



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

**PRONÓSTICO PARA LA ESTIMACIÓN DE METANO EN UN
SISTEMA DE EXTRACCIÓN FORZADA DE BIOGÁS**

T E S I S

**PARA OBTENER EL TÍTULO DE: LICENCIADO EN
MATEMÁTICAS APLICADAS Y COMPUTACIÓN
MATEMÁTICAS APLICADAS Y COMPUTACIÓN**

P R E S E N T A:



**ANTONIO ROJAS ALEXIS DAVID
DOMÍNGUEZ OCAMPO EMITH NICOLAS**

TUTOR DE TESIS:

DOCTOR ARTURO RODRIGUEZ GARCIA

2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

En primera instancia queremos agradecer al Maestro Jorge Sánchez Gómez (expresidente de la Federación Mexicana de Ingeniería Sanitaria), a la Dra. María Neftalí Rojas Valencia (Académica del Instituto de Ingeniería, UNAM) y a la Maestra Flor Semadeni (Ingeniera y experta en el manejo de biogás), que nos apoyaron en la obtención de información y explicación sobre el lugar, teniendo una disposición favorable para nuestra investigación.

Queremos agradecer a nuestro asesor de Tesis Arturo Rodríguez García (Dr. en Ciencias de la Computación), quien apoyo y oriento durante todo el proyecto de investigación, puesto que nos permitió comprender y razonar el uso de las redes neuronales, permitiendo llegar a generar un buen análisis que facilito un resultado favorable en nuestro proyecto.

Dominguez Ocampo Emith Nicolás

Este trabajo es dedicado para mi familia, que me apoyó incondicionalmente. A mi papá, Nicolás, que con altas y bajas siempre estuvo ahí, y me enseñó a ser el hombre y la persona que soy.

A mí mamá, Carmen, que gracias a su cariño incondicional y sus consejos me enseñaron a querer seguir avanzando en la vida.

Gracias a ambos, sé que sin su apoyo no hubiera podido llegar hasta donde estoy. Gracias por corregirme, gracias por confiar en mí y gracias por creer siempre en mí.

A mis hermanos y a mi sobrina, por siempre tener su cariño y su apoyo en mis momentos más difíciles.

A mi amigo y hermano Alexis, gracias por confiar en mí, tú más que nadie sabe todo lo que nos tocó pasar, estoy orgulloso de lo que somos y de lo que seremos.

En general gracias a la gente cercana a mí que siempre me ha insistido en ser mejor persona moral y profesional.

Antonio Rojas Alexis David

Quiero agradecer las diferentes oportunidades que se presentan en la vida, puesto que estas me permitieron tener la oportunidad de realizar este trabajo de investigación. Agradezco a la institución, a la escuela y a todo aquel personal docente, que me aportaron las bases esenciales para poder emprender mi vida profesional.

Gracias a mis padres, David y Julieta, que desde la infancia me han apoyado e incentivado a seguir mis sueños y luchar por cumplir mis metas, sin su apoyo incondicional, no estaría redactando en estos momentos. Gracias a mi futura esposa, puesto que ella me ha permitido seguir mejorando como persona y como ser humano.

Gracias a mi amigo, hermano y compañero de tesis Nicolás, que con sus esfuerzos, ideas y compromiso realizamos este gran trabajo para iniciar una nueva e increíble etapa.

Índice

Agradecimientos.....	3
Índice.....	5
Listado de Ilustraciones.....	7
Listado de Tablas	10
Resumen	11
Objetivos.....	12
Glosario.....	13
Introducción.....	17
Capítulo 1 Relleno Sanitario Urbano Bordo Xochiaca, antecedentes, procesos y obtención de metano.....	22
1.1 Relleno Sanitario Urbano Bordo Xochiaca.....	22
1.2 Biogás en el Relleno Sanitario Urbano Bordo Xochiaca.....	27
1.3 Metano en la composición del biogás	30
Capítulo 2 Métodos de pronósticos para series de tiempo.....	32
2.1 Serie de tiempo para la predicción de metano.....	32
2.2 Modelos para predicción a series de tiempo	34
2.2.1 Definición de pronóstico	34
2.2.2 Promedio móvil ponderado	35
2.2.3 Holt-Winters	36
2.2.4 Box-Jenkins	39
2.2.5 Red neuronal artificial.....	44
Capítulo 3 Recolección de datos	49
3.1 Obtención de los datos.....	49
3.2 Características de los datos.....	50
3.2.1 Metano sin tratamiento de datos atípicos.....	51
3.2.2 Metano con tratamiento de datos atípicos.....	56
3.3 Tratamiento del metano	61
Capítulo 4 Aplicación de los modelos de pronóstico a los datos.....	62
4.1 Combinaciones posibles.....	62
4.2 Consideración de los días de la semana	62
4.3 Series analizadas.....	63

4.4 Elección de la serie	69
4.5 Método promedio móvil ponderado	69
4.5.1 Aplicación del método promedio móvil ponderado	69
4.5.2 Comportamiento de la estimación del modelo promedio móvil ponderado	71
4.5.3 Pronóstico promedio móvil ponderado	72
4.6 Método Holt-Winters	72
4.6.1 Aplicación del método Holt-Winters	72
4.6.2 Comportamiento de la estimación del modelo Holt-Winters	75
4.6.3 Pronóstico Holt-Winters	79
4.7 Método Box-Jenkins	80
4.7.1 Aplicación del método Box-Jenkins	80
4.7.2 Comportamiento de la estimación del modelo Box-Jenkins	83
4.7.2.1 Residuales, parsimonia y errores de los residuales de metano sin tratamiento	83
4.7.2.2 Residuales, parsimonia y errores de los residuales de metano con tratamiento	86
4.7.3 Pronóstico Box-Jenkins	90
4.8 Método red neuronal artificial	91
4.8.1 Aplicación del método red neuronal artificial	91
4.8.2 Comportamiento de la estimación del modelo red neuronal artificial	95
4.8.3 Pronóstico red neuronal artificial	100
Capítulo 5 Análisis de los resultados	102
5.1 Resultado de la elección de la serie de metano	102
5.2 Resultados del modelo promedio móvil ponderado	103
5.3 Resultados del modelo Holt-Winters	104
5.4 Resultados del modelo Box-Jenkins	106
5.5 Resultados del modelo red neuronal artificial	110
5.6 Resultados generales de las metodologías	113
Conclusiones	117
Bibliografía	119

Listado de Ilustraciones

Ilustración 1: Mapa relleno sanitario Bordo Xochiaca (<i>Dirección General de Bibliotecas, 2014</i>). _____	23
Ilustración 2: Mapa de la red extracción forzada de biogás (<i>Dirección General de Bibliotecas, 2014</i>). _____	24
Ilustración 3: Relleno Sanitario (<i>WordPress, 2014</i>). _____	25
Ilustración 4: Metano en el biogás (<i>ENERGÌA, 2020</i>). _____	30
Ilustración 5: Series de tiempo. _____	34
Ilustración 6: Etapas de la modelo metodología. _____	40
Ilustración 7: Identificación de los parámetros AR. _____	41
Ilustración 8: Identificación de los parámetros MA. _____	42
Ilustración 9: Estructura neuronal (Crispi, 2019). _____	45
Ilustración 10: Red neuronal artificial para estimación. _____	46
Ilustración 11: Distribución de metano sin tratamiento. _____	51
Ilustración 12: Dispersión de metano sin tratamiento. _____	52
Ilustración 13: Serie de tiempo de metano sin tratamiento. _____	53
Ilustración 14: Componentes de metano sin tratamiento. _____	54
Ilustración 15: Periodograma de metano sin tratamiento. _____	55
Ilustración 16: Funciones de correlación de metano sin tratamiento. _____	56
Ilustración 17: Distribución de metano con tratamiento. _____	57
Ilustración 18: Dispersión de metano con tratamiento. _____	57
Ilustración 19: Serie de tiempo de metano con tratamiento. _____	58
Ilustración 20: Componentes de metano con tratamiento. _____	58
Ilustración 21: Periodograma de metano con tratamiento. _____	59
Ilustración 22: Funciones de correlación de metano con tratamiento. _____	60
Ilustración 23: Comparativo de series. _____	60
Ilustración 24: Diagrama de elección de muestras. _____	63
Ilustración 25: Componentes de metano viernes sin tratamiento. _____	63
Ilustración 26: Comportamiento de ruido blanco de metano viernes sin tratamiento. _____	64
Ilustración 27: Componentes de metano viernes con tratamiento. _____	64
Ilustración 28: Comportamiento de ruido blanco de metano viernes con tratamiento. _____	65
Ilustración 29: Componentes de metano sábado con tratamiento. _____	66
Ilustración 30: Comportamiento de ruido blanco de metano sábado con tratamiento. _____	66
Ilustración 31: Componentes de metano domingo con tratamiento. _____	67
Ilustración 32: Comportamiento de ruido blanco de metano domingo con tratamiento. _____	67
Ilustración 33: Componentes de metano diario con tratamiento. _____	68
Ilustración 34: Comportamiento de ruido blanco de metano diario con tratamiento. _____	68

Ilustración 35: Estimación del modelo promedio móvil ponderado. _____	70
Ilustración 36: Comportamiento de la irregularidad de metano con y sin tratamiento. _____	71
Ilustración 37: Pronóstico del modelo promedio móvil ponderado. _____	72
Ilustración 38: Pruebas de la raíz unitaria de metano sin tratamiento. _____	73
Ilustración 39: Pruebas de la raíz unitaria de metano con tratamiento. _____	73
Ilustración 40: Comportamiento de la estacionalidad de metano con y sin tratamiento. _____	73
Ilustración 41: Estimación del modelo Holt-Winters de metano sin tratamiento. _	74
Ilustración 42: Estimación del modelo Holt-Winters de metano con tratamiento. _	74
Ilustración 43: Correlación de residuales del modelo Holt-Winters de metano sin tratamiento. _____	75
Ilustración 44: Correlación de residuales del modelo Holt-Winters de metano con tratamiento. _____	75
Ilustración 45: Prueba Ljung-Box a los residuales del modelo Holt-Winters de metano con y sin tratamiento. _____	76
Ilustración 46: Normalidad de residuales del modelo Holt-Winters de metano sin tratamiento. _____	76
Ilustración 47: Normalidad de residuales del modelo Holt-Winters de metano con tratamiento. _____	77
Ilustración 48: Distribución de residuales del modelo Holt-Winters de metano sin tratamiento. _____	77
Ilustración 49: Distribución de residuales del modelo Holt-Winters de metano con tratamiento. _____	78
Ilustración 50: Pronóstico de modelo Holt-Winters de metano sin tratamiento. __	79
Ilustración 51: Pronóstico de modelo Holt-Winters de metano con tratamiento. _	79
Ilustración 52: Prueba de Dickey-Fuller de metano sin tratamiento. _____	80
Ilustración 53: Prueba de Dickey-Fuller de metano con tratamiento. _____	81
Ilustración 54: Funciones de correlación de metano sin tratamiento. _____	81
Ilustración 55: Funciones de correlación de metano con tratamiento. _____	82
Ilustración 56: Estimaciones de los modelos Box-Jenkins de metano con y sin tratamiento. _____	82
Ilustración 57: Comportamiento de media cero en residuales del modelo Box- Jenkins de metano sin tratamiento. _____	83
Ilustración 58: Prueba de heterocedasticidad de residuales del modelo Box- Jenkins de metano sin tratamiento. _____	83
Ilustración 59: Correlación de residuales del modelo Box-Jenkins de metano sin tratamiento. _____	84
Ilustración 60: Normalidad de residuales del modelo Box-Jenkins del metano sin tratamiento. _____	85
Ilustración 61: Valores atípicos en residuales del modelo Box-Jenkins de metano sin tratamiento. _____	85
Ilustración 62: Correlación de estimación SARIMA de metano sin tratamiento. __	86

Ilustración 63: Comportamiento de media cero en residuales del modelo Box-Jenkins de metano con tratamiento.	86
Ilustración 64: Prueba de heterocedasticidad de residuales del modelo Box-Jenkins de metano con tratamiento.	87
Ilustración 65: Correlación de residuales del modelo Box-Jenkins de metano con tratamiento.	87
Ilustración 66: Normalidad de residuales del modelo Box-Jenkins de metano con tratamiento.	88
Ilustración 67: Valores atípicos en residuales del modelo Box-Jenkins de metano con tratamiento.	89
Ilustración 68: Correlación de estimación SARIMA de metano con tratamiento.	89
Ilustración 69: Pronóstico del modelo Box-Jenkins de metano sin tratamiento.	90
Ilustración 70: Pronóstico del modelo Box-Jenkins de metano con tratamiento.	91
Ilustración 71: Normalidad y escalamiento de los datos de entrada.	92
Ilustración 72: Bases de entrenamiento y validación para la red neuronal artificial.	92
Ilustración 73: Creación de la red neuronal artificial.	93
Ilustración 74: Diagrama de la red neuronal artificial.	93
Ilustración 75: Datos que se utilizarán para entrenamiento y validación.	93
Ilustración 76: Estimación del modelo red neuronal artificial de metano sin tratamiento.	94
Ilustración 77: Estimación del modelo red neuronal artificial de metano con tratamiento.	94
Ilustración 78: Correlación de residuales del modelo red neuronal artificial de metano sin tratamiento.	95
Ilustración 79: Correlación de residuales del modelo red neuronal artificial de metano con tratamiento.	96
Ilustración 80: Normalidad de residuales del modelo red neuronal artificial de metano sin tratamiento.	97
Ilustración 81: Normalidad de residuales del modelo red neuronal artificial de metano con tratamiento.	97
Ilustración 82: Media cero y prueba de heterocedasticidad en los residuales del modelo red neuronal artificial de metano sin tratamiento.	98
Ilustración 83: Media cero y prueba de heterocedasticidad en los residuales del modelo red neuronal artificial de metano con tratamiento.	99
Ilustración 84: Pronóstico del modelo red neuronal artificial de metano sin tratamiento.	100
Ilustración 85: Pronóstico del modelo red neuronal artificial de metano con tratamiento.	100
Ilustración 86: Equivalencia de energías dado el porcentaje de metano.	115

Listado de Tablas

Tabla 1: Procesos de las metodologías. _____	19
Tabla 2: Características generales del biogás (Varnero Moreno, 2011). _____	28
Tabla 3: Estructura de datos obtenidos por las macros. _____	50
Tabla 4: Comportamiento estadístico de metano sin tratamiento. _____	51
Tabla 5: Comportamiento estadístico de metano con tratamiento. _____	56
Tabla 6: Pesos del suavizamiento. _____	61
Tabla 7: Distribución de metano viernes sin tratamiento. _____	64
Tabla 8: Distribución de metano viernes con tratamiento. _____	65
Tabla 9: Distribución de metano sábado con tratamiento. _____	66
Tabla 10: Distribución de metano domingo con tratamiento. _____	68
Tabla 11: Distribución de metano con tratamiento. _____	69
Tabla 12: Pesos encontrados con la herramienta Solver. _____	70
Tabla 13: Pesos elegidos a criterio propio. _____	70
Tabla 14: Error de modelo promedio móvil ponderado. _____	72
Tabla 15: Coeficientes de estimación de metano con y sin tratamiento. _____	80
Tabla 16: Errores de los modelos Holt-Winters. _____	80
Tabla 17: Error del modelo SARIMA de metano sin tratamiento. _____	90
Tabla 18: Error del modelo SARIMA de metano con tratamiento. _____	91
Tabla 19: Error del modelo red neuronal artificial de metano sin tratamiento. __	101
Tabla 20: Error del modelo red neuronal artificial de metano con tratamiento. _	101
Tabla 21: Comparativo de series aceptadas. _____	103
Tabla 22: Resultado de la metodología promedio móvil ponderado. _____	104
Tabla 23: Resultado de la metodología Box-Jenkins. _____	107
Tabla 24: Resultado de la metodología red neuronal artificial. _____	111

Resumen

En esta tesis se comparan metodologías para estimaciones a corto o mediano plazo del comportamiento de la serie de tiempo del gas metano, estas metodologías son promedio móvil ponderado, Box-Jenkins, Holt-Winters y red neuronal artificial. Estas metodologías tienen como fin generar una comparación entre ellas, así como, llegar a generar un proceso de implementación rápida para estimaciones del metano y facilitar su interpretación. Las estimaciones obtenidas serán aceptadas de mejor forma si cumplen con reducir la irregularidad que maneja la muestra en los 3 años de información.

En el trabajo se menciona la ubicación del Relleno Sanitario Urbano Bordo Xochiaca, algunas de sus características y funciones que se llevan a cabo en ese sitio, ventajas y desventajas de las metodologías implementadas y la obtención de metano dentro de una máquina de extracción forzada de biogás que se obtiene diariamente durante muestras automáticas. Se describe a su vez la estructura y el tratamiento a la que fue sometida la serie de tiempo. Es importante decir que se realizaron las pruebas para criterios de ruido blanco y ruido blanco gaussiano en cada modelo expuesto.

Se busca obtener un pronóstico que permita generar un reemplazo o valor esperado de metano aun cuando la máquina tenga fallas y no obtenga alguna información en determinado tiempo. El comportamiento del metano durante el tiempo observado indica claramente un cambio de nivel en su proceso estocástico, teniendo un margen de incertidumbre que permite ser analizado y tratado para reducir el componente irregular de una serie de tiempo. También se busca resaltar las ventajas de cada una de las metodologías respecto a sus componentes de series de tiempo ya que estas fueron implementadas con la mayor cantidad de pruebas que permitieran generar una mejor aceptación de las estimaciones.

El generar redes neuronales artificiales produce una menor variabilidad en los componentes de series de tiempo aplicados en algunas otras áreas de la ciencia, esperando que esta metodología de *machine learning* favorezca la estimación de metano.

Los modelos aplicados son propuestas iniciales, pero no optimas, por lo que, su comparativo solo permitirá saber cuál metodología reduce mejor la irregularidad que compone al metano.

Objetivos

Se analizará una serie con el fin de proponer una estimación mediante metodologías para series de tiempo, esto para efectuar comparaciones en los errores obtenidos y facilitar la elección de las decisiones que el experto del sitio llegue a elegir.

Se pretende dar a conocer al lector, que la aplicación de los modelos puede llegar a estimar componentes del biogás o muestras obtenidas por los expertos, basándose en procesos matemáticos de series de tiempo.

Se busca mediante las metodologías aplicadas, minimizar el error de estimación ante la irregularidad de la muestra.

Por otra parte, de la muestra de metano obtenida se busca llegar a generar un tratamiento que llegue a mejorar el modelo implementado.

Se considera que la red neuronal artificial deba ser una estimación más precisa que el resto de las metodologías, favoreciendo las decisiones del experto en el Relleno Sanitario Urbano Bordo Xochiaca.

Por último, se aplicará un modelo que reemplace los criterios expertos y justifique un remplazo al valor una vez la máquina de extracción falle.

Glosario

Análisis espectral: El análisis espectral tiene que ver con vibraciones y oscilaciones, las cuales, en su forma más pura, son ondulaciones sin quiebres abruptos que se repiten periódicamente a través del tiempo, manteniendo siempre la misma amplitud y frecuencia de oscilación, como la señal emitida por un diapasón. Estas ondulaciones se conocen como ondas senoidales (o cosenoidales) y son los componentes de prácticamente todas las señales conocidas, como el ruido blanco, la voz, o la salida de un modelo ARMA. El análisis espectral separa una señal en las diversas ondas senoidales que la conforman (Montenegro, 2009).

ARIMA: Describir una serie de observaciones después de que haya sido diferenciada d veces, a fin de extraer las posibilidades de no estacionariedad. Esta fórmula se puede aplicar a cualquier modelo. Si hay algún componente p , d , q , igual a cero, se elimina el término correspondiente de la fórmula general (de la Fuente Fernández, 2013).

Axón: El axón o neurita es definido en anatomía como una prolongación que se deriva del soma o cuerpo neuronal y que tiene como meta lograr la sinapsis neuronal. Así mismo sabemos que el axón está recubierto en algunas zonas con las células de Schwann, lo cual permite que la conducción nerviosa se dé más rápido y efectivo que en las zonas donde no hay mielina (FisioOnline, 2017).

Bacterias aerobias: Se denominan organismos aerobios o aeróbicos a los organismos que pueden vivir o desarrollarse en presencia de oxígeno diatómico.¹ El concepto se aplica no solo a organismos sino también a los procesos implicados ("metabolismo aerobio") y a los ambientes donde se llevan a cabo. Un "ambiente aerobio" es aquel rico en oxígeno, a diferencia del anaerobio, donde el oxígeno está ausente, o uno microaerófilo, donde el oxígeno se encuentra en muy bajos niveles (Universal, 2020).

Bacterias anaerobias: Son microorganismos que son capaces de sobrevivir y multiplicarse en ambientes que no tienen oxígeno. Por ejemplo, pueden proliferar en tejido humano lesionado que no esté recibiendo un flujo de sangre rica en oxígeno. Este tipo de bacterias causan infecciones como el tétanos y la gangrena. Las infecciones anaerobias normalmente causan abscesos (acumulación de pus), y la muerte del tejido. Muchas bacterias anaerobias producen enzimas que destruyen el tejido, y a veces libera toxinas poderosas (MedlinePlus, 2019).

Deep learning: El *Deep learning* o aprendizaje profundo se define como un algoritmo automático estructurado o jerárquico que emula el aprendizaje humano con el fin de obtener ciertos conocimientos. Destaca porque no requiere de reglas programadas previamente, sino que el propio sistema es capaz de «aprender» por sí mismo para efectuar una tarea a través de una fase previa de entrenamiento. A su vez, también se caracteriza por estar compuesto por redes neuronales artificiales entrelazadas para el procesamiento de información. Se emplea principalmente para la automatización de análisis predictivos (SmartPanel, 2019).

Dendritas: Son partes de las neuronas que se encuentran repartidas por todo el cuerpo, es decir, tanto en el cerebro y la médula espinal como en aquellas que están en los ganglios, los órganos internos, los músculos, etc. En concreto, las dendritas son pequeñas ramificaciones que salen del cuerpo celular (la parte de la neurona en la que se encuentra el núcleo de la célula). En comparación con el axón, las dendritas suelen ser más cortas y finas, de manera que terminan más cerca del cuerpo celular (Triglia, 2018).

Dióxido de carbono (CO₂): Es un gas incoloro, denso y poco reactivo. Forma parte de la composición de la tropósfera (capa de la atmósfera más próxima a la Tierra) actualmente en una proporción de 350 ppm. (partes por millón). Su ciclo en la naturaleza está vinculado al del oxígeno. El balance del dióxido de carbono es sumamente complejo por las interacciones que existen entre la reserva atmosférica de este gas, las plantas que lo consumen en el proceso de fotosíntesis y el transferido desde la tropósfera a los océanos (Braga, 2019).

Hidrocarburo: Son compuestos orgánicos cuya estructura molecular se forma de la unión entre átomos de hidrógeno y carbono. La fórmula básica de los hidrocarburos es la siguiente: C_xH_y . Estos compuestos orgánicos pueden encontrarse en diferentes estados de materia: líquido, gaseoso (gas natural o por condensación) y eventualmente sólido (Hidrocarburos, 2020).

Lixiviados: Son los resultantes del filtrado lento de agua a través de los materiales del vertedero (percolación). Esta adquiere, por arrastre, materiales de los residuos y productos de la descomposición de estos que quedan suspendidos o disueltos dando lugar a líquidos con diferente agresividad y un alto coste de tratamiento. La presencia de lixiviados suele venir asociada a la presencia de gases producto de la descomposición de los residuos, tema que no trataremos en este post pero que deberá tenerse en cuenta cuando se diseñe un vertedero (Maquinaria, 2016).

Metano (CH₄): El metano es el hidrocarburo alcano más sencillo, cuya fórmula química es CH₄. Cada uno de los átomos de hidrógeno está unido al carbono por medio de un enlace covalente. Es una sustancia no polar que se presenta en forma de gas a temperaturas y presiones ordinarias. Es incoloro, inodoro e insoluble en agua. En la naturaleza se produce como producto final de la putrefacción anaeróbica de las plantas. Este proceso natural se puede aprovechar para producir biogás. Muchos microorganismos anaeróbicos lo generan utilizando el CO₂ como aceptor final de electrones. El gas natural lo contiene en diversas proporciones según el yacimiento de donde es extraído, desde el 83% al 97%. El gas natural comercializado es mayoritariamente metano con algunos otros hidrocarburos añadidos en pequeña proporción, como etano, propano, butano y algo de nitrógeno. No obstante, en las últimas décadas ha cobrado importancia la explotación comercial del gas metano de carbón, como fuente de energía. El metano es un gas de efecto invernadero relativamente potente que contribuye al calentamiento global del planeta Tierra ya que tiene un potencial de calentamiento global de 23. Esto significa que en una medida de tiempo de 100 años cada kg de CH₄ calienta la Tierra 23 veces más que la misma masa de CO₂, sin embargo, hay

aproximadamente 220 veces más dióxido de carbono en la atmósfera de la Tierra que metano por lo que el metano contribuye de manera menos importante al efecto invernadero.

Mineralización: Es el proceso por el cuál la materia orgánica se degrada a través de los microorganismos del suelo y se transforma en CO₂, H₂O y los compuestos minerales básicos tales como N₂, H₂S, Fe(II) etc.... dependiendo de que existan unas condiciones ambientales (aeróbicas o anaeróbicas) y microbiológicas (Bragado Alcaraz, 2018).

Neurona: Es un tipo de célula que representa la unidad estructural y funcional del sistema nervioso. Su función consiste en transmitir información a través de impulsos nerviosos, desde un lugar del cuerpo hacia otro. Estos impulsos nerviosos son impulsos químicos y eléctricos. En las neuronas más grandes las velocidades de conducción del impulso nervioso (velocidad de conducción axonal) pueden alcanzar hasta 120 m/s (Ambientech, 2020).

PH: El PH es una medida la cual indica cuál es la cantidad presente en una muestra de iones de hidrógeno. Es decir, estas siglas hacen referencia al Potencial de Hidrógeno o también Potencial de Hidrogeniones que hay en una solución o en una sustancia.

Red Neuronal Artificial: Son un modelo inspirado en el funcionamiento del cerebro humano. Está formado por un conjunto de nodos conocidos como neuronas artificiales que están conectadas y transmiten señales entre sí. Estas señales se transmiten desde la entrada hasta generar una salida (Sandra, 2019).

RSM (Relleno Sanitario Municipal): Se denomina relleno sanitario al espacio donde se depositan los residuos sólidos de una ciudad después de haber recibido determinados tratamientos. Para impedir que se contamine el subsuelo, se impermeabiliza el terreno con polietileno de alta densidad u otra sustancia y se coloca arcilla.

SARIMA: Proceso estadístico que se utiliza para hacer frente a la estacionalidad que se repite en todas las observaciones en una serie temporal. Seasonal Autoregressive Integrated Moving Average Model – Modelo Estacional Autorregresivo integrado de Media Móvil.

- Son compuestos orgánicos formados únicamente por átomos de hidrógeno y carbono.
- Son hidrofóbicos, esto es, insolubles en agua.
- Son lipofílicos, es decir, solubles en solventes orgánicos.

Test de Grubs: También conocida como prueba residual máxima normalizada o prueba de desviación extrema, es una prueba que se usa para detectar valores atípicos en datos univariados.

Transformación de datos: En estadística, la transformación de datos se efectúa para asegurarse que estos tienen una distribución normal (lo que puede ser un remedio para analizar datos que tienen otros tipos de distribución, falta de normalidad, linealidad, y homocedasticidad). Con la transformación de los datos estos se preparan para diferentes tipos

de análisis. Cuando se usan este tipo de pruebas estadísticas la transformación de datos puede ser necesaria ya que este tipo de análisis asumen que los datos deben ser lineales, normales y tienen homocedasticidad en su varianza.

Vertedero: Son sitios designados para tirar basura, basura u otros tipos de desechos sólidos. Históricamente, son los medios más comunes para eliminar los desechos sólidos que se entierran o se dejan apilar en pilas. El relleno sanitario es un fenómeno común en todo el mundo, especialmente debido a la mayor cantidad de desechos de nuestros hogares, escuelas, oficinas, hospitales y mercados. Algunos vertederos están bien gestionados y diseñados como parte de la gestión integrada de residuos (Decología, 2019).

Introducción

El Relleno Sanitario Urbano Bordo Xochiaca fue uno de los rellenos sanitarios más importantes en la Ciudad de México y el área metropolitana, la significancia de su terreno y la cantidad de toneladas de basura que recibió en toda su vida útil fue, principalmente, la fuente de producción de biogás. Geográficamente, la posición del Relleno Sanitario Urbano Bordo Xochiaca en su IV etapa es la siguiente.

Se ubica al noreste de la Ciudad de México en el municipio de Texcoco, Estado de México (sobre el kilómetro 1.2 de la carretera Peñón-Texcoco). A unos 11 km en línea recta del Centro Histórico, y a no más de tres kilómetros del Aeropuerto Internacional de la Ciudad de México (AICM), la Alameda Oriente y el centro comercial Ciudad Jardín Bicentenario. El terreno del Bordo Xochiaca posee una extensión aproximada de 420 hectáreas, en una poligonal de 5 lados: 1.92 km al noreste (colindando con el Circuito Exterior Mexiquense), 2.2 km al sureste, 1.7 km al suroeste, 1.4 km al noroeste y 1.2 km al norte (hacia la carretera Peñón Texcoco) (GUTIÉRREZ, 2013).

El biogás que se obtiene en cualquier relleno sanitario es importante, ya que inicialmente son lugares de disposición final de los residuos sólidos, adicionalmente una vez clausurado el relleno sanitario este gas es monitoreado, tratado y obtenido diariamente. El biogás se compone normalmente entre el 45% y 60% de metano, en donde los niveles de aprovechamiento de este combustible ecológico son entre 55% y 60%, los demás componentes que lo estructuran son, principalmente, el dióxido de carbono (CO₂) y Oxígeno (O₂). El monitoreo que se le da al biogás es la razón principal para poder tener una concentración de información que facilite la obtención de la muestra a estudiar. Teniendo en cuenta que las muestras son tomadas diarias en intervalos de horas constantes en el Relleno Sanitario Urbano Bordo Xochiaca.

El metano por su parte es uno de los gases más abundantes en la atmósfera siendo este un gas que contribuye al calentamiento global. Dentro del relleno sanitario el metano obtenido es enviado al quemador que se encuentra catalizando a este gas en todo momento convirtiéndolo en CO₂. El metano al ser catalizado puede llegar a generar una conversión energética ya sea para obtención de electricidad o poder calorífico.

Algunas definiciones para las metodologías de estimación y la composición de los gases obtenidos dentro del sitio estudiado serán conceptos que podremos encontrar en esta tesis.

Como selección del tema de esta tesis, debemos tener en cuenta primero, que se llegó a generar la búsqueda de artículos que relacionaran los temas de predicción y nos dieran referencias para generar algún tipo de criterio que favoreciera a los resultados que se obtuvieron, esto fue considerado inicialmente, pero durante la recolección de lecturas nos dimos cuenta que no había una gran cantidad de artículos de donde generar estas comparaciones, por lo que el elegir este tema tiene como apoyo el dar a conocer un sitio

geográfico dentro de la CDMX que actualmente funciona y genera energías alternativas, y que podemos llegar a tener un alcance de información para sus predicciones.

Hay que considerar que el Relleno Sanitario Urbano Bordo Xochiaca tiene un acercamiento a la Ciudad de México, que se encuentra en tercer lugar de las más grandes por extensión territorial en el mundo por lo que los expertos del sitio son ingenieros que llevan años en proyectos de rellenos sanitarios para la recolección de basura o lugares parecidos en ciudades dentro del país. Los expertos llevan el control de monitorear y de que no existan anomalías, entre ellas, esta cuando la máquina que registra las muestras de metano no llega a generar información, causa principal del análisis de esta tesis.

Cualquier muestra que cumpla con ser una serie de tiempo, permite ser analizada para aplicar metodologías orientadas a series estacionarias. Para esta investigación se nos proporcionaron datos necesarios, pero no regulares, esto quiere decir que no se comportan como una distribución normal. Se propondrá un tratamiento a la serie de tiempo por medio de la metodología de promedio móvil ponderado donde esta suaviza a la serie. El procedimiento que se llevó a cabo para proponer un tratamiento fue el análisis espectral de series de tiempo ya que este tiene como aplicaciones generar diferentes gráficos y pruebas que nos ayudan a visualizar con más nitidez y acercamiento los componentes de una serie de tiempo para poder definir qué posibles tratamientos son óptimos y también para que una vez aplicados estos tratamientos volver a cotejar sus componentes de serie de tiempo, y validar su mejoramiento o no.

Para la investigación realizada se tiene como propósito llegar a generar un proceso automático en ciertos intervalos de tiempo, ya sea cada semana o cada mes, pero durante los acercamientos al sitio se nos mencionaba que no se tenían equipos que soportaran ciertos programas, por lo que, el generar macros no era una opción.

Dentro de lo establecido para la tesis; dado que los datos se encuentran dañados se permitió generar un ajuste de modificación a la serie que llega a ser favorable y de aceptación para futuros valores de metano obtenido en el Relleno Sanitario Urbano Bordo Xochiaca. De igual forma se buscaron implementar modelos matemáticos para pronosticar la producción de metano en el Relleno Sanitario Urbano Bordo Xochiaca. Estas dos son el objetivo principal de la tesis. En esta investigación podrás encontrar algunas comparaciones de modelos, ante una misma variable de interés que es obtenida de un sitio geográficamente accesible y orientado con fines ecológicos.

Las metodologías que se llevan a la práctica permiten generar estimaciones a series de tiempo estacionarias, facilitando poder ajustar los componentes que la componen. Las metodologías utilizadas en esta tesis fueron el promedio móvil ponderado, Holt-Winters, Box Jenkins y red neuronal artificial. Todas fueron partícipes de los mismos análisis, pruebas y resultados para series de tiempo, esto se realizó con el propósito de llevar un proceso en cascada en todas las metodologías. La serie de tiempo que utilizamos, por su naturaleza de la muestra, tiene

valores atípicos que requieren ajustarse. La serie es estacionaria y de igual forma llega a tener estacionalidad como componente de ajuste para las metodologías. Se considera en inicio que la red neuronal debe generar una mejor estimación con un error favorable para la entrega a los expertos del relleno sanitario esto debido a la aplicación de la metodología de secuencia, y que Holt-Winters debe llegar a tener un mejor ajuste que Box-Jenkins y que promedio móvil ponderado dado que esta metodología esta dirigida a datos cambiantes. La comparación de los residuales es el factor de elección del mejor modelo que se realizó y mediante el valor de los errores se aceptó una metodología para estimaciones en el relleno sanitario. Una de la hipótesis considerada inicialmente es el periodo posible para elegir para conseguir una serie de tiempo estacionaria ya sea diaria, semanal, quincenal y mensual.

Uno de los objetivos esperados de investigación es llegar a obtener un pronóstico de metano con el error R cuadrada mayor y un error medio absoluto pequeño, generar un resultado que se entregue al experto cada cierto tiempo y poder seguir utilizándolo en predicciones futuras o en alguna otra variable de interés, ya sea la metodología o el pronóstico. Para considerar que el objetivo favorezca al experto del sitio, se analizaron primero las ventajas y desventajas de las metodologías, para identificar qué componentes suavizan, posteriormente se ejecutó la metodología, gráficos de distribución y dispersión, tablas de comportamiento, gráficos de residuales, pruebas a residuales y errores de cada estimación, con el fin de comparar cada resultado, estos serán los objetivos secundarios que dan la pauta de elección para las estimaciones.

Describiendo un poco el proceso de cada metodología podríamos dividir subprocesos de cada una de ellas como en la *Tabla 1: Procesos de lo muestra*.

Metodologías	Identificación	Estimación	Verificación	Pronostico
Promedio móvil ponderado	Elección del numero de ponderaciones. Componentes de series de tiempo.	Calculo de la ecuación de suavizamiento.	Ajuste de los pesos con el menor error.	Resultados de la serie a corto plazo.
Holt-Winters	Componentes de serie de tiempo. Método aditivo o multiplicativo.	Patrón estacional, tendencial y constante.	Normalidad de residuales y menor error.	Resultados de la serie a corto y mediano plazo.
Box-Jenkins	Componentes de serie de tiempo. Funciones de autocorrelación. Comportamiento de estacionariedad.	Coefficientes de los términos autorregresivos y de media móviles	Comprobación de comportamiento de residuales como ruido blanco gaussiano.	Resultados de la serie a corto y mediano plazo.
Red neuronal artificial	Tipo de red neuronal. Función secuencial.	Metodo de activación. Epocas de entrenamiento.	Ajuste de los pesos con el menor error.	Resultados de la serie a corto, mediano y largo plazo.

Tabla 1: Procesos de las metodologías.

Algo importante que debe ser mencionado es que el experto dentro del relleno sanitario pueda tener un valor pronosticado que le permita reemplazar en caso de que la máquina llegara a fallar y con ello tomar alguna decisión que afecte de forma directa el monitoreo o mantenimiento del sistema.

La aportación que se aplicó en todo el proceso, la recolección de información, el acudir al sitio, la investigación de artículos relacionados a temas del biogás o de algún relleno sanitario, claridad del tema para sus aplicaciones, conceptos de programación entre otras tantas definiciones son la relevancia de poder llegar a generar un pronóstico para apoyar las decisiones en un sitio de interés ecológico.

La aplicación de esta tesis tiene un formato de entrega con los expertos del sitio, puesto que se tiene tanto los ejecutables como los procesos para replicar esto en tiempo real. Al tener un pronóstico y aplicarlo en el sitio es posible tener un resultado favorable ante la falta de información que llega a ser constante.

La investigación contribuye a generar un proceso y aplicación para datos reales que pueden ser tratados junto con herramientas de software libre como Rstudio, Excel o Python. Los resultados y procesos apoyan a que los expertos tengan un criterio o justificación para tomar una decisión para futuras eventualidades. El aplicar una serie de tiempo a muestras de biogás visualiza lo relevante de las aplicaciones de nuestra carrera en otras áreas de investigación. Ahora bien, se describirán en términos generales el contenido de cada capítulo a continuación.

En el capítulo 1 se describen las características de un relleno sanitario, así como la ubicación del Relleno Sanitario Urbano Bordo Xochiaca. También la generación del biogás es estos sitios y los componentes que la conforman. Dentro de los componentes del biogás se describe al gas metano y sus características generales. Por último, se podrán observar ventajas y desventajas generales de los temas abordados.

En la sección 2 explica las definiciones de los temas matemáticos requeridos para generar un modelo de pronóstico, por lo que las definiciones y características de las series de tiempo; qué es un pronóstico, proceso de las metodologías, ventajas y desventajas de cada una de ellas y qué relevancia se tiene ante la variable de interés, van a ser comentadas dentro de esta sección.

Para el capítulo 3 se presenta la descripción de los datos para poder obtener la muestra, así como la distribución de estos, de igual importancia se explica la serie de tiempo con o sin tratamiento de datos atípicos, pruebas de raíz unitaria y comparativos de las series en el tiempo. Dentro de los últimos puntos, se describe cómo se utilizó el criterio de ponderación para suavizar a la serie de metano.

Con respecto a la sección 4 se puntualiza, en primera instancia, las combinaciones posibles en la periodicidad de la serie de tiempo, la variable con y sin tratamiento y el día de la semana a considerar. Posteriormente se generan las metodologías para los pronósticos esperados de metano. La estructura general de este capítulo está descrita por la aplicación del método, posteriormente el comportamiento de la estimación, después la validación del modelo y por último el pronóstico.

En cuanto a la última sección se tienen los resultados de cada una de las metodologías expuestas anteriormente, la relación que se tiene es idéntica ante la descripción del capítulo

4 ya que se describe cada uno de los subtemas para llegar a tener un resultado comparativo de las metodologías aplicadas. En este capítulo se describe también cómo los valores de los errores son importantes para la aceptación del modelo.

Capítulo 1 Relleno Sanitario Urbano Bordo Xochiaca, antecedentes, procesos y obtención de metano

1.1 Relleno Sanitario Urbano Bordo Xochiaca

Actualmente la basura es uno de los mayores problemas que intervienen en la contaminación del medio ambiente, lo cual ha causado que durante los últimos años se busque la implementación de distintas estrategias que den soluciones al reciclaje, reducción y reutilización de la basura. Muchas de las soluciones propuestas ante este problema global son aquellas en donde estos desechos provenientes del petróleo sean fuentes de energía que permita reemplazar gran parte de las fuentes no renovables y que afecten directamente la degradación de este residuo, transformándolo en un compuesto con características favorables que apoyan al medio ambiente, estas propuestas son implementadas en las grandes ciudades, en sitios geográficamente estudiados y manipulados para la disposición final de los residuos que desechamos los seres humanos.

Consideramos basura todo aquello que ya no nos sirve. Cuando regresamos de compras, las bolsas y los empaques los tiramos a la basura, porque ya no nos sirven. También arrojamos al bote de basura la fruta que ya no podemos comer porque se echó a perder, el vidrio que se rompió y el trapo que ya no utilizamos (Acurio, 1997). Acostumbramos a tirar la basura en cualquier lugar y nos consideramos educados si tenemos el hábito de depositar la basura en su lugar. Creemos que con esto ya hemos cumplido. Sin embargo, el problema de la basura no se resuelve con esto. Cuando la envolvemos en bolsas y la sacamos para que el barrendero de la ciudad la recoja, lo que estamos haciendo es "pasar el paquete" de la basura al servicio de limpieza (Reyes, 2004).

En la actualidad existen lugares destinados para la acumulación de basura, la cual dada las distintas necesidades de una comunidad pueden variar sus características y ubicaciones, teniendo como resultado distintos tipos de basureros que pueden llegar a tener materiales orgánicos e inorgánicos o combinación de ambas (Yauli, 2012). Con el fin de llevar un control de los desechos, cada comunidad tiene normas y criterios que facilitan la reutilización y mantenimiento de estos sitios, ya que se encuentran normalmente a las orillas de las ciudades, pueblos, municipios, etc. A estos sitios se le atribuyen distintos nombres que dan al final una idea general de "tiraderos a cielo abierto", pero que a la larga se vuelven una fuente considerable de componentes orgánicos que a su vez pueden llegar a generar energía eléctrica.

Un Relleno Sanitario Urbano (RSU), es un sitio geográfico de disposición final en el cual está la basura de una comunidad, ciudad o municipio cercanos a este. Las características y funciones de un RSU varían depende al residuo, zona geográfica o por el fin de generar metano entre otras orientaciones siempre con dirección a la basura y al entorno en el que se encuentra.

Los RSU, son aquellos lugares que cumplen con normas oficiales de higiene y mantenimiento en zonas geográficas establecidas para su proceso de control de la basura. En México existen distintos lugares con estas características, por lo cual, para centrar el objetivo de predicción para la toma de decisiones de esta tesis, nos enfocaremos en el Relleno Sanitario Urbano de Bordo Xochiaca (RSU BX) del cual se obtiene la información requerida (una serie de tiempo) y el cual se encuentra ubicado en Nezahualcóyotl, Estado de México.

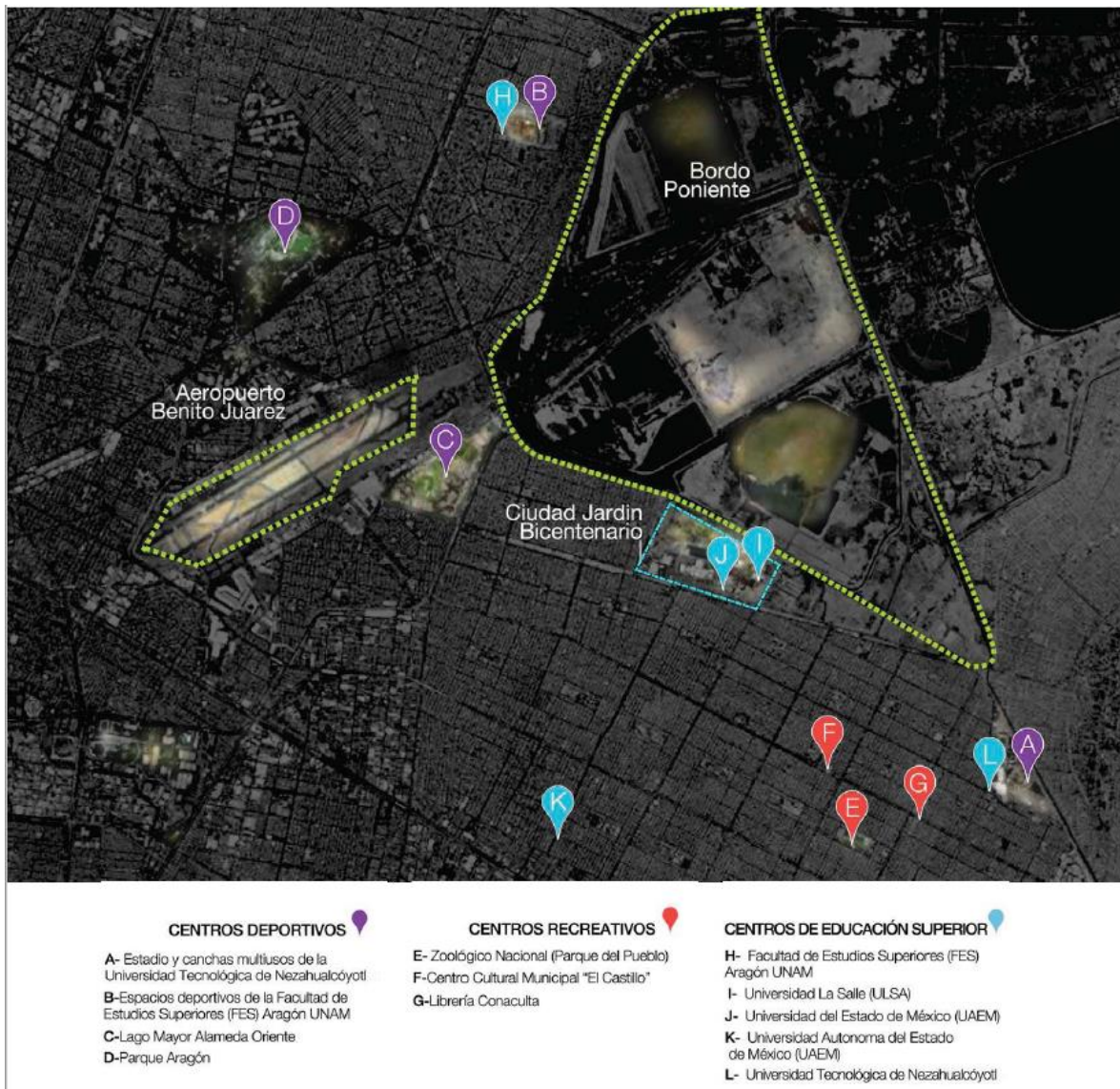
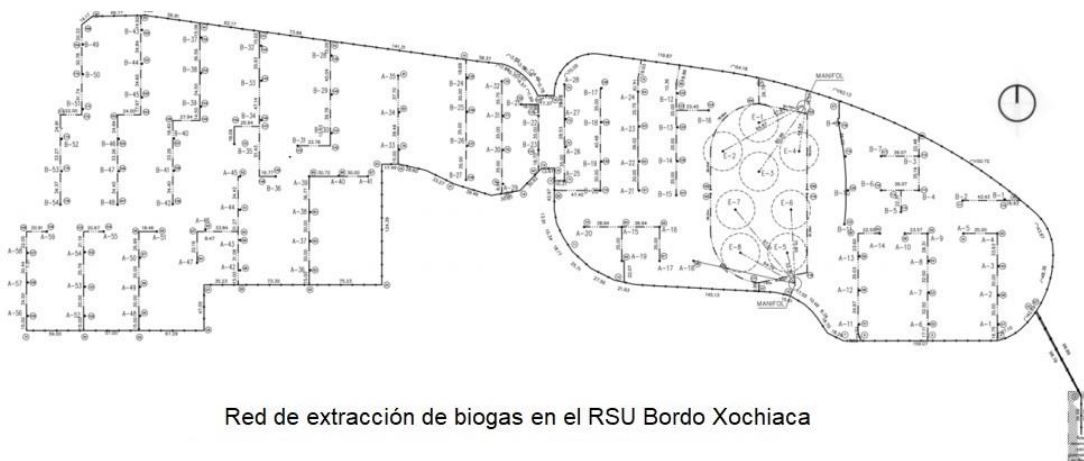


Ilustración 1: Mapa relleno sanitario Bordo Xochiaca (*Dirección General de Bibliotecas, 2014*).

Ahora bien, regresando al contexto de un “tiradero a cielo abierto” podemos decir que todo basurero con cierto seguimiento puede llegar a convertirse en un contenedor de gases, pero si desde el inicio de la construcción del RSU y gracias a un estudio geográfico previo, un

“tiradero a cielo abierto” puede llegar a generar una gran concentración de gases favorables para el medioambiente y con base a las normas de cumplimiento llegar a un funcionamiento adecuado para no ser un peligro tanto para la comunidad cercana a ella como a la fauna y flora en donde está ubicada, todo esto para tener resultados que generen energías alternas a las actuales que son en muchas formas causantes de problemas ambientales.

El relleno sanitario es una técnica de disposición final de los residuos sólidos en el suelo que no causa molestia ni peligro para la salud o la seguridad pública; tampoco perjudica el ambiente durante su operación ni después de su clausura. Esta técnica utiliza principios de ingeniería para confinar la basura en un área lo más estrecha posible, cubriéndose con capas de tierra diariamente y compactando la para reducir su volumen. Además, prevé los problemas que puedan causar los líquidos y gases producidos por efecto de la descomposición de la materia orgánica.



Red de extracción de biogás en el RSU Bordo Xochiaca

Ilustración 2: Mapa de la red extracción forzada de biogás (*Dirección General de Bibliotecas, 2014*).

Para la construcción del RSU BX, se consideran cumplir características que deben tener los RSU, tanto para el ecosistema, tales como el posible efecto que pueda ocasionar al ser humano. Estas características tienen que ver con la topografía del lugar, sistemas de drenajes para los lixiviados, producción diaria de biogás e infraestructura del sitio. Para el efecto de esta tesis los lixiviados no serán analizados. Más adelante, se describe cómo se obtiene el biogás y sus características para poder obtener al metano, gas que más compone a este combustible ecológico.

Ahora bien, a los espacios en donde existe un almacén con un cierto volumen de basura que posteriormente pasó a ser un RSU, se debe generar una estructura que permita el tratamiento y funcionamiento para la reutilización del gas generado en el suelo. La ilustración 1: Relleno Sanitario, visualiza la estructura general de los RSU.

Cómo es un relleno sanitario



Ilustración 3: Relleno Sanitario (WordPress, 2014).

Las funciones del RSU son fundamentales para un mayor control, tanto para las decisiones de producción de biogás como para la armonía del ecosistema, dichas decisiones van desde la creación de planes que minimicen los factores tanto internos como externos que causan daños al ecosistema o sociedad cercana a estos, así como las herramientas para la medición y captación del biogás. La información de los residuos alojados, tratamiento de lixiviados, análisis de riesgo en los pozos, estructura de los pozos, maquinarias de extracción, etc., permiten un mejor abastecimiento de biogás, así como la producción de metano diario producido en estos sitios ecológicos.

Las características y usos finales de los RSU son factores importantes actualmente en la sociedad, puesto que permiten tener un mejor control para la basura en lugares con una gran cantidad de población, permitiendo tener mejor uso ecológico para las futuras generaciones. Por lo que genera un nuevo enfoque a las posibles soluciones para la producción de energía eléctrica ecológica en los RSU.

Ventajas:

- La inversión inicial de capital es inferior a la que se necesita para instaurar el tratamiento de residuos mediante plantas de incineración o de composta.
- Tiene menores costos de operación y mantenimiento que los métodos de tratamiento.
- Un relleno sanitario es un método completo y definitivo, dada su capacidad para recibir todo tipo de RSM.
- Genera mano de obra poco calificada, disponible en abundancia en los países en desarrollo.

- Recupera gas metano en los rellenos sanitarios que reciben más de 500 toneladas/día, lo que puede constituir una fuente alternativa de energía para algunas ciudades.
- Su lugar de emplazamiento puede estar tan cerca del área urbana como lo permita la existencia de lugares disponibles, lo que reduce los costos de transporte y facilita la supervisión por parte de la comunidad.
- Permite recuperar terrenos que se consideraban improductivos o marginales, tornándose útiles para la construcción de parques, áreas recreativas y áreas verdes, etc.
- Un relleno sanitario puede comenzar a funcionar en corto tiempo como método de eliminación de residuos.
- Se considera flexible porque puede recibir mayor cantidad adicional de residuos con poco incremento de personal.

Desventajas:

- La adquisición del terreno es difícil debido a la oposición de los vecinos al sitio seleccionado, fenómeno conocido como NIMBY (*not in my back yard*), por diversas razones: la falta de conocimiento sobre la técnica del relleno sanitario, se asocia el término relleno sanitario al de botadero a cielo abierto, la evidente desconfianza mostrada hacia las administraciones locales que no garantizan la calidad ni sostenibilidad de la obra y la falta de saneamiento legal del lugar.
- El rápido proceso de urbanización, que limita y encarece el costo de los pocos terrenos disponibles, lo que obliga a ubicar el relleno sanitario en sitios alejados de la población.
- La vulnerabilidad de la calidad de las operaciones del relleno y el alto riesgo de transformarlo en un botadero a cielo abierto, principalmente por la falta de voluntad política de las administraciones municipales para invertir los fondos necesarios a fin de asegurar su correcta operación y mantenimiento.
- No se recomienda el uso del relleno clausurado para construir viviendas, escuelas, etc.
- La limitación para construir infraestructura pesada por los asentamientos y hundimientos después de haber clausurado el relleno sanitario.
- Se requiere un monitoreo luego de la clausura del relleno sanitario, no sólo para controlar los impactos ambientales negativos, sino también para evitar que la población use el sitio indebidamente.
- Puede ocasionar impacto ambiental de largo plazo si no se toman las previsiones necesarias en la selección del sitio y no se ejercen los controles para mitigarlos.
- En rellenos sanitarios de gran tamaño conviene analizar los efectos del tráfico vehicular, sobre todo de los camiones que transportan los residuos por las vías que confluyen al sitio y que producen polvo, ruido y material volante. En el vecindario el

impacto lo generan los líquidos, gases y malos olores que pueden emanar del relleno.

- Los predios o terrenos situados alrededor del relleno sanitario pueden devaluarse.
- En general, no puede recibir residuos peligrosos (WordPress, 2014).

1.2 Biogás en el Relleno Sanitario Urbano Bordo Xochiaca

Como se ha mencionado anteriormente el RSU tiene propiedades favorables para una solución distinta de crear energías ecológicas y de sustituir energías no renovables que dañan al planeta, considerando lo anterior, esta tesis busca favorecer las decisiones para el comportamiento de metano mediante una estimación a corto plazo dentro del RSU BX.

Entonces se sabe que con base en los desperdicios que se generan en las ciudades o zonas urbanas de gran cantidad de población la solución de obtener biogás una vez clausurado para su uso y mantenimiento al RSU se requiere de herramientas que faciliten la obtención de este gas ecológico. Hay que considerar que, en México, la población en zonas rurales generalmente tiene sus propias fosas en donde queman o solo tiran la basura, el efecto de obtener biogás sería una manera favorable de ayudar al medio ambiente de estos lugares que dañan al ecosistema.

Bordo Xochiaca se encuentra actualmente clausurado, por lo cual, está siendo controlado y monitoreado para observar el comportamiento de biogás diariamente, la información para la predicción de metano fue proporcionada por ingenieros expertos en estos sitios y con lo cual ellos buscan que se apoye al mejoramiento y orientación de decisiones diarias a fin de aprovechar la producción de metano durante los años que aún tiene por producir.

La instalación de la red actualmente en Bordo Xochiaca permite gestionar y observar tanto la infraestructura como la producción de los gases, como a su vez de líquidos, robos de tuberías entre otras cosas. Es un trabajo de cuidar situaciones correspondientes al entorno social del lugar geográfico, ya que al estar expuestos en zonas de gran tamaño visualizar un robo o detectar una falla de instalación puede tardar dos o más días. Es considerable también mencionar que al igual que las fallas externas, algunas fallas internas pueden causar, como en este caso de investigación, que la máquina se descompone en lapsos de tiempo medianamente largos y que los datos de estos períodos serán tratados como la transformación necesaria para la estimación de metano en el RSU BX.

Para tener en cuenta cómo se genera el biogás se requiere que la basura o los residuos en el subsuelo se descompongan, se debe saber que la descomposición de los residuos como lo pueden ser plásticos, madera, vidrio, etc. que se producen diariamente en las zonas pobladas son introducidas en capas que favorezcan su deterioro, este proceso puede ser visualizado en la ilustración 1: Relleno Sanitario, de donde en las capas inferiores por donde los líquidos

son extraídos, los residuos empiezan a ser descompuestos y al no haber oxígeno (condiciones anaerobias) en estas capas la producción de biogás es más abundante, puesto que la descomposición de los residuos es más rápida.

Cualquier sustancia o compuesto se denomina como biodegradable si puede ser descompuesto por la acción de microorganismos, ya sea en condiciones aerobias o anaerobias. Una sustancia puede ser convertida por dichos microorganismos hasta productos finales de carbono CO_2 y CH_4 , esta sustancia final puede llegar a ser usada como una fuente de energía; a este proceso se le conoce como mineralización. No todos los microorganismos poseen el metano o la capacidad para generar dichos productos finales, de la misma manera que no todos los microorganismos dentro de un reactor pueden usar los sustratos en su forma original; a la generación de sustancias intermediarias dentro del proceso biológico se le conoce como biodegradación primaria (WordPress, 2014).

El biogás es una mezcla gaseosa formada principalmente de metano y dióxido de carbono, pero también contiene diversas impurezas. La composición del biogás depende del material digerido y del funcionamiento del proceso. Cuando el biogás tiene un contenido de metano superior al 45% es inflamable. El biogás tiene propiedades específicas que se indican en la Tabla 2: Características generales del biogás.

Composición	55-70% metano (CH_4) 30-45% dióxido de carbono (CO_2) <5% trazas de otros gases
Poder calorífico	6.0 - 6.5 $Kwhm^3$
Limite de explosión	6 - 12 % de biogás en el aire
Temperatura de ignición	650 - 750°C (con el contenido de CH_4 mencionado)
Presión crítica	74-88 atm
Temperatura crítica	-82.5°C
Densidad normal	1.2 $kg\ m^{-3}$
Olor	Huevo podrido (el olor del biogás desulfurado es imperceptible)
Masa molar	16.043 $kgmol^{-1}$

Tabla 2: Características generales del biogás (Varnero Moreno, 2011).

Las características de biogás para la obtención de metano varían depende la zona geográfica en la que se encuentre, ya que al ser producida por descomposición anaerobia (sin oxígeno), más las temperaturas externas de la zona elegida contribuyen a su generación, por lo que estas características son distintas en cada RSU creado. Como se ha mencionado, los riesgos tanto de un RSU y la producción de biogás al ser inflamable en proporciones de metano mayores al 45% de la composición total de biogás, son las principales razones por las que el

control y mantenimiento de estos deben ser constantes y parcialmente duraderas mientras exista biogás en el subsuelo.

Como sabemos el biogás permite poder obtener una forma distinta de energía eléctrica y esto a su vez ayuda a generar alternativas de protección para el medio ambiente, la reducción de basura con el objetivo ecológico es esencial entenderlo y practicarlo como a su vez la reutilización, reducción y reciclaje. Si tenemos la opción de crear algún tipo de RSU podríamos tener planes a largo plazo para atacar estos problemas.

Ventajas del biogás:

- El biogás es ecológico.
- La generación de biogás reduce la contaminación del suelo y el agua.
- La generación de biogás produce fertilizante orgánico.
- Es una tecnología simple y de bajo costo que fomenta una economía circular.
- Es una fuente amplia de energía.
- Menor contaminante.
- Importante sustituto de los derivados del petróleo.
- Puede usarse también para la generación de electricidad.
- Fuente de ingreso de agricultores.
- La producción de residuos es reversible a energía limpia.
- Da un propósito útil a los vertederos.
- Alternativa de uso en cocinas.

Desventajas del biogás:

- Pocos avances tecnológicos.
- Contiene impurezas.
- Efecto de la temperatura en la producción de biogás.
- Menos adecuado para áreas metropolitanas densas.
- Sistema de almacenamiento complejo y difícil de maniobrar.
- Emisión de dióxido de carbono CO_2 (Cruzito, 2014).

Hasta ahora hemos hablado de la importancia de un RSU para la disposición final de los residuos que desechamos diariamente, también describimos cómo se genera en el subsuelo el biogás y el lugar para RSU BX para la aplicación de la tesis. También consideremos que el biogás es clasificado como un gas favorable para energías alternas y para el medio ambiente, por lo que en este estudio solo nos centraremos en el metano dado que las demás variables como el flujo de la red, la temperatura, el CO_2 entre otros, se considerarán para investigaciones futuras.

1.3 Metano en la composición del biogás

El metano es un gas y es, después del bióxido de carbono, el compuesto por carbono más abundante en la atmósfera. También se le conoce como el hidrocarburo alcano más sencillo que existe, es decir, que están compuestos por átomos de carbono e hidrógeno. Es uno de los gases con mayor abundancia y en especial es un gas problemático ya que en proporción, el metano atrapa 28 veces más calor que su equivalente de bióxido de carbono y 32 veces si se considera su interacción con aerosoles. Actualmente el metano es aproximadamente el 20% de los gases de efecto invernadero que se producen por acción humana.

El metano que se encuentra en el RSU BX tiene una estimación de producción de aproximadamente 20 años, por lo cual es importante considerar las posibles producciones mediante un pronóstico que tenga una estimación favorable y de mejoramiento en la obtención de datos, considerando los posibles tratamientos implementados en las metodologías elegidas.

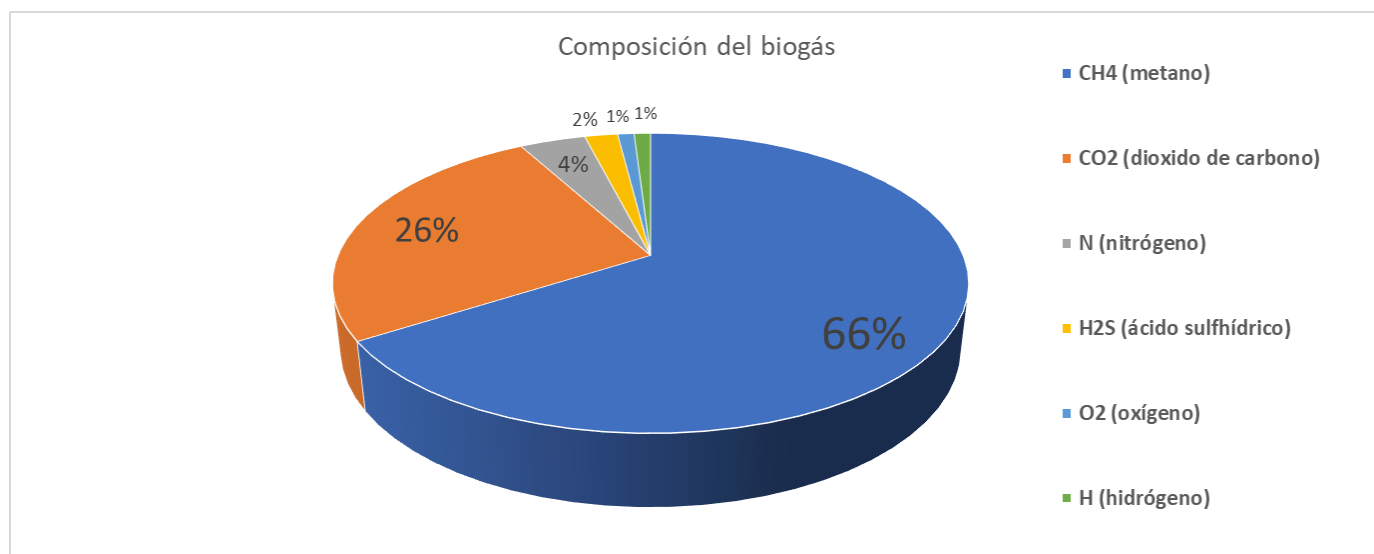


Ilustración 4: Metano en el biogás (ENERGÌA, 2020).

Bordo Xochiaca maneja su propio monitoreo y extracción del metano, ingenieros especializados dentro del RSU controlan el catalizador y mantienen un constante seguimiento de las instalaciones, generando un mejor flujo de biogás y más producción. Estas instalaciones permiten registrar la producción diaria de metano que permitirá analizar los datos y obtención de la serie de tiempo descrita en el capítulo 3 de esta tesis y con ello realizar una estimación, apoyando a las decisiones internas que mejoren el funcionamiento de la máquina y que a partir de esto se pueda generar energía eléctrica.

El metano al ser el dato de investigación se debe interpretar como una serie de tiempo por lo cual el proceso y seguimiento de los ingenieros encargados del sitio causa que se tenga una información real. El programa de monitoreo tiene como objetivo lo siguiente:

- Llevar a cabo la operación y mantenimiento, tanto del sistema de extracción, conducción y combustión del biogás; como del sistema de captación, acondicionamiento y filtración de lixiviados.
- Realizar las mediciones del flujo, composición y presión del biogás, en toda la red de extracción de la zona deportiva, que permite mantener las condiciones de presión, flujo, contenido de oxígeno y porcentaje de metano en la mezcla de biogás, para lograr su combustión en forma eficiente, en las condiciones operacionales que requiere el quemador.
- Medir los parámetros básicos que permitan determinar el buen funcionamiento del sistema de captación, acondicionamiento y filtración de lixiviados, como es el pH y los sólidos; así como realizar los muestreos de los lixiviados frescos y acondicionamientos para su envío al laboratorio, con la frecuencia que se requiera.
- Ejecutar un programa de mediciones, que permita registrar el porcentaje de metano y nivel de explosividad en diferentes puntos de la zona deportiva y sus colindancias, definidos precisamente con base en un escenario de alto riesgos, derivado de condiciones de operación críticas, por el paro de los sistemas de extracción y combustión de biogás, debido a eventualidades que invariablemente, derivan en dicho supuesto.
- Identificar las zonas de riesgo por donde pueda presentarse cualquier liberación no-controlada de biogás, en la periferia de las plataformas donde se desplantan las instalaciones deportivas del RSU BX, en las propias instalaciones deportivas en el entorno del hospital vivo; así como en las colindancias con el Campus de la Universidad Autónoma del Estado de Morelos, en la Colonia del Sol y la Zonas Comercial y de Servicios en general (Dirección General de Bibliotecas, 2014).

Capítulo 2 Métodos de pronósticos para series de tiempo

2.1 Serie de tiempo para la predicción de metano

Existen diferentes formas de generar una estimación o predicción de datos, pero en la mayoría de las ocasiones los datos contienen comportamientos irregulares que causan incertidumbre a la hora de pronosticar un modelo, por lo que se debe considerar la metodología que minimice la irregularidad, optimice el proceso de predicción y que a su vez facilite la interpretación de la información.

El RSU BX produce biogás y como se menciona en el capítulo 1 la composición de este gas en su mayoría es el metano, del cual se puede obtener por medio de un sistema de extracción que es controlada por los especialistas del sitio y gracias a esto se mantiene funcionando en la actualidad. Como consecuencia del funcionamiento se obtiene un conjunto de datos los cuales serán utilizados como una muestra de análisis, serie de tiempo.

Una serie de tiempo puede interpretar el comportamiento histórico de una variable de interés y mediante características matemáticas puede ser utilizada para estimaciones futuras favoreciendo la toma de decisiones. Para una buena toma de decisión se requiere que el analista que implementa la metodología, cuál sea ésta, cuente con la habilidad de interpretar y solucionar comportamientos no esperados dentro del proceso elegido, esto es de los principales puntos a analizar a la hora de poder trabajar con series de tiempo.

La información de metano como se ha mencionado será una serie de tiempo que dependerá de la muestra tomada de la máquina de extracción, dado que las muestras son tomadas continuamente. Para la explicación de la muestra tomada será descrita en el capítulo 3 de la tesis. La forma de estimación de los datos será interpretada bajo los siguientes conceptos.

Una serie de tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo ordenados cronológicamente y espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí. El principal objetivo de una serie de tiempo X_t donde $t = 1, 2, 3, \dots, n$ es un análisis para hacer pronósticos (Villavicencio, 2004).

Las series de tiempo tienen componentes que determinan un riesgo de estimación a los datos. Los datos al ser observados y analizados tendrán componentes que, de no ser tratados, puedan generar resultados inesperados, y, como consecuencia de esto, retrasar las decisiones ante situaciones reales. Si el modelo no es favorable con los datos obtenidos, se deberá replantear todo el proceso y en la mayoría de las ocasiones no optimiza tiempos, costos o el factor humano.

Una serie de tiempo tiene un total de cuatro componentes, por lo que a cada uno de ellos se le busca realizar un tratamiento. Esto dependerá de la metodología implementada de donde en su mayoría de las veces se aplican suavizamientos, favoreciendo el comportamiento de los

datos. Estos cuatro componentes que se buscan identificar son la tendencia, la estacionalidad, los ciclos y las irregularidades.

La descomposición clásica es un método que se basa en el supuesto que la serie de datos se pueden desagregar en componentes: tendencia, ciclicidad, estacionalidad e irregularidad; que se describen a continuación.

1. Tendencia

Una serie de tiempo con tendencia es aquella que contiene un componente de largo plazo que representa el crecimiento o declinación de la serie a través de un período amplio.

2. Variación estacional

Se define como estacional una serie de tiempo con un patrón de cambio en periodos menores o iguales a un año. Por lo regular, el desarrollo de una técnica de pronóstico estacional comprende la selección de un método multiplicativo o uno de adición y estimar después índices estacionales a partir de la historia de la serie.

3. Ciclo

El efecto cíclico se define como la fluctuación en forma de onda alrededor de la tendencia. Los patrones cíclicos tienden a repetirse en los datos cada dos, tres o más años. Es difícil establecer un modelo para estos patrones cíclicos, ya que no son estables y son difíciles de observar.

4. Irregular

El componente irregular de la serie de tiempo es el factor residual, es decir, “todo lo que sobra” y toma en consideración las desviaciones de los valores reales de la serie de tiempo en comparación con los esperados; es el elemento aleatorio.

Cada uno de los componentes de la serie de tiempo se busca ser analizado, por lo que se busca como objetivo minimizar el error de estimación ante la comparación de las metodologías descritas durante este capítulo. Para los datos de producción de metano los componentes tendrán su propio ajuste, esto dependerá del método de pronóstico ya que a cada uno de ellos afecta a uno o a varios de los componentes de la serie de tiempo.

Para identificar algunas de las variaciones de una serie de tiempo se pueden utilizar gráficas, pruebas o tablas (OLAYA, 2003).

A continuación, se visualizan algunas series de tiempo.

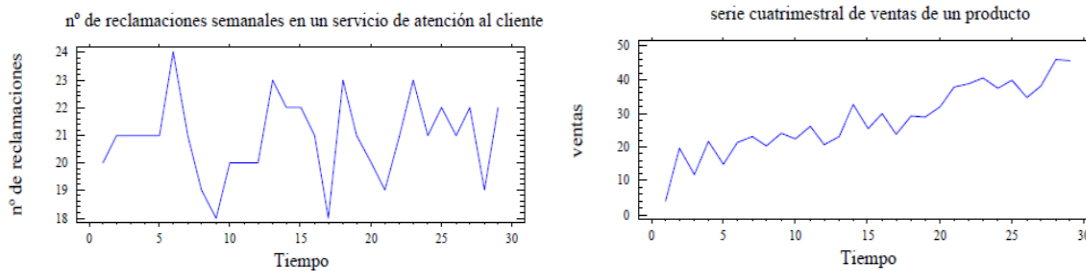


Ilustración 5: Series de tiempo.

Como se observa en las gráficas anteriores, las series de tiempo llegan a tener componentes notorios que las permiten identificar

Ahora bien, si se interpretaran las gráficas como series de tiempo podríamos encontrar patrones estacionales, tendencias e irregularidades durante el tiempo.

Los pronósticos que permitan identificar los componentes de tendencia, estacionalidad e irregularidad serán considerados para nuestros datos de interés ya que estos se encuentran como factores que afectan a nuestra muestra analizada. Se espera que las metodologías permitan eliminar la mayoría de los componentes que afectan a nuestra serie de tiempo. No olvidar que la ciclicidad es un factor que normalmente interpreta modelos económicos, por lo que el impacto de estimación en nuestra serie no será de análisis. A continuación, se describen algunas características que definen a estos modelos de predicción, considerando la importancia de pronosticar.

2.2 Modelos para predicción a series de tiempo

2.2.1 Definición de pronóstico

Un pronóstico es el resultado de un análisis espectral dada una muestra obtenida, con el objetivo de generar un nivel de aceptación para la toma de decisiones en eventos futuros. Sin dejar de considerar la incertidumbre del análisis implementado.

Pronosticar es el arte y ciencia de predecir acontecimientos futuros. Las decisiones siempre se toman con información insuficiente y con un margen de incertidumbre mayor o menor, dependiendo del tiempo y los recursos que se destinan a la búsqueda y el análisis de la información. Debemos entonces establecer algunas suposiciones y actuar. Sabemos que nunca ocurrirá exactamente lo pronosticado; entonces, ¿para qué pronosticar? Porque

necesitamos una cierta base, aunque sea mínima para generar, un criterio, una justificación que permita tomar una decisión a partir del pronóstico que generemos, favoreciendo y distribuyendo recursos financieros, tecnológicos y humanos (Cohen, 2000).

El hecho de generar un pronóstico que nos describa la producción de metano es favorecer el mantenimiento de la máquina de extracción que da paso al quemador de biogás, el cual cataliza el metano y lo transforma en energía eléctrica todo esto bajo el control y supervisión de los especialistas del sitio.

2.2.2 Promedio móvil ponderado

El promedio móvil ponderado es una metodología de suavizamiento que permite estimar a corto plazo una serie de tiempo, esta metodología tiene como particularidad utilizar el promedio de rangos de datos asignados por el analista y a cada uno de ellos asignarle una ponderación, de donde a esa ponderación la suma de sus k factores deberán ser siempre 1, agregando que al dato más reciente se le asigne el factor con mayor peso. Esta metodología permite con regularidad eliminar la tendencia de una serie de tiempo como a su vez reduce la irregularidad de los datos.

Este método de pronóstico es una variación del promedio móvil. Mientras, en el promedio móvil simple se le asigna igual importancia a cada uno de los datos que componen dicho promedio, en el promedio móvil ponderado podemos asignar cualquier importancia (peso) a cualquier dato del promedio (siempre que la suma de los pesos sea equivalente al 100%). Es una práctica regular aplicar el factor de ponderación (peso) mayor al dato más reciente. El pronóstico de promedio móvil ponderado es óptimo para patrones de demanda aleatorios o nivelados donde se pretende eliminar el impacto de los elementos irregulares históricos mediante un enfoque en periodos de demanda reciente, dicho enfoque es superior al del promedio móvil simple (López, 2019).

La ecuación de obtención del modelo es el siguiente:

$$Y_t = \sum_{t=1}^n X_{t-1} * C_i$$

$$i = 1, 2, \dots, n$$

De donde:

- Y_t es el promedio de datos en el periodo t.
- C_i Valor del factor de ponderación.
- X_{t-1} Valor inmediato anterior dado el tiempo t-1.
- N es el número de factores.

A continuación, mencionaremos algunas ventajas y desventajas de la metodología implementada.

Ventajas:

- Fácil de implementar.
- Tecnología accesible, solo hojas de cálculo.
- La interpretación de datos va referente a la media de información.
- No requiere de muchos datos históricos para tener un mejor comportamiento.
- El pronóstico obtenido es más eficiente que un modelo experto.

Desventajas:

- Ante valores atípicos el modelo puede no ser preciso en su estimación.
- Las tecnologías actuales mejoran las estimaciones, reduciendo el error.
- Al ser un modelo genérico, no hay un proceso común que permita adaptarse de manera precisa a los datos, determinando posibles irregularidades en sus estimaciones.
- No considera variables externas o dependientes de la serie de tiempo, teniendo como irregularidad la posible falta de correlación de los datos.
- Estima en periodos muy cortos, posiblemente de uno a dos tiempos dada la frecuencia de información obtenida.

Este método fue considerado por su ponderación constante y creciente según el tiempo más reciente. Permitiendo generar comparaciones con las siguientes metodologías de estimación.

2.2.3 Holt-Winters

Holt-Winters es una metodología de suavizamiento exponencial triple que permite detectar la tendencia y estacionalidad de una serie de tiempo. Como objetivo de estudio permitirá suavizar el comportamiento estacional identificado en el metano registrado en el RSU BX.

Holt-Winters es un proceso que parte del método de Holt de donde solo considera la tendencia y el nivel de una serie de tiempo por lo que Winters agrega un tercer suavizamiento el cual permite atacar al componente estacional, facilitando la detección creciente o constante en la serie de tiempo analizada.

En la mayoría de las ocasiones es un método para pronosticar la demanda a corto plazo y útil cuando los datos contienen una tendencia y un patrón estacional. Es popular porque es simple y es fácilmente automatizado. También es capaz de adaptarse a los cambios en los niveles. A continuación, se describen los componentes que se buscan suavizar en la serie.

- Nivel: Es el nivel constante que queda después de que hemos desestacionalizado la serie y se elimina el efecto de factores aleatorios (ruido).
- Tendencia: Este es el cambio en el nivel subyacente que se espera que se produzca entre hoy y el próximo mes. Refiere al delta de cambio (índice de incremento o decremento) entre los periodos de la serie “pendiente”.
- Índice de estacionalidad: Es la ponderación que se le da al valor pronosticado de acuerdo con el peso correspondiente con base al mismo periodo o periodos coincidentes.

Para cada componente hay una constante de suavización, que cae entre cero y uno. Las constantes de suavización más grandes significan más peso para los últimos datos (FUENTE, 2015).

Este método tiene dos principales modelos que dependen del tipo de estacionalidad observada y encontrada en la serie de tiempo.

El primero de ellos es el modelo multiplicativo estacional, este modelo presupone que a medida que se incrementan los datos, también se incrementa el patrón estacional. La mayoría de las gráficas de series de tiempo muestran este patrón. En este modelo, la tendencia y los componentes de estación se multiplican y luego se suman al componente de error.

El segundo modelo es el aditivo estacional, el cual parte de que los datos tienen o se visualizan efectos de los factores individuales ya que estos se pueden llegar a diferenciar y con ello agruparse para modelar a la serie de tiempo. Un modelo aditivo es opcional para los procedimientos de descomposición y para el método de Winters (Villaverde, 2017).

El método de Holt-Winters está constituido por la siguiente ecuación de donde se considera el proceso multiplicativo para representar las definiciones de cada una de las variables.

En este modelo la serie tiene una tendencia al menos localmente, y un patrón estacional creciente.

$$S_t = (xt/Ct-L) + (1 - \alpha)(St-1 - Bt-1)$$

$$B_t = (S1 - St-1) + (1 - \beta)-1$$

$$C_t = \gamma (xt/ St) + (1 - \gamma)-L$$

$$Ft+K = (St + KBt)Ct+K- L$$

De donde:

- S_t es el suavizado del patrón constante.
- B_t es el suavizado del patrón tendencial.
- C_t es el suavizado del patrón estacional.
- α, β, γ son constantes de suavizamiento $0 < \alpha, \beta, \gamma < 1$.
- L es la longitud de la estacionalidad (Fuente, 2015).

Ahora se considera el proceso aditivo para representar las definiciones de cada una de las variables.

En este modelo la serie tiene una tendencia al menos localmente, y un patrón estacional creciente.

$$S_t = (xt - Ct-L) + (1 - \alpha)(St-1 - Bt-1)$$

$$B_t = (S1 - St-1) + (1 - \beta)-1$$

$$C_t = \gamma (xt - St) + (1 - \gamma)-L$$

$$Ft+K = (St + KBt +Ct+K- L)$$

De donde:

- S_t es el suavizado del patrón constante.
- B_t es el suavizado del patrón tendencial.
- C_t es el suavizado del patrón estacional.
- α, β, γ son constantes de suavizamiento $0 < \alpha, \beta, \gamma < 1$.
- L es la longitud de la estacionalidad.

Este modelo puede llegar a identificar la estacionalidad permitiendo generar una estimación más real dada las tres ponderaciones aplicadas, las cuales ayudan a reducir el error esperado en el modelo. Algunas de las ventajas y desventajas de esta metodología son las siguientes:

Ventajas:

- Requiere muy pocos datos históricos para actualizar un pronóstico de un periodo al siguiente solo son necesarios la demanda del último periodo y el pronóstico de este. Este modelo incorpora en el nuevo pronóstico todas las demandas anteriores.
- El modelo es eficaz, sencillo y fácil de entender.
- Se puede computarizar para diferentes productos, sus partes o sus elementos.
- Sirve en los sectores de manufactura y de servicios.
- Requiere suavizamiento constante en rangos de 0 a 1.
- Determina automáticamente el coeficiente de suavizamiento con base a los errores de los periodos previos.

Desventajas:

- Detecta tendencias, pero las retarda.
- No funciona bien con funciones erráticas, donde existen datos muy aleatorios con baja autocorrelación (Rojas, 2019).

Con base a lo analizado y observando el comportamiento del método Holt-Winters, permitirá de inicio identificar la estacionalidad en la serie de tiempo y el cambio de nivel para los datos de metano, en caso de tener datos aleatorios o ruido blanco se espera que el modelo sea eficiente y no se busque algún otro método a implementar que ajuste mejor a la serie.

2.2.4 Box-Jenkins

Box-Jenkins es una metodología que actualmente es utilizada con menos regularidad puesto que su proceso conlleva a un análisis amplio debido a que parte de un proceso débilmente estacionario de segundo orden, de donde las funciones de media y covarianza no cambien con el tiempo, como a su vez debe cumplir con ciertos comportamientos en sus residuales ya que estos se deben comportar como ruido blanco gaussiano, en otras palabras, los residuales son series de tiempo que tienen una media cero, varianza constante, no correlacionados y se distribuyen normalmente.

En particular, la metodología de Box-Jenkins exige que la serie de tiempo sea débilmente estacionaria, lo cual significa que ni el nivel de la serie (su media) ni su variabilidad depende del tiempo, mientras que la covarianza solamente depende de la separación k entre las variables; de este modo, graficar una cierta cantidad de observaciones contiguas de la serie produciría la misma imagen que si se graficara el mismo número de observaciones contiguas, pero en k periodos antes o después (Guerrero, 2003).

Cabe mencionar que, para todas las metodologías de estimación utilizadas en esta tesis, parten de que la serie llega a tener comportamientos atípicos de la muestra obtenida, el cual se busca mejorar, ya que al obtener la información del RSU BX la máquina de extracción de biogás llega a fallar en algunos momentos causando registros de valor cero en periodos largos de tiempo.

La metodología de Box-Jenkins consta de cuatro etapas de las cuales cada una de ellas permite y depende del proceso anterior para tener una estimación con el menor error posible. Esta metodología permite tener un proceso en orden y en caso de no tener una estimación favorable se pueda regresar y replantear el modelo propuesto. Box-Jenkins es una de las metodologías más elaboradas para la estimación de series de tiempo, por lo cual es recomendable para el análisis en el RSU BX, ya que los datos serán tratados para que los valores atípicos sean los menos visibles en la predicción, ya que esta metodología es sensible

ante valores irregulares. Como dato adicional una metodología Box-Jenkins estima modelos ARIMA y también SARIMA.

Las etapas que se deben seguir en la elaboración de un modelo ARIMA o SARIMA con fines predictivos son las siguientes:

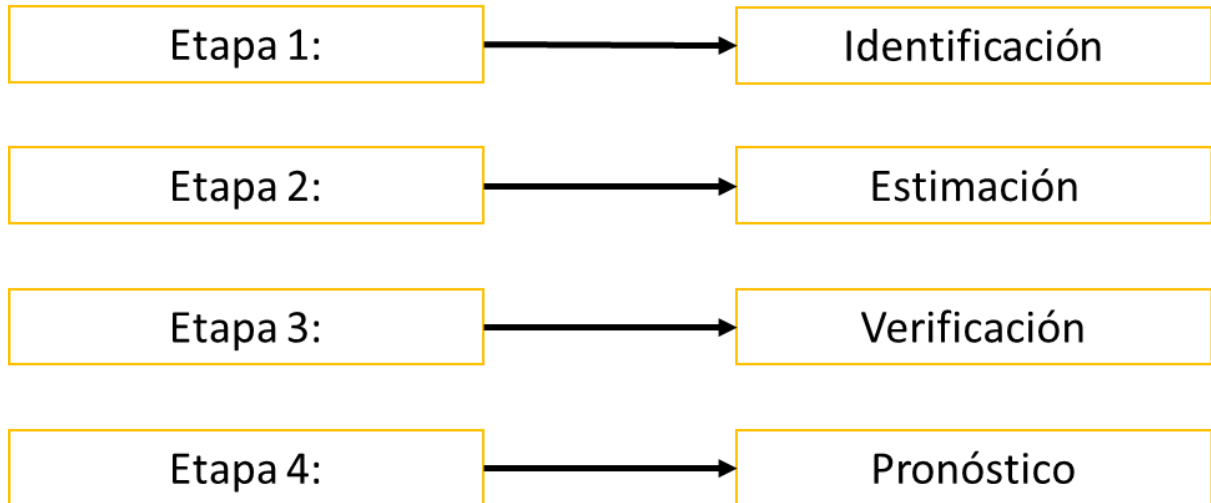


Ilustración 6: Etapas de la modelo metodología.

Etapa 1. Identificación:

Esta fase consiste en detectar el tipo de proceso estocástico que ha generado en los datos. Esto significa encontrar los valores adecuados de la parte autorregresiva, las diferencias a la serie y la parte invertible del modelo ARIMA o SARIMA. Las herramientas fundamentales en la identificación son el correlograma muestral y el correlograma parcial muestral.

Etapa 2. Estimación:

En esta etapa se estiman los coeficientes de los términos autorregresivos y de media móvil incluidos en el modelo, cuyo número de rezagos p y q ya han sido identificados en la etapa anterior.

Etapa 3. Verificación:

En esta etapa se busca evaluar si el modelo estimado se ajusta a los datos en forma razonablemente buena, ya que es posible que exista otro modelo ARIMA que también lo haga. A esta etapa también se le conoce como validación o comprobación en la cual se efectúan algunas pruebas antes de hacer uso del modelo para la predicción. Como último en esta etapa se comprueba que los residuales sean ruido blanco gaussiano.

Etapa 4. Pronóstico:

Para pronosticar un periodo futuro a partir del modelo seleccionado; es decir aquel que es “el mejor” resultante de las etapas anteriores, es importante considerar si la variable original fue diferenciada (Rosales, 2018).

Las gráficas para identificar los parámetros de MA llevan el nombre de función de autocorrelación (ACF), mientras que para AR son función de autocorrelación parcial (PACF). Con base a lo anterior se considera las siguientes características de cada una de las etapas que permitan entender más este proceso.

De la etapa 1 se considera encontrar los coeficientes autorregresivos AR (p), media móvil MA (q) así como las diferencias efectuadas I (d) y de existir los valores S (s) que es el componente estacional de la serie de tiempo.

Los coeficientes de correlación parcial modelo (AR) expresa el grado de relación lineal existente entre dos variables tras eliminar de ambos la dependencia lineal atribuible a terceras variables. La siguiente gráfica muestra la función de autocorrelación parcial del metano.

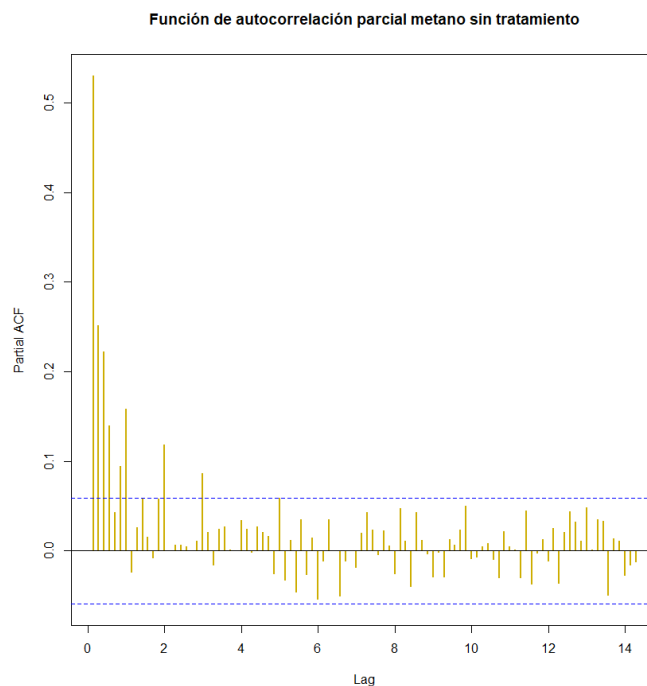


Ilustración 7: Identificación de los parámetros AR.

Para el caso de los modelos MA se tiene como objetivo identificar en la serie la correlación muestral. Las estimaciones MA proporcionan información sobre qué tanta interdependencia existe entre una cierta vecindad de puntos de la serie. La siguiente gráfica muestra la función de autocorrelación del metano.

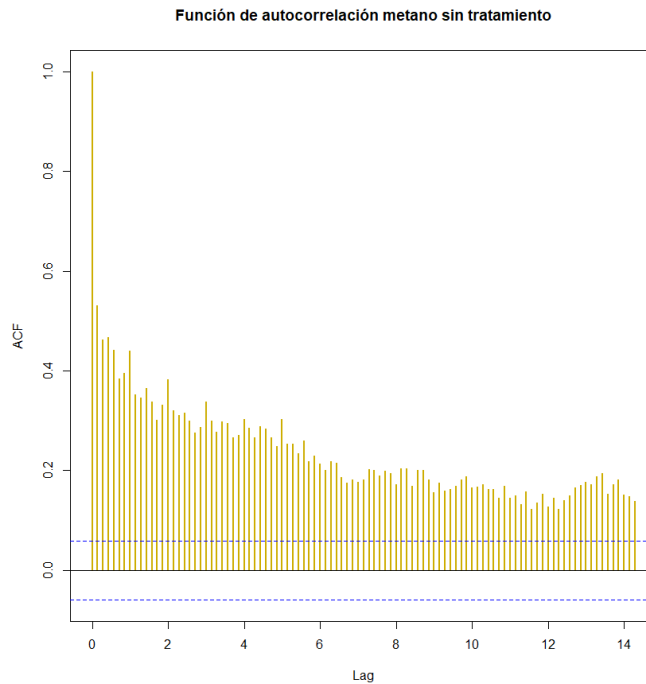


Ilustración 8: Identificación de los parámetros MA.

Los valores de diferencia “I” son aquellas que se realizan a la serie de tiempo la cual permite ajustar la tendencia e irregularidad que pudieran tener en la muestra, considerar que, si hay existencia de estacionalidad en la serie, las diferencias estacionales ayudan a estabilizar los datos y permiten generar un modelo más favorable dado al periodo en el cual se analiza.

Los valores “S” como tal no son valores dados, si no que al observar a la serie y se llegue a identificar comportamientos estacionales, el modelo tendrá que ser orientado a ajustar la estacionalidad y con ellos los parámetros AR, MA y sus diferencias realizadas.

Para la etapa 2 partimos de lo visualizado en las gráficas de identificación de los parámetros MA y AR, con ello identificamos los posibles valores significativos (barras visualizadas por encima de la desviación estándar ± 2), ya que todas aquellas barras que estén dentro de este rango no determinan coeficientes de relación para un modelo ARIMA o SARIMA ya que son ruido.

La ecuación de estimación de modelos ARIMA o SARIMA están representadas de la siguiente forma:

Ecuación ARIMA

$$\varphi(B)\nabla^d y_t = \delta + \theta(B)\varepsilon_t$$

Ecuación SARIMA

$$\Phi(B^s)\phi(B)\nabla^D\nabla^d Y_t = \mu + \theta(B^s)\theta(B)\varepsilon_t$$

De donde:

- ϕ representa los coeficientes AR.
- ∇ son las diferencias realizadas.

- y_t es la estimación del próximo periodo.
- δ es una constante.
- θ representa los coeficientes MA.
- B es el rezago distribuido.
- ε el error en el tiempo t (Ruiz Hernández, Barrios Puente, & Gómez Gómez, 2019).

Para la etapa 3 se debe considerar una validación de los parámetros estimados esto se refiere al análisis a los residuales obtenidos en el modelo, parte del error obtenido debe pasar por un proceso de distintas pruebas que permitan descartar correlaciones, valores atípicos, así como la afirmación de la normalidad y media cero en los residuales. Como dato adicional en la validación existen modelos sobre estimados de los cuales se les recomienda aplicar la parsimonia en las estimaciones realizadas, esto es fundamental para un buen pronóstico.

Para la etapa 4 del cual es visualizar un pronóstico, se requiere que tanto las estimaciones como el ruido blanco tengas las características requeridas, ya que de ser incorrectas el pronóstico dará posiblemente errores significativos o en su defecto la estimación estará fuera de un dato cercano a lo esperado, por lo cual, las etapas anteriores son críticas para su predicción.

A continuación, se mencionan algunas ventajas y desventajas de la metodología Box-Jenkins (modelos ARIMA o SARIMA).

Ventajas:

- Contiene un gran grado de ajuste que proporciona a la mayoría de las series de tiempo.
- Se implementa mediante software gratuito.
- Reduce los componentes principales de las series de tiempo.
- A diferencia de la modelación clásica en donde se ajusta una serie a un modelo matemático ya establecido, los modelos ARIMA se ajustan a una serie en particular.
- Los conceptos que se utilizan para la modelación ARIMA, se derivan de sólidas teorías de la probabilidad clásica y de la estadística matemática.
- Los modelos ARIMA son una familia de modelos, no simplemente un único modelo, Box y Jenkins desarrollaron una completa estrategia que sirve de guía para escoger un apropiado modelo dentro de esta gama de modelos.
- Un apropiado modelo ARIMA produce óptimas predicciones.
- La metodología Box-Jenkins permite generar modelos SARIMA, los cuales atacan principalmente la estacionalidad de una serie de tiempo.

Desventajas:

- Al ser una metodología amplia implica mayor tiempo dado los pasos y los detalles al observar las series de tiempo. Lo anterior mencionado complica el tiempo óptimo para una toma de decisiones si no se tuviera una implementación computacional adecuado.

- En ocasiones el analista interpreta valores correlacionados sin mucho peso para un pronóstico favorable.
- Recién iniciado en la teoría y la ejecución de la metodología Box-Jenkins, la correcta elección de un modelo adecuado en los primeros intentos no siempre es la óptima (Cortés Patiño, 2011).

2.2.5 Red neuronal artificial

Una de las metodologías más utilizadas recientemente y que a su vez es una de las herramientas de estimación más poderosas son las redes neuronales artificiales, de las cuales, se describirán algunas de sus características a continuación.

Como sabemos, actualmente existen diferentes tipos de métodos que facilitan la manipulación de los datos, tanto para poder entender o interpretar ciertos comportamientos como a su vez predecir estos mismos, entre estas metodologías se encuentra *Deep learning*, el cual busca imitar el funcionamiento del cerebro humano.

El cerebro humano tiene una estructura de comunicación maravillosa, ya que este puede estar conectado entre millones de neuronas que se envían información mediante impulsos electromagnéticos. Cada impulso contiene información que permite interpretar, recordar, imaginar entre otras muchas acciones, todo esto para el funcionamiento correcto del cuerpo humano. Cabe mencionar que es importante describir algunas partes de una red neuronal previo a la interpretación en una computadora.

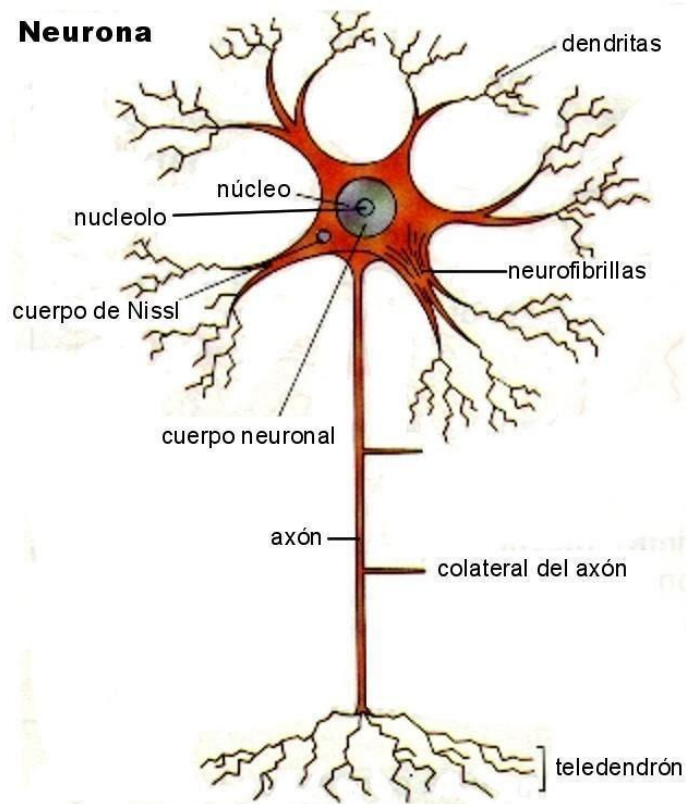


Ilustración 9: Estructura neuronal (Crispi, 2019).

Como se observa en la ilustración 9: Estructura neuronal, podemos ver la estructura de una red neuronal del ser humano y como componentes principales para la descripción de una red neuronal orientada a una estimación serían el axón y las dendritas, las cuales se conectan y transmiten la información al núcleo de la neurona.

Una neurona por sí sola no tiene ninguna estructura capaz de poder realizar lo que todas juntas pueden llegar a obtener, el proceso de intercambiar información entre neuronas se le conoce como sinapsis. La incógnita de convertir una red neuronal como herramienta de análisis ante una computadora se ha convertido en un hecho real en la actualidad ya que se ha mejorado en los últimos años esta técnica de estimación. Ahora bien, para poder observar una red neuronal e interpretar o imitar al cerebro humano se tiene la estructura mostrada en la ilustración 10: Red neuronal artificial para estimación.

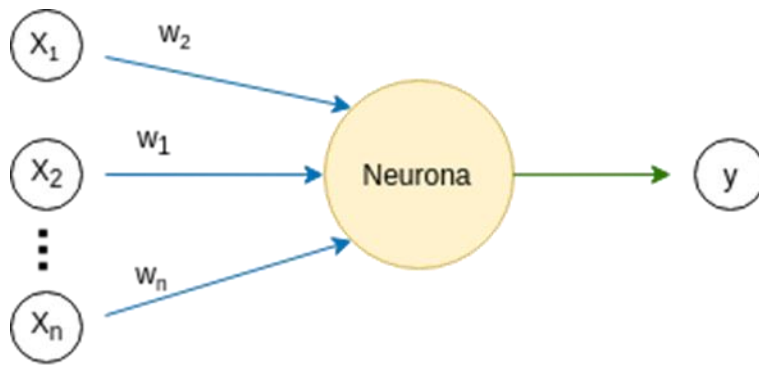


Ilustración 10: Red neuronal artificial para estimación.

Esta sería una representación gráfica de una neurona artificial. Las flechas representan dendritas, y las X como información que le pasan otras neuronas, entonces nuestras flechas serían la sinapsis, en este caso la señal que le pasamos es básicamente un valor de entrada. Entonces, a partir de estos valores, generan un valor de salida. Estas entradas pueden ser cualquier cosa, pero por ahora las llamaremos, variables independientes y a la salida la llamaremos variable dependiente y a las sinapsis las llamaremos pesos (Crispi, 2019).

El objetivo de una red neuronal artificial es permitir un ajuste a los pesos que son asignados inicialmente por el analista y estos a su vez ser mejorados durante el entrenamiento. Para los pasos que se buscan implementar se dividen en dos puntos, ya que estos en primera instancia son la suma del producto de cada variable versus el peso asignado. Una vez realizado este punto se busca el método o función de acción que permita tener una salida esperada.

Para el caso de estimación de metano obtenido en el RSU BX se buscará una función de activación que permita describir el comportamiento de la serie de tiempo y con ello obtener los próximos días del metano que favorezcan en las decisiones de la máquina de extracción en caso de que llegara a fallar.

La función de activación dada la suma ponderada de pesos versus las variables de entrada de una red neuronal artificial se muestra en la siguiente ecuación.

$$\varphi\left(\sum_{i=0}^m \omega_i * x_i\right)$$

De donde:

- φ es la función de activación.
- ω son los pesos de cada relación hacia un nodo.
- X son las variables de entrada.

La función tangente hiperbólica transforma los valores introducidos a una escala (-1,1), donde los valores altos tienen de manera asintótica a 1 y los valores muy bajos tienden de manera asintótica a -1.

$$f(x) = \frac{2}{1 + e^{-2x}} - 1$$

Función tangente hiperbólica

Características de la función tangente hiperbólica:

- Lenta convergencia.
- Centrada en 0.
- Está acotada entre -1 y 1.
- Se utiliza para decidir entre una opción y la contraria.
- Buen desempeño en redes recurrentes (Calvo, 2018).

Otro punto importante de este método es que es rápido, por la manera en que funciona se tiende a pensar que no, porque se ajusta por cada dato, pero en realidad como no tiene que cargar todos los datos en memoria, para ajustar los pesos y esperar hasta recorrer todas las filas de nuestros datos, va uno por uno, por lo que es un algoritmo más liviano y rápido (Crispi, 2019).

A continuación, se describen algunas ventajas y desventajas de la metodología de una red neuronal artificial.

Ventajas:

- Aprendizaje: Las redes neuronales artificiales (RNA) tienen la habilidad de aprender mediante una etapa que se llama etapa de aprendizaje. Esta consiste en proporcionar a la RNA datos como entrada a su vez que se le indica cuál es la salida (respuesta) esperada.
- Auto organización: Una RNA crea su propia representación de la información en su interior, descartando al usuario de esto.
- Tolerancia a fallos: Debido a que una RNA almacena la información de forma redundante, ésta puede seguir respondiendo de manera aceptable aun si se daña parcialmente.
- Flexibilidad: Una RNA puede manejar cambios no importantes en la información de entrada, como señales con ruido u otros cambios en la entrada (ej. si la información de entrada es la imagen de un objeto, la respuesta correspondiente no sufre cambios si la imagen cambia un poco su brillo o el objeto cambia ligeramente)
- Tiempo real: La estructura de una RNA es paralela, por lo cual, si esto es implementado con computadoras o en dispositivos electrónicos especiales, se pueden obtener respuestas en tiempo real.

- Son excelentes como clasificadores/reconocedores de patrones y pueden ser usadas donde las técnicas tradicionales no funcionan.
- Las redes neuronales artificiales pueden manejar excepciones y entradas de datos anormales, muy importante para sistemas que manejan un amplio rango de datos (sistemas de radar y sonar).

Desventajas:

- Complejidad de aprendizaje para grandes tareas, cuantas más cosas se necesiten que aprenda una red, más complicado será enseñarle.
- Tiempo de aprendizaje elevado. Esto depende de dos factores: primero si se incrementa la cantidad de patrones a identificar o clasificar y segundo si se requiere mayor flexibilidad o capacidad de adaptación de la red neuronal para reconocer patrones que sean sumamente parecidos, se deberá invertir más tiempo en lograr que la red converja a valores de pesos que representan lo que se quiera enseñar.
- No permite la interpretación de lo que se ha aprendido, la red por sí sola proporciona una salida, un número, que no puede ser interpretado por ella misma, sino que se requiere de la intervención del programador y de la aplicación en sí para encontrarle un significado a la salida proporcionada.
- Elevada cantidad de datos para el entrenamiento, cuanto más flexible se requiere que sea la red neuronal, más información tendrá que enseñarle para que realice de forma adecuada la identificación.
- Otros problemas con las redes neuronales son la falta de reglas definitorias que ayuden a realizar una red para un problema dado (Wikidot, 2015).

Capítulo 3 Recolección de datos

3.1 Obtención de los datos

Para explicar lo siguiente debemos de poner en contexto al lector. El tiradero Neza 1 fue uno de los primeros tiraderos clandestinos, a cielo abierto, grandes y famosos en toda América Latina dentro del Bordo de Xochiaca, este último se construyó en 1945, e históricamente recibió basura: de 1986 a 1991, por lo que, su superficie creció a 80 hectáreas y sumó 600 mil toneladas más de basura durante sus últimos años (López Piña & Quintero Soto, 2009).

Llegó a recibir 14 mil toneladas diarias, hasta su cierre en el 2006 contando con 420 hectáreas y almacenando más de 30 millones de toneladas. Como se ha mencionado, el objetivo de la tesis es predecir el comportamiento de metano puesto que este, está clasificado como un gas de efecto invernadero, ya que al estar en descomposición la basura confinada en los RSU la producción de este gas es favorable para análisis.

Hablando estructuralmente, el sistema de extracción forzada de biogás fue un proyecto de reconstrucción del tiradero Bordo Xochiaca. La estructura de conexión de 175 pozos en el RSU BX permite canalizar la composición del biogás y así mismo al estar conectadas entre sí, son enviadas hacia un catalizador la ilustración 2: Mapa de la red extracción forzada de biogás muestra esta conexión de pozos.

El programa de monitoreo tiene como objetivo llevar las mediciones del flujo, composición y presión del biogás, en toda la red de extracción; por lo que permite mantener las condiciones de presión, flujo, contenido de oxígeno y porcentaje de metano en la mezcla para la generación de biogás con la finalidad de poder lograr la combustión en forma eficiente, en las condiciones operacionales que requiere el quemador (Ahumada de Toledo, 2014).

Estos dos extractores llevan todo el biogás a un quemador principal que es de tipo cerrado, para que las reducciones de emisiones de gases efecto invernadero puedan ser monitoreadas. El poder medir este biogás representa una medida de seguridad como a su vez el proceso de control en el RSU BX.

Por medio de la iniciativa académica se consiguió hacer contacto con los responsables del lugar, expertos en el sistema. Una vez contactados y en el sitio se nos mencionó que el quemador principal es de donde se obtienen las constantes muestras, y que los periodos de estas muestras son cada cierto tiempo, siendo la principal fuente de obtención de la información, la información no es más que la composición del biogás, y está estructurada en hojas de datos con formatos Excel.

Para facilitar la manipulación de datos se utilizaron programas que facilitaran el ordenamiento de metano por lo que se utilizaron macros que son archivos de ejecución.

Objetivo de las macros (creación de los dos módulos):

- Realizar la unión de información de cada una de las hojas del archivo con formato Excel, esto consiste en traer de forma iterativa a través de un rango de celdas los días transcurridos durante la semana n, todo esto por el total de hojas acumuladas por el extractor de biogás y enviadas a una nueva hoja con los datos acumulados.
- De la hoja creada se extraen los datos y se almacenan, como criterio de esto, se consideran las horas con menor cantidad de valores iguales a 0 o "N/A" (días en que la máquina no funcionó y se almacena el dato). Para completar el objetivo y optimizar el proceso de las macros en los módulos diseñados, se consideró el número de las columnas con información requerida para el análisis de metano, esto quiere decir, que se descartaron columnas que no eran de interés. El proceso iterativo fue encontrar la fecha y contar la posición en la que se encontraba un valor mayor a 0 o distinto de "N/A" está fue la estructura lógica para tener una muestra favorable del biogás.

En la *Tabla 3: Estructura de datos obtenidos por las macros* se muestra la estructura general del formato final de las ejecuciones de las macros.

Dia	Hora	CH4	CO2	O2	Bal	Temperatura	Flujo
05/11/2017	12:20	42.2%	23.0%	1.30%	33.5%	1718° F	469 m3/Hr

Tabla 3: Estructura de datos obtenidos por las macros.

3.2 Características de los datos

Del proceso iterativo se extrae de la base creada el dato de metano (CH₄), que se le atribuye un comportamiento de serie de tiempo. Como característica de la muestra se observó que la composición del biogás contiene impurezas, como se muestra en la variable Bal de la tabla anterior, considerar que la impureza del biogás depende del porcentaje de esta variable ya que si se encuentra superior al 5% causa un deterioro en la composición del biogás.

A continuación, se describe el comportamiento de la serie de tiempo la cual es analizada en dos muestras, estas son metano sin tratamiento de datos atípicos y con tratamiento de datos atípicos. Hay que considerar que los términos de metano con y sin tratamiento son conceptos que serán mencionados frecuentemente en el texto ya que son la aplicación de la transformación de la serie. En el subtema final de este capítulo se describe la metodología aplicada para el tratamiento de la serie de tiempo.

3.2.1 Metano sin tratamiento de datos atípicos

Aquí se describen los datos estadísticos de la serie de tiempo, se considera que la correlación de metano durante el tiempo observado es significativa para estimar a corto o mediano plazo.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
1,096	36.70%	8.6%	18.2%	30.5%	36.7%	41.30%	58.60%

Tabla 4: Comportamiento estadístico de metano sin tratamiento.

De la tabla se observa que la serie en un intervalo de tiempo de tres años tiene un promedio del 36.7% de composición con base al total del biogás extraído, esta extracción y comportamiento del metano contiene también un mínimo del 18.2% y un máximo del 58.6%, por lo que se tiene una muestra de 1,096 datos el cual equivale a un dato tomado diariamente del 1 de enero del 2015 al 31 de diciembre del 2017.

Las gráficas observadas permiten visualizar y apoyar la decisión de ajuste por el analista, también si se realizan pruebas que verifiquen lo observado ya que en las gráficas se tendría una justificación para las decisiones de ajuste. Para describir metano sin tratamiento se grafica un histograma, puesto que este permite observar la acumulación, dispersión y la forma de distribución de nuestra variable de interés.

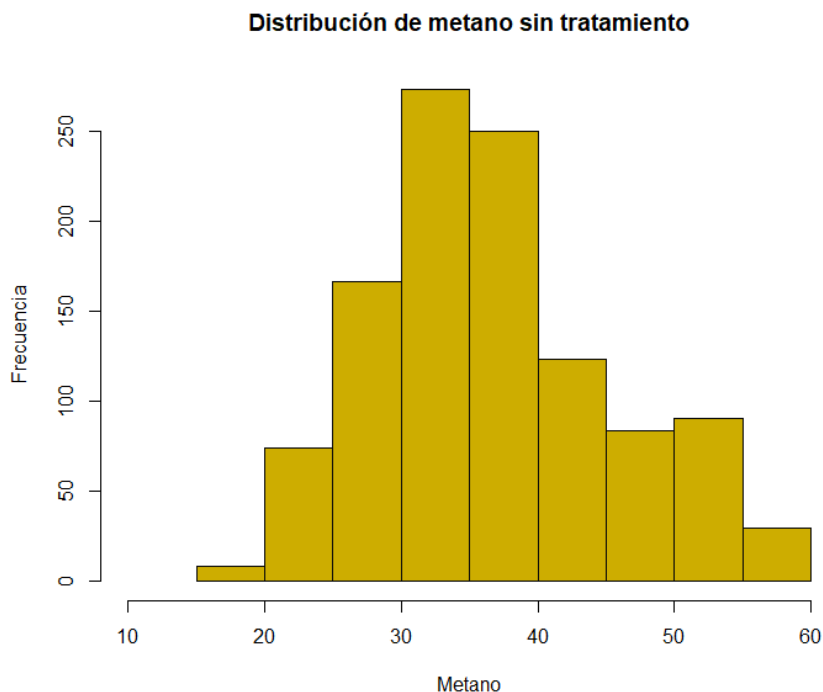


Ilustración 11: Distribución de metano sin tratamiento.

A continuación, podemos observar el gráfico de caja, el cual nos permite en la mayoría de las ocasiones identificar datos irregulares y/o atípicos, a este gráfico se le atribuye el nombre de gráfico caja-bigote ya que se espera que la mayor concentración de datos esté dentro de la caja y los datos atípicos en el desplazamiento de las líneas tanto a la derecha como a la izquierda, de ahí surge la comparación de un bigote.

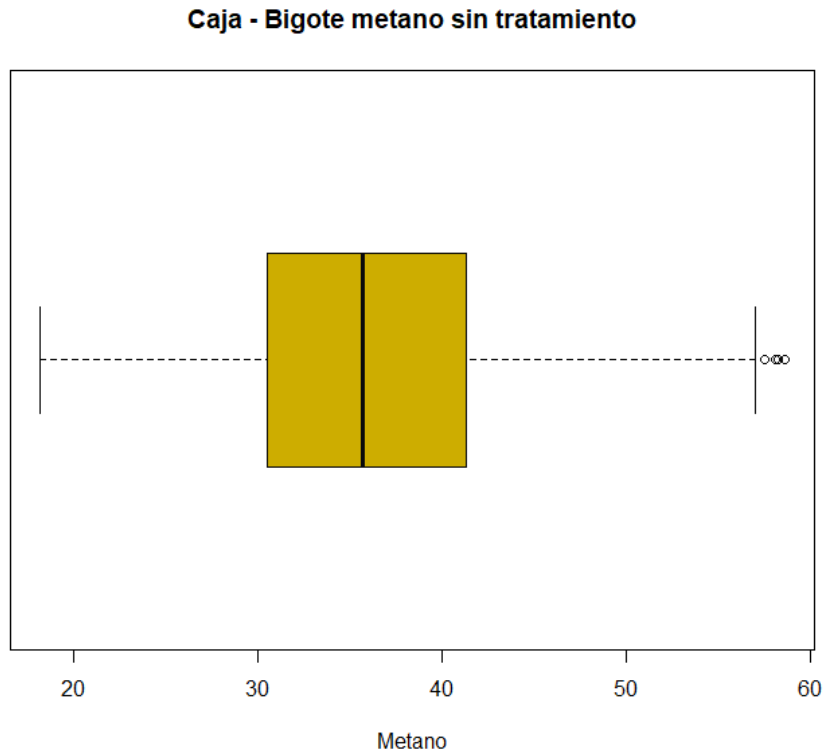


Ilustración 12: Dispersión de metano sin tratamiento.

Una vez observada la dispersión e irregularidad de los datos en los gráficos anteriores, visualizamos a la serie en una gráfica sobre el tiempo, en la cual se observará su comportamiento durante los años analizados, esto permitirá ver los componentes de la serie de tiempo y facilitar la identificación para las metodologías a implementar.

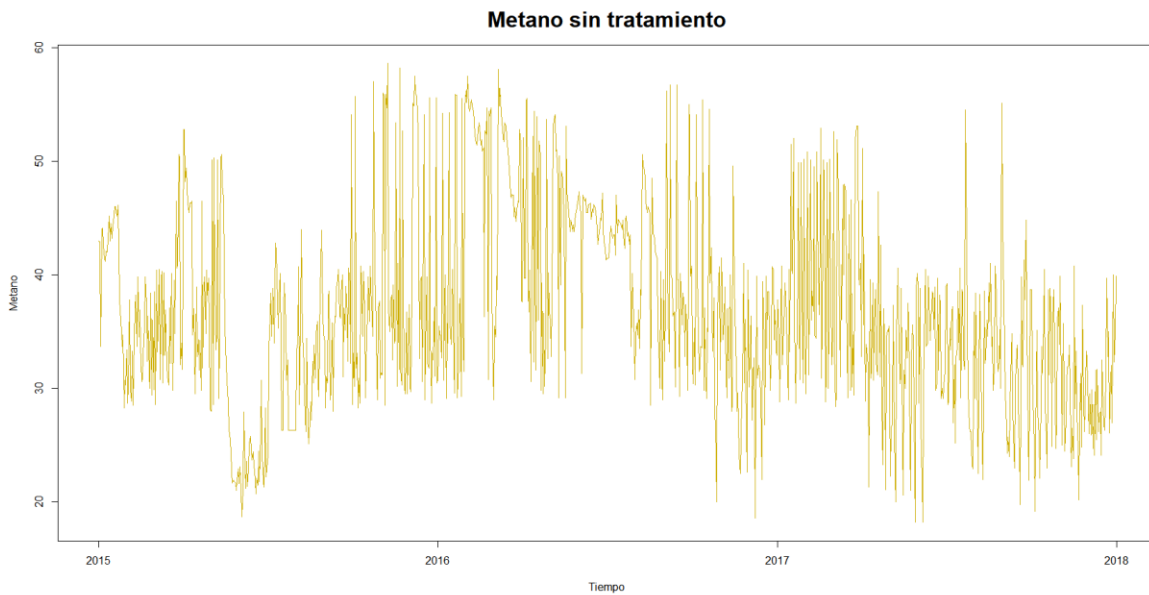


Ilustración 13: Serie de tiempo de metano sin tratamiento.

Es importante mencionar que para el comportamiento de correlación las gráficas de periodograma integrado, estacionalidad, tendencia, irregularidad, función de correlación y correlación parcial de metano sin tratamiento de valores atípicos pueden llegar a ayudar a mejorar la estimación de los modelos.

La visualización de los componentes de una serie de tiempo es importante de identificar en la mayoría de los métodos implementados, en la siguiente gráfica se puede observar la estacionalidad, tendencia e irregularidad de la serie.

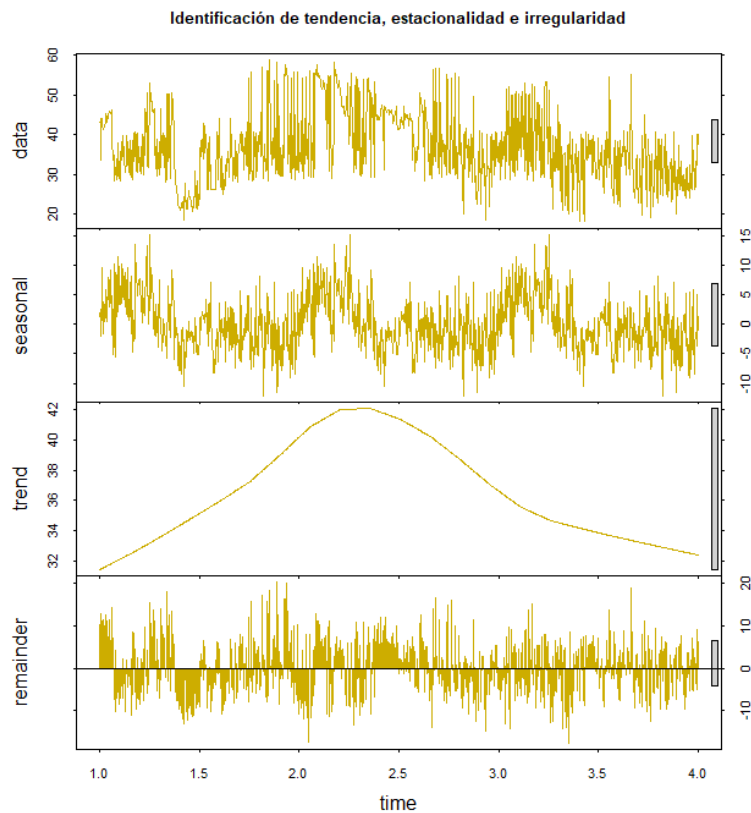


Ilustración 14: Componentes de metano sin tratamiento.

Como se puede observar en el periodo de los 3 años la tendencia de la serie no es constante, también se observa un comportamiento estacional durante los inicios de cada año, otro componente observado es la irregularidad de los datos ya que este tiene un comportamiento de picos que salen de la desviación, esto posiblemente son los atípicos de los días en que la máquina no funciona.

En la siguiente gráfica analizaremos el análisis espectral mediante el periodograma integrado puesto que este ayuda a saber si la serie no se comporta como ruido blanco, ya que si fuera así los datos estarían dentro de las líneas diagonales y por lo tanto no podría ser predecible.

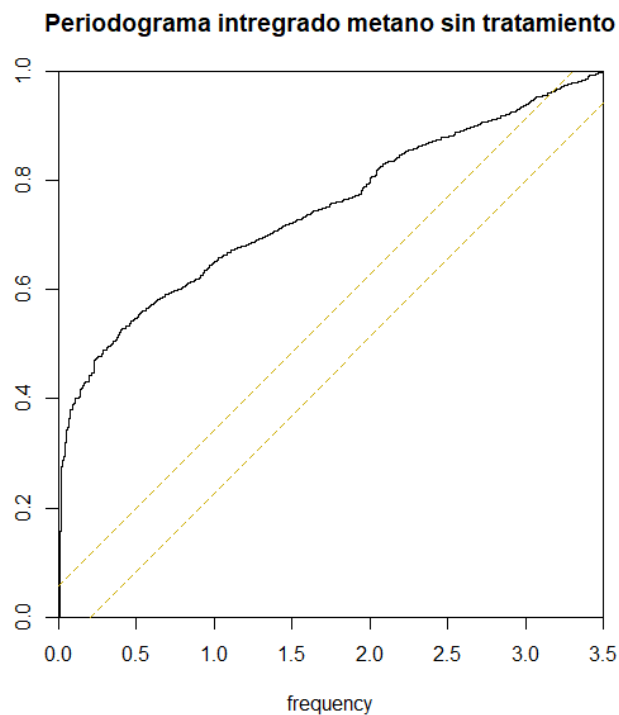


Ilustración 15: Periodograma de metano sin tratamiento.

La gráfica nos indica que es una serie posible para predecir y no se comporta como ruido blanco.

Es importante aclarar que la muestra tomada del RSU BX en cierto punto es irregular, por lo que se busca mejorar esa inconsistencia en los datos para un mejor ajuste de los modelos. Ahora visualizamos la función de correlación y correlación parcial de metano sin tratamiento.

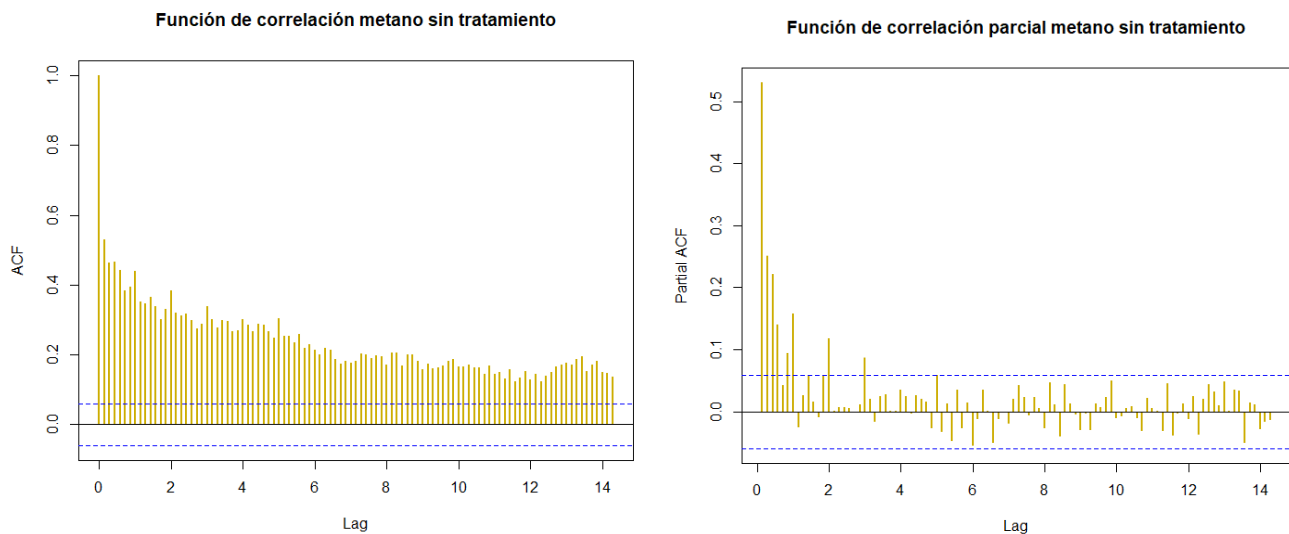


Ilustración 16: Funciones de correlación de metano sin tratamiento.

Se observa que la función de autocorrelación parcial decrece de forma fragmentada, se ha mencionado en el capítulo 2 que los gráficos de correlación muestral facilitan la predicción para los modelos ARIMA o SARIMA.

3.2.2 Metano con tratamiento de datos atípicos

Como se observó anteriormente el metano sin tratamiento de datos atípicos, sería quizá poco predecible si la usamos sin ajustar los componentes no estables observados, por ello una de las propuestas de la tesis para mejorar la toma de decisiones del RSU BX es reducir la incertidumbre de los datos de metano, por lo que se genera un ajuste de suavizamiento a la serie. En primer aspecto visualizamos la descripción estadística de la muestra suavizada, posteriormente los gráficos con el tratamiento aplicado para que al final se describa el tratamiento realizado.

Se generan las mismas gráficas para comparativo de los datos y se espera que con este tratamiento la incertidumbre de los componentes y los valores atípicos de la serie de tiempo reduzcan. La descripción estadística es la siguiente.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
1,096	36.70%	6.7%	20.5%	32.3%	36.7%	41.80%	55.30%

Tabla 5: Comportamiento estadístico de metano con tratamiento.

Como se observa, los datos siguen siendo 1,096, tenemos una diferencia en los valores mínimos y máximos, así como en la desviación de estos. Se podría indicar que la serie se

escalono de cierta forma con el tratamiento aplicado. El histograma del metano con tratamiento se muestra a continuación, así como la gráfica de caja.

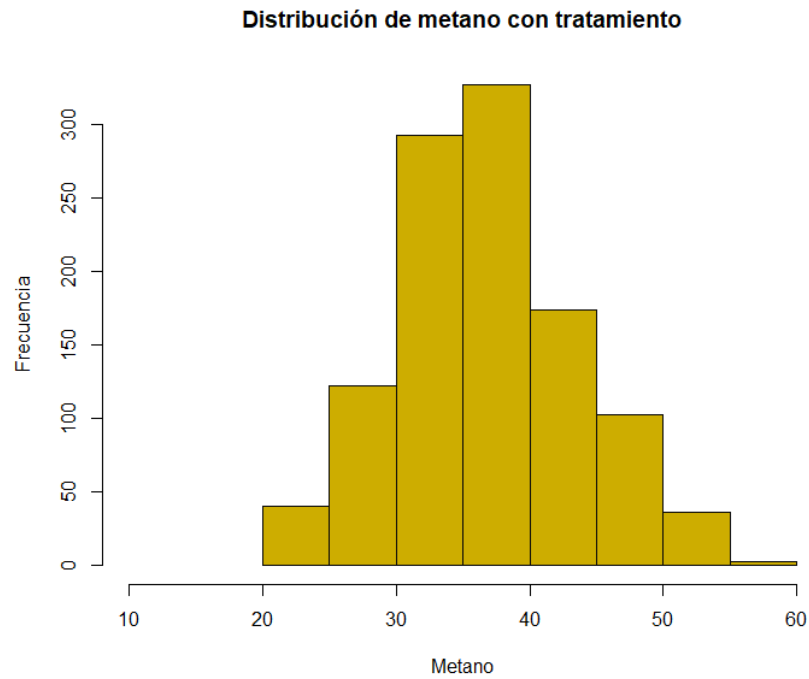


Ilustración 17: Distribución de metano con tratamiento.

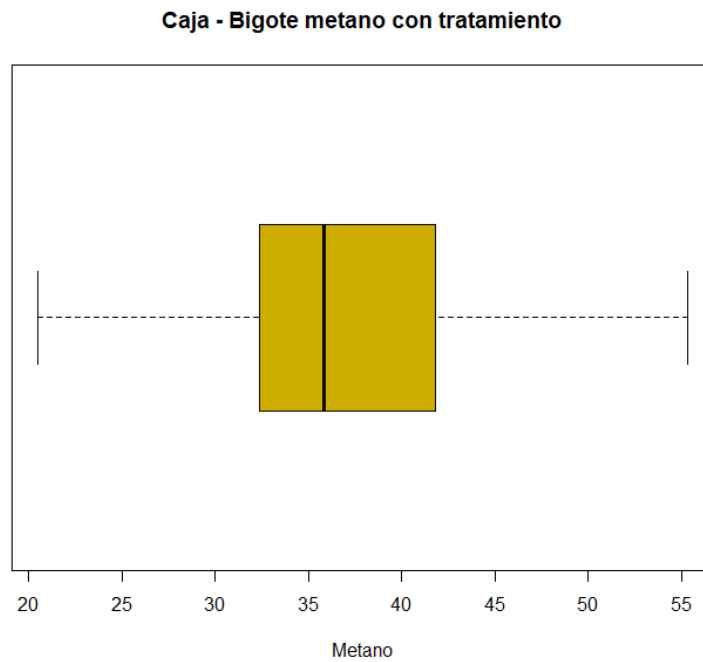


Ilustración 18: Dispersión de metano con tratamiento.

Una vez observada la dispersión y distribución de los datos visualizamos a la serie en una gráfica sobre el tiempo, se espera tener menor irregularidad como a la que se observaba en metano sin tratamiento de datos atípicos.

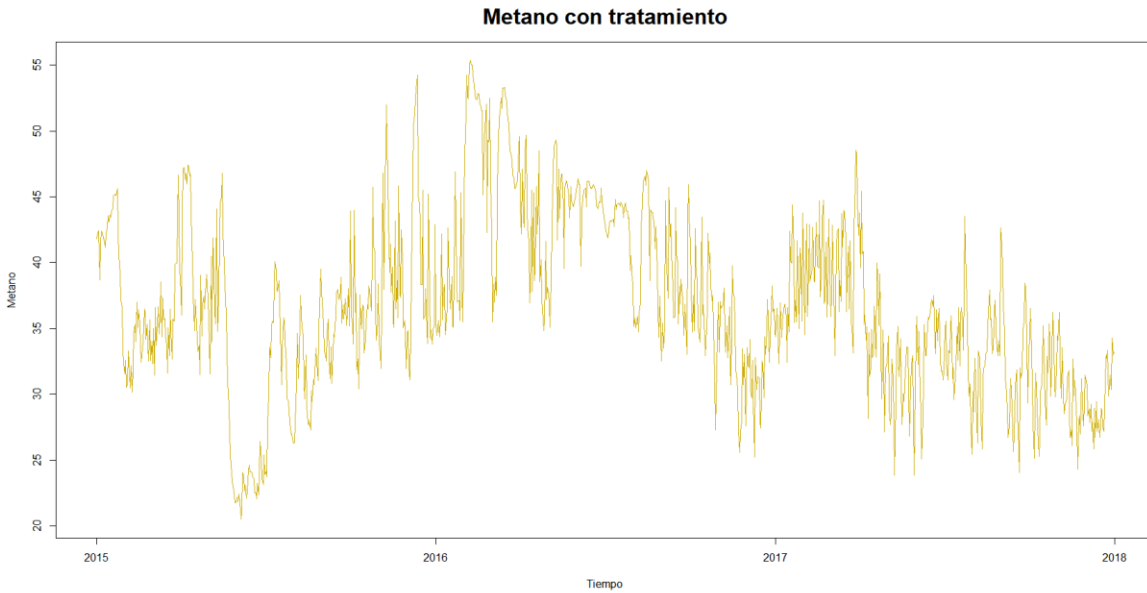


Ilustración 19: Serie de tiempo de metano con tratamiento.

Como siguiente gráfico veremos los componentes de la serie para diferenciar a la serie metano sin tratamiento.

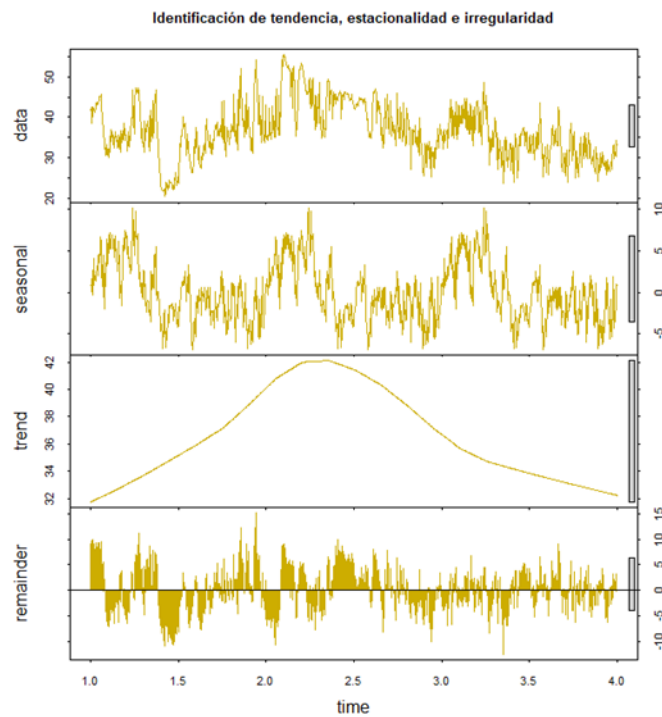


Ilustración 20: Componentes de metano con tratamiento.

Como se observa la estacionalidad e irregularidad de la serie fue mejorada, mientras que la tendencia aún sigue apareciendo durante los años observados, se espera que al aplicar diferencias a la serie esta se minimice, pero todo esto es parte del capítulo 4 en donde se describen las metodologías elegidas. Ahora se observa el análisis espectral con el periodograma de metano con tratamiento.

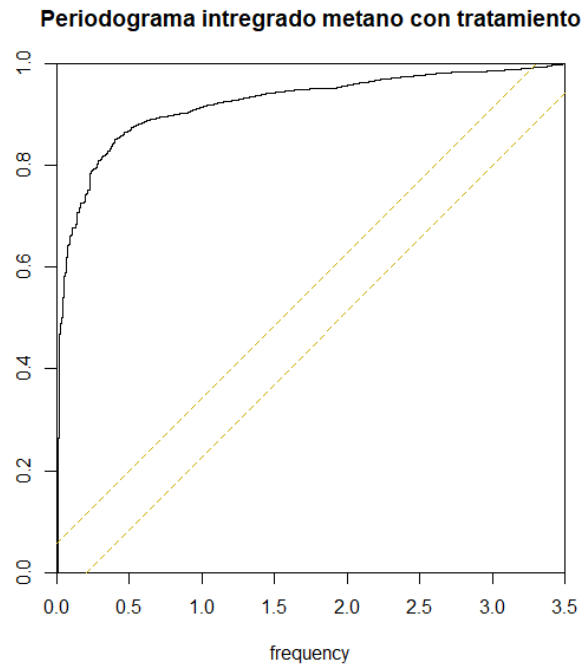


Ilustración 21: Periodograma de metano con tratamiento.

Por último, visualizamos las funciones de correlación de metano con tratamiento, a estas gráficas las segmentamos de la misma forma para poder visualizar y comparar la correlación muestral directa e indirecta de la serie.

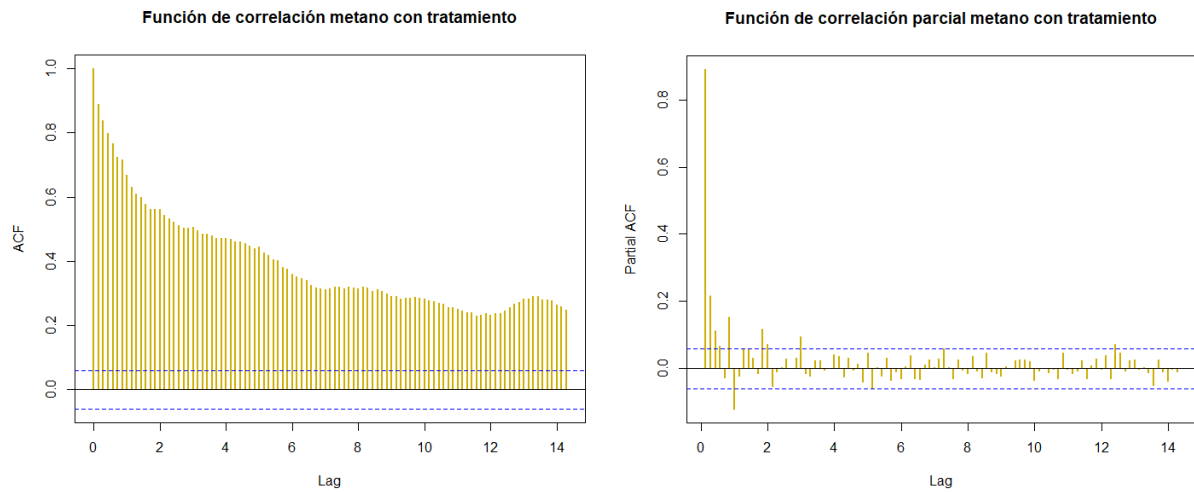


Ilustración 22: Funciones de correlación de metano con tratamiento.

Cómo comparativo de metano con y sin tratamiento de datos atípicos, se tienen las siguientes gráficas, las cuales permiten ver comparativamente el tratamiento aplicado.

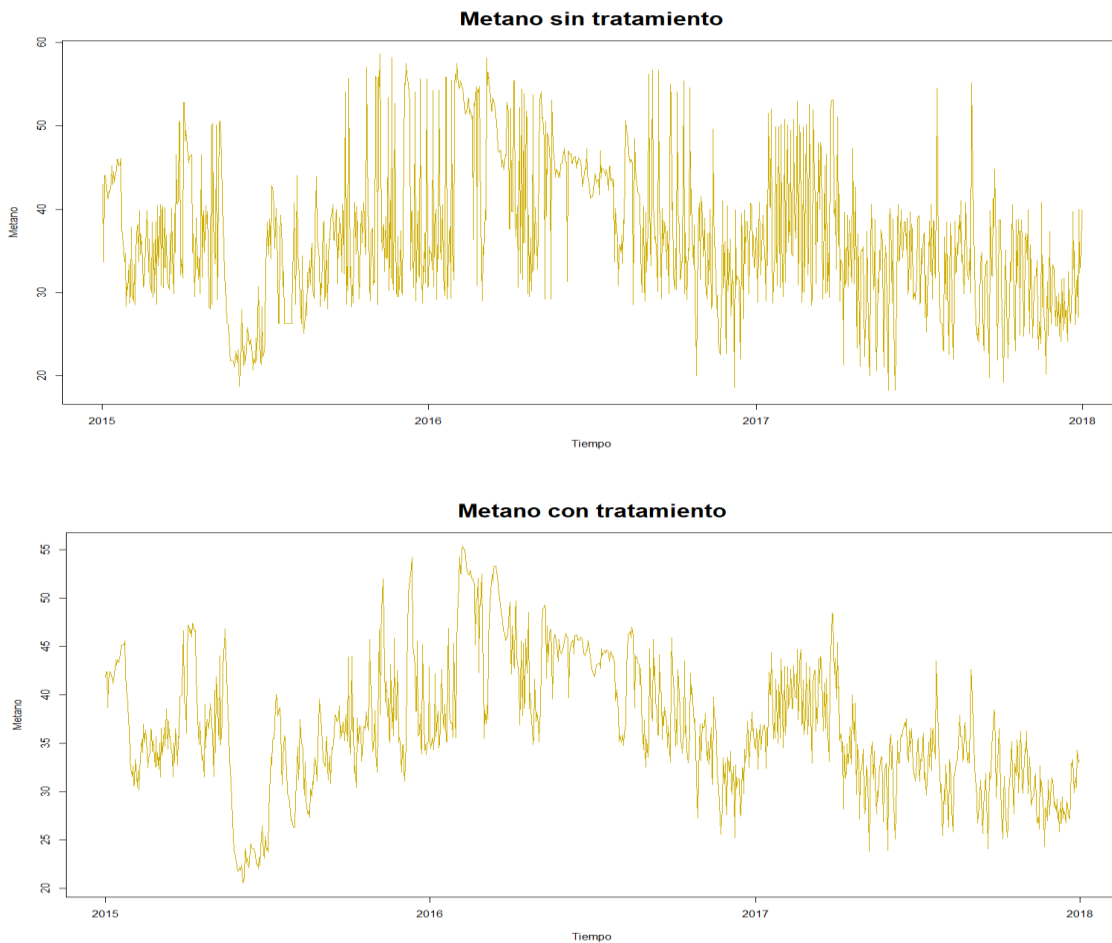


Ilustración 23: Comparativo de series.

3.3 Tratamiento del metano

La muestra obtenida en el RSU BX fue evaluada por un posible comportamiento de valores atípicos. Esto es más favorable cuando se generan validaciones de valores extremos que superan tres veces la desviación estándar. Como herramienta de apoyo se utiliza la prueba de Grubbs de donde inicialmente se espera que no se comporte de manera normal la serie, por lo que facilita la decisión de generar algún proceso que trate a los datos atípicos y a su vez suavice el componente irregular, esto permite considerar que al generar algún tratamiento los modelos que estimemos tengan mejor predicción al tiempo n.

Se debe verificar si los datos muestran una distribución normal antes de aplicar la prueba de Grubbs. La prueba de Grubbs siempre verifica el valor que muestra la mayor desviación absoluta de la media. Si se identificó y eliminó un valor atípico, la prueba no debe repetirse sin adaptar el valor crítico. La aplicación de la prueba es bastante simple y directa: se busca el máximo de las diferencias absolutas entre los valores x y la media. El resultado se divide por la desviación estándar de la muestra. Si la estadística de prueba resultante es mayor que el valor crítico, el valor correspondiente puede considerarse un valor atípico.

Para considerar la toma de decisiones del RSU BX aplicaremos un promedio móvil ponderado, esto sabiendo que la prueba de Grubbs salió con valores atípicos una vez aplicada.

Para apoyar y optimizar los valores de los pesos, Excel maneja una función de nombre Solver que permite mediante condiciones de valores iniciales poder encontrar objetivos mínimos o máximos, que en este caso será la producción de metano. Lo anterior se utilizó para poder encontrar siete pesos elegidos representando los días de la semana.

Coeficientes	Valor de la ponderación
β_1	0.03
β_2	0.06
β_3	0.08
β_4	0.11
β_5	0.13
β_6	0.17
β_7	0.42

Tabla 6: Pesos del suavizamiento.

Capítulo 4 Aplicación de los modelos de pronóstico a los datos

4.1 Combinaciones posibles

Debido a que la información viene distribuida de manera diaria es posible tomar muestras con o sin tratamiento de datos atípicos y frecuencias periódicas. Se debe mencionar que no se toman los promedios de las frecuencias sino el día transcurrido del periodo observado, considerando esto, se muestran las posibles combinaciones de series de metano.

- Día de la semana:
Grupo = {Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo}
Total = 7
- Tratamiento:
Grupo = {0,1}
Total = 2
- Frecuencia
Grupo = {1, 7, 14, 30}
Total = 4

Posibles Combinaciones: $7 \times 2 \times 4 = 56$

4.2 Consideración de los días de la semana

Teniendo presente que el metano se encuentra en la composición del biogás e históricamente tiene una media del 36.7% en la muestra diaria, se considera como elección inicial tomar un cierto universo de variables con el menor número de datos reemplazados en el tratamiento, lo que nos da como resultado los días de la semana viernes, sábado y domingo como a su vez la muestra diaria.

Por lo que tenemos un total de 24 combinaciones, tan solo por tomar estos tres días. A estas muestras las simulamos para descartar 13 series, de las 11 sobrantes se les aplicaron validaciones para descartar ruido blanco, como a su vez, se graficaron los componentes de cada una de ellas, permitiendo quedarnos con solo 5 muestras aceptables.

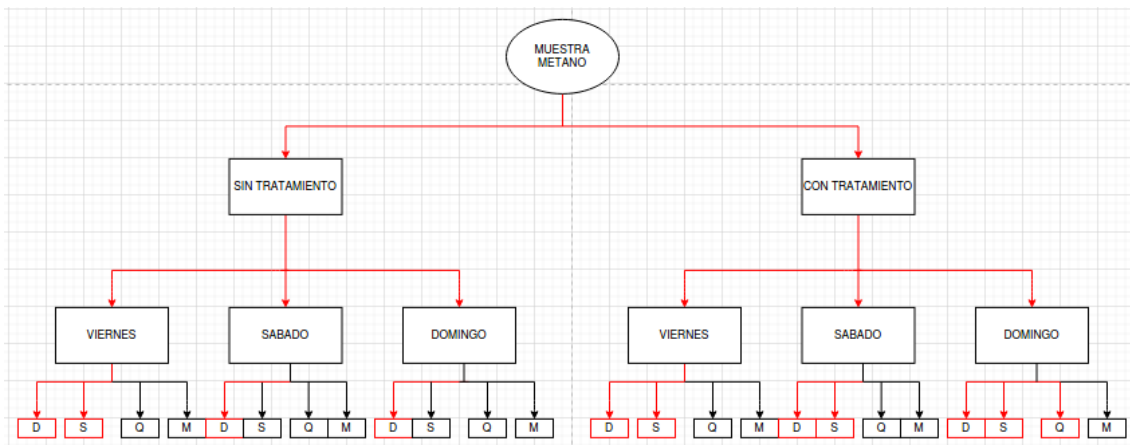


Ilustración 24: Diagrama de elección de muestras.

4.3 Series analizadas

Las cinco combinaciones seleccionadas tienen los siguientes comportamientos y son comparadas unas con otras.

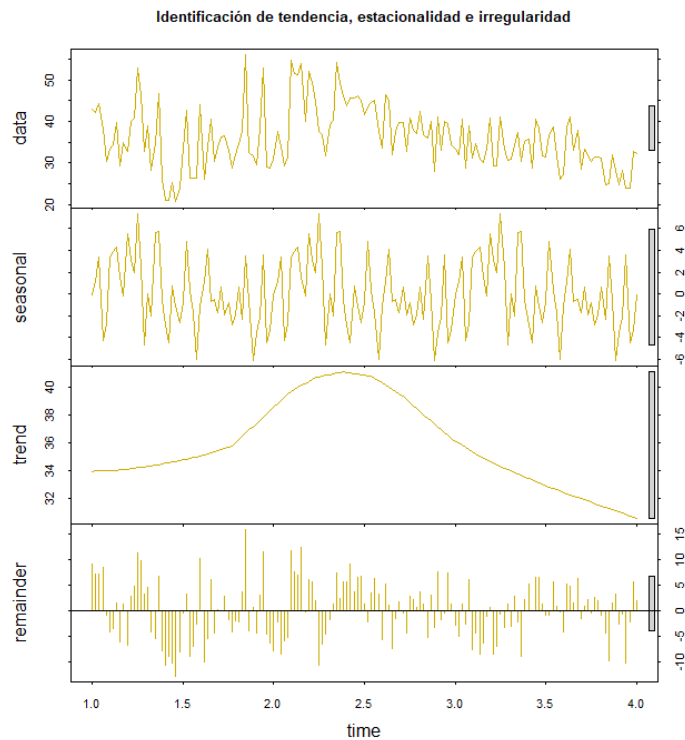


Ilustración 25: Componentes de metano viernes sin tratamiento.

Periodograma integrado metano viernes sin tratamiento

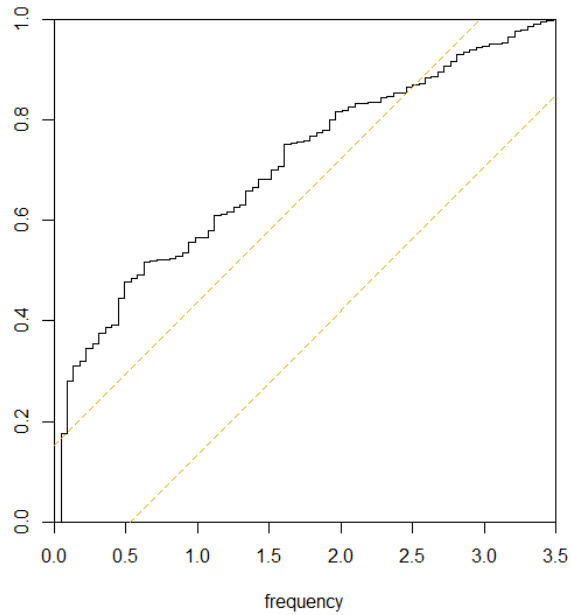


Ilustración 26: Comportamiento de ruido blanco de metano viernes sin tratamiento.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
157	35.90%	7.50%	20.70%	30.90%	35.90%	40.50%	55.90%

Tabla 7: Distribución de metano viernes sin tratamiento.

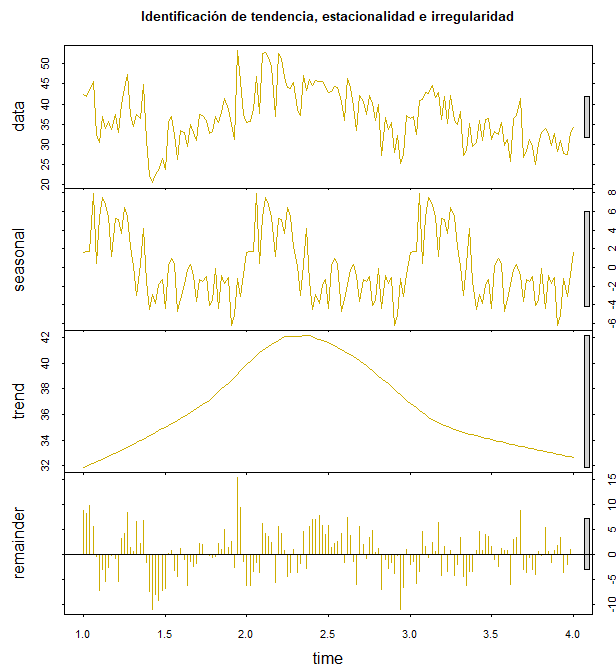


Ilustración 27: Componentes de metano viernes con tratamiento.

Periodograma integrado metano viernes con tratamiento

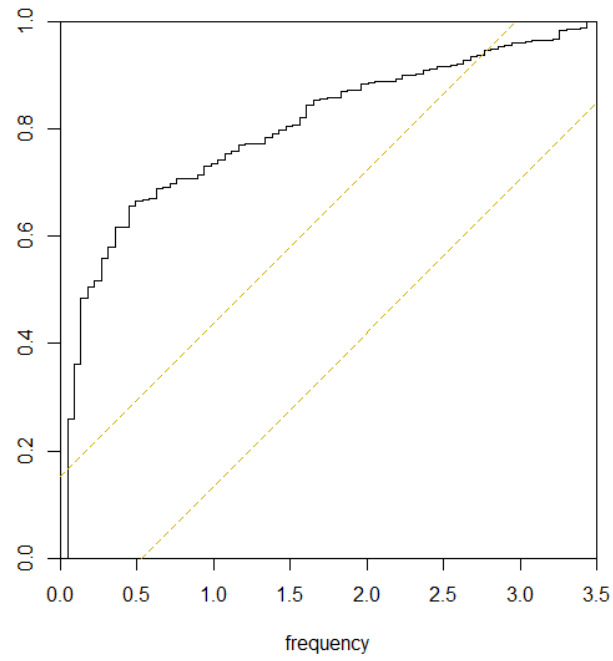


Ilustración 28: Comportamiento de ruido blanco de metano viernes con tratamiento.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
157	36.70%	6.80%	20.50%	32.40%	36.70%	42.00%	53.10%

Tabla 8: Distribución de metano viernes con tratamiento.

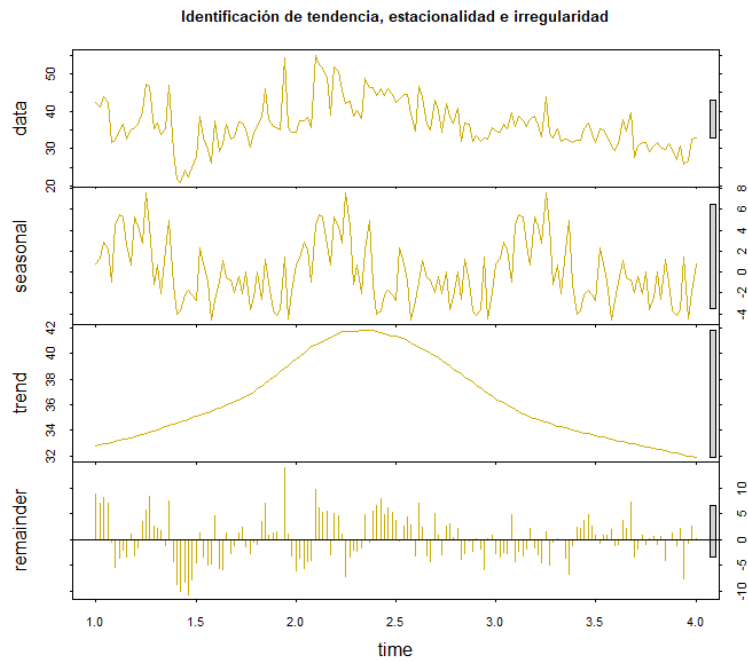


Ilustración 29: Componentes de metano sábado con tratamiento.

Periodograma integrado metano sabado con tratamiento

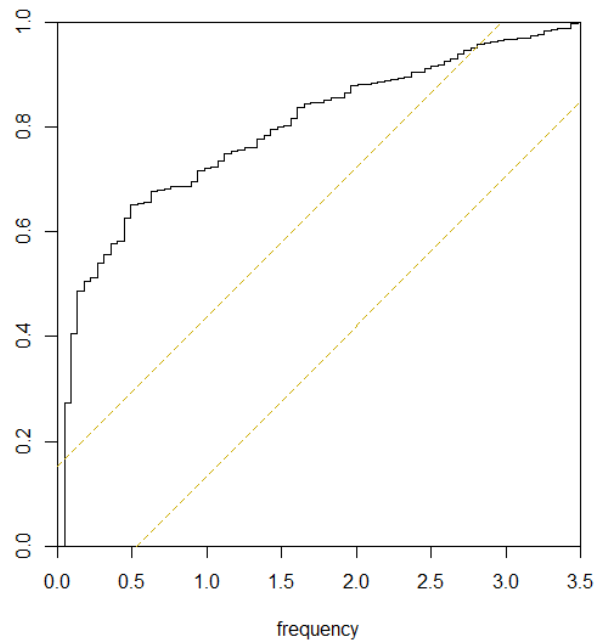


Ilustración 30: Comportamiento de ruido blanco de metano sábado con tratamiento.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
157	36.50%	6.40%	21.00%	32.20%	36.50%	39.70%	54.90%

Tabla 9: Distribución de metano sábado con tratamiento.

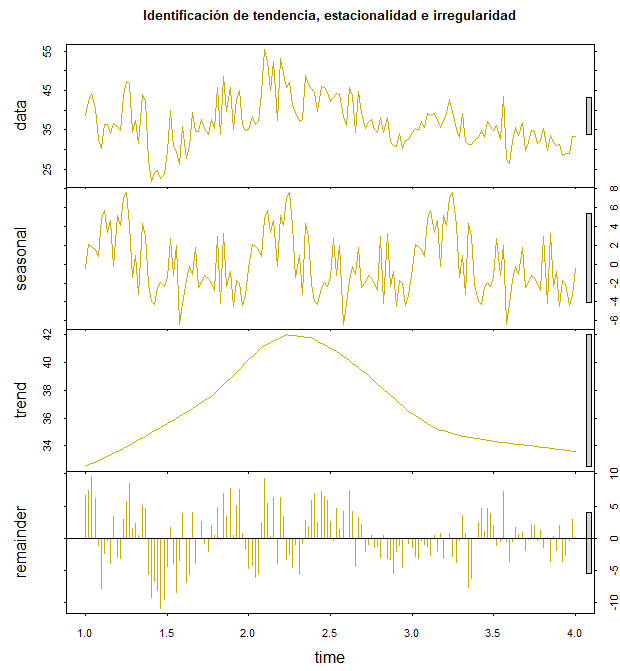


Ilustración 31: Componentes de metano domingo con tratamiento.

Periodograma integrado metano domingo con tratamiento

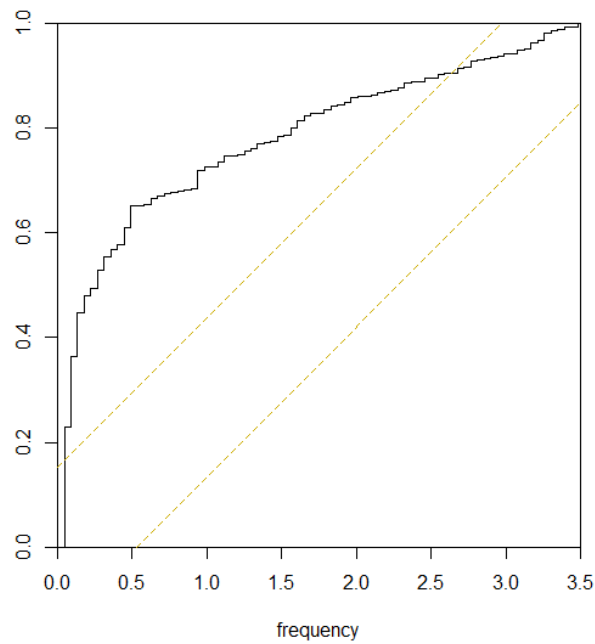


Ilustración 32: Comportamiento de ruido blanco de metano domingo con tratamiento.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
157	36.80%	6.30%	21.70%	35.50%	36.80%	40.00%	55.30%

Tabla 10: Distribución de metano domingo con tratamiento.

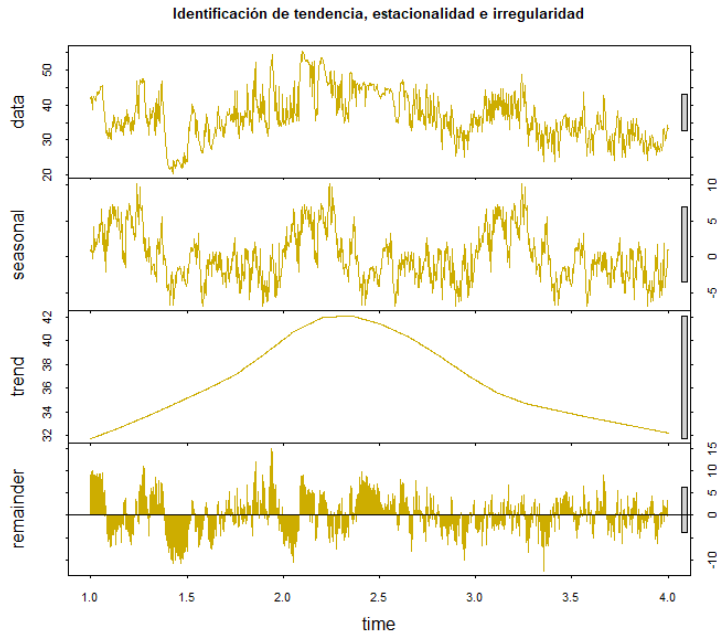


Ilustración 33: Componentes de metano diario con tratamiento.

Periodograma integrado metano diario con tratamiento

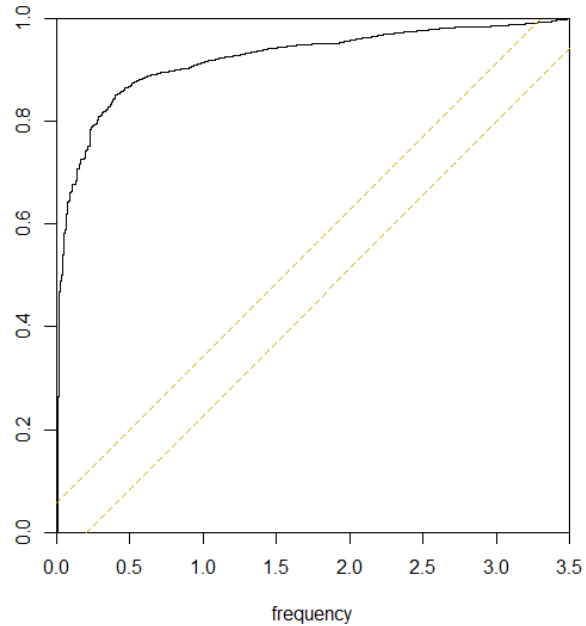


Ilustración 34: Comportamiento de ruido blanco de metano diario con tratamiento.

Datos	Media	Desviación	Mínimo	Cuartil 25%	Cuartil 50%	Cuartil 75%	Máximo
1096	36.70%	6.70%	20.50%	32.30%	36.70%	41.80%	55.30%

Tabla 11: Distribución de metano con tratamiento.

4.4 Elección de la serie

Para iniciar los modelos de predicción se toma la muestra diaria del metano, para esto la cantidad de datos fue una de las razones de la elección de esta muestra. Dado que nuestros datos diarios por naturaleza son más densos que los tomados por días de la semana, los procesos de las metodologías (promedio móvil ponderado, Holt-Winters, Box-Jenkins y redes neuronales artificiales) muestran lo siguiente.

Todos los resultados observados serán descritos en el capítulo 5, esto con el fin de poder generar un seguimiento de comparaciones y resultados de cada metodología aplicada.

4.5 Método promedio móvil ponderado

4.5.1 Aplicación del método promedio móvil ponderado

Se sabe que los pesos asignados son parte fundamental de esta metodología, por lo que no hay una fórmula que nos ayude a tener los pesos óptimos, sin embargo, podemos utilizar diferentes técnicas. De primer momento, es usual que el dato más reciente sea al que se le asigne el mayor peso posteriormente las ponderaciones más alejadas al dato reciente, tengan menor peso y por último las ponderaciones siempre sumen 100% en el total de pesos considerados. Dado esto, asignamos pesos a criterio propio para poder observar los errores y junto con la ayuda de Solver (herramienta de Excel) facilitamos la obtención del error mínimo.

Recordemos que una de las ventajas de este método es que puede ser computarizado de manera sencilla y rápida ya que para realizar un pronóstico solo es cuestión de hacer uso de la función SUMAPRODUCTO en Excel de donde la primera matriz (fija) serán los pesos que hayamos seleccionado y la segunda matriz serán los datos que queremos pronosticar. No olvidemos que esta segunda matriz será la matriz que estaremos moviendo para poder realizar nuestro pronóstico. A continuación, se muestran los pesos seleccionados en la metodología y adicional las ponderaciones hacen referencia a los días de la semana.

Número de ponderación	Valor de la ponderación
1	0.03
2	0.06
3	0.08
4	0.11
5	0.13
6	0.17
7	0.42

Tabla 12: Pesos encontrados con la herramienta Solver.

Número de ponderación	Valor de la ponderación
1	0.01
2	0.03
3	0.06
4	0.1
5	0.11
6	0.12
7	0.47

Tabla 13: Pesos elegidos a criterio propio.

En la gráfica siguiente tenemos la estimación del modelo promedio móvil ponderado.

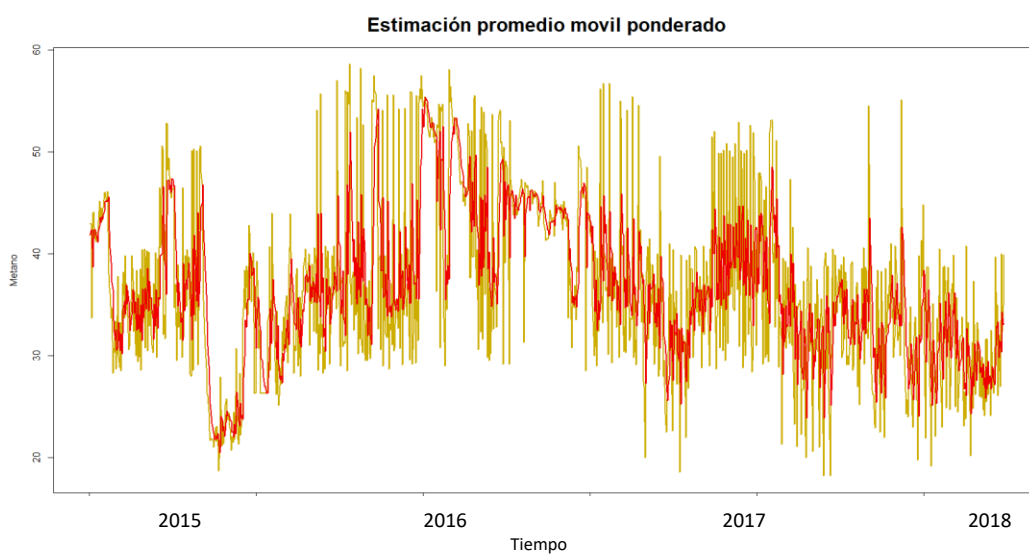


Ilustración 35: Estimación del modelo promedio móvil ponderado.

4.5.2 Comportamiento de la estimación del modelo promedio móvil ponderado

El componente irregular de las series se visualiza a continuación.

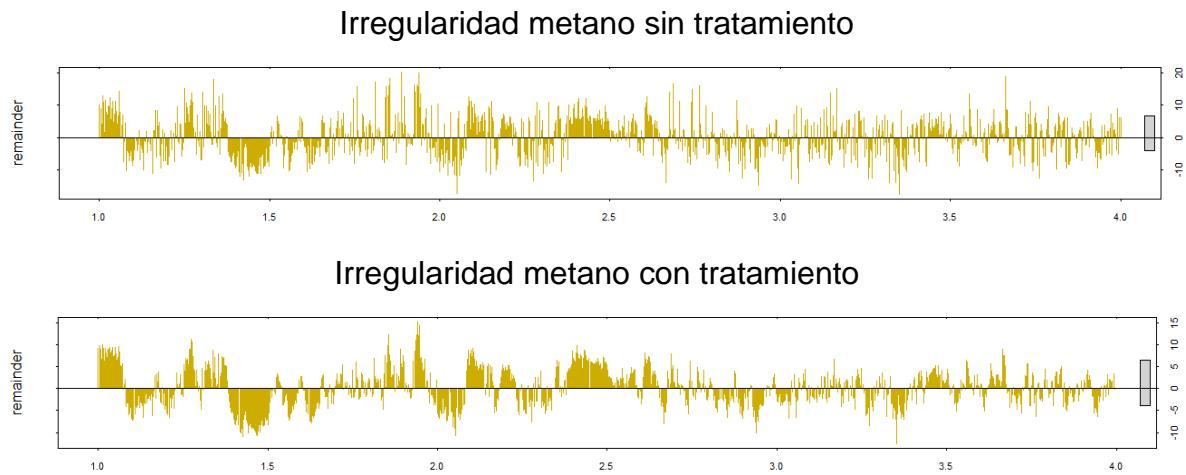


Ilustración 36: Comportamiento de la irregularidad de metano con y sin tratamiento.

Es importante considerar que el principal comportamiento del suavizamiento con el método de promedio móvil pondera es reducir el impacto que se tiene cuando la máquina de extracción forzada de biogás llega a fallar y no registra información alguna.

4.5.3 Pronóstico promedio móvil ponderado

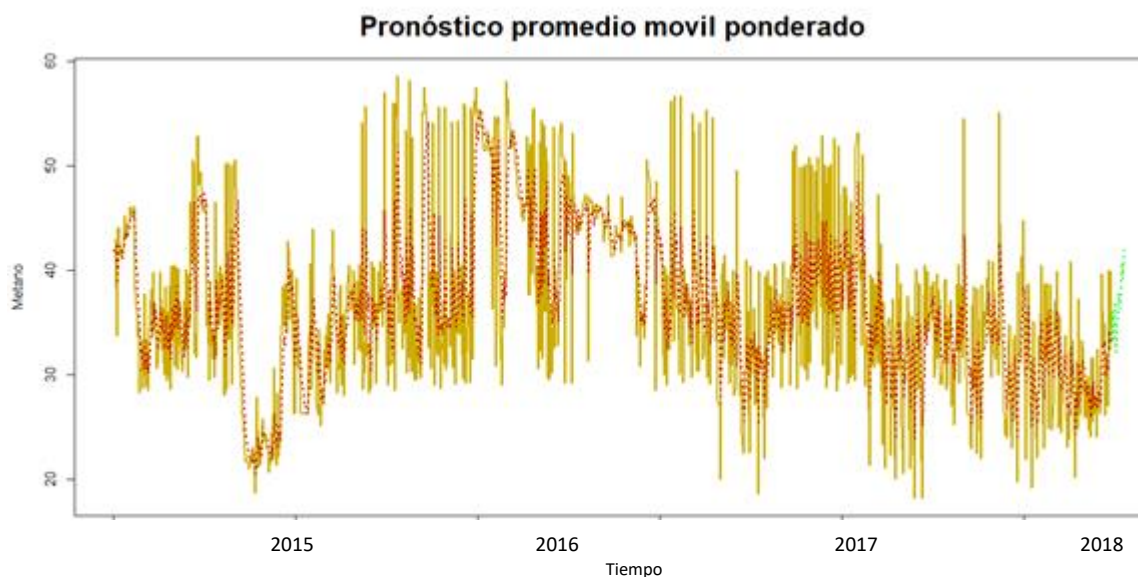


Ilustración 37: Pronóstico del modelo promedio móvil ponderado.

Por lo tanto, se observa el pronóstico de los primeros días del 2018, los errores mostrados hacen la referencia al método ponderado de la herramienta de Solver, esto para poder tener un criterio de aceptación ante un valor expuesto por un analista.

Tatramiento	MAE	MAPE	R ²
Con	5.11	14.45	0.3632

Tabla 14: Error de modelo promedio móvil ponderado.

4.6 Método Holt-Winters

4.6.1 Aplicación del método Holt-Winters

Para la estimación de Holt-Winters primero se considera el comportamiento de la tendencia y de la irregularidad del metano, para poder observar estos componentes realizamos la prueba de raíz unitaria para comprobar la tendencia de la serie con y sin tratamiento.

Resultado de metano sin tratamiento, aplicación del test “PP.test()” y “adf.test()”.

```
> PP.test(Prediccion)
Phillips-Perron Unit Root Test
data: Prediccion
Dickey-Fuller = -18.373, Truncation lag parameter = 7, p-value = 0.01

> adf.test(Prediccion)
Augmented Dickey-Fuller Test
data: Prediccion
Dickey-Fuller = -3.6557, Lag order = 10, p-value = 0.02733
alternative hypothesis: stationary
```

Ilustración 38: Pruebas de la raíz unitaria de metano sin tratamiento.

Resultado de metano con tratamiento, aplicación del test “PP.test()” y “adf.test()”.

```
> PP.test(Prediccion)
Phillips-Perron Unit Root Test
data: Prediccion
Dickey-Fuller = -7.3666, Truncation lag parameter = 7, p-value = 0.01

> adf.test(Prediccion)
Augmented Dickey-Fuller Test
data: Prediccion
Dickey-Fuller = -4.1696, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Ilustración 39: Pruebas de la raíz unitaria de metano con tratamiento.

Para observar el comportamiento estacional de las series las graficamos sobre el tiempo.

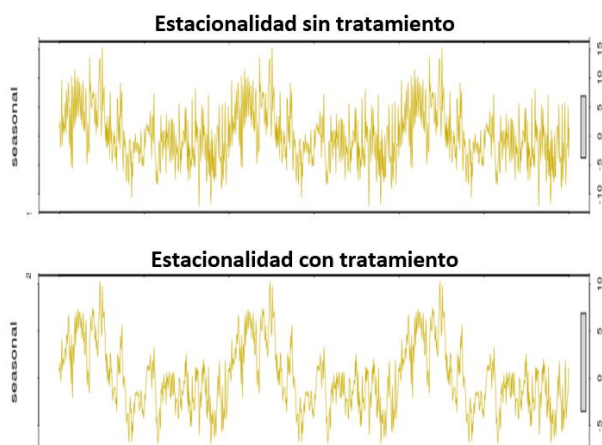


Ilustración 40: Comportamiento de la estacionalidad de metano con y sin tratamiento.

El método de Holt-Winter puede ser utilizado ya sea por un modelo de adición o multiplicación de los componentes, en este caso será seleccionado un método aditivo para la predicción de metano con y sin tratamiento de datos.

Primero observaremos al metano sin tratamiento con su estimación de Holt-Winters aditivo.

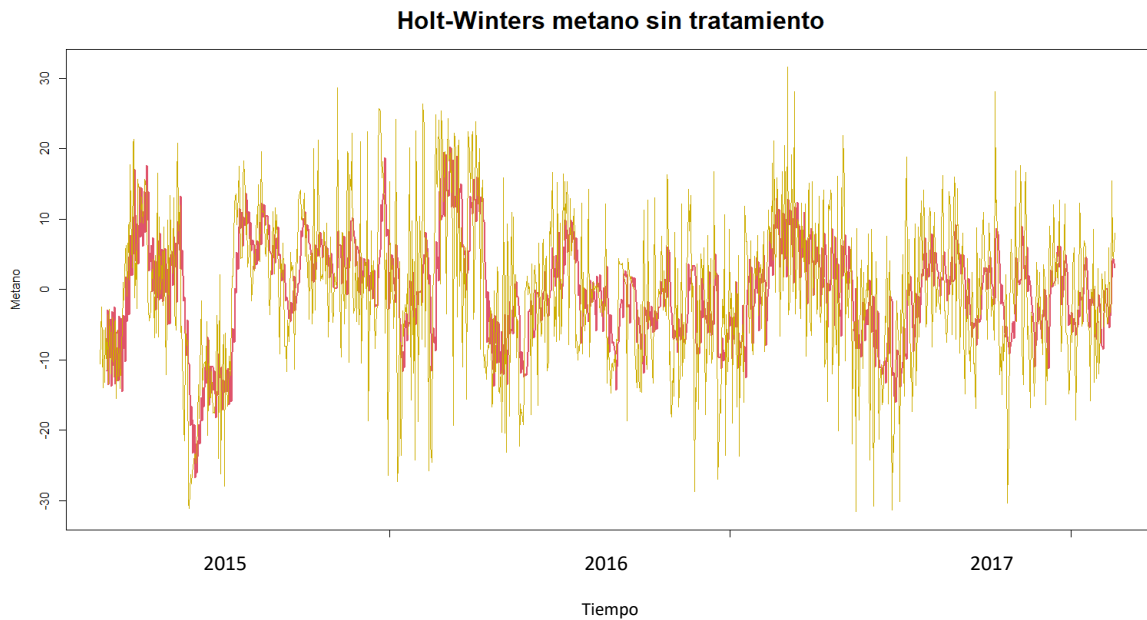


Ilustración 41: Estimación del modelo Holt-Winters de metano sin tratamiento.

En segundo caso observaremos la estimación de metano con tratamiento, aplicando el método aditivo de la misma forma.

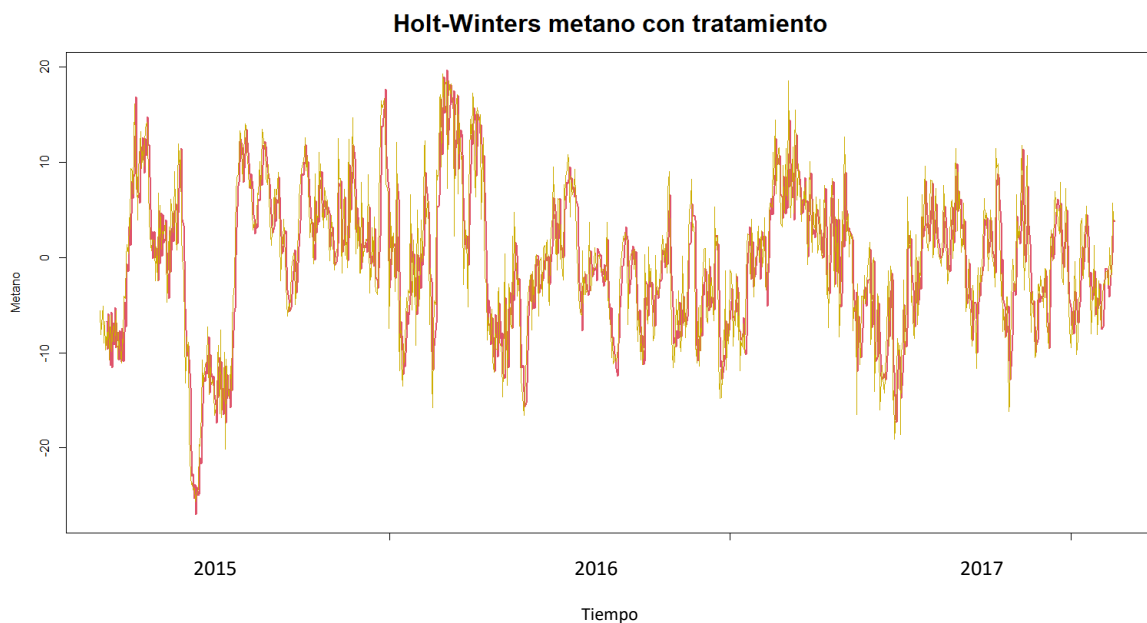


Ilustración 42: Estimación del modelo Holt-Winters de metano con tratamiento.

4.6.2 Comportamiento de la estimación del modelo Holt-Winters

Tenemos ahora que observar de cada una de las estimaciones el comportamiento de los residuales, así como las gráficas de función de correlación muestral.

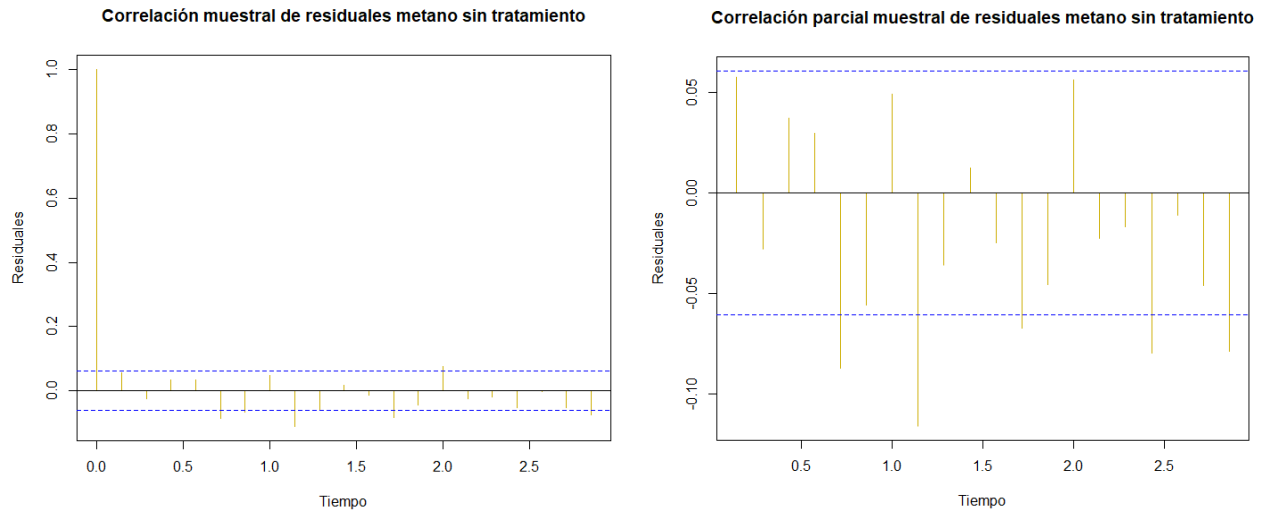


Ilustración 43: Correlación de residuales del modelo Holt-Winters de metano sin tratamiento.

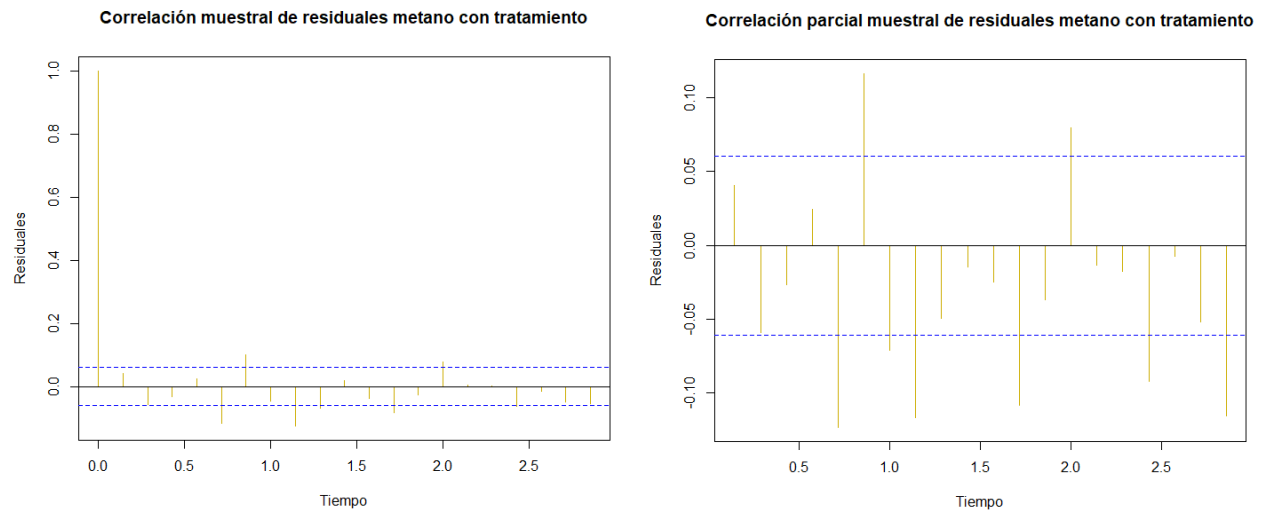


Ilustración 44: Correlación de residuales del modelo Holt-Winters de metano con tratamiento.

Se aplica la prueba de Ljung-Box para validar lo observado en las gráficas anteriores, esto permite validar el supuesto de correlación muestral entre los residuales.

Metano sin tratamiento

Metano con tratamiento

```
> Box.test(HoW$residuals, lag=20, type="Ljung-Box") > Box.test(HoW$residuals, lag=20, type="Ljung-Box")
```

Box-Ljung test

Box-Ljung test

data: HoW\$residuals
X-squared = 66.093, df = 20, p-value = 7.807e-07

data: HoW\$residuals
X-squared = 81.783, df = 20, p-value = 1.952e-09

Ilustración 45: Prueba Ljung-Box a los residuos del modelo Holt-Winters de metano con y sin tratamiento.

Ahora visualizamos la normalidad de la serie mediante las siguientes gráficas.

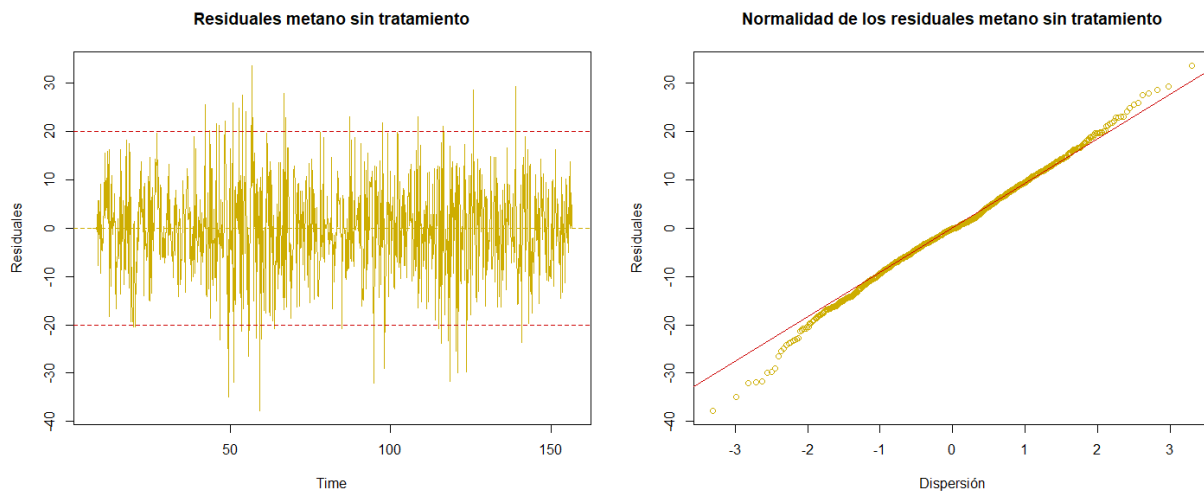


Ilustración 46: Normalidad de residuales del modelo Holt-Winters de metano sin tratamiento.

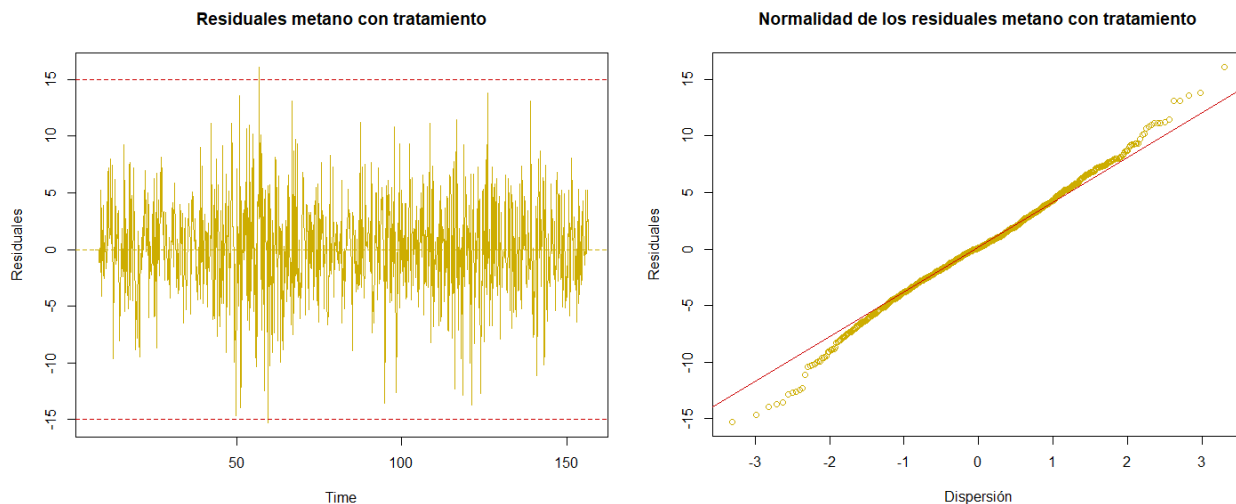


Ilustración 47: Normalidad de residuales del modelo Holt-Winters de metano con tratamiento.

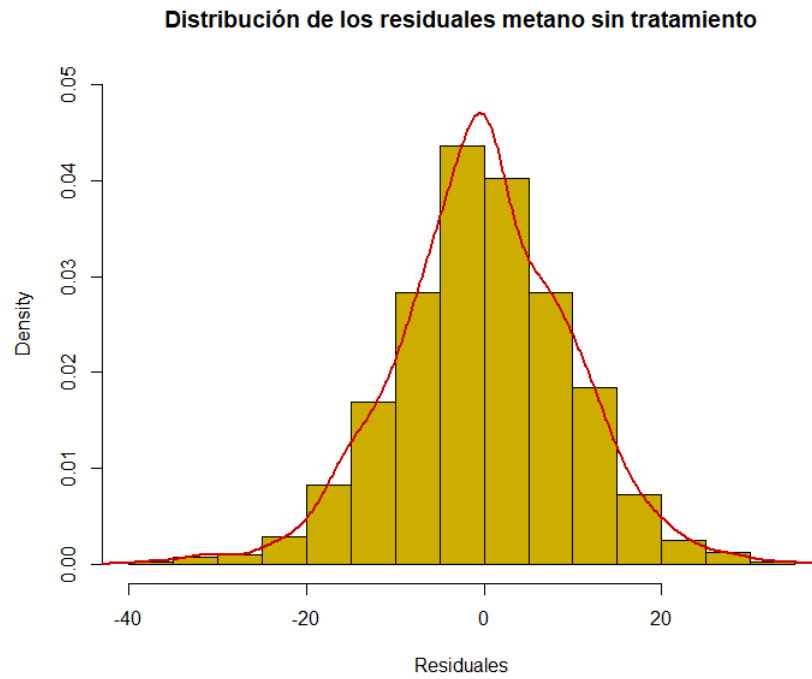


Ilustración 48: Distribución de residuales del modelo Holt-Winters de metano sin tratamiento.

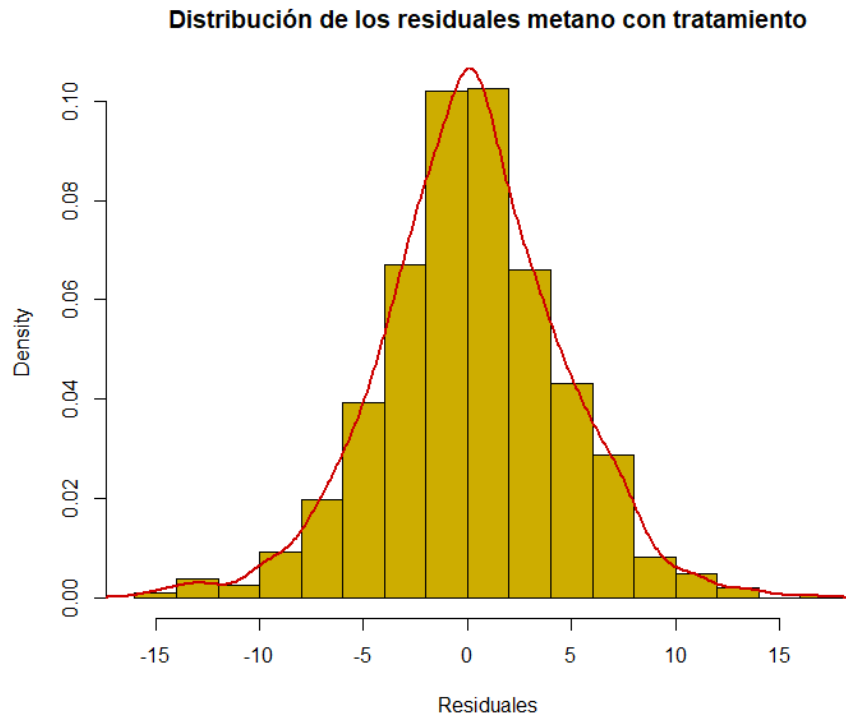


Ilustración 49: Distribución de residuales del modelo Holt-Winters de metano con tratamiento.

4.6.3 Pronóstico Holt-Winters

Se obtiene el pronóstico y se visualiza en las siguientes gráficas, a su vez en las tablas se muestran los coeficientes de las series. Las estimaciones están a 15 días para ambos modelos.

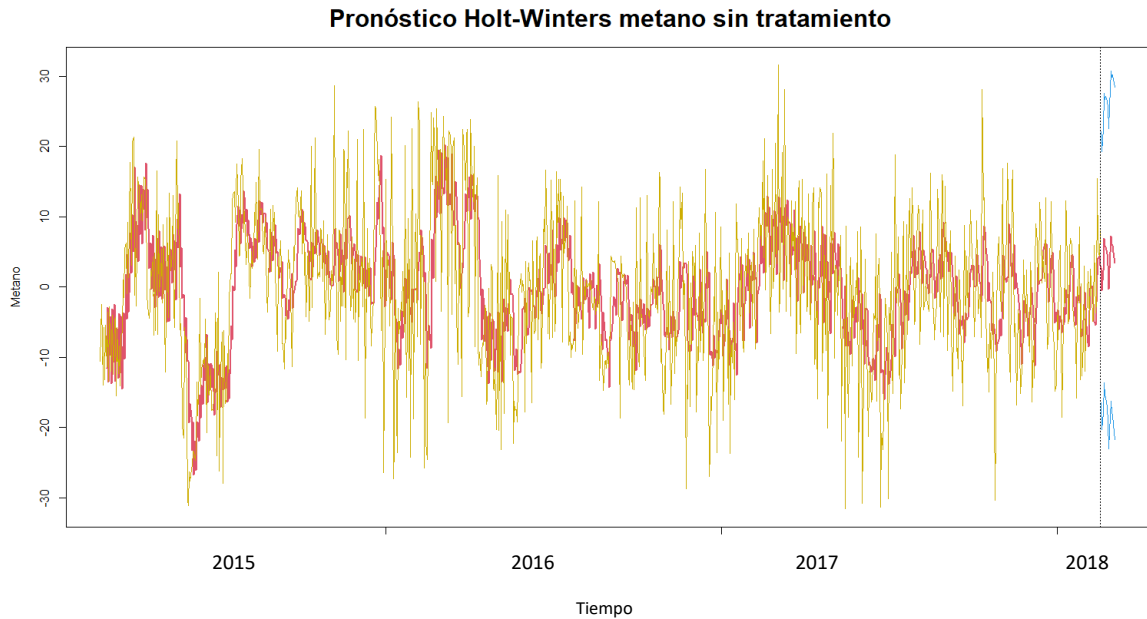


Ilustración 50: Pronóstico de modelo Holt-Winters de metano sin tratamiento.

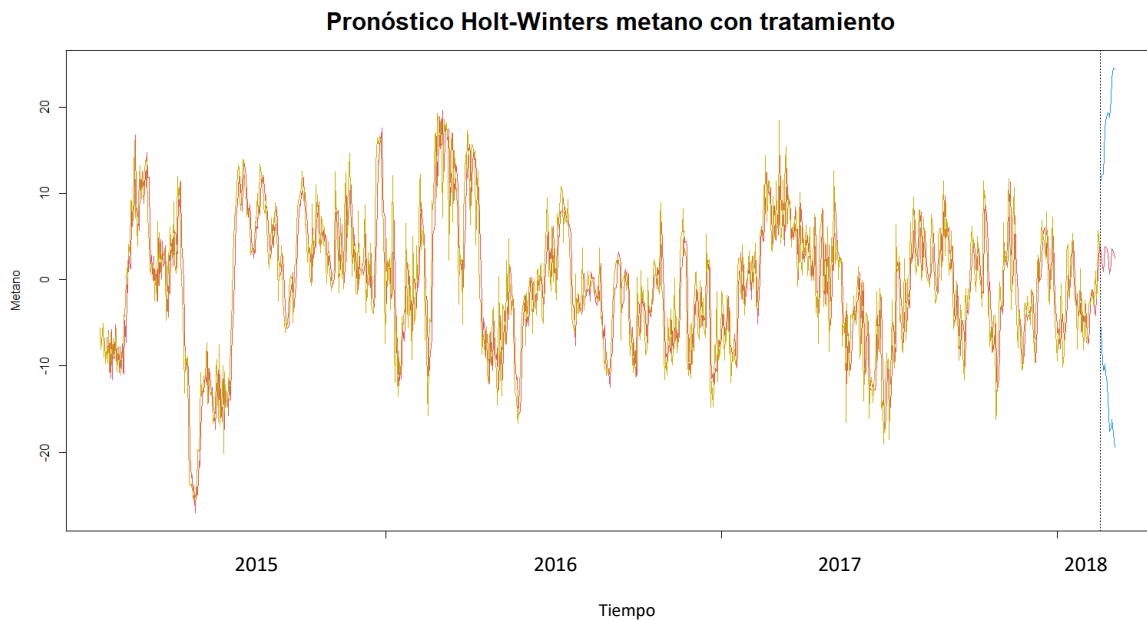


Ilustración 51: Pronóstico de modelo Holt-Winters de metano con tratamiento.

Tratamiento	a	b	S1	S2	S3	S4	S5	S6	S7
Sin	4.78	0.03	-2.09	-5.49	-2.65	1.92	0.6	0.57	-0.47
Con	1.91	-0.03	1.14	0.23	-0.93	0.28	2.1	2.08	1.92

Tabla 15: Coeficientes de estimación de metano con y sin tratamiento.

Por último, tenemos los errores de cada uno de los modelos aplicados.

Tatramiento	MAE	MAPE	R ²
Con	3.32	14.45	0.6864
Sin	7.57	14.45	0.1991

Tabla 16: Errores de los modelos Holt-Winters.

4.7 Método Box-Jenkins

4.7.1 Aplicación del método Box-Jenkins

Box-Jenkins es una de las metodologías con mayor grado de implementación debido a las validaciones requeridas para un buen modelo, para la estimación de metano se tendrán las gráficas y pruebas que permitan facilitar la interpretación de esta metodología. Cabe mencionar que en el capítulo 2 se hace la referencia que la metodología de Box-Jenkins es un proceso que estima modelos ARIMA o SARIMA.

Primero tendremos la comparación de las pruebas de raíz unitaria, esto considerando una diferencia aplicada a cada una de las series.

Diferencia estacional metano sin tratamiento $k = 0$

```
> adf.test(diff(Prediccion, alternative="stationary", k=0))
```

Augmented Dickey-Fuller Test

```
data: diff(Prediccion, alternative = "stationary", k = 0)
Dickey-Fuller = -13.969, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Ilustración 52: Prueba de Dickey-Fuller de metano sin tratamiento.

Diferencia estacional metano con tratamiento K = 7

```
> adf.test(diff(Prediccion, lag =7))
```

Augmented Dickey-Fuller Test

```
data: diff(Prediccion, lag = 7)
Dickey-Fuller = -9.9095, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Ilustración 53: Prueba de Dickey-Fuller de metano con tratamiento.

Ahora se generan las gráficas de correlación muestral y correlación parcial muestral de nuestras series.

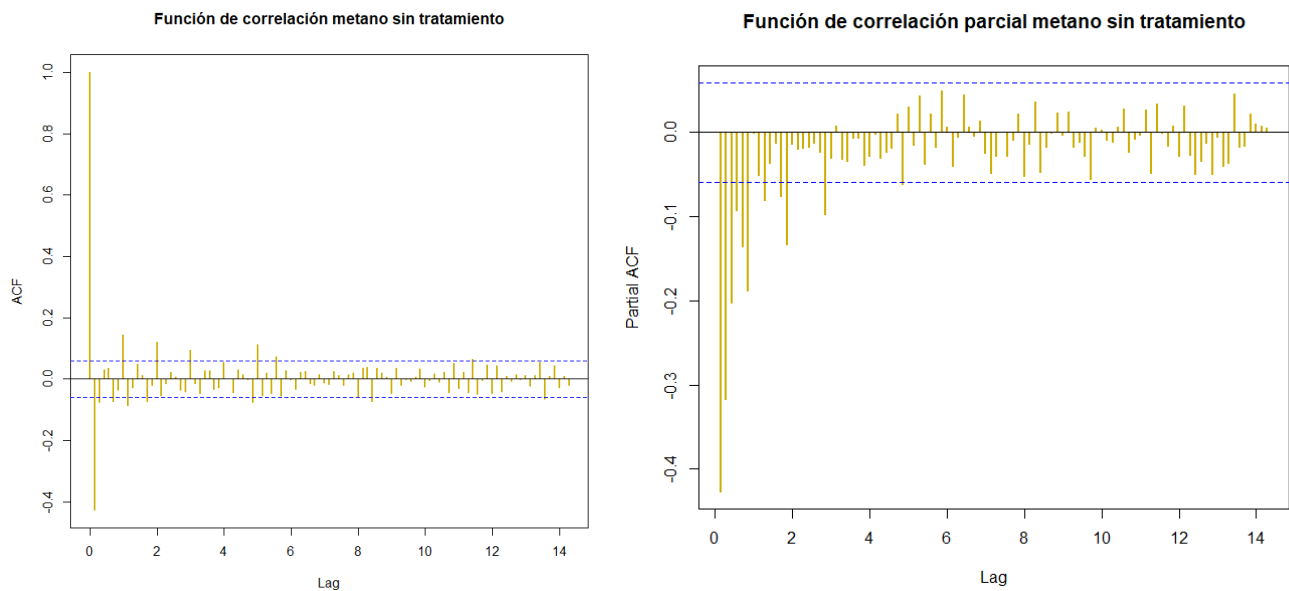


Ilustración 54: Funciones de correlación de metano sin tratamiento.

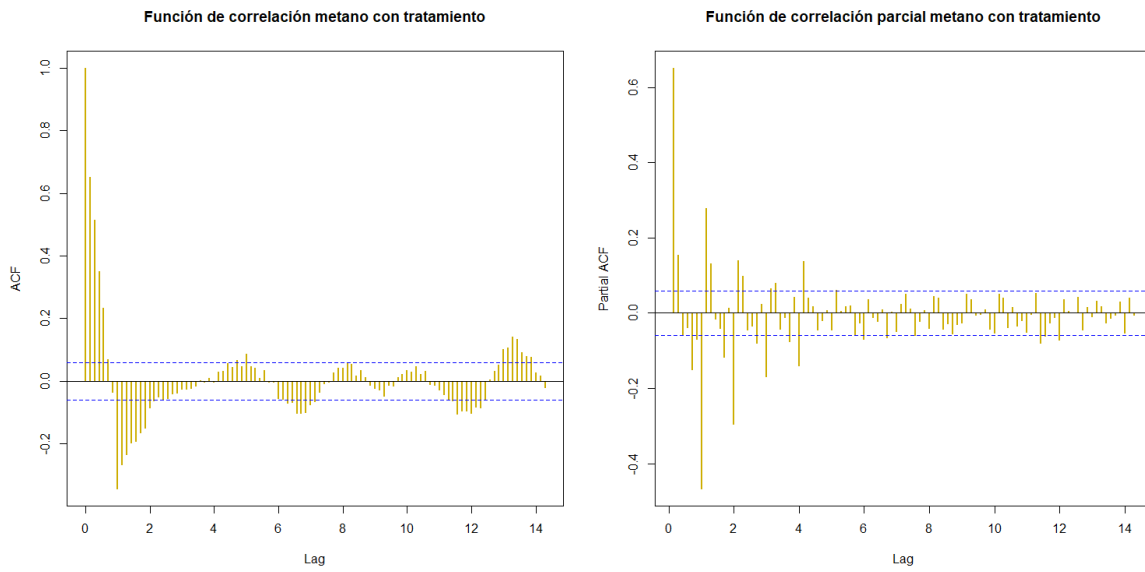


Ilustración 55: Funciones de correlación de metano con tratamiento.

Ahora se muestran las estimaciones de los modelos con y sin tratamiento.

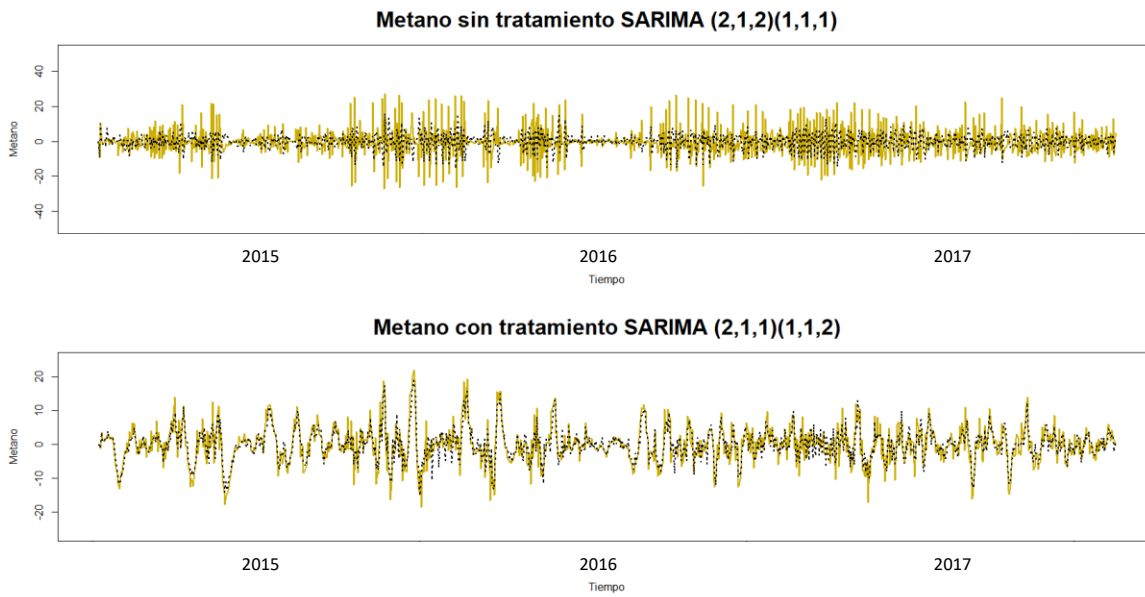


Ilustración 56: Estimaciones de los modelos Box-Jenkins de metano con y sin tratamiento.

4.7.2 Comportamiento de la estimación del modelo Box-Jenkins

4.7.2.1 Residuales, parsimonia y errores de los residuales de metano sin tratamiento

Consideramos como primer punto la media cero en los residuales.

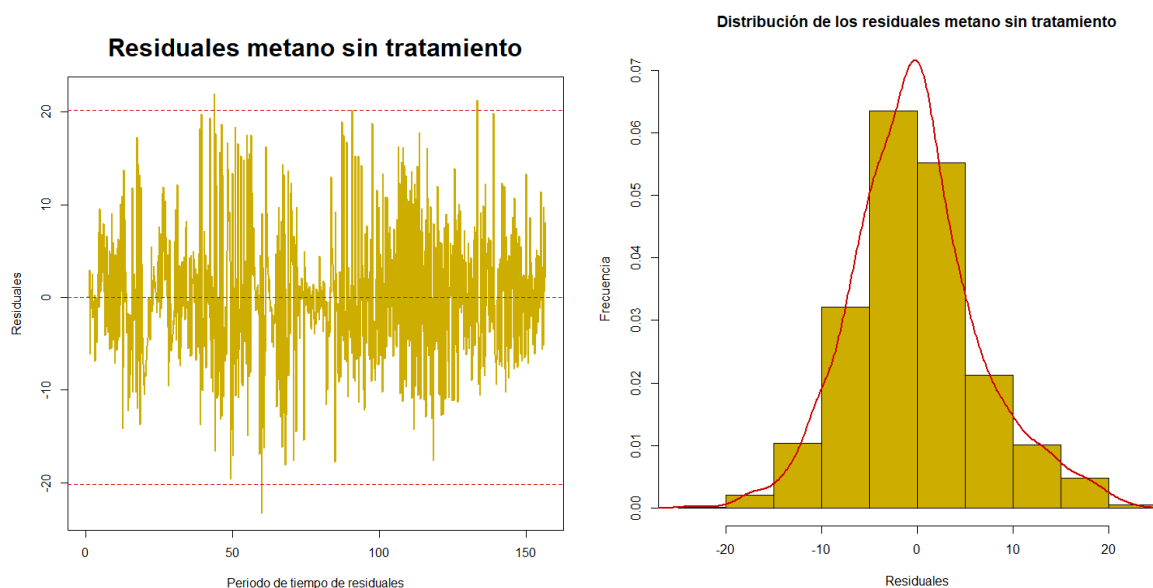


Ilustración 57: Comportamiento de media cero en residuales del modelo Box-Jenkins de metano sin tratamiento.

Como siguiente validación será la heterocedasticidad de la serie, la cual facilita la comprensión de que si la varianza crece con el tiempo o no.

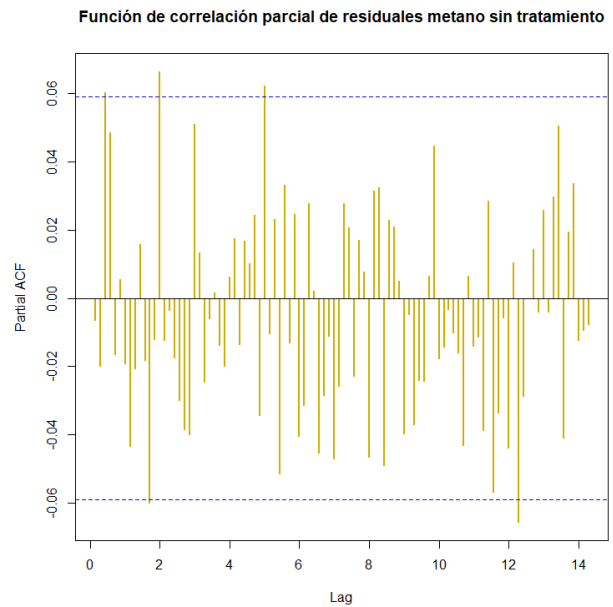
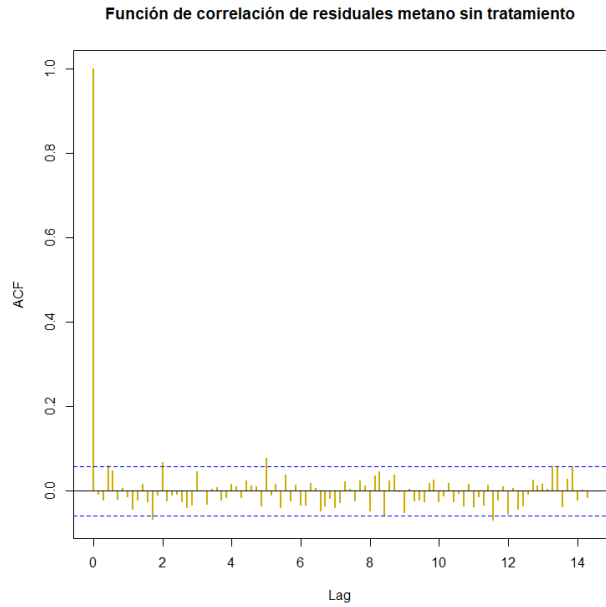
```
> leveneTest(residuales,Factor)
Levene's Test for Homogeneity of Variance (center = median)
  Df F value Pr(>F)
group 1 3.7236 0.05391 .
     1093

> bartlett.test (residuales,Factor)
Bartlett test of homogeneity of variances

data:  residuales and Factor
Bartlett's K-squared = 0.037068, df = 1, p-value = 0.8473
```

Ilustración 58: Prueba de heterocedasticidad de residuales del modelo Box-Jenkins de metano sin tratamiento.

Ahora bien, la correlación muestral de residuales en las siguientes graficas en donde se espera observar que no haya barras significativas, serán la una de las principales validaciones aplicadas a nuestra serie de tiempo.



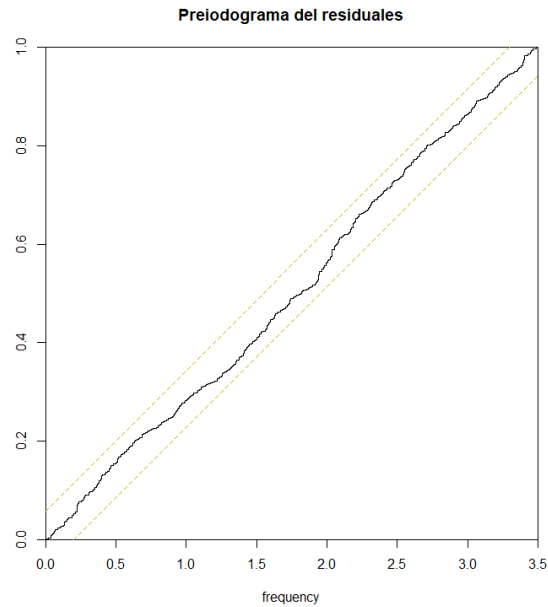
```
> dwtest (residuales~Factor)

Durbin-Watson test

data: residuales ~ Factor
DW = 2.0116, p-value = 0.564
alternative hypothesis: true autocorrelation is greater than 0
```

Ilustración 59: Correlación de residuales del modelo Box-Jenkins de metano sin tratamiento.

Como siguiente validación de esta metodología es la normalidad de los residuales, ya que se busca que llegue a tener un comportamiento de ruido blanco gaussiano.



```
> ad.test(residuales)
```

Anderson-Darling normality test

```
data: residuales
A = 8.351, p-value < 2.2e-16
```

```
> lillie.test(residuales)
```

Lilliefors (Kolmogorov-Smirnov) normality test

```
data: residuales
D = 0.074246, p-value = 1.41e-15
```

Ilustración 60: Normalidad de residuales del modelo Box-Jenkins del metano sin tratamiento.

La siguiente gráfica es para observar si no existen datos atípicos en los residuales durante el tiempo.

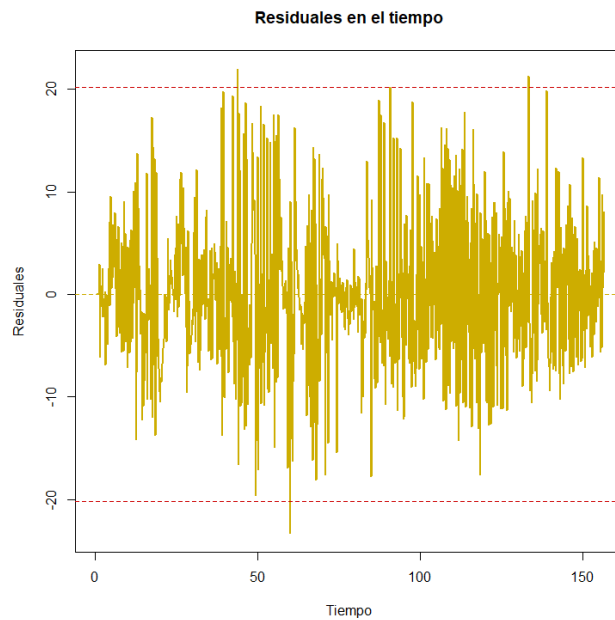


Ilustración 61: Valores atípicos en residuales del modelo Box-Jenkins de metano sin tratamiento.

La parsimonia del modelo es un tema para no sobre ajustar al modelo, mostramos los coeficientes para el modelo de metano sin tratamiento.

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar1	0.210584	0.035576	5.9193	3.234e-09	***
ar2	0.077687	0.034179	2.2729	0.0230300	*
ma1	-1.918579	0.019669	-97.5419	< 2.2e-16	***
ma2	0.918604	0.019518	47.0646	< 2.2e-16	***
sar1	0.116144	0.032855	3.5350	0.0004077	***
sma1	-0.995527	0.030175	-32.9917	< 2.2e-16	***

Ilustración 62: Correlación de estimación SARIMA de metano sin tratamiento.

4.7.2.2 Residuales, parsimonia y errores de los residuales de metano con tratamiento

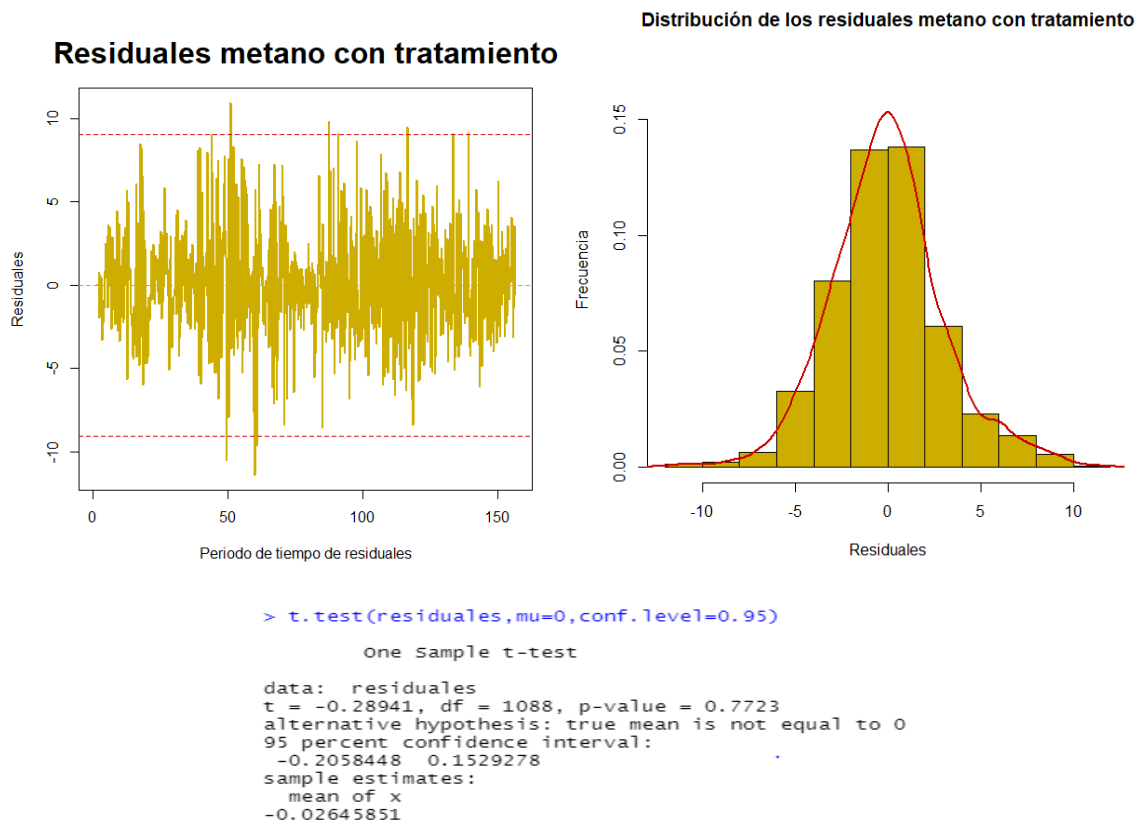


Ilustración 63: Comportamiento de media cero en residuales del modelo Box-Jenkins de metano con tratamiento.

Ahora observamos la heterocedasticidad de la serie.

```
> leveneTest(residuales,Factor)
Levene's Test for Homogeneity of Variance (center = median)
  Df F value Pr(>F)
group 1 0.3592 0.5491
 1087

> bartlett.test (residuales,Factor)
Bartlett test of homogeneity of variances
data: residuales and Factor
Bartlett's K-squared = 0.44709, df = 1, p-value = 0.5037
```

Ilustración 64: Prueba de heterocedasticidad de residuales del modelo Box-Jenkins de metano con tratamiento.

La correlación en la serie con tratamiento se muestra de la siguiente forma.

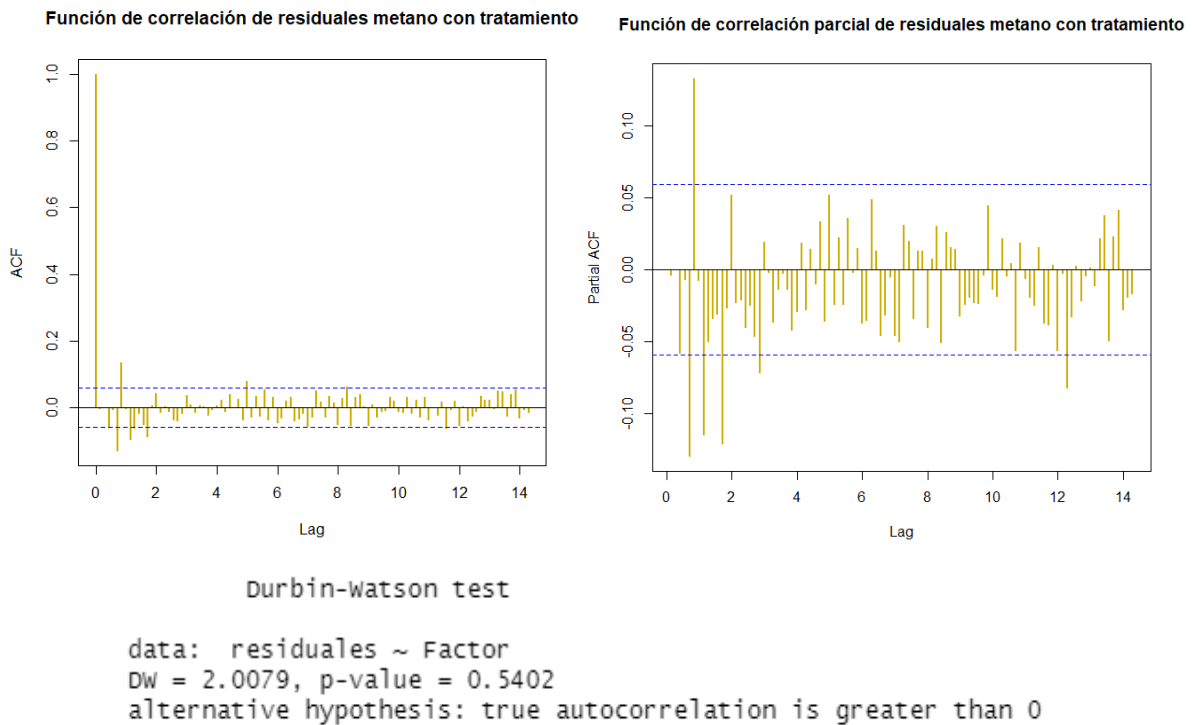
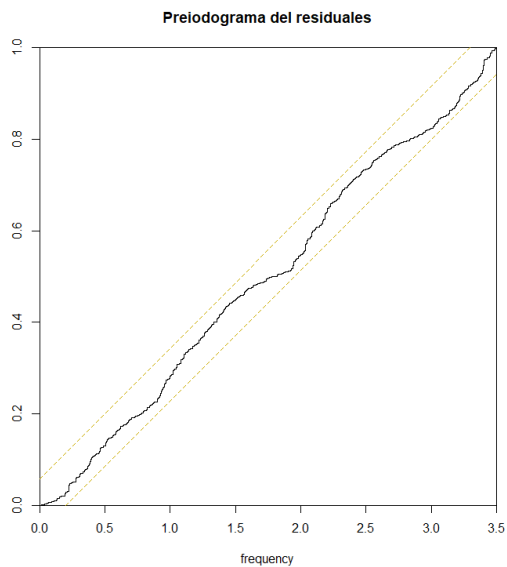


Ilustración 65: Correlación de residuales del modelo Box-Jenkins de metano con tratamiento.

La normalidad de los residuales de la serie con tratamiento.



Lilliefors (Kolmogorov-Smirnov) normality test

data: residuales
D = 0.05286, p-value = 1.454e-07

Anderson-Darling normality test

data: residuales
A = 3.1547, p-value = 6.535e-08

Ilustración 66: Normalidad de residuales del modelo Box-Jenkins de metano con tratamiento.

Datos atípicos en los residuales con tratamiento de datos.

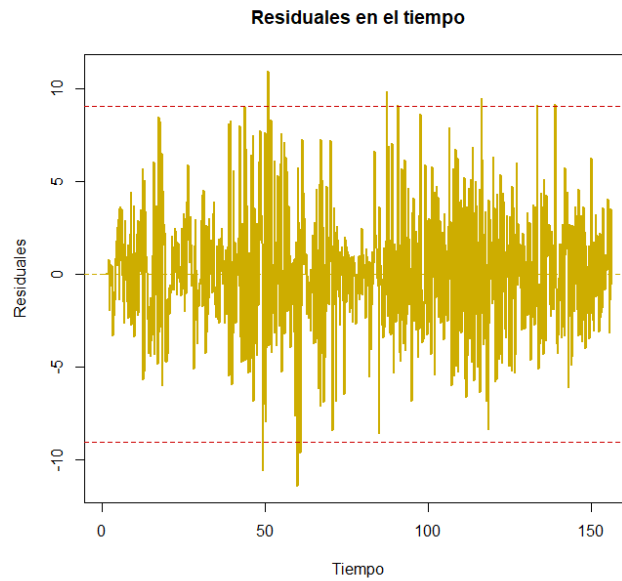


Ilustración 67: Valores atípicos en residuales del modelo Box-Jenkins de metano con tratamiento.

Parsimonia del modelo de la serie con tratamiento.

	Estimate	Std. Error	z value	Pr(> z)	
ar1	0.027375	0.389373	0.0703	0.94395	
ar2	-0.072169	0.101235	-0.7129	0.47592	
ma1	-0.336149	0.402616	-0.8349	0.40377	
sar1	-0.069881	0.033372	-2.0940	0.03626	*
sma1	-1.956792	0.016335	-119.7907	< 2e-16	***
sma2	0.957091	0.016070	59.5564	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Ilustración 68: Correlación de estimación SARIMA de metano con tratamiento.

4.7.3 Pronóstico Box-Jenkins

Se genera la estimación a 30 días para el comportamiento de metano con y sin tratamiento.

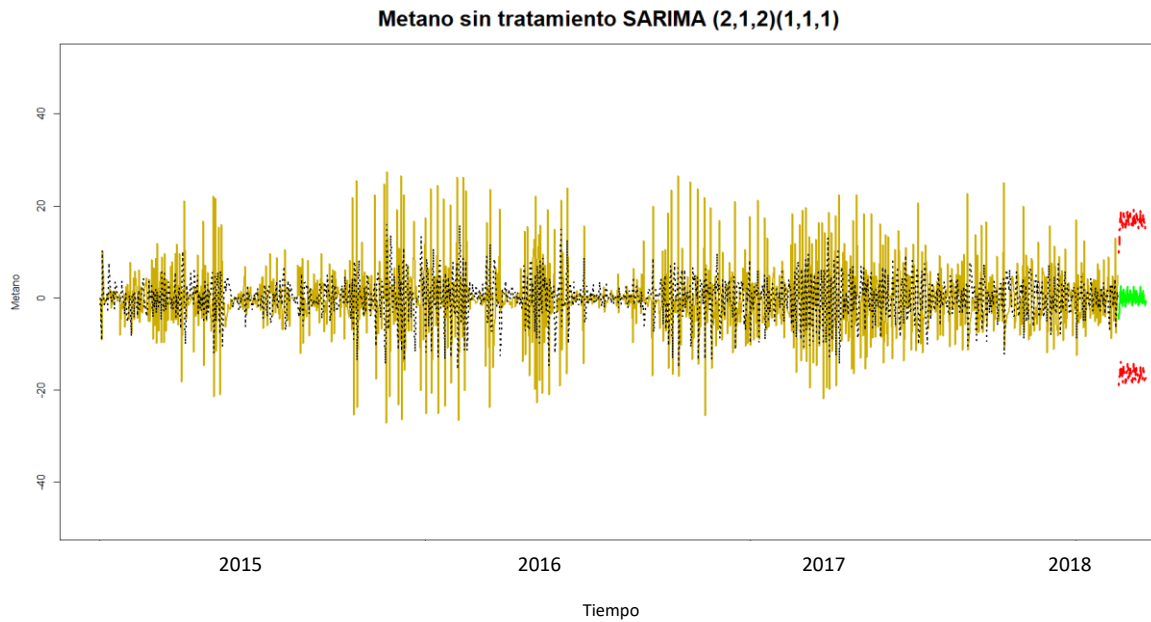


Ilustración 69: Pronóstico del modelo Box-Jenkins de metano sin tratamiento.

Obtenemos los errores de la estimación a 30 días de la serie de metano sin tratamiento.

Tatramiento	MAE	MAPE	R ²
Sin	5.29	238.7	0.2745

Tabla 17: Error del modelo SARIMA de metano sin tratamiento.

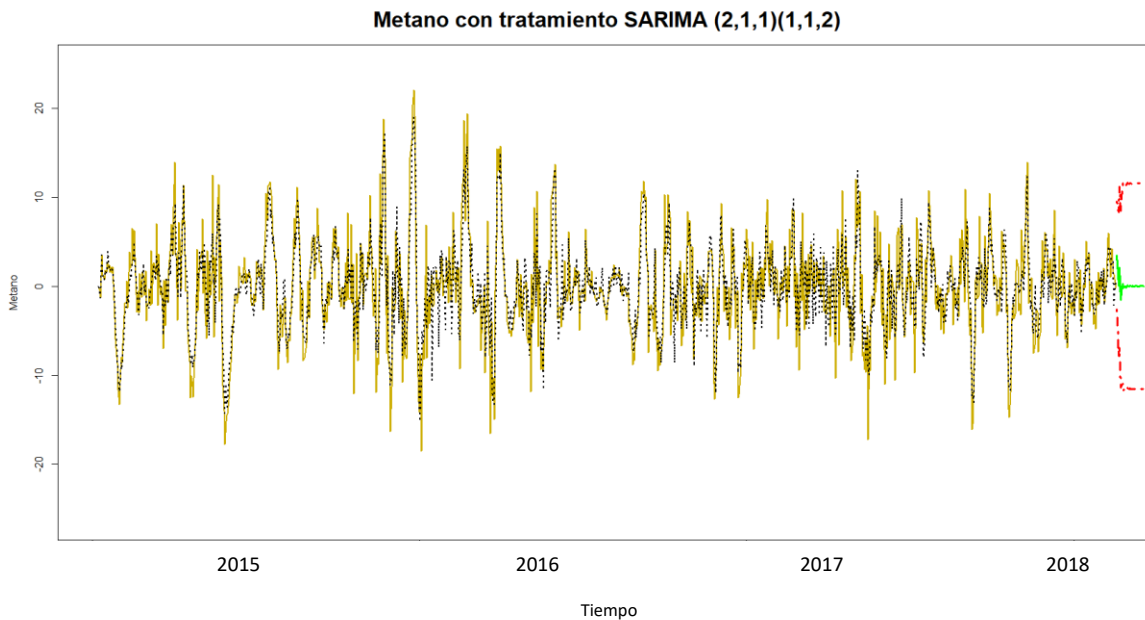


Ilustración 70: Pronóstico del modelo Box-Jenkins de metano con tratamiento.

Obtenemos los errores de la estimación a 30 días de la serie de metano con tratamiento.

Tatramiento	MAE	MAPE	R ²
Con	2.28	298.3	0.6945

Tabla 18: Error del modelo SARIMA de metano con tratamiento.

4.8 Método red neuronal artificial

4.8.1 Aplicación del método red neuronal artificial

Como primer paso para la ejecución de la red neuronal se tiene que considerar que la serie con y sin tratamiento tenga un comportamiento normal y estacionario, partiendo de esto, el siguiente punto se basa en generar un ajuste a los datos ya que estos serán enviados a un modelo que buscará una secuencia de patrones de entrada y salida, método de secuencia en el entrenamiento del modelo.

El segundo paso para la ejecución de la red neuronal artificial es tener en cuenta el método de activación del modelo, por lo cual para el caso de esta será una función que de un rango entre los valores de -1 y 1.

Ahora bien, para poder entrenar a la red neuronal artificial con la serie de tiempo se genera un escalonamiento distribuido entre -1 y 1, de donde se utilizará la función MinMaxScaler().

Posteriormente ocupamos la función `series_to_supervised` la cual recibe cuatro parámetros descritos a continuación:

- `Data`: Secuencia de observaciones como una lista o matriz 2D NumPy.
- `Num_In` = Número de observaciones de retraso como entrada.
- `Num_Out` = Número de observaciones de salida.
- `DropNaN` = Si se borran o no las filas con valores NaN.

Esta función devuelve un solo valor:

- `Return` = Pandas DataFrame de la serie transformada.

Para esta función utilizamos una herramienta de Pandas llamada `shift` la cual empuja y copia el arreglo hacia adelante o hacia atrás recorriéndolo y posicionando el valor anterior como una nueva columna.

```
# Normalidad de Los datos
scaler = MinMaxScaler(feature_range=(-1, 1))
values=values.reshape(-1, 1) # esto lo hacemos porque tenemos 1 sola dimension
scaled = scaler.fit_transform(values)
# Muestra del aprendizaje supervisado
reframed = series_to_supervised(scaled, PASOS, 1)
reframed.head()
```

Ilustración 71: Normalidad y escalamiento de los datos de entrada.

Es importante segmentar la muestra en validación y entrenamiento, esto es recomendable para que los datos sean comparables. Por otra parte, es importante considerar que se genere una estructura de entrada para la red neuronal artificial. Esta estructura se puede ver reflejada en lo que corresponden a nuestros datos de entrenamiento transformados.

```
# Estructura de Tridimensional de Los datos.
values = reframed.values
reshape input to be 3D [samples, timesteps, features]
x_train = x_train.reshape((x_train.shape[0], 1, x_train.shape[1]))
x_val = x_val.reshape((x_val.shape[0], 1, x_val.shape[1]))
print(x_train.shape, y_train.shape, x_val.shape, y_val.shape)
```

Ilustración 72: Bases de entrenamiento y validación para la red neuronal artificial.

Una vez creada la base de entrada se genera la red neuronal artificial.

```
def crear_modeloFF():
    model = Sequential()
    model.add(Dense(N, input_shape=(1,PASOS),activation='relu'))
    model.add(Flatten())
    model.add(Dense(1, activation='tanh'))
    model.compile(loss='mean_absolute_error',optimizer='Adam',metrics=["mse"])
    model.summary()
    return model
```

Ilustración 73: Creación de la red neuronal artificial.

La variable “N” es igual al número de nodos en la capa oculta. Para los datos del RSU BX se tiene la siguiente cantidad a entrenar y la cantidad de datos a validar.

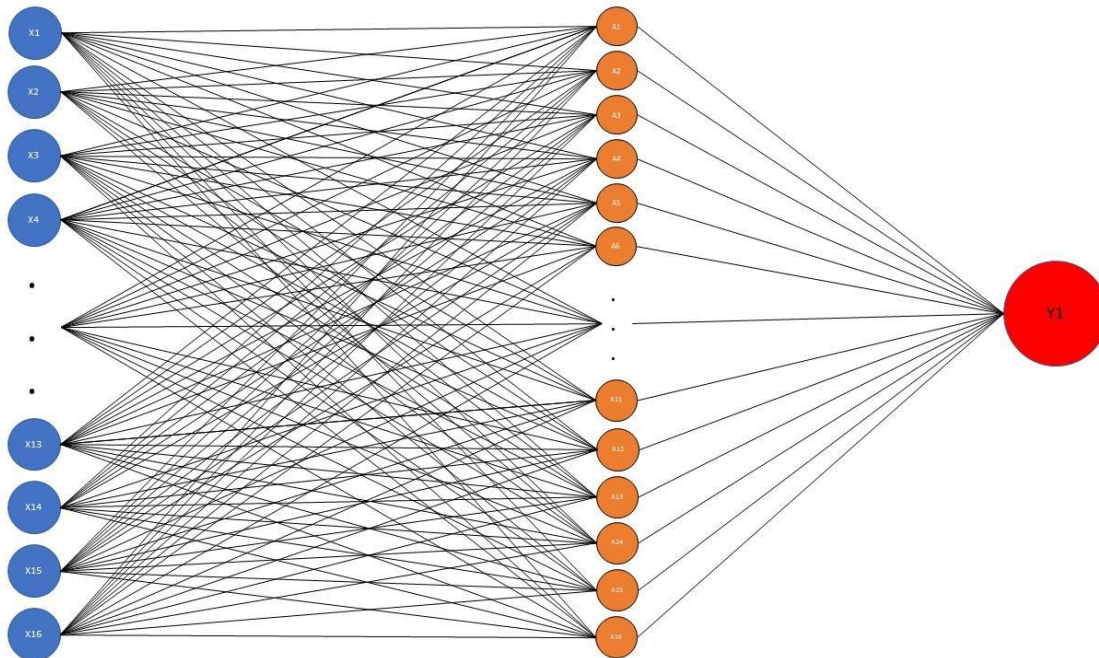


Ilustración 74: Diagrama de la red neuronal artificial.

(1081, 1, 16) (1081,) (35, 1, 16) (35,)

Ilustración 75: Datos que se utilizarán para entrenamiento y validación.

Ahora generamos las estimaciones dado el entrenamiento ejecutado con las funciones anteriores de la red neuronal artificial.

Estimación de la red neuronal metano sin tratamiento

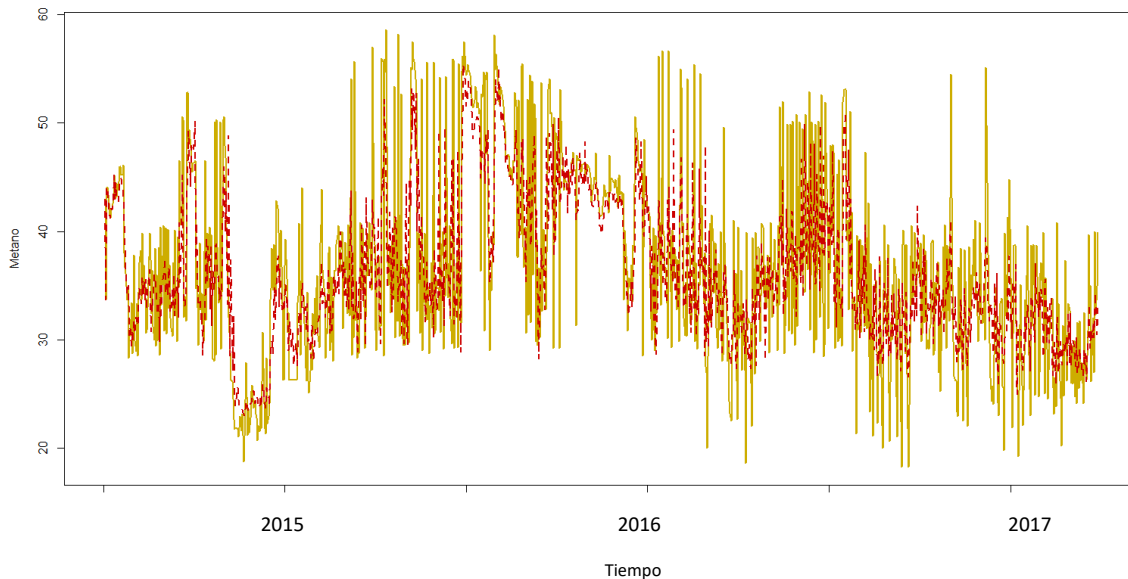


Ilustración 76: Estimación del modelo red neuronal artificial de metano sin tratamiento.

Estimación de la red neuronal metano con tratamiento

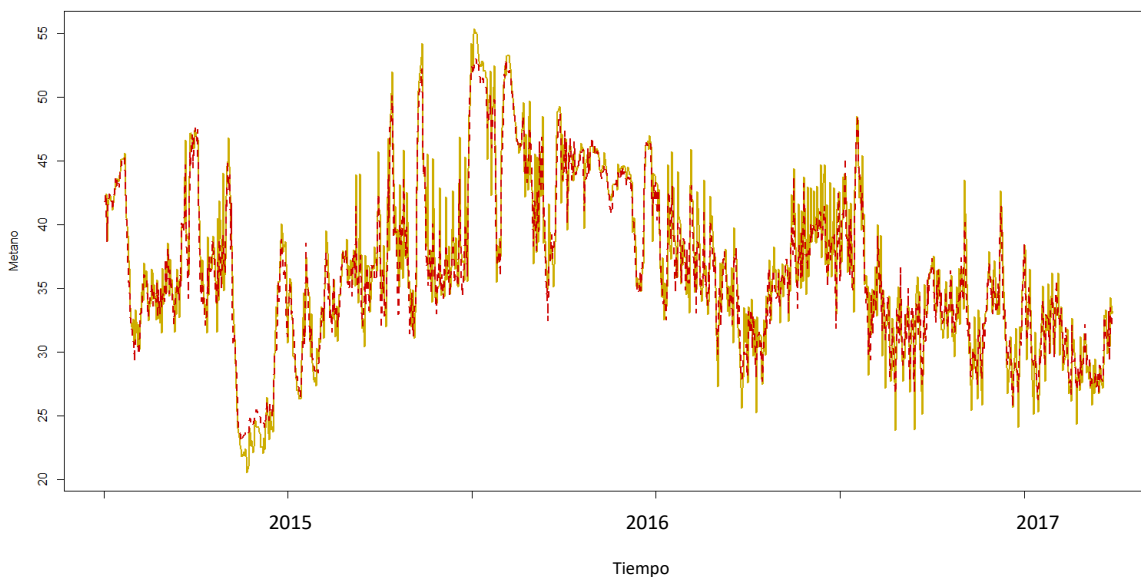


Ilustración 77: Estimación del modelo red neuronal artificial de metano con tratamiento.

4.8.2 Comportamiento de la estimación del modelo red neuronal artificial

Observamos el comportamiento de los residuales durante el periodo de los tres años. Como primer punto partiremos para mostrar la correlación de los residuales, así como la prueba de “Ljung-Box” para la validación del comportamiento de ruido blanco.

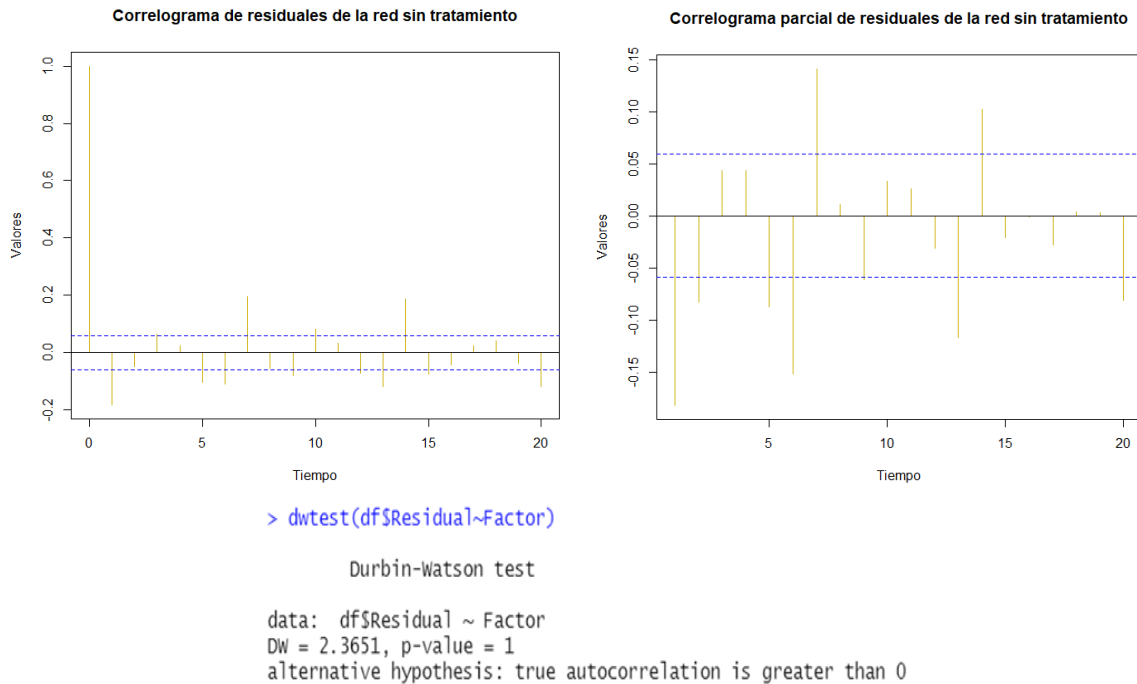
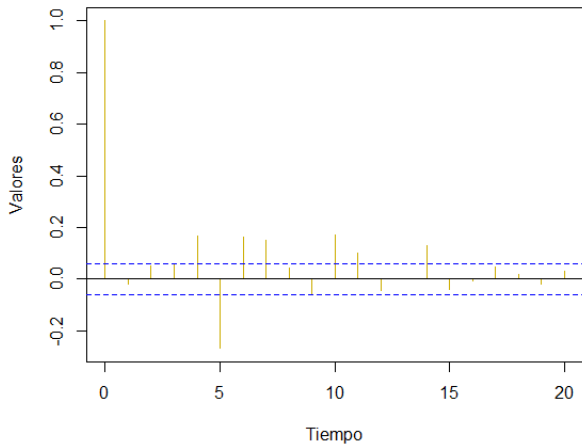
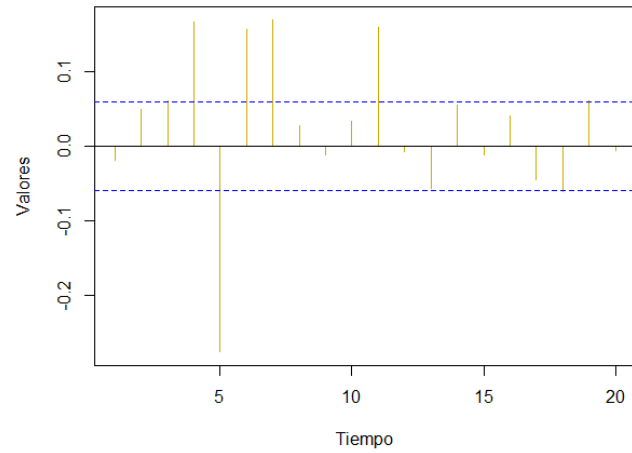


Ilustración 78: Correlación de residuales del modelo red neuronal artificial de metano sin tratamiento.

Correlograma de residuales de la red con tratamiento



Correlograma parcial de residuos de la red con tratamiento



```
> dwtest (df$Residual~Factor)
```

Durbin-Watson test

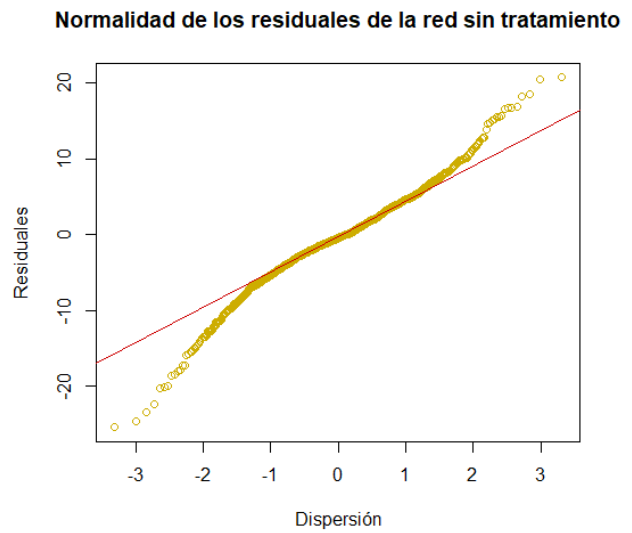
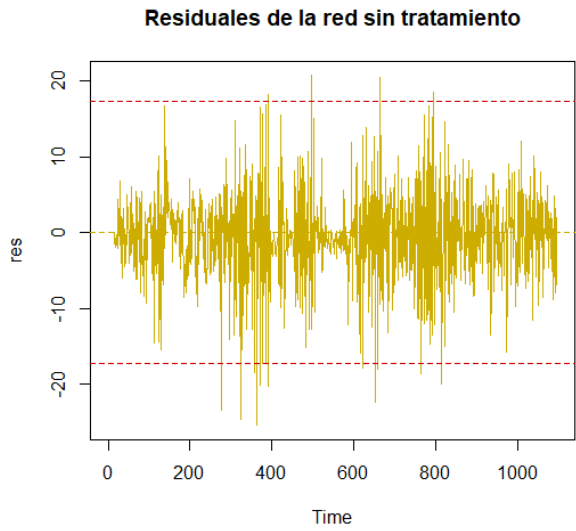
data: df\$Residual ~ Factor

DW = 2.042, p-value = 0.7473

alternative hypothesis: true autocorrelation is greater than 0

Ilustración 79: Correlación de residuales del modelo red neuronal artificial de metano con tratamiento.

Ahora observamos la normalidad de las estimaciones obtenidas de las series.

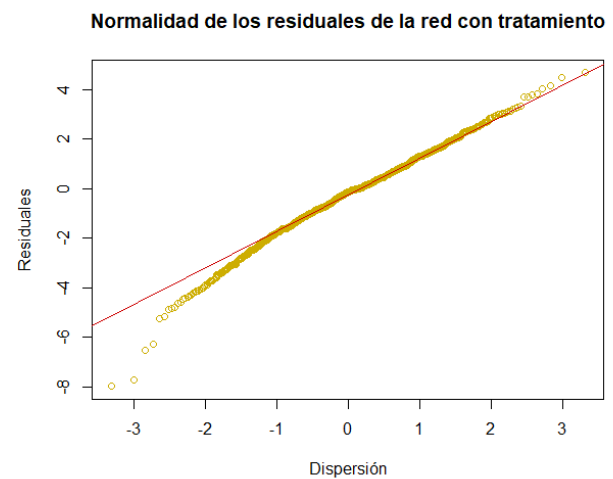
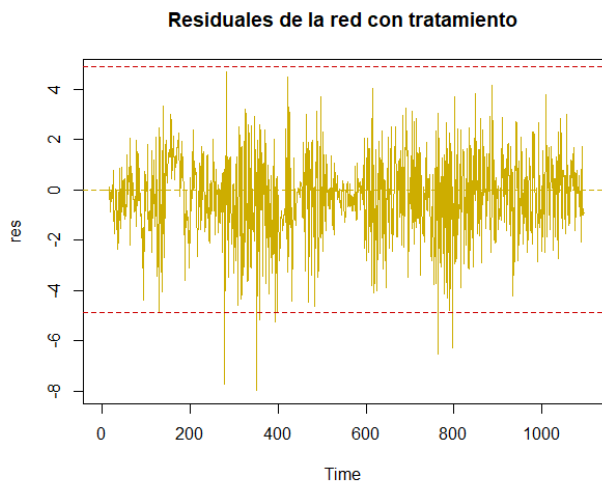


```
> lillie.test(residuales)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuales
D = 0.04173, p-value = 0.0001205
```

Ilustración 80: Normalidad de residuales del modelo red neuronal artificial de metano sin tratamiento.



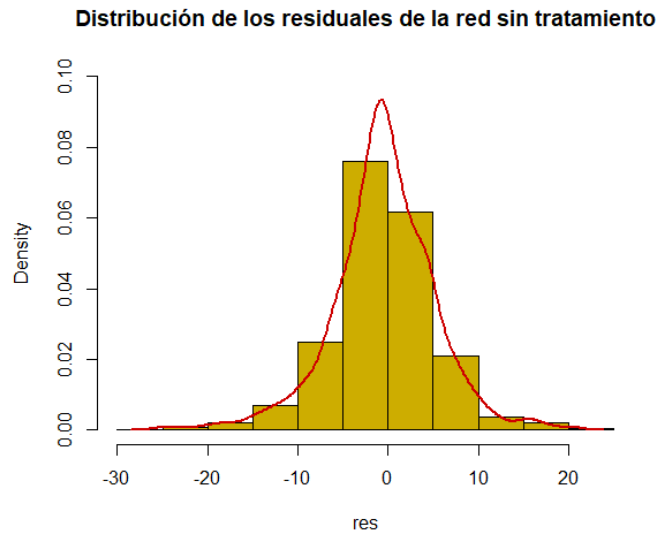
```
> lillie.test(residuales)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuales
D = 0.04173, p-value = 0.0001205
```

Ilustración 81: Normalidad de residuales del modelo red neuronal artificial de metano con tratamiento.

También daremos el comportamiento de la heterocedasticidad y media cero en los residuales.



```
> t.test(res,mu=0,conf.level=0.95)
```

One Sample t-test

```
data: res
t = -2.7265, df = 1095, p-value = 0.006504
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.8182540 -0.1333925
sample estimates:
mean of x
-0.4758233
```

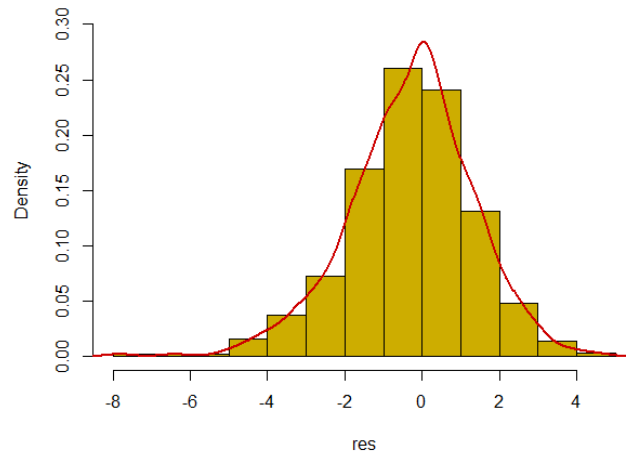
```
> bartlett.test (res,Factor)
```

Bartlett test of homogeneity of variances

```
data: res and Factor
Bartlett's K-squared = 0.10208, df = 1, p-value = 0.7494
```

Ilustración 82: Media cero y prueba de heterocedasticidad en los residuales del modelo red neuronal artificial de metano sin tratamiento.

Distribución de los residuales de la red con tratamiento



```
> t.test(res,mu=0,conf.level=0.95)
```

One Sample t-test

```
data: res
t = -6.0011, df = 1095, p-value = 2.66e-09
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.3921097 -0.1988799
sample estimates:
mean of x
-0.2954948
```

```
> bartlett.test (res,Factor)
```

Bartlett test of homogeneity of variances

```
data: res and Factor
Bartlett's K-squared = 1.2153, df = 1, p-value = 0.2703
```

Ilustración 83: Media cero y prueba de heterocedasticidad en los residuales del modelo red neuronal artificial de metano con tratamiento.

4.8.3 Pronóstico red neuronal artificial

Las estimaciones están a 30 días y se muestran a continuación.

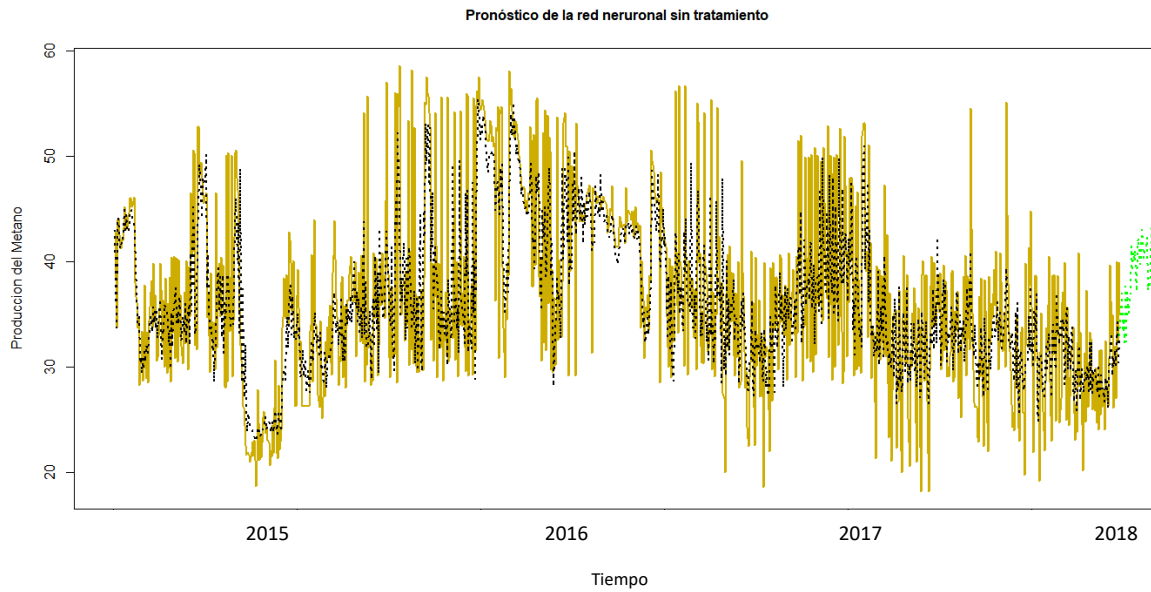


Ilustración 84: Pronóstico del modelo red neuronal artificial de metano sin tratamiento.

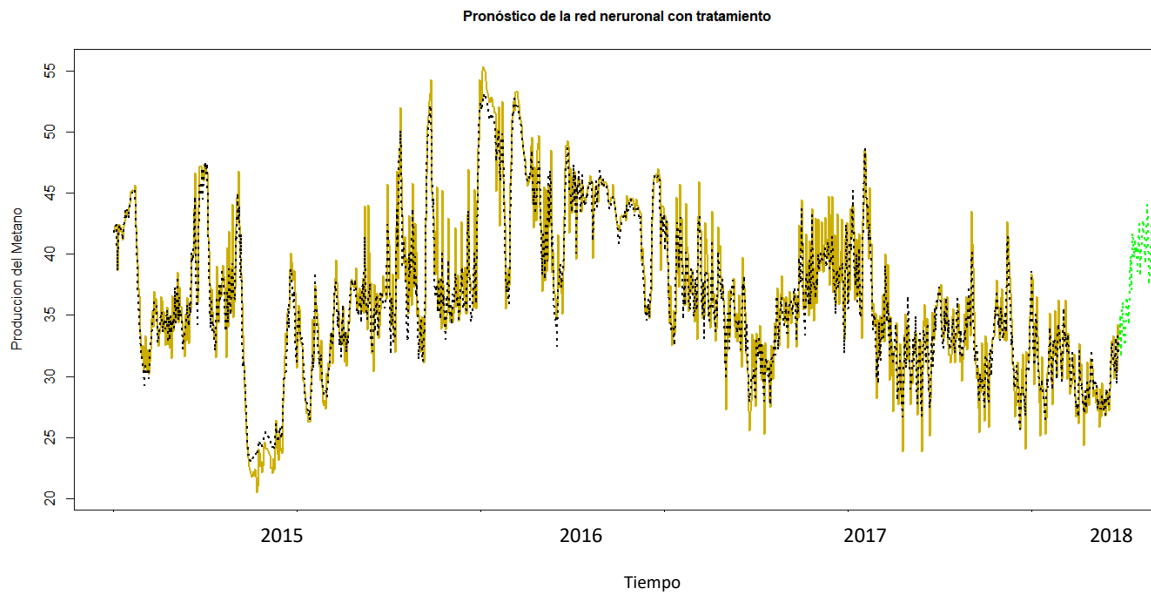


Ilustración 85: Pronóstico del modelo red neuronal artificial de metano con tratamiento.

Los errores de los modelos estimados se muestran en las tablas siguientes.

Tatramiento	MAE	MAPE	R ²
Sin	4.23	11.91	0.5532

Tabla 19: Error del modelo red neuronal artificial de metano sin tratamiento.

Tatramiento	MAE	MAPE	R ²
Con	1.25	3.5	0.9392

Tabla 20: Error del modelo red neuronal artificial de metano con tratamiento.

Capítulo 5 Análisis de los resultados

5.1 Resultado de la elección de la serie de metano

Para poder trabajar con la serie y que nuestros resultados sean lo más homogéneo posible se toman los criterios siguientes:

- Serie con la menor cantidad de datos sustituidos.
- Que el componente irregular de la serie de tiempo sea lo más estable posible. Esto se mide bajo lo observado en los residuales de las diferentes metodologías.
- Serie con el mayor número de datos históricos para entrenamiento y validación.
- Serie estacionaria.

Por lo que, estos criterios son considerados como la premisa principal para una mejor estimación y predicción del metano en el RSU BX. Simultáneamente, como segunda premisa se toma el análisis exploratorio de los datos que se pudieron visualizar en el capítulo 4 y que a su vez fue necesaria con el objetivo de poder dar a entender que aun teniendo diferentes combinaciones no todas las series cumplían con el criterio principal de ser ajustadas ante los valores faltantes de la máquina de extracción forzada de biogás.

Recordemos que, en el diagrama de elección de muestras, se observa el árbol de decisiones para los casos con menos datos sustituidos, permitiendo la elección de las series más significativas y que a su vez se descarten las que menos cumplían los puntos de la homogeneidad esperada.

Para la profundidad del árbol observado en la ilustración 24: Diagrama de elección de muestras, se consideran tres niveles: el primero de ellos son los datos con y sin tratamiento, en segundo nivel los días de la semana con menos datos modificados (viernes, sábado y domingo) y por último la frecuencia periódica establecida. Esto nos ayudó no solo para facilitar la decisión de la muestra a estimar, sino que también busca que los expertos del RSU puedan llegar a tener diferentes criterios para la predicción una vez se apliquen los modelos.

Ahora como resultados de las series mostradas en el capítulo 4 previo a la elección de la muestra, se describen los resultados y la razón por las que no fueron elegidas aun siendo muestras ajustadas con el tratamiento y estacionarias inicialmente.

Periodo de la serie	Serie Analizada		Componentes observados				Razón aceptación/rechazo	Características de la muestra
	Sin	Con	Tendencia	Estacionalidad	Ciclicidad	Irregularidad		
Viernes	X		Decrece en los últimos años	Se repite el periodo en un tiempo aproximado de seis meses		Mayor concentración en los primeros años de la muestra (Fallas de la máquina)	Rechazada al generar más de dos diferencias estacionales, la serie se vuelve ruido blanco.	*157 datos. *Media de 35.6% *Desviación estándar de 7.5%
Viernes		X	Decrece en los últimos años	Se repite el periodo en un tiempo aproximado de cada año		Menor incertidumbre dado el tratamiento aplicado	Rechazada al generar más de dos diferencias estacionales, la serie se vuelve ruido blanco.	*157 datos. *Media de 36.7% *Desviación estándar de 6.8%
Sábado		X	Creciente en el último año	Se repite el periodo en un tiempo aproximado de cada año		Menor incertidumbre dado el tratamiento aplicado, concentración de irregularidad en el año 2016	Rechazada al generar más de 3 diferencias estacionales, la serie se vuelve ruido blanco.	*157 datos. *Media de 36.5% *Desviación estándar de 6.4%
Domingo		X	Decrece en los últimos años	Se repite el periodo en un tiempo aproximado de cada año		Menor incertidumbre dado el tratamiento aplicado, concentración de irregularidad en el año 2015-2016	Rechazada al generar más de 1 diferencias estacionales, la serie se vuelve ruido blanco.	*157 datos. *Media de 36.8% *Desviación estándar de 6.3%
Diario		X	Decrece en los últimos años	Se repite el periodo en un tiempo aproximado en semanas		Menor incertidumbre dado el tratamiento aplicado, mas notorio que las series anteriores	Aceptada por cantidad de datos y al generar diferencias no se vuelve ruido blanco	*1096 datos. *Media de 36.7% *Desviación estándar de 6.7%

Tabla 21: Comparativo de series aceptadas.

5.2 Resultados del modelo promedio móvil ponderado

Recordemos que en el capítulo 2 la metodología de promedio móvil ponderado fue descrita con el objetivo de poder predecir una serie de tiempo, pero esto, no solo tuvo la función de predicción, sino que también fue utilizada para suavizar a la serie, ya que esta tiene datos atípicos. En el capítulo 3 se describió el proceso como un método de suavizamiento mientras que en el capítulo 4 fue descrito como una metodología de predicción, pero en ambos casos los pesos de ponderamiento son los mismos.

En vista de que la metodología fue utilizada para dos propósitos, en la Tabla 22: Resultado de la metodología promedio móvil ponderado, se tendrá la comparación del resultado de cada una de ellas.

Resultado del modelo promedio móvil ponderado	
Metodología para suavizamiento	Metodología para estimación
* Al aplicar el suavizamiento a la muestra, se tiene una mejor distribución del componente irregular durante el tiempo, esto fue comprobado en la comparación de gráficas y en las pruebas de ruido blanco.	* La serie ponderada utilizó un total de siete pesos que asimilan los días de la semana, esto considerando la predicción de un octavo día.
* Favorece la estimación de las metodologías implementadas.	* El error de la serie varía muy poco entre el modelo experto y solver por lo que las estimaciones son muy débiles.
* La metodología nos permitió reemplazar un modelo experto con las herramientas básicas que se manejan en el RSU BX.	El modelo no requirió de diferencias, puesto que parte de un suavizamiento a los datos reales y con ello se aplicarán las metodologías siguientes.

Tabla 22: Resultado de la metodología promedio móvil ponderado.

Dado lo anterior se pudo observar que los errores de este pronóstico comparándolos con los elegidos de forma analítica y los obtenidos con Solver no tienen mucha diferencia, esto debido a la regla de que el dato más reciente se le asigne el mayor peso y vaya bajando gradualmente causando que la metodología sea relativamente tomada a criterio del analista. Es una realidad que este modelo favorece el cálculo del dato más reciente, pero no diferencia los componentes que pueden tener las series de tiempo, de ahí que, se considera a los 7 pesos ponderados no como los óptimos, pero ya que se busca el seguimiento diario, estos pesos llegan a favorecer la predicción de metano a corto tiempo.

5.3 Resultados del modelo Holt-Winters

En relación con la metodología de promedio móvil ponderado implementado a la muestra de metano, partimos en dos a nuestra serie de tiempo, una sin tratamiento, la cual contiene los picos irregulares reemplazados por los días en que la máquina no funcionaba y la segunda muestra que es la serie con tratamiento y de la cual se espera que las predicciones se ajusten de mejor manera.

La metodología de Holt-Winters fue elegida para la estimación en el RSU BX ya que esta se ajusta mejor al componente irregular, y adicionalmente a la estacionalidad constante o

creciente, ahora bien, la metodología nos permitió poder comparar las dos muestras y estimarlas en un lapso de 15 días, esto porque, si la predecimos con mayor tiempo se comporta como una constante. Los resultados de estos dos modelos con la metodología Holt-Winters fue segmentada de la siguiente forma.

La aplicación de las pruebas de raíz unitaria que se implementaron fue dos, el primera prueba de Philips-Perron y la segunda es el aumentado de Dickey – Fuller, puesto que estos dos facilitan saber si la serie requiere diferencias para eliminar el impacto de la tendencia.

Las diferencias aplicadas para la serie con y sin tratamiento fueron auto correlacionadas en la parte regular como a su vez se realizó una diferencia anual en periodos semanas para el impacto de la estacionalidad.

La estacionalidad de la serie, nos indica que esta metodología debe de ser aplicada de forma aditiva, puesto que no tiene una variación creciente en la estacionalidad durante el tiempo observado. Tanto como en la serie con y sin tratamiento podemos ver que las estimaciones se ajustan a la serie correspondientes, esto nos indica que el método aditivo fue seleccionado correctamente, esto se puede observar en las ilustraciones 50 y 51 del capítulo 4.

Partiendo de la estimación de la serie sin tratamiento con una diferencia estacional anual, se obtiene que los datos no llegan a adaptarse para los casos atípicos, mientras que la serie con tratamiento y con la misma diferencia estacional anual llegan a ser más cercana, casi idéntica a la muestra.

La función de correlación y correlación parcial de los residuales ayudan a identificar si llegan a presentar alguna correlación muestral, las dos funciones de correlación de las series ayudaron a observar barras significativas que nos indican que nuestros residuales tienen alguna correlación muestral para ambos casos, considerando realizar más diferencias a las series, pero para una entrega de posibles estimaciones, no fueron realizadas de primera instancia.

Como validación complementaria del modelo estimado es importante verificar como anteriormente se mencionó, que los residuales tanto en las funciones de correlación como en las pruebas tengan un comportamiento de ruido blanco, es por esto, que se utiliza la prueba de Ljung-Box de donde se tiene la prueba de hipótesis siguiente, si $X^2 > 0.1$ se concluye que los residuales son significativamente diferentes de cero, lo que indica que los resultados de las pruebas y gráficos de las dos series con y sin tratamiento no son ruido blanco.

Existe una clase de pruebas, denominados Pormanteau, que consideran las primeras correlaciones de los residuales en forma simultánea para comprobar si el conjunto es diferente de cero. A esta familia de pruebas pertenece el de Ljung Box, cuyo estadístico, Q , se calcula como (Makridakis, Wheelwright, & Hyndman, 1998, pág. 319) $Q = n \sum p^k 2h; k = 1$ en donde n es el número de datos de la serie y h es el máximo lag considerado (comúnmente 20). Si los residuales son ruido blanco, Q se distribuye aproximadamente como X^2 con $(h - m)$ grados

de libertad, en donde m es el número de parámetros involucrados. Si $X^2 h-m > 0.1$, se concluye que los ρ^k no son significativamente diferentes de cero.

La serie de tiempo sin tratamiento de datos tuvo un menor valor en la X^2 determinado que si de este modelo se aplica la estimación tendría mejor comportamiento en la serie con tratamiento de datos.

Una de las validaciones más importantes de los residuales es el comportamiento de la normalidad. Holt-Winters es una metodología que se ajusta a datos cambiantes por lo que estas dos validaciones a los residuales nos permiten tener más claro que el metano al estar frecuentemente monitoreado tenga presente los posibles efectos de irregularidad humana o tecnológica que cada muestra obtenida pueda llegar generar, comparando este criterio, la serie sin tratamiento se encuentra sesgada en su normalidad, mientras que para la serie con tratamiento la normalidad es un tanto más central a una media cero, esto se debe al ajuste realizada en la serie con tratamiento.

Como resultado de las estimaciones y del proceso de Holt-Winters se tiene que la predicción a 15 días llega a ser muy parecida, la diferencia entre cada uno de ellos son los pesos asignados para los componentes de la tendencia, estacionalidad y los niveles cambiantes en el tiempo, así como el error R cuadrada y el error medio absoluto.

Ahora bien, el error R cuadrada nos determina que tanta variabilidad se tendría en las estimaciones siguientes, mientras que el error medio absoluto permite identificar la aceptación del modelo cuando los errores son pequeños. La serie de tiempo con tratamiento llega a estimar con menor error al metano en los próximos 15 días, pero aún sigue siendo un modelo poco estable puesto que los errores observados son muy altos y la variabilidad de nuestros resultados es muy grande.

5.4 Resultados del modelo Box-Jenkins

Esta metodología lleva un proceso explícitamente ordenado debido a que se cumplen ciertos pasos en cadena para llegar a la estimación esperada, en el capítulo 2 se pudo observar que la metodología Box-Jenkins estima modelos ARIMA y a su vez modelos SARIMA, por lo que los resultados para la predicción de metano fueron los siguientes.

Como primera instancia se realizó las pruebas estadísticas para la posible presencia de la tendencia en nuestra serie diaria la cual es muy notoria, la prueba de Dickey-Fuller, el cual nos indica si se requieren generar diferencias al no ser una serie estacionaria, y para ello la prueba es aceptada cuando nuestro (p-value) > 0.05 . Observamos que la prueba de Dickey-Fuller en ambas series no tiene una raíz unitaria significativa, puesto que se realizó una diferencia a la serie sin tratamiento en la parte regular, mientras que se hizo una diferencia con $K=0$ es la parte estacional, de donde el valor K nos indica el orden de la diferencia

realizada en la parte estacional, por otro lado, a la serie con tratamiento se le aplicó una diferencia en su parte regular y una diferencia con orden $K=7$ en su parte estacional, esto debido a que la serie esta considerada con un tratamiento de 7 días para predecir un octavo por lo que nos generó una mejor estimación con esta diferencia realizada a la parte estacional.

Es importante mencionar que los modelos aplicados fueron SARIMA, ya que la estacionalidad de las series con y sin tratamiento aun realizadas las diferencias presentan barras significativas en las funciones de correlación. Comentado lo anterior, como resultado de la aplicación del modelo para cada una de las series la Tabla 23: Resultado de la metodología Box-Jenkins, muestra el resultado de las diferencias realizadas en las dos series.

Box-Jenkins	
Sin Tratamiento	Con Tratamiento
La función de correlación nos ayudó a identificar en los lag 24 y 48 aproximadamente barras significativas que salen del rango límite, por otro lado, y no menos importante observamos en la función de correlación parcial que en su parte regular decrecen y se observa a su vez que en los lag 24 aproximadamente, hay barras significativas.	La función de correlación con tratamiento, se observa un comportamiento distinto, primero en la función de correlación en la parte regular decrece mientras que en la parte estacional hay barras significativas muy notorias, por otro lado, en la función de correlación parcial observamos en su parte regular que se trunca y en la parte estacional alrededor del lag 42 aparece la última barra significativa.

Tabla 23: Resultado de la metodología Box-Jenkins.

Dado lo anterior, el objetivo principal es buscar que los modelos cumplan con la parsimonia, esto en términos coloquiales nos indica no llegar a generar modelos sobreestimados.

- Modelo de metano sin tratamiento SARIMA(2,1,1)(1,1,1)
- Modelo de metano con tratamiento SARIMA(2,1,1)(1,1,2)⁷

En la estimación de los modelos observamos que las series respectivamente con las diferencias aplicadas y sus estimaciones de los modelos SARIMA, llega a comportarse de mejor forma a la serie con tratamiento.

El siguiente paso de la metodología de Box-Jenkins es el comportamiento de los residuales, puestos que estos indican si la estimación seleccionada llega a ser la más adecuada posible, los residuales en la metodología deben ser considerados como ruido blanco gaussiano, es decir, deben distribuirse normalmente con media cero, varianza constante y no deben existir correlaciones significativas entre ellos.

Residuales de metano sin tratamiento: En la primera parte verificamos que los residuales tengan un comportamiento de media cero, por lo que en las gráficas de validación de la media cero en los residuales podemos observar que si bien no se comporta con esta característica la prueba t.test nos indica no hay una media cero. En la gráfica de distribución tenemos un

intervalo significativo con valores atípicos y con mayor cantidad de datos negativos, mientras que a lo largo del tiempo la irregularidad es notoria en los residuales.

Residuales de metano con tratamiento: El comportamiento de la media cero visualizado en las gráficas de los residuales y la prueba t.test nos indica que el resultado del valor de la media es diferente a cero, pero más aceptable que la de residuales sin tratamiento. En el gráfico de distribución a comparación de los residuales sin tratamiento se encuentra con un mejor ajuste a los valores atípicos y más cercana a la media cero, por lo que la serie de residuales en el tiempo no tiene una irregularidad tan notoria como la de los residuales sin tratamiento.

Heterocedasticidad de los residuales en metano sin tratamiento: Como se ha venido explicando es importante que la varianza se comporte constante en los residuales. Para ello se utilizó la prueba de Bartlett y la prueba de Levene, los cuales tiene como parámetros iniciales los residuales y un factor dividido en subgrupos para poder aplicarlos. Esta prueba tiene una hipótesis que si el (p-value) > 0.05 no existe heterocedasticidad por lo que en las gráficas de validación de la media cero en los residuales observados sus p-value no llegan a ser mayores a 0.05 por lo que no existe la heterocedasticidad en nuestros residuales, en otras palabras, la varianza de los residuales es constante.

Heterocedasticidad de los residuales en metano con tratamiento: Considerando lo anterior y replicando el proceso para los residuales con tratamiento, la varianza se debe comportar constante a lo largo del tiempo. Para ello se utilizó de igual forma la prueba de Bartlett y la prueba de Levene. Estas pruebas como fueron mencionados anteriormente llevan como hipótesis que si el (p-value) > 0.05 no existe heterocedasticidad por lo que en las gráficas de validación de la media cero en los residuales observados no llegan a ser mayores a 0.05 en su p-value, y de la misma forma no existe la heterocedasticidad en nuestros residuales con tratamiento, en otras palabras, la varianza de los residuales es constante en él tiempo.

Como siguiente paso vamos a verificar el supuesto de no correlación mutua de los residuales. Este paso lo podemos dividir en dos; uno donde podemos utilizar las funciones de correlación y correlación parcial, donde en ambas gráficas no debemos observar barras significativas y la segunda es una prueba que pertenece a la familia Portmanteau y que de este se elige la prueba Durbin-Watson el cual considera una hipótesis nula en donde la correlación no existe en los residuales durante el tiempo.

En caso de existir correlación entre los residuales, se concluye que el modelo está sobre estimado, es decir, se tienen que agregar parámetros ya sean un modelo de media móvil o de autorregresivo, sin embargo, esta premisa hace que en un momento dado la parsimonia del modelo sea errónea.

Correlación muestral nula en los residuales de metano sin tratamiento: En las gráficas de validación de la correlación de los residuales, tenemos barras significativas en los primeros periodos, sin embargo, hay que mencionar que, del universo de las series ensayadas, esta serie fue la que tuvo la menor correlación en sus residuales. De igual importancia observamos

la prueba Durbin-Watson que comprueba si los residuos se comportan como ruido blanco, si se tiene un estadístico de prueba DW cercano a 2 aceptamos la hipótesis y en nuestro caso el resultado de las correlaciones es significativamente diferentes de cero, por lo tanto, nos indica que hay correlaciones significativas en los residuales de esta serie sin tratamiento.

Correlación muestral nula en los residuales de metano con tratamiento: En las gráficas de validación de la correlación de los residuales, tenemos también barras significativas en los primeros periodos. De igual importancia observamos la prueba Durbin-Watson que obtuvo un resultado de las correlaciones significativamente diferente de cero, por lo tanto, nos indica que también que hay correlaciones significativas en los residuales para ambos modelos.

Normalidad en los residuales de metano sin tratamiento: En la gráfica de periodograma integrado, la prueba de Kolmogorov.S. y Anderson.D. para el supuesto de normalidad, observamos que hay ligeramente valores que salen de los intervalos, sin embargo se considera como aceptable para el modelo, el cual no deberá causar problema en lo absoluto, para la parte estadística se realizaron las pruebas de Kolmogorov-Smirnov y de Anderson-Darling el cual se acepta cuando el valor de significancia (p -value) > 0.05 que indicara que existe normalidad en la serie.

Normalidad en los residuales de metano con tratamiento: En la gráfica de periodograma integrado, la prueba de Kolmogorov.S. y Anderson.D. para el supuesto de normalidad, observamos no hay valores que salgan del intervalo de aceptación, para la parte estadística se realizaron también las pruebas de Kolmogorov-Smirnov y de Anderson-Darling. El p -value de los residuales fue de igual forma menor a 0.05 pero no tan alejados como los residuales sin tratamiento lo cual nos da como resultado que la normalidad sea más aceptable, aunque no llega a cumplirla para la serie con tratamiento aplicado.

Identificación de valores atípicos en los residuales de metano sin tratamiento: Después de analizar a los residuales para que estos cumplan con ser ruido blanco gaussiano se valida que no existan valores atípicos en el intervalo de tiempo analizado, para ello en la gráfica de residuales sobre el tiempo, se obtienen tres veces la desviación estándar y estas se grafican como líneas punteadas de donde los valores fuera de ese intervalo se consideran como valores atípicos. La serie presenta esta irregularidad en los residuales y nos indica que la serie sin tratamiento debe ser ajustada para su estimación, lo cual para mejoramientos futuros el RSU serian considerados y aplicados.

Identificación de valores atípicos en los residuales de metano con tratamiento: Para esta serie también se valida que no existan valores atípicos en el intervalo de tiempo analizado, para ello en la gráfica de residuales sobre el tiempo, se obtiene tres veces la desviación estándar y estas se grafican como líneas punteadas de donde los valores fuera de ese intervalo calculado se consideran como valores atípicos, por lo que la serie presenta una menor irregularidad en los residuales a comparación de la serie sin tratamiento, sin embargo existen valores que requieren ajustarse y de la misma manera serán considerados para siguientes mejoras de los modelos

Ahora bien, una vez analizados los residuales de los modelos SARIMA se genera la validación de la parsimonia con los pesos AR, MA, SAR y SMA. En la serie sin tratamiento la parsimonia para los pesos solo llega a ser menor para la parte SAR, mientras que para la serie con tratamiento los pesos no tienen significancia por lo que podrían no ser necesarios en caso de querer estimar mejor el modelo, podríamos no tomar en cuenta a la parte invertible de la serie con tratamiento para futuras estimaciones.

Por último, realizaremos la validación de admisibilidad que en otras palabras cumplan con ser estacionarias e invertibles para los modelos autorregresivos y de medias móviles. Este punto resulta ser de interés ya que recordemos que un modelo MA por construcción a partir de un proceso lineal asegura que es estacionario sin embargo debe comprobarse que sea invertible. Por otro lado, un modelo AR donde sus parámetros garantizan que dicho modelo es invertible, pero al igual se debe comprobar la parte estacionaria, es importante aclarar que esto se aplica al tener la combinación de un modelo ARMA. Las ecuaciones de admisibilidad aplicadas fueron favorables para ambos modelos cumpliendo con ser invertibles y estacionarias, esta validación no se expone en el capítulo 4 pero es aceptable en su comprobación.

Una vez ya realizados y verificados todos estos pasos vamos a describir el resultado de nuestro pronóstico que, en este caso, será a una predicción de 30 días. En el pronóstico metano sin tratamiento se considera como valor constante de predicción después de los primeros 15 días ya que se vuelve una línea recta. Por otro lado, el error R cuadrado nos indica para la serie sin tratamiento tiene una mayor variabilidad en las estimaciones generadas y un error medio absoluto grande en las estimaciones generadas, mientras que, para metano con tratamiento, aunque después de los primeros 15 días esta se vuelva una línea constante el R cuadrado es significativamente mejor, puesto que la variabilidad es menor para los valores esperados del modelo teniendo una mejor aceptación de comportamiento en un corto tiempo y bajando el error medio absoluto que la serie sin tratamiento.

5.5 Resultados del modelo red neuronal artificial

Como se sabe una red neuronal artificial está dada por una estructura de entradas y salidas de donde existe un método que da la activación para la salida de la red. El resultado de nuestra red neuronal artificial para la estimación de metano en el RSU BX inicia considerando que nuestra serie fue escalonada entre el rango -1 a 1, esto mediante las funciones en *python* para metodologías de *machine learning*, lo que permitió elegir la función tangente hiperbólica, puesto que esta es recomendada para redes neuronales que buscan estimaciones a series de tiempo. Es importante aclarar que los métodos de activación dependen del resultado esperado. No olvidar que para poder generar una estimación con una red neuronal artificial esta debe cumplir con ser estacionaria y distribuirse de forma normal, por lo que los criterios

explicados en capítulo 3 dan como aceptada que las series de metano con y sin tratamiento cumplen con estos dos puntos críticos.

Como recordatorio del objetivo en los resultados de esta tesis, el generar una red neuronal artificial debe permitir una estimación más precisa que las metodologías anteriores, favoreciendo las decisiones para la regulación de la máquina de extracción en el RSU BX.

Ahora bien, considerando que la función de escalamiento y método de activación de salida de la red neuronal artificial fue la tangente hiperbólica, la series fueron generadas en un *dataframe*, donde se buscó que en la muestra de entrenamiento tuviera 16 días, para pronosticar un 17 y para ello la función *series_to_supervised* facilitó la generación de esta muestra tanto para la serie con y sin tratamiento, ya que se buscó generar una comparación exclusiva como resultado de la red neuronal artificial. La función *serie_to_supervised* toma primero la muestra y la transforma en un *dataframe*, de donde dependerá de las variables de entrada para considerar la forma de recorrer una fila hacia delante o hacia atrás. Los datos de entrada fueron los valores escalonados de -1 a 1 de metano con y sin tratamiento. Se eligió que se recorra de forma hacia delante, por lo que los datos de entrada serán dieciséis días, esto fue considerado ya que el comportamiento del dato t-1 llegará a ser el que genera un error cuadrático medio menor durante las épocas de asignación de pesos que la red neuronal artificial realiza. Una vez esto, se generó la muestra de entrenamiento y la muestra de validación para la red neuronal artificial, se tuvo un total de 1,060 datos de entrenamiento y 35 de validación, facilitando el resultado esperado de las dos series estimadas.

Por último, se considera en la creación de la red neuronal artificial los siguientes puntos; 1) Se generan un total de 16 nodos en una sola capa oculta, aunque este no fue el óptimo, se tuvo el menor error con este número de nodos colocados; 2) Se tienen 16 nodos de entrada resultantes de las muestras de entrenamiento y validación; 3) El método para el ajuste de pesos es “relu” considerado con un buen proceso para el ajuste en cada proceso del entrenamiento; 4) Solo una variable de salida con el método de activación tangente hiperbólica; 5) El error cuadrático medio será el que se evalúa en la ejecución de la red pero la R cuadrada y el error medio absoluto serán las que se comparen con los diferentes modelos.

Los resultados de las series con y sin tratamiento de la red neuronal artificial fueron las siguientes.

Red neuronal artificial	
Sin Tratamiento	Con Tratamiento
En la estimación de la serie sin tratamiento, podemos observar que esta tiene una mejor adaptación ante los picos atípicos durante algunos intervalos de tiempo, no es idéntica a la serie pero si está mejor adaptada que los modelos anteriores, puesto que este sin la necesidad de generar diferencias se ajusta de mejor forma.	En la estimación de la serie con tratamiento, al igual que la serie anterior esta es muy cercana a los datos sin necesidad de generar diferencias, pero con mejor comportamiento que la serie sin tratamiento estimada con la red neuronal, esto mejora en muchas escalas las estimaciones esperadas con esta metodología y facilita tener una mejor herramienta para modelos en el RSU BX.

Tabla 24: Resultado de la metodología red neuronal artificial.

Como resultado de la estimación de las series, se generó la validación de los residuales que de la misma forma a los métodos anteriores se buscó que tengan un comportamiento de no correlación muestral, normalidad y media cero.

Correlación de residuales metanos sin tratamiento de la red neuronal artificial: La correlación de residuales de la serie sin tratamiento y la prueba de “Ljung-Box”, se puede observar que la serie requiere de diferencias para poder tener quizá una menor correlación tanto en la estacionariedad como en la invertibilidad ya que en ambas gráficas se visualizan barras significativas, también en caso de no generar diferencias, se puede llegar a generar más capas ocultas que permitan tener una mejor precisión, por otro lado, la prueba aplicada permite validar esta correlación de residuales al tener un $p_value < 0.1$.

Correlación de residuales metanos con tratamiento de la red neuronal artificial: La correlación de residuales de la serie con tratamiento y la prueba de “Ljung-Box”, se puede observar que la serie aun teniendo barras significativas estas no indican correlación en los residuales por lo que no se requiere generar diferencias ni reajustar el modelo de primera instancia, teniendo un $p_value > 0.1$ rechazando la hipótesis nula de correlación diferente de cero.

Normalidad de los residuales metanos sin tratamiento de la red neuronal artificial: Ahora se analiza la normalidad de los residuales de la red sin tratamiento, se valida que no llega a tener una distribución normal y mediante la prueba de Kolmogorov-Smirnov se comprueba que el $p_value < 0.05$.

Normalidad de los residuales metanos con tratamiento de la red neuronal artificial: Por otro lado, en la normalidad de los residuales de la red con tratamiento, llega a tener una mejor distribución, pero al generar la prueba de Kolmogorov-S el p_value sigue siendo menor a 0.05, pero con una menor diferencia que los residuales de la serie sin tratamiento.

Comportamiento de media cero en los residuales metano sin tratamiento de la red neuronal artificial: Por último, para la media cero y la prueba de heterocedasticidad en los residuales de la red sin tratamiento, se generaron dos pruebas de validación, el primero de ellos es para el cálculo de la media cero en donde se obtuvo un valor cercano a cero, por otro lado la segunda prueba permite saber si existe heterocedasticidad por lo que el $p_value > 0.05$ se acepta que los residuales no tienen este comportamiento y al final el diagrama de distribución permite ver que existe un sesgo de lado izquierdo teniendo valores de -30 y no centrándose a la media cero.

Comportamiento de media cero en los residuales metano con tratamiento de la red neuronal artificial: Considerablemente en la media cero y la prueba de heterocedasticidad en los residuales de la red con tratamiento, se generaron las mismas validaciones y la misma gráfica de distribución del cual para el primero la media observada es de -0.2954 muy cercano a cero e idéntica a la de sin tratamiento, en la prueba de heterocedasticidad tiene un $p_value > 0.05$ que de igual forma se acepta que no hay heterocedasticidad en los residuales, y ya como última validación, la gráfica de distribución no presenta sesgos como en los residuales de la

serie sin tratamiento acercándose más al comportamiento de la media cero esperado pero no aceptado para esta serie.

La red neuronal artificial llega a ser predecible hasta por un lapso de 30 días mejorando las estimaciones de los modelos anteriores. El pronóstico de las dos series se puede observar en las ilustraciones 84 y 85 en las cuales observamos una mejoría en la serie de metano con tratamiento vs los modelos ya probados y debido a la naturaleza de su proceso resulta ser más eficiente estimar a esta serie. La R cuadrada es más cercana a 1 en la serie con tratamiento y representa la mejor aceptación dada la variabilidad de estimación en los 30 días, a su vez el error cuadrático absoluto es el más pequeño de todos los modelos aplicados.

5.6 Resultados generales de las metodologías

La implementación de las cuatro metodologías de predicción se considera favorable para el análisis implementado con el fin de tener un procedimiento factible y adecuado para los expertos en el RSU BX. La exploración espectral fue fundamental para poder obtener los resultados esperados, sin embargo, como se ha venido mencionando, los datos no fueron en su generalidad fáciles de interpretar y como a su vez en las metodologías los errores de predicción fueron elevados.

Se obtiene que la metodología generada con promedio móvil ponderado permitió que la serie fuera suavizada favorablemente y con esto poder obtener resultados interpretables de las metodologías Box-Jenkins, Holt-Winters y red neuronal artificial, pero a sí misma la metodología no facilita un pronóstico ya que solo los datos fueron históricos en el mismo periodo de tiempo. Si la obtención de los datos hubiera sido continua, se pensaría que el tratamiento no hubiera sido necesario, pero en este caso la serie tuvo datos faltantes, por lo que sí requirió un ajuste en el componente irregular de la serie de tiempo, sin embargo, la interpretación de una manera más natural (sin ajuste de datos) no fue posible y los resultados en las metodologías de Holt-Winters, Box-Jenkins y red neuronal artificial con la metodología de promedio móvil ponderado resultaron ser favorables para la posible tomas de decisiones.

Por otro lado, la metodología de Holt-Winters tenía como antecedente teórico aceptar datos cambiantes y el ajuste al componente estacional en las series de tiempo. Para el caso de metano no fue posible tener este comportamiento ya que, aun generando diferencias de correlación en la parte estacional y regular, esta tiene un bajo nivel de aceptación al tener una R cuadrada muy bajo con posible alto nivel de variabilidad en las estimaciones futuras, Por lo que la metodología a corto plazo de promedio móvil ponderado sería mejor que la estimación de Holt-Winters, considerando a las dos series obtenidas (con y sin tratamiento). De la misma forma, para el componente estacional se eligió un método aditivo debido a que se observa un comportamiento constante en las tendencias de la serie, ya que ayudó a tener un ajuste más significativo en la estimación de metano sin tratamiento.

Ahora bien, para la metodología de Box-Jenkins se tiene presente la predicción de modelos SARIMA, ya que estos permitieron el suavizamiento del componente estacional observado en cada serie obtenida. Mencionando lo anterior, se concluye que las series requieren diferencias tanto para el ajuste de la tendencia y cambio de nivel, como a su vez en la parte estacional repetida semanal y anualmente. Las diferencias generadas fueron necesarias, pero no suficientes, por lo que se le realizó una diferencia regular y estacional.

Por otro lado, la parte de los residuales tanto en la serie con y sin tratamiento no llegaron a cumplir el comportamiento de ruido blanco gaussiano que solicita la metodología de Box-Jenkins, es importante mencionar que los criterios de media cero, heterocedasticidad, normalidad y la no correlación de los de residuales, no lleguen a cumplir las pruebas requeridas, y por ello se tenga que generar aún más diferencias a las series ya diferenciadas teniendo como mayor problema llegar a ser ruido blanco. Se tiene en consideración a la parsimonia, ya que esta no fue sobreestimada favoreciendo la predicción y el error de estimación R cuadrada, ya que los valores de los errores fueron superiores a cero teniendo un mejor margen de aceptación que las metodologías de promedio móvil ponderado y Holt-Winters.

Por el contrario, en la red neuronal artificial es importante considerar tanto el proceso para la creación de la red, así como los parámetros de entrenamiento y posteriormente el comportamiento de los residuales en las dos series obtenidas, ya que estas tuvieron diferente peso en la hora de estimación, así como en los errores resultantes.

La función de activación que utilizamos facilitó el tratamiento de los datos, ya que se generó un ajuste en la escala de las series de metano, una vez escalonados, la agrupación de los datos de entrada para el entrenamiento que fueron en intervalos de 16 días tuvo mayor efecto que las metodologías anteriores. Las dos estimaciones resultaron ser adecuadas debido a que se ajustan en ambas series, ya que se comportaron de mejor manera que las metodologías tradicionales, definiendo a la red neuronal artificial como un buen modelo para estimaciones de metano en el RSU BX.

Sin embargo, en sus residuales observamos barras significativas en sus funciones de correlación y correlación parcial, pero de alguna manera menos relevantes a comparación de los residuales de Box-Jenkins, Holt-Winters y promedio móvil ponderado. La normalidad y la media cero de los residuales se diferencia por el sesgo de la muestra distribuida normalmente, favoreciendo al comportamiento a metano con tratamiento. Adicionalmente las estimaciones fueron mayores a 30 días, teniendo la metodología adecuada posible a implementar en el RSU BX. Las R cuadradas tuvieron como menor valor al que se generó con la mejor estimación en Box-Jenkins y, mientras que en metano con tratamiento obtuvo el valor más alejado a cero en la R cuadrada. Por lo que, resulta útil aplicar una metodología que pronostique intervalos largos y tengan una menor variabilidad en los pronósticos futuros, aunque como sugerencias se podría generar un proceso que optimice los pesos de la red teniendo un mejor ajuste en el modelo.

Si se considera como dato esperado la predicción de metano dentro de la composición del biogás, se piensa que puede tener una mejor utilidad ante porcentajes del 45% al 65%, por lo tanto se considera que la red neuronal artificial fue la que más aceptación tiene para la toma de decisiones en el RSU BX, ya que esta red predice un máximo del 44% de metano durante los próximos 30 días, teniendo al biogás por debajo del rango de utilidad pero con la ausencia de valores faltantes en la toma de la muestra que la máquina de extracción obtiene. Este comportamiento puede verse afectado por la irregularidad de predicción ante los modelos implementados.

Un primer factor puede ser la estacionalidad de los datos a nivel de operación debido a situaciones climatológicas o en épocas de mayor temperatura por decir algunos de los factores impredecibles, que pueden ser causantes de esas irregularidades estacionales.

Un segundo factor sería que metano con tratamiento para el método Holt-Winters fuera irregular, ya que el reemplazo de los datos faltantes hace que la serie sea poco estable.

Como tercer factor a la metodología de Box-Jenkins le causa que el reemplazo de los datos, en los residuales, lleguen a ser todos rechazados como ruido blanco gaussiano.

Si el efecto de los valores atípicos tuviera una regularidad en la muestra obtenida de metano a lo largo del tiempo, posiblemente nos permitan generar y seleccionar uno o varios modelos, y los comportamientos de estos quizá favorezcan algún método.



Ilustración 86: Equivalencia de energías dado el porcentaje de metano.

La toma de las decisiones considerada por nosotros está segmentada por dos puntos, ya que al tener entre un 60% de producción de metano en la composición del biogás esta puede ser transformada a un tipo de energía como se muestra en la ilustración 86: Equivalente de energías dado el porcentaje de metano.

Comportamiento de nuestro modelo: Con los datos obtenidos y el modelo diseñado con red neuronal artificial se puede considerar como decisión del experto que al tener la posible producción de metano estos puedan realizar ajustes en el monitoreo o en la irregularidad de la máquina, puesto que, si en la extracción llegara a fallar la máquina, como regularmente sucede, esta predicción puede llegar a ser un valor de reemplazo en el dato no obtenido.

- Porcentaje de metano en la estimación de la red neuronal artificial: Debido al resultado obtenido en la red podríamos mencionar la decisión de poder llegar a un porcentaje de energía un tanto bajo (máximo 44% obtenido de metano en los próximos 30 días) pero de igual manera estable y aprovechable. Sin embargo, los datos predichos pudieran llegar a dar un indicador acumulado de producción de metano.

El metano al no superar el 50% de la composición general del biogás en la estimación de los próximos 30 días es considerado como un gas de nivel bajo. Se espera que se tenga a su vez una réplica para algún otro dato, como la temperatura o concentración de CO₂, así como la producción de lixiviados, ya que las metodologías pudieran llegar a ser utilizadas para diferentes comportamientos dentro del RSU BX.

Conclusiones

Para concluir, la propuesta considerada como tratamiento para la serie de tiempo fue promedio móvil ponderado ya que este se ajustó a tal grado que reduce el componente irregular y atípico de la muestra de metano, para mejorar las validaciones y estimaciones de los modelos creados. La propuesta considerada favoreció las estimaciones a corto plazo. Sin embargo, fue importante contrastar los modelos con o sin el tratamiento, debido a que el resultado forma parte del proceso de elección del pronóstico que el experto deberá tomar.

En relación con el uso de estas metodologías el aplicar los modelos en sitios relacionados a un RSU u otros, son considerados alcanzables puesto que estas predicen a mediano y corto plazo ya que se ajustan a los componentes de una serie de tiempo serán mediante la creación de programas que automaticen las aplicaciones necesarias para cada objetivo buscado. Como resultado de esto, las muestras obtenidas que incluyen composición del biogás, flujos y temperaturas son predecibles si se desarrollan mediante un modelo univariado o multivariado de cada variable que la muestra tenga.

Las metodologías aplicadas en esta tesis fueron evaluadas y comparadas mediante los residuales obtenidos y sus estimaciones a corto o mediano plazo, ya que se busca que se comporten como ruido blanco y además ser aceptadas mediante los valores de R cuadrada y el error medio absoluto, por lo que, las estimaciones de las redes neuronales artificiales y Box-Jenkins tuvieron mejor aceptación para estimar al metano en un máximo de 30 días.

El factor principal para aceptar los objetivos esperados fue el analizar los datos mediante análisis espectrales y componentes de una serie de tiempo. Con esto nos referimos a que aun teniendo una variabilidad y periodicidad de datos donde estos, a la hora de la extracción no fueron correctos por el fallo de la máquina, permitieron generar modelos de predicción favorables para la toma de decisiones del RSU BX.

Los residuales fueron en gran parte el problema de aceptación en las pruebas requeridas de la verificación necesarias de todos los modelos. Los valores atípicos obtenidos por la máquina de extracción de biogás en el RSU BX tuvo un peso significativo de predicción en la metodología de todos los modelos que no fueron posibles ajustar en 100%.

Por último, el mejor método de predicción es la red neuronal artificial con tratamiento, con esta conclusión se plantea generar un proceso de optimización para posibles ejecuciones futuras como uso en el RSU BX. Las herramientas de ejecución para el tratamiento y proceso de estimación son generalmente lenguajes de programación de software libre, tales como Rstudio y Python y de otro tipo como Excel, por lo que los procesos finales deban ser ejecutables en herramientas prácticas como son las macros las cuales se tienen en el sitio de monitoreo del RSU BX, ya que estas permitirán generar presentaciones para decisiones rápidas. Adicional a la implementación, los modelos de estimación generados pueden ser utilizados para algunos otros componentes o patrones que se busquen para el mejoramiento del sitio.

La interpretación de los modelos para una conclusión que facilite los procesos podría ser mensuales, ya que se tiene un buen rango de aceptación con la red neuronal artificial. Para el caso de aceptación de los modelos es que lleguen a tener una R cuadrada lo más alejada de cero se tendrá presente que la variabilidad de los resultados futuros de la muestra analizada no será tan alejada de los datos históricos ya que la maquina aun si falla podría considerar un menor impacto en el reemplazo del valor por la estimación calculada.

Bibliografía

- Ahumada de Toledo, F. (2014). *Centro Deportivo de Alto Rendimiento, Ciudad Jardín Bicentenario, Nezahualcóyotl*. México: Universidad Nacional Autónoma de México.
- Ambientech. (01 de 09 de 2020). <https://ambientech.org/>. Obtenido de <https://ambientech.org/la-neurona>
- Braga, L. (2019). *CONICET MENDOZA*. Obtenido de <https://www.mendoza.conicet.gov.ar/portal/enciclopedia/terminos/DioxiCar.htm>
- Bragado Alcaraz, R. (04 de 11 de 2018). *Tiloom.com*. Obtenido de <https://www.tiloom.com/mineralizacion-de-materia-organica/>
- Cortés Patiño, J. (2011). *Repositorio Digital Facultad de Ingeniería, UNAM*. Obtenido de <http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/363/A7.pdf>
- de la Fuente Fernández, S. (2013). *MODELO ARIMA(p,d,q)(P,D,Q)s*. Madrid España: Universidad Autónoma de Madrid.
- Decología. (18 de 08 de 2019). *Decología.com*. Obtenido de <https://decologia.info/medio-ambiente/vertederos/>
- Dirección General de Bibliotecas, U. (2014). *Ciudad Jardín Bicentenario*. CDMX: Dirección General de Bibliotecas de la UNAM.
- FisioOnline. (2017). *FisioOnline*. Obtenido de <https://www.fisioterapia-online.com/glosario/axon-o-neurita>
- Hidrocarburos. (03 de 12 de 2020). *Significados.com*. Obtenido de <https://www.significados.com/hidrocarburos/>
- López Piña, J., & Quintero Soto, L. (10 de 04 de 2009). *Revista unam.mx*. Obtenido de *Revista Digital Universitaria*: <http://www.revista.unam.mx/vol.10/num4/art19/int19.htm>
- Maquinaria, A. (08 de 01 de 2016). *Aristegui*. Obtenido de <https://www.aristegui.info/los-lixiviados-caracteristicas-y-actuaciones/>
- MedlinePlus. (1 de 10 de 2019). *MedlinePlus Información de salud para usted*. Obtenido de <https://medlineplus.gov/spanish/ency/article/002230.htm>
- Montenegro, A. (2009). *Análisis Espectral*. Departamento de Economía. Bogotá: Facultad de Ciencias Económicas y Administrativas.

- Ruiz Hernández, J., Barrios Puente, G., & Gómez Gómez, A. (03 de 2019). *Scielo.org*. Obtenido de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342019000200225
- Sandra. (22 de 10 de 2019). *ATRIA INNOVATION*. Obtenido de <https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/>
- SmartPanel. (25 de 02 de 2019). *SmartPanel*. Obtenido de <https://www.smartpanel.com/que-es-deep-learning/>
- Triglia, A. (10 de 12 de 2018). *Psicología y Mente*. Obtenido de ¿Qué son las dendritas de las neuronas?: <https://psicologiaymente.com/neurociencias/dendritas>
- Universal, E. L. (11 de 11 de 2020). *Enciclopedia Libre Universal*,. Obtenido de https://es.wikipedia.org/wiki/Organismo_aerobio
- Varnero Moreno, M. (2011). *MANUAL DEL BIOGÁS*. Santiago Chile: FAO.
- Wikidot. (2015). *Wikidot.com*. Obtenido de <http://redes-neuronales.wikidot.com/definicion-ventajas-desventajas>
- WordPress. (02 de 02 de 2014). *WordPress.com*. Obtenido de <https://rellenosanitarios.wordpress.com/>
- Reyes, J. A. (23 de abril de 2004). <http://www.paot.org.mx>. Obtenido de http://www.paot.org.mx/contenidos/paot_docs/pdf/basura_df.pdf
- GUTIÉRREZ, G. B. (2 de Diciembre de 2013). Academia. Obtenido de [www.academia.edu:http://www.academia.edu/12429445/INVESTIGACION%3A%93N_RELLENO_SANITARIO_BORDO_PONIENTE_EC%3A%93POLIS](http://www.academia.edu/12429445/INVESTIGACION%3A%93N_RELLENO_SANITARIO_BORDO_PONIENTE_EC%3A%93POLIS)
- Jaramillo, J. (2002). GUÍA PARA EL DISEÑO, CONSTRUCCIÓN Y OPERACIÓN DE RELLENOS SANITARIOS MANUALES. Colombia: OPS/CEPIS.
- Cruzito. (1 de 1 de 2014). Energía Solar Hoy. Obtenido de Energía Solar Hoy: <https://energiasolarhoy.com/biogas-ventajas-y-desventajas/>
- ENERGÌA, E. (2020). eco-gmbh. Recuperado el 20 de 01 de 2020, de <https://www.eco-gmbh.eu/es/temas/biogas.html>
- Villavicencio, J. (s.f.). Introducción a Series de Tiempo. En J. Villavicencio, Introducción a Series de Tiempo (pág. 33).
- FUENTE, Y. T. (2015). "METODOLOGÍA DE INVENTARIOS Y PRONÓSTICOS PARA DISTRIBUIDORES DE PRODUCTOS DE TELECOMUNICACIÓN.". CMDX: UNIVERSIDAD NACIONAL AUTÓNOMA.

- Villaverde, N. R. (7 de junio de 2017). rpubs. Obtenido de rpubs.com:
<https://rpubs.com/nanrosvil/283121>
- Rojas, C. (2019). ACADEMIA. Obtenido de ACADEMIA:
https://www.academia.edu/19489098/MODELO_HOLT_WINTERS_MULTIP_LICATIVO
- Guerrero, V. M. (2003). Análisis estadístico de series de tiempo económicas. México: Thomson.
- Rosales, R. (2018). Universidad de los Andes Facultad de Economía. Obtenido de Universidad de los Andes Facultad de Economía:
https://economia.uniandes.edu.co/files/profesores/ramon_rosales_alvarez/docs/econometria2/Salidas%20y%20Ejercicios/EJC202220Metodologa20Box20-20Jenkins.pdf
- Crispi, J. (12 de junio de 2019). medium.com. Obtenido de
<https://medium.com/@jcrispis56/una-introducci%C3%B3n-completa-a-redes-neuronales-con-python-y-tensorflow-2-0-b7f20bcfeb5>
- Calvo, D. (07 de Diciembre de 2018). Diego Calvo . Obtenido de Clasificación de sonido con Redes Neuronales Convolucionales:
<https://www.diegocalvo.es/author/diegocalvo/>
- Cohen, I. R. (s.f.). Gestión de Recursos – FI UBA:. Obtenido de Pronósticos:
<http://materias.fi.uba.ar/7628/PronosticosTexto.pdf>
- López, B. S. (30 de junio de 2019). <https://www.ingenieriaindustrialonline.com/>. Obtenido de Promedio móvil ponderado:
<https://www.ingenieriaindustrialonline.com/pronostico-de-la-demanda/promedio-movil-ponderado/>
- OLAYA, P. C. (Diciembre de 2003). ESTUDIO DEL PRONÓSTICO DE LA DEMANDA DE ENERGIA ELECTRICA,. Obtenido de Scientia et Technica Año IX: file:///C:/Users/aantonio/Downloads/Dialnet-ESTUDIODELPRONOSTICODELADEMANDADEENERGIAELECTRICAU-4845141.pdf
- Wei, W. W. (2006). Time Series Analysis. Univariate and Multivariate methods. U.S.A.: Addison Wesley.
- Acurio, G., A. Rossin, P.F. Texeira y F. Zepeda. Diagnóstico de la Situación del Manejo de Residuos Sólidos Municipales en América Latina y el Caribe. BID, Organización Panamericana. Washington, D.C. 1997.
- LAURA, Yauli; PIEDAD, Ana. Manual para el Manejo de Desechos Sólidos en la Unidad Educativa Darío Guevara, Parroquia Cunchibamba, Cantón Ambato, Provincia de Tungurahua. 2012. Tesis de Licenciatura.