*H6) + (-0.9554*H7) + (-0.7993*H8) + (-0.8835*H9) + (-0.9591*H10) - 0.6308)$$

$$H25 = \sigma ((-0.7339*H1) + (-0.0834*H2) + (-1.0489*H3) + (0.3756*H4) + (-0.0956*H5) + (-1.2746*H6) + (-0.6360*H7) + (-0.9327*H8) + (-1.2956*H9) + (-0.8589*H10) + 0.9119)$$

Tercera capa interna (5 neuronas)

$$H26 = \sigma ((1.3637*H11) + (-0.4936*H12) + (0.1519*H13) + (0.0457*H14) + (-0.2954*H15) + (0.0159*H16) + (0.75*H17) + (1.0629*H18) + (-0.6772*H19) + (-1.5043*H20) + (0.2393*H21) + (-0.1163*H22) + (-0.8120*H23) + (0.0946*H24) + (0.8643*H25) - 2.1663)$$

$$H27 = \sigma ((-1.222*H11) + (0.3144*H12) + (-1.1395*H13) + (0.2707*H14) + (0.2002*H15) + (1.5191*H16) + (-0.1935*H17) + (0.7407*H18) + (0.8944*H19) + (-0.2019*H20) + (-1.4663*H21) + (-0.038*H22) + (-0.3043*H23) + (0.3884*H24) + (0.2002*H25) + 0.057)$$

$$H28 = \sigma ((1.0882*H11) + (-0.0269*H12) + (-0.0967*H13) + (0.4605*H14) + (0.3959*H15) + (1.7529*H16) + (1.2973*H17) + (0.9763*H18) + (0.2533*H19) + (-0.6403*H20) + (-1.0910*H21) + (-0.9308*H22) + (-0.8661*H23) + (0.5057*H24) + (-0.5789*H25) - 0.2149)$$

$$H29 = \sigma ((1.3250*H11) + (0.2703*H12) + (0.0171*H13) + (-0.7838*H14) + (-0.7004*H15) + (-0.7522*H16) + (-0.1696*H17) + (-0.6491*H18) + (0.1667*H19) + (1.099*H20) + (0.3055*H21) + (-0.5575*H22) + (-1.0709*H23) + (0.1975*H24) + (0.1220*H25) - 0.1719)$$

$$H30 = \sigma ((0.3803*H11) + (0.6560*H12) + (-0.2808*H13) + (0.3084*H14) + (-0.692*H15) + (2.2581*H16) + (-1.6966*H17) + (0.5795*H18) + (0.037*H19) + (-1.4118*H20) + (-0.3472*H21) + (-0.711*H22) + (-0.5006*H23) + (-0.6699*H24) + (-0.3931*H25) - 0.9308)$$

$$\begin{aligned} \mathbf{Salida} = \sigma & ((-0.9749 * H26) + (0.3254 * H27) + (-0.9075 * H28) \\ & + (1.3258 * H29) + (-0.1403 * H30) + 0.1743)) \end{aligned}$$

Ecuación 11 RNA Modelo 10-15-5

En donde,

$H1$ = ecuación de la primera neurona de la RNA Modelo 10-15-5

$H2$ = ecuación de la segunda neurona de la RNA Modelo 10-15-5

$H3$ = ecuación de la tercera neurona de la RNA Modelo 10-15-5

Hn = ecuación de la n neurona de la RNA Modelo 10-15-5 (hasta $n = 30$)

x = precipitación pluvial [mm/día]

$Salida$ = turbidez [NTU]

σ = función log sigmoide $\left[\frac{1}{1+e^{-z}} \right]$

4.2.2.2.1. Índice de sensibilidad del número de redes neuronales para la relación turbidez y precipitación

Al realizar el índice de sensibilidad del número de neuronas y al compararlo con los índices de los parámetros calculados anteriormente, se puede observar en la *Tabla 13* que se obtuvo un valor muy superior, es decir, este parámetro es el más significativo de los estudiados.

Tabla 13 Índice de Sensibilidad (número de neuronas) para la relación entre turbidez y precipitación

Análisis de Sensibilidad utilizando el Índice de Sensibilidad “S”	
Parámetro	Índice “S”
Número de neuronas	0.00206

Elaboración propia

Cabe recalcar que el índice de sensibilidad para el número de neuronas se lo realizó considerando el rango del parámetro de 2 a 30 (número de neuronas analizadas previamente) como valor de entrada y sus respectivos R² como valor de salida; siguiendo la Ecuación 8.

4.2.3. Modelo entre turbidez y precipitación con la metodología de superficie de respuesta

Adicionalmente, con el uso de la metodología de superficie de respuesta mediante el software Minitab, se obtuvo como resultado una ecuación para calcular la turbiedad en función de la precipitación. En cuanto a los errores presentados en esta ecuación obtenida, se tuvo para la etapa de calibración del modelo un R² de 0.34. En tanto, para la validación del modelo el coeficiente disminuyó a 0.19, pudiendo representar únicamente el 19 % de los datos de la variable de respuesta, considerándose de igual forma un modelo malo. Los errores de los datos se pueden observar en la Tabla 14.

Tabla 14 Errores del modelo turbidez y precipitación con la metodología de superficie de respuesta

Calibración			Validación		
RMSE	R ²	EF	RMSE	R ²	EF
0.26	0.34	0.93	3.19	0.19	0.8

Elaboración propia

La Ecuación 12 obtenida con este software se presenta a continuación:

$$\mathbf{Turbidez} = 23.61 + 2.776P + 0.0405P^2$$

Ecuación 12 Cálculo de la turbidez con la metodología de superficie de respuesta del proceso

En donde,

Turbidez = valor de salida del nivel de turbidez del agua cruda [NTU]

P = precipitación pluvial [mm/día]

4.3. Respuesta del esfuerzo del tratamiento químico según turbidez del agua cruda y precipitación pluvial en la PTAP-M

4.3.1. Análisis del esfuerzo del tratamiento químico en la PTAP-M

Previamente se efectuó la transformación de los valores del policloruro de aluminio en sulfato de aluminio; esto con el fin de trabajar únicamente con datos de sulfato de aluminio como unidad base del esfuerzo del tratamiento químico o el uso de coagulante según la turbidez del agua cruda. Se obtuvo un valor promedio de sulfato de aluminio consumido igual a 241.27 Kg/día durante todo el período estudiado (1 de enero de 2014 – 31 de mayo de 2020). El valor máximo de coagulante consumido fue igual a 1069.23 Kg/d; este valor fue para el 25 de junio de 2015 que tuvo un nivel de turbidez igual a 333.9 NTU y una precipitación total de 35.37 mm. Cabe recalcar que para ese nivel de turbidez específicamente la planta hizo uso del sulfato de aluminio y PAC conjuntamente (300 Kg/d de sulfato de aluminio y 300 Kg/d de PAC equivalentes a 769.23 Kg/d de sulfato de aluminio). Verificando el día antes y después a tal fecha, se observó que; el día 24 de junio de 2015 tuvo un valor de turbidez de 65.6 NTU para una precipitación igual a 22.59 mm (ambos datos menores en comparación con los valores del 25 de junio de 2015); para este escenario la cantidad de coagulante requerida fue de 281.54 Kg/d (haciendo uso únicamente del sulfato de aluminio). En cuanto al siguiente día del suceso, el 26 de junio de 2015, se observó que existió una precipitación de 16.34 mm con una turbidez igual a 368.79 NTU; para estas condiciones se utilizó 769.23 Kg/d de coagulante en unidades de sulfato de aluminio porque se optó por el uso exclusivo de PAC (300 Kg/d).

A oposición, el menor valor de la cantidad de coagulante requerida corresponde a 78.04 Kg/d del 12 de noviembre de 2015 (mismo año del suceso con mayor cantidad de coagulante utilizada); para una turbidez de 13.04 NTU y con una precipitación total de 3.96 mm. Al observar la data anterior y posterior a este día se evidenció que el día anterior (11 de noviembre de 2015) tuvo una precipitación mínima de 0.45 mm, una turbidez de 15.28 NTU y un uso de sólo sulfato de aluminio igual a 92.86 Kg/d. Al día siguiente (13 de noviembre de 2015), se presentó una precipitación 1.52 mm y una turbidez de 25.65 NTU; para lo cual se utilizó, nuevamente, sólo el sulfato de aluminio igual a 164.97 Kg/d. Siendo visible un comportamiento relativamente esperado y evidente, pero no predecible al cien por ciento al no mantener constante una relación directamente proporcional para todas las fechas en mención con respecto a la sucesión de comportamiento de las variables: a la cantidad de precipitación pluvial, el nivel de turbidez del agua cruda y el esfuerzo de tratamiento químico.

Acorde a la literatura y parcialmente a lo comprobado, el uso de la cantidad del coagulante es directamente proporcional a la turbidez; del mismo modo, las observaciones indican que este uso de coagulante o esfuerzo de tratamiento químico también estaría relacionado directamente con las precipitaciones. Sin embargo, la precipitación más alta se presentó el 18 de octubre de 2019 con 50.24 mm, a esta se le acreditó una turbidez de 290.88 NTU y un uso de 446.98 Kg/d de coagulante, recalando que se hizo uso de los dos tipos de coagulantes con el que trabaja la PATAP-M (sulfato de aluminio con 228.9 Kg/d y PAC con 85.05 Kg/d). En contraposición, con el valor ya mencionado de 35.37 mm de precipitación que ocasionó una turbidez de 333.9 NTU (valor superior con 50.24 mm de precipitación) y hubo un ascenso en el uso del esfuerzo químico o uso de coagulante a 1069.23 Kg/d. Los datos expuestos muestran que para el día con mayor precipitación (18 de octubre de 2019) se evidenció un nivel de turbidez y uso de coagulante menor en comparación al día con una menor precipitación (25 de junio de 2015); lo que se esperaba un comportamiento opuesto al percibido.

Se supone por la evidencia científica y la experiencia, que el aumento de la turbidez en el agua cruda de ingreso conduce al aumento del uso de coagulantes diariamente. Aunque se evidencia que, para el caso de los datos de cantidad de coagulante usado y turbidez de la PTAP-M no existe una relación definida, pero en la mayoría sigue la tendencia de al aumentar la turbidez se aumenta el esfuerzo de tratamiento químico. Por el contrario, para la relación de la cantidad de coagulante usado y precipitación; estas variables muestran un comportamiento independiente entre sí al no estar vigorosamente ligadas con el nivel de turbidez.

Según los resultados (cálculos realizados en base a la sección “2.6.2.2. *Parámetros adicionales*”), la dosis del coagulante varía entre un máximo de 144.31 mg/l para una turbidez de 403.87 NTU y un mínimo de 8.69 mg/l para una turbidez de 13.09 NTU. Estos valores implican la influencia del caudal de ingreso de agua cruda. Complementario a estos valores, los valores de R_{WT} ; relación entre la dosis del coagulante y la calidad de agua cruda según la turbidez, trabajadas en unidades de $mg/l * NTU$ muestra la coherencia y precisión de la dosificación (*Ecuación 2*). Los resultados de la R_{WT} son los siguientes: el menor valor de R_{WT} fue de 0.063 $mg/l * NTU$, se le atribuye al 7 de diciembre de 2014 con una dosificación igual a 22.95 mg/l y una turbidez igual a 366.32 NTU. Mientras que el mayor valor de R_{WT} es 10.39 $mg/l * NTU$ para una dosificación igual a 44.07 mg/l y una turbidez de 4.24 NTU, siendo un suceso del 30 de noviembre de 2016. Estos resultados muestran que, por un lado, se utilizó 0.063 mg de coagulante por litro de agua cruda y por unidad de turbidez; por el otro lado, se utilizó 10.39 mg de coagulante por litro de agua cruda y por unidad de turbidez. Se deduce que

en el segundo caso se está aplicando un mayor esfuerzo de tratamiento químico por utilizar mayor cantidad de coagulante, pero la realidad de la relación R_{WT} es que para el primer caso el valor de la turbidez es significativamente alto, 366.32 NTU, (usando exactamente 194.61 Kg/d de coagulante) en comparación con el valor de la turbidez despreciable de 4.24 NTU (usando exactamente 347.5 Kg/d de coagulante). Esto demuestra a simple vista una incongruencia en la precisión de la dosificación porque la calidad del agua cruda es mejor en el segundo caso (4.24 NTU) donde se utiliza mucha más cantidad de coagulante (mayor esfuerzo de tratamiento químico).

Con todo lo analizado, las observaciones difieren de la experiencia de los técnicos de la PTAP-M, ya que, la dosificación del coagulante se realiza principalmente en base a la experiencia profesional, la misma que considera el nivel de turbidez y las condiciones meteorológicas del momento. Una precipitación fuerte en la zona donde está ubicada la PTAP-M en un lapso de aproximadamente cuatro horas implica un aumento de la turbidez del agua cruda que ingresa a la planta, lo que significa incrementar la dosis a utilizar. Este pronóstico se verifica posteriormente realizando la prueba de jarras con el agua cruda cuya turbiedad ha incrementado. Sin embargo, como muestran los datos, esta relación entre las variables no siempre es exacta, precisa y significativa; lo que representa que existen otros factores de influencia para la turbidez y su potencial dosificación. Un aspecto que debe considerarse a futuro es el tamaño de la zona de captación aguas arriba de la PTAP-M, en sí el área de estudio es muy pequeña. La superficie de las fuentes potenciales de sedimentos y los factores de movilidad (precipitaciones) de dichos sedimentos y partículas es menor o insignificante para llegar a predecir el esfuerzo del tratamiento químico de la PTAP-M para una turbidez y momento determinado.

4.3.2. Enfoque de modelización de regresión simple y múltiple entre el esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial

Suponiendo que la relación entre el esfuerzo de tratamiento químico, calidad del agua cruda y precipitación es positiva y se mantiene relativamente estrecha como se supone; se espera obtener resultados igualmente positivos con respecto a un modelo idóneo que sea capaz de predecir la cantidad de coagulante necesario a partir del valor de turbidez y precipitación del momento.

Para este punto se conformó tres tipos de relaciones en donde la cantidad de coagulante usada o el esfuerzo de tratamiento químico es el punto en común. De forma independiente se

relacionó la cantidad de coagulante y la turbidez, esperando que esta relación sea la más estrecha y altamente significativa estadísticamente, en comparación con la relación de cantidad de coagulante y la variable precipitación; debido a que, la predicción del esfuerzo de tratamiento químico partiendo de la precipitación se podría considerar independiente porque influyen un sin número de otros factores externos. Por último, englobando todas las variables se relacionó y trabajó con la cantidad de coagulante según los correspondientes datos de turbidez y precipitación. Recalcando que esta relación incluye las tres variables, estas se relacionaron de la siguiente manera: cantidad de coagulante como variable dependiente, la turbidez como variable independiente; en tanto, la precipitación, como variable independiente adicional o variable de control significativa, considerándola así porque la precipitación es una condición climática que reduce el riesgo de una estimación sesgada.

La *Tabla 15* y *Tabla 16* exponen la calibración y validación respectivamente de los siete modelos de distinta naturaleza para cada una de las tres relaciones expuestas anteriormente. En primera instancia, considerando la parte de calibración de los modelos (*Tabla 15*) para los tres conjuntos: Cantidad de coagulante – Turbidez, Cantidad de coagulante – Precipitación y Cantidad de coagulante – Turbidez – Precipitación se obtuvo el mejor resultado de ajuste de R^2 con el modelo de tipo lineal con 0.32, 0.10 y 0.32 respectivamente. Corroborando lo supuesto al ajustar un 32 % entre la relación de cantidad de coagulante y la turbidez, ante un bajo 10 % que ajusta la cantidad de coagulante y la precipitación, pero con un mismo 32 % de ajuste al incluir y trabajar con las tres variables en la relación cantidad de coagulante predicho, turbidez y precipitación conjuntamente.

Tabla 15 Resultados de los errores de las etapas de calibración de los modelos de regresión para el esfuerzo de tratamiento químico, turbidez y precipitación

		Calibración			
<i>Conjunto</i>	Modelo	Ecuación	R²	RMSE	EF
<i>Cantidad de coagulante – Turbidez</i>	Lineal	$y = 1.09704x + 194.2202$	0.32	0.41	0.87
	Lineal sin intercepto	$y = 2.60847x$	0.32	3.67	-0.29
	Cuadrático	$y = -897.720x^2 + 2670.521x + 240.804$	0.11	7446.35	-7246300
	Cuadrático sin intercepto	$y = -897.720x^2 + 2670.521x$	0.11	7452.33	-7246500
	Exponencial	$y = 194.96e^{0.003x}$	0.16	1.81	0.95
	Cúbico	$y = 377.586x^3 - 897.720x^2 + 2670.521x + 240.804$	0.03	65604.29	-46117000
	Cúbico sin intercepto	$y = 377.586x^3 - 897.720x^2 + 2670.521x$	0.03	65598.32	-46117000

Continuación

Tabla 15 Resultados de los errores de las etapas de calibración de los modelos de regresión para el esfuerzo de tratamiento químico, turbidez y precipitación

	Lineal	$y = 5.5045x + 210.4339$	0.10	3.12	0.28
Cantidad de coagulante – Precipitación	Lineal sin intercepto	$y = 20.4199x$	0.10	5.62	-22.18
	Cuadrático	$y = 179.658x^2 + 1526.291x + 240.804$	0.09	2735.26	-13909607
	Cuadrático sin intercepto	$y = 179.658x^2 + 1526.291x$	0.09	2729.277	-13894296
	Exponencial	$y = 204.88e^{0.0142x}$	0.10	2.63	-2.35
	Cúbico	$y = 94.338x^3 + 179.658x^2 + 1526.291x + 240.804$	0.07	24022.62	-61503269
	Cúbico sin intercepto	$y = 94.338x^3 + 179.658x^2 + 1526.291x$	0.07	24035.82	-6828493
	Lineal	$y = 1.03925T + 1.03577P + 190.95928$	0.32	0.84	0.60
Cantidad de coagulante – Turbidez – Precipitación	Lineal sin intercepto	$y = 1.76782T + 10.47024P$	0.29	1.34	-13.61
	Cuadrático	$y = -820.145(T + P)^2 + 2682.233(T+P) + 240.804$	0.12	31016.31	-1433100
	Cuadrático sin intercepto	$y = -820.145(T + P)^2 + 2682.233(T+P)$	0.12	32056.31	-1484300
	Cúbico	$y = 94.338(T + P)^3 + 179.658(T+P)^2 + 1526.291(T+P) + 240.804$	0.03	166775	-10888700
	Cúbico sin intercepto	$y = 94.338(T + P)^3 + 179.658(T+P)^2 + 1526.291(T+P)$	0.03	187675	-1179870

* T =valor de turbidez

* P =valor de precipitación

Elaboración propia

Efectivamente la primera relación entre la cantidad de coagulante y turbidez resultó comparable con la tercera entre cantidad de coagulante, turbidez y precipitación porque las dos mostraron con el modelo lineal que se puede ajustar el 32 % de la variabilidad de la cantidad de coagulante (Tabla 15). Con el fin de desigualar estas dos relaciones y verificar la calidad de los modelos lineales se verificaron los valores de RMSE y EF. Los valores obtenidos para el RMSE fueron 0.41 (Cantidad de coagulante – Turbidez) y 0.84 (Cantidad de coagulante – Turbidez – Precipitación); siendo el primer valor el más idóneo porque considera netamente la turbidez con la cantidad de coagulante, conociendo que estos valores son estrechos y directamente proporcionales. Por último, de la misma forma, los valores de EF obtenidos fueron de 0.87 (Cantidad de coagulante – Turbidez) y 0.60 (Cantidad de coagulante – Turbidez – Precipitación); los dos valores entran dentro del rango apto de 0.5 – 1 para considerar que el

modelo arroja una buena simulación. En definitiva, el modelo lineal de la relación cantidad y coagulante y turbidez tiene los mejores resultados por la evidente relación entre las variables, pero la calidad del modelo lineal que engloba las tres variables (Cantidad de coagulante – Turbidez – Precipitación) es equiparable porque sus diferencias son mínimas en comparación con los demás modelos que presentan resultados pesimistas.

En segunda instancia y haciendo contraste con la calibración, en la validación se modificó la interpretación de los resultados y la calidad de los modelos (*Tabla 16*). De igual forma para las tres relaciones o tres conjuntos los mejores resultados se exponen en el modelo de naturaleza lineal y lineal sin intercepto si se considera al estadístico R^2 , pero considerando los demás errores (RMSE y EF), la interpretación deriva en que los mejores modelos para los tres conjuntos son de naturaleza únicamente lineal. Para el primer conjunto Cantidad de coagulante – Turbidez se obtuvo los valores de 0.34, 2.16 y 0.97 para los estadísticos R^2 , RMSE y EF respectivamente; mostrando que el modelo logra explicar el 34 % de la variabilidad total de los datos, ajustándose de forma no aceptable. El RMSE de 2.16 no es idóneo por no estar relativamente cercano a 0 y el EF de 0.97 está expresando que el modelo es bueno prediciendo los datos de coagulante en base a la turbidez. En el caso del segundo conjunto (Cantidad de coagulante – Precipitación), como era de esperar, el modelo lineal sostiene un ajuste bajo de solo el 4 % de R^2 (valor que disminuyó en un 6 % del modelo lineal en calibración), un valor de 5.31 para el RMSE y con un valor de 0.91 para el EF; aunque el EF se mantiene dentro del rango óptimo, el RMSE muestra un valor relativamente significativo que indica la sobreestimación y subestimación de los datos de coagulante y el R^2 bajo indica que el modelo no alcanza ajustar ni una veintava parte de la totalidad de los datos. Finalmente, para la relación Cantidad de coagulante – Turbidez – Precipitación se considera mejores resultados, pero está compensado con el conjunto Cantidad de coagulante – Turbidez porque mantuvo resultados equivalentes al presentar 0.35, 2.21 y 0.99 para los estadísticos R^2 , RMSE y EF respectivamente.

Tabla 16 Resultados de los errores de las etapas de validación de los modelos de regresión para el esfuerzo de tratamiento químico, turbidez y precipitación

Conjunto	Modelo	Ecuación	R ²	Validación	
				RMSE	EF
Cantidad de coagulante – Turbidez	Lineal	$y = 1.09704x + 194.2202$	0.34	2.16	0.97
	Lineal sin intercepto	$y = 2.60847x$	0.35	3.05	-21.08
	Cuadrático	$y = -897.720x^2 + 2670.521x + 240.804$	0.12	592341	-746236
	Cuadrático sin intercepto	$y = -897.720x^2 + 2670.521x$	0.13	592353	-74747130
	Exponencial	$y = 194.96e^{0.003x}$	0.12	4.42	0.99
	Cúbico	$y = 377.586x^3 - 897.720x^2 + 2670.521x + 240.804$	0.06	29214731	-5653538297
	Cúbico sin intercepto	$y = 377.586x^3 - 897.720x^2 + 2670.521x$	0.06	29214719	-5652464191
Cantidad de coagulante – Precipitación	Lineal	$y = 5.5045x + 210.4339$	0.04	5.31	0.91
	Lineal sin intercepto	$y = 20.4199x$	0.04	9.28	-26.14
	Cuadrático	$y = 179.658x^2 + 1526.291x + 240.804$	0.04	1349	-8169
	Cuadrático sin intercepto	$y = 179.658x^2 + 1526.291x$	0.04	1337	-6929.64
	Exponencial	$y = 204.88e^{0.0142x}$	0.12	6.61	0.98
	Cúbico	$y = 94.338x^3 + 179.658x^2 + 1526.291x + 240.804$	0.03	4521.338	-10619.44
	Cúbico sin intercepto	$y = 94.338x^3 + 179.658x^2 + 1526.291x + 240.804$	0.03	4509.35	-9198.223
Cantidad de coagulante – Turbidez – Precipitación	Lineal	$y = 1.03925T + 1.03577P + 190.95928$	0.35	2.21	0.99
	Lineal sin intercepto	$y = 1.76782T + 10.47024P$	0.27	3.36	-20.84
	Cuadrático	$y = -820.145(T + P)^2 + 2682.233(T+P) + 240.804$	0.14	625282.7	-85754825
	Cuadrático sin intercepto	$y = -820.145(T + P)^2 + 2682.233(T+P)$	0.14	625294.74	-85887169
	Cúbico	$y = 94.338(T + P)^3 + 179.658(T+P)^2 + 1526.291(T+P) + 240.804$	0.07	9407924.29	-904738071
	Cúbico sin intercepto	$y = 94.338(T + P)^3 + 179.658(T+P)^2 + 1526.291(T+P)$	0.07	9407912	-904308419

* T=valor de turbidez

* P=valor de precipitación

Elaboración propia

Definitivamente, se observó que en la relación Cantidad de coagulante – Turbidez los estadísticos (R^2 , RMSE y EF) se alteraron al ir de calibración a validación; mejorando con respecto al R^2 y empeorando para el RMSE; en resumen 0.32, 0.41 y 0.87 para la calibración y 0.34, 2.16 y 0.97 para la validación. Igualmente, para la relación Cantidad de coagulante – Turbidez – Precipitación los valores mejoraron para el R^2 y decayeron para el RMSE, mientras que para el EF se mantienen dentro del rango idóneo; en resumen 0.32, 0.84 y 0.66 para la calibración y 0.35, 2.21 y 0.99 para la validación (R^2 , RMSE y EF). El hecho de que el valor del R^2 mejore en ambas relaciones significa un buen ajuste a los datos, pero la disminución del RMSE denota que los datos predichos se alejan de la media.

4.3.3. Modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial con Redes Neuronales Artificiales

4.3.3.1. Análisis de sensibilidad de parámetros para el modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial

Este apartado mantiene el mismo fundamento inicial expuesto en la sección “4.2.2.1. Análisis de sensibilidad de parámetros para el modelo entre turbidez y precipitación”, basado en los resultados de Chang & Liao (2012). Se consideró primero los parámetros de tiempo de entrenamiento y velocidad de aprendizaje, manteniendo así, como constante el número de neuronas (3).

Como se exhibe en la *Tabla 17*, con respecto al parámetro de tiempo de entrenamiento, se puede considerar que el comportamiento va dirigido de la siguiente manera: a medida que el valor del tiempo de entrenamiento aumenta, el valor del R^2 de validación incrementa del mismo modo; es decir, la calidad del modelo mejora y disminuye su error. Este comportamiento se asemeja con los resultados presentados por Chang & Liao (2012), quienes mostraron que el error de simulación disminuye cuando aumentan los tiempos de entrenamiento o la velocidad de aprendizaje. Este enunciado se ejemplifica numéricamente (*Tabla 17*); para un tiempo de entrenamiento de 500 se tiene un R^2 de 0.48, mientras que aumentando el tiempo de entrenamiento a 20000 se tiene un R^2 de 0.58; manifestándose que el incremento del R^2 resulta relativamente importante entre los cinco escenarios analizados ante el significativo incremento del valor del tiempo de entrenamiento en cada escenario.

Tabla 17 Cambios en la precisión de la predicción bajo diferentes parámetros RNA para la relación esfuerzo tratamiento químico, turbidez y precipitación

a. Escenarios con varios tiempos de entrenamiento

<i>Escenario</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Número de neuronas	3	3	3	3	3
Tiempo de entrenamiento	500	5000	10000	15000	20000
Velocidad de aprendizaje	0.1	0.1	0.1	0.1	0.1
Función de transferencia	S	S	S	S	S
R ² calibración	0.42	0.42	0.42	0.42	0.49
R ² validación	0.48	0.49	0.49	0.48	0.58

b. Escenarios con varias velocidades de aprendizaje

<i>Escenario</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
Número de neuronas	3	3	3	3
Tiempo de entrenamiento	20000	20000	20000	20000
Velocidad de aprendizaje	0.001	0.01	0.1	0.5
Función de transferencia	S	S	S	S
R ² calibración	0.419	0.422	0.419	0.401
R ² validación	0.458	0.466	0.486	0.479

c. Escenario con distinta función de transferencia

<i>Escenario</i>	<i>1</i>
Número de neuronas	3
Tiempo de entrenamiento	20000
Velocidad de aprendizaje	0.1
Función de transferencia	T
R ² calibración	0.422
R ² validación	0.48

Elaboración propia

Del mismo modo, se evidenció que la velocidad de aprendizaje mantiene cierto patrón, al aumentar el valor de la velocidad de aprendizaje de 0.001 hasta 0.1, los valores de R² del modelo van en aumento de 0.458 hasta 0.486. Pero, con valores de la velocidad de aprendizaje mayores a 0.1, como el 0.5, el valor de R² tiende a disminuir nuevamente. Por lo tanto, el mejor valor para este parámetro en este caso es 0.1; concordando con el mismo valor de la velocidad de aprendizaje para el caso de Chang & Liao (2012).

El parámetro de la función de transferencia (seleccionado para las capas ocultas y capas de salida) se considera, en base a la literatura pertinente, un parámetro predominante para el correcto funcionamiento de cualquier modelo; por aquello, resulta de suma importancia seleccionar la mejor función de transferencia según la situación (Chang & Liao, 2012). En la selección de las funciones de transferencia, cuando se utilizan funciones inadecuadas, los errores de predicción aumentará considerablemente y la calidad del modelo será baja o inapropiada para predecir el comportamiento de las variables. Se demostró estadísticamente (*Tabla 17*) que el modelo de RNA para las variables turbidez, precipitación o esfuerzo de tratamiento químico presentó los mejores resultados con la función de transferencia log sigmoide.

Con todo el análisis desglosado hasta el momento, la *Tabla 18* presenta el análisis de la sensibilidad de los parámetros para la relación entre esfuerzo del tratamiento químico, turbidez y precipitación. Según el índice de sensibilidad “S”, el parámetro más sensible (entre los considerados para la sección) es la velocidad de aprendizaje por obtener el valor absoluto más alto. Entonces, concretamente la predicción del esfuerzo del tratamiento químico es más sensible al parámetro velocidad de aprendizaje que al parámetro tiempo de entrenamiento.

Tabla 18 Índice de Sensibilidad (tiempo de entrenamiento y velocidad de aprendizaje) para la relación esfuerzo de tratamiento químico, turbidez y precipitación

Análisis de Sensibilidad utilizando el Índice de Sensibilidad “S”	
<i>Parámetro</i>	<i>Índice “S”</i>
Tiempo de entrenamiento	0.000509
Velocidad de aprendizaje	0.00433

Elaboración propia

4.3.3.2. Ensayo de Redes Neuronales Artificiales para el modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial

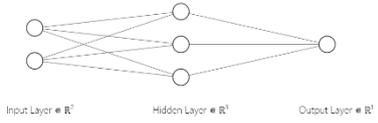
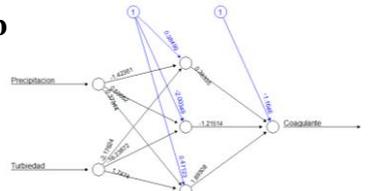
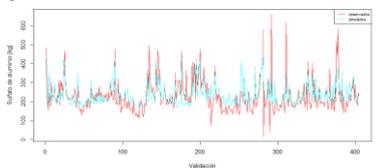
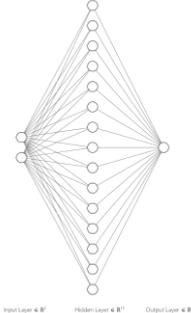
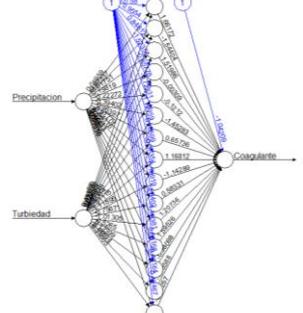
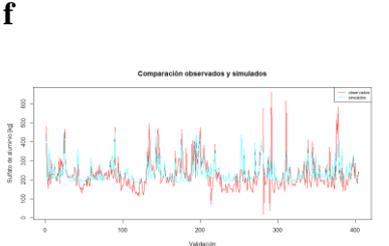
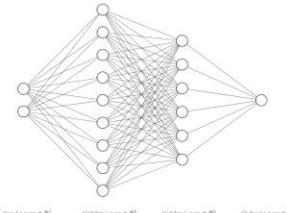
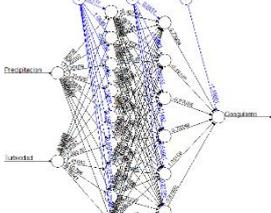
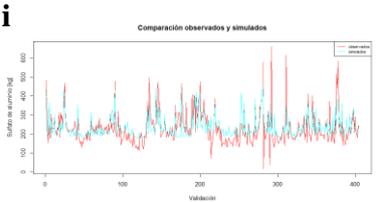
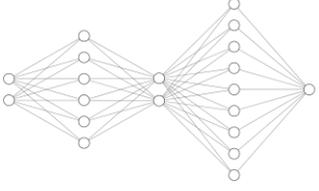
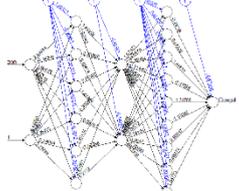
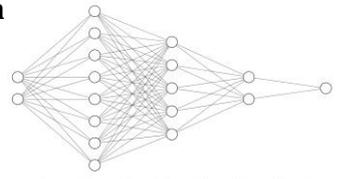
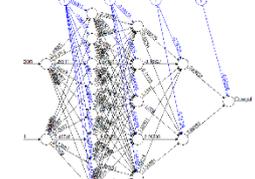
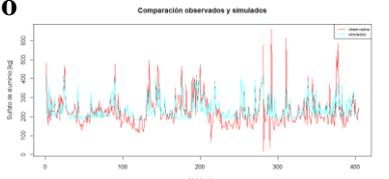
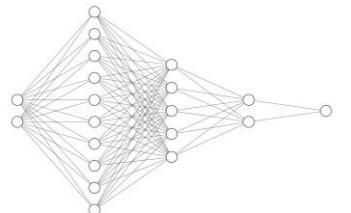
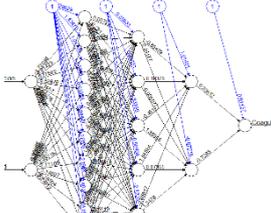
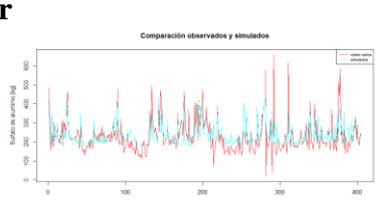
Con el fin de obtener un modelo adecuado para la relación entre el esfuerzo de tratamiento químico, la calidad del agua y la precipitación de la PTAP-M se entrenó redes neuronales artificiales con un variado número de capas internas y neuronas; desde 1 hasta 4 capas internas y desde 3 hasta 15 neuronas por capa interna; recalando que estos arreglos de capas internas y neuronas fueron optados para ser ensayados en base a literatura pertinente. En total se entrenó 7 modelos diferentes (con un error del 0.1 y un máximo de 20000 ciclos de iteraciones); en la capa de entrada de todos los modelos se trabajó con dos entradas correspondientes al nivel de

turbidez y al valor de precipitación; la capa de salida es un único valor correspondiente al uso de coagulante que refleja el esfuerzo de tratamiento químico en la PTAP-M.

Como se muestra en la *Tabla 19*, para cada modelo se presentó el bosquejo de estructura base de la RNA (secciones a, d, g, j, m, p, s), de la misma forma, el sistema neuronal final de la RNA (secciones b, e, h, k, n, q, t) con los valores de las neuronas, las capas y las redes; es decir, las entradas, la parte algorítmica y la salida (trabajadas con la función de transferencia sigmoide entre las capas intermedias y la de salida). Conjuntamente, en la tercera columna de la *Tabla 19* se adjuntó las secciones c, f, i, l, o, r, u; en estas gráficas se evidencia el contraste del esfuerzo del tratamiento químico observado (color rojo) contra el esfuerzo del tratamiento químico predicho (color cian) por cada modelo de RNA en la etapa de validación o verificación. En dichas secciones (c, f, i, l, o, r, u) se percibe que todos los 7 modelos mantienen la línea de predicción similar entre sí; determinando que los modelos no predicen de manera satisfactoria los valores máximos y mínimos del esfuerzo de tratamiento químico; en otras palabras, los modelos resultaron pocos sensibles a valores extremos. Esto se corrobora porque, en la calibración de los 7 modelos todos presentan un valor del R^2 igual a 0.48 o 0.49; lo que se resume en que los modelos pueden predecir sólo el 48 y 49 % del comportamiento de la variable dependiente a partir de las variables independientes; aunque no es un valor insignificante se lo considera bajo por no alcanzar ni la mitad de predicción. En la parte de validación (*Tabla 16*), el R^2 aumentó de manera importante en 8 hasta 9 puntos más; llegando a predecir el 57 % en cuatro de los modelos (15, 6-2-9, 8-5-2 y 10-5-2), el 58 % en los tres modelos restantes (3, 9-6 y 8-6-4-2).

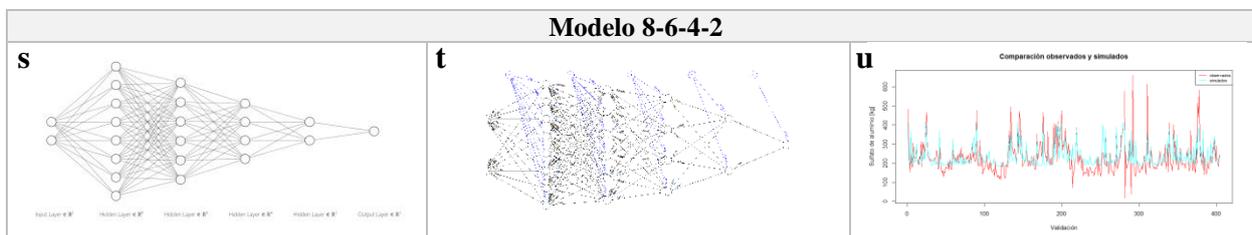
Enfatizando en los modelos con los mejores resultados globales; uno de los modelos seleccionados fue el Modelo 3, se lo entrenó con una única capa interna de tres neuronas. Este modelo presentó uno de los mejores rendimientos entre los demás modelos según los estadísticos. Se consideró que el Modelo 3 mostró un desempeño pobremente aceptable al tener un R^2 de 0.58 (*Tabla 20*), porcentaje que indica que el modelo explica el 58 % de la variabilidad total de los datos y que se está ajustando de forma aceptable a los datos por ser un porcentaje que engloba más de la mitad de todos los datos. Sin embargo, el valor del R^2 no resulta suficiente para calificar la efectividad del modelo, porque al tener un R^2 vagamente significativo y conociendo que se trabaja con los predictores estadísticamente significativos (turbidez y precipitación), se predice la relación o asociación entre un cambio del valor de los predictores y su posterior cambio en el valor de respuesta del esfuerzo de tratamiento químico.

Tabla 19 Resumen de los modelos de RNA entrenados para la relación esfuerzo de tratamiento químico, turbidez y precipitación

Estructura base de la RNA	Sistema global final de la RNA	Esfuerzo del tratamiento químico observado y predicho
Modelo 3		
<p>a</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^3$ Output Layer $\bullet R^1$</p>	<p>b</p> 	<p>c</p> 
Modelo 15		
<p>d</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^{15}$ Output Layer $\bullet R^1$</p>	<p>e</p> 	<p>f</p> 
Modelo 9-6		
<p>g</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^9$ Hidden Layer $\bullet R^6$ Output Layer $\bullet R^1$</p>	<p>h</p> 	<p>i</p> 
Modelo 6-2-9		
<p>j</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^6$ Hidden Layer $\bullet R^2$ Hidden Layer $\bullet R^9$ Output Layer $\bullet R^1$</p>	<p>k</p> 	<p>l</p> 
Modelo 8-5-2		
<p>m</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^8$ Hidden Layer $\bullet R^5$ Hidden Layer $\bullet R^2$ Hidden Layer $\bullet R^2$ Output Layer $\bullet R^1$</p>	<p>n</p> 	<p>o</p> 
Modelo 10-5-2		
<p>p</p>  <p>Input Layer $\bullet R^2$ Hidden Layer $\bullet R^{10}$ Hidden Layer $\bullet R^5$ Hidden Layer $\bullet R^2$ Hidden Layer $\bullet R^2$ Output Layer $\bullet R^1$</p>	<p>q</p> 	<p>r</p> 

Continuación

Tabla 19 Resumen de los modelos de RNA entrenados para la relación esfuerzo de tratamiento químico, turbidez y precipitación



* **Tabla 19:** Estructura base de la RNA de cada modelo (a, d, g, j, m, p, s). Sistema global final de la RNA de cada modelo (b, e, h, k, n, q, t). Esfuerzo del tratamiento químico observado y predicho de cada modelo (c, f, i, l, o, r, u).

Elaboración propia

Entonces, adicionalmente se considera el RMSE y el índice EF para un análisis más completo del modelo (Tabla 20, sección “Validación”). Con un valor de 1.68 para el RMSE, reiterando que para esta medida se busca el valor más cercano a 0, se confirma la existencia de un ajuste aceptable con respecto a la precisión del modelo para predecir los valores respuesta en función de los valores observados; esta medida es considerada un criterio importante a tener presente para los modelos de predicción. Con un valor de 0.85 para el EF, se expresa una buena simulación según lo establecido en el Capítulo 2 (Marco teórico); este índice a más de considerar la variabilidad del modelo como lo hace el R^2 , considera el valor de respuesta como tal para verificar que el valor sea cercano al observado o está siendo sobreestimado o subestimado. En este caso, EF de 0.85 ($EF = 0.5-1$) muestra que el modelo tiene un rendimiento aceptable, pero se debe considerar que existen valores de respuesta del esfuerzo del tratamiento químico que se están sobreestimando y subestimando; lo que se refleja en el valor del RMSE.

Con resultados similares sobresale el Modelo 8-6-4-2, la arquitectura de este modelo es de cuatro capas internas; la primera de 8 neuronas, una segunda capa con 6 neuronas, la tercera capa con 4 neurona y una última con 2 neuronas (Tabla 19, sección (s, t)). Este modelo presentó un R^2 de 0.58 (mismo valor que en el Modelo 3), un valor de RMSE de 1.81 (más alto que en el Modelo 3) y un 0.88 (más alto que en el Modelo 5) para el índice EF (Tabla 20, sección “Validación”). De la misma forma, el valor del R^2 se mantiene significativo como en el anterior caso; el valor del RMSE empeoró ínfimamente por la sobreestimación y subestimación de los valores de respuesta que se ve reflejado, de igual forma, en el aumento del valor del EF a comparación del Modelo 3; por lo que se considera que el modelo presenta un rendimiento hasta cierto punto aceptable.

Tabla 20 Resultados de los errores de las etapas de calibración y validación de los modelos de RNA para la relación esfuerzo de tratamiento químico, turbidez y precipitación

Relación entre variable dependiente del uso de coagulante y las variables independientes de turbidez y precipitación (Llaucay + Nudpud)

Modelo	Calibración			Validación		
	R ²	RMSE	EF	R ²	RMSE	EF
Modelo 3	0.49	0.05	0.53	0.58	1.68	0.85
Modelo 15	0.48	0.39	0.64	0.57	2.07	0.86
Modelo 9-6	0.49	0.07	0.42	0.58	1.92	0.84
Modelo 6-2-9	0.48	0.02	0.72	0.57	2.03	0.85
Modelo 8-5-2	0.49	0.11	0.48	0.57	1.89	0.83
Modelo 10-5-2	0.49	0.16	0.22	0.57	1.88	0.83
Modelo 8-6-4-2	0.49	0.11	0.71	0.58	1.81	0.88

Elaboración propia

En definitiva, se infiere que el Modelo 3 aunque con menor complejidad en su estructura, presenta un desempeño mejorado, casi imperceptible, en comparación con el Modelo 8-6-4-2. Pero, se desalienta la futura aplicación de los modelos en la PTAP-M por la baja relevancia estadística en la predicción (suponiendo resultados optimistas y seguros con un 80 – 90 % de predicción); se debería considerar otros factores que son significativos para las relaciones establecidas y que aporten con la información relevante y necesaria a la RNA para el desarrollo de un modelo con mejor capacidad.

La Ecuación 13 para el Modelo 3 de RNA para la relación calidad del agua, esfuerzo de tratamiento químico y precipitación se presenta a continuación:

$$H1 = \sigma((0.7622 * P) + (1.2116 * T) - 1.5421)$$

$$H2 = \sigma((0.1244 * P) + (-9.0135 * T) - 1.07595)$$

$$H3 = \sigma((-0.658 * P) + (7.974 * T) + 1.1179)$$

$$Salida = \sigma((0.3317 * H1) + (-0.708 * H2) + (0.53425 * H3) - 0.1871)$$

Ecuación 13 RNA Modelo 3

En donde,

$H1$ = ecuación de la primera neurona de la RNA Modelo 3

$H2$ = ecuación de la segunda neurona de la RNA Modelo 3

$H3$ = ecuación de la tercera neurona de la RNA Modelo 3

P = precipitación pluvial [mm/día]

T = turbidez [NTU]

Salida = cantidad de coagulante [Kg]

$$\sigma = \text{función log sigmoide} \left[\frac{1}{1+e^{-z}} \right]$$

De igual forma, se presenta a continuación el sistema de ecuaciones para el Modelo 8-6-4-2:

Primera capa interna (8 neuronas)

$$H1 = \sigma((-0.2946*P) + (-0.29457*T) - 0.6447)$$

$$H2 = \sigma((-2.9932*P) + (7.9641*T) - 0.9566)$$

$$H3 = \sigma((0.5951*P) + (6.4299*T) - 0.2491)$$

$$H4 = \sigma((-0.5838*P) + (6.7209*T) - 0.2744)$$

$$H5 = \sigma((0.0218*P) + (2.7638*T) - 1.1016)$$

$$H6 = \sigma((1.6442*P) + (10.944*T) + 1.1953)$$

$$H7 = \sigma((0.2352*P) + (2.2781*T) - 1.4920)$$

$$H8 = \sigma((-0.64*P) + (3.934*T) + 0.12)$$

Segunda capa interna (6 neuronas)

$$H9 = \sigma((1.3404*H1) + (1.6459*H2) + (0.7077*H3) + (-1.6613*H4) + (-2.4424*H5) + (0.8845*H6) + (-1.68*H7) + (0.91*H8) + (1.54*H9)) + (1.99*H10) - 0.95362)$$

$$H10 = \sigma((-1.26*H1) + (-0.29*H2) + (0.54*H3) + (-1.18*H4) + (-1.05*H5) + (1.03*H6) + (-0.8864*H7) + (-0.9322*H8) - 1.04)$$

$$H11 = \sigma((1.3658*H1) + (4.4804*H2) + (-1.3292*H3) + (2.4521*H4) + (0.94023*H5) + (-0.6542*H6) + (-0.9144*H7) + (0.3461*H8) - 1.156)$$

$$H12 = \sigma((-0.1434*H1) + (-0.6903*H2) + (1.2304*H3) + (1.10*H4) + (1.2843*H5) + (0.3969*H6) + (-0.8549*H7) + (0.6694*H8) + 1.312)$$

$$H13 = \sigma((-3.7335*H1) + (1.0988*H2) + (1.1408*H3) + (0.6408*H4) + (0.4388*H5) + (0.3014*H6) + (0.7097*H7) + (-0.501*H8) - 0.3256)$$

$$H14 = \sigma((-1.2167*H1) + (0.9441*H2) + (0.7508*H3) + (-1.21*H4) + (1.6654*H5) + (1.5221*H6) + (-1.0467*H7) + (1.6343*H8) - 1.457)$$

Tercera capa interna (4 neuronas)

$$H15 = \sigma((-0.5925*H9) + (0.308*H10) + (0.245*H11) + (-0.04*H12) + (0.095*H13) + (-0.618*H14) + 1.2902)$$

$$H16 = \sigma ((1.34719 * H9) + (-2.20 * H10) + (-0.55 * H11) + (-1.61 * H12) + (0.49 * H13) + (0.29 * H14) - 0.3226)$$

$$H17 = \sigma ((0.084 * H9) + (2.50 * H10) + (0.18 * H11) + (2.79 * H12) + (0.701 * H13) + (1.443 * H14) - 2.25)$$

$$H18 = \sigma ((0.317 * H9) + (1.065 * H10) + (0.347 * H11) + (-2.479 * H12) + (-1.495 * H13) + (-0.249 * H14) - 0.4015)$$

Cuarta capa interna (2 neuronas)

$$H19 = \sigma ((0.687 * H15) + (-0.536 * H16) + (-0.571 * H17) + (1.9596 * H18) - 0.68)$$

$$H20 = \sigma ((-0.5432 * H15) + (0.4176 * H16) + (-0.471 * H17) + (1.3963 * H18) + 0.409)$$

$$\text{Salida} = \sigma ((-0.4267 * H16) + (-1.306 * H17) + 1.1)$$

Ecuación 14 RNA Modelo 8-6-4-2

En donde,

$H1$ = ecuación de la primera neurona de la RNA Modelo 8-6-4-2

$H2$ = ecuación de la segunda neurona de la RNA Modelo 8-6-4-2

$H3$ = ecuación de la tercera neurona de la RNA Modelo 8-6-4-2

Hn = ecuación de la n neurona de la RNA Modelo 8-6-4-2 (hasta $n = 20$)

P = precipitación pluvial [mm/día]

T = turbidez [NTU]

$Salida$ = cantidad de coagulante [Kg]

σ = función log sigmoide $\left[\frac{1}{1+e^{-z}} \right]$

4.3.3.2.1. Índice de sensibilidad del número de redes neuronales para la relación esfuerzo de tratamiento químico, turbidez y precipitación

Haciendo mención nuevamente a Chang & Liao (2012); en su artículo se demostró que para la predicción de la turbidez del agua cruda en base a los datos pluviométricos de aguas arriba, el parámetro con mayor sensibilidad es el número de neuronas antes que los dos parámetros de tiempo de entrenamiento y velocidad de aprendizaje. Para el actual caso, se determinó que la velocidad de aprendizaje fue el parámetro más sensible con un índice “S” igual a 0.00433

(Tabla 18). En contra del pronóstico, el parámetro número de neuronas obtuvo un índice “S” de 0.00353 (Tabla 21) sutilmente menor al valor de la velocidad de aprendizaje (0.00433). Revelando que para el modelo entre la relación esfuerzo del tratamiento químico, turbidez y precipitación los parámetros más sensibles por igual son el número de neuronas y la velocidad de aprendizaje. Se subraya que el valor del índice “S” toma el valor absoluto y, en el caso del parámetro número de neuronas, el valor negativo indica que la entrada y salida del modelo están inversamente relacionadas.

Tabla 21 Índice de Sensibilidad (número de neuronas) para la relación esfuerzo de tratamiento químico, turbidez y precipitación

Análisis de Sensibilidad utilizando el Índice de Sensibilidad “S”	
Parámetro	Índice “S”
Número de neuronas	-0.00353

Elaboración propia

4.3.4. Modelo entre esfuerzo de tratamiento químico, turbidez del agua cruda y precipitación pluvial con la metodología de superficie de respuesta

Para la relación de esfuerzo de tratamiento químico, calidad del agua y precipitación también se optó por aplicar el método del diseño de superficie de respuesta, mediante el software Minitab; obteniéndose así la Ecuación 15 para la predicción de la cantidad de coagulante a partir de un valor de precipitación y turbidez de un determinado momento para la PTAP-M.

$$C = 177.11 + 0.511 * P + 1.759 * T - 0.0733 * P^2 - 0.002434 * T^2 + 0.01704 * PT$$

Ecuación 15 Cantidad de coagulante modelada con la metodología de superficie de respuesta

En donde,

C = valor de salida de la cantidad de coagulante necesario para un nivel de turbidez del agua cruda y precipitación pluvial de entrada de un determinado momento [kg]

T = nivel de turbidez del agua cruda [NTU]

P = precipitación pluvial [mm/día]

La validación de la calidad del modelo se muestra en la Tabla 22 con valores de los errores similares entre la etapa de calibración y la etapa de validación.

Tabla 22 Errores del modelo esfuerzo de tratamiento químico, turbidez y precipitación con la metodología de superficie de respuesta

Calibración			Validación		
R ²	RMSE	EF	R ²	RMSE	EF
0.51	1.06	0.99	0.51	1.59	0.70

Elaboración propia

En la calibración se obtuvo un valor de R² igual a 0.51 y este valor se mantiene constante para la validación del modelo. Un valor bastante bueno de 1.59 para el RMSE (RMSE=0) y un valor de 0.70 para el EF (EF=0.5 – 1) en la validación indica que el modelo está obteniendo valores aceptables, pero predice sólo el 51 % de la variabilidad de los datos; pudiendo considerarlo como un modelo poco adecuado para la predicción de la cantidad de coagulante (Tabla 22).

4.3.5. Elasticidades de la calidad del agua cruda en la PTAP-M

Con respecto a las elasticidades de la turbidez obtenidas se observó que la elasticidad de la calidad del agua cruda que ingresa a la PTAP-M varía entre 0.97 y 1.87, con un error estándar de 0.036 que significa que mientras más bajo es este valor todos los valores de elasticidad se encuentran más precisos con respecto a la media de los datos. Por un lado, el valor de la elasticidad de la turbidez de 0.97 implica que una disminución del 1 % de la turbidez del agua cruda resulta en una reducción del 0.97 % del esfuerzo del tratamiento químico.

Por otro lado, la elasticidad de 1.87 significa que una disminución del 1 % de la turbidez del agua cruda de ingreso a la PTAP-M es igual a la reducción del 1.87 % del esfuerzo del tratamiento químico o uso de coagulantes. En este caso, con el rango de elasticidad obtenida se consuma que, con el mínimo valor de 0.97, la relación entre calidad del agua y esfuerzo de tratamiento químico es inelástico porque un cierto cambio porcentual de la turbidez causa un bajo cambio porcentual en la cantidad usada de coagulante. Por el contrario, el valor de 1.87 fue el valor máximo para el presente estudio, se considera que la relación entre las variables de calidad del agua y esfuerzo de tratamiento químico es elástica ya que un cierto cambio porcentual en la turbidez provoca un alto cambio porcentual en el uso de coagulante.

4.4. Comparación de los resultados con la literatura

4.4.1. Comparación del modelo entre la relación turbidez y precipitación

Los resultados mostraron que no existe una buena predicción de la turbidez del agua cruda, en función de la precipitación pluvial registrada en las dos estaciones meteorológicas ubicadas en la zona de las captaciones del agua destinada a la potabilización. Para el caso de la regresión

exponencial se obtuvo un R^2 de 0.21, superior a los resultados obtenidos por Göransson et al. (2013) en su artículo donde obtienen un coeficiente de 0.12 al realizar el análisis entre estas variables de estudio. Cabe mencionar que existe una diferencia significativa entre los valores de turbidez analizados, en el caso del artículo la turbidez máxima con la que se trabajó fue de 45 NTU, en cambio, en la PTAP-M se encontró una turbiedad superior a los 800 NTU.

Por otro lado, al analizar el artículo realizado por López et al. (2015) y comparar la precipitación con la turbidez en la PTAP de Ibagué se obtuvo una relación de 0.54 y se observó que al incrementar la cantidad de lluvia el valor de turbidez también aumentaba, algo muy distinto en el presente estudio, donde no existió una relación definida ya que en muchos de los casos, la turbidez aumentaba con valores de precipitación bajos, de manera contraria muchas de las veces la turbidez fue baja con valores altos de precipitación. El máximo valor promedio anual de turbidez en el artículo fue de 43 NTU y el valor máximo de la precipitación fue de 300 mm/año. Además de ello, los autores señalaron que la precipitación es de suma importancia en muchos aspectos relacionados con el abastecimiento y calidad del agua, debido a que las fuentes de abastecimiento son alimentadas de manera directa o indirecta por la precipitación pluvial, destacando que la calidad que presenta el agua cruda puede estar influenciada por la alteración del suelo, degradación de la microcuenca, erosión cercana a los cauces de agua y el uso de productos químicos producto de las actividades económicas que se desarrollan en la zona.

En este contexto, Zúñiga et al. (2016) obtuvieron que, en los meses con mayor turbidez del agua cruda existió valores altos de precipitación, lo cual evidenció el arrastre de los sedimentos hacia los ríos, sin embargo, para algunos meses analizados en los cuales existió una alta turbidez la precipitación fue escasa, similar a lo sucedido en el presente estudio. Los autores explicaron que esto pudo deberse a que, algunas precipitaciones no comprometen toda el área de la cuenca del río Combeima (área de estudio), donde existe zonas con procesos erosivos avanzados y al estar en contacto con la lluvia puede liberar sedimentos que eventualmente llegan al cuerpo de agua, de esta forma se menciona, que una alta turbidez en la captación de la PTAP es el resultado de dos factores: precipitación alta y áreas alteradas (erosionadas). La relación entre la variable turbidez y precipitación pluvial obtenidas en este estudio fue un R^2 de 0.45.

Por otro lado, en cuanto al uso de las redes neuronales artificiales para la predicción de la turbidez, el resultado fue menos eficiente que el obtenido al utilizar la regresión exponencial,

se observó un R^2 equivalente al 19 %, un valor muy distinto al obtenido por Chang & Liao (2012), que lograron obtener un nivel de predicción de alrededor del 80 % de los datos. Los parámetros de sensibilidad fueron los mismos en ambos casos, a diferencia del tiempo de entrenamiento que para el presente estudio fue de 20000 y en el artículo se trabajó con 500. En el artículo propuesto por Sari et al. (2017) se realizó un monitoreo de las cuatro estaciones climáticas, donde se analizó el nivel del agua en los ríos que varió entre 0.53 y 2.8 m, turbidez que osciló entre 1.2 y 1161.7 NTU y los sólidos disueltos con una media de 0.19 g/L. El valor de la turbidez estuvo fuertemente influenciado por la precipitación, los resultados mostraron una buena relación entre la turbidez y el nivel del agua con un R^2 de 0.73, presentando una relación positiva al utilizar RNA, el modelo que presentó este resultado fue entrenado con 20 neuronas en la capa interna, velocidad de aprendizaje 0.001 y tiempo de entrenamiento 200000, utilizando la serie de entrenamiento total.

En este punto del estudio se tiene asegurado la existencia de la relación entre la calidad del agua cruda (turbidez) y varios factores como la precipitación. Esto explica la variación de la turbidez durante el año, además, debido a estas relaciones, la calidad del agua cruda al ingreso de la PTAP-M se considera mala, aceptable o buena en base al nivel de turbidez (uno de los parámetros medidos y aplicado en la presente tesis), por ende, este aumento de la turbidez conduce al aumento del esfuerzo de tratamiento químico.

4.4.2. Comparación del modelo entre la relación de esfuerzo de tratamiento químico, turbidez y precipitación

La relación múltiple entre el esfuerzo del tratamiento químico, turbidez y precipitación se investigó y utilizó varios métodos para modelar el esfuerzo del tratamiento químico o uso de coagulante en base al conjunto de entrada de turbidez y precipitación. La modelización con mejor resultado se presentó mediante el uso de RNA. La aplicación de las RNA es compleja en el sentido de cómo trabaja en sí la red neural por qué; por un lado, trabaja como un modelo de “caja negra” y por el otro lado, busca imitar la función de las neuronas en el cerebro humano. Con el entrenamiento de los diferentes modelos de RNA se interpretó que; mientras más sean las variables de entrada (turbidez y precipitación), mayor será el número de neuronas por capas internas que puedan ser consideradas. Por la complejidad de su funcionamiento, se supone que al relacionar las tres variables (esfuerzo de tratamiento químico, turbidez y precipitación) para encontrar un modelo que prediga la cantidad de coagulante necesario para un valor de precipitación y turbidez al ingreso de la planta de forma inmediata, se obtuvo resultados lo

suficientemente optimistas para hacer uso de esta herramienta en la PTAP-M. El mejor modelo de predicción formado mediante las RNA tiene un R^2 de 0.58, atribuible al Modelo 3 para el presente caso; con estructura: capa de entrada con dos neuronas, capa oculta con tres neuronas y capa de salida con una neurona.

A diferencia de lo mencionado, en el estudio de Mondorf (2020) se relacionó las mismas variables de precipitación, calidad del agua cruda (nivel de turbidez) y el esfuerzo de tratamiento químico en base a diagramas de series de tiempo y análisis de regresión simple y múltiple. Mondorf (2020) encontró mediante las series de tiempo que, las relaciones establecidas muestran algunos grados de correlación, es decir que para la turbidez del agua cruda y el uso de coagulante parece seguir el curso estacional según la precipitación; variando el valor de R^2 desde 0 hasta 0.71, implica que el 71 % de la variable dependiente es explicada por las variables independientes y que sólo el 29 % se podría explicar mediante otros factores. A partir de esta información se desarrolló una ecuación para el modelo de regresión múltiple.

Con mejores resultados, pero con una relación menos compleja (se considera las características físico-químicas del agua cruda y la dosificación de coagulantes), los estudios de Peña (2016) y Peña Rojas & Flores del Pino (2014) plantearon optimizar la dosificación de coagulantes en una PTAP de Perú a partir de la turbidez, el pH, la conductividad, el color y los sólidos disueltos totales; mediante el uso de redes neuronales artificiales como herramienta. De esta forma se determinó que la red de cuatro capas mostró el mejor resultado: capa de entrada con cinco neuronas, la primera capa oculta de seis neuronas, la segunda capa oculta de cinco neuronas y la capa de salida con una neurona; obteniéndose una correlación de 98.4 % en el entrenamiento, con asertividad de 96.9 % y demostrando que los factores con mayor influencia son el color, turbidez y pH. En definitiva, mediante el uso de RNA, para este caso, se predijo la dosis óptima de coagulante con una probabilidad de error máximo de 1.6 %.

Otro caso con resultados positivos y más óptimo que el del presente estudio, es el estudio realizado por Baouab & Cherif (2018), que a partir del uso de redes neuronales artificiales trabajaron con el modelo L45-MOD (específico para el agua cruda y para turbidez inferior a 45.5 NTU) y con el modelo ATP-MOD (para el agua cruda con turbidez superior a 45.5 NTU). Estos modelos presentaron buenas relaciones, con un coeficiente de correlación que superó el 80 % (81 % para L45-MOD). Sumado a esto último, en el artículo realizado por Baouab & Cherif (2018) se demostró que el rendimiento de la red neuronal artificial depende de la base de datos de entrenamiento, por lo que, para el presente caso puede ser un motivo por el cual los

modelos no presentan resultados tan significativos. Además, un punto a favor de este estudio fue la capacidad del modelo para ser extrapolado y adoptado de forma fácil y segura por otras plantas de tratamiento que tengan distintas condiciones o con procesos de potabilización variados. Con el conjunto de ecuaciones resultantes se plantearon implementar a futuro un software para los operadores de las plantas de tratamiento estudiadas para sustituir la dosificación manual y proporcionar una dosis de coagulante en tiempo real. A diferencia del presente caso, donde los resultados no tienen la asertividad relevante que se espera y los modelos planteados están en función de la situación específica de la PTAP-M.

Para el caso de las elasticidades del agua cruda en base a la turbidez; en primera instancia, Price & Heberling (2018) señalaron que las elasticidades de la turbidez en la literatura oscilan de un -0.11 a un 0.30. Del mismo modo, el estudio de Mondorf (2020) presentó elasticidades cercanas a estos valores en mención, con una elasticidad estadísticamente significativa de 0.389; implicando que una disminución del 1 % en la turbidez del agua conduce a una reducción del 0.389 % en el esfuerzo del tratamiento químico. Para la presente tesis, se obtuvo elasticidades entre 0.97 y 1.87; haciendo alusión que en el peor de los casos una disminución del 1% de la turbidez del agua cruda resulta en una reducción del 0.97 % del esfuerzo del tratamiento químico y en el caso más favorable, una disminución del 1% de la turbidez del agua cruda de ingreso a la planta reduce un 1.87 % del esfuerzo del tratamiento químico. Las respuestas de los distintos casos de estudio en función a las diferentes turbiedades del agua cruda, expresada por el uso de coagulante, no son comparables entre sí.

El rango de variación tan extremo de las elasticidades mostrada para la presente tesis podría deberse a un mal manejo de productos químicos, básicamente un uso excesivo o innecesario de coagulantes como lo ha demostrado los resultados de la relación R_{WT} ; teniendo en cuenta que la reducción de las dosis excesivas de coagulante genera un ahorro sustancial de tiempo y capital a los operadores. Haciendo inferencias sobre el supuesto uso excesivo de productos químicos en la PTAP-M se expone dos consideraciones: primeramente, durante el análisis de la data se observó que existían valores incongruentes de uso de PAC sumamente elevados para días con un valor de turbidez, e incluso precipitación, bajos. Como segunda consideración está el hecho de que los operadores de la PTAP-M preparan los tanques de sulfato de aluminio y PAC previo a la necesidad; según estimación de los propios operadores, un tanque de sulfato de aluminio preparado tiene un tiempo de duración de 13 a 16 horas aproximadamente (fundamentando que este coagulante es el de uso principal), mientras que un tanque de PAC preparado tiene un tiempo de duración aproximado sobre las 24 horas. Por lo que, de forma

errónea se registra estos valores preparados, específicamente de PAC, como dato de consumo de un solo día. En el momento de la transformación del PAC a unidades equivalentes de sulfato de aluminio se consideró distribuir el valor elevado de consumo de PAC al resto de días posteriores, pero esto significaría un error mucho más notable al suponer que para todos estos días se requirió el uso de PAC adicional al sulfato de aluminio.

5. CONCLUSIONES

Con el desarrollo del trabajo de titulación “Impacto de la precipitación pluvial en la turbidez del agua superficial usada para la potabilización en el cantón Azogues durante los años 2014 – 2020” se puede concluir lo siguiente:

Las variables utilizadas para la presente tesis fueron: precipitación pluvial en la estación Llaucay y estación Nudpud, turbidez del agua cruda y esfuerzo del tratamiento químico (cantidad de coagulante). Para el período de estudio comprendido entre enero de 2014 y mayo de 2020, se recopiló 2341 datos diarios de las variables mencionadas, de dicha información únicamente se consideró como válida al 87 %. Respecto a estas variables se evidenció una elevada variabilidad de los datos registrados, teniendo en cuenta la desviación estándar y la media, las que presentaron mayor variabilidad fueron: turbiedad del agua cruda con una desviación estándar de 60.17 y esfuerzo del tratamiento químico expresado como sulfato de aluminio con una desviación de 100.36. Con lo que se puede decir que tanto la turbidez como el esfuerzo del tratamiento químico no se mantienen constantes y presentan extrema variación durante la operación de la PTAP-M. Finalmente, en cuanto a la comprobación de la normalidad de los datos, se determinó que todas las variables analizadas no siguieron una distribución normal, según los respectivos test de normalidad realizados, pero se mantuvo el comportamiento de la normalidad según el fundamento del teorema del límite central.

Para el segundo objetivo, las variables seleccionadas fueron precipitación pluvial y la turbidez del agua cruda las cuales se evaluaron mediante el análisis de regresión lineal, redes neuronales artificiales (RNA) y la metodología de superficie de respuesta. De los resultados se desprende que las relaciones establecidas muestran un grado bajo de correlación entre las variables analizadas. Los coeficientes de determinación (R^2) obtenidos, indican la fuerza de las correlaciones. Para el caso de la calibración; los valores obtenidos variaron desde 0.07 hasta 0.23, el modelo más óptimo para el conjunto de Turbiedad – Nudpud – Llaucay fue el de tipo lineal por presentar los mejores estadísticos: 0.22, 0.32 y 0.76 para el R^2 , RMSE y EF respectivamente. Para el caso de la validación del modelo, el mejor R^2 fue de 0.21 con el modelo exponencial para el mismo conjunto de Turbiedad – Nudpud – Llaucay. El valor del coeficiente de determinación implica que, (considerando la validación) el 21 % de la variable dependiente podría explicar la variable independiente, mientras que el 79 % podría explicarse por otros factores.

Al utilizar las RNA para el análisis de relación entre las variables, en primera instancia se realizó un análisis de sensibilidad de los parámetros óptimos dentro de la red para obtener los mejores resultados, la mejor combinación de parámetros para el modelo entre la turbidez del agua cruda y la precipitación fluvial fue: tiempo de entrenamiento de 20000, velocidad de aprendizaje de 0.1, en tanto, la función de transferencia de la capa oculta y capa de salida fue función de transferencia log sigmoide. En cuanto al número de neuronas en la capa interna, se determinó en base a los errores que los modelos que presentaron los mejores resultados fueron: el Modelo 10-15-5 con un R^2 de calibración y validación de 0.22 y 0.19, un RMSE de 2.06 y 0.9 y EF de 1 y 0.89 respectivamente; el Modelo 2 con un coeficiente de determinación para calibración de 0.22, un RMSE de 2.02 y EF de 0.99 y para el caso de la validación un R^2 de 0.17, un RMSE 0.84 y EF de 0.96. La ecuación final de los modelos considerados como los mejores se presentan en la *Ecuación 10* y en la *Ecuación 11*. Además, se consideró al número de neuronas en la capa interna como el parámetro más sensible dentro de la red neuronal, ya que, al calcular el índice de sensibilidad este parámetro presenta un valor muy superior a los parámetros mencionados anteriormente, el valor obtenido fue de 0.0021.

En cuanto al uso de la metodología de superficie de respuesta para predecir la turbidez con la precipitación pluvial, se obtuvo que el modelo es capaz de predecir en la etapa de calibración el 34 % de los datos, en cambio, para la etapa de validación el cálculo muestra un coeficiente de determinación de 0.19.

Por lo tanto, los resultados de los análisis realizados muestran claramente que no existe una relación significativa entre la turbidez del agua cruda y la precipitación pluvial y a su vez el mejor modelo obtenido en la etapa de validación en base a los errores calculados se puede decir que fue el modelo exponencial con el uso de regresión lineal; es evidente que la diferencia con el resto de los modelos con cada una de las metodologías citadas anteriormente no es significativa. Es por ello, que se recomienda para futuros usos la metodología de superficie de respuesta, ya que mediante esta se obtuvo el valor más alto para la calibración y para la validación varió únicamente en 0.02 del modelo realizado con regresión lineal, además, presenta ventajas considerables al comparar con los otros modelos, las principales son: el tiempo del cálculo y la facilidad del software Minitab.

Para defender el fundamento del tercer objetivo se precisó del análisis previo del esfuerzo del tratamiento químico de la PTAP-M. Se afirma que la cantidad de coagulante utilizada está relacionada con el nivel de turbidez del agua cruda de forma moderada; en tanto, de una forma

menos marcada casi independiente, la relación entre la cantidad de coagulante y las precipitaciones es ajena. Este argumento está reflejado en los coeficientes de determinación obtenidos más altos que indican la fuerza de las relaciones entre las variables; considerando la relación de tipo lineal e indicando una pendiente positiva se tiene para la relación Cantidad de coagulante – Turbidez un R^2 de 0.34 y para la relación Cantidad de coagulante – Precipitación un R^2 de 0.04. Siendo estos dos grados de correlación catalogadas como bajos; valores casi inapreciables. Esta falta de relación y congruencia entre los parámetros seleccionados (esfuerzo del tratamiento químico, turbidez y precipitación) pueden deberse a la interacción e influencia de otros factores externos a los considerados en la presente tesis.

Una vez definida las relaciones entre el esfuerzo del tratamiento químico, turbidez y precipitación se modeló este comportamiento para explicar el vínculo de las tres variables en un único modelo matemático. Se elaboró, una vez más, modelos de regresión múltiple, modelos con redes neuronales artificiales (RNA) y mediante la metodología de superficie de respuesta. En los tres casos, la variable dependiente representa el esfuerzo del tratamiento químico con la cantidad de coagulante requerida en función a la variación de la calidad del agua cruda. La calidad del agua cruda representa a la variable independiente a partir del nivel de turbidez del agua de ingreso, la variable control fue la precipitación de los puntos de captación Llaucay y Nudpud. Con esto aclarado y con los resultados estadísticos de la calidad de los modelos; se deriva que el modelo más confiable y recomendado es el Modelo 3 desarrollado bajo las RNA con un R^2 de validación igual a 0.58. Encasillando este resultado en un grado de correlación entre moderado a aceptable; implicando que el 58 % de la variable dependiente (esfuerzo del tratamiento químico) está y podría ser explicada por las variables independientes (turbidez del agua cruda y precipitación); mientras que el restante 42 % de la variable dependiente podría explicarse por factores ajenos.

En mención a la elasticidad del agua cruda de la PTAP-M o la elasticidad de la turbidez, la evaluación proyectó elasticidades entre 0.97 y 1.87, con una media de 1.004. Este último valor implica que a una reducción del 1 % del valor de la turbidez del agua cruda se reduce el 1.004 % del esfuerzo del tratamiento químico o la cantidad de coagulante requerido para esa turbidez. Estos valores son significativos para la PTAP-M porque se trata de una reducción en la cantidad de productos químicos que deriva directamente en la disminución del gasto económico; sin embargo, la validez de los resultados está limitada al estudio. Las elasticidades que se calculan son válidas para ser consideradas exclusivamente dentro del rango de los datos utilizados para el mismo cálculo. Dicho esto, las elasticidades obtenidas en este estudio son admitidas para el

rango de turbidez de 3 NTU a 805 NTU, además ya se sabe que las elasticidades no son relevantes para cambios extremos de la calidad del agua cruda.

De forma global se concluye que sólo en el momento que se determine de forma fidedigna, cien por ciento segura y con un error mínimo la base adecuada para describir la relación entre la precipitación y la turbidez del agua cruda mediante la conjugación idónea de las variables significativas a través de la modelación, será posible establecer sistemas de alerta temprana para niveles de turbidez altos en las plantas de tratamiento y datos inmediatos para la cantidad de coagulante necesaria a utilizar para el momento determinado. En muchos estudios si se ha logrado establecer esta relación de forma concreta para brindar esta herramienta a las plantas de tratamiento, pero en el caso de la PTAP-M los resultados encontrados no son los esperados ni los sospechados; por lo que se debería explorar más a fondo con una nueva disponibilidad del conjunto de datos más actuales y fiables y, no está por demás, considerar otros factores externos que se conoce que influyen significativamente en la conducta estudiada.

6. RECOMENDACIONES

Finalizando el trabajo las recomendaciones pertinentes al estudio son las siguientes:

- Aludiendo a EMAPAL como responsable de la PTAP-M, contar con un mantenimiento preventivo y correctivo de los equipos y aparatos de las estaciones meteorológicas, con el fin de reducir los inconvenientes y posibles errores en la información obtenida. De igual forma asegurar y mantener calibrados los equipos de las estaciones meteorológicas.
- Realizar un mantenimiento conveniente en todo el tramo de las líneas de conducción del agua cruda para evitar la influencia de factores externos en la calidad del agua cruda.
- Realizar estudios detallados de la influencia del reservorio, el cual se encuentra a 1 Km antes del ingreso del agua cruda a la PTAP-M, ya que este puede estar desempeñando el papel de un sedimentador, afectando de manera considerable el valor de la turbidez a la entrada de la planta.
- Adicional al punto anterior se recomienda medir la turbidez del agua cruda en el punto de ingreso al reservorio y medir nuevamente la turbidez con la que sale el agua cruda hacia la planta.
- Medir la turbidez del agua cruda en el punto más próximo a la entrada a la planta de tratamiento de agua potable y comparar con los datos medidos en la salida del agua cruda del reservorio, recomendación dirigida directamente para la PTAP-M.
- Se recomienda para estudios futuros incluir variables significativas que puedan influenciar en la turbidez para el análisis, como, por ejemplo: cobertura vegetal de la zona aledaña a las captaciones, erodabilidad o erosión del suelo, considerando que: cuanto mayor sea la erodabilidad mayor porcentaje de erosión se presentará en el suelo, intensidad de la lluvia, uso de suelo, entre otras.

7. BIBLIOGRAFÍA

- Aragón. (2015). *Estadística en el área de las Ciencias Sociales y Administrativas*.
- Arboleda, J. (1992). *Teoría y práctica de la purificación del agua*. 1–72.
<https://cidta.usal.es/cursos/etap/modulos/libros/teoria.pdf>
- ASCE. (2000). Artificial neural networks in hidrology. I: preliminary concepts. *Taiyangneng Xuebao/Acta Energiae Solaris Sinica*, 35(2), 9.
- Baouab, M. H., & Cherif, S. (2018). Prediction of the optimal dose of coagulant for various potable water treatment processes through artificial neural network. *Journal of Hydroinformatics*, 20(6), 1215–1226. <https://doi.org/10.2166/hydro.2018.014>
- Barrenechea, A. (2015). *Aspectos fisicoquímicos de la calidad del agua*.
<http://www.ingenieroambiental.com/4014/uno.pdf>
- Basogain, X. (2015). *Redes neuronales artificiales y sus aplicaciones*.
- Bautista-Díaz, M. L., Victoria-Rodríguez, E., Vargas-Estrella, L. B., & Hernández-Chamosa, C. C. (2020). Pruebas estadísticas paramétricas y no paramétricas: su clasificación, objetivos y características. *Educación y Salud Boletín Científico Instituto de Ciencias de La Salud Universidad Autónoma Del Estado de Hidalgo*, 9(17), 78–81.
<https://doi.org/10.29057/icsa.v9i17.6293>
- Cárdenas, J. D. L., & Patiño, G. F. E. (2010). *Universidad de Cuenca Facultad de Ingeniería Escuela de Ingeniería Civil*. 206.
<https://dspace.ucuenca.edu.ec/bitstream/123456789/725/1/ti853.pdf>
- Cardona, D. (2011). *Caracterización del agua cruda del río la vieja como fuente superficial para el proceso de potabilización de Emcartago S.A. E.S.P. July*, 1–7.
- Carreño, R., Montenegro, K., & Mosquera, J. (2019). Guía De Agua Segura. *Ministerio de Salud Pública*, 72. <https://www.salud.gob.ec/wp-content/uploads/2020/12/Guia-Agua-Segura.pdf>
- Chang, C.-L., & Liao, C.-S. (2012). Parameter Sensitivity Analysis of Artificial Neural Network for Predicting Water Turbidity. *World Academy of Science, Engineering and Technology International Journal of Geological and Environmental Engineering*, 6(10), 657–660. <https://publications.waset.org/5674/parameter-sensitivity-analysis-of-artificial->

neural-network-for-predicting-water-turbidity

- Crittenden, J. C. T. R. R. H. D. H. K. T. G. (2012). *MWH's Water Treatment* (3rd ed). John Wiley & Sons, Inc.
- Duarte, O., Díaz, E., & Ricciardi, C. (1991). *Pluviometría*. 1–15. http://www.fca.uner.edu.ar/files/academica/deptos/catedras/riego/Archivos/Cap_03_-_Pluviometria.pdf
- Fernández, A. (2012). *El agua: un recurso esencial*. <https://www.redalyc.org/pdf/863/86325090002.pdf>
- GAD de Azogues. (2018). *Plan de Desarrollo y Ordenamiento Territorial Azogues*. 383.
- Gebbie, P. (2001). 64 th Annual Water Industry Engineers and Operators Conference All Seasons International Hotel-Bendigo. In *Conference All Seasons International Hotel-Bendigo* (Vol. 5).
- Gorchev, H. G., & Ozolins, G. (1984). WHO Guidelines for drinking- water quality. *WHO Chronicle*, 38(3), 104–108.
- Idrovo, C. (2010). Optimización de la planta de tratamiento de Uchupucun Universidad de Cuenca Ecuador. *Universidad de Cuenca Ecuador*, 1–244. <https://dspace.ucuenca.edu.ec/bitstream/123456789/2426/1/tq1066.pdf>
- INEN 1108. (2020). Agua Potable. Requisitos. Nte Inen 1108. *Instituto Ecuatoriano de Normalización*, 5, 1–10. <http://normaspdf.inen.gob.ec/pdf/nte/1108-5.pdf>
- Isasi, P., & Galván, I. (2004). *Redes de neuronas artificiales un enfoque práctico*.
- Jiménez, B., Valiente, E., Ponce, G., Villanueva, S., Botello, A., López, B., Herrera, G., Carrillo, J., & Cardona, A. (2010). *Calidad del agua: un enfoque multidisciplinario*. www.clacso.edu.ar
- Khuri, A. I., & Mukhopadhyay, S. (2010). Response surface methodology. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(2), 128–149. <https://doi.org/10.1002/wics.73>
- Larrañaga, P., Inza, I., & Moujahid, A. (2012). *Redes Neuronales*.
- Martos, Á. (2015). *La importancia del agua en nuestro planeta*. https://tauja.ujaen.es/bitstream/10953.1/2374/1/MARTOS_LÓPEZ%2C_ÁNGEL.pdf

- Matich, D. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*.
- Ministerio de Salud, & Control de Calidad del Agua. (2013). *Manual Práctico de Análisis de Agua*.
https://bvsms.saude.gov.br/bvs/publicacoes/manual_practico_analisis_agua_4_ed.pdf
- Ministerio del Ambiente del Ecuador. (2011). Norma de Calidad Ambiental y de descarga de efluentes : Recurso Agua. *TULAS Texto Unificado de Legislación Secundaria Del Ministerio Del Ambiente*, 8–9.
- Mondorf, D. (2020). *The Water-Energy-Food Security Nexus in the Kenyan-Ugandan border region Impact of surface water quality deterioration through agricultural practices on drinking water*.
- Montgomery, D. C., & Runger, G. C. (1994). Applied Statistics and Probability for Engineers. In *European Journal of Engineering Education* (Vol. 19, Issue 3).
<https://doi.org/10.1080/03043799408928333>
- Montoya, C., Loaiza, D., Torres, P., Cruz, C., & Escobar, J. (2011). *Efecto del incremento en la Turbiedad del agua cruda sobre la eficiencia de procesos convencionales de potabilización*.
- Morales, E., Solano, M., Morales, R., Reyes, L., Barrantes, K., Achí, R., & Chacón, L. (2019). Evaluation of climatic seasonal influence on the quality of water for human consumption in a San José (Costa Rica) supply system, 2017-2018. *Rev Costarr Salud Pública*, 28(1), 77–87.
- Morales, P. (2011). *Correlación y regresión, simple y múltiple*.
<http://www.upcomillas.es/personal/peter/investigacion/Regresion.pdf>
- Municipio de Azogues, EMAPAL EP, CG Paute, & PROMAS. (2003). *Plan de manejo de la microcuenca del río Tabacay*.
- Myers, R. H., Khuri, A., & Carter, W. H. (1989). Response surface methodology: 1966-1988. *Technometrics*, 31(2), 137–157. <https://doi.org/10.1080/00401706.1989.10488509>
- Naciones Unidas. (2021). Informe de las Naciones Unidas sobre el Desarrollo de los Recursos Hídricos en el Mundo 2021: El valor del agua. In *UNESCO*. UNESCO.
- Novoa, G., & Guillén, B. (2018). Desarrollo de una estación agro-meteorológica automática

- remota para el levantamiento de información climática en la cuenca del río Pisque. In *Tesis* (Vol. 1). <http://dspace.ups.edu.ec/bitstream/123456789/5081/1/UPS-CYT00109.pdf>
- Okun, D., & Schulz, C. (1990). *Tratamiento de aguas superficiales para países en desarrollo* (3rd ed., Vol. 2). Limusa México.
- OMS. (2019). Guías para la calidad del agua de consumo humano. *Organización Mundial de La Salud*, 4, 608. <https://apps.who.int/iris/bitstream/handle/10665/272403/9789243549958-spa.pdf?sequence=1&isAllowed=y>
- Orellana, J. (2016). Tratamiento De Las Aguas. *Ingeniería Sanitaria- UTN - FRRO*, 1–123.
- Pardo Merino, A., & Ruiz Díaz, M. Á. (2005). *Análisis de datos con SPSS 13 Base*.
- Peña, A. (2016). *Uso de redes neuronales artificiales para optimizar la dosificación de coagulantes en la Planta de Tratamiento de Agua Potable - Huancayo*.
- Peña Rojas, A., & Flores del Pino, L. (2014). Redes neuronales para el tratamiento de agua potable en zona de altitud del Perú. *Ambiente y Desarrollo*, 18(35), 108. <https://doi.org/10.11144/javeriana.ayd18-35.rnta>
- Price, J., & Heberling, M. (2018). The Effects of Source Water Quality on Drinking Water Treatment Costs: A Review and Synthesis of Empirical Literature. *Ecological Economics*, 151(November 2017), 195–209. <https://doi.org/10.1016/j.ecolecon.2018.04.014>
- Quispe, R. (2015). Evaluación de la contaminación por metales pesados en cuerpos de agua dejados por la minería aurífera, en la comunidad nativa de Tres Islas del departamento de Madre de Dios. *Repositorio de La Universidad Nacional Amazonica de Madre de Dios*, Noviembre, 128.
- Rock, C., & Rivera, B. (2014). *La calidad del agua, E. coli y su salud*. <https://extension.arizona.edu/sites/extension.arizona.edu/files/pubs/az1624s.pdf>
- Rodríguez Jaume, M. J., & Mora Catalá, R. (2001). *Estadística Informática: casos y ejemplos con el SPSS*.
- Rojo, J. (2008). Regresión lineal múltiple. *VDI-Z Integrierte Produktion*, 150(10), 22–23.
- Rolando Mora, Gamboa, J. C., & Fernández, M. V. (2006). Zonificación de la susceptibilidad

al deslizamiento: Resultados obtenidos para la península de papagayo mediante la modificación del Método Mora-Vahrson. In *III Curso Internacional sobre microzonificación y su Aplicación en la mitigación de desastres* (pp. 38–43).

Romero, J. (2009). *Calidad del Agua* (tercera).
https://www.academia.edu/38972813/Calidad_del_Agua_ROMERO_ROJAS

Romero, Z. (2015). “*Atlas del Cantón Azogues.*” 73.

Salazar, L. (2020). *Tratamiento de agua para purificación. Teoría y principios de diseño.*
<https://es.scribd.com/document/456314499/Libro-Tratamiento-de-Agua-v19>

SAMSA. (2008). *Proceso de potabilización del agua.* Stationery Office.
[http://www.elaguapotable.com/Proceso potabilización\(Sansa\).pdf](http://www.elaguapotable.com/Proceso%20potabilizaci3n(Sansa).pdf)

Sari, V., dos Reis Castro, N. M., & Pedrollo, O. C. (2017). Estimate of Suspended Sediment Concentration from Monitored Data of Turbidity and Water Level Using Artificial Neural Networks. *Water Resources Management*, 31(15), 4909–4923.
<https://doi.org/10.1007/s11269-017-1785-4>

Secretaría de Economía. (2001). *Análisis de agua-determinación de cloruros totales en aguas naturales, residuales y residuales tratadas-método de prueba.*
<https://www.gob.mx/cms/uploads/attachment/file/166789/NMX-AA-073-SCFI-2001.pdf>

Segerer, C., & Villodas, R. (2013). Las Precipitaciones. *Hidrología I*, 2. <http://www.conosur-rirh.net/ADVF/documentos/hidro1.pdf>

Sierra, C. (2011). *Calidad del agua: Evaluación y diagnóstico* (L. López (ed.); Vol. 2).
[https://www.academia.edu/9511155/Calidad_del_agua_evaluación_y_diagnóstico](https://www.academia.edu/9511155/Calidad_del_agua_evaluaci3n_y_diagn3stico)

Silicatos y Derivados S.A. (2011). Sulfato de aluminio. *Sidesa*, 52 55, 1–12.
[http://www.aniq.org.mx/pqta/pdf/Respaldo/Sulfato de Al.Sol. Libre de Fe\(toda LIT\).pdf](http://www.aniq.org.mx/pqta/pdf/Respaldo/Sulfato%20de%20Al.Sol.Libre%20de%20Fe(toda%20LIT).pdf)

Tapia, J. D. (2016). Estructura de la precipitación y su variabilidad espacio-temporal a lo largo de una gradiente altitudinal. *Universidad de Cuenca*, 1–71.
<http://dspace.ucuenca.edu.ec/bitstream/123456789/24348/1/tesis.pdf?fbclid=IwAR2CUCJ-QUXrAVvTXPhx5Jbfi1JEI0VDqDRfCjV14sZ8aQj0EfGJPOGFR1s>

UNESCO, & ONU-Agua. (2020). *Informe Mundial de las Naciones Unidas sobre el*

Desarrollo de los Recursos Hídricos 2020: Agua y Cambio Climático. UNESCO.
www.unesco.org/open-access/

Vargas, S. (2011). Estudio Y Diseño De La Captación, Conducción, Planta De Tratamiento Y Distribución Del Sistema De Agua Potable De La Comunidad De Ambatillo Alto En La Parroquia De Ambatillo, Provincia De Tungurahua, Para Su Posterior Construcción. *Repo.Uta.Edu.Ec*, 1, 130.
<http://repo.uta.edu.ec/bitstream/handle/123456789/5301/Mg.DCEv.Ed.1859.pdf?sequence=3>

WHO. (2017). Water quality and Health - review of turbidity: Information for regulators and water suppliers. *Who/Fwc/Wsh/17.01*, 10.
https://www.who.int/water_sanitation_health/publications/turbidity-information-200217.pdf
http://www.who.int/water_sanitation_health/publications/turbidity-information-200217.pdf

ANEXOS

Anexo A: Recorrido in situ de los puntos de captación, líneas de conducción y reservorio dentro del área de estudio de la microcuenca del río Tabacay



Anexo 1 Quebrada Nudpud, microcuenca del río Tabacay



Anexo 2 Sistema de captación de agua cruda de la quebrada Nudpud, microcuenca del río Tabacay

UCUENCA



Anexo 3 Zona aledaña al sistema de captación de agua cruda de la quebrada Nudpud, microcuenca del río Tabacay



Anexo 4 Quebrada Llaucay, microcuenca del río Tabacay



Anexo 5 Sistema de captación de agua cruda de la quebrada Llaucay, microcuenca del río Tabacay



Anexo 6 Zona aledaña al sistema de captación de agua cruda de la quebrada Llaucay, microcuenca del río Tabacay



Anexo 7 Tramos de la línea de conducción a cielo abierto del agua cruda proveniente de la quebrada Nudpud, microcuenca del río Tabacay



Anexo 8 Tramo de la línea de conducción cubierta del agua cruda proveniente de las quebradas Nudpud y Llaucay, microcuenca del río Tabacay



Anexo 9 Tramo de la línea de conducción cubierta con maleza, hojarasca y restos de flora deforestada, microcuenca del río Tabacay



Anexo 10 Reservoirio de agua cruda en la PTAP-M (desembocadura de la línea de conducción de agua cruda de Nudpud y Llaucay). Ingreso del agua cruda de la línea de conducción al reservorio (círculo amarillo)



Anexo 11 Continuación del reservorio de agua cruda en la PTAP-M



Anexo 12 Sección final del reservorio de agua cruda en la PTAP-M. Ingreso del agua cruda del reservorio a la tubería (1 km) para ingresar a la PTAP-M

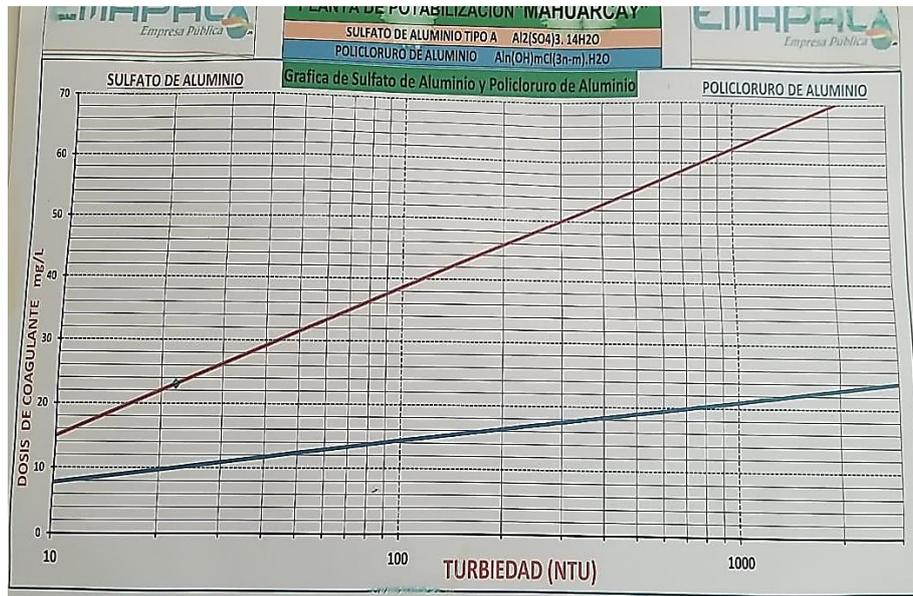


Anexo 13 Ingreso del agua cruda de la tubería (1 km) a la PTAP-M. Punto de medición del nivel de turbidez del agua cruda



Anexo 14 Inicio del proceso de potabilización del agua cruda en la PTAP-M

Anexo B: Curva de dosificación del sulfato de aluminio y policloruro de aluminio de la PTAP-M



Anexo 15 Gráfica de sulfato de aluminio y PAC, curva de dosificación de la PTAP-M