



UNIVERSIDAD TÉCNICA DEL NORTE

**FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
CARRERA DE INGENIERÍA INDUSTRIAL**

**TRABAJO DE GRADO PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO INDUSTRIAL**

**TEMA:
MODELO PARA LA PREDICCIÓN Y EVALUACIÓN DE LA
DEMANDA DE AGUA POTABLE DE EMAPALA EP MEDIANTE
REDES NEURONALES ARTIFICIALES.**

AUTOR: GRANIZO ROJAS RONNIE JOSÉ

DIRECTOR: ING. YAKCLEEM MONTERO SANTOS MSC.

IBARRA – ECUADOR

2020

UNIVERSIDAD TÉCNICA DEL NORTE

BIBLIOTECA UNIVERSITARIA

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE

IDENTIFICACIÓN DE LA OBRA.

En cumplimiento del Art. 144 de la Ley de Educación Superior, hago la entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
CÉDULA DE IDENTIDAD:	2100772983		
APELLIDOS Y NOMBRES:	GRANIZO ROJAS RONNIE JOSÉ		
DIRECCIÓN:	AV. QUITO KM 39 CASCALES		
EMAIL:	rjgranizor@utn.edu.ec		
TELÉFONO FIJO:		TELÉFONO MÓVIL:	0979712171
DATOS DE LA OBRA			
TÍTULO:	MODELO PARA LA PREDICCIÓN Y EVALUACIÓN DE LA DEMANDA DE AGUA POTABLE DE EMAPALA EP MEDIANTE REDES NEURONALES ARTIFICIALES.		
AUTOR (ES):	GRANIZO ROJAS RONNIE JOSÉ		
FECHA: AAAAMMDD			
SÓLO PARA TRABAJOS DE GRADO			
PROGRAMA:	<input checked="" type="checkbox"/> PREGRADO <input type="checkbox"/> POSGRADO		
TITULO POR EL QUE OPTA:	INGENIERO INDUSTRIAL		
ASESOR/DIRECTOR:	ING. YAKCLEEM MONTERO SANTOS MSC.		

UNIVERSIDAD TÉCNICA DEL NORTE
FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
CARRERA DE INGENIERÍA INDUSTRIAL

CONSTANCIAS

El autor (es) manifiesta (n) que la obra de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es (son) el titular (es) de los derechos patrimoniales, por lo que asume (n) la responsabilidad sobre el contenido de la misma y saldrá (n) en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 23 días del mes de septiembre de 2020



Nombre: Granizo Rojas Ronnie José
Cédula: 210077298-3

**UNIVERSIDAD TÉCNICA DEL NORTE
FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
CARRERA DE INGENIERÍA INDUSTRIAL**

CERTIFICACIÓN

Ing. Montero Santos Yackleem MSc., director del Trabajo de Grado desarrollado por el señor estudiante Granizo Rojas Ronnie José.

CERTIFICA

Que, el proyecto de Trabajo de Grado Titulado "MODELO PARA LA PREDICCIÓN Y EVALUACIÓN DE LA DEMANDA DE AGUA POTABLE DE EMAPALA EP MEDIANTE REDES NEURONALES ARTIFICIALES". ha sido elaborado en su totalidad por el señor estudiante Granizo Rojas Ronnie José bajo mi dirección para la obtención del título de Ingeniero Industrial. Luego de ser revisado, considero que se encuentra concluido y cumple con las exigencias y requisitos académicos de la Facultad de Ingeniería en Ciencias Aplicadas, Carrera de Ingeniería Industrial, autorizo su presentación y defensa para que pueda ser juzgado por el tribunal correspondiente.



.....
Ing. Yackleem Montero Santos MSc.

DIRECTOR DE TESIS

**UNIVERSIDAD TÉCNICA DEL NORTE FACULTAD DE INGENIERÍA EN
CIENCIAS APLICADAS CARRERA DE INGENIERÍA INDUSTRIAL**

DEDICATORIA

A mis padres José Luis Granizo e Irma Rojas, quienes son el motor de mi vida, han estado conmigo en las buenas y malas, los que me han apoyado incondicionalmente en cada una de mis etapas estudiantiles y personales, los que me han enseñado que con esfuerzo, dedicación y respeto se puede llegar muy lejos, son mi ejemplo para seguir superándome cada día.

A mi hermano Joel Granizo, por darme ánimos para seguir adelante, a mis tíos Jhofre y Francisco quienes me han sabido brindar apoyo moralmente, los consejos que me otorgaban he sabido recibirlos de buena manera y ponerlos en práctica, muchas gracias.

Y a mi familia y amigos quienes también me han sabido brindar su apoyo y se les agradece de todo corazón.

RONNIE JOSÉ GRANIZO ROJAS

**UNIVERSIDAD TÉCNICA DEL NORTE FACULTAD DE INGENIERÍA EN
CIENCIAS APLICADAS CARRERA DE INGENIERÍA INDUSTRIAL**

AGRADECIMIENTO

Primeramente, agradezco a Dios por obsequiarme un día más de vida, salud, educación, una familia, por cuidarme y guiar mi camino para poder realizar de buena manera mis acciones. Agradezco a mis padres, a mi familia y amigos quienes han estado conmigo en trascurso de mi carrera y de mi vida personal y que han aportado cada uno con su granito de arena para poder forjar la persona que soy hoy en día.

A la empresa EMAPALA EP, por abrirme las puertas, brindándome su confianza y conocimientos en el ámbito de agua potable y de esta manera realizar con éxitos el presente trabajo de grado.

A la Universidad Técnica del Norte, y en especial a la carrera de Ingeniería Industrial quien juntamente con sus docentes, me supieron brindar sus conocimientos y experiencias para poder triunfar en la vida profesional y personal.

A mis compañeros y amigos Daniel y Roberth quienes me han apoyado incondicionalmente y han estado en los buenos y malos momentos a lo largo de mi vida universitaria y personal.

A los ingenieros Yakcleem Montero Santos MSc., y Leandro Lorente MSc., quienes me apoyaron y guiaron con las tutorías académicas para poder generar un trabajo de titulación de calidad.

RONNIE JOSÉ GRANIZO ROJAS

INDICE

CONSTANCIAS	III
CERTIFICACIÓN.....	IV
DEDICATORIA.....	V
AGRADECIMIENTO.....	VI
CAPÍTULO I.....	1
1. GENERALIDADES	1
1.1. Problema.....	1
1.2. Objetivos	2
1.2.1. Objetivo General.....	2
1.2.2. Objetivos Específicos	2
1.3. Justificación.....	2
1.4. Metodología	3
1.4.1. Métodos de Pronóstico de Demanda	4
1.4.3. Método Deductivo	5
1.4.4. Método Descriptivo	5
1.4.5. Metodología para la Aplicación de Redes Neuronales Artificiales.....	5
1.5. Alcance.....	6
CAPÍTULO II.....	7
2. FUNDAMENTACIÓN TEÓRICA	7
2.1. Normativa Legal.....	7
2.2. Producción de Agua Potable	8
2.2.1. Red de Abastecimiento o Distribución	9

2.2.2.	Captación	10
2.2.3.	Almacenamiento	10
2.2.4.	Transformación.....	10
2.2.5.	Almacenamiento de agua potable.....	10
2.2.6.	Distribución	11
2.2.7.	Demanda de Agua Potable	11
2.3.	Pronóstico.....	11
2.3.1.	Aplicaciones de los pronósticos	13
2.3.2.	Clasificación de los métodos de los pronósticos	14
2.3.3.	Métodos cualitativos.....	14
2.3.3.1.	Consenso de un panel.....	14
2.3.3.2.	Analogía histórica	14
2.3.3.3.	Investigación de mercados.....	15
2.3.4.	Métodos cuantitativos.....	15
2.3.4.1.	Métodos de pronósticos para serie de tiempo	15
2.3.4.2.	Modelos causales	16
2.3.5.	Selección de un método de pronóstico	16
2.3.6.	Series de tiempo.....	17
2.3.7.	Mediciones del error de un pronóstico	19
2.4.	Machine Learning	22
2.4.1.	Regresión Lineal Múltiple	23
2.4.2.	Support Vector Machine.....	23
2.5.	Redes Neuronales Artificiales.....	24
2.5.1.	Ventajas de las redes neuronales	25

2.5.2.	Aprendizaje adaptativo	25
2.5.3.	Auto-organización	25
2.5.4.	Tolerancia a fallos	25
2.5.5.	Operación en tiempo real.....	26
2.5.6.	Fácil inserción dentro de la tecnología existente.....	26
2.5.7.	Característica de las Redes Neuronales Artificiales	26
2.5.8.	Aplicaciones de las Redes Neuronales Artificiales	27
2.5.9.	Estructura básica de una red neuronal	28
2.5.9.1.	Analogía con el cerebro	28
2.5.9.2.	Diagrama de una neurona artificial PE	28
2.6.	Modelos neuronales.....	29
2.7.	Taxonomía de modelos de las redes neuronales	31
2.8.	Aprendizaje o Entrenamiento.....	32
2.8.1.	Aprendizaje supervisado.....	32
2.8.2.	Aprendizaje reforzado	33
2.8.3.	Aprendizaje no supervisado.....	33
2.9.	Arquitectura de una red neuronal	34
2.10.	Algoritmo de retropropagación (BP).....	38
2.11.	R Studio	38
CAPÍTULO III		40
3. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL.....		40
3.1.	Descripción de la empresa.....	40
3.2.	Localización de la empresa	40
3.3.	Misión.....	41

3.4. Visión	41
3.5. Objetivos Estratégicos	41
3.6. Líneas Estratégicas	42
3.7. Organigrama	43
3.8. Descripción del Proceso Productivo	44
3.9. Producción de la empresa y consumo anual.....	45
3.10. Datos de la población	47
3.11. Volumen de almacenamiento	48
3.12. Plantas de tratamiento.....	48
3.13. Análisis de autocorrelación de los sistemas de agua potable	49
3.13.1. Análisis de autocorrelación de la producción mensual de agua potable	49
3.13.2. Análisis de autocorrelación de la distribución de agua potable.....	50
3.13.3. Análisis de autocorrelación de los clientes mensuales	55
3.13.4. Análisis de la serie de tiempo de la demanda de agua potable.....	56
3.13.5. Análisis de estacionalidad de las series de tiempo	57
CAPÍTULO IV	63
4. DESARROLLO DEL MODELO DE PRONÓSTICO.....	63
4.1. Método para la elaboración del pronóstico	63
4.2. Identificación de la variable	64
4.3. Recolección de datos	64
4.4. Estructuración del modelo.....	66
4.4.1. Análisis de la serie temporal.....	66
4.4.2. Análisis de tendencia y estacionalidad	66
4.5. Pronóstico de la red neuronal óptima.....	67

4.6. Estructura de la red MLP	69
4.7. Pronóstico con la mejora del entrenamiento de la red	71
CAPÍTULO V	73
5. ANÁLISIS DE LOS RESULTADOS OBTENIDOS.....	73
5.1. Comportamiento de la demanda.....	73
5.2. Comparación de modelos	74

INDICE DE FIGURAS

Figura 2.1 Sistema de abastecimiento de agua	9
Figura 2.2 Serie de tiempo con tendencia	18
Figura 2.3 Serie de tiempo estacionaria	18
Figura 2.4 Serie de tiempo estacional.....	19
Figura 2.5 Componentes de una neurona	28
Figura 2.6 Diagrama de una neurona artificial PE	29
Figura 2.7 Modelo de las unidades de las neuronas o procesamiento.....	31
Figura 2.8 Taxonomías	32
Figura 2.9 Aprendizaje supervisado	33
Figura 2.10 Aprendizaje reforzado.....	33
Figura 2.11 Aprendizaje supervisado	34
Figura 2.12 Arquitectura de una red neuronal.....	35
Figura 2.13 Redes Feed-forward	36
Figura 2.14 Redes recurrentes	36
Figura 2.15 Redes LSTM	37
Figura 2.16 Redes GRU	38
Figura 2.17 Software R Studio	39
Figura 3.1 Localización de la empresa EMAPALA EP	41
Figura 3.2 Estructura Orgánica de EMAPALA	43
Figura 3.3 Autocorrelación de la Producción de agua potable.....	50
Figura 3.4 Autocorrelación de la distribución de la Red Principal de agua potable	50
Figura 3.5 Autocorrelación de la distribución del Sistema Aeropuerto	51
Figura 3.6 Autocorrelación de la distribución del Sistema San Pedro	51
Figura 3.7 Autocorrelación de la distribución del Sistema Venezuela	52

Figura 3.8 Autocorrelación de la distribución del Sistema de los Tanques Elevados.....	52
Figura 3.9 Prueba Dickey Fuller en el Nivel del Sistema de los Tanqueros	54
Figura 3.10 Prueba Dickey Fuller en la primera Diferencia del Sistema de los Tanqueros	55
Figura 3.11 Autocorrelación de la distribución del Sistema de los Tanqueros	55
Figura 3.12 Autocorrelación de los clientes mensuales de agua potable	56
Figura 3.13 Autocorrelación de la demanda de agua potable	56
Figura 3.14 Análisis de estacionalidad de la producción mensual 2017-2019.....	57
Figura 3.15 Análisis de estacionalidad de la distribución de la Red Principal.....	58
Figura 3.16 Análisis de estacionalidad de la distribución Sistema Aeropuerto	58
Figura 3.17 Análisis de estacionalidad de la distribución Sistema San Pedro	59
Figura 3.18 Análisis de estacionalidad de la distribución Sistema Venezuela.....	59
Figura 3.19 Análisis de estacionalidad de la distribución Tanqueros	60
Figura 3.20 Análisis de estacionalidad de la distribución Tanques Elevados.....	60
Figura 3.21 Análisis de estacionalidad de los clientes mensuales 2017-2019	61
Figura 3.22 Análisis de estacionalidad de la Demanda de agua potable.....	62
Figura 4.1 Método para la elaboración del pronóstico	64
Figura 4.2 Análisis de la serie temporal de la demanda de agua potable	66
Figura 4.3 Análisis de tendencia	67
Figura 4.4 Análisis de estacionalidad	67
Figura 4.5 Codificación para el pronóstico de la red neuronal óptima	68
Figura 4.6 Pronóstico de la red MLP.....	69
Figura 4.7 Estructura de las capas de la red	70
Figura 4.8 MSE de la red neuronal óptima.....	70
Figura 4.9 Configuración en la mejora de entrenamiento de la red	71

Figura 4.10 Mejora del entrenamiento de la red.....	71
Figura 4.11 Pronóstico con la mejora del entrenamiento de la red MLP	72
Figura 5.1 Comparación de la Demanda de agua potable	74
Figura 5.2 Comparación entre modelos.....	75

INDICE DE TABLAS

Tabla 1.1 Método General de Pronóstico	4
Tabla 2.1 Método General de Diseño de un Pronóstico	12
Tabla 3.1 Producción de agua potable 2017-2018	46
Tabla 3.1 Producción de agua potable 2019	47
Tabla 3.3 Proyección poblacional del cantón Lago Agrio 2017-2020	48
Tabla 4.1 Consumo mensual 2017-2019	65
Tabla 4.2 Clientes mensuales 2017-2019	66
Tabla 4.3 Pronóstico de la demanda de agua potable 2020-2021	72
Tabla 5.1 Comparativa de la Demanda de agua potable	74
Tabla 5.2 Comparación entre modelos	76

ANEXOS

Anexo 1. Codificación de la serie temporal de la demanda de agua potable	83
Anexo 2. Entrenamiento de la Red Neuronal Artificial	83
Anexo 3. Proyección de SPSS.....	85
Anexo 4. Proyección de Crystal Ball	85
Anexo 5. Pronóstico de la producción de agua potable 2020-2021	86
Anexo 6. Sectorización de captación, tratamiento y distribución de agua potable	86

RESUMEN

En la presente investigación hace referencia a la aplicación de redes neuronales artificiales en el pronóstico de la demanda de agua potable de la empresa EMAPALA EP de la ciudad de Lago Agrio, juntamente con la utilización del software de programación R Studio. Cabe recalcar que, para una correcta toma de decisiones de cualquier empresa, los pronósticos son parte fundamental en la planificación, ejecución y control de las actividades que se estén desarrollando. También se debe tomar en cuenta el tiempo que se va a pronosticar la variable, para lo cual se debe identificar los datos históricos y su clasificación ya que pueden estar en meses o años, es importante que la información esté en una sola unidad de tiempo ya que estos datos serán los valores de entrada del modelo. En la ejecución de la investigación se pudo obtener un Error Cuadrático Medio de 0.0048, representando un alto nivel de confiabilidad, es decir los valores pronosticados tienen un comportamiento similar a los reales. Para fortalecer la hipótesis de que el pronóstico realizado con redes neuronales es mejor, se comparó con otros métodos estadísticos tradicionales y los resultados obtenidos corroboran la suposición presentada con anterioridad. Los pronósticos llevan a la empresa a visualizar como se comportará la demanda de su servicio en un futuro, de esta manera la información obtenida de las predicciones aportará al mejoramiento de la producción, a la continua distribución de agua potable y con ello la satisfacción por parte de los clientes.

ABSTRACT

In this research, he refers to the application of artificial neural networks in the forecast of the demand for drinking water of the company EMAPALA EP of the city of Lago Agrio, together with the use of the R Studio programming software. It should be noted that, for the correct decision-making of any company, forecasts are a fundamental part in the planning, execution and control of the activities that are being developed. The time that the variable is going to forecast must also be taken into account, for which the historical data and its classification must be identified since they can be in months or years, it is important that the information is in a single unit of time already that these data will be the input values of the model. In the execution of the investigation, it was possible to obtain a Mean Square Error of 0.0048, representing a high level of reliability, that is, the predicted values have a similar behavior to the real ones. To strengthen the hypothesis that the forecast made with neural networks is better, it was compared with other traditional statistical methods and the results obtained corroborate the assumption presented previously. The forecasts lead the company to visualize how the demand for its service will behave in the future, in this way the information obtained from the forecasts will contribute to the improvement of production, the continuous distribution of drinking water and therefore satisfaction on the part of customers.

CAPÍTULO I

1. GENERALIDADES

1.1.Problema

La Empresa Municipal de Agua Potable y Alcantarillado está dedicada a la dotación, implementación, prestación, operación y mantenimiento de los servicios de agua potable, saneamiento y alcantarillado del cantón Lago Agrio. EMAPALA EP cuenta con un sistema de red de recurso hídrico competente en la zona urbana, no obstante, zonas rurales donde su población cada vez se incrementa tienen la necesidad de estos servicios básicos, esenciales para el desarrollo y bienestar de la ciudadanía.

Según la proyección demográfica 2010-2020 (INEC, 2010). La población del Cantón Lago Agrio para el último año proyectado será de 119,594 habitantes, con una tasa de crecimiento de 2,18%. Debido al aumento anual de la población que presenta el cantón y para garantizar el bienestar de la comunidad Lagoagrensé, se ha visto la necesidad y la importancia de la realización de un pronóstico de la demanda de este servicio básico.

Uno de los principales enfoques de los pronósticos realizados por las empresas de agua potable, es hacer coincidir la cantidad de suministro hídrico que la empresa ofrece con la cantidad demandada por la población, garantizando la proporción un servicio de calidad a los consumidores. La capacidad de respuesta ante las necesidades de consumo de forma eficiente se ha visto de forma inconsistente, debido a ello la flexibilidad y adaptabilidad de la empresa es de gran importancia, sobre todo en zonas rurales donde el crecimiento poblacional se incrementa e influye en la distribución del servicio de agua potable. Viéndose en la necesidad de incluirse en la planificación de distribución del servicio hídrico para un desarrollo controlado y eficiente (Caiado, 2009), (Babel, Gupta, & Pradhan, 2007).

Sin una correcta planificación de la demanda de Agua Potable y Alcantarillado, la administración y distribución de los recursos sería descompensada, no se priorizaría los lugares con mayor necesidad y por consiguiente la insatisfacción de los clientes sería mayor. Hay que tener en cuenta los controles de este servicio, debido a que en la red de distribución problemas como: fugas, taponamientos e instalaciones clandestinas imposibilitan la llegada de agua potable a todos los sectores, generando inconsistencias en la planificación del recurso.

1.2. Objetivos

1.2.1. Objetivo General

Diseñar un modelo para el pronóstico de la demanda de agua potable de la empresa EMAPALA EP que garantice la distribución eficaz y eficiente del recurso hídrico mediante redes neuronales artificiales.

1.2.2. Objetivos Específicos

- Elaborar el marco teórico referencial de la investigación mediante la revisión de las técnicas y algoritmos empleados para el pronóstico de la demanda de agua potable.
- Diagnosticar la situación actual del consumo de agua potable de la empresa para lograr una mejor planificación y control del recurso hídrico.
- Desarrollar el modelo para el pronóstico de la demanda de agua potable de la empresa EMAPALA EP.
- Analizar los resultados obtenidos en la planificación de la demanda de agua potable a partir del desarrollo del pronóstico.

1.3. Justificación

La captación y distribución del agua han sido unos de los principales pilares para la supervivencia y desarrollo de las poblaciones, actualmente las empresas encargadas del abastecimiento de este recurso hídrico buscan las maneras de cómo satisfacer de agua potable a toda la población. Es de vital importancia de que estas entidades encargadas del suministro puedan adaptarse con facilidad y dotar de este recurso hídrico a los habitantes, en especial a los sectores en donde su población tiene un crecimiento considerable.

Según el objetivo 1 del Plan Nacional Toda una Vida, el Estado reconoce y garantiza una vida digna con iguales oportunidades para todas las personas, esto incluye la dotación de los servicios básicos fundamentales para el desarrollo poblacional como son: el Agua Potable, Saneamiento y Alcantarillado (Senplades, 2017).

Según la Asamblea General de las Naciones Unidas, mediante la resolución 64/292. El derecho al agua y saneamiento, aprobada el 28 de julio del 2010, reconoce que el derecho al agua potable y

el saneamiento es un derecho humano esencial para el pleno disfrute de la vida y de todos los derechos humanos.

El desconocimiento de los sectores en donde se incrementa el crecimiento poblacional y la incorrecta planificación para el abastecimiento de este recurso ha generado la necesidad de contar con una herramienta que permita identificar, pronosticar y evaluar la demanda del recurso hídrico. Con la ayuda del software R para la ejecución del modelo de pronósticos, lo que se pretende es reducir la incertidumbre de lo que pueda suceder en un futuro cercano, basándose en información que se asemeja a la realidad, generando un alto nivel de confianza y satisfacción de los clientes.

Los pronósticos exactos también contribuyen en la toma de decisiones (Ghiassi, Zimbra, & Saidane, 2008), (Jain & Ormsbee, 2002), (Manuel Herrera, Torgo, Izquierdo, & Pérez-García, 2010). Comúnmente el agua suministrada se distribuye en base a la experiencia de los trabajadores encargados, no obstante, un pronóstico preciso y confiable de la demanda de agua potable ayuda a la entrega de un servicio confiable y continuo (Tiwari & Adamowski, 2014).

Las Redes Neuronales Artificiales (ANN) se han visto como una tecnología para minería de datos, debido a que permite modelar de manera efectiva y eficiente problemas grandes y complejos (Salas, 2004). Anteriormente eran utilizados con fines académicos y ciencia, pero en la actualidad es un método nuevo y novedoso para las industrias y empresas públicas, las mismas que están optando por salir de lo convencional e integrar métodos de pronósticos que les permitan tomar buenas decisiones, obtener solvencia y una correcta planificación de sus recursos.

1.4. Metodología

El enfoque que se utilizará para la adquisición de información será Cuantitativo, debido a que en el estudio se analizará datos numéricos por medio de estadísticas para dar solución a preguntas de investigación o realizar verificaciones de hipótesis.

Se utilizarán una serie de métodos y técnicas, mismas que permitirán alcanzar los objetivos establecidos y los resultados esperados en la presente investigación. El principal método que se llevará a cabo es el Método General de Pronóstico, mismo que identifica una serie de pasos organizados para el correcto desarrollo de la proyección.

N°	Paso	Descripción
1	Identificar el horizonte temporal	Corto, mediano y largo plazo
2	Observar los hechos	Identificar las variables a predecir y sus incidencias
3	Seleccionar una técnica	Seleccionar las técnicas que se ajusten mejor al tipo de problema presentado
4	Aplicar técnicas seleccionadas	Desarrollo del pronóstico

Tabla 1.1 Método General de Pronóstico

Fuente: (Siadén & Andrea, 2012)

1.4.1. Métodos de Pronóstico de Demanda

Método General del pronóstico

- Método de Diseño de Modelos de Pronóstico con Redes Neuronales Artificiales

Técnicas

Redes Neuronales Artificiales

Herramientas

- Software R Studio

1.4.2. Metodología bibliográfica

Es un proceso sistemático en el cual se analizará, selecciona, evalúa y se aplica todo el material ya sea físico o digital que va a ser utilizado como fuente de teoría para la realización de la investigación.

1.4.3. Método Deductivo

Mediante la realización de la aplicación de Redes Neuronales Artificiales para el pronóstico de la demanda, se podrá analizar la proyección estimada para la distribución del recurso hídrico.

1.4.4. Método Descriptivo

Se analiza la situación actual de la empresa mediante la observación detallada y el registro de las actividades que se desarrollan en la misma.

1.4.5. Metodología para la Aplicación de Redes Neuronales Artificiales

La metodología consta de los siguientes pasos:

1. Selección de la variable
2. Recolección de datos
3. Preprocesamiento de los datos
4. Definición del conjunto de Entrenamiento, Validación y Prueba
 - a) Conjunto de Entrenamiento
 - b) Conjunto de Validación
 - c) Conjunto de Prueba
5. Selección de la arquitectura de Redes Neuronales Artificiales
 - a) Número de Neuronas de Entrada
 - b) Número de Capaz Ocultas
 - c) Número de Neuronas Ocultas
 - d) Número de Neuronas de Salida
 - e) Función de Transferencia
6. Criterios de Evaluación
7. Entrenamiento de las Redes Neuronales Artificiales
 - a) Número de Iteraciones
 - b) Tasa de Aprendizaje y Momentum
8. Implementación del Modelo de Redes Neuronales Artificiales

1.5.Alcance

La relevancia de la investigación se enfoca en la necesidad de realizar el pronóstico de la demanda de agua potable de la empresa EMAPALA EP, mediante la utilización de redes neuronales artificiales, juntamente con el apoyo del software R Studio, herramienta utilizada para los pronósticos esperados en dependencia del crecimiento de la población y la demanda de obtener este recurso.

Para garantizar el correcto modelamiento del pronóstico, se observará el crecimiento poblacional del cantón Lago Agrio, se recopilará los datos históricos de la demanda de consumo de 36 meses, se establecerá el tiempo de la proyección esperada, la tendencia del consumo del recurso hídrico de la población, las estacionalidades que presentan en ciertas épocas del año y los errores de la proyección del modelo de la demanda.

La ejecución total del proyecto dispondrá de un periodo de 6 meses, en donde se trabajará juntamente con el departamento de Operaciones y Mantenimiento para la recolección y análisis de la información. Al culminar el tiempo establecido, se entregará a la empresa EMAPALA EP, el modelo y la información del pronóstico determinado para el abastecimiento de Agua Potable al cantón Lago Agrio.

El modelo de pronósticos contribuirá al mejoramiento de la planificación de agua potable y apoyará a la investigación del crecimiento poblacional para los años venideros, de la misma manera aportará a la correcta toma de decisiones y desenvolvimiento de las áreas que conforma la empresa.

CAPÍTULO II

2. FUNDAMENTACIÓN TEÓRICA

La información presentada se enfoca en la normativa legal, el estudio del proceso de captación y distribución de agua, de los métodos y técnicas empleados en la ejecución de los pronósticos juntamente con la utilización de redes neuronales artificiales. Teniendo en cuenta la importancia que conlleva en la planificación de las empresas o entidades destinadas a la distribución del recurso hídrico potable.

2.1. Normativa Legal

Para la ejecución del trabajo de grado, la fundamentación legal es una de las principales temáticas que se deben llevar a cabo debido a que ayuda a sustentar, fortalecer y justificar la necesidad de la realización de este proyecto, ya que se enfoca principalmente en el mejoramiento de la calidad de vida de la población.

Una buena calidad de vida es un derecho que todas las personas deben poseer, debido a ello el Estado reconoce en la Constitución de la República del Ecuador, Art. 12 Sección primera que “El derecho humano al agua es fundamental e irrenunciable. El agua constituye patrimonio nacional estratégico de uso público, inalienable, imprescriptible, inembargable y esencial para la vida ([CONST], 2008, art.12).

Uno de los objetivos más importante para el bienestar de la población ecuatoriana es el objetivo 1 del Plan Nacional Toda una Vida, en el cual el Estado reconoce y garantiza una vida digna con iguales oportunidades para todas las personas, esto incluye la dotación de los servicios básicos fundamentales para el desarrollo poblacional como son: el Agua Potable, Saneamiento y Alcantarillado (Senplades, 2017).

Ecuador prohíbe la privatización del agua, ya que reconoce que es un recurso fundamental para la economía, el ambiente y principalmente la vida. Debido a que el recurso hídrico debe permanecer público se creó la Ley de Recursos Hídricos Uso y Aprovechamiento del Agua en el Registro Oficial No. 305, en la cual se estipula que el “Estado garantiza el derecho humano al agua como el derecho de todas las personas a disponer de agua limpia, suficiente, salubre, aceptable, accesible y asequible para uso personal y doméstico en cantidad, calidad, continuidad y cobertura, entre otros aspectos”(“Asamblea Nacional de la República del Ecuador," 2014).

Contar con un recurso hídrico potable de calidad es un derecho de la ciudadanía, por esta razón EMAPALA trabaja juntamente con la Agencia Nacional de Regulación, Control y Vigilancia Sanitaria (ARCSA), la cual tiene como misión encargarse de la protección de la salud de la población mediante la vigilancia y el control de los establecimientos y productos de consumo humano.

El recurso hídrico debe ser de calidad y contar con todos los parámetros establecidos permitiéndola ser apta para el consumo del ser humano, por esta razón la empresa de agua potable trabaja de la mano con la Agencia de Regulación y Control del Agua (ARCA), encargada de regular y llevar un control de la gestión integral anclada a este recurso como lo es el agua. Al mismo tiempo se encarga de verificar el aprovechamiento económico y la calidad de los servicios públicos anexados al agua, incentivando al cuidado al uso responsable de este patrimonio vital ("ARCA," 2020).

2.2. Producción de Agua Potable

Es el agua apta para el consumo del ser humano, el agua potable debe estar exento de bacterias y microorganismos que pueden generar problemas a la salud de las personas. Para que el líquido que es consumido por los seres humanos sea distribuido por todos los hogares, debe primero pasar por una planta de tratamiento o potabilizadora, esta a su vez limpia y trata el agua hasta que esté apta para el total y normal consumo por parte de los consumidores.

La dotación del servicio de agua potable confiable y segura a toda la población es uno de los principales aspectos de una empresa que provee el abastecimiento del líquido vital (Bata, Carriveau, & Ting, 2020).

El sistema de abastecimiento de agua debe constar de cinco partes esenciales para el suministro de agua potable:

- Red de abastecimiento
- Captación
- Almacenamiento de agua bruta
- Transformación de agua
- Almacenamiento de agua tratada
- Distribución por medio de conducciones

En la figura 2.1, se puede identificar las distintas fases fundamentales que intervienen en el tratamiento y distribución del agua.



Figura 2.1 Sistema de abastecimiento de agua

Fuente: (Aristegui, 2016)

2.2.1. Red de Abastecimiento o Distribución

En la actualidad la necesidad de poder contar con un servicio de agua potable se ha incrementado y con ello la importancia de construir sistemas de redes de tuberías enlazadas entre sí que permitan llevar el agua potable desde la zona de captación hasta las zonas de consumo del recurso hídrico potable. Cabe recalcar que el agua suministrada debe ser debidamente tratada y potabilizada, dando cumplimiento con las normativas de salubridad y de calidad.

Sin embargo existen algunos factores propios de influencia de la región que participan en la demanda y por ende la distribución de agua potable como son la clase de consumidor, las tarifas de consumo, crecimiento demográfico, recursos económicos, avances tecnológicos y las estrategias que se plantee la empresa para poder tener una manejo óptimo de todo el sistema hídrico potable (Bata et al., 2020).

2.2.2. Captación

La captación del agua es el inicio de la red de abastecimiento del recurso, es la acción de obtener convenientemente el agua, para la realización del proceso de captación, se debe conocer si el agua es origen subterráneo o superficial. La recepción del recurso hídrico es ejecutada por medio de pozos o galerías de extracción, y en caso de que las aguas se encuentren en la superficie terrestre, son captadas por medio de bocatomas, galerías filtrantes, perpendiculares o paralelas. Se debe tener siempre en cuenta el curso del agua para garantizar que la mayor cantidad de agua sea obtenida.

2.2.3. Almacenamiento

Se denomina al proceso y consecuencia de depositar el agua en un determinado espacio. Las instalaciones que son adecuados para el almacenaje de agua son denominadas embalses, los cuales conservan el agua procedente de ríos, arroyos, lluvia, etc. Además, permiten garantizar el agua que va a ser utilizada para su potabilización y distribución a toda la población.

2.2.4. Transformación

Es la parte más compleja del sistema debido a que transforman el agua bruta en agua potable y apta para el consumo humano. El tratamiento que se realiza se modifica dependiendo de la calidad de agua bruta que se encuentre. La potabilización consta de las siguientes partes:

Reja: Imposibilita el acceso de ‘material grueso’ y lo aparta una vez en la superficie.

Desarenador: Impide el paso de materiales en suspensión.

Floculadores: Se encarga de añadir los productos químicos para separar materiales finos y sustancias en suspensión coloidal.

Decantadores y filtros: Los decantadores, llamados también sedimentadores, aíslan una parte de material fino. Los filtros son útiles para retirar el material en suspensión.

Filtros: Son utilizados para apartar completamente el material sobrante.

2.2.5. Almacenamiento de agua potable

Se denomina al proceso de depositar el agua ya tratada en espacios específicos. Para el agua ya potabilizada existen dos tipos de tanques en cuales se almacena como son: los tanques apoyados en el suelo, que son colocados a un extremo de la planta de tratamiento y distribución, o los tanques elevados, que son colocados en lugares con una altitud mayor a la red de alimentación.

2.2.6. Distribución

Generalmente la distribución se inicia desde los tanques de almacenaje de agua potable. La red de distribución para abastecer a toda la población del recurso hídrico potable consta de componentes como son: Estación de bombeo, tuberías principales, secundarias y terciarias, tanques de almacenamientos ubicados en zonas intermedias, válvulas que dan apertura a la operación de la red y sectorizar el suministro en caso de excepción, dispositivos de macro y micro medición, medidores domiciliarios.

2.2.7. Demanda de Agua Potable

Se puede definir como la cantidad de agua que es requerida en las tomas para el consumo de una localidad o área de proyectos que se están ejecutando. Se consideran algunos tipos de usuarios como: domésticos, industrias, comerciales, zonas turísticas. En cada uno de estos tipos de consumidores existen y se tienen en cuenta el agua no contabilizada, esto se da a fugas, conexiones clandestinas, etc. La estimación de la demanda de agua potable se realiza generalmente en base al crecimiento de la población que constantemente está en acenso, al cobertura esperada y del consumo que se tiene previsto suministrar (Siadén & Andrea, 2012).

2.3.Pronóstico

El pronóstico es definido desde varias perspectivas, pero a su vez consta de tres características específicas que permiten dar un concepto acertado:

- Futuro: Es el cálculo de los datos futuros de una variable estudiada.
- Incertidumbre: Al momento de realizar un pronóstico se debe tener en cuenta que se va siempre a tener un margen de error, para lo cual el pronóstico busca que este valor tenga un margen mínimo.
- Juicio personal: En la realización de un pronóstico, se debe tener presente la persona que va a realizar dicha predicción. La experiencia y un buen juicio permiten a la persona recolectar los datos y determinar que método será el que se va a utilizar, de la misma manera lo hará para los resultados obtenidos y su interpretación.

Cabe recalcar que los pronósticos cuentan con una base general, es decir poseen una estructura que puede ser utilizada para cualquier tipo de predicciones como se establece en la tabla 1.

Nº	Pasos	Descripción
1	Identificar el horizonte temporal	La cual puede ser corto, mediano o largo plazo
2	Observar los hechos	Identificar las variables a predecir y sus dependencias.
3	Seleccionar una técnica	Seleccionar las técnicas que más se ajustan al tipo de problema presentado.
4	Aplicar técnica seleccionada	Implementar el pronóstico de la técnica seleccionada.

Tabla 2.1 Método General de Diseño de un Pronóstico

Fuente: (Siadén & Andrea, 2012)

Entonces podemos conceptualizar que un pronóstico es la estimación del valor futuro de una variable mediante la aplicación de métodos novedosos y procedimientos que aporten en la disminución del margen de error, teniendo en cuenta el buen juicio y la experiencia del responsable de realizar dicha estimación (Gutiérrez, 2013).

Pronosticar es muy importante en tipos de empresas, ya que la estimación de hechos futuros se puede incorporar al proceso de toma de decisiones (Bowerman, Richard, II, & Anne, 2007).

Para la realización de una predicción efectiva, es necesario contar con la mayor cantidad de información histórica del producto o servicio que se va a analizar, esto fundamentará el trabajo realizado y garantizará que los datos predictivos sean óptimos y con un margen de error mínimo. Según (Bata et al., 2020), existen 3 factores fundamentales que deben ser considerados para la realización de un pronóstico relativamente preciso. El principal factor es la utilización de la demanda histórica de agua potable, según (Guo et al., 2018), (Arandia, Ba, Eck, & McKenna, 2016), quienes utilizaron únicamente este insumo para sus pronósticos, los resultados obtenidos se mostraron razonables y precisos.

Autores como (Rice, Carriveau, Ting, & Bata, 2017), señalan que para un buen pronóstico además de la demanda histórica del agua potable es necesario la utilización de variables como el clima, la estacionalidad de los datos de temperatura, la evaporización, los días y horas a la semana.

Otro de los factores determinantes es el horizonte de temporal, este a su vez posee dos dimensiones de tiempo, el primero es para determinar el periodo pasado que se necesita para el pronóstico, es decir los datos históricos. El segundo es para determinar cuánto va a durar el pronóstico, cabe recalcar que, para obtener un buen pronóstico, se debe contar con una cantidad de datos históricos considerables.

El tercer factor determinante es el método modelo en sí. Los modelos de Redes Neuronales (ANN) han demostrado tener una gran viabilidad para la realización de pronósticos, lo han demostrado en aplicaciones tanto lineales como no lineales debido a que poseen una gran capacidad para aprender constantemente, entrenar por sí mismo y mejorar los sistemas para la predicción de la demanda de agua potable (Bata et al., 2020).

Según (Pacchin, Gagliardi, Alvisi, & Franchini, 2019), el pronóstico preciso de la demanda de agua tiene un protagonismo importante en la optimización del diseño, operación y gestión de las municipalidades encargadas del suministro del recurso hídrico.

La proyección, constituye uno de los principales pilares para una empresa u organización, debido a que influye en la planificación y por ende en la economía de la misma. Tiene una gran incidencia en factores determinantes como son: la capacidad instalada, la producción, la capacidad financiera, la estructura organizacional, la contratación de maquilado, etc.

2.3.1. Aplicaciones de los pronósticos

Los pronósticos pueden ser utilizados principalmente por las empresas para estimar la diversidad de variables que pueden tener, se les puede denominar micro pronósticos, debido a que generan un gran impacto en una empresa o en una entidad ya sea pública como privada. Algunas áreas en donde se pueden generar los pronósticos son: producción, mercadotecnia, finanzas, recursos humanos y planificación estratégica.

En algunas de las ocasiones se necesita hacer proyecciones de variables a nivel macro como el índice de empleo en un país o el nivel de inflación o el PIB, aquello se lo denomina macro pronósticos, debido a que se trata a un impacto general ya sea de un país, ciudad o región.

2.3.2. Clasificación de los métodos de los pronósticos

Para la realización de las predicciones pueden utilizarse una serie de métodos y técnicas que son clasificadas en dos tipos que son: métodos cuantitativos y cualitativos.

2.3.3. Métodos cualitativos

Se basa principalmente del conocimiento de los expertos, los cuales pueden realizar predicciones con su experiencia, juicio y otros aspectos no numéricos que les permiten tener una validez en sus predicciones. Esta clase de métodos se utilizan cuando no existen datos, no hay confiabilidad en la información o su acceso está restringida (Gutiérrez, 2013).

Algunos de los métodos que tienen mayor relevancia son:

2.3.3.1. Consenso de un panel

Este método se basa en reunir a diferentes expertos para la realización de análisis de variables que tengan aspectos en común y dialogar en el futuro de estas, de esta manera que se podrá llegar a un consenso en donde se podrá identificar el pronóstico de las variables discutidas. Todo el proceso para la toma de decisiones debe ser guiado por una persona que será el moderador, así se alcanzará el objetivo propuesto (Gutiérrez, 2013).

La ventaja de esta metodología es, que se pueda reunir a una serie de personas expertas que puedan compartir sus experiencias y las soluciones que podrían aportar para tener una visión positiva. Pero en caso de que uno de los que conforman el grupo tenga un punto de vista diferente a los demás integrantes, va a presentarse ciertas inconformidades que imposibilitaría un consenso entre todas las partes participantes.

2.3.3.2. Analogía histórica

Generalmente es utilizado al momento de pronosticar la demanda que va a tener un determinado producto o servicio, a su vez no cuenta con la suficiente información necesaria y genera una incertidumbre en los resultados que se van a generar en el futuro. La base de este método supone que los pronósticos realizados de algún producto o servicio que está ingresando un mercado, posee una similitud con otro que ya está siendo comercializado y del cual se posee información histórica. De este modo se puede decir que se va a tener una demanda similar en el futuro a pronosticar (Gutiérrez, 2013).

2.3.3.3. Investigación de mercados

Hoy en día para la introducción de nuevos productos y servicios al mercado deben realizar posteriormente un estudio de mercado y, la necesidad de analizar e identificar la aceptación que va a tener por parte de los consumidores. Este método es de gran utilidad para realización de métodos cualitativos. Algunas de las técnicas que son utilizadas permiten tener una visión de la futura demanda del producto o servicio que va a ser comercializado por primera vez en el mercado (Gutiérrez, 2013).

2.3.4. Métodos cuantitativos

Para la utilización de este tipo de método, es necesario contar con datos históricos que posean validez para predecir valores a futuro de variables a considerar. Par la manipulación de este método, no se necesita que la persona encarda de hacerlo tenga experiencia, debido a que no necesita tener un resultado basado en la intuición, no obstante, es necesario de cierta manera tener una perspectiva cualitativa que permitirá analizar la información y tomar decisiones.

Estos métodos pueden ser clasificados en dos grupos los cuales son: Pronósticos de serie de tiempos y modelos causales.

2.3.4.1. Métodos de pronósticos para serie de tiempo

Esta clase de métodos se basa en la suposición de que el comportamiento de los datos históricos seguirá portando un mismo patrón, en un futuro con la realización de la proyección trata de identificar un proceder oculto en la serie de tiempo. Se debe identificar cuando un método de este tipo va a ser útil o no, esto depende del comportamiento de los datos históricos en un futuro, si el histórico es el mismo, el método es acertado caso contrario el grado de utilidad sería baja

Algunos de los métodos de pronósticos de series de tiempo son:

- Spectral Analysis
- Box-Jenkins
- State space

2.3.4.2. Modelos causales

Se debe tener en cuenta que este método es utilizado cuando la información recopilada no se toma en consideración las variables de tiempo, de la misma manera se toman en cuenta los datos recopilados aproximadamente en la misma locación del tiempo. Lo que se pretende identificar es, si existen otras variables que puedan tener relación con la variable de interés, se debe tener en cuenta ciertas medidas de comportamiento que va a poseer la variable que se pretende pronosticar.

Por lo general son llamados causales, debido a que sus variables presentan una relación causa-efecto, esto nos quiere decir que la causa es generada en dependencia del comportamiento de una o varias variables, y el comportamiento de otra variable tomado como medida puede generar el efecto.

En la clasificación de los modelos causales se identifican principalmente:

- Modelos de regresión
- Modelos econométricos

2.3.5. Selección de un método de pronóstico

Hoy en día para la realización de pronósticos, existen varios métodos que pueden identificar, analizar y solucionar las diferentes dudas que pueden generarse en las empresas acerca del comportamiento de la demanda en el futuro de los diferentes productos y servicios que son comercializados en el mercado. No se puede decir con seguridad que un método sea mejor que otro, ya que cada uno cuenta con características que le permiten (Mohamad, 2016a) generar una predicción, pero para elegir el mejor método a utilizar dependerá de ciertos factores como:

- Disponibilidad de datos: Con este factor parte la investigación, debido a que se identificará si consta o no con información relevante a la variable a analizar, y se decidirá si es un método cualitativo o cuantitativo.
- Precisión deseada: Se debe identificar el grado de exactitud que se pretende alcanzar, para lograr una buena precisión en el pronóstico, mismo que contará con el análisis de varios métodos de predicción. En otras ocasiones que presentan un mayor grado de complejidad, se necesitará una gran cantidad de tiempo y medios.

- Uso que se le dará al pronóstico: Frecuentemente las empresas tienen un mayor interés en la ejecución de pronósticos, pero siempre hay que tener en cuenta el propósito y el impacto que se obtendrá al momento de su finalización.
- Disponibilidad de recursos: Al momento de la realización de los pronósticos es esencial conocer la cantidad de recursos con los que se cuenta. En ocasiones habrá predicciones que no necesitarán en gran proporción el recurso económico, pero sí el recurso tiempo, o el pronóstico tenga una gran complejidad que sea necesario necesitar el bien humano y equipos especializados.
- Importancia del pasado para estimar el futuro: Para realizar una estimación hacia el futuro, es muy importante conocer el trayecto que posee una variable, de tal manera se determinará si la información histórica es de utilidad o no para la realización del pronóstico.
- Persona que va a realizar el pronóstico: Se debe tener en cuenta que el responsable de la realización de los pronósticos debe contar con intuición para seleccionar el método a utilizar y poder obtener el resultado esperado.

2.3.6. Series de tiempo

Para la realización de un buen pronóstico es necesario contar con datos confiables, secuenciales y ordenados. Los datos pueden ser recolectados diarios, semanales, mensuales, anuales, etc. Así como se lo realiza para el pronóstico de la empresa de agua potable, se identifica la cantidad de recurso hídrico que ha suministrado a la ciudad en los 3 años pasados. Los datos recogidos son mensuales, es decir, deben seguir un orden para poder almacenar correctamente la información que va a ser analizada y posteriormente procesada.

Para la identificación del método de pronóstico se analiza los patrones con los que se podrían comportar los datos, como se muestra en las figuras 2.2-2.4.

- Estacionaria: Esta serie de tiempo se identifica por estar dentro de una línea horizontal o estable a lo largo de un periodo de tiempo.
- Con tendencia: Esta serie de tiempo se caracteriza por poseer una tendencia que se puede comportar tanto en forma descendente o ascendente de acuerdo con los datos presentados.

- Estacional: Esta serie de tiempo posee un patrón con una conducta repetitiva con lapsos de tiempos menores o iguales a un año. Además, estas estacionalidades son divididas de igual magnitud.

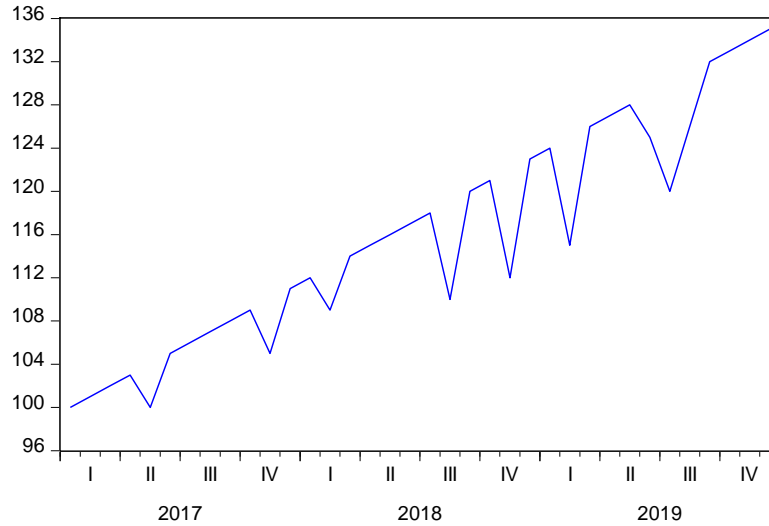


Figura 2.2 Serie de tiempo con tendencia

Fuente: Propia

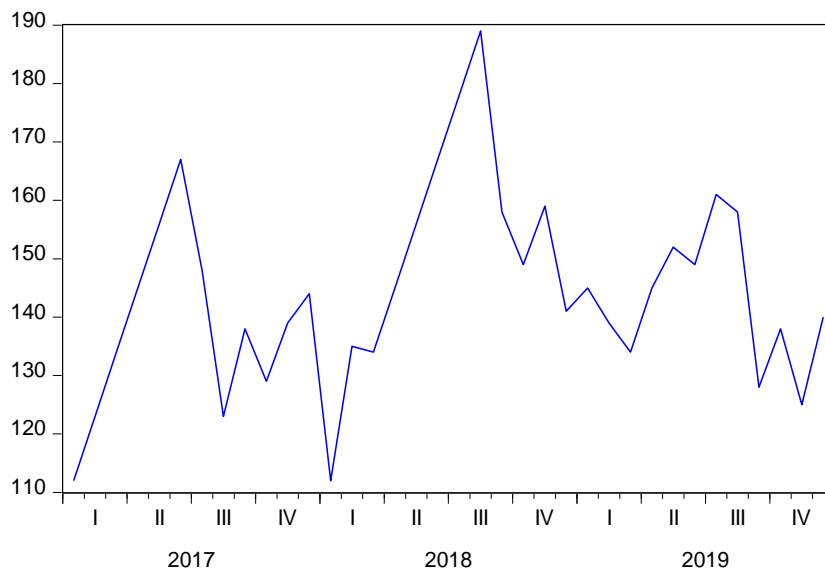


Figura 2.3 Serie de tiempo estacionaria

Fuente: Propia

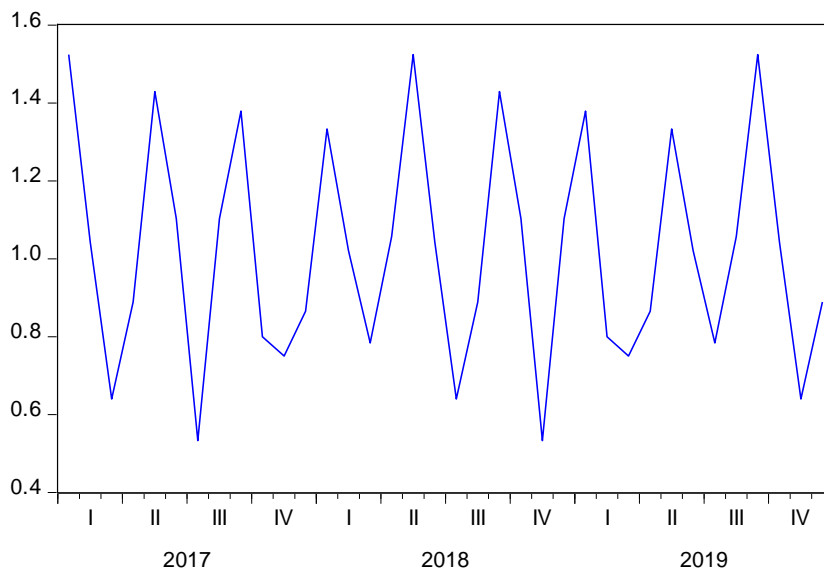


Figura 2.4 Serie de tiempo estacional

Fuente: Propia

2.3.7. Mediciones del error de un pronóstico

En un pronóstico siempre se va a contar con un margen de error variable, de esto dependerá el grado de precisión que tendrá la estimación realizada. Los datos obtenidos por la estimación serán más confiables si el margen de error es mínimo. Por esta razón se debe realizar varias mediciones del error relacionado con las predicciones resultantes por medio de un método específico.

Se define al error como la variabilidad que presenta el valor real con lo pronosticado, el error se lo puede calcular con la fórmula 1.

$$e_t = Y_t - F_t \quad (1)$$

e_t = Error del pronóstico en el periodo t

Y_t = Valor real de la variable en el periodo t

F_t = Valor pronosticado de la variable en el periodo t

Para la realización de los pronósticos de las series de tiempo aplican mediciones de errores, los cuales servirán de utilidad para la elaboración de pronósticos por medio de modelos causales, las fórmulas 2, 3, 4,5 y 6 son utilizadas para calcular los errores comunes en un pronóstico.

- Error medio (Mean error, o ME)

$$ME = \frac{\sum(Y_t - F_t)}{n} = \frac{\sum e_t}{n} \quad (2)$$

- Error absoluto medio (mean absolute error, o MAE)

$$MAE = \frac{\sum|Y_t - F_t|}{n} = \frac{\sum|e_t|}{n} \quad (3)$$

- Error porcentual medio (mean percentage error, o MPE)

$$MPE = \frac{\sum\left(\frac{Y_t - F_t}{Y_t}\right)}{n} = \frac{\sum\left(\frac{e_t}{Y_t}\right)}{n} \quad (4)$$

- Error porcentual absoluto medio (mean absolute percentage error, o MAPE)

$$MAPE = \frac{\sum\left|\frac{Y_t - F_t}{Y_t}\right|}{n} = \frac{\sum\left|\frac{e_t}{Y_t}\right|}{n} \quad (5)$$

- Error cuadrado medio (mean squared error, o MSE)

$$MSE = \frac{\sum(Y_t - F_t)^2}{n} = \frac{\sum e_t^2}{n} \quad (6)$$

Cada uno de los errores que se encuentran posee diferentes interpretaciones al momento de realizar el análisis de los pronósticos obtenidos o los modelos empleados. Algunos errores como el error medio (ME) y el error porcentual promedio (MPE), permiten identificar la orientación que se espera obtener en el pronóstico, ya que valorará si el promedio del pronóstico está por encima o debajo del valor esperado.

Los errores que permiten identificar el tamaño del error del pronóstico son los errores: (MAE) Absoluto medio y (MAPE) Porcentual absoluto medio. Estos errores trabajan solamente con la magnitud del error (valor absoluto), esta a su vez, permite dar una noción de la exactitud que se pretende esperar del pronóstico realizado. Pueden ser muy útiles al momento de realizar comparaciones de los resultados previstos entre diferentes métodos de estimación.

Existe un error que es más comúnmente utilizado para la comparación de resultados y es el (MSE) el error cuadrático medio, ya que identifica al momento de la comparación qué valor es el más pequeño para realizar un mejor pronóstico, debido a que tiene semejanza con lo establecido a cerca de la varianza y por ende con la desviación estándar.

2.3.8. Errores Bayesianos

El criterio de información bayesiana (BIC), es utilizado para la ejecución de cálculos de diferentes modelos como una función de bondad de ajuste, a su vez es utilizado para la selección del mejor modelo entre un grupo limitado de modelos. Su fundamentación se presenta en la probabilidad logarítmica (LLF) y guarda una gran relación con el criterio de información Akaike. Cabe mencionar que el BIC únicamente se puede utilizar cuando los valores de la variable dependiente de los modelos a ser analizados son similares, de esta manera todas las variables consideradas van a ser tomadas en cuenta (Mohamad, 2016b).

Para la identificación del mejor modelo se lo realiza con la siguiente fórmula, en la cual el modelo con el menor resultado obtenido será el óptimo y el utilizado, ya que su ajuste y número de variables interpretativas será inferior.

$$BIC = k * \ln n - 2 * \ln(L) \quad (7)$$

k = Número de variables del modelo

$\ln(L)$ = Función de log – verosimilitud para el modelo estadístico

2.3.9. Error Akaike

El criterio de investigación Akaike (AIC), posee una similitud con los errores Bayesianos (BIC), en la cual permite distinguir y verificar el mejor modelo entre la competencia, los criterios que este error analiza son la bondad de ajuste y la cantidad de parámetros que contiene (Ingdal, Johnsen, & Harrington, 2019). Según (Mohamad, 2016a), AIC no se puede definir como una prueba de hipótesis, sino se puede analizar como un apoyo para la identificación del mejor modelo de una herramienta en específico, en este caso el software R Studio. Cabe mencionar que el modelo que tenga el AIC más bajo, será el elegido para la ejecución de proyectos propuestos.

El error Akaike se define como:

$$AIC = 2k - 2 * \ln(L) \quad (8)$$

En donde:

k = es el número de parámetros de modelo

$\ln(L)$ = es la función de log – verosimilitud para el modelo estadístico

Para el análisis de conjunto de valores más pequeños, el AIC es aplicado en orden secundario.

$$AICc = AIC + \frac{2k(k+1)}{N-k-1} = \frac{2*N*k}{N-k-1} - 2 * \ln(L) \quad (9)$$

En donde:

N = Tamaño de muestra de datos

k = Número de parámetros del modelo

2.4. Machine Learning

A lo largo de la evolución el ser humano ha sido capaz de crear nuevos métodos, técnicas y disciplinas para lograr la simplicidad de las cosas, uno de los cuales es el aprendizaje automático o Machine Learning perteneciente al campo de la inteligencia artificial que permite a los sistemas computarizados sin intervención del ser humano, poder aprender a analizar y modelar la información correspondiente a patrones, tendencias y la relación existente entre datos.

Cabe recalcar que la información recolectada para ser previamente analizada es histórica, es decir; datos que son propios de un determinado producto o servicio que son archivados y pueden ser utilizados para diferentes propósitos que se requiera.

Según (Pal, Maity, Ratnam, Nonaka, & Behera, 2020), la predicción fundamentada en Machine Learning ha demostrado tener una gran confiabilidad en la realización de pronósticos en diferentes escenarios, pese a ello el método posee una capacidad limitada para la comprensión de procesos subyacentes, cabe mencionar la destreza de los algoritmos de ML, ya que se originan de la utilización de los datos identificados para el entrenamiento.

Debido al gran impacto que ha tenido en el ámbito empresarial, Machine Learning es utilizada en entidades tanto públicas como privadas, los beneficios que pueden tener con esta disciplina son: Predicción de tendencias, impulsa la innovación, mejoramiento del segmento del público, reducción de costos, mejoramiento en publicidad, mejoramiento de la atención y satisfacción del cliente.

El aprendizaje supervisado también cuenta con una serie de técnicas de clasificación que permite que un sistema computarizado pueda aprender adecuadamente y obtener datos con un rango mínimo de error. Algunas de estas técnicas son: Árboles de decisión, Reglas de asociación,

Algoritmos genéticos, Máquinas de soporte de vectores, Algoritmo de agrupamiento, Redes bayesianas, Regresión Lineal Múltiple, Máquinas de Vectores de Soporte y Redes neuronales artificiales, esta última técnica va a ser utilizada para la investigación de la predicción de la demanda de agua potable (L. L. Lorente-Leyva et al., 2019),(Herrera-Granda et al., 2019),(L. L. Lorente-Leyva, Patino-Alarcon, D. R., Montero-Santos, Y., Herrera-Granda, I. D., Peluffo-Ordonez, D. H., Lastre-Aleaga, A.M., Cordoves-Garcia, A., 2020).

A continuación, se presenta algunos métodos para pronosticar la demanda de agua potable:

2.4.1. Regresión Lineal Múltiple

Este método es una de las formas de regresión lineal que tiene la capacidad de replicar la relación que existe entre dos o más variables independientes conocidas como predictores y una variable dependiente. Es utilizada para predecir el valor de salida de una variable que depende de las variables establecidas a la entrada del modelo (Gupta & Saraswat, 2020).

Este modelo de regresión lineal múltiple tiene una respuesta (variable dependiente de salida) y k número de predictores es decir (variables independientes de entrada), se lo expresa en la figura 7 como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (10)$$

$Y =$ Variable dependiente de salida

$X_i =$ Variables dependientes de entrada (predictores) Representa un modelo residual

$\beta_0 + \beta_1 + \beta_2 + \dots + \beta_k =$ Coeficientes del modelo de regresión

2.4.2. Support Vector Machine

Este método se fundamenta en el aprendizaje supervisado, se desarrolla en base en la teoría de aprendizaje estadístico el cual cuenta con la capacidad de solucionar problemas de clasificación binaria no lineal. Para obtener una separación de clase adecuada, este método lo que hace es generar un grupo de posibles hiperplanos en un espacio de dimensiones superior. Para obtener una separación adecuada se lo obtiene a través de un hiperplano teniendo en cuenta la mayor distancia a los puntos de datos de entrenamiento más cercanos de cualquier clase (Gupta & Saraswat, 2020).

2.5. Redes Neuronales Artificiales

Las redes neuronales tratan de imitar el comportamiento del cerebro humano, el cual le caracteriza por la experiencia y la recolección del conocimiento a partir de un conjunto de datos. Podemos decir que este sistema emula gráficamente la configuración neuronal del cerebro mediante una simulación ordenada, a través de su modelamiento de estructura de su procesamiento con cierta capacidad de operación en paralelo, o por medio de la construcción física de sistemas cuya elaboración tiene similitud con la estructura neuronal biológica.

También son conocidas como redes interconectadas masivamente en paralelo con elementos simples y con una organización jerárquica, mismos que pueden interactuar con elementos del mundo, tal cual lo realiza el sistema nervioso biológico (López & Fernández, 2008).

Las redes neuronales poseen la capacidad de solucionar problemas de manera individual o grupal que tienen la aptitud de combinadas con otros métodos para acciones que se requiera clasificar, diagnosticar, identificar, predecir u optimizar, el balance de los datos y conocimientos tiene tendencia hacia los datos. Pueden existir necesidad de aprendizaje en el lapso que se va a ejecutar y en la cual puede existir una cierta tolerancia de fallos, las redes neuronales se acoplan con la realización de reajustes de manera constante en sus pesos de sus interconexiones (Salas, 2004).

Según (Alvisi & Franchini, 2017), las Redes Neuronales Artificiales se basan en las redes neuronales biológicas ya que puede recibir información, analizar las señales de entrada y determinar un resultado o salida. Uno de los tipos de ANN más comunes son el Perceptrón multicapa (MLP), en las cuales las neuronas están ordenas por capas (capas de entrada) que reciben la información y realizan la ponderación de los datos introducidos, luego la información pasa al procesamiento de la misma mediante funciones preestablecidas, esto se lo ejecuta por medio de (capas ocultas), una vez procesada la información se envía a la (capa de salida) para la verificación del resultado de la información.

Según (Ratnam, Dijkstra, & Behera, 2020), las Redes Neuronales Artificiales son básicamente modelos que poseen un aprendizaje automático desarrollados para el modelamiento de los procesos en las neuronas, poseen una condición no lineal y han sido utilizados por las ciencias climatológicas en la realización de pronósticos. Generalmente la estructura de estos modelos cuenta con una capa

de entrada, varias capas ocultas y una capa de salida. El número de neuronas en las capas ocultas determina la confiabilidad de las predicciones.

El propósito de la realización del modelo de Redes Neuronales Artificiales es la obtención del pronóstico de la demanda del recurso hídrico urbano, teniendo en cuenta las variables a utilizarse ya que servirán para la obtención del horizonte temporal del pronóstico y el nivel de confianza con el que contará el modelo (L. L. Lorente-Leyva et al., 2019).

2.5.1. Ventajas de las redes neuronales

Debido a la semejanza relevante que poseen las redes neuronales artificiales con el cerebro humano, se ha generado una serie de ventajas puesto que tienen la capacidad de adaptar la experiencia y estudiarla, procesar los casos históricos y generalizarlos con los futuros casos, aislar cualidades indispensables por medio de entradas que se establecen como información irrelevante (Matich, 2001).

2.5.2. Aprendizaje adaptativo

Esta es una de las ventajas más relevantes que presenta las redes neuronales, debido a que pueden realizar las tareas mediante un previo entrenamiento o experiencia. Las redes neuronales son sistemas adaptativos ya que cuentan con capacidad de autoajuste en los elementos que componen el procesamiento (Matich, 2001).

2.5.3. Auto-organización

Esta ventaja va de la mano con el aprendizaje adaptativo ya que le permite emplear sus capacidades de adaptación para reorganizar los datos que obtienen del aprendizaje y/o operación. Al momento de incurrir en el aprendizaje, se modificará cada elemento procesal, mientras que la auto organización modifica totalmente la red neuronal para poder cumplir con los objetivos propuestos (Matich, 2001).

2.5.4. Tolerancia a fallos

Las redes neuronales fueron los primeros sistemas computacionales en poder soportar fallos, aunque exista un fallo en una cantidad pequeña de neuronas y su sistema se vea comprometido, no sufriría una caída imprevista. Su capacidad a la toleración de fallas se debe a que la información almacenada, está distribuida por las conexiones entre las neuronas, dando como resultado una redundancia en esta clase de almacenamiento (Matich, 2001).

2.5.5. Operación en tiempo real

En todas las áreas de aplicación, es necesario que los procesos puedan realizarse de una manera inmediata. Esto es posible ya que las redes neuronales se adaptan rápidamente gracias a su implementación paralela. El cambio de los pesos de las conexiones de las redes debe ser mínimo para que puedan funcionar sin interferencias en tiempo real (Matich, 2001).

2.5.6. Fácil inserción dentro de la tecnología existente

El mejoramiento de los sistemas existentes se puede dar mediante la utilización de las redes neuronales, ya que estas tienen la capacidad de comprobarse, verificarse y trasladarse a un soporte físico. Esto permitirá evitar dificultades para la inserción de las redes neuronales y garantice una posterior evaluación del sistema para desarrollos más amplios (Matich, 2001).

2.5.7. Característica de las Redes Neuronales Artificiales

Las redes neuronales artificiales están inspiradas en las neuronas biológicas que poseen el cerebro de los seres humanos, su estructura es similar y tiene la capacidad de imitar las funciones más comunes que el cerebro humano realiza (Olabe, 1998).

Al ser tan similares la red neuronal biológica con la red neuronal artificial presentan una serie de características:

- **Aprender:** La obtención del conocimiento se debe realizar por medio del estudio, actividad constante y experiencia, las redes neuronales se modifican su conducta dependiendo del ambiente. Se les puede ingresar un grupo de inputs y se encargan de procesarla y obtener outputs con el menor error posible.
- **Generalizar:** Expandir o incrementar, en las redes neuronales artificiales es común debido a su composición y naturaleza. Las RNA tienen la capacidad de proveer, respuestas concretas a las entradas que poseen diminutas variaciones debido a los factores de ruido o distorsión.
- **Abstraer:** Apartar o separar directamente las cualidades de un elemento. Las redes neuronales artificiales tienen la capacidad de retirar las propiedades de un grupo de entradas que no presentan similitud o aspectos en común.

2.5.8. Aplicaciones de las Redes Neuronales Artificiales

- **Clasificación:** Permite reconocer los datos obtenidos y clasificarlos según la clase a la que pertenezca. De la misma manera posee la capacidad de poder realizar para un problema establecido estructuras que especifican la identificación de sus características.
- **Asociación:** El funcionamiento de esta red es como el proceso de restablecer los datos a partir de la entrada establecida relacionada con los datos previamente almacenados.
- **Agrupamiento:** Los datos de entrada no son clasificados, a medida que se van obteniendo la señal, la red es la encargada de generar la clasificación.
- **Generalización:** Las redes tienen la capacidad de identificar componentes ocultos que guardan relación entre diferentes elementos y pueden deducir las relaciones numéricas que existen en los datos presentados en situaciones similares.
- **Optimización:** Al momento de identificar la opción más factible entre un grupo de alternativas, las redes neuronales artificiales facilitan esa selección de una manera rápida y con las características necesarias para dar solución al problema establecido.
- **Predicción y control:** Las redes neuronales artificiales poseen una gran flexibilidad para analizar los datos ya que el sistema va aprendiendo y puede acoplarse a los cambios que se puedan presentar.

En la literatura (Kaastra & Boyd, 1996), se establece un procedimiento para diseñar un modelo para una red neuronal artificial que permita realizar predicciones. El modelo consta de ocho pasos los cuales son:

1. Seleccionar la variable
2. Recolectar los datos
3. Preprocesamiento de datos
4. Conjunto de entrenamiento, prueba y validación
5. Paradigmas de redes neuronales
 - 5.1. Número de capas ocultas
 - 5.2. Número de neuronas ocultas
 - 5.3. Número de neuronas de salida
 - 5.4. Funciones de transferencia

6. Criterio de evaluación
7. Entrenamiento de red neuronal
 - 7.1. Número de iteraciones de entrenamiento
 - 7.2. Tasa de aprendizaje e impulso
8. Aplicación

2.5.9. Estructura básica de una red neuronal

2.5.9.1. Analogía con el cerebro

La unidad fundamental del sistema nervioso es la neurona, específicamente del cerebro. Las neuronas son unidades de procesamiento e información de señales entre sí. La combinación de las entradas debe ser suficientemente fuerte para que la señal de salida de una neurona se active (Campos, 2018).

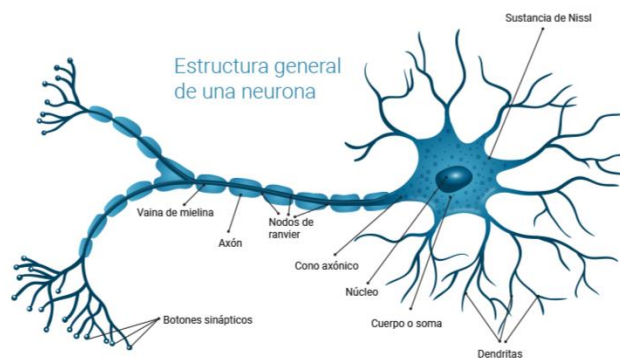


Figura 2.5 Componentes de una neurona

Fuente: (Campos, 2018)

2.5.9.2. Diagrama de una neurona artificial PE

La unidad similar a la neurona del ser humano es la neurona del ser humano, el cual es llamada EP (Elemento procesador), esta parte procesadora cuenta con una variedad de entradas mismas que son combinadas principalmente con una suma de básica. Los elementos procesadores pueden tener conexiones con las entradas de diferentes neuronas artificiales por medio de nexos ponderados, los cuales pertenecen a la eficacia de la sinapsis de las conexiones neuronales (Khepri, 2018).

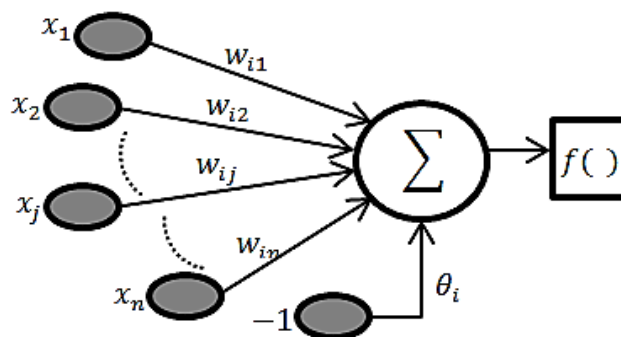


Figura 2.6 Diagrama de una neurona artificial PE

Fuente: (Khepri, 2018)

Una red neuronal consiste en un conjunto de unidades elementales PE conectadas de una forma concreta. El interés de las ANN no reside solamente en el modelo del elemento PE sino en las formas en que se conectan estos elementos procesadores. Generalmente los elementos PE están organizados en grupos llamados niveles o capas. Según (García, 2019), las capas son una agrupación de neuronas en las que las capas anteriores a ella son descritas como su entrada, esto no pasa lo mismo con la primera capa del modelo, la cual la entrada viene siendo los datos que se ingresa al sistema y de esta capa su salida viene siendo la entrada de una posterior.

2.6. Modelos neuronales

Estos modelos aceptan ciertas simplificaciones del patrón biológico para el desarrollo matemático, para una correcta realización de un modelo de red neuronal, en su metodología debe utilizar arquitecturas y reglas de aprendizaje para permitir la obtención de la estructura estadística que está presentes en los datos. Los elementos que debe poseer para este proceso son: la estructuración de los nodos, topología de la red y el algoritmo de aprendizaje son utilizados para la “estimación” de los pesos o parámetros de la red (Torra Porrás, 2004).

Los modelos neuronales presentan ventajas esenciales para su funcionamiento como lo son el aprendizaje, la generalidad y la robustez en la presencia del “ruido” que presentan los datos. Podemos decir que la recolección de la información se lo puede obtener por medio del aprendizaje de la estructura interna de la información, de tal manera que son las conexiones o los pesos entre los elementos que forman las redes neuronales en las cuales se conserva los datos.

Los aspectos que se toman en mayor relevancia en los modelos de red neuronal son:

- Uno de los aspectos relevantes es la arquitectura, y que la constitución entre las conexiones de las unidades neuronales. La forma que presenta ha permitido generar una variedad de posibles modelos los cuales han sido clasificados de acuerdo con ciertas variables en común.

En un primer acercamiento, se puede identificar una clasificación en dependencia del tipo de salidas (output) que es generada por el modelo, se los puede dividir en dos tipos: modelos deterministas y modelos estocásticos. En el caso del modelo determinista cada una de las neuronas precede una ley del modelo, como se puede observar en la ecuación 7.

$$y = f(\sum_{i=1}^n w_i x_i) \quad (11)$$

En la cual $f(.)$ es la función que permite la activación, para el caso del modelo estocástico, la interpretación para la activación de la red se da como una probabilidad de una situación lógica, como se observa en las fórmulas 9 y 10.

$$P = (y \equiv 1) = f(\sum_{i=1}^n w_i x_i) \quad (12)$$

$$P = (y \equiv 0) = 1 - P(y \equiv 1) = 1 - f(\sum_{i=1}^n w_i x_i) \quad (13)$$

Donde la salida (output) se interpreta como un valor continuo entre [0,1] a su vez es interpretado como la probabilidad.

- La tipología es otro de los aspectos que existen en las unidades de las neuronas artificiales. En este modelo existen neuronas visibles y neuronas ocultas, las neuronas visibles se encuentran a los extremos de la red, es decir tanto en los inputs (entradas), como en los outputs (salidas). Las neuronas ocultas son las que generalmente se encuentran en el medio y son las encargadas de la detención de la representación interna de la información, estas neuronas no pueden tener conexiones directas con las neuronas visibles.

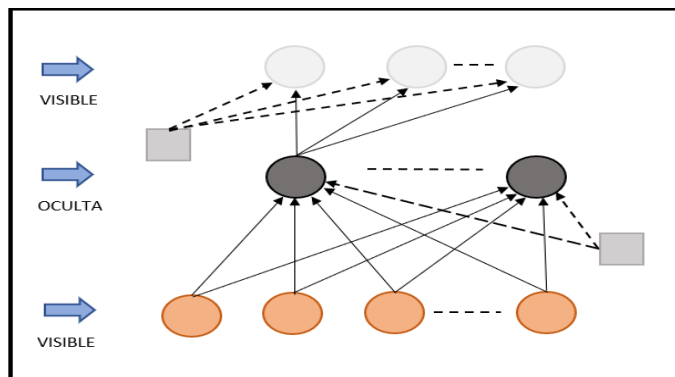


Figura 2.7 Modelo de las unidades de las neuronas o procesamiento

Fuente: (Torra Porrás, 2004)

- El tipo de conexiones tienen una gran incidencia en el establecimiento entre las unidades de procesamiento o neuronas. Estos modelos se pueden transferir en una única dirección descritos como feed-forward y los modelos que se propagan en diferentes direcciones, inclusive con procesos de retroalimentación denominados recurrentes.
- El aspecto de los modelos de aprendizaje puede clasificarse en dos tipos. Supervisado, es el que se encarga de adaptar las conexiones de cada neurona para obtener un resultado confiable y que se acerque lo más posible al resultado que queremos obtener, quien dirige este tipo de paradigma es denominado supervisor. Por otro lado, el aspecto No Supervisado, no consta de un supervisor y su finalidad es el establecimiento de la representación de las entradas (inputs) en una nueva dimensión con base estadístico (Torra Porrás, 2004).

2.7. Taxonomía de modelos de las redes neuronales

Debido a que hoy en día existe una variabilidad de modelos de redes neuronales artificiales, se da la necesidad de realizar una clasificación de acuerdo con los modelos que tengan similitud, se clasifican en tres grupos: en concordancia con el aprendizaje (learning paradigm), de cómo es su arquitectura (network architecture), y en función al área de las aplicaciones.

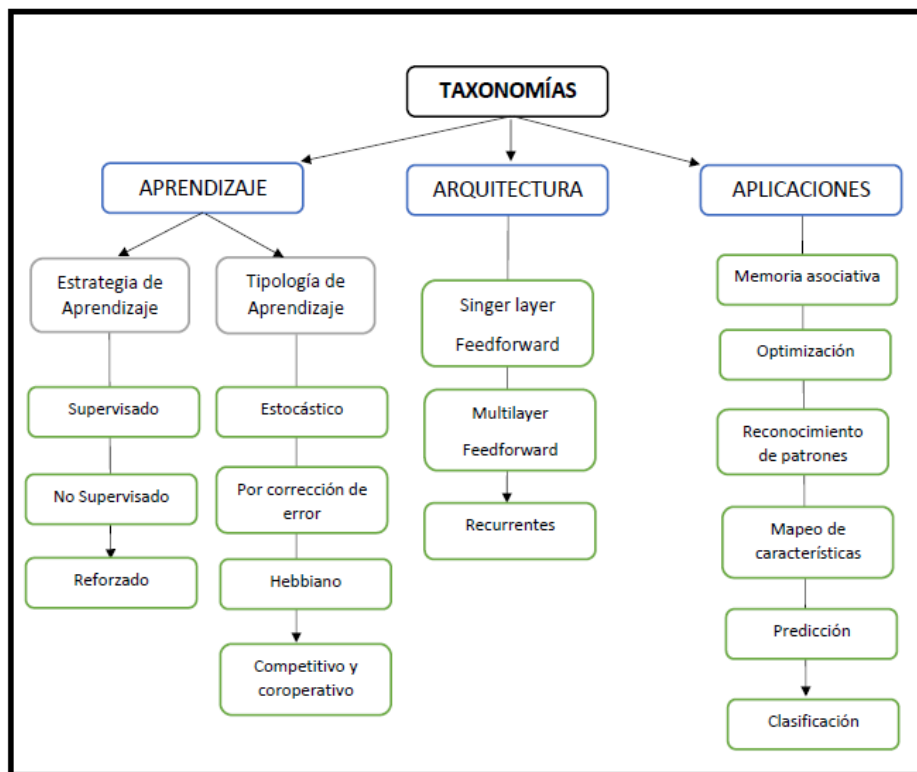


Figura 2.8 Taxonomías

Fuente: (Torra Porras, 2004)

2.8. Aprendizaje o Entrenamiento

Las redes neuronales artificiales cuentan con una serie de cualidades, de las cuales, la principal que se destaca es su competencia de poder aprender por medio de un conjunto de un modelo de entrenamiento, el cual que permite procesar la información hasta que pueda encontrar un modelo que se pueda acoplar a los datos entrenados. Este procedimiento se lo llama entrenamiento de la red y puede darse de dos formas, entrenamiento supervisado y entrenamiento no supervisado (Torra Porras, 2004).

2.8.1. Aprendizaje supervisado

Tiene como finalidad entrenar a la red neuronal por medio de un conjunto de datos de entrenamiento, los cuales constan de datos de entradas y salidas. El propósito del algoritmo de entrenamiento pretende adaptar la red, para que la salida producida por las redes neuronales

artificiales pueda ser la más cercana posible a los datos que se pretende obtener por medio de la información introducida en la entrada del modelo (Torra Porras, 2004).

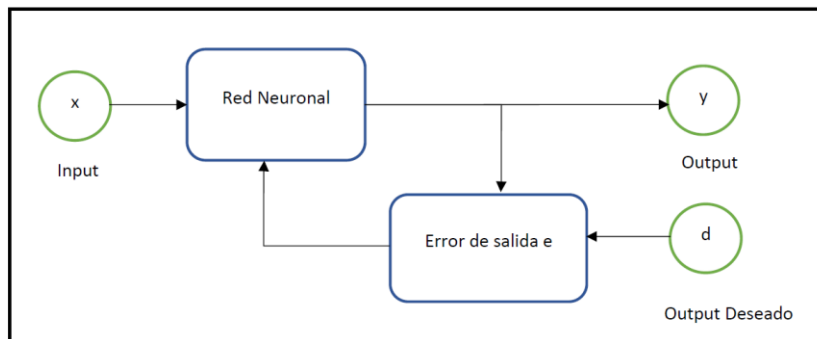


Figura 2.9 Aprendizaje supervisado

Fuente: (Torra Porras, 2004)

2.8.2. Aprendizaje reforzado

Se trata de un aprendizaje supervisado, su principal característica es la falta de disposición de un modelo completo del comportamiento que se pretende obtener, esto quiere decir que, en la fase del entrenamiento, la salida que se obtendrá de la red no va a ser la que realmente se desea.

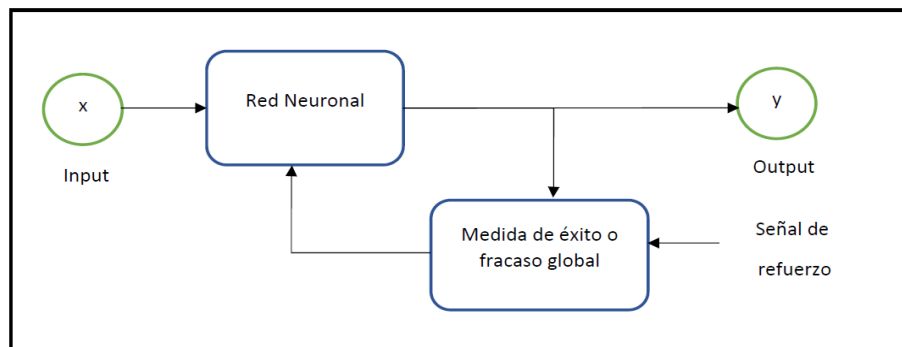


Figura 2.10 Aprendizaje reforzado

Fuente: (Torra Porras, 2004)

2.8.3. Aprendizaje no supervisado

Este tipo de aprendizaje cuenta solamente con una unidad de patrones a las redes neuronales artificiales. El propósito de este entrenamiento es adecuar el peso de la red, para que pueda encontrar un orden el cual le permita compilar toda la información (Torra Porras, 2004).

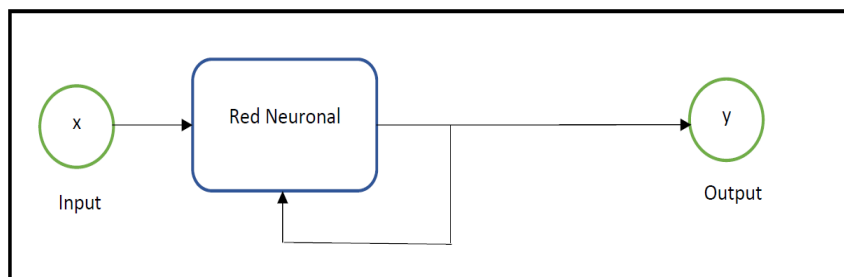


Figura 2.11 Aprendizaje supervisado

Fuente: (Torra Porras, 2004)

2.9. Arquitectura de una red neuronal

Las redes neuronales tienen un comportamiento de dos maneras el individual y el grupal, se puede estudiar las interconexiones que existen entre ellas. Se ha identificado que las neuronas se agrupan por capas, cada uno de estos conjuntos poseen una cantidad de neuronas con un comportamiento similar y una variabilidad en la cantidad.

Cada una de las capas que constituye la red neuronal está conectada a la inmediata posterior total o parcial, la última capa está excepta a ser conectada debido a que esta constituye el fin o la salida de la red neuronal. Están constituidas generalmente por tres capas, las cuales son:

- **Capa de entrada:** La cantidad y tipo de neuronas por la cual esta capa está constituida, depende específicamente de la información proporcionada por el problema.
- **Capa oculta o intermedia:** Aquí la cantidad de capas dependen de la complejidad del problema de la red, es decir puede ser más de una capa, en esta sección de capas se puede entrenar y obtener propiedades de adaptación, generalidad, extracción de diferentes características que permiten mejorar el modelamiento de la red neuronal.
- **Capa de salida final:** La configuración de esta capa depende de la forma en la que se espera la salida de la red.

Se pueden conectar, relacionar, distribuir cada una de las capas que a su vez están constituidas por el tipo de neuronas con sus mismas características. La arquitectura de las redes neuronales puede estar constituidas como:

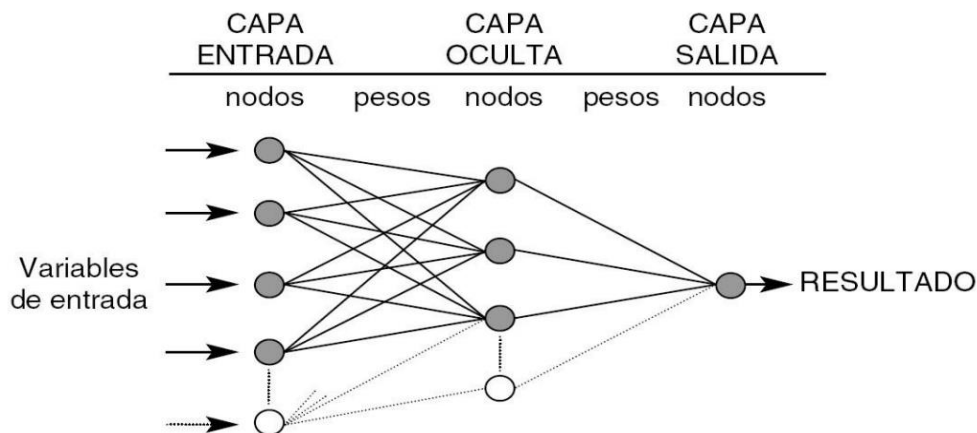


Figura 2.12 Arquitectura de una red neuronal

Fuente: (RODRÍGUEZ AEDO, 2016)

Las redes neuronales artificiales se ha determinado que se clasifican de dos formas como son: según el número de capas de neuronas en la red (consta de una sola capa o multicapa) y también puede clasificarse de acuerdo a los patrones de conexión que existe entre las neuronas y la dispersión de los datos. De acuerdo a esta última clasificación se puede limitar dos tipos de arquitectura.

- **Redes Feed-forward**

En esta clasificación de redes, la circulación de los datos desde la entrada hasta la salida es siempre en dirección hacia delante, ya que el procesamiento de la información puede presentarse en múltiples capas de las neuronas, pero nunca se va a retroceder. Es decir, las conexiones que se presentan desde la salida de una neurona a la entrada de otra en la misma capa o en la capa que la precede (Torra Porrás, 2004).

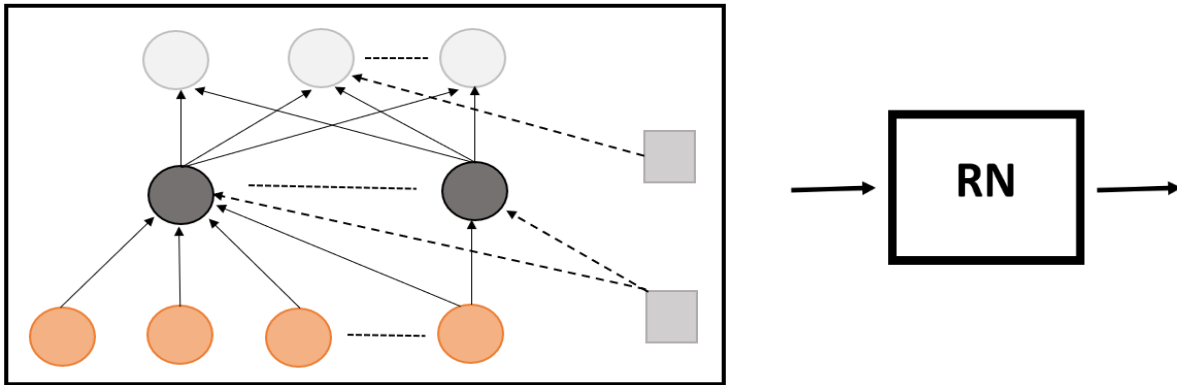


Figura 2.13 Redes Feed-forward

Fuente: (Torra Porras, 2004)

- **Redes Recurrentes**

Se asemeja a la red Feed-forward, pero la diferencia que existe entre estas dos es que las redes recurrentes si pueden tener conexiones hacia atrás.

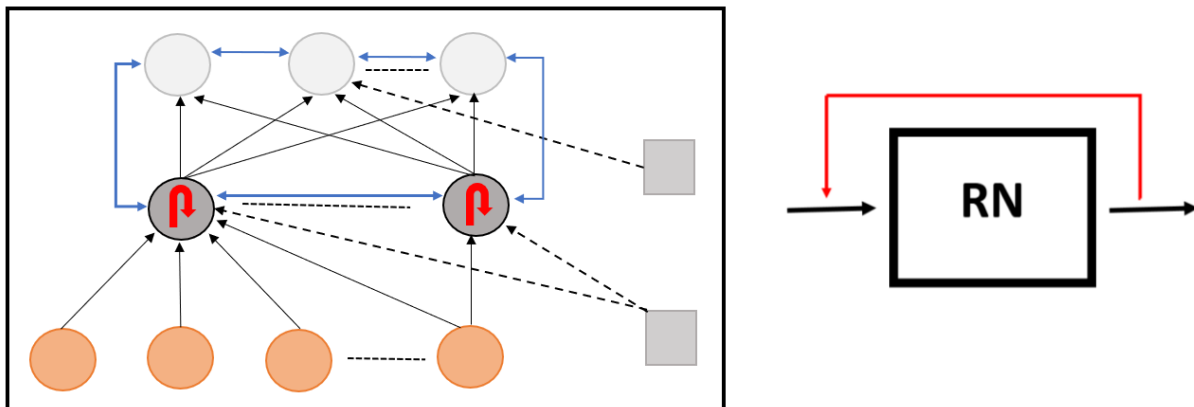


Figura 2.14 Redes recurrentes

Fuente: (Torra Porras, 2004)

- **Redes LSTM**

Son consideradas un tipo de redes especiales ya que pueden trabajar a corto y largo plazo, estas en particular su principal función es evitar las dependencias a largo plazo, no les cuesta recordar información durante extensos periodos de tiempo ya que esas redes fueron predeterminadas para

realizar lograr ese objetivo. A diferencia de una red estándar que sus módulos cuentan con una estructura básica, con una sola capa de tanh, las redes recurrentes en especial las LSTM cuentan con cuatro capas de redes neuronales las cuales se comportan de una manera bastante peculiar (Oinkina, 2015).

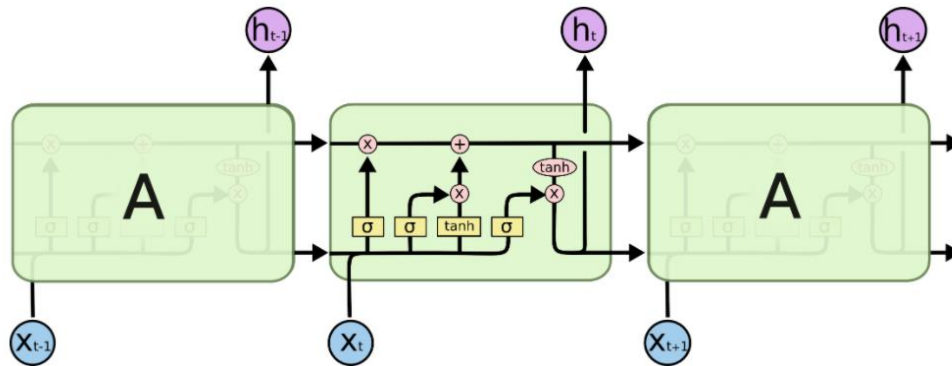


Figura 2.15 Redes LSTM

Fuente: (Oinkina, 2015)

- **Redes GRU**

Son consideradas redes recurrentes, también conocidas como redes con compuertas (Gated Recurrent Neural Networks). Se identifican por ser más eficientes en su ejecución, poseer menos parámetros y un costo computacional menor que las LSTM, obteniendo una mejora en la memoria a largo plazo. Se podría decir que estas redes están en medio de las RRN y las LSTM, debido a su fácil y rápida capacitación están siendo utilizadas cada vez más (Calvo, 2018).

Posee dos compuertas en la unidad de memoria por las cuales la información introducida es analizada y distribuida por todo el sistema:

- Puerta de actualización: Aquí se analiza y se establece la cantidad de información de las celdas anteriores que se debe conservar.
- Puerta de reajuste: Se identifica la mejor manera para introducir una nueva información al sistema y a su vez que esta se acople a las celdas con información establecida con anterioridad.

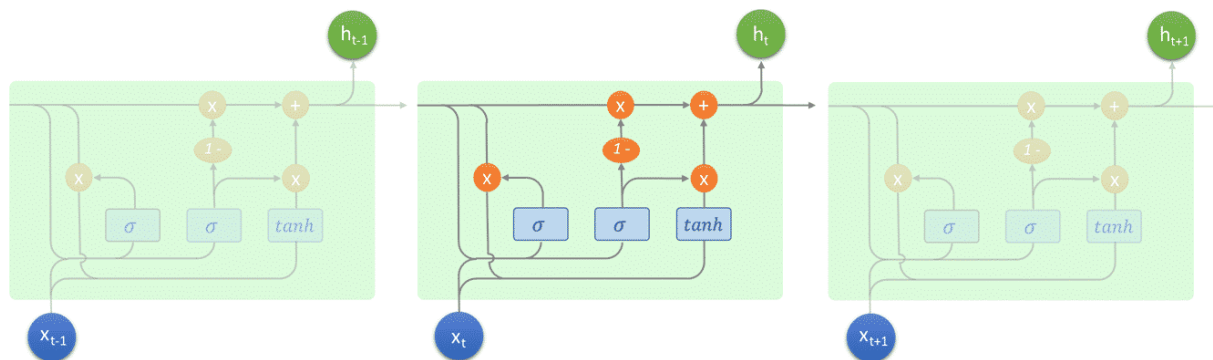


Figura 2.16 Redes GRU

Fuente: (Calvo, 2018)

2.10. Algoritmo de retropropagación (BP)

Es un algoritmo que se basa principalmente en el aprendizaje, el cual consiste en propagar el error producido y enviarlo hacia atrás, desde la salida hacia la entrada de la capa, esto permitirá que se puedan ajustar los pesos de los datos y reducir el error encontrado.

Como se lo menciono anteriormente la red “Backpropagation” consiste en el aprendizaje de una agrupación de pares de entrada y salida, para que se de el correcto funcionamiento utilizan un ciclo de propagación-adaptación de dos fases (RODRÍGUEZ AEDO, 2016).

- En la primera fase el patrón que va a servir de estímulo para la entrada de la capa de la red siempre va a ser hacia adelante, los datos introducidos van pasando capa por capa hasta llegar a la salida, obteniendo de esta manera el error al momento de la comparación entre la que se desea obtener.
- En la segunda fase, el error encontrado a la salida se transmite hacia atrás y pasa por todas las neuronas que se encuentran en las capas intermedias, las cuales tienen una incidencia en la salida. A estas neuronas se les otorga un porcentaje del error aproximado por la intervención en la primera salida.

2.11. R Studio

R Studio es una herramienta de gran utilidad en el área de la estadística, unas de las características principales es su codificación abierta, la gratuidad para su utilización, cuenta con una

gran capacidad de análisis de datos, gran poder para resolver cálculos estadísticos y con la ayuda de su editor de programas, tiene la capacidad de generar nuevas funciones y ampliar las actuales (Craveri, Kern, & Pacini, 2019).

Las ventajas de esta herramienta la han convertido en la opción más viable para el aprendizaje de la estadística y matemática. Gracias a que esta herramienta se ha vuelto conocida se han ejecutado interfaces gráficas de usuario GUI, la cual permite el mejoramiento de la interacción con el usuario.

Según (Ortiz) y (Salim, Gopal, & Ayub, 2018), R Studio es un (IDE) entorno de desarrollo integrado para el software R. Para la ejecución y correcta programación posee una consola, editor de resaltado de sintaxis que permite la ejecución directa de código, así como herramientas para el trazado, el historial, la depuración y la administración del espacio de trabajo. RStudio es un código abierto y comercial y puede ser utilizado en el escritorio de Windows, Mac y Linux o en un navegador conectado a RStudio Server o RStudio Server Pro (Debian / Ubuntu, RedHat / CentOS y SUSE Linux).

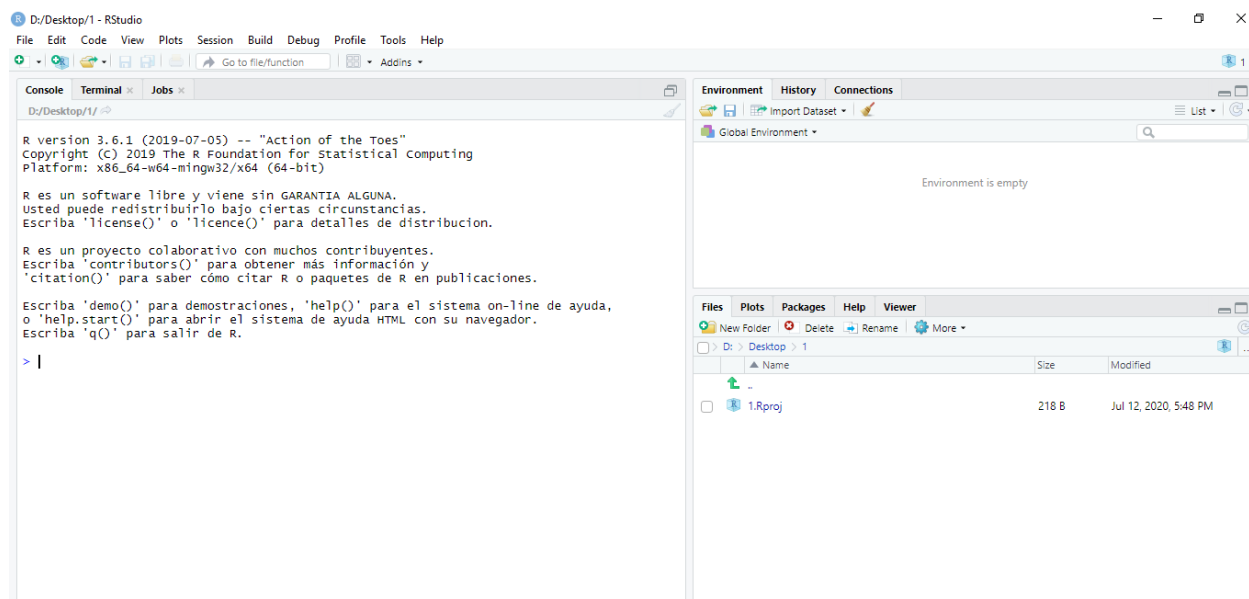


Figura 2.17 Software R Studio

Fuente: R Studio

CAPÍTULO III

3. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL

3.1. Descripción de la empresa

EMAPALA EP es una empresa eficiente creada el 31 de marzo del 2015 dedicada a la prestación de servicios básicos de Agua Potable y Alcantarillado, y en la recaudación de tasas retributivas de calidad y con estándares requeridos para la preservación de la salud de los habitantes del cantón Lago Agrio.

La empresa suministra estos servicios a diferentes sectores del cantón como son: centro poblado de Nueva Loja, recinto Puerto Rico, Recinto San Vicente, centro poblado de la parroquia Pacayacu, comunidad de La Granito, Y de Harberth, Primero de Mayo, Los Laureles, Virgen del Rosario, centro poblado de la parroquia Dureno, recintos de la parroquia El Eno, Aukayacu, Amazonas, San Bartolo, recintos de la parroquia General Farfán, 5 de Agosto, Corazón Orense, Trampolín del Triunfo, Puerto Ecuador, Puerto Camacho, 9 de Marzo, Nuevo Mundo y los centros binacionales de atención fronterizo.

Para la comodidad de los pobladores del cantón, EMAPALA EP ha instalado puntos específicos de cobro de los servicios suministrados como son: edificio principal de EMAPALA EP, edificio municipal del cantón Lago Agrio, Red Facilito, parroquia Pacayacu y la parroquia de Dureno. Permitiendo que lo recaudado sea utilizado para el mejoramiento de los servicios y por consecuente la satisfacción de los contribuyentes.

La empresa se rige a las leyes y normativas vigentes a nivel nacional, a su vez trabaja con las entidades gubernamentales como son: La secretaría del Agua, la Agencia Nacional de Regulación, Control y Vigilancia Sanitaria, Agencia de Regulación y Control del Agua, Asociación de Municipalidades Ecuatorianas y la Empresa Municipal de Agua Potable y Alcantarillado de Ibarra.

3.2. Localización de la empresa

La Empresa Municipal de Agua Potable y Alcantarillado - EMAPALA EP, para la comodidad de los usuarios sus oficinas se encuentran ubicadas en el Cantón Lago Agrio en las calles Av. Quito y 20 de junio, junto al parque recreativo Nueva Loja, como se muestra en la figura 3.1.

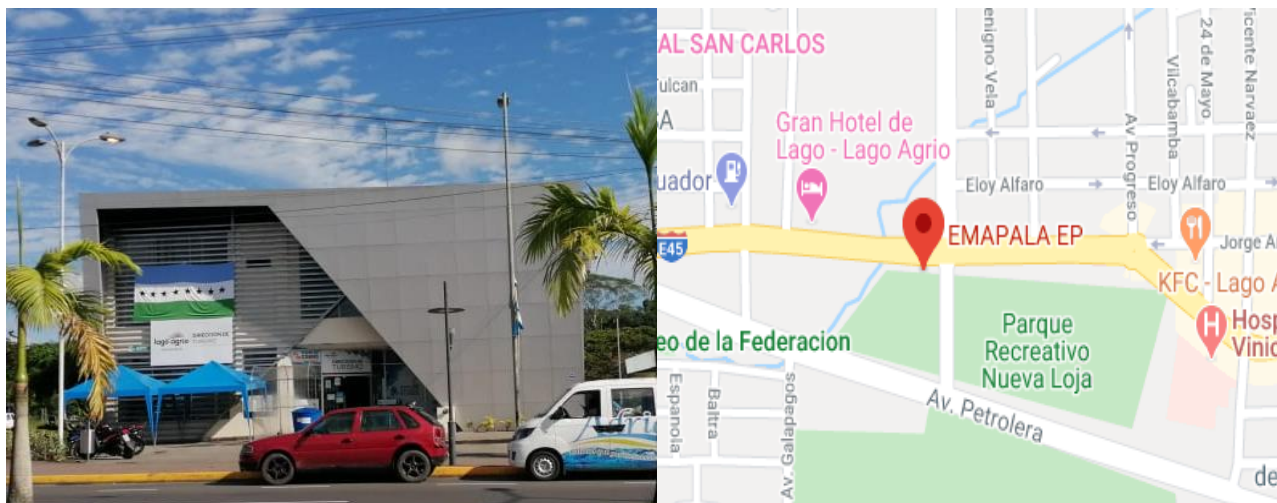


Figura 3.1 Localización de la empresa EMAPALA EP

Fuente: Propia, Google Maps

3.3. Misión

La Empresa Pública Municipal de Agua Potable y Alcantarillado de Lago Agrio tiene como finalidad servir a la colectividad en el abastecimiento de Agua Potable y Alcantarillado del cantón, preservando la salud de la población y el medio ambiente.

3.4. Visión

Ser una empresa eficiente y eficaz, que genere rentabilidad social en sus inversiones, capaz de dotar de un servicio de agua potable y alcantarillado continuo, oportuno y de calidad a beneficio del Cantón Lago Agrio.

3.5. Objetivos Estratégicos

Comunidad:

- Asegurar la eficiencia (continuidad, cobertura y calidad) en la prestación de los servicios públicos operados por la Empresa teniendo en cuenta la sostenibilidad ambiental.

Financiera:

- Asegurar la disponibilidad de ingresos mediante la gestión efectiva de los procesos y logro de otras fuentes de financiación institucional.

Formación y Crecimiento:

- Contar con personal de servidores y obreros idóneos y capacitados que cumplan los perfiles asignados a los puestos institucionales con un alto sentido de pertenencia, servicio y valoración a esta empresa.

3.6. Líneas Estratégicas

Las líneas estratégicas que guiarán en gran medida en todas las acciones que se llevarán a cabo dentro de la planeación estratégica y toma de decisiones de la Empresa Pública Municipal de Agua Potable y Alcantarillado de Lago Agrio, son:

1. **Sostenibilidad financiera:** Administrar eficientemente los recursos para alcanzar estándares institucionales positivos.
2. **Mejoramiento comercial:** Incrementar el número de conexiones con equipos de medición, a fin de recaudar eficazmente la facturación de los servicios.
3. **Excelencia operacional:** Ejecutar impecablemente los procesos de la cadena de valor para proveer agua potable de calidad continuamente.
4. **Planificación empresarial:** Incorporar una planificación estratégica con proyectos de corto y largo plazo que integre, coordine y controle la gestión institucional de la empresa.
5. **Responsabilidad Social:** Gestionar los riesgos ambientales, de salud y de seguridad demostrando preocupación por la comunidad y su entorno.
6. **Compromiso con los empleados y trabajadores:** Contar con el número óptimo de funcionarios y trabajadores, emprendiendo el mejoramiento de sus competencias y dotarles de las herramientas necesarias para su correcto desempeño laboral.

3.7. Organigrama

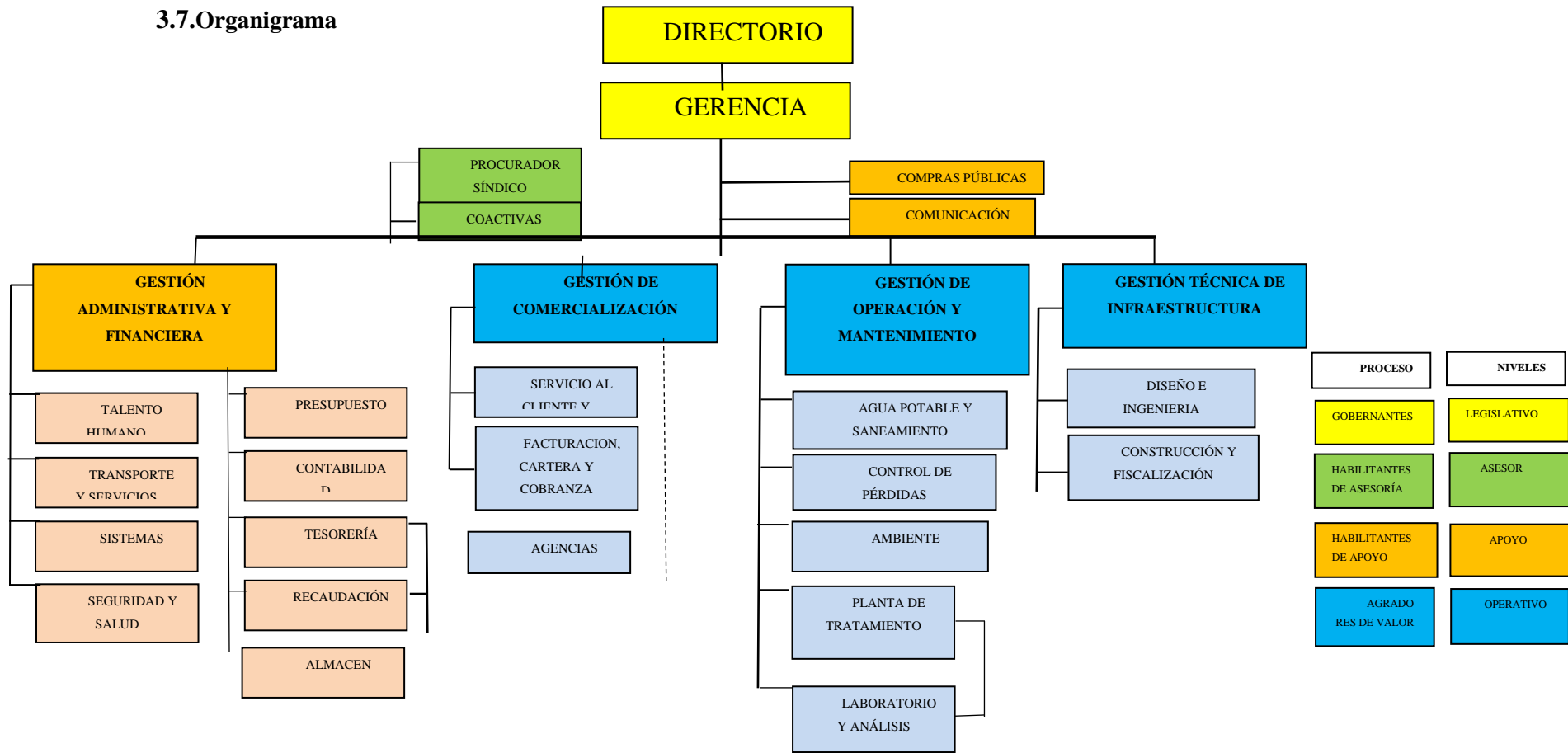


Figura 3.2 Estructura Orgánica de EMAPALA

Fuente: EMAPALA EP

3.8. Descripción del Proceso Productivo

Zona de Captación

Para la obtención del agua de su respectivo tratamiento y distribución a toda la población cuentan con una barcaza flotante con 3 unidades de bombeo, cada una con capacidad de 150 litros /s, poseen un motor de 200 hp c/u controlados con variadores de frecuencia con arranque suave de 200 hp que permiten regular el caudal que llega a la planta, generalmente funcionan 2 unidades de bombeo y 1 unidad queda en backup. Actualmente la capacidad de bombeo a la planta desde la captación es de 200 litros/s.

Zona de Tratamiento y Distribución

El recurso hídrico para ser apta para el consumo de las personas debe pasar por una serie de procesos que le permitirán contar con los estándares de calidad requeridos para su distribución.

El primer proceso denominado Aireación, se lo realiza por medio de 4 bandejas de oxigenación que permiten separar las partículas de hierro del agua, este proceso se lo realiza comúnmente cuando se tiene las captaciones de aguas subterráneas. Este sistema de oxigenación en la planta de agua de EMAPALA EP ayuda en una pequeña parte a la potabilización del recurso hídrico debido a que la empresa posee una zona de captación de agua superficial.

Para el control de la cantidad de agua cruda que ingresa a la planta poseen un Flujoímetro que les indica los litros/s en tiempo real que llega desde la captación, esta a su vez tiene la capacidad de 300 l/s, pero la planta tiene un soporte máximo de 240 l/s. Este instrumento a su vez cuenta con un Totalizador que se encarga de contabilizar los m³ de agua que ingresa a la planta desde que se instaló.

El siguiente proceso la Cámara de Mezcla es el primer punto en el cual que se introduce al agua el químico denominado Policloruro de Aluminio, permitiendo separar las partículas de lodo del agua. Luego el agua pasa a su segundo punto de aplicación de químico denominado Floculación, la sustancia introducida denominada Nalco que ayuda al Policloruro hacer los flops del agua que al igual que el primer químico introducido ayuda a la separación de partículas de lodo del agua en proceso de tratamiento.

El proceso para que el químico de peso al Flop y este caiga al fondo de las rejillas (Galletas) se llama Sedimentación, ayudando al recurso hídrico a ingresar por unos orificios que permiten el traslado del agua hacia los filtros.

Realizado el respectivo filtrado pasa a una cámara en la cual se introduce el tercer químico llamado Cloro Gas, luego de que se realiza dicha mezcla pasa sus 3 tanques de almacenamiento cada uno con capacidad de 1000 m³ para su respectiva distribución. Para garantizar que el agua potable llegue a todos los sectores, posee 7 unidades de bombeo de 100 l/s c/u, permitiéndole a la empresa tener un continuo abastecimiento.

La empresa EMAPALA para poder llevar un registro y control de los sistemas cuenta con un software llamado SCADA, que permite controlar y supervisar los procesos desde la zona de captación hasta la distribución, su centro de mando se encuentra en la planta de Agua, de allí se controla el arranque, la regulación de frecuencia de los motores, almacenamiento de datos históricos del nivel del río y del caudal, datos históricos de la corriente y voltaje de los motores.

3.9. Producción de la empresa y consumo anual

La Empresa Municipal de Agua Potable y Alcantarillado de Lago Agrio cuenta con dos plantas de agua potable, las que producen 300 l/s y 100 l/s. La empresa EMAPALA EP para el correcto seguimiento y control de la información, llevan a cabo un registro mensual del consumo de agua potable y sus clientes, permitiendo identificar el comportamiento que ha tenido la producción a lo largo del año. A continuación, en la tabla 3.1 y 3.2, se presenta el consumo mensual del servicio hídrico potable en m³ de los 3 últimos años, la producción, los índices de agua no contabilizada y sus clientes.

MESES	PLANTA ANTIGUA Y NUEVA 2017		IANC	PLANTA ANTIGUA 2018		PLANTA NUEVA 2018		TOTAL 2018		IANC
	m ³ /MES	LT/S	%	m ³ /MES	LT/S	m ³ /MES	LT/S	m ³ /MES	LT/S	%
ENERO	525373	202,3	59,9	51840	20	541498	208,5	593338	228,4	43,18
FEBRERO	407242	156,8	41,42	98632	38	417544	160,8	516176	198,7	47,56
MARZO	516118	198,7	54,45	118963	45,8	480236	184,9	599199	230,7	60,69
ABRIL	473260	182,2	49,81	140789	54,2	480897	185,1	621686	239,3	58,48
MAYO	527147	203	55,56	114048	43,9	450789	173,6	564837	217,5	50,26
JUNIO	523994	201,7	51,57	121824	46,9	445896	171,7	567720	218,6	50,62
JULIO	528208	203,4	56,31	111456	42,9	430897	165,9	542353	208,8	52,73
AGOSTO	543014	209,1	57,75	120896	46,6	456887	176,3	577783	222,9	47,26
SEPTIEMBRE	510551	196,6	43,29	128784	49,7	457124	176,4	585908	226	51,34
OCTUBRE	526748	202,8	57,18	119789	46,2	446789	172,4	566578	218,6	39,21
NOVIEMBRE	459938	177,1	50,92	118978	45,9	439871	169,7	558849	215,6	51,99
DICIEMBRE	549369	211,5	56,15	125734	48,5	426987	164,7	552721	213,2	49,03

Tabla 3.1 Producción de agua potable 2017-2018

Fuente: EMAPALA EP

MES	PLANTA ANTIGUA 2019		PLANTA NUEVA 2019		TOTAL 2019		IANC
	m ³ /MES	LT/S	m ³ /MES	LT/S	m ³ /MES	LT/S	%
ENERO	151200	58,3	479120	184,8	630320	243,2	52,02
FEBRERO	149760	57,8	426628	164,6	576388	222,4	61,43
MARZO	147600	56,9	469570	181,2	617170	238,1	46,44
ABRIL	144720	55,8	463948	179	608668	234,8	48,3
MAYO	146160	56,4	478134	184,5	624294	240,9	51,06
JUNIO	147600	56,9	474086	182,9	621686	239,8	58
JULIO	146354	56,5	509812	196,7	656166	253,1	51,06
AGOSTO	138650	53,5	495230	191,1	633880	244,6	53,81
SEPTIEMBRE	137725	53,1	504020	194,5	641745	247,6	53,32
OCTUBRE	145530	56,1	477838,66	184,4	623368,54	240,5	43,32
NOVIEMBRE	134694	52	486325	187,6	621019	239,6	44,15
DICIEMBRE	140112	54,1	489500	188,8	629612	242,9	49,36

Tabla 3.1 Producción de agua potable 2019

Fuente: EMAPALA EP

3.10. Datos de la población

La población del cantón Lago Agrio con el pasar de los años se ha incrementado continuamente en las zonas rurales y urbanas, según las proyecciones presentadas por el Instituto de Nacional de Estadísticas y Censos, para el año 2020 la ciudad contará con 119594 habitantes, a continuación, en la tabla 5, se presenta la proyección de habitantes por clasificación de edad desde el 2017 hasta el 2020.

Cantón	GRUPOS DE EDADES	2017	2018	2019	2020
LAGO AGRIO	< 1 año	2.364	2.373	2.380	2.387
	1 - 4	9.988	10.028	10.065	10.101
	5 - 9	12.586	12.652	12.714	12.773
	10 - 14	12.423	12.540	12.640	12.728
	15 - 19	11.452	11.678	11.883	12.063
	20 - 24	10.007	10.256	10.510	10.762

25 - 29	9.216	9.418	9.630	9.855
30 - 34	8.733	8.950	9.161	9.368
35 - 39	7.549	7.790	8.029	8.262
40 - 44	6.512	6.740	6.976	7.219
45 - 49	5.542	5.737	5.937	6.146
50 - 54	4.587	4.768	4.950	5.129
55 - 59	3.611	3.793	3.976	4.161
60 - 64	2.654	2.801	2.958	3.122
65 - 69	1.844	1.937	2.037	2.146
70 - 74	1.312	1.369	1.428	1.491
75 - 79	905	955	1.004	1.053
80 y Más	669	717	770	828
TOTAL	111.956	114.503	117.048	119.594

Tabla 3.3 Proyección poblacional del cantón Lago Agrio 2017-2020

Fuente: (SNI, 2017)

3.11. Volumen de almacenamiento

El recurso hídrico ingresado por todos los procesos obligatorios para su potabilización es llevado a la zona de almacenaje, la planta de agua potable cuenta con 3 tanques de almacenamiento, cada uno con una capacidad de 1000 m³.

Por medio de 7 unidades de bombeo de 100 litros/s c/u, la planta distribuye a los distintos sistemas localizados alrededor de toda la ciudad como es el sistema Aeropuerto, sistema San Pedro, Sistema Venezuela, además abastece a los tanqueros, tanques elevados y la mayor parte a la red principal. El agua potable también es utilizada para uno de los procesos más importantes en el mantenimiento como es el retro lavado, que permite ingresar en el sistema agua a contraflujo logrando la limpieza de las distintas partículas de suciedad que se presentan en el sistema de potabilización.

3.12. Plantas de tratamiento

Según ("I," 2014), Título II Recursos Hídricos, en el capítulo I, Definición, Infraestructura y Clasificación de los recursos Hídricos. Art. 11 Establece como infraestructuras hidráulicas a todas las obras realizadas para la intervención en todos los procesos para el mejoramiento y tratamiento de agua, como es la captación del recurso hídrico, el control, regulación, almacenamiento y

distribución. Las empresas encargadas de este servicio básico ya sean públicas como privadas deben acatarse a esta ley, ya que el interés de la correcta utilización del líquido vital es de carácter comunitario.

EMAPALA EP cuenta con dos plantas de tratamiento, la planta Antigua con una capacidad de 100 l/s y la planta Nueva con una capacidad de 300 l/s, ambas utilizadas para el mejoramiento del agua apta para el consumo humano, se encuentran ubicadas en la Vía Aguarico, frente al río utilizado para la captación del recurso hídrico.

3.13. Análisis de autocorrelación de los sistemas de agua potable

La revisión de los patrones se desarrolla en dependencia de tres puntos específicos en el seguimiento del agua potable: producción general mensual, la distribución por cada sistema o red y el consumo o demanda facturado. Cabe recalcar que se realiza un análisis de la producción y distribución debido a que permite verificar si la cantidad de consumo de agua potable por parte de la población ha aumentado o disminuido.

Para el desarrollo de la autocorrelación de las series presentadas se realiza con la ayuda del software IBM SPSS, en el cual se establece la cantidad de datos que se va a analizar, en este caso son 36 datos de la demanda de agua potable de los años 2017 al 2019. Se identifica la variable, se selecciona la opción de predicciones- autocorrelaciones, se escoge la base de datos a examinar y se procede a ejecutar.

Las autocorrelaciones dan a conocer si las variables analizadas poseen tendencia ya sea positiva, negativa o si son estacionarias, esto se lo logra a través de los retardos, ya que permiten verificar el grado de afectación que tiene el tiempo (t) por un tiempo (t-1). Si el valor de autocorrelación no disminuye y no adopta un valor de cero en el retardo 4, la serie de tiempo posee tendencia, caso contrario es estacionaria.

3.13.1. Análisis de autocorrelación de la producción mensual de agua potable

En el análisis de la autocorrelación se identifica que la serie de tiempo de la producción general de agua potable posee una tendencia creciente significativa, debido a que los valores obtenidos de la autocorrelación son lejanos a cero en los primeros retardos de la serie, se lo observa en la figura 3.3 Análisis de auto correlación de la producción general de agua potable.

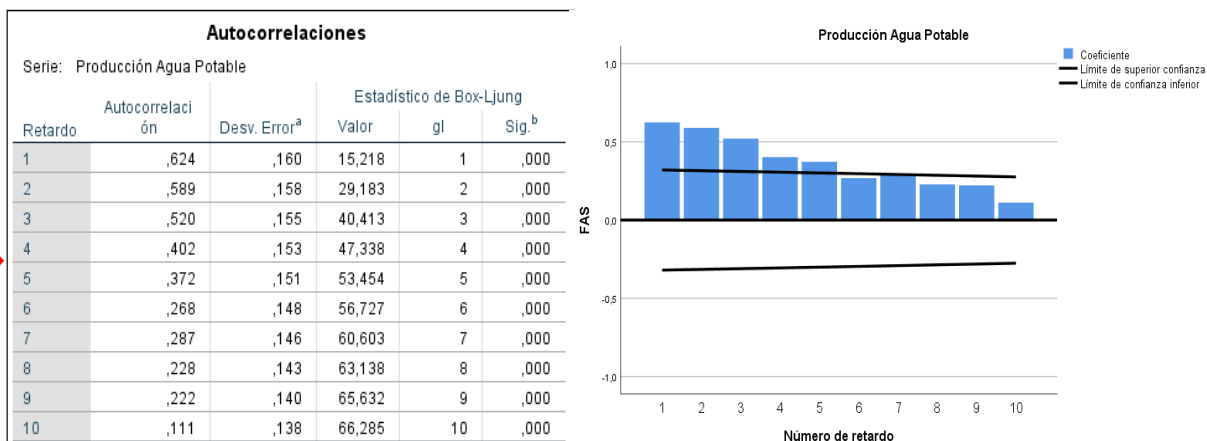


Figura 3.3 Autocorrelación de la Producción de agua potable

Fuente: IMB SPSS

3.13.2. Análisis de autocorrelación de la distribución de agua potable

El análisis de la serie de tiempo se analiza de manera independiente por cada sistema que la empresa distribuye. En el análisis desarrollado para la Red Principal se identifica una tendencia positiva debido a que se ha venido incrementando la demanda en este sector de distribución, su coeficiente de correlación en el cuarto retardo es (0,402), de la misma manera para el sistema Aeropuerto, se verifica que posee una tendencia creciente, su coeficiente de correlación es (0,438). A continuación, en las figuras 3.4 y 3.5 se observa la autocorrelación de los dos sistemas Red Principal y Aeropuerto.

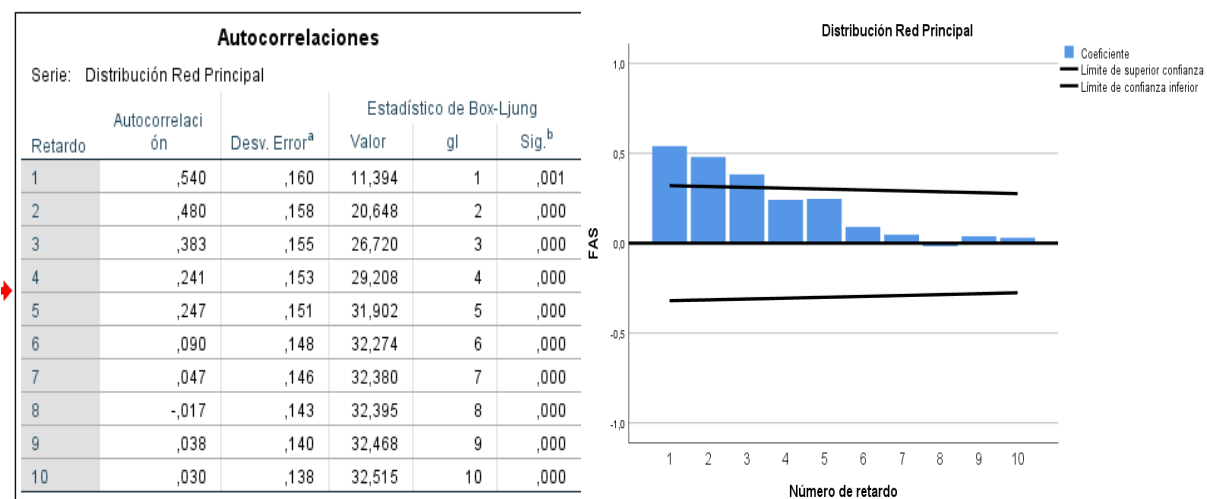


Figura 3.4 Autocorrelación de la distribución de la Red Principal de agua potable

Fuente: IMB SPSS

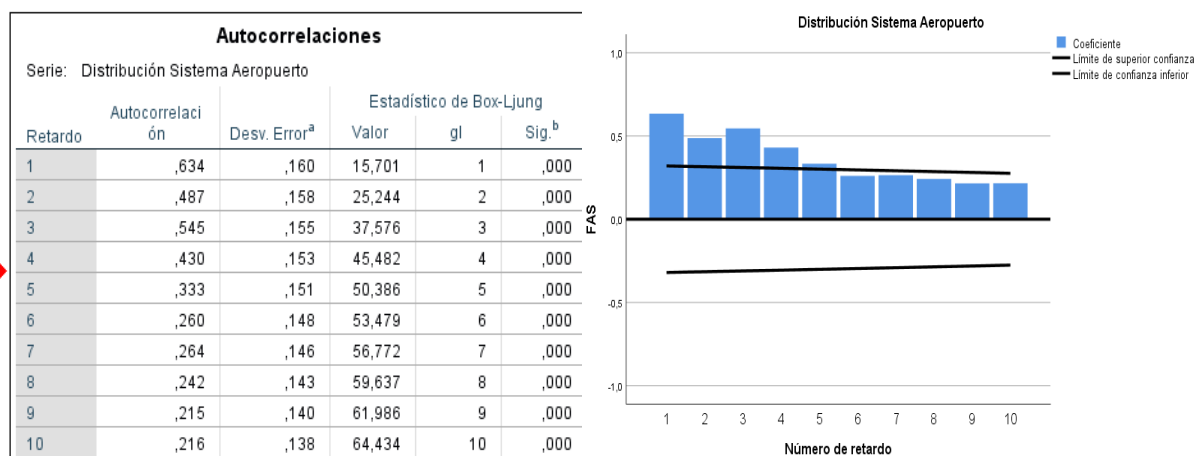


Figura 3.5 Autocorrelación de la distribución del Sistema Aeropuerto

Fuente: IMB SPSS

En la red de distribución San Pedro, presenta una tendencia creciente, su coeficiente de correlación en el cuarto retardo es de (0,295). Para el sistema Venezuela, se puede observar que la información procesada de los años 2017-2018-2019, ha generado una correlación en el cuarto retardo de (0,431), obteniendo una tendencia creciente para las dos series de tiempo. A continuación, se observa la autocorrelación en las figuras 3.6 y 3.7, distribución del sistema San Pedro y Venezuela.

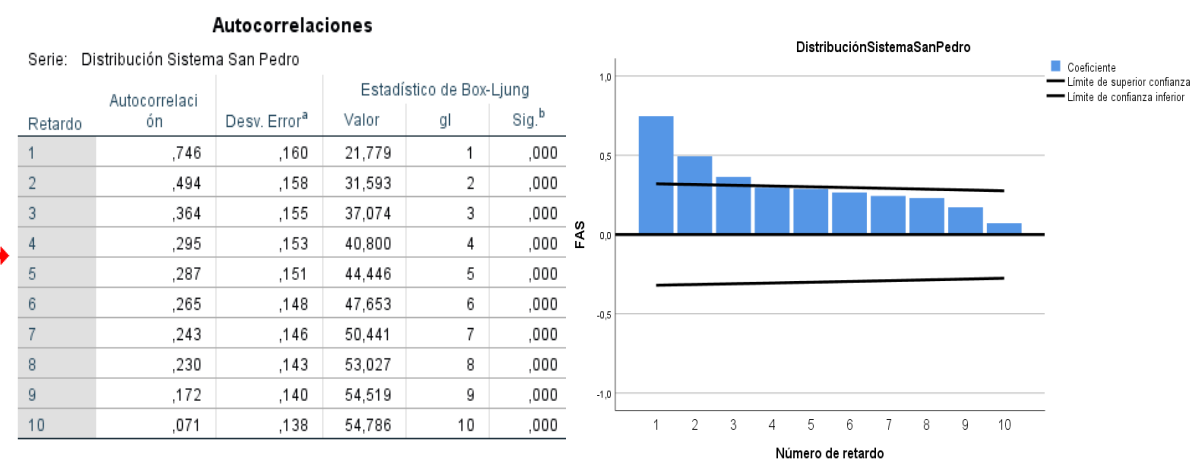


Figura 3.6 Autocorrelación de la distribución del Sistema San Pedro

Fuente: IMB SPSS

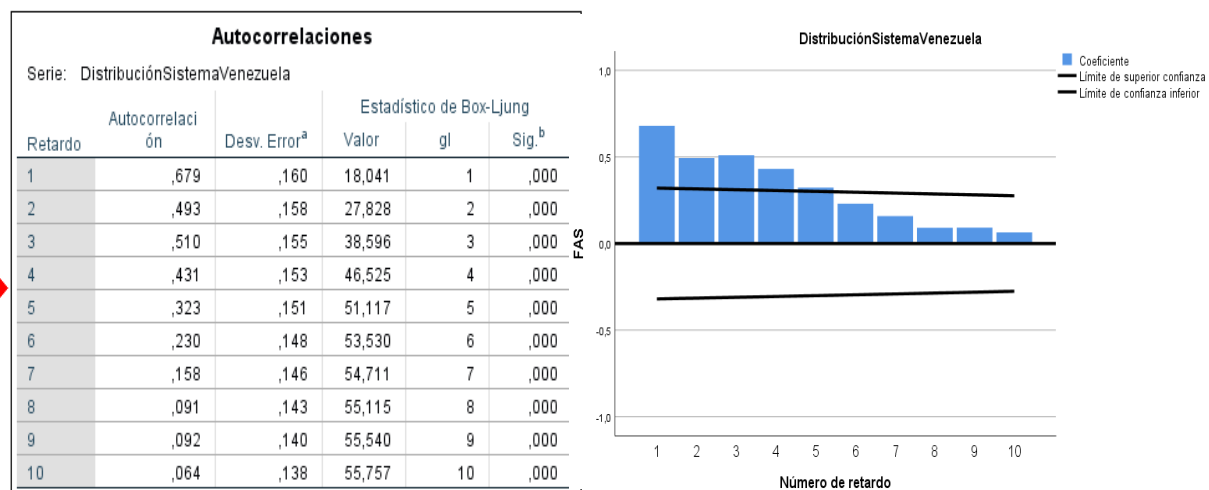


Figura 3.7 Autocorrelación de la distribución del Sistema Venezuela

Fuente: IMB SPSS

En el sistema Tanques Elevados se analiza la autocorrelación con un retardo de 16 en la cual se identifica que antes del retardo 4 el valor es (0,291), no llega o asemeja a cero, dando como resultado una serie de tiempo con tendencia negativa ya que se ha venido reduciendo la cantidad de metros cúbicos distribuida por los 3 tanques que contemplan este sistema tal como se muestra en la figura 3.8, distribución de los Tanques elevados.

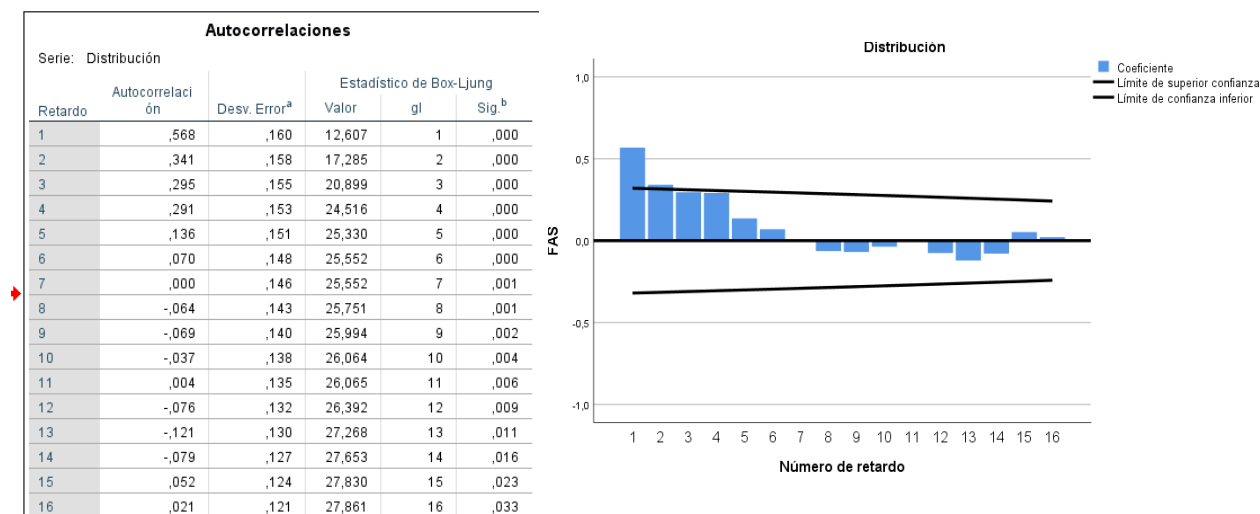


Figura 3.8 Autocorrelación de la distribución del Sistema de los Tanques Elevados

Fuente: IMB SPSS

Para el sistema de distribución de los Tanqueros además del análisis de autocorrelación realizado por el software antes mencionado, se utiliza el software EViews juntamente con la prueba de Dickey Fuller. En la cual se puede identificar que la serie de tiempo es estacionaria. Esto se lo puede verificar con el análisis de hipótesis que el programa previamente lo ejecuta, dando como resultado tanto en el nivel inicial como en la primera diferencia una estacionariedad en los datos de la distribución de Tanqueros.

$H_0 =$ La serie de distribución de tanqueros no es estacionaria (presenta raíz unitaria)

$H_i =$ La serie de distribución de tanqueros es estacionaria

El resultado de la prueba de hipótesis se identifica que las probabilidades a nivel como en la primera diferencia es 0,010 y 0,000, menores que el nivel de significación prefijado el cual es del 5% o 0,05. El t-estadístico de Dickey-Fuller se encuentra en la zona de Rechazo y se puede concluir con seguridad que se rechaza la hipótesis nula, dando como estacionaria la serie de tiempo estudiado. A continuación, en las figuras 3.9, 3.10 y 3.11 se observa el análisis del sistema de distribución de los Tanqueros.

Null Hypothesis: DISTRIBUCION has a unit root Exogenous: Constant Lag Length: 0 (Automatic – based on SIC, maxlag=9)				
		t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic		-4.516916	0.0010	
Test critical values:				
	1% level	-3.632900		
	5% level	-2.948404		
	10% level	-2.612874		
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(DISTRIBUCION) Method: Least Squares Date: 08/13/20 Time: 17:34 Sample (adjusted): 2017M02 2019M12 Included observations: 35 after adjustments				
Variable	Coefficien t	Std. Error	t-Statistic	Prob.
DISTRIBUCION (-1)	0.764943	0.169351	-4.516916	0.0001
C	3455.124	773.8867	4.464638	0.0001

R-squared	0.382052	Mean dependent var	6.657143
Adjusted R-squared	0.363326	S.D. dependent var	796.006
S.E. of regression	635.1487	Akaike info criterion	15.8010
Sum squared resid	13312656	Schwarz criterion	4
Log likelihood	274.5182	Hannan-Quinn criter.	15.8899
F-statistic	20.40253	Durbin-Watson stat	2
Prob(F-statistic)	0.000076		1.95279

Figura 3.9 Prueba Dickey Fuller en el Nivel del Sistema de los Tanqueros

Fuente: EViews

Null Hypothesis: D(DISTRIBUCION) has a unit root
 Exogenous: Constant
 Lag Length: 2 (Automatic – based on SIC, maxlag=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.779390	0.0000
Test critical values:		
1% level	-3.653730	
5% level	-2.957110	
10% level	-2.617434	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(DISTRIBUCION,2)
 Method: Least Squares
 Date: 08/13/20 Time: 17:40
 Sample (adjusted): 2017M05 2019M12
 Included observations: 32 after adjustments

Variable	Coefficien t	Std. Error	t-Statistic	Prob.
D (DISTRIBUCION (-1))	2.393813	0.414198	-5.779390	0.0000
D (DISTRIBUCION (-1),2)	0.839660	0.302725	2.773668	0.0098
D(DISTRIBUCION (-2),2)	0.345926	0.176603	1.958781	0.0602
C	17.24957	128.2372	-0.134513	0.8940

R-squared	0.740004	Mean dependent var	16.43750
Adjusted R-squared	0.712147	S.D. dependent var	1352.06

S.E. of regression	725.4106	Akaike info criterion	16.1278
Sum squared resid	14734173	Schwarz criterion	16.3110
Log likelihood	254.0451	Hannan-Quinn criter.	16.1885
F-statistic	26.56459	Durbin-Watson stat	1.97590
Prob(F-statistic)	0.000000		3

Figura 3.10 Prueba Dickey Fuller en la primera Diferencia del Sistema de los Tanqueros

Fuente: EViews

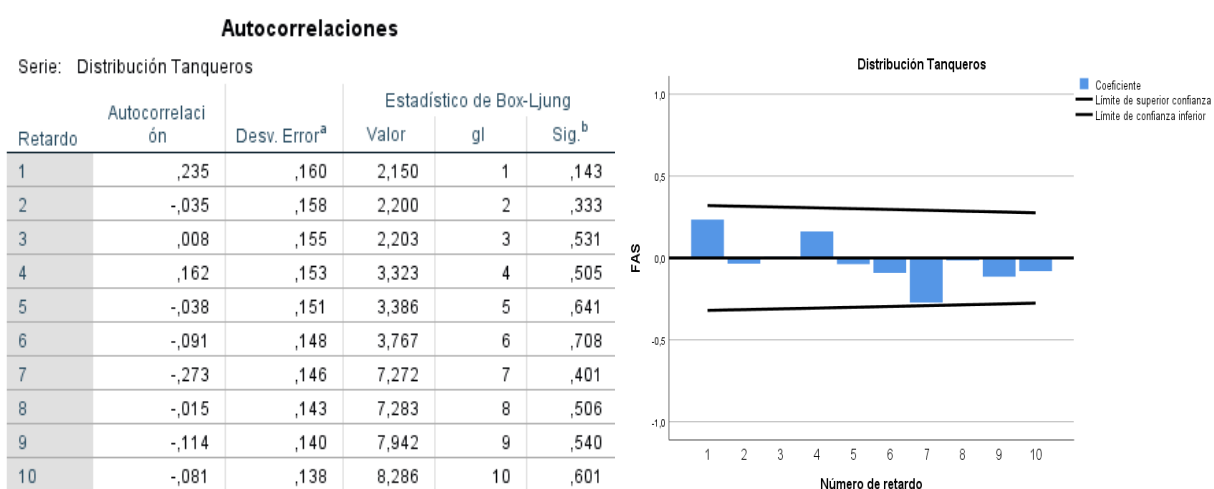


Figura 3.11 Autocorrelación de la distribución del Sistema de los Tanqueros

Fuente: IMB SPSS

3.13.3. Análisis de autocorrelación de los clientes mensuales

De la misma manera para el análisis de los clientes mensuales de los últimos tres años se utiliza el software ya antes mencionado, el análisis de autocorrelación proporciona un resultado en los cuales los históricos presentados han venido obteniendo una tendencia creciente, el valor de autocorrelación hasta el retardo 4 es (0,687), valor diferente o cerca de cero. Esto justifica que efectivamente el número de clientes que se benefician del recurso hídrico potable han venido incrementándose año tras año. A continuación, se muestra en la figura 3.12 el análisis de los clientes mensuales.

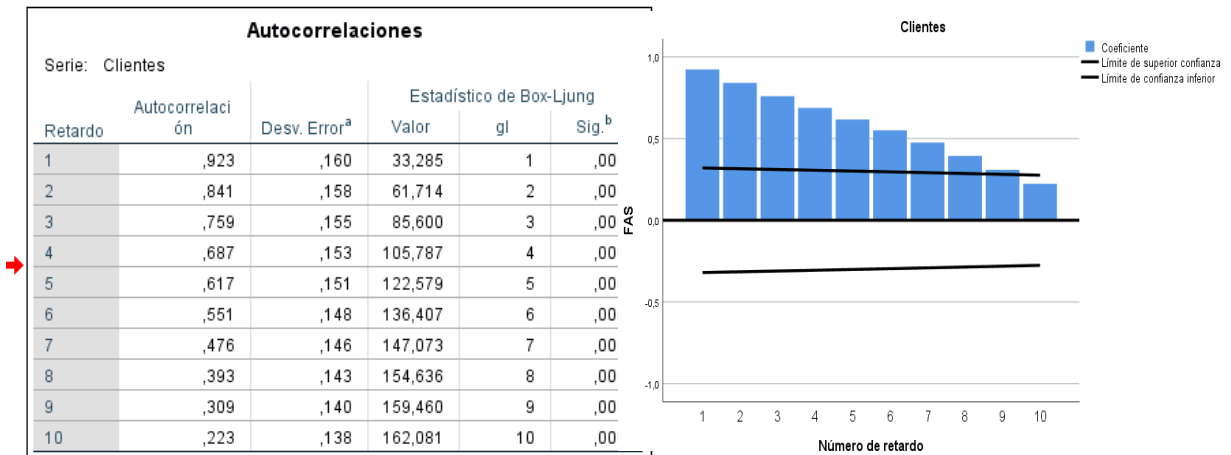


Figura 3.12 Autocorrelación de los clientes mensuales de agua potable

Fuente: IMB SPSS

3.13.4. Análisis de la serie de tiempo de la demanda de agua potable

Para la demanda de agua potable por parte de los clientes, se identifica que la serie de tiempo cuenta con una tendencia positiva, para corroborar la información se analiza los retardos de la serie de tiempo y se observa que hasta el retardo 4 el coeficiente es de (0,414), como se observa la autocorrelación en la figura 3.13, de la demanda de agua potable.

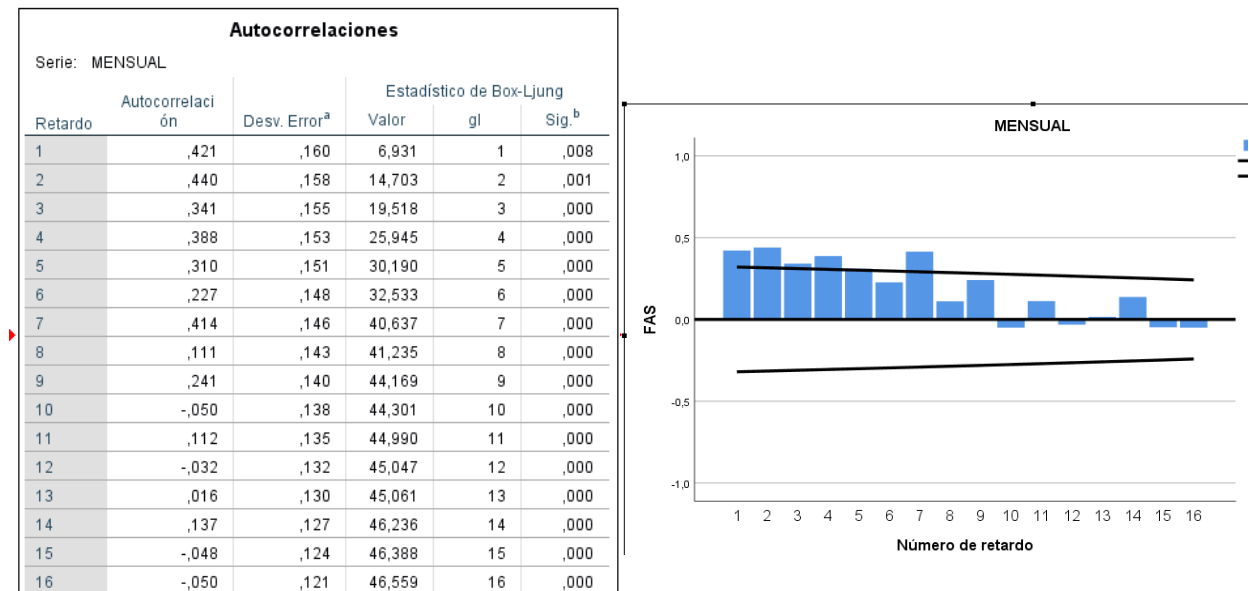


Figura 3.13 Autocorrelación de la demanda de agua potable

Fuente: IMB SPSS

3.13.5. Análisis de estacionalidad de las series de tiempo

El análisis de estacionalidad se desarrolla de los datos recolectados de los años 2017 al 2019 de la producción, sistemas de distribución y consumo mensual de agua potable, de igual manera se analiza por medio del software R Studio. Se puede observar en la figura 3.14 en los estudios pertinentes, que la producción de agua potable no se identifica estacionalidad en los periodos de tiempo, debido a que el comportamiento de la producción ha sido diferente en los datos examinados. En el gráfico en el eje X se analiza los meses de producción y en el eje Y la producción de agua potable.

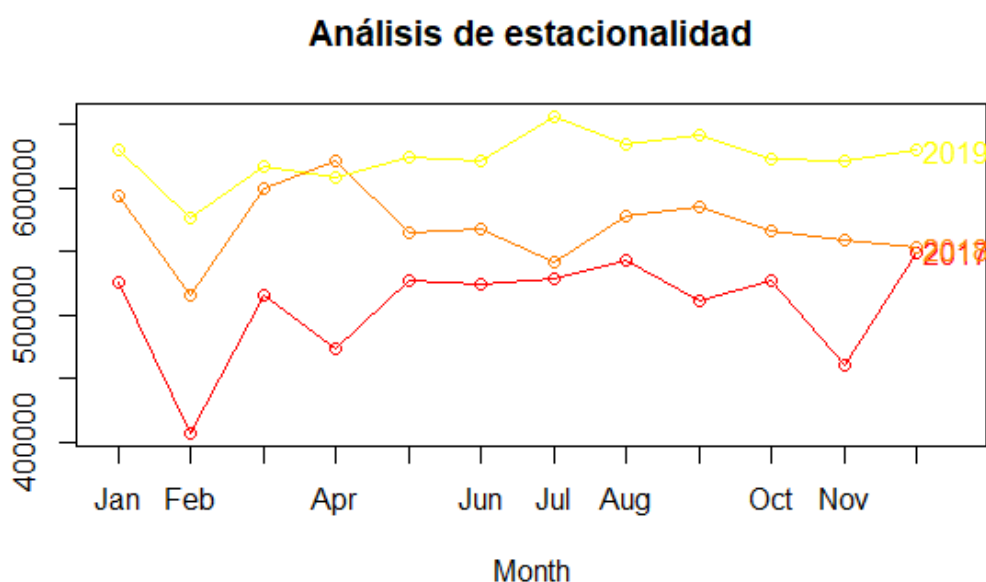


Figura 3.14 Análisis de estacionalidad de la producción mensual 2017-2019

Fuente: R Studio

En los sistemas de distribución de agua potable, se ha identificado en ciertas épocas de los años analizados en donde existen comportamientos similares en el abastecimiento del recurso hídrico. Para la Red principal se observa que no presenta estacionalidad en ningún periodo del año. En el sistema Aeropuerto no existe estacionalidad, en los sistemas San Pedro y Venezuela no presentan una variación periódica similar en la repartición del líquido vital. A continuación, en las figuras 3.15-3.18, se observa la estacionalidad de estos sistemas de distribución. En el gráfico en el eje X se analiza los meses de distribución de agua potable y en el eje Y la distribución de agua potable.

Análisis de estacionalidad

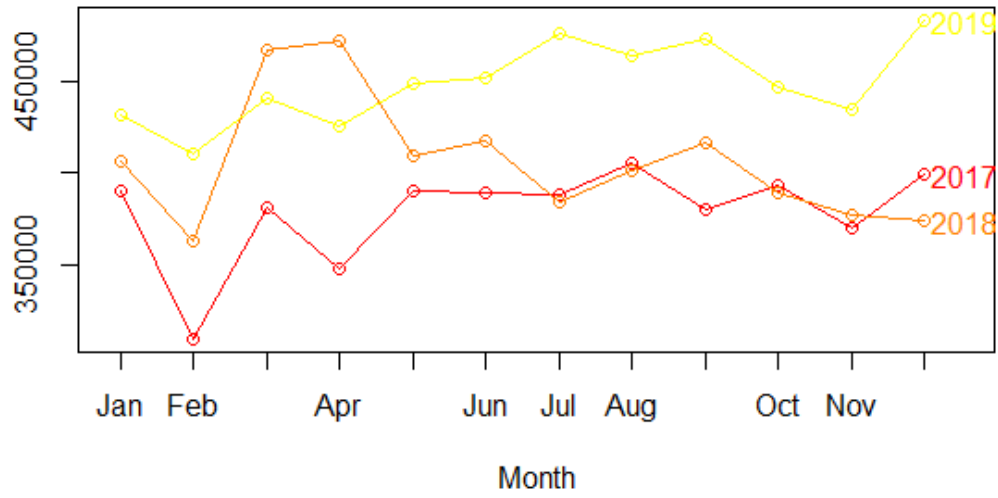


Figura 3.15 Análisis de estacionalidad de la distribución de la Red Principal

Fuente: R Studio

Análisis de estacionalidad

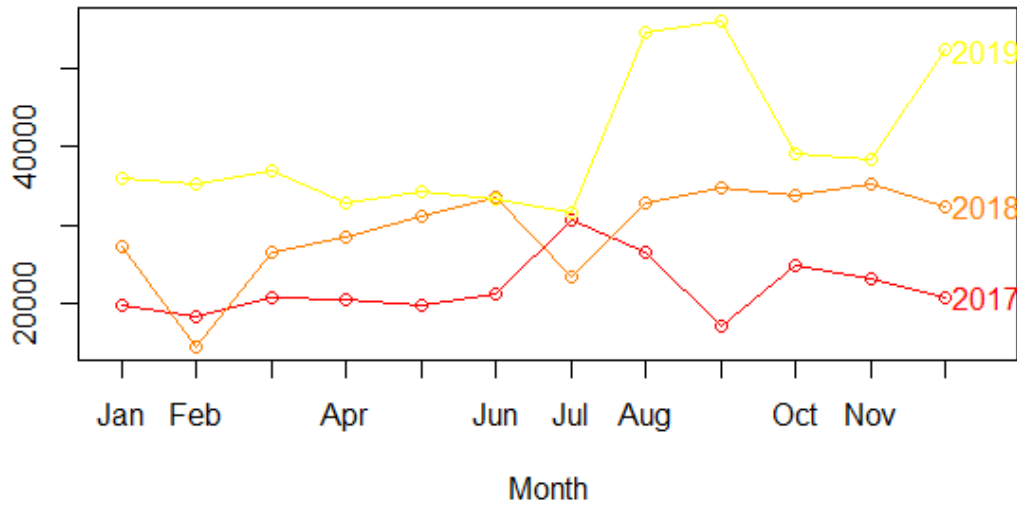


Figura 3.16 Análisis de estacionalidad de la distribución Sistema Aeropuerto

Fuente: R Studio

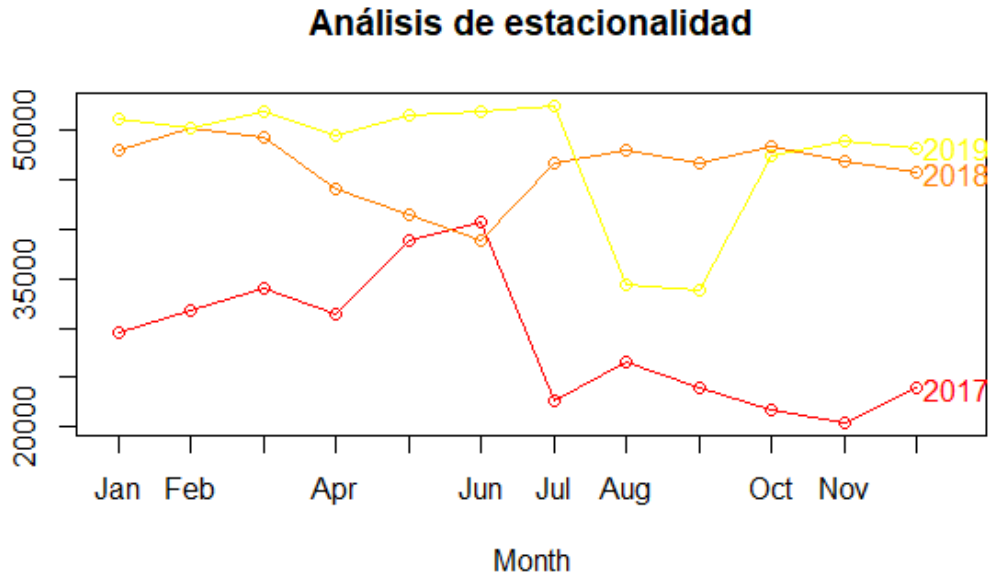


Figura 3.17 Análisis de estacionalidad de la distribución Sistema San Pedro

Fuente: R Studio

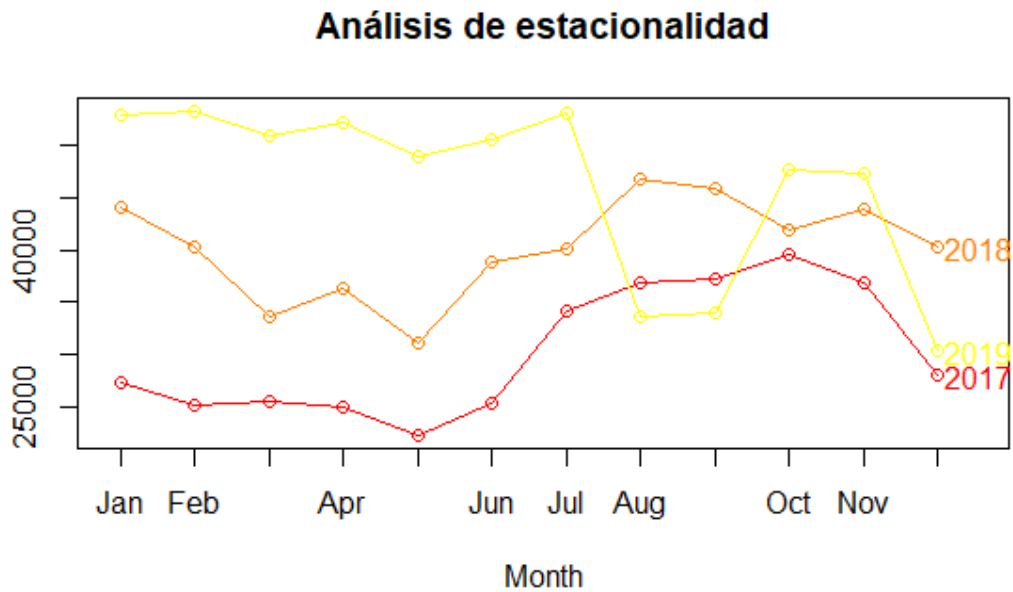


Figura 3.18 Análisis de estacionalidad de la distribución Sistema Venezuela

Fuente: R Studio

En el sistema de la distribución de los Tanqueros la estacionalidad ocurre en varios periodos del año como son marzo y agosto. En cuanto al sistema de reparto de los Tanques elevados no posee una similitud de distribución en el en los meses de los históricos analizados. Los análisis de estacionalidad de los sistemas de distribución se observan en las figuras 3.19-3.20.

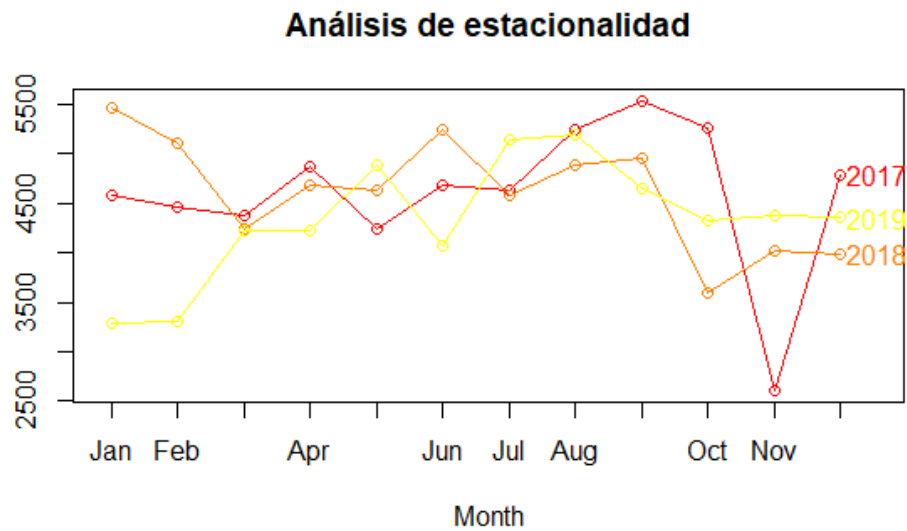


Figura 3.19 Análisis de estacionalidad de la distribución Tanqueros

Fuente: R Studio

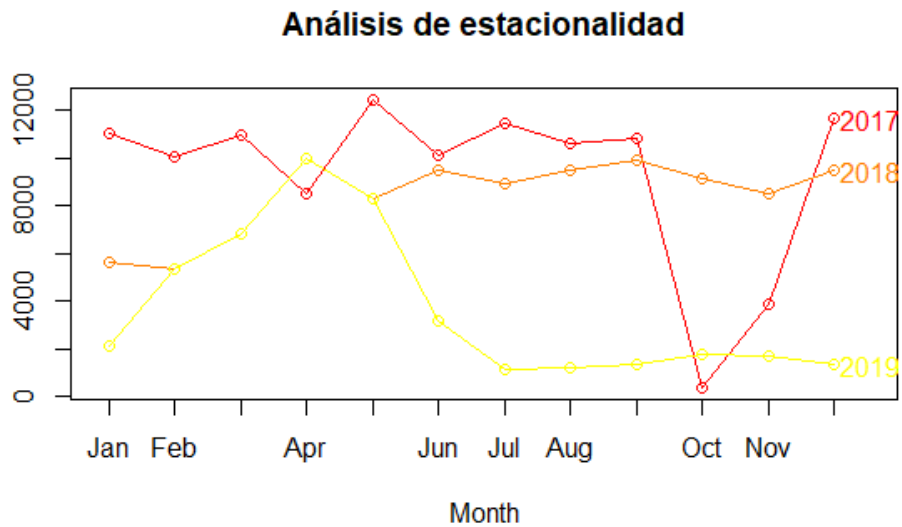


Figura 3.20 Análisis de estacionalidad de la distribución Tanques Elevados

Fuente: R Studio

La base de datos recolectada de los clientes mensuales de los tres últimos periodos históricos 2017-2019, arroja un resultado en el cual se identifica que no posee estacionalidad ya que los clientes han venido incrementándose progresivamente. El análisis de estacionalidad del consumo mensual de agua potable, como se lo demuestra en la figura 3.21. En el gráfico en el eje X se analiza los meses que los clientes hacen uso del servicio y en el eje Y la cantidad de clientes que hacen uso del agua potable.

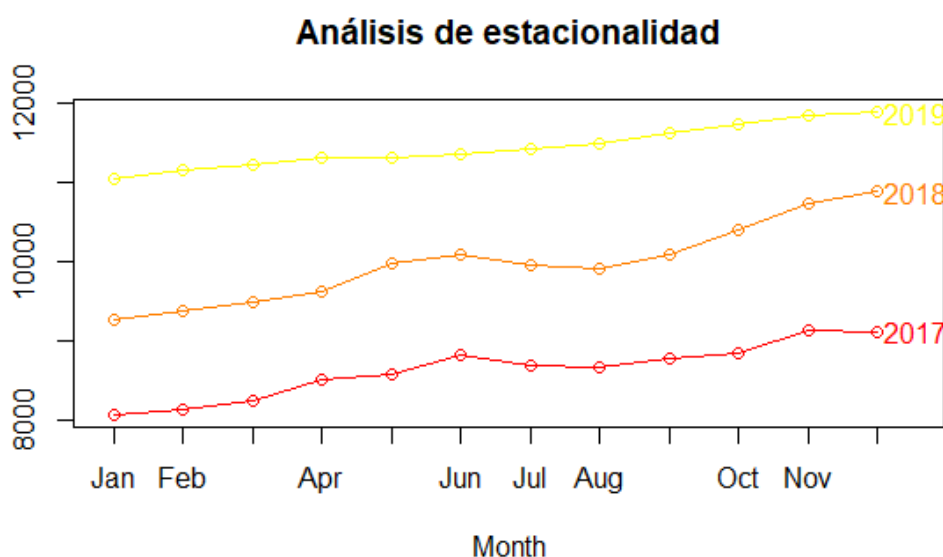


Figura 3.21 Análisis de estacionalidad de los clientes mensuales 2017-2019

Fuente: R Studio

En el caso de los datos históricos de la demanda de agua potable, el consumo por parte de los clientes de los tres años pasados se identifica que existe estacionalidad el mes de septiembre y una pequeña similitud en el mes de febrero, para mejor visualización se puede verificar en la figura 3.22, del análisis de estacionalidad de la demanda de agua potable. En el gráfico en el eje X se analiza los meses demandados del servicio hídrico potable y en el eje Y la demanda en metros cúbicos de agua potable.

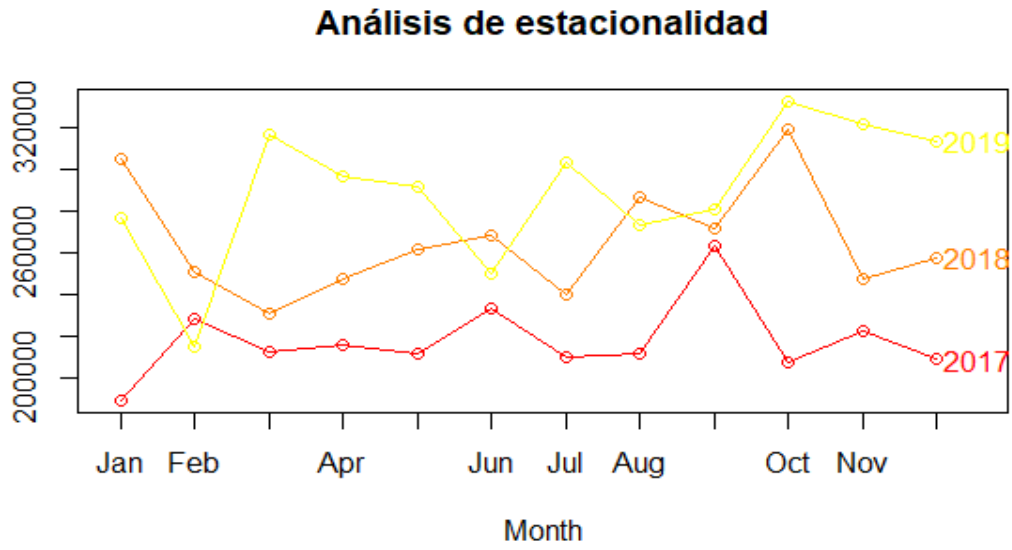


Figura 3.22 Análisis de estacionalidad de la Demanda de agua potable

Fuente: R Studio

CAPÍTULO IV

4. DESARROLLO DEL MODELO DE PRONÓSTICO

Para la estructuración del modelo de redes neuronales, se definen los datos que van a ser introducidos en la capa de entrada, en este caso es la demanda mensual de agua potable de los últimos años que comprenden 2017 hasta el 2019. Se analiza la autocorrelación y la estacionalidad de la serie a pronosticar, las pruebas en la red neuronal definen el número óptimo de neuronas que van a ser ingresadas a las 3 capas ocultas definidas en el modelo.

Para que el pronóstico posea un valor de error cercano a cero, se realiza el entrenamiento de las redes neuronales, de igual manera el número de entrenamientos dependerá de que tan rápido disminuya el error, la cual permita visualizar una predicción con mayor exactitud.

4.1. Método para la elaboración del pronóstico

Para llevar a cabo la realización del pronóstico en la figura 4.1 se establece lo principal para su ejecución, los pasos a seguir como es la identificación del horizonte temporal, es decir si es pronóstico va a ser a corto, mediano o largo plazo juntamente con las variables que van a predecir, en este caso la demanda en metros cúbicos de agua potable. Al momento de tener identificado las variables se establece la técnica que se vaya a utilizar, que va a ser para esta investigación las Redes Neuronales Artificiales. La aplicación de la técnica se obtendrá la tendencia, estacionalidad y la autocorrelación. De la misma manera se establecerá los parámetros de para el entrenamiento y validación de la red neuronal para la identificación del pronóstico y la minimización del error obtenido.

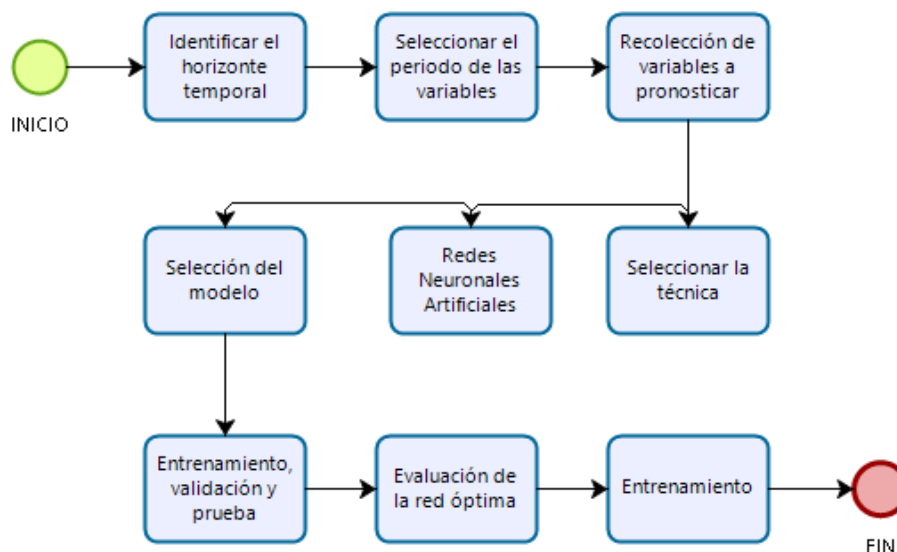


Figura 4.1 Método para la elaboración del pronóstico

Fuente: Propia

4.2. Identificación de la variable

Para garantizar la correcta ejecución y obtención de los pronósticos, se debe definir correctamente los datos de la demanda de agua potable, ya que con estos valores se comenzará a trabajar en el modelo establecido.

- Variable por pronosticar: Demanda de agua potable de EMAPALA EP
- Tipo de variable: Numérica o cuantitativa
- Unidad de medida: Metro cúbico (m^3)

4.3. Recolección de datos

La empresa EMAPALA EP posee un registro de la cantidad de clientes que hacen el uso del servicio básico y la cantidad en metros cúbicos que demandan, esto es obtenido por medio del software SCADA que permite ordenar y controlar los procesos a distancia, a su vez permite generar los reportes de la producción y consumo de agua potable. A continuación, en la tabla 4.1 y 4.2 se observan los clientes abonados y el consumo.

Meses	Consumo		
	2017 <i>m</i>³	2018 <i>m</i>³	2019 <i>m</i>³
Enero	189011	304903	276701
Febrero	228047	250927	214928
Marzo	212167	230896	316891
Abril	215287	247203	296983
Mayo	211103	261716	291878
Junio	233179	268416	249745
Julio	209940	239772	303207
Agosto	211474	286573	273650
Septiembre	263315	271552	281032
Octubre	207550	319730	332641
Noviembre	222366	247178	321461
Diciembre	208760	257637	313581

Tabla 4.1 Consumo mensual 2017-2019

Fuente: EMAPALA EP

Meses	Cientes		
	2017	2018	2019
Enero	8080	9261	11053
Febrero	8151	9385	11157
Marzo	8240	9482	11226
Abril	8512	9618	11320
Mayo	8580	9984	11312
Junio	8820	10087	11349
Julio	8697	9968	11424
Agosto	8682	9904	11498
Septiembre	8782	10098	11613
Octubre	8840	10402	11720
Noviembre	9132	10724	11843
Diciembre	9121	10878	11884

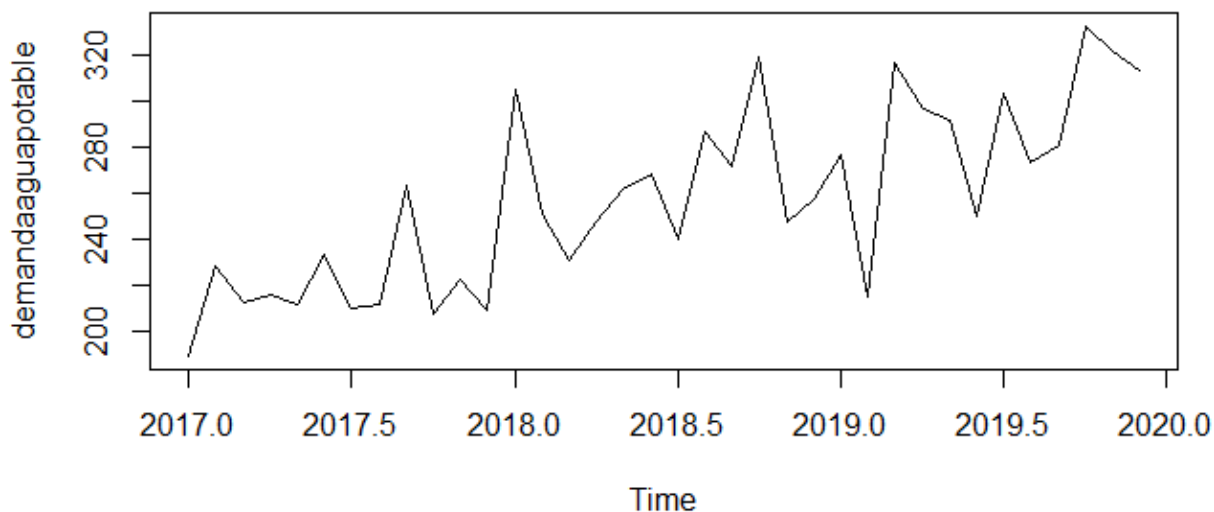
Tabla 4.2 Clientes mensuales 2017-2019

Fuente: EMAPALA EP

4.4. Estructuración del modelo

4.4.1. Análisis de la serie temporal

La serie temporal es utilizada para analizar cómo se comportan la o las variables en el transcurso del tiempo, como se lo ha explicado anteriormente esta serie temporal pertenece a la demanda de agua potable de los años 2017-2019, ordenados mensualmente y su unidad de medida es en metros cúbicos. A continuación, en la figura 4.2, se observa el comportamiento que ha tenido la demanda de agua potable a través de los años.

**Figura 4.2** Análisis de la serie temporal de la demanda de agua potable

Fuente: R Studio

4.4.2. Análisis de tendencia y estacionalidad

Los datos históricos obtenidos de la demanda de agua potable han venido incrementándose año tras año debido al aumento del consumo por parte de los clientes, de esta manera la serie de tiempo ha obtenido una tendencia positiva desde el 2017 hasta el 2019. Se observa en la figura 4.3 del análisis de tendencia.

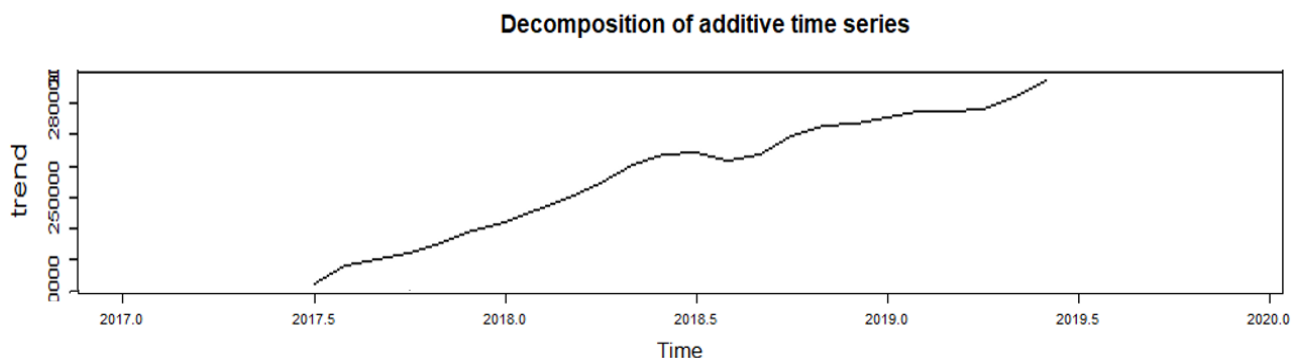


Figura 4.3 Análisis de tendencia

Fuente: R Studio

Mediante la descomposición de la serie de tiempo se observa como a lo largo de los 3 años históricos analizados, los patrones de consumo de agua potable por parte de la población del cantón son similares en ciertas épocas del año, dando como resultado una estacionalidad significativa. Se lo puede identificar en la figura 4.4 del análisis de estacionalidad.

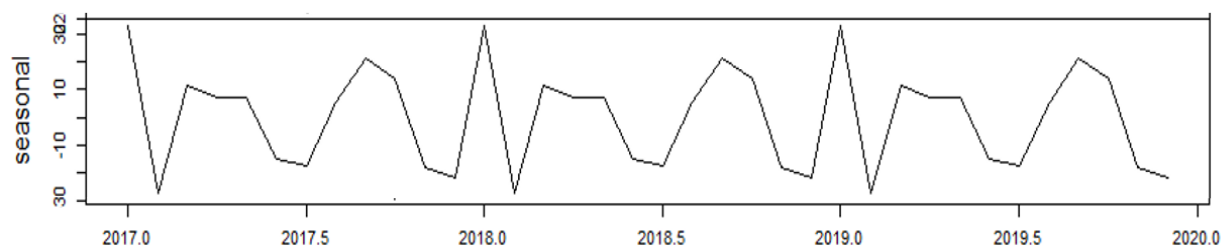


Figura 4.4 Análisis de estacionalidad

Fuente: R Studio

4.5. Pronóstico de la red neuronal óptima

El objetivo principal es realizar el pronóstico de la demanda de agua potable, para lo cual, la base de datos ingresada es la cantidad en metros cúbicos históricos del consumo, las variables de entrada van a ser 12, por los meses de cada año recolectado desde el 2017 hasta el 2019.

La red neuronal óptima es la primera con la que se comienza el modelo, en la cual se establece las repeticiones que van a tener para la disminución MSE, se establece el número de nodos en las 3 capas ocultas del modelo Autorregresivo (AR).

El número de redes a entrenar dependerá de que tan grande o pequeño sea el error, es decir si aumenta el número de repeticiones, disminuirá el MSE, es decir si reps= 4 (mse=0,5), si reps=40 (mse=0,1), reps= 100 (mse=0,07), reps=200 (mse=0,05). Para el modelo se utilizará 200 repeticiones.

Para la obtención del pronóstico de 2020 y 2021 en la codificación de la red MLP se colocará (y*2), esto hace referencia de que no se realizará el pronóstico para un solo año, sino para dos. En este modelo con la función Sel.Lag= False, anula la selección automática de los retrasos, de esta manera los retrasos se establecen de forma manual del 1 hasta el 1:12, siendo estos los regresores del modelo.

De esta manera se comienza con la primera interacción (fit1) y se procede a identificar el pronóstico obtenido, el error MSE, la gráfica de pronóstico de la red MLP y la estructura de las capas y nodos respectivamente.

```
#####RED NEURONAL ÓPTIMA
# 1
fit1 <- mlp(y*2, reps=200, hd = c(26,28,24), sel.lag=FALSE, lags = 1:12)
print(fit1)
p2<-forecast(fit1)
plot(p2)
plot(fit1)
```

Figura 4.5 Codificación para el pronóstico de la red neuronal óptima

Fuente: R Studio

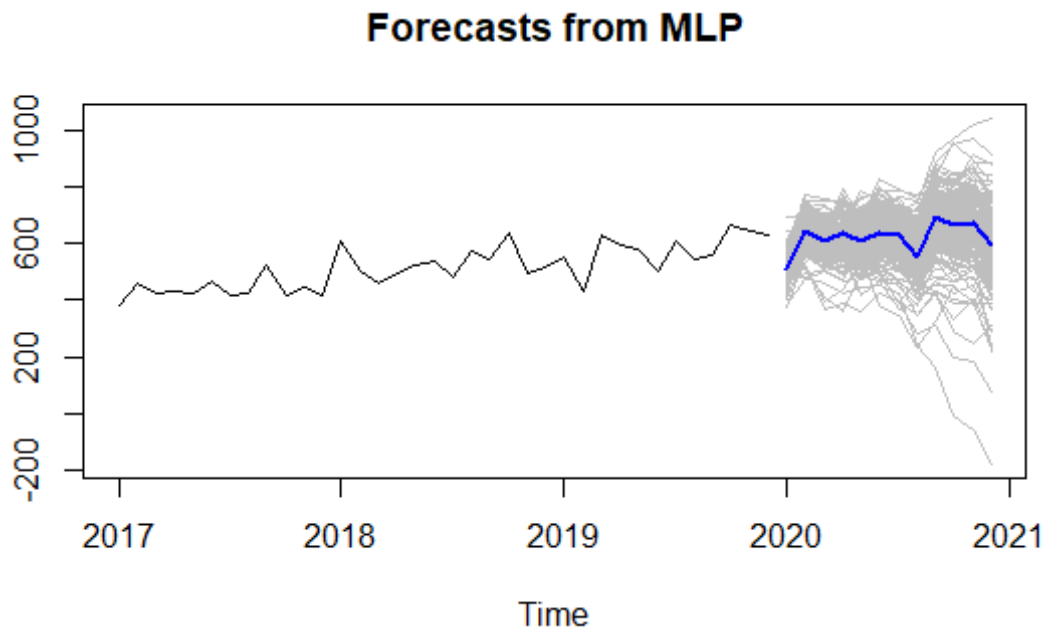


Figura 4.6 Pronóstico de la red MLP

Fuente: R Studio

4.6. Estructura de la red MLP

El perceptrón multicapa para la elaboración del modelo está conectado totalmente, a su vez consta de 1 capa de entrada, 3 capas ocultas y 1 capa de salida.

- La capa de entrada está compuesta por 12 nodos circulares
- Las capas ocultas están compuestas por 26, 28 y 24 nodos
- La capa de salida consta de un solo nodo

Cabe señalar que cada nodo circular es una neurona artificial y cada línea es la conexión que va desde la salida de una neurona a la entrada a la otra, generando el resultado de la información procesada, en este caso el pronóstico de la demanda de agua potable.

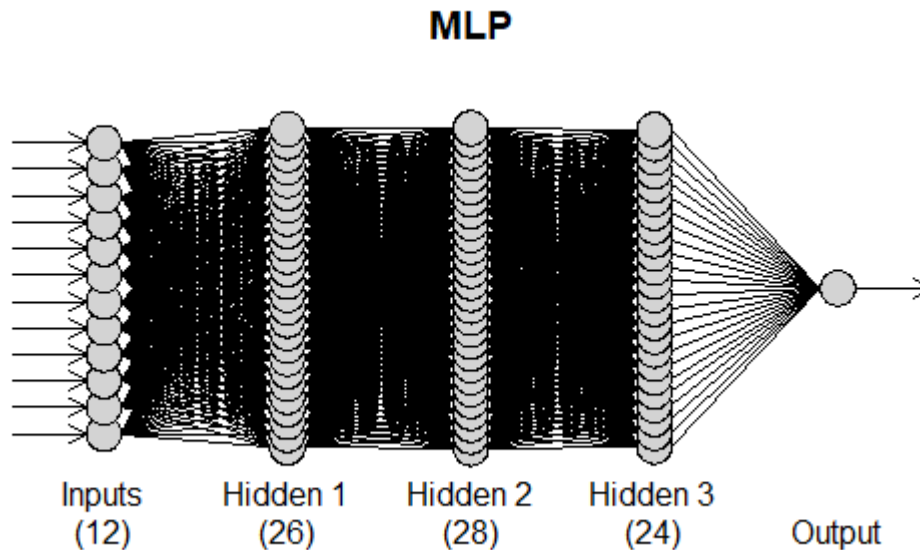


Figura 4.7 Estructura de las capas de la red

Fuente: R Studio

El resultado de la red óptima del modelo fit1 es un MSE de 0.0309, siendo resultado de la combinación y comportamiento de la red.

```
> #####RED NEURONAL ÓPTIMA
> # 1
> fit1 <- mlp(y*2, reps=200, hd = c(26,28,24), sel.lag=FALSE, lags = 1:12)
MLP fit with (26,28,24) hidden nodes and 200 repetitions.
Series modelled in differences: D1.
Univariate lags: (1,2,3,4,5,6,7,8,9,10,11,12)
Forecast combined using the median operator.
MSE: 0.0309.
```

Figura 4.8 MSE de la red neuronal óptima

Fuente: R Studio

Como para la red óptima en la figura 4.4, se codificó para 2 años de pronósticos, al momento de realizar las mejoras de los entrenamientos, en la primera mejora (fit2), se debe indicar que el número de datos a pronosticar debe ser para 2 años ($h=24$), como se observa en la figura 4.7, de la configuración de la mejora de los entrenamientos.


```

##Mejora de entrenamiento
##
fit2<-mlp(y, model=fit1,retrain=20)
print(fit2)
plot(fit2)
plot(forecast(fit2,h=24))
forecast(fit2,h=24)

```

Figura 4.9 Configuración en la mejora de entrenamiento de la red

Fuente: R Studio

4.7. Pronóstico con la mejora del entrenamiento de la red

Para que el error obtenido al inicio del modelo pueda disminuir, se procede a entrenar nuevamente, pero partiendo de los resultados obtenidos de la red óptima (fit1), para ello se comienza con el fit2 que es el primer entrenamiento partiendo del fit1, se realiza 20 repeticiones de entrenamiento a las redes. En la obtención del mejor pronóstico y el menor error posible, se entrenó 30 veces la red óptima dando como resultado en el fit10 un MSE de 0.0048, al no haber mejoraría en la solución, es decir, la obtención del menor error se mantiene, entonces el modelo, el algoritmo, la red, deja de iterar, ya que al comparar con el valor próximo no mejora, entonces se estable el valor presentado como la mejor solución.

```

> fit10<-mlp(y, model=fit9, retrain=TRUE)
> print(fit10)
MLP fit with (26,28,24) hidden nodes and 200 repetitions.
Series modelled in differences: D1.
Univariate lags: (1,2,3,4,5,6,7,8,9,10,11,12)
Forecast combined using the median operator.
MSE: 0.0048.
> plot(fit10)
> plot(forecast(fit10,h=24))
> forecast(fit10,h=24)

```

Figura 4.10 Mejora del entrenamiento de la red

Fuente: R Studio

Forecasts from MLP

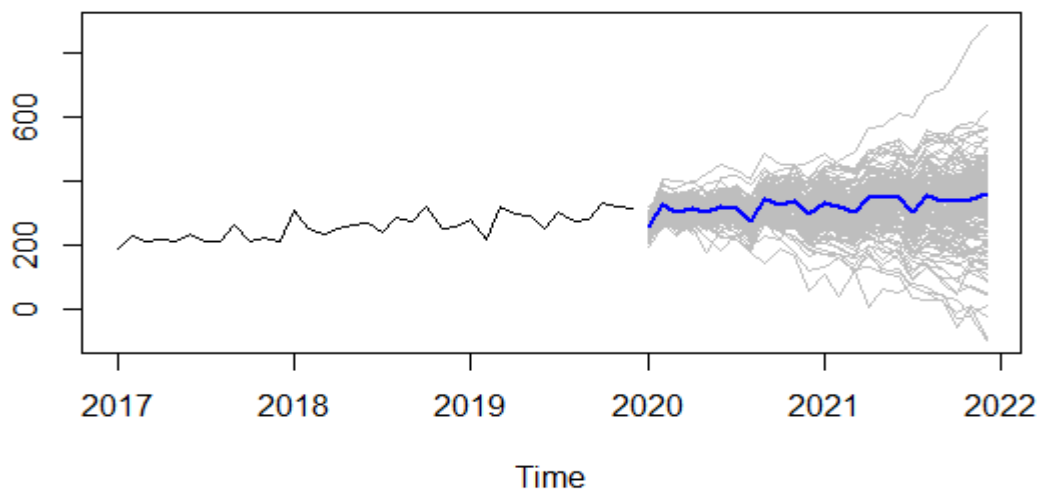


Figura 4.11 Pronóstico con la mejora del entrenamiento de la red MLP

Fuente: R Studio

El resultado del pronóstico obtenido para el 2020-2021 de la demanda de agua potable se lo muestra en la tabla 4.3.

Meses	Demanda Pronosticada 2020	Demanda Pronosticada 2021
	m^3	m^3
Enero	260050,10	327214,40
Febrero	319713,20	316629,60
Marzo	305147,30	302864,60
Abril	314652,60	346805,90
Mayo	306564,40	346143,30
Junio	317216,10	346254,50
Julio	311650,20	302724,70
Agosto	272307,00	348844,00
Septiembre	338984,70	335032,80
Octubre	329859,30	335717,10
Noviembre	333719,20	340542,00
Diciembre	293434,50	349149,00

Tabla 4.3 Pronóstico de la demanda de agua potable 2020-2021

Fuente: Propia

CAPÍTULO V

5. ANÁLISIS DE LOS RESULTADOS OBTENIDOS

La obtención del pronóstico mediante la utilización de Redes Neuronales Artificiales se analiza por medio del mejor MSE, este es el error cuadrático medio de un estimador, es decir calcula la diferencia cuadrática promedio de los valores pronosticados. Con esto se puede decir que mientras menor sea el valor de este error, mejor será el resultado de la predicción.

Se comparó el modelo de RNA con SARIMA (1,1,1) (1,0,1) y el Aditivo de Winters, modelos que identifican la estacionalidad y tendencia de la serie de tiempo que se esté analizando, juntamente con la utilización de los softwares SPSS y Crystal Ball. El resultado de la predicción de la demanda de agua potable para el año 2020-2021 es en función de metros cúbicos.

5.1. Comportamiento de la demanda

Se identifica la demanda que ha tenido EMAPALA EP desde el 2017 hasta el 2019, a su vez se observa el comportamiento pronosticado para el 2020, en la cual se puede decir que el consumo de agua potable por parte de los usuarios ha venido incrementándose año tras año. Como se observa en la tabla 5.1 y en la figura 5.1.

Meses	Demanda 2017	Demanda 2018	Demanda 2019	Demanda Pronosticada 2020	Demanda Pronosticada 2021
	m^3	m^3	m^3	m^3	m^3
Enero	189011,00	304903,00	276701,00	260050,10	327214,40
Febrero	228047,00	250927,00	214928,00	319713,20	316629,60
Marzo	212167,00	230896,00	316891,00	305147,30	302864,60
Abril	215287,00	247203,00	296983,00	314652,60	346805,90
Mayo	211103,00	261716,00	291878,00	306564,40	346143,30
Junio	233179,00	268416,00	249745,00	317216,10	346254,50
Julio	209940,00	239772,00	303207,00	311650,20	302724,70
Agosto	211474,00	286573,00	273650,00	272307,00	348844,00
Septiembre	263315,00	271552,00	281032,00	338984,70	335032,80

Octubre	207550,00	319730,00	332641,00	329859,30	335717,10
Noviembre	222366,00	247178,00	321461,00	333719,20	340542,00
Diciembre	208760,00	257637,00	313581,00	293434,50	349149,00

Tabla 5.1 Comparativa de la Demanda de agua potable

Fuente: Propia

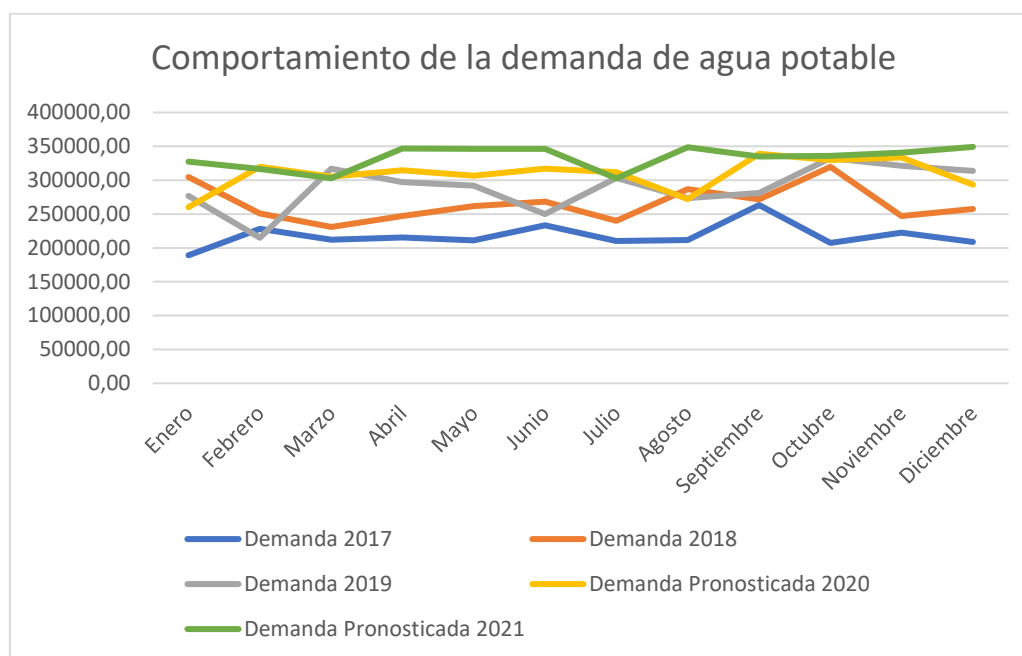


Figura 5.1 Comparación de la Demanda de agua potable

Fuente: Propia

5.2. Comparación de modelos

Para garantizar la confiabilidad de los resultados obtenidos de las Redes Neuronales Artificiales se realiza la comparación, obteniendo de esta manera los pronósticos por cada uno de ellos y sus respectivos errores MSE.

En la figura 5.2, se observa cómo se comportan cada una de las proyecciones obtenidas, en el modelo SARIMA (1,1,1) (1,0,1) el pronóstico obtenido se podría decir que se mantiene en la mitad de los dos modelos también analizados, en el modelo Aditivo de Winters, el pronóstico se mantiene en todos los meses con valores superiores que los dos modelos analizados. Y en cuanto a las Redes

Neuronales Artificiales el pronóstico es más real y se asemeja a lo consumido en años anteriores, teniendo en cuenta el crecimiento de la demanda por parte de sus consumidores.

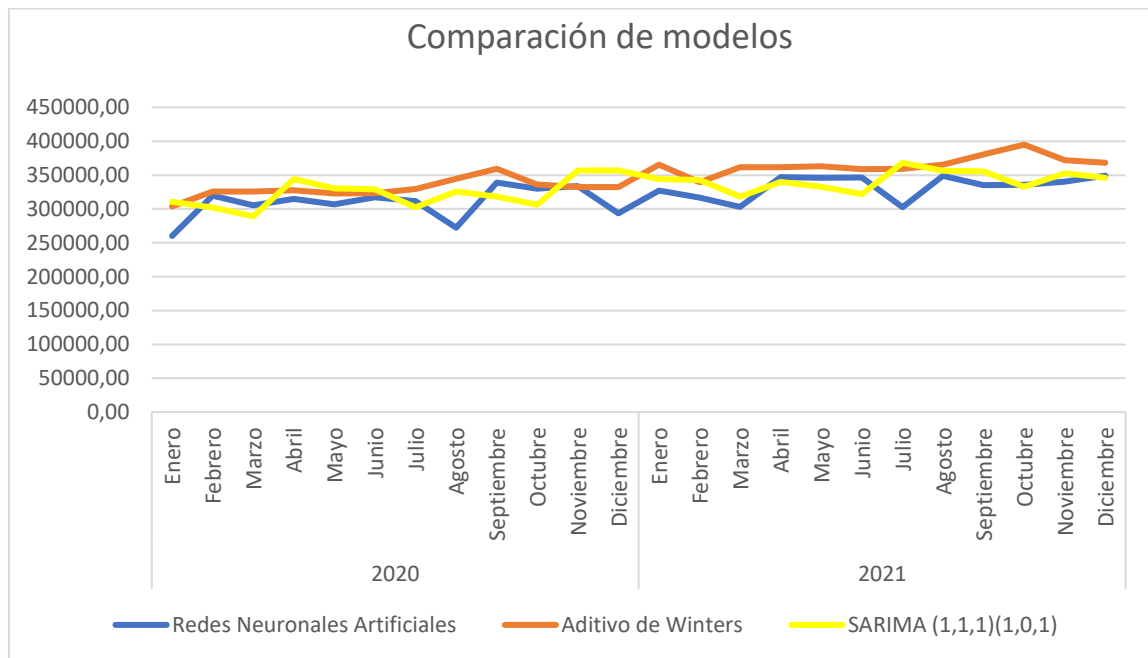


Figura 5.2 Comparación entre modelos

Fuente: Propia

El resultado obtenido de la ejecución de los modelos se pudo identificar que el mejor pronóstico es de las Redes Neuronales Artificiales con un $MSE= 0.0048$, mientras que el SARIMA (1,1,1) (1,0,1) realizado por el software Crystal Ball, un $MSE= 544.60$ y el modelo Aditivo de Winters, ejecutado por el software SPSS obtuvo un $MSE= 573.93$. A continuación se observa en la tabla 5.2, la comparación de los modelos. Cabe recalcar que la comparación de estos modelos entre sí fue debido a que poseen el criterio de identificar la estacionalidad y tendencia que existe en la serie de tiempo.

		R STUDIO	IBM SPSS	ORACLE CRYSTAL BALL
		Redes Neuronales Artificiales	Aditivo de Winters	SARIMA (1,1,1) (1,0,1)
Pronósticos	Meses	Demanda Pronosticada		
		m^3	m^3	m^3
2020	Enero	260050,10	329301,1	285.168,80
	Febrero	319713,20	303730,1	310.635,61
	Marzo	305147,30	325747,43	302.228,94
	Abril	314652,60	325587,1	289.177,20
	Mayo	306564,40	327328,43	344.335,47
	Junio	317216,10	322876,1	330.424,71
	Julio	311650,20	323402,43	329.205,31
	Agosto	272307,00	329661,77	302.347,16
	Septiembre	338984,70	344395,77	325.841,67
	Octubre	329859,30	359069,77	318.107,34
	Noviembre	333719,20	336097,77	306.132,30
	Diciembre	293434,50	332422,1	356.752,56
2021	Enero	327214,40	365155,23	343.986,23
	Febrero	316629,60	339584,23	342.867,13
	Marzo	302864,60	361601,56	318.218,45
	Abril	346805,90	361441,23	339.780,20
	Mayo	346143,30	363182,56	332.682,13
	Junio	346254,50	358730,23	321.692,21
	Julio	302724,70	359256,56	368.148,24
	Agosto	348844,00	365515,89	356.432,12
	Septiembre	335032,80	380249,89	355.405,08
	Octubre	335717,10	394923,89	332.784,10
	Noviembre	340542,00	371951,89	352.572,09
	Diciembre	349149,00	368276,23	346.057,94
ERROR	RSME	0,069	23,957	23,55
	MSE	0,0048	573,93	554,60

Tabla 5.2 Comparación entre modelos

Fuente: Propia

CONCLUSIONES

- Mediante la realización del marco teórico referencial se pudo distinguir información relevante de técnicas y algoritmos empleadas en la elaboración, predicción, y evaluación del modelo, logrando de esta manera resaltar la importancia y la precisión de las redes neuronales artificiales utilizadas en este trabajo de grado.
- El análisis de la situación actual permite obtener datos de consumo y realizar el análisis estadístico de los mismos, además se comprueba la inexistencia de pronósticos de la demanda del recurso hídrico. Se puede observar también que la empresa trabaja con reglamentos de entidades regulatorias como el ARCA, ARCSA y la Secretaría del Agua, garantizando el servicio básico a 23 sectores del cantón Lago Agrio.
- Con el apoyo del software R Studio se pudo aplicar la técnica de Redes Neuronales Artificiales para la elaboración del pronóstico de la demanda de agua potable, tomando en cuenta la codificación correcta del modelo y la mejor solución de entrenamiento de la red, logrando de esta manera obtener una base de datos que posea similitud a lo consumido por parte de la población.
- Los resultados obtenidos por medio de las Redes Neuronales fueron analizados y comparados con otros modelos para identificar la mejor solución y el menor error de pronóstico, por parte del modelo SARIMA (1,1,1) (1,0,1), alcanzó un MSE de 554.60, para el modelo Aditivo de Winters un MSE de 573.93 y las RNA se obtuvo un MSE de 0.0048, corroborando la hipótesis que las Redes Neuronales Artificiales son la mejor opción para la predicción de la demanda de agua potable.

RECOMENDACIONES

- Tomar en cuenta que para una investigación fundamentada se debe recolectar la mayor cantidad de información posible, especialmente de libros y artículos relacionados al proyecto que se esté realizando, cabe mencionar que el estudio de pronósticos está en constante actualización y por esta razón su importancia de la utilización en la ejecución de trabajos científicos.
- Se recomienda tomar en consideración el pronóstico de la demanda de agua potable presentado, el cual será un apoyo en la planificación de presupuestos, futuros proyectos y actividades previstos por la empresa para el mejoramiento de este servicio básico, especialmente en la formalización del agua no contabilizada.
- Identificar correctamente la o las variables de entrada que van a ser analizadas en el modelo, realizar correctamente la codificación y ejecutar la cantidad de entrenamientos necesarios para obtener un error mínimo y por consiguiente un buen pronóstico.
- Se recomienda comparar los resultados obtenidos por medio de las Redes Neuronales Artificiales con métodos de pronósticos que posean iguales condiciones para visualizar el comportamiento de los datos y obtener una mejor interpretación de la información.

BIBLIOGRAFÍA

- Alvisi, S., & Franchini, M. (2017). Assessment of predictive uncertainty within the framework of water demand forecasting using the Model Conditional Processor (MCP). *Urban Water Journal*, 14(1), 1-10.
- Arandia, E., Ba, A., Eck, B., & McKenna, S. (2016). Tailoring seasonal time series models to forecast short-term water demand. *Journal of Water Resources Planning and Management*, 142(3), 04015067.
- ARCA. (2020). Retrieved from <http://www.regulacionagua.gob.ec/>
- Arístegui, M. (Producer). (2016). ARÍSTEGUI MAQUINARIA. Retrieved from <https://www.aristegui.info/como-funciona-una-red-de-abastecimiento-de-agua-potable/>
- Asamblea Nacional de la República del Ecuador. (2014). In. Quito, Ecuador.
- Bata, M. t. H., Carriveau, R., & Ting, D. S.-K. (2020). Short-term water demand forecasting using nonlinear autoregressive artificial neural networks. *Journal of Water Resources Planning and Management*, 146(3), 04020008.
- Bowerman, B. L. C., Richard, T., II, K., & Anne, B. (2007). *Pronósticos, series de tiempo y regresión: un enfoque aplicado*/Bruce L. Bowerman, Richard T O'Connell, Anne B. Koehler.
- Calvo, D. (Producer). (2018). Red Neuronal Recurrente – RNN. Retrieved from <https://www.diegocalvo.es/red-neuronal-recurrente/>
- Campos, M. (Producer). (2018, Febrero). SOLDAI. Retrieved from <https://medium.com/soldai/inspiraci%C3%B3n-biol%C3%B3gica-de-las-redes-neuronales-artificiales-9af7d7b906a>
- Craveri, A., Kern, S. R., & Pacini, C. D. (2019). *Implementación de R Studio en el área Estadística en las carreras de Ingeniería*. Paper presented at the I Congreso Internacional de Ingeniería Aplicada a la Innovación y Educación-Asamblea General de ISTEAC (Córdoba, Argentina, 20 al 22 de noviembre de 2019).
- García, O. (Producer). (2019). XERIDIA. *Redes Neuronales artificiales: Qué son y cómo se entrenan*. Retrieved from <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>

- Guo, G., Liu, S., Wu, Y., Li, J., Zhou, R., & Zhu, X. (2018). Short-term water demand forecast based on deep learning method. *Journal of Water Resources Planning and Management*, 144(12), 04018076.
- Gupta, Y., & Saraswat, A. (2020). *Machine Learning Techniques for Short-Term Forecasting of Wind Power Generation*. Paper presented at the International Conference on Advanced Machine Learning Technologies and Applications.
- Gutiérrez, A. F. (2013). *Manual de pronósticos para la toma de decisiones*: Editorial Digital del Tecnológico de Monterrey.
- Herrera-Granda, I. D., Chicaiza-Ipiales, J. A., Herrera-Granda, E. P., Lorente-Leyva, L. L., Caraguay-Procel, J. A., García-Santillán, I. D., & Peluffo-Ordóñez, D. H. (2019). *Artificial Neural Networks for Bottled Water Demand Forecasting: A Small Business Case Study*. Paper presented at the International Work-Conference on Artificial Neural Networks.
- I. (2014). In. Quito, Ecuador: Ediciones Legales.
- Ingdal, M., Johnsen, R., & Harrington, D. A. (2019). The Akaike information criterion in weighted regression of immittance data. *Electrochimica Acta*, 317, 648-653.
- Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial. *Neurocomputing*, 10, 215-236.
- Khepri, W. (Producer). (2018). William Khepri. *Redes Neuronales, ¿qué son?* Retrieved from <https://medium.com/@williamkhepri/redes-neuronales-que-son-a64d022298e0>
- Lorente-Leyva, L. L., Patino-Alarcon, D. R., Montero-Santos, Y., Herrera-Granda, I. D., Peluffo-Ordóñez, D. H., Lastre-Aleaga, A.M., Cordoves-Garcia, A. (2020). Artificial Neural Networks in the Demand Forecasting of a Metal-Mechanical Industry. *Journal of Engineering and Applied Sciences*. 81-87.
- Lorente-Leyva, L. L., Pavón-Valencia, J. F., Montero-Santos, Y., Herrera-Granda, I. D., Herrera-Granda, E. P., & Peluffo-Ordóñez, D. H. (2019). *Artificial Neural Networks for Urban Water Demand Forecasting: A Case Study*. Paper presented at the Journal of Physics: Conference Series.
- López, R. F., & Fernández, J. M. F. (2008). *Las redes neuronales artificiales*: Netbiblo.
- Matich, D. J. (2001). *Redes Neuronales: Conceptos básicos y aplicaciones*. Universidad Tecnológica Nacional, México.

- Mohamad (Producer). (2016a). NUMXL SOPORTE. *Apéndice B: Criterio de Información Akaike (AIC)*. Retrieved from <https://support.numxl.com/hc/es/articles/215531083-Ap%C3%A9ndice-B-Criterio-de-Informaci%C3%B3n-Akaike-AIC->
- Mohamad (Producer). (2016b). NUMXL SOPORTE. *Apéndice C: Criterio de Información Bayesiano / Schwarz (BIC/SIC)*. Retrieved from <https://support.numxl.com/hc/es/articles/214065006-Ap%C3%A9ndice-C-Criterio-de-Informaci%C3%B3n-Bayesiano-Schwarz-BIC-SIC->
- Oinkina (Producer). (2015). blog de colah. *Comprensión de las redes LSTM*. Retrieved from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Olabe, X. B. (1998). Redes neuronales artificiales y sus aplicaciones. *Publicaciones de la Escuela de Ingenieros*.
- Ortiz, G. B. APLICACION DE LA ANALITICA DE SOFTWARE Y LA INTELIGENCIA DE SOFTWARE A DATOS DE CALIDAD DE SISTEMAS DESARROLLADOS EN IBM RPG USANDO LA HERRAMIENTA QSOURCE Y RSTUDIO.
- Pacchin, E., Gagliardi, F., Alvisi, S., & Franchini, M. (2019). A comparison of short-term water demand forecasting models. *Water resources management*, 33(4), 1481-1497.
- Pal, M., Maity, R., Ratnam, J., Nonaka, M., & Behera, S. K. (2020). Long-lead prediction of ENSO modoki index using machine learning algorithms. *Scientific reports*, 10.
- Ratnam, J., Dijkstra, H., & Behera, S. K. (2020). A machine learning based prediction system for the Indian Ocean Dipole. *Scientific reports*, 10(1), 1-11.
- Rice, D. C., Cariveau, R., Ting, D. S.-K., & Bata, M. t. H. (2017). Evaluation of crop to crop water demand forecasting: Tomatoes and bell peppers grown in a commercial greenhouse. *Agriculture*, 7(12), 104.
- RODRÍGUEZ AEDO, N. C. (2016). " PRONOSTICO DE DEMANDA DE AGUA POTABLE MEDIANTE REDES NEURONALES".
- Salas, R. (2004). Redes neuronales artificiales. *Universidad de Valparaíso. Departamento de Computación, 1*.
- Salim, N., Gopal, K., & Ayub, A. (2018). *Experiential Statistics Learning with RStudio: Study on Students' Engagement*. Paper presented at the Journal of Physics: Conf. Series.
- Senplades, S. (2017). Plan Nacional de Desarrollo 2017-2021 Toda una Vida. *Plan Nacional de Desarrollo 2017-2021 Toda una Vida*.

- Siadén, V., & Andrea, Y. (2012). Aplicación de las redes neuronales artificiales para el pronóstico de la demanda de agua potable en la Empresa Epsel SA de la ciudad de Lambayeque.
- SNI (Producer). (2017). Sistema Nacional de Información. Retrieved from <https://sni.gob.ec/proyecciones-y-estudios-demograficos>
- Torra Porras, S. (2004). *Siniestralidad en seguros de consumo anual de las entidades de previsión social, La. Perspectiva probabilística y econométrica. Propuesta de un modelo econométrico neuronal para Cataluña*: Universitat de Barcelona.
- Tiwari, M. K., & Adamowski, J. F. (2014). Medium-term urban water demand forecasting with limited data using an ensemble wavelet–bootstrap machine-learning approach. *Journal of Water Resources Planning and Management*, 141(2), 04014053.
- Zhang, X., Yue, M., Yao, Y., & Li, H. (2018). Regional annual water consumption forecast model. *Desalination and Water Treatment*, 114, 51-60.

ANEXOS

Anexo 1. Codificación de la serie temporal de la demanda de agua potable

```

library(nnfor)
library(ggplot2)
library(TSstudio)
library(forecast)

###cargando la base de datos
library(readxl)
setwd("D:/Desktop/TESIS/MATRICES")
base_datos <- read_excel("Demandam3.xlsx")
#base_datos <- read_excel("D:/Desktop/TESIS/MATRICES/Demandam3.xlsx",
#                          sheet = "Hoja1")
view(base_datos)
#Demanda de Agua Potable
##Análisis de datos
##Diagra de autocorrelacion
#vamos a convertir la base de datos de agua potable en serie temporal(ts)
demandaaguapotable=ts(base_datos$Demandam3,freq=12,start=c(2017,1))
demandaaguapotable
plot(demandaaguapotable)

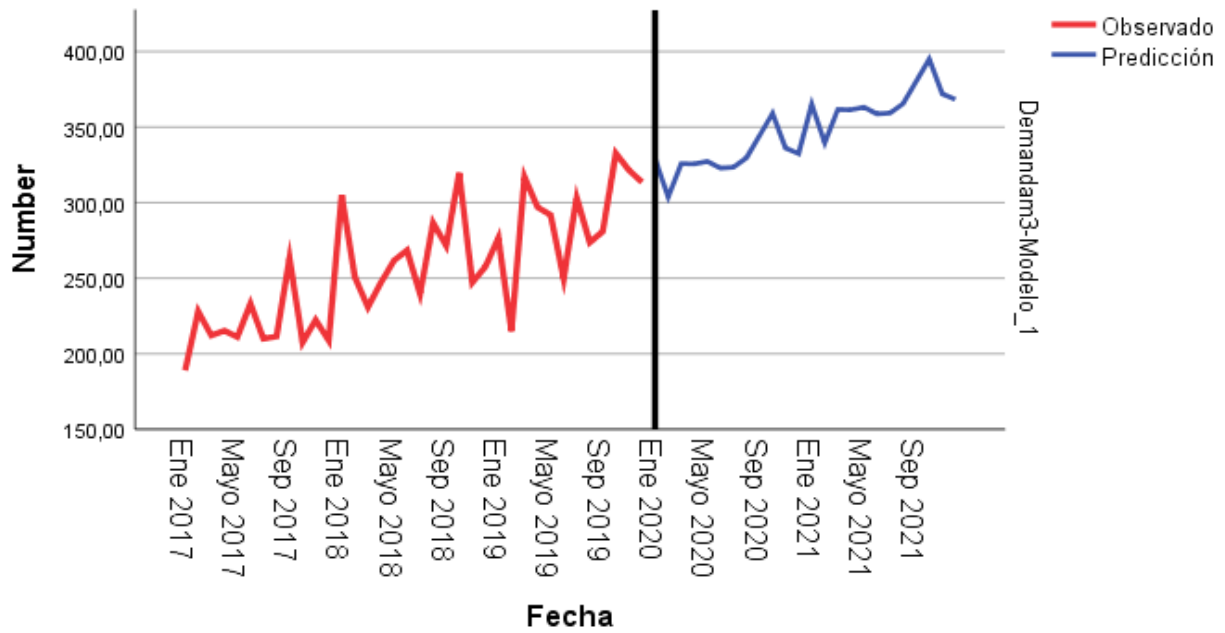
```

Anexo 2. Entrenamiento de la Red Neuronal Artificial

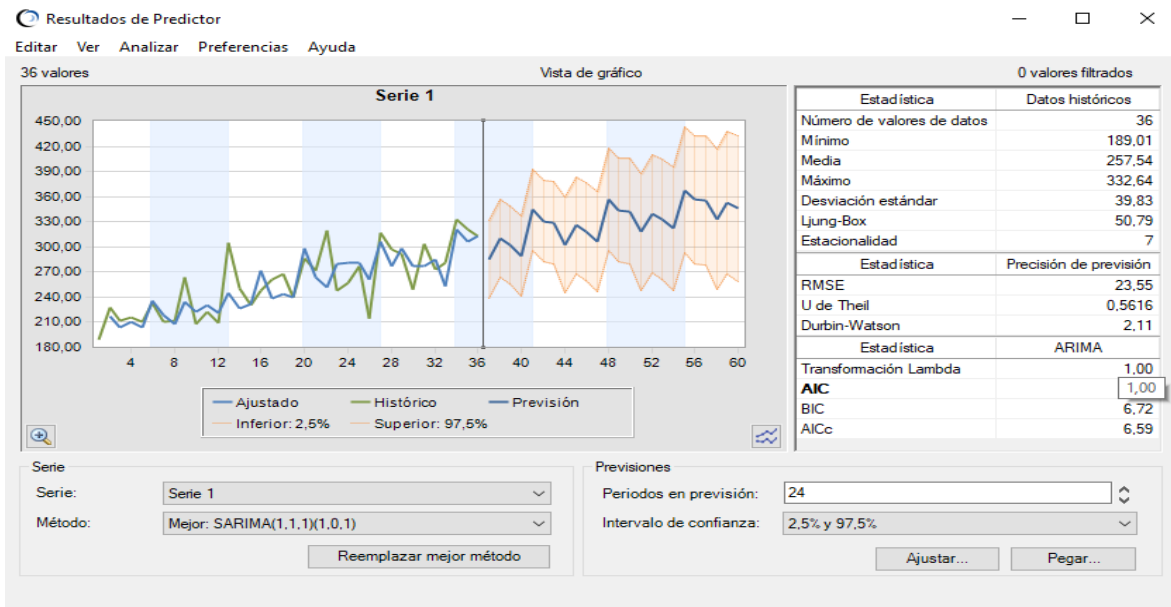
Resultado de los análisis con MLP		
Número de observaciones:		36 obs
Variables:		1
Número de nodos por capa	Entrada	12
	Ocultas	26/28/24
	Salida	1
Código del entrenamiento	Nº de Neuronas	MSE
Fit1	1	0,0359
Fit2	2	0,009
Fit3	3	0,0075
Fit4	4	0,0065
Fit5	5	0,0088
Fit6	6	0,008
Fit7	7	0,0064
Fit8	8	0,0073

Fit9	9	0,007
Fit10	10	0,0048
Fit11	11	0,0061
Fit12	12	0,0076
Fit13	13	0,0082
Fit14	14	0,0081
Fit15	15	0,0086
Fit16	16	0,0054
Fit17	17	0,0057
Fit18	18	0,0083
Fit19	19	0,0084
Fit20	20	0,0063
Fit21	21	0,0092
Fit22	22	0,0088
Fit23	23	0,0081
Fit24	24	0,0068
Fit25	25	0,0076
Fit26	26	0,0086
Fit27	27	0,0066
Fit28	28	0,0069
Fit29	29	0,006
Fit30	30	0,007

Anexo 3. Proyección de SPSS



Anexo 4. Proyección de Crystal Ball



Anexo 5. Pronóstico de la producción de agua potable 2020-2021

Meses	Producción Pronosticada 2020	Producción Pronosticada 2021
	m^3	m^3
Enero	682358,9	713083,5
Febrero	623557,8	654449,6
Marzo	681969,6	705911,8
Abril	673919,6	690187,1
Mayo	673607,4	710477,8
Junio	656986,4	697975,6
Julio	691065,8	736645,2
Agosto	673104,2	707064
Septiembre	686828	718000,5
Octubre	665797,2	698336,4
Noviembre	662816,5	699297,7
Diciembre	649638,9	682659,8

Anexo 6. Sectorización de captación, tratamiento y distribución de agua potable



