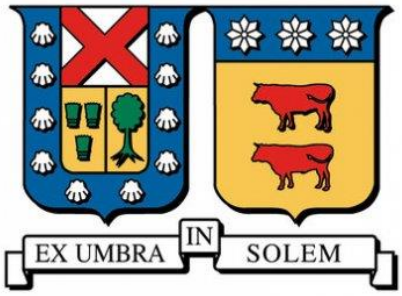

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INDUSTRIAS

COMPARACIÓN DE UN MODELO HÍBRIDO OBTENIDO DE LA MEZCLA DE VECTORES AUTORREGRESIVOS Y LA METODOLOGÍA DE REDES NEURONALES ARTIFICIALES ANN-VAR Y UN MODELO ECONOMÉTRICO DE VECTORES AUTORREGRESIVOS (VAR) PARA LA PREDICCIÓN DEL NIVEL DE MP2.5 EN SANTIAGO DE CHILE

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

AUTOR

ESTEFANIA ISABEL ROJAS HERRERA

PROFESOR GUÍA:

WERNER KRISTJANPOLLER

PROFESOR CORREFERENTE:

KEVIN MICHELL

VALPARAISO, NOVIEMBRE, 2018.

AGRADECIMIENTOS

Agradezco a Dios reconfortarme y ayudarme día a día, a mi familia y amigos por la paciencia y apoyo que me han brindado durante mi instancia en la universidad y a lo largo de mi vida, también a todos aquellos que pasaron por mi vida para dejar parte de ellos, sin embargo ya no están cerca. Agradezco a mi madre por motivarme en momentos de tristezas y a mi padre por entregarme su cariño y enseñarme lo que es el esfuerzo. Por último, agradezco a mi profesor guía por la disposición y ayuda brindada y a la Universidad Técnica Federico Santa María, por darme la oportunidad de convertirme en una ingeniera civil industrial.

RESUMEN EJECUTIVO

El trabajo realizado compara modelos para la predicción de material particulado MP2.5 y entrega como resultado que los modelos híbridos mejoran la precisión en el pronóstico del nivel de material particulado al mezclar modelos autorregresivos, factores meteorológicos y redes neuronales. En general, el mejor pronóstico de ANN para el caso de material particulado por hora está dado por modelo híbrido encontrado con autorregresivo 2, ventana móvil de 50, 2 capas y 20 neuronas, el mejor modelo horario se tiene que el MSE presenta una mejor de 65,7512 a 39,8348, esto es cerca del 39,41%, este modelo no incluye las variables meteorológicas como entrada adicional a la red, ya que para el caso del modelo por hora, incluirlas no entregaba buenos indicadores. Para el caso por día los modelos híbridos entregan mejores resultados incluyendo todas las variables y con ventana móvil de 20 días, pero además aquellos que tenían como entrada a la red los pronósticos del VAR y los factores meteorológicos; lo modelos presentados entregan mejores indicadores que el modelo econométrico, con una mejora del MSE de alrededor del 11% y del MAPE de aproximadamente un 17%. Además, puede verse durante el análisis que los tamaños de ventana móvil más reducidos fueron los que entregaron mejores resultados. Y Por otro lado, al variar el número de neuronas no se muestran mejorías, no pudiendo determinar algún número de neuronas por sobre otro, es decir, no es de gran importancia en la cantidad de neuronas en los modelos realizados.

El uso de modelos híbridos ha resultado ser una metodología bastante eficaz al mejorar pronósticos. El mantener pronósticos cercanos ayuda en el contexto medioambiental a prevenir gran cantidad de problemas.

ABSTRACT

Este estudio compara un modelo econométrico con modelos híbridos para obtener aquel que entregue la mayor precisión en el pronóstico del material particulado, MP2.5, para un día hacia adelante y una hora adelante, con el objetivo de un análisis más completo en temas de contaminación ambiental. Específicamente, se hace uso de diferentes configuraciones de Red Neuronal Artificial (ANN) junto a un modelo econométrico, VAR y datos meteorológicos. Los modelos híbridos son capaces de capturar la relación no lineal que el modelo VAR no logra incluir. La motivación para usar modelos híbridos esta en que la integración de diferentes modelos logra recoger más información y por lo tanto mejora la capacidad de predecir en contraste con modelos separados. La capacidad predictiva de los modelos es comparada a través del MSE y MAPE. Los resultados arrojan que el modelo híbrido VAR-ANN supera al econométrico, tanto en el pronóstico por hora, como por día.

Keywords: MP2.5, VAR, Contaminación ambiental, Red Neuronal Artificial

Índice de contenidos

Índice de ilustraciones

Índice de tablas

1. Problema de investigación

El nivel de contaminación en Chile ha ido en aumento en los últimos años a causa del mayor uso de calefacción domiciliaria junto con el aumento del tránsito vehicular y el desarrollo industrial. En el país existen una serie de ciudades que presentan un alto grado de contaminación, siendo una de las más contaminadas la ciudad de Santiago debido a que es esta donde hay una mayor concentración de fuerza laboral del país.

La contaminación del aire resulta de la introducción de agentes físicos, químicos o biológicos que alteran las características naturales del aire y que traen consigo resultados perjudiciales para la salud humana y equilibrio biológico. El aumento de la contaminación es causante de diversos trastornos en la salud de las personas estudios han demostrado que el estar expuestos a ambientes con niveles de contaminación están asociados al deterioro de función pulmonar, enfermedades respiratorias como asma, entre otras (Romero, Diego, & Álvarez, 2006).

El predecir distintos índices para poder tomar acciones ante algún eventual suceso es un tema en aumento ya que al tener noción de lo que podría pasar en el futuro permitiría evitar resultados indeseados, para esto se utilizan diversos métodos, como son los econométricos y en los últimos años se ha hecho uso de la inteligencia artificial, ya que permiten manejar muchos datos y aprender en base a estos para estimar un comportamiento; el uso de estas herramientas se ha visto en diversas áreas, tanto para determinar la calidad del agua utilizando inteligencia artificial (Lopez, Figueroa , & Corrales, 2015), como para determinar la volatilidad de distintos índices bursátiles como oro mediante modelos híbridos (Kristjanpoller & Minutolo, 2015), inclusive para predecir

el valor del IPC mediante métodos de redes neuronales diferenciales (Cabrera & Ortiz, 2012) entre otros.

En Chile se han realizado algunos pronósticos de los niveles de contaminación en el aire, como por ejemplo utilizando un modelo de inteligencia artificial de redes neuronales (Salini & Perez, 2006); realizando una comparación de cuatro métodos, como son análisis discriminante no-paramétrico, redes neuronales, regresión lineal múltiple y modelos MARS encontrando aquel que entregue un resultado más acertado (Silva, Alvarado, Montaño, & Pérez, 2003).Sin embargo no se han realizado estudios al respecto utilizando modelos híbridos, los cuales son la mezcla de métodos de inteligencia artificial con métodos genéricos de predicción, estos han mostrado una significativa disminución en el error asociado al pronóstico (Kristjanpoller & Minutolo, 2015) mostrando ser una herramienta altamente recomendada por sus buenos resultados, por esta razón y sumado a los buenos resultados obtenidos por un reciente trabajo de investigación en el cual se hace uso de un novedoso modelo hibrido para realizar un pronóstico de las concentraciones de partículas en el aire en las ciudades de Beijing, Tianjin y la provincia de Hebei (Feng, y otros, 2015), es que se busca realizar un modelo similar para la ciudad de Santiago.

Mediante lo expuesto queda preguntarse, ¿Cuáles son las consecuencias en la salud de las personas de la elevada contaminación?, ¿Qué contaminantes son los más dañinos? Y finalmente, a la hora de predecir, ¿Es mejor un modelo hibrido ante uno econométrico?

2. Objetivos

2.1 Objetivo general

Evaluar la calidad de dos modelos para predecir el nivel de contaminación del aire en Santiago, mediante la comparación de un modelo econométrico con un modelo hibrido, para seleccionar el modelo de predicción más adecuado que pueda ser utilizado para tomar medidas preventivas y evitar problemas de salud en la población.

2.2 Objetivos específicos

- Analizar las diversas consecuencias de los elevados niveles de contaminación en el aire en la salud de las personas, mediante estudios ya realizados sobre el tema, para tener noción de la importancia a la hora de tomar medidas preventivas respecto a la contaminación.
- Identificar los contaminantes más dañinos y aquellos que se presenten en mayor cantidad en el aire, mediante el escrutinio de artículos relacionados, para seleccionar los datos más adecuados que sirvan de entrada al modelo.
- Realizar un modelo econométrico de predicción de contaminantes en el aire, utilizando como entrada las variables más representativas, para obtener un punto de comparación con el modelo hibrido.
- Generar un modelo hibrido y compararlo con el modelo econométrico mientras se realizan modificaciones, integrando variables y variando métodos de resolución, para obtener un modelo mejorado, utilizando la minimización del error como indicador.

3. Marco Teórico

3.1 Estudios previos

Se han realizado varios estudios respecto del tema de la contaminación, respecto de la importancia que tiene dadas sus repercusiones, utilizando distintos métodos. Por ejemplo El 2003 (Silva, Alvarado, Montaño, & Pérez, 2003) realizan un estudio en el que comparan cuatro procedimientos predictivos, para la ciudad de Santiago de Chile. Los métodos utilizados son análisis discriminante o paramétrico, redes neuronales, regresión lineal múltiple y modelos MARS. De donde se pudo destacar el desempeño observado en la metodología no paramétrica, ya que esta no requiere supuestos distribucionales lo que lo sitúa en un contexto más realista y flexible. Algunos inconvenientes que se plantean son que el método MARS necesita que se fijen parámetros e especifiquen su complejidad, en el caso de las redes neuronales, su arquitectura, numero de nodos de la capa oculta y la cantidad de capas juegan un rol fundamental, además los MARS requieren de especificar el número máximo de funciones base, el orden máximo de interacción entre las variables predictoras, por lo tan tanto tener conocimiento previo sobre el tema ayuda para cada problema en particular.

En (Salini & Perez, 2006) Se diseña una red neuronal artificial (RNA) para hacer predicciones de valores de concentraciones horarias de material particulado fino en la atmósfera. El tipo de modelo de RNA usado fue uno de multicapas, alimentado hacia adelante y entrenado mediante la técnica de propagación hacia atrás. Se probaron redes sin capa oculta y con una y dos capas ocultas. El mejor modelo resultó ser con una capa oculta. Además, el perceptrón simple con función lineal fue superior en efectuar predicciones de

un paso en adelante que el método de persistencia. Se observó, además, que el perceptrón simple mejora su capacidad de predecir cuándo se usa una función de transferencia no lineal. Se puede afirmar que cuando efectos no lineales no son demasiado importantes en la modelación, las redes de multicapas no son significativamente mejores que el perceptrón. Sin embargo, como ocurrió en este caso, cuando estos efectos no lineales pasan a ser importantes, las redes de multicapas son mejores en cuanto a su capacidad de predicción, respecto de los modelos lineales.

(Niu, Wang, Sun, & Li, 2016) Utiliza un modelo hibrido que mezcla el principio de ´´descomposición y conjunto´´ con una meta heurística denominada el optimizador de lobo gris (GWO) para mejorar la precisión de la predicción de PM 2.5 diaria, este algoritmo de predicción inspirado en el lobo, simula el comportamiento de caza y liderazgo del lobo gris, y ha sido verificado por ser más competitivo que otros algoritmos se plantea como un marco prometedor para la predicción, pero es obvio que el número de componentes independientes descompuestos en etapa de descomposición de datos es importante para los resultados de los pronósticos. Ya que demasiados componentes pueden dar lugar a la complejidad del modelo y la acumulación de errores.

En (Perez & Gramsch, 2016) se presentan los resultados de un modelo de predicción concentraciones PM2.5 por hora en Santiago de Chile. El estudio se concentra en la comparación entre los valores observados de modelo y en la estación de monitoreo con las concentraciones más altas (estación Navia Cerro) para el período de tiempo entre abril y agosto, ya que esta es la estación donde se presentan mayores concentraciones por lo general. El modelo de predicción es una red neural de alimentación hacia adelante, las

variables de entrada son valores horarios pasados de las concentraciones de PM10 y PM2.5 medidos en la estación de ciudad con los valores más altos, las concentraciones de una estación vecina y algunas variables meteorológicas observadas y pronosticados. La formación se realiza con datos de 2010 y 2011 y el modelo se prueba con valores de 2012. La exactitud de la predicción es significativamente mejor que las diferentes formas de persistencia y puede ser considerado como una herramienta útil para anticipar episodios.

En (Niu, Gan, Sun, & Li, 2017) se muestra un estudio de descomposición en modo empírico conjunto y máquina vectorial de soporte cuadrado mínimo (EEMDLSSVM) basado en la reconstrucción del espacio de fase (PSR) se propone para un día de anticipación de la predicción de concentración de PM 2.5, de acuerdo con la aplicación de un paradigma de aprendizaje de descomposición-conjunto. Los principales métodos del modelo propuesto incluyen principalmente: en primer lugar, EEMD (Descomposición de modo empírico de conjunto) para descomponer los datos originales de concentración de PM 2.5 en algunas funciones modelo intrínsecas (IMFS); segundo, se aplica PSR (reconstrucción del espacio de fases) para determinar la forma de entrada de cada componente extraído; tercero, LSSVM (máquina vectorial de soporte cuadrado mínimo), es una herramienta de pronóstico eficaz, se utiliza para predecir todos los componentes reconstruidas de forma independiente; finalmente, se emplea otro LSSVM para agregar todos los componentes previstos en resultados de conjunto a una predicción final. Los resultados muestran que este modelo propuesto puede superar a los modelos de comparación y puede significativamente mejorar la predicción del rendimiento en términos de mayor exactitud de predicción.

En (Feng, y otros, 2015) se plantea un nuevo modelo híbrido que combina análisis de la trayectoria de masa de aire y la transformación wavelet para mejorar la red neuronal artificial (ANN) , se realiza un pronóstico las concentraciones medias diarias de PM 2.5 con dos días de anticipación. El modelo fue desarrollado a partir de 13 diferentes estaciones de control de la contaminación del aire en Beijing, Tianjin, y provincia Hebei (área Jing-Jin-Ji). La trayectoria masa de aire se utiliza para reconocer distintos corredores para el transporte de aire "sucio" y aire "limpio" de las estaciones seleccionadas. Con cada corredor, se construyó una red de estación triangular basado en las trayectorias de las masas de aire y las distancias entre los sitios vecinos. La velocidad y dirección del viento también se consideraron como parámetros en el cálculo de este indicador de contaminación atmosférica. Por otra parte, la serie de tiempo original de PM 2.5 concentración se descompuso mediante la transformación wavelet en unas pocas sub-series con menor variabilidad. La estrategia de predicción se aplicó a cada uno de ellos y, posteriormente se agruparon los resultados de predicción individuales de pronóstico diario meteorológicos, los respectivos predictores de contaminantes se utilizaron como entrada a un tipo red neural de retro propagación de múltiples capas (MLP). La verificación experimental del modelo propuesto se llevó a cabo durante un período de más de un año (entre septiembre de 2013 y octubre de 2014). Se ha encontrado que la trayectoria modelo geográfico y la transformación wavelet pueden ser herramientas eficaces para mejorar los pronósticos de PM 2.5.

En (Lv, Cobourn, & Bai, 2016) se utilizan modelos de regresión empíricos para pronosticar la cantidad de PM 2.5 y O 3. Las concentraciones de contaminación del aire se desarrollaron y evaluaron en tres grandes ciudades de China, Beijing, Nanjing y

Guangzhou. Los modelos de predicción son modelos de regresión no lineal empíricos diseñados para uso en una plataforma de recuperación de datos y la previsión automatizado. El modelo incluye una variable de la calidad del aire en contra del viento, PM24, para tener en cuenta para el transporte regional de PM 2.5, y una variable de persistencia (día anterior de concentración de PM 2.5. El modelo de PM2.5 arrojo buenos resultados, con coeficiente de determinación (R 2) valores de 0,54, 0,65 y 0,64, y el error medio normalizado (NME) valores de 0,40, 0,26 y 0,23, respectivamente, para las tres ciudades. El modelo de O3 también se comportó adecuadamente. El rendimiento global de pronóstico de PM 2.5 y O 3 durante el año de ensayo varió de regular a buena, dependiendo de la ubicación. Las previsiones eran algo diferentes en comparación con retrospectivas de ese mismo año, dependiendo de la exactitud de los datos de entrada meteorológicos Pronosticados. Para los pronósticos críticos, que son aquellos que sobrepasan los estándares, el modelo de PM 2,5 arrojo buenos resultados.

En (Jian, Zhao, Zhu, Zhang, & Bertolatti, 2012) con el fin de investigar el efecto de los factores meteorológicos en partículas submicrométricas (partícula ultra fina (UFP) y partículas 1,0 (PM 1.0)) se llevó a cabo un modelo de estudio en Hangzhou, una ciudad con un rápido aumento de vehículos en carretera. Se utilizó para este fin modelo estadístico autorregresivo integrados de media móvil (ARIMA).Los resultados ARIMA indicaron que la presión barométrica y la velocidad del viento fueron anti-correlacionada y la temperatura y la humedad relativa se correlacionó positivamente con las concentraciones de Número de UFP y PM 1.0. Estos datos sugieren que los factores meteorológicos son significativos predictores en la previsión de las concentraciones atmosféricas de partículas

submicrométricas. Este estudio proporciona un marco que puede ser aplicado en estudios futuros, con los datos de series de tiempo a gran escala, para predecir el impacto de los factores meteorológicos en las concentraciones de partículas submicrónicas en las ciudades de rápido desarrollo como lo es China.

El material articulado atmosférico (PM) es uno de los contaminantes que pueden tener un significativo impacto en la salud humana, dado esto, en (Biancofiore, y otros, 2017) se realiza un estudio, los datos son recopilados durante tres años en un área urbana de la costa del Adriático estos se analizan utilizando tres modelos: un modelo de regresión lineal múltiple, un modelo de red neural con y sin arquitectura recursiva. Los parámetros meteorológicos medidos y la concentración de PM10 se utilizan como entrada para pronosticar la concentración diaria promedio de PM10 de uno a tres días antes. Todas las simulaciones muestran que la red neuronal con la arquitectura recursiva tiene mejores actuaciones en comparación tanto con el modelo de regresión lineal múltiple y el modelo de red neural sin la arquitectura recursiva. Se observó que la inclusión de concentración de monóxido de carbono (CO) como parámetro adicional de entrada en el modelo ha servido para mejorar el pronóstico. Por último, todos los modelos se utilizan para predecir la concentración de PM2.5, usando como entrada los datos meteorológicos, la PM10 y la concentración de CO, para simular la situación cuando no se observa PM2.5. La comparación entre PM2.5 observada y pronosticada muestra que la red neuronal es capaz de predecir las concentraciones de PM2.5 incluso si PM2.5 no está incluido entre los parámetros de entrada.

En (Elangasinghe, Singhal, Kim N, & Salmond, 2014), se propone una metodología para extraer la información clave de los parámetros meteorológicos normalmente disponibles y el patrón de emisión de las fuentes presentes durante todo el año (por ejemplo, emisiones de tráfico) para construir una herramienta de pronóstico de contaminación del aire que sea fiable. La metodología se prueba mediante el modelado de las concentraciones de NO2 en un sitio cerca de una carretera principal en Auckland, Nueva Zelanda. El modelo básico consiste en un modelo ANN para predecir las concentraciones de NO2 utilizando ocho variables de predicción: velocidad del viento, estas son: dirección del viento, la radiación solar, temperatura, humedad relativa, así como "hora del día", "día de la semana" y "mes del año" que representa las variaciones en el tiempo de las emisiones de acuerdo con sus escalas de tiempo correspondientes. De las tres técnicas de optimización de analizadas en este estudio, siendo estas, un algoritmo genético, la selección hacia adelante, y la eliminación hacia atrás, la técnica de algoritmo genético dio predicciones resultantes con un error absoluto medio más pequeño. La naturaleza de la función no lineal interna del modelo de red neural optimizada genéticamente entrenados se extrajo entonces basada en la respuesta del modelo a las perturbaciones a las variables predictoras individuales a través de análisis de sensibilidad. Un modelo simplificado, basado en la eliminación sucesiva de las variables meteorológicas de predicciones menos significativas, fue entonces desarrollado hasta que la eliminación subsiguiente diera como resultado una disminución significativa en el rendimiento del modelo. Se encontró que el modelo ANN desarrollado para superar a un modelo de regresión lineal basado en los mismos parámetros de entrada. El enfoque propuesto ilustra cómo la técnica de modelado ANN se puede utilizar para identificar las variables meteorológicas claves necesarias para

capturar adecuadamente la variabilidad temporal de las concentraciones de contaminación del aire para un escenario específico. Este estudio revela que, eligiendo cuidadosamente las entradas para representar mensual, diaria y patrones de emisión por hora y relaciones para la velocidad del viento, dirección del viento, ture tempera atmosférica, la humedad relativa y la radiación solar, un modelo simple de ANN puede dar una predicción fiable de las concentraciones de dióxido de nitrógeno. Una vez que la mayoría de los factores que influyen han sido identificados a través de análisis de sensibilidad, incluso una sola red se puede utilizar para múltiples contaminantes sobre la base de los parámetros que influyen comúnmente. Una metodología similar se podría aplicar a otros escenarios en los que los patrones meteorológicos y de emisión de contaminantes, aunque de diferentes maneras, como, por diferentes escenarios de emisiones de carretera o de fuentes industriales.

3.2 Situación en chile

En Chile se mantienen registros de la situación medioambiental del país, tanto en formato de informes como también en los Sistemas de Información del Ministerio del Medio Ambiente que están disponibles en sitios web, entre ellos http://sinca.mma.gob.cl/ y http://sinia.mma.gob.cl/.

Ilustración 1: Ciudades con mayor contaminación. Fuente:(CENTRO DE ANÁLISIS DE POLÍTICAS PÚBLICAS, 2016).

Una amplia porción del territorio nacional está actualmente afectado por problemas de Contaminación, véase ilustracion1. La contaminación del aire además de dañar la salud de personas y animales, afecta la vegetación y el suelo, deteriora materiales, disminuye la

visibilidad y además contribuye al cambio climático. Es por esto que la calidad del aire en el país es un tema con tanto énfasis en materia de gestión ambiental, creándose para esto planes de descontaminación y trabajos con las comunidades para de esta forma mejorar la eficiencia de los hogares en términos energéticos. Se detectan tres grandes fuentes de contaminación del aire en el país: Las actividades industriales, la calefacción de las viviendas y los medios de transporte, para esto ha habido importantes avances como es el caso de regulaciones aplicadas para el sistema de transporte público, restricciones a vehículos catalíticos con mayor antigüedad y mayores exigencias en el control de emisiones por parte de las Plantas de Revisión Técnica. Otras medidas que se han tomado son incrementar significativamente el número de estaciones de monitoreo del país. Se continúan realizando recambios de calefactores y estableciendo normas de calidad y emisión para las principales fuentes industriales emisoras de contaminantes, como es el caso de termoeléctricas y fundiciones de cobre. También, se busca normar el límite de emisiones para actividades industriales no menos relevantes, como es el caso de las calderas y equipos electrógenos. Chile cuenta con la Estrategia de Descontaminación Atmosférica 2014-2018, la cual contempla 14 nuevos planes de descontaminación para este período, estrategia que sido un factor clave en la disminución de episodios críticos, la ilustración 2 muestra los planes de descontaminación vigentes. (CENTRO DE ANÁLISIS DE POLÍTICAS PÚBLICAS, 2016)

Planes de descontaminación vigentes

Año PDA	Comuna o zona fuente emisora	Declaración de Zona Saturada por:
1992	Complejo Industrial Ventanas	SO2 y MP10 en 1994
1993	Fundición Chuquicamata(actualizado 2001)	MP10 y SO2 en 1991
1995	Fundición Hernán Videla Lira	SO2 en 1993
1998	María Elena y Pedro de Valdivia (actualizado 2004)	MP10 en 1993
1998	Fundición de Caletones	MP10 y SO2 en 1994
1998	Fundición de Potrerillos	SO2 y MP10 en 1997
1998	Región Metropolitana (actualizado 2004 y 2009)	MP10, CO y O3 en 1996
2009	Temuco y Padre Las Casas	MP10 en 2005.
2010	Tocopilla	MP10 en 2007.
2013	Rancagua y 17 comunas del valle central de la VI Región	MP10 en 2009
2014	Andacollo	MP10 en 2009
2015	Temuco y Padre Las Casas	MP2.5 en 2013
2016	Talca y Maule	MP10 en 2010
2016	Chillán y Chillán Viejo	MP10 en 2013
2016	Osorno	MP10 en 2012
2016	Coyhaique	MP10 en 2012, MP2.5 en 2016

*** PDA en elaboración

Ilustración 2: Planes de descontaminación vigente. Fuente:(CENTRO DE ANÁLISIS DE POLÍTICAS PÚBLICAS, 2016).

3.3 Situación en Santiago de chile

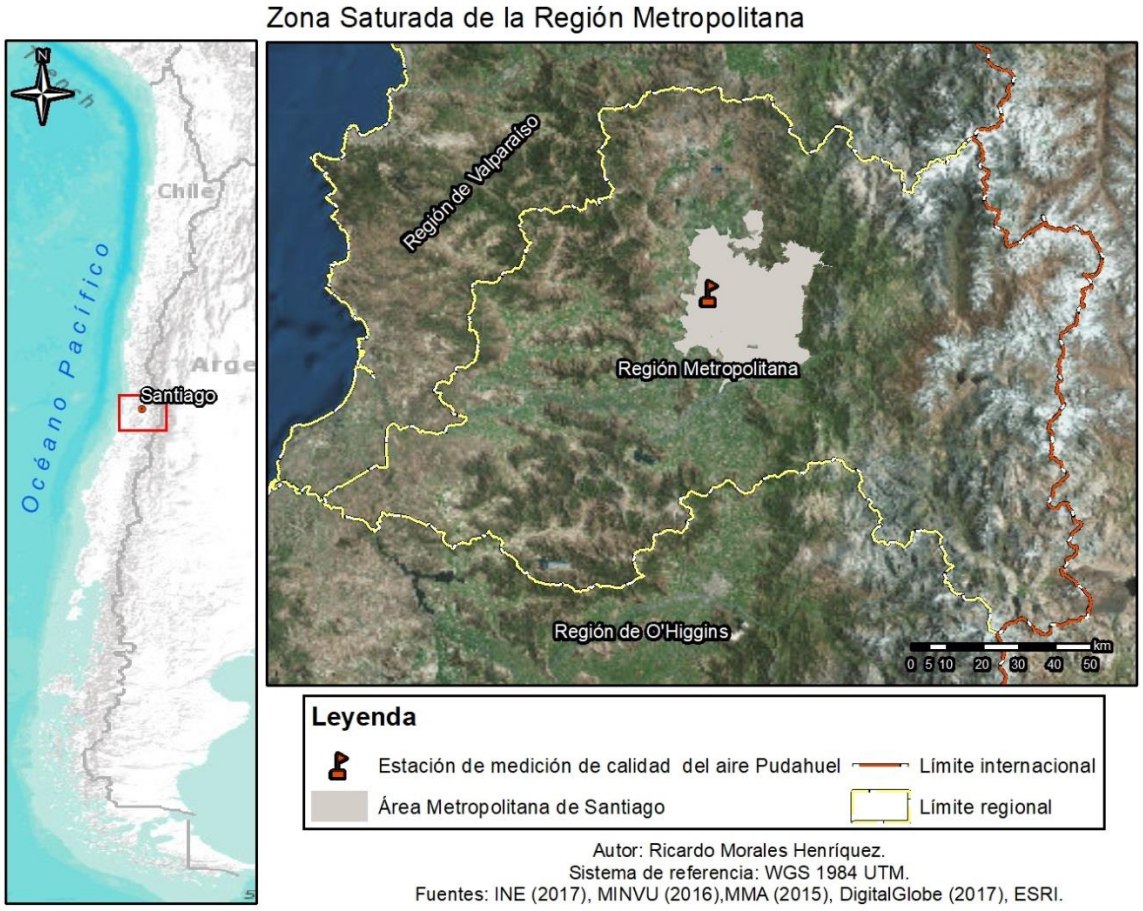

Ilustración 3: Cuenca de Santiago. Fuente: Elaboración propia.

La Región Metropolitana (33,5ºS, 70,8ºW) está ubicada entre la cordillera de Los Andes y la cordillera de la Costa. Por el norte, el cordón montañoso de Chacabuco marca el límite con la región de Valparaíso y por el sur, los cerros de Angostura y Chada constituyen el límite con la Región del Libertador General Bernardo O'Higgins, tiene una longitud de 80 km en la dirección norte-sur y de 35 km de este a oeste. Predominan los relieves montañosos que encierran hacia el centro de la región una amplia zona donde se emplaza el área metropolitana de Santiago. Su clima es de tipo mediterráneo. Es conocida como la cuenca de Santiago, y sus características dificultan la circulación de vientos y la renovación

del aire. Durante el período otoño-invierno las características topoclimáticas de Santiago generan condiciones desfavorables para la dispersión de contaminantes, ya que su topografía limita las capacidades de ventilación debido a las barreras orográficas, y la disminución de las temperaturas en los meses de otoño-invierno no permiten la elevación de las masas de aire en la cuenca. (Alvarado Zuñiga, 2010).

Las concentraciones de material particulado respirable (MP10 y MP2.5) tienen un ciclo estacional marcado, con valores más altos en otoño-invierno y menores en los meses de primavera y verano. Esta situación se muestra en la ilustración 4, a modo de ejemplo, con los promedios mensuales entre los años 1997 y 2015 de la estación Pudahuel.

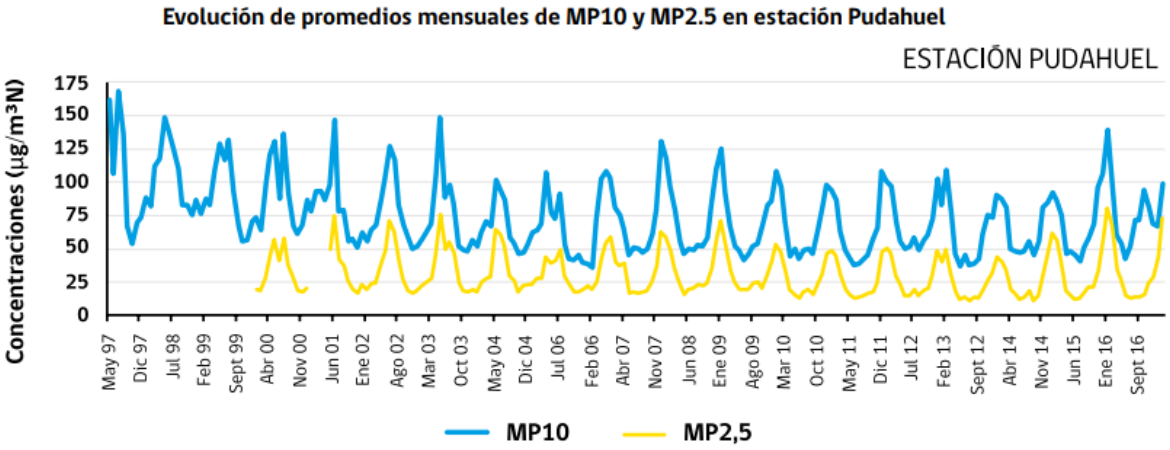

Ilustración 4: Evolución MP2.5 y MP10. Fuente:(Chile, 2016).

El alto nivel de contaminación mediante material particulado es un problema frecuente en la región metropolitana en especial durante el invierno. La principal causa de esta contaminación seria la elevada emisión de origen antropogénico tanto orgánicos como inorgánicos además de que la situación geográfica dificulta la dispersión de los contaminantes. Sumado a ello la densidad de la población, la contaminación debido al transporte público y situación meteorológica de la región (Salini & Perez, 2006).

En la figura 5 se puede apreciar una comparación entre desde el 2000 al 2015 de las estaciones las Condes y Pudahuel que permiten hacer una comparación entre los años 2000 y 2015. Se observan los percentiles 98 de las concentraciones diarias de MP2.5, notándose que las estaciones P. O'Higgins, La Florida y Las Condes tienen una disminución, mientras que Pudahuel tiene un aumento. Todas las estaciones, tanto en el año 2000 como en el 2015 están por sobre el valor de la norma diaria, presentando Pudahuel y Cerro Navia tienen un valor muy por encima de la norma. (Chile, 2016)

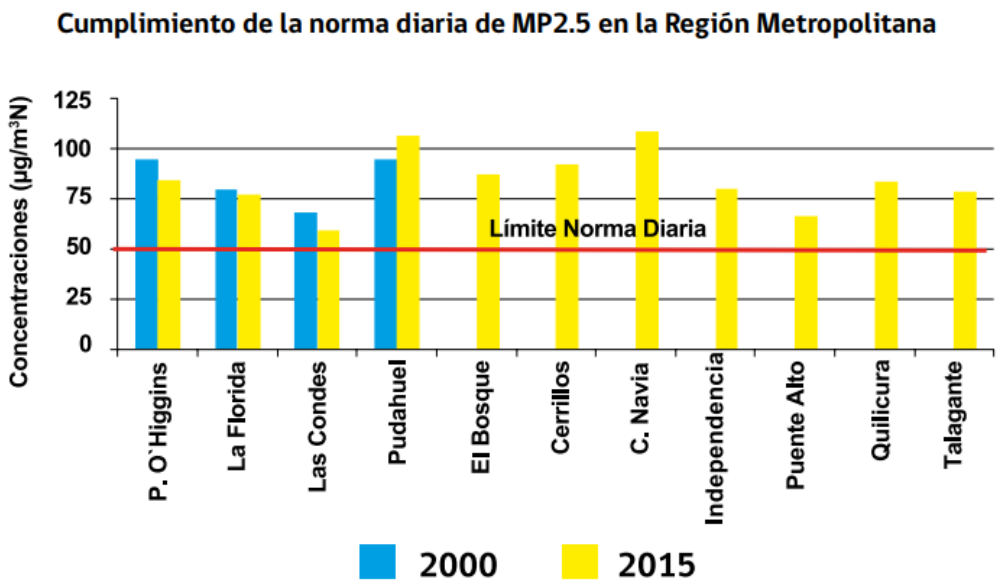

Ilustración 5: Cumplimiento de la norma diaria de MP2.5. Fuente: (Chile, 2016).

3.4 Efectos de la contaminación en la salud

Los problemas asociados a la salud de las personas son variados, estudios epidemiológicos realizados han demostrado que el estar expuesto a ambientes contaminados trae consigo un incremento en la incidencia de asma, el deterioro de la función pulmonar se puede ver drásticamente aumentado además del aumento en la gravedad de enfermedades respiratorias en niños y adolescentes; esto puede ocurrir incluso si la exposición está por debajo de lo que las normas internacionales consideran tolerables. (Romero, Diego, & Álvarez, 2006)

Los contaminantes presentes junto a sus derivados traen efectos nocivos en la salud ya que estos alteran moléculas que son indispensables para la realización de procesos bioquímicos y fisiológicos del ser humano. Los factores que limitan el riesgo de del perjuicio toxico son las propiedades físico- químicas, las dosis de las sustancias que entran en contacto con los tejidos críticos y las respuestas de estos a dichas sustancias. (Romero, Diego, & Álvarez, 2006)

Muchos efectos irritativos en el aparato respiratorio pueden ser consecuencia de las sustancias emitidas por los vehículos, principalmente el nitrógeno, ozono, oxidantes fotoquímicos, bióxido de azufre y partículas. (Romero, Diego, & Álvarez, 2006).

Gracias a las investigaciones realizadas sobre la contaminación atmosférica y su efecto en la salud se ha determinado que el riesgo por partículas inhalables depende de la penetración y cantidad de depósitos de estas partículas en las diferentes partes del aparato respiratorio y la respuesta biológica a estas. Las partículas que permanecen en los pulmones pueden ocasionar diversos efectos adversos para la salud de las personas, entre

ellos tenemos la Interferencia con los mecanismos de limpieza del tracto respiratorio que impide o dificulta la eliminación de partículas nocivas. Produce una irritabilidad e incluso generación de procesos cancerígenos debido a la toxicidad excesiva de algunas de estas partículas.

En la ilustración 6, se puede observar el número de casos asociados a mortalidad y morbilidad debido a la exposición a contaminación atmosférica por MP2.5 en el año 2015 según grupo etario. (MMA, 2017)

MORTALIDAD Y MORBILIDAD ASOCIADA A LA EXPOSICIÓN A MP2,5			
TIPO DE EVENTO	EVENTO	GRUPO DE EDAD	CASOS
Mortalidad Prematura	Cardiopulmonar	Mayores de 30 años	3.723
Admisiones Hospitalarias	Cardiovasculares	Mayores de 18 años	1.709
Admisiones Hospitalarias	Pulmonar crónica	Mayores de 18 años	231
Admisiones Hospitalarias	Neumonía	Mayores de 65 años	1.049
Admisiones Hospitalarias	Ataques de asma	0-64 años	152
Visita a Sala de Emergencias	Bronquitis aguda	0-17 años	108.100
Restricción de Actividad	Días de pérdida de trabajo	18-64 años	870.756
Restricción de Actividad	Días de actividad restringida	18-64 años	3.861.706
Restricción de Actividad	Días de actividad restringida menor	18-64 años	7.273.360

Ilustración 6: Mortalidad y morbilidad asociada a la exposición de MP2.5. Fuente: (MMA, 2017).

3.5 Contaminantes más dañinos

Los principales agentes contaminantes de la atmósfera se pueden clasificar en contaminantes primarios y secundarios.

Los contaminantes primarios son aquellas sustancias emitidas por la fuente contaminante y vertidas directamente a la atmósfera, siendo las más comunes el monóxido de carbono (CO), los hidrocarburos (NOx), los óxidos sulfurosos (SOx), dióxido de carbono (CO2), los compuestos orgánicos volátiles (COV), el material particulado PM, en general la mayoría de los hidrocarburos y las partículas, los demás se presentan menos frecuentemente. Los contaminantes secundarios se forman en el aire a partir de reacciones químicas o fotoquímicas entre los contaminantes primarios y los compuestos que están en la atmósfera. (Salini Calderon, 2009)

Aquellos contaminantes que tienen mayor capacidad para afectar la salud de las personas son aquellos que provienen de emisiones primarias o transformaciones atmosféricas. Estos son monóxido de carbono, óxidos de nitrógeno, hidrocarburos no quemados, ozono, plomo, oxidantes fotoquímicos y en menor medida partículas suspendidas de bióxido de azufre junto a los compuestos orgánicos volátiles, siendo los vehículos motorizados una de las fuentes más importantes de muchos de estos contaminantes**.** (Romero, Diego, & Álvarez, 2006)

Existe una especial atención a nivel de salud de la población por las sustancias químicas que presentan propiedades carcinogénicas o mutagénicas estos corresponden a los contaminantes orgánicos atmosféricos. Dichas sustancias pueden ser absorbidas directamente en su fase gaseosa o también acumularse en el aparato respiratorio cuando

están asociadas a partículas muy pequeñas (PM10: fracción respirable) Fuentes de PM10 son la combustión y los procesos industriales. (Salini & Perez, 2006)

Las partículas más grandes, sobre 5 µm son filtradas por la acción conjunta de los cilios del conducto nasal y la mucosa que cubre la cavidad nasal y la tráquea. Sin embargo, las partículas de diámetro entre 0,5 y 5 µm pueden depositarse en los bronquios e incluso en los alvéolos pulmonares, no obstante estas son eliminadas por los cilios de bronquios y bronquiolos al cabo de algunas horas. Las partículas menores a 0,5 µm tienen la facilidad de penetrar profundamente hasta depositarse en los alvéolos pulmonares, permaneciendo bastante tiempo, inclusive años, debido a que no existe un mecanismo que facilite la eliminación. Por esta razón el material particulado con un diámetro igual o menor a 2,5 µm (PM2,5) es el que más peligroso ya que este provoca mayor repercusión a la salud humana. (Alvarado Zuñiga, 2010)

En la ilustración 7, se muestra, a modo de comparación, el tamaño del material particulado de 2,5 µm y 10 µm respecto del diámetro de un cabello humano promedio.

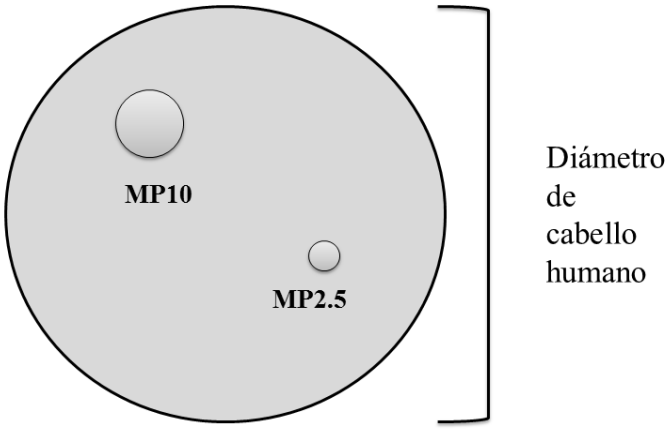

Ilustración 7: Relación tamaño cabello humano y material particulado. Fuente: Elaboración propia.

3.6 Series de tiempo

Según (Gujarati, 2010) Una serie de tiempo se puede definir como un conjunto de observaciones sobre los valores de una variable en diferentes momentos, es decir, medida en el tiempo. Se pueden distinguir dos tipos de modelos de series de tiempo:

Los Modelos deterministas que son de métodos de extrapolación sencillos en los que no se hace referencia a las fuentes o naturaleza de la aleatoriedad subyacente en la serie. Su simplicidad relativa generalmente va acompañada de menor precisión. Un ejemplo de esto son los modelos de promedio móvil en los que se calcula el pronóstico de la variable a partir de un promedio de los valores inmediatamente anteriores. Y los **Modelos estocásticos** que se basan en la descripción simplificada del proceso aleatorio subyacente en la serie ,es decir, se asume que la serie observada se extrae de un grupo de variables aleatorias con una cierta distribución conjunta difícil de determinar, por lo que se construyen modelos aproximados que sean útiles para la generación de pronósticos. (Rios, 2008)

La serie podría ser estacionaria o no estacionaria. Una Serie es no estacionaria si sus características de media, varianza y covarianza cambian a través del tiempo dificultando su modelamiento. Aun así en variadas ocasiones si la serie es diferenciada una o más veces la serie resultante ser estacionaria. Una Serie se considera estacionaria si su media y su varianza son constantes en el tiempo y además si el valor de la covarianza entre dos periodos depende sólo de la distancia o rezago entre estos dos periodos, y no del tiempo en el cual se calculó la covarianza. (Gujarati, 2010)

El análisis de regresión basado en información de series de tiempo supone implícitamente que las series de tiempo en las cuales se basa son estacionarias. Sin embargo la mayoría de las series no son estacionarias. Para medir la estacionariedad de forma informal se utiliza el correlograma, sin embargo si se desea hacer de forma formal se suelen utilizar las pruebas de Dickey-Fuller y Dickey-Fuller Aumentada.

La regresión de una variable de serie de tiempo sobre una o más variables de series de puede entregar resultados sin sentido o espurios. Este fenómeno es conocido como regresión espuria. Para evitar esto se debe conocer si existe integración entre las variables. Esto quiere decir que una combinación lineal de dos o más variables puede ser estacionaria aunque las variables no lo sean a nivel individual. Para detectar la cointegración se suele ocupar las pruebas de Engle- Granger (EG) y Engle-Granger aumentada (EGA) (Gujarati, 2010).

En general se utiliza la metodología de Box-Jenkins para el modelamiento estocástico de series de tiempo, ya que puede manejar cualquier serie, estacionaria o no estacionaria, y por haber sido implementado en numerosos programas computacionales. Los pasos básicos de la metodología de Box-Jenkins son:

1. Verificación de la estacionariedad de la serie. Si esta no es estacionaria, diferenciarla hasta alcanzar estacionariedad.
2. Identificar un modelo tentativo.
3. Estimar el modelo.
4. Verificar el diagnostico (si este no es adecuado, volver al paso 2).
5. Usar el modelo para pronosticar.

La idea es identificar el proceso estocástico que ha generado los datos, estimar los parámetros que caracterizan dicho proceso, verificar que se cumplan las hipótesis que han permitido la estimación de dichos parámetros. Si dichos supuestos no se cumplieran, la fase de verificación sirve como retroalimentación para una nueva fase de identificación. Cuando se satisfagan las condiciones de partida, se puede utilizar el modelo para pronosticar. (Rios, 2008)

3.7 Vectores Autorregresivos

Fueron introducidos por Sims en 1980 el término "autorregresivo" se refiere a la aparición del valor rezagado de la variable dependiente en el lado derecho, y el término "vector" se atribuye a que tratamos con un vector de dos o más variables, esta metodología se asemeja a la de las ecuaciones simultáneas, pues considera diversas variables endógenas de manera conjunta. En este tipo de modelo cada variable endógena es explicada por sus valores rezagados y por los valores rezagados de las demás variables endógenas dentro del modelo. (Gujarati, 2010)

Puede expresarse de forma matricial como:

$$Y_t = C + \Phi_1 Y_{t-1} + \cdots + \Phi_p Y_{t-p} + a_t$$

P muestra el rezago del sistema, es decir, el periodo de tiempo hasta el que se va a considerar que influyen sobre sí mismas y sobre las demás variables.

Las matrices Φ_n tienen la siguiente forma:

$$\Phi_n = \left[\begin{pmatrix} \Phi_{11}^n & \cdots & \Phi_{1k}^n \\ \vdots & \ddots & \vdots \\ \Phi_{k1}^n & \cdots & \Phi_{kk}^n \end{pmatrix} \right]$$

Con dimensión (k x k) con k número de variables en el sistema e Y_{t-n}

$$Y_{t-n} = \begin{bmatrix} z_{t-n} \\ x_{t-n} \\ v_{t-n} \end{bmatrix}$$

Con dimensiones (1 x k)

De la misma forma que las series individuales, los VAR deben ser estacionarios. Para estimar los VAR se aplica MCO aunque transformado matricialmente para un sistema de ecuaciones. (Hidalgo, 2014)

3.8 Redes neuronales

Las Redes Neuronales Artificiales (RNA) o sistemas conexionistas son sistemas de procesamiento de la información cuya estructura y funcionamiento se comportan de forma aproximadamente análoga a las redes de neuronas biológicas. Están constituidos por un conjunto de elementos simples de procesamiento llamados neuronas, estas están conectadas entre sí por conexiones que tienen un valor numérico modificable llamado peso. Lo que realiza la neurona es un procesamiento simple que consiste generalmente en la suma de los valores de entrada que recibe de otras neuronas con que tiene conexión, luego compara estos valores con un valor umbral y si este es superado o igualado se envía una salida a las unidades siguientes con las que tiene conexión. Tanto las entradas que la unidad recibe como las salidas que envía dependen a su vez del peso de las conexiones por las cuales se realizan dichas operaciones. (Mercado Polo, Pedraza Caballero, & Martínez Gómez, 2015).

3.8.1 Estructura

En cuanto a la estructura de una red neuronal podemos distinguir 3 niveles, cada nivel realiza una función distinta. Véase ilustración 8.

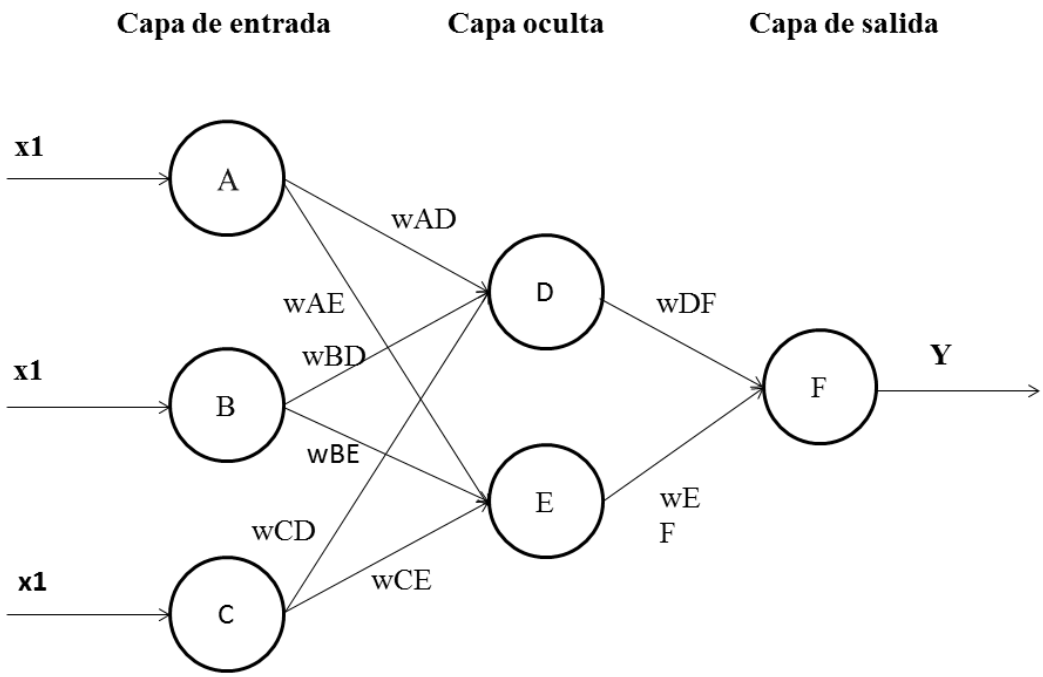

Ilustración 8: Estructura de red neuronal. Fuente: Elaboración propia

Capa de ingreso o input: Este es el primer nivel de la red. La capa de input recibe la información desde fuentes externas y las lleva hasta el siguiente nivel. En esta capa no se procesa información.

Algunas de las funciones de entrada más comúnmente utilizadas y conocidas son:

- **Sumatoria de las entradas pesadas**: es la suma de todos los valores de entrada a la neurona, multiplicados por sus correspondientes pesos.
- **Productoria de las entradas pesadas:** es el producto de todos los valores de entrada a la neurona, multiplicados por sus correspondientes pesos.

- **Máximo de las entradas pesadas:** solamente toma en consideración el valor de entrada más fuerte, previamente multiplicado por su peso correspondiente.

Capas ocultas: procesan la información mandada por las capas de entrada. La información es procesada gracias a una función matemática que trabaja sobre los datos ingresados. Esta función previamente es definida como la función de activación de la neurona. La red puede tener más de una capa oculta.

Capas de salida u output: El último componente que una neurona necesita es la función de salida. La función de salida determina qué valor se transfiere a las neuronas vinculadas. Si la función de activación está por debajo de un umbral determinado, ninguna salida se pasa a la neurona siguiente.

Dos de las funciones de salida más comunes son:

- Ninguna: este es el tipo de función más sencillo, tal que la salida es la misma que la entrada. Es también llamada función identidad.
- Binaria: $\begin{cases} 1, & \text{si act} \geq u \\ 0, & \text{si no} \end{cases}$

  Donde u es el umbral (Matich, 2001).

3.8.2 Funciones de activación

Una neurona artificial puede estar activa o inactiva. La función activación calcula el estado de actividad de una neurona, el rango de esta normalmente va de (0 a 1) o de (–1 a 1), dependiendo del tipo de función de activación. El valor 0 o -1 corresponde al estado inactivo y 1 a activo.

Las funciones de activación utilizadas comúnmente son:

1) Función lineal:

$$f(x) = \begin{cases} -1\ , & x \leq \frac{1}{a} \\ a * x\ , & \frac{-1}{a} < x < \frac{1}{a} \\ 1\ , & x \geq \frac{1}{a} \end{cases}$$

Cuando a = 1 la salida es igual a la entrada.

2) Función sigmoidea: Los valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1.

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

3) Función tangente hiperbólica:

Los valores de salida de la función tangente hiperbólica están comprendidos dentro de un rango que va de -1 a 1.

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

Función Tangente hiperbólica Sigmoidea: Los valores entregados están comprendidos entre -1 y 1. Incluye ventajas de las funciones Tanh y Sig para realizar pronósticos en redes de múltiples capas. (Matich, 2001)

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

3.8.3 Métodos de aprendizaje

Una red neuronal debe aprender a calcular la salida correcta para cada conjunto de entrada. Este proceso de aprendizaje es denominado: proceso de entrenamiento. Durante el proceso de aprendizaje, los pesos van sufriendo modificaciones hasta llegar a un estado de estabilidad.

Hay dos métodos de aprendizaje importantes que pueden distinguirse:

- Aprendizaje supervisado.
- Aprendizaje no supervisado.

Aprendizaje supervisado.

El aprendizaje supervisado es aquel que es controlado por un agente externo que determina la respuesta que debería generar la red a partir de cierta entrada. En caso de que ésta no coincida con la salida esperada, se debe proceder a modificar los pesos de las conexiones para que de esta forma la salida se aproxime más a la que se tiene por objetivo.

Aprendizaje por corrección de error.

Este consiste en ir ajustando los pesos de las conexiones de la red en función de la diferencia entre los valores deseados y los obtenidos a la salida de la red, es decir, en función del error. Uno de los algoritmos más conocidos es la regla de aprendizaje Delta o regla del mínimo error cuadrado (LMS Error: Least Mean Squared Error), este utiliza la diferencia con la salida objetivo o target, pero toma en consideración a todas las neuronas predecesoras que tiene la neurona de salida. De esta forma es posible cuantificar el error global cometido en cualquier instante durante el entrenamiento de la red, lo gran importancia, ya que mientras mayor información se tenga sobre el error cometido, más rápido se podrá la red ir aprendiendo. Importante mencionar la regla de aprendizaje de

propagación hacia atrás, que es una generalización de la regla de aprendizaje Delta, es la primera regla de aprendizaje que permitió realizar cambios sobre los pesos en las conexiones de la capa oculta. (Matich, 2001)

Aprendizaje no supervisado.

Las redes con aprendizaje no supervisado no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas. En este caso la red no recibe información que le indique si la salida entregada es correcta o no.

3.8.4 Tipos de redes

En cuanto a tipos de redes se tienen las Redes monocapa, donde se establecen conexiones entre las neuronas que pertenecen a la única capa que constituye la red. Y las Redes multicapa que como su nombre lo dice son aquellas que disponen de un conjunto de neuronas agrupadas en múltiples capas o también denominados niveles (véase ilustraciones 9 y 10). Por lo general, todas las neuronas de una capa reciben señales de entrada desde otra capa anterior, y envían señales de salida a una capa posterior. A estas conexiones se las conoce como conexiones hacia adelante o feedforward (Matich, 2001).

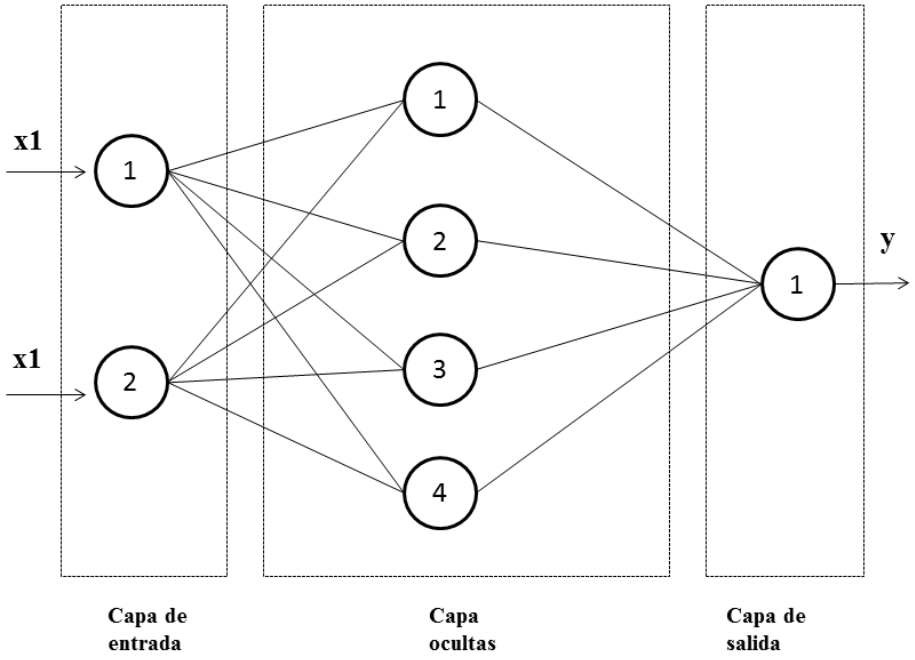

Ilustración 9: red neuronal multicapa1. Fuente: Elaboración Propia

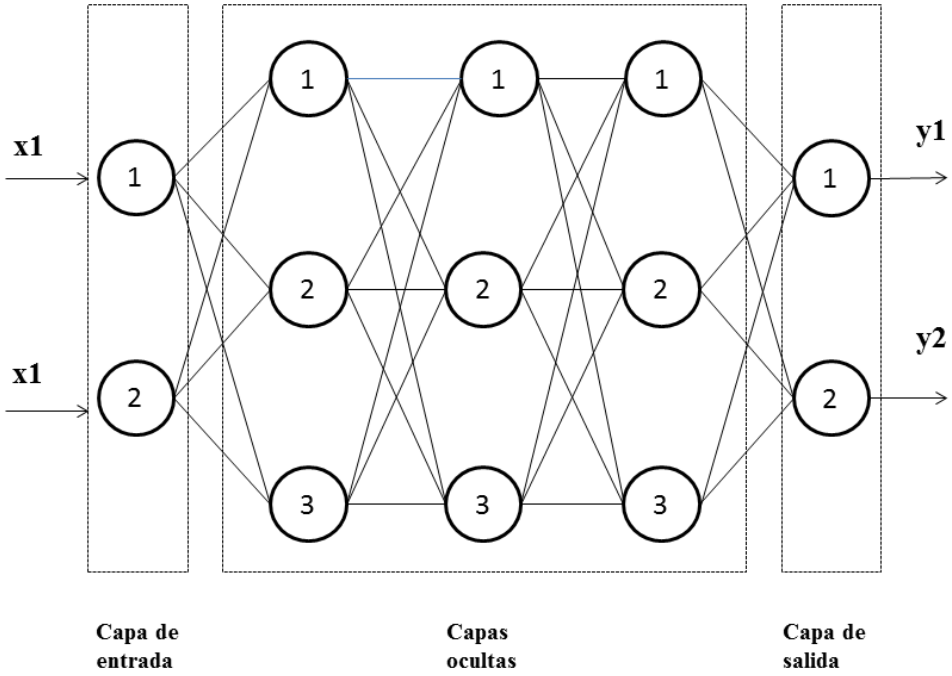

Ilustración 10: Red neuronal multicapa2. Fuente: Elaboración Propia

Además, en un gran número de estas redes existe la posibilidad de conectar la salida de las neuronas de capas posteriores a la entrada de capas anteriores, estas conexiones son denominadas se conocen como conexiones hacia atrás o feedback. Estas dos posibilidades permiten distinguir entre dos tipos de redes con múltiples capas: las redes con conexiones hacia adelante o redes feedforward, y las redes que tienen conexiones tanto hacia adelante como hacia atrás, también llamadas redes feedforward/feedback. (Matich, 2001)

3.8.5 Conexión entre neuronas.

La conectividad entre los nodos de una red neuronal está relacionada con la forma en que las salidas de las neuronas están canalizadas para convertirse en entradas de otras neuronas. La señal de salida de un nodo puede ser una entrada de otro elemento de proceso, o incluso ser una entrada de sí mismo, conocida como conexión autorrecurrente. Cuando ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes, la red se describe como de conexión hacia delante. Cuando las salidas pueden ser conectadas como entradas de neuronas de niveles previos o del mismo nivel, incluyéndose ellas mismas, la red es de conexión hacia atrás. Las redes de propagación hacia atrás que tienen lazos cerrados son denominadas: sistemas recurrentes. (Matich, 2001)

Redes de propagación hacia atrás (backpropagation).

El nombre de backpropagation viene de la forma en que el error es propagado hacia atrás a través de la red neuronal, es decir, el error se propaga hacia atrás desde la capa de salida. Esto permite que los pesos sobre las conexiones de las neuronas ubicadas en las capas ocultas se modifiquen durante el entrenamiento. El cambio de los pesos en las conexiones de las neuronas además de influir sobre la entrada global, influye en la activación y por consiguiente en la salida de una neurona. Por lo tanto, es de gran utilidad considerar las variaciones de la función activación al modificarse el valor de los pesos. (Matich, 2001)

3.9 Medidas de rendimiento

3.9.1 Error cuadrático medio (MSE)

Para medir la magnitud de los errores y determinar si es posible el pronóstico exacto, se puede considerar el valor del error cuadrático medio (MSE). Esta medida es el promedio de los errores cuadráticos de todos los pronósticos. Errores grandes en la predicción indicarían que el componente irregular es muy grande y la técnica de predicción no logra capturarla de manera apropiada. El control de los errores indica si el modelo se ajusta al patrón de la serie. El mejor modelo resultará ser aquel cuyo valor de MSE sea menor. (Delgadillo-Ruiz, Ramírez-Moreno, Leos-Rodríguez, Salas González, & Valdez-Cepeda, 2016)

$$MSE = \frac{1}{N}\sum_{1}^{N}(y_t - \hat{y}_t)^2$$

3.9.2 Error Porcentual Absoluto Medio (MAPE)

Es la media de los errores porcentuales en valor absoluto, por lo tanto no considera si el error es hacia arriba o hacia abajo, simplemente si es error, sin importar la dirección. El MAPE suele ser útil para expresar de una forma simple, en términos porcentuales genéricos, el error cometido. El MAPE es una de las estadísticas más usadas para comparación de pronósticos, y se utiliza junto con el MSE para comparar modelos, midiendo el desempeño en el ajuste y pronóstico sobre las estimaciones de los modelos. Es mejor pronóstico aquel que tenga menor MAPE. (Valencia, Vanegas, Correa, & Restrepo, 2017)

$$MAPE = \frac{1}{N}\sum_{1}^{N}\frac{|y_t - \hat{y}_t|}{|y_t|}$$

4. Metodología y Data

En esta sección se detalla la metodología empleada, primeramente se muestran los datos y su estructura, posteriormente se explica la construcción del modelo econométrico y finalmente la construcción del modelo hibrido realizado.

Se modelara el comportamiento del material particulado tanto por hora como por día, para de esta forma tener un reporte más completo que sirva como fuente de análisis y toma de decisión.

4.1 Datos

Los datos empleados son extraídos de la base de datos de la página del Sistema de Información Nacional del Aire SINCA. Tanto para la realización de los modelos horarios como para los diarios se utilizaron los datos entregados de la comuna de Pudahuel, región metropolitana, para el generar el modelo por hora se usan datos correspondientes desde el 6 de mayo del 2017 al 6 de mayo del 2018, medidos cada una hora, sin embargo a estos datos se le quitaron aquellas horas y días en que no existía registro de alguno de los parámetros, quedando en total una cantidad de 6661 mediciones.

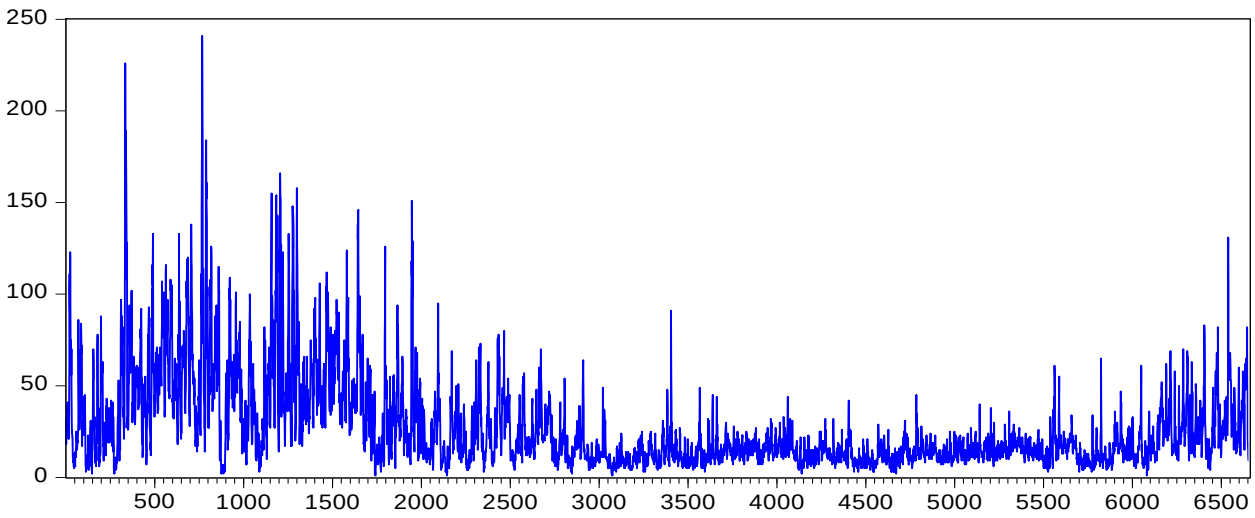

Ilustración 11:MP2.5. Fuente: Elaboración Propia

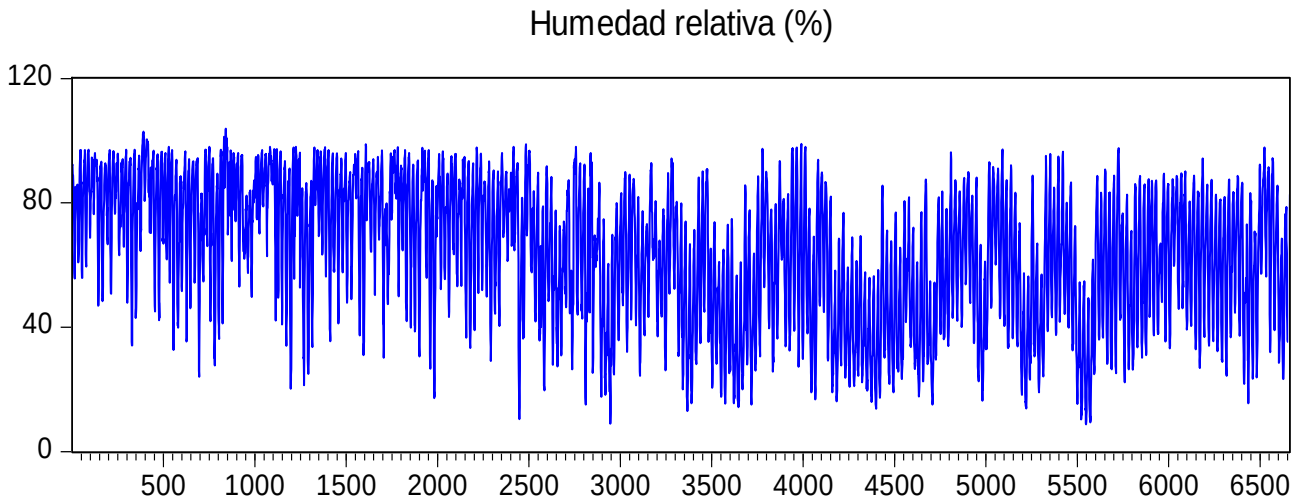

Ilustración 12: Humedad Relativa. Fuente: Elaboración propia

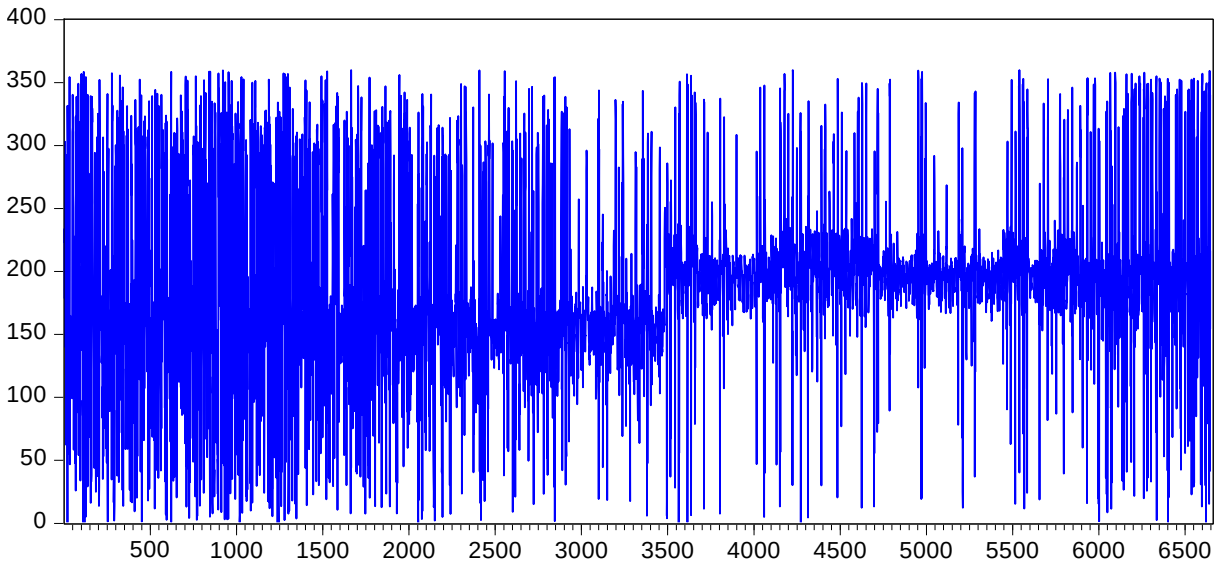

Ilustración 13: Dirección del Viento. Fuente: Elaboración propia

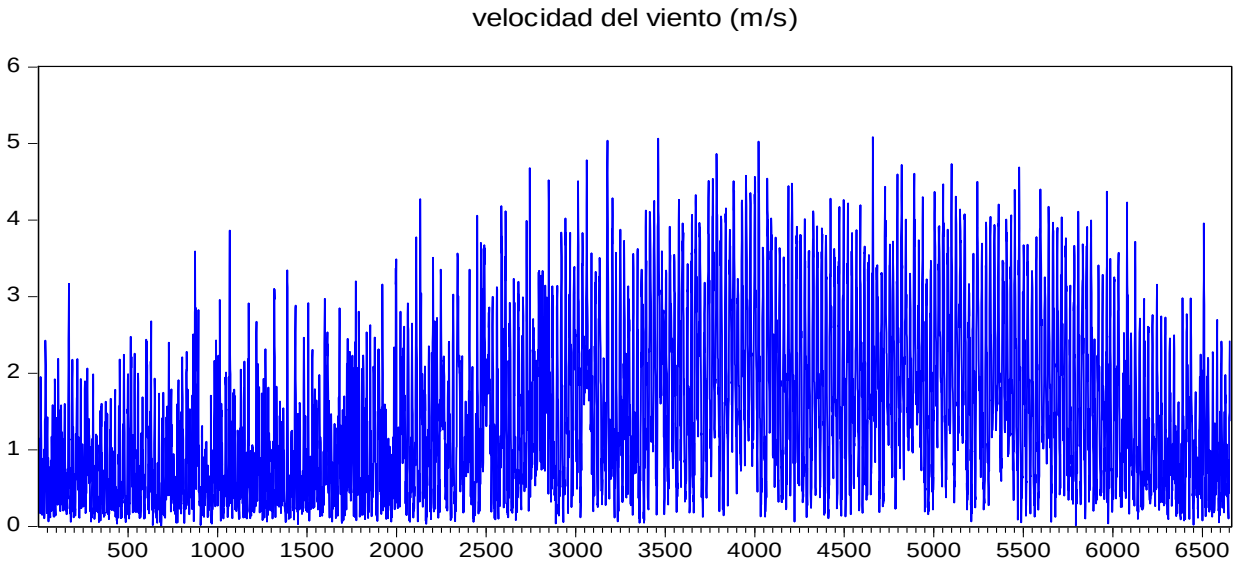

Ilustración 14: Velocidad del viento. Fuente: Elaboración propia

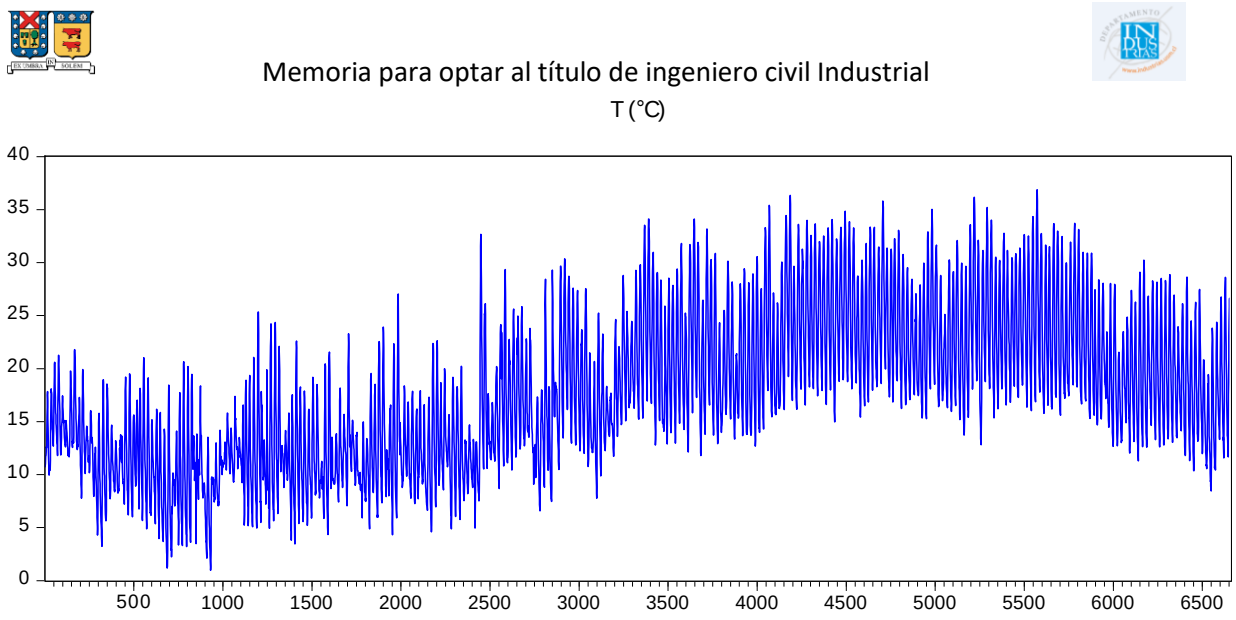

Ilustración 15: Temperatura. Fuente: Elaboración propia

En las ilustraciones 11 ,12 ,13 ,14 y 15 se puede apreciar el comportamiento de cada una de las variables, podemos observar que la concentración de material particulado es mayor en los meses correspondientes tanto a otoño e invierno, meses de abril, mayo, junio, julio, agosto. La humedad relativa tiene comportamiento más menos similar a lo largo del año. La velocidad del viento es mayor en los meses correspondientes a época veraniega debido a que las diferencias de presiones provocan mayor movimiento de las masas de aire. En cuanto a la temperatura claramente es mayor en los meses de primavera verano.

Para los modelos diarios se utilizaron los datos de tres años desde el mes de mayo del año 2015 hasta mayor del 2018, obteniéndose un total de 793 observaciones debido a que fue necesario eliminar los días en que el registro de alguna de las variables no estuviera registrado.

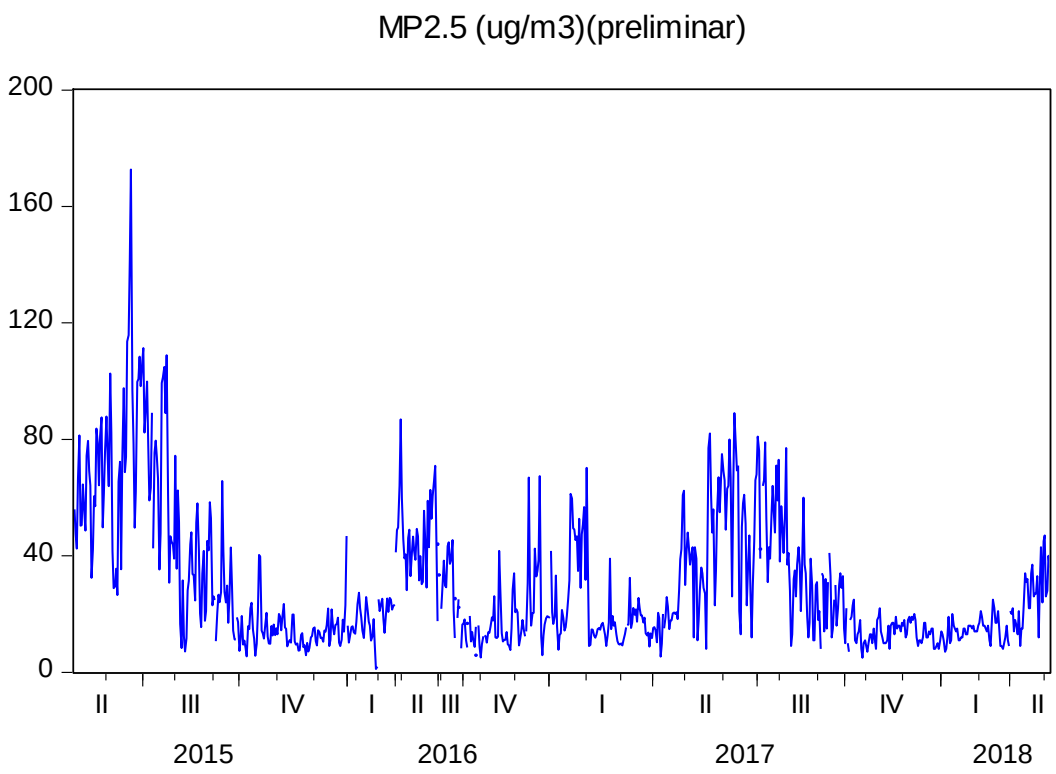

Ilustración 16:MP2.5 diario. Fuente: Elaboración propia.

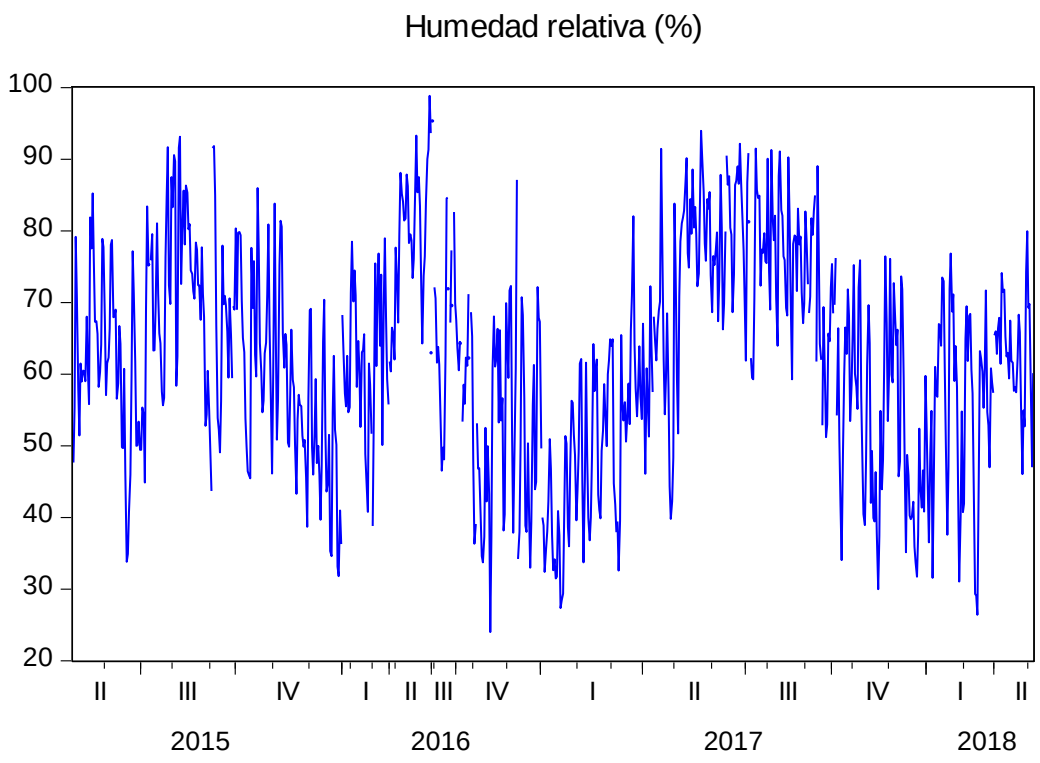

Ilustración 17: Humedad relativa. Fuente: Elaboración propia.

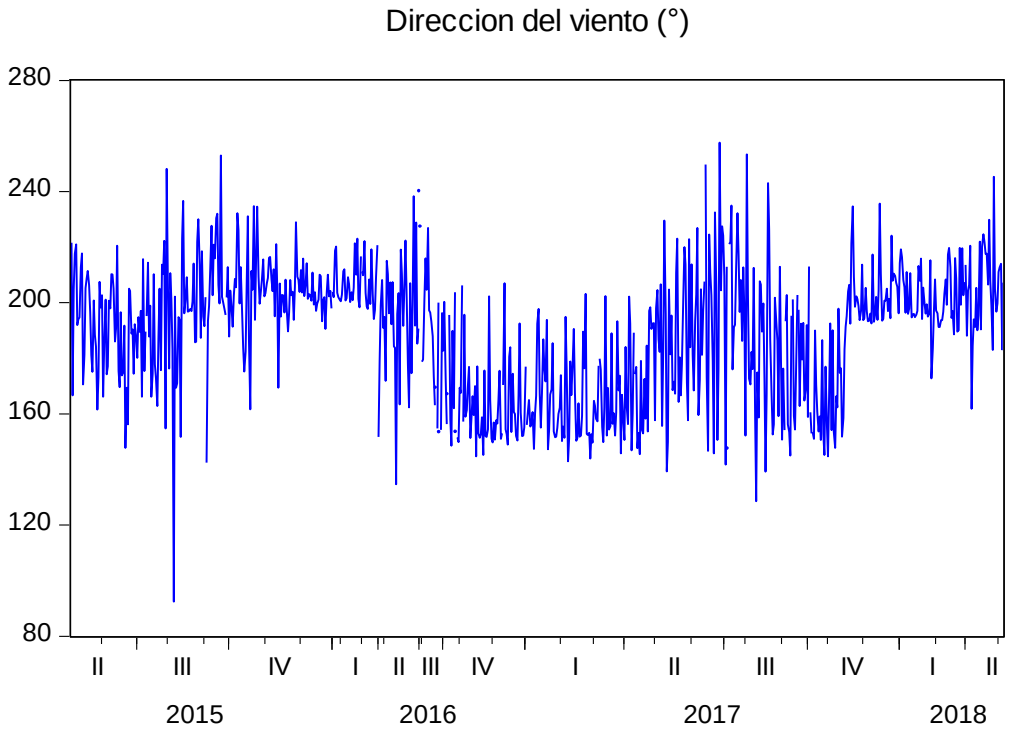

Ilustración 18: Dirección del viento diaria .Fuente: elaboración propia.

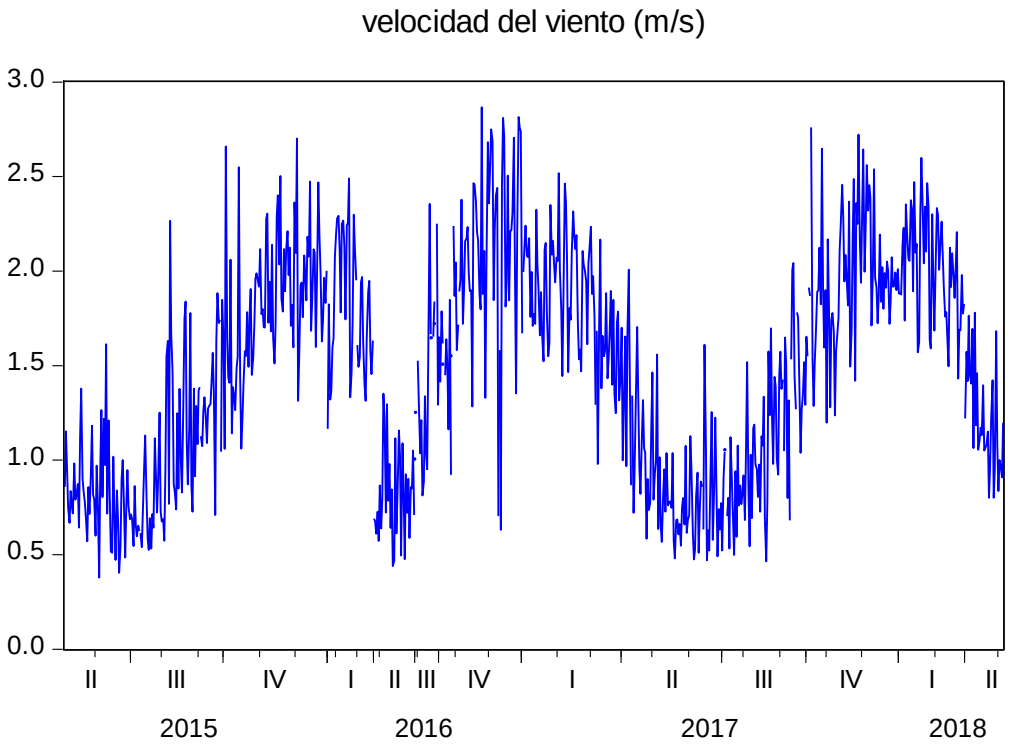

Ilustración 19: Velocidad del viento diaria .Fuente: Elaboración propia.

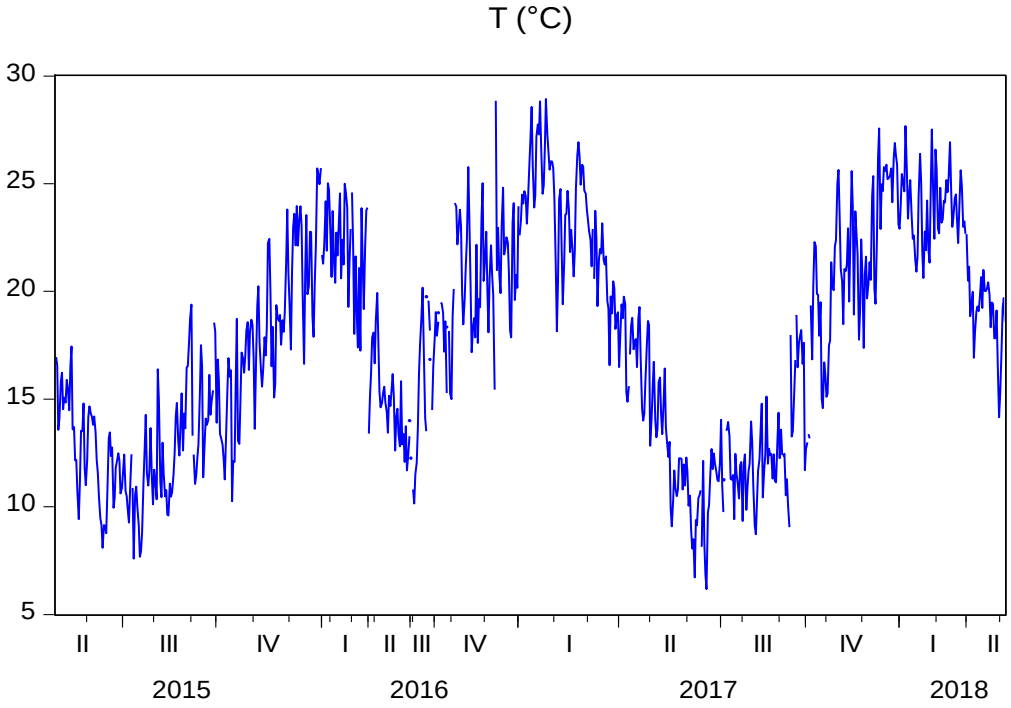

Ilustración 20: Temperatura diaria. Fuente: Elaboración propia.

Las ilustraciones 16, 17,18,19 y 20 se observa el comportamiento del material particulado, humedad relativa, dirección del viento, velocidad del viento y temperatura respectivamente desde el 2015 al 2018, observándose un comportamiento acorde a la estación del año correspondiente, por ejemplo se observa que la temperatura se eleva a finales del 4 trimestre del año y primer trimestre del año siguiente, ya que estos son los meses de verano, así como también la velocidad del viento es mayor en estos meses producto de la diferencia en las presiones de masas de aire.

4.2 Modelo econométrico propuesto

Se propone un modelo de Vectores Autorregresivos, VAR, que permitirá pronosticar con una hora de anticipación el nivel del material particulado de menos de 2,5 micrómetros en la ciudad de Santiago de Chile. Para la realización de este se tendrán como variables de adicionales la humedad relativa, la temperatura, la dirección del viento y la velocidad del viento, es decir las condiciones climatológicas del día.

El primer paso es la obtención de los datos de la sección de estadísticas de la página del Sistema de Información Nacional de Calidad del Aire SINCA, extrayendo cada serie en formato Excel para luego unificar estas en un solo archivo para conformar la base de dato; es aquí donde se realiza una revisión de datos y se eliminan aquellos días y horas en donde alguno de los valores de las variables no este medido.

El paso siguiente es comprobar la situación de las series, es decir, verificar si estas son estacionarias, debido a que esto es una característica necesaria en los modelos de vectores autorregresivos. Para esto se aplica la prueba de raíz unitaria, Dickey-Fuller aumentada.

Teniendo las series de forma estacionaria se procede a estimar el Modelo VAR, se realiza la prueba de causalidad de Granger para comprobar que las variables son necesarias para ser incluidas en el modelo.

Luego se determina la cantidad de rezagos óptimos mediante el criterio de Akaike. Con lo que finalmente se obtiene el Modelo.

Para la realización del modelo econométrico de material particulado por hora se utilizaron 5000 datos para la creación del modelo, y los datos restantes para pronostico y

cálculo de indicadores generando un resultado, es decir, utilizando una ventana móvil de 5000.

En cuanto al modelo diario propuesto, este corresponde a un VAR que permitirá pronosticar con un día de anticipación el nivel del material particulado de menos de 2,5 micrómetros, para la realización de este se tendrán como variables adicionales la humedad relativa, la temperatura, la dirección del viento y la velocidad del viento, al igual que en el modelo por hora se analiza la estacionariedad de las variables con las pruebas correspondientes.

Para la realización del modelo econométrico de material particulado por día se utilizaron 600 datos para crear el modelo y los restantes para pronósticos y comparación, es decir, utilizando una ventana móvil de 600.

4.3 Modelo híbrido propuesto

El modelo híbrido propuesto para pronosticar con una hora de anticipación la concentración de MP2.5 resulta de la combinación de un modelo de vectores autorregresivos VAR, con un modelo de red neuronal de múltiples capas, entrenada con un aprendizaje de backpropagation. De esta forma se pretende capturar el componente lineal mediante el modelo VAR y el componente no lineal mediante la Red neuronal.

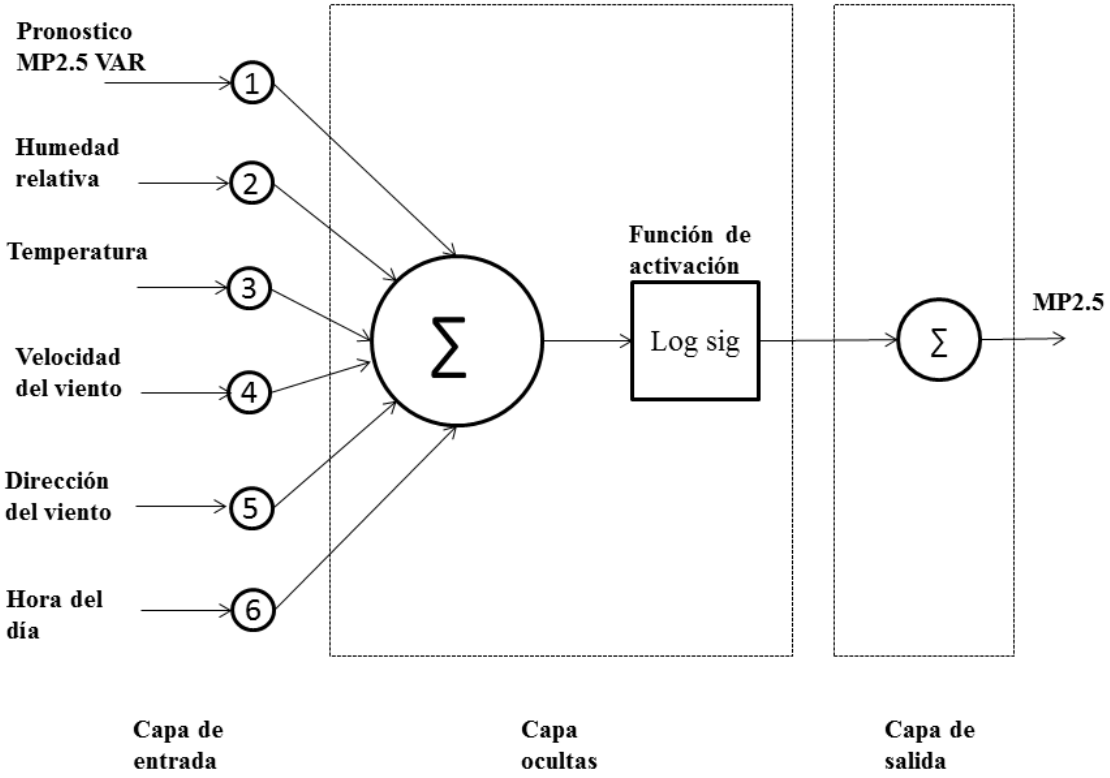

Ilustración 21: Configuración Red Neuronal MP2.5 por hora. Fuente: Elaboración propia.

La ilustración 21, muestra la configuración de la red neuronal, en esta se aprecian las condiciones meteorológicas como variables de entrada más el pronóstico generado por el modelo VAR, además se añade la variable hora como entrada adicional a la red neuronal.

El primer paso es recoger los valores entregados por el modelo VAR previamente realizado, ya que este será incluido como entrada a la red neuronal.

Para obtener el pronóstico se utilizaran como variables de entrada a la red el pronóstico del VAR, la humedad relativa, temperatura, velocidad del viento y dirección del viento, sumado a los datos autorregresivos del material particulado, además se pretende analizar el modelo hibrido incluyendo la variable hora y sin incluir esta. Debido al uso de

rezagos en el pronóstico del modelo VAR es que existe perdida de datos, por esta razón finalmente la red neuronal trabaja con 6634 registros horarios.

Se realizaran varios modelos utilizando distinta cantidad de autorregresivos utilizando2 , 5 y 10; tamaño de la ventana móvil 20, 50,100 300 y 500; cantidad de capas 1,2,3 y 5 y número de neuronas5, 10, 15, 25 y 30, incluyendo o no la variable hora, para de esta forma encontrar el mejor modelo.

Finalmente se compara el modelo hibrido con el modelo econométrico mediante contraste del error cuadrado medio (MSE) y Error Porcentual Absoluto Medio (MAPE).

Al igual que el modelo por hora, el modelo por día utiliza como entrada el pronóstico del VAR, la humedad relativa, temperatura, dirección del viento y velocidad del viendo, además de los valores pasados de la serie. Sensibilizando respecto a cantidad de autorregresivos utilizando 1, 5 y 6, utilizando como tamaño de ventana móvil 20, 30, 50 y 60 ; número de capas 1, 2, 3 y 5 y en cuanto a cantidad de neuronas 5, 10, 15, 20 y 25.

Se produce una pérdida de los primeros datos debido a los rezagos del modelo VAR quedando con 787 registros diarios como entrada a la red.

La ilustración 22, muestra la configuración de la red neuronal para pronosticar el MP2.5.

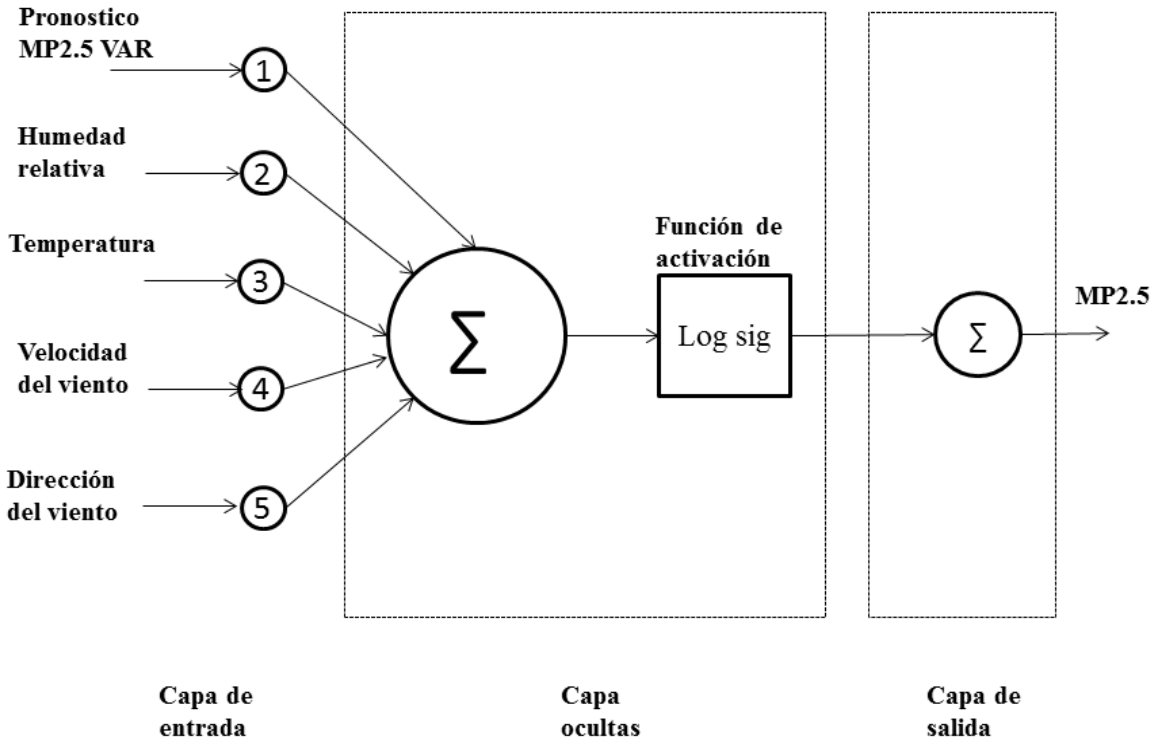

Ilustración 22: Configuración Red Neuronal MP2.5 por día Fuente: Elaboración propia.

5. Resultados

5.1 Modelo econométrico

5.1.1 Modelo econométrico Por Hora

Primeramente al realizar la prueba de Dickey-Fuller Aumentada para comprobar la estacionariedad de las variables, se obtuvo que tanto como el MP2.5 como la temperatura, humedad relativa, velocidad del viento y dirección del viento eran series estacionarias (véase anexos 1, 2, 3, 4, 5) por lo cual no fue necesario realizar ninguna modificación a estas.

Teniendo ya las series de forma estacionaria se procedió a realizar la prueba de causalidad de Granger para comprobar si una variable seria para predecir otra variable, se obtuvo que todas las variables eran necesarias para ser incluidas en el modelo acorde con lo observado en estudios acerca del tema de que los factores meteorológicos son bastante importantes para la predicción de material particulado.

Posteriormente se procedió a calcular la cantidad de rezagos necesarios para el correcto pronóstico del modelo, esto realizando la prueba de AKAIKE. Esta entrego como conclusión que la cantidad adecuada era de 27 rezagos. Lo que podría parecer un número elevado, sin embargo dado la cantidad de datos y que los registros están medidos cada una hora es un número completamente aceptable.

El pronóstico para el material particulado por hora entregado por el modelo VAR se puede apreciar en la figura ilustración 23, estos pronósticos entregan un MSE de 65,7512 y un MAPE de 0,2913, resultados que en general son bastantes buenos pronósticos.

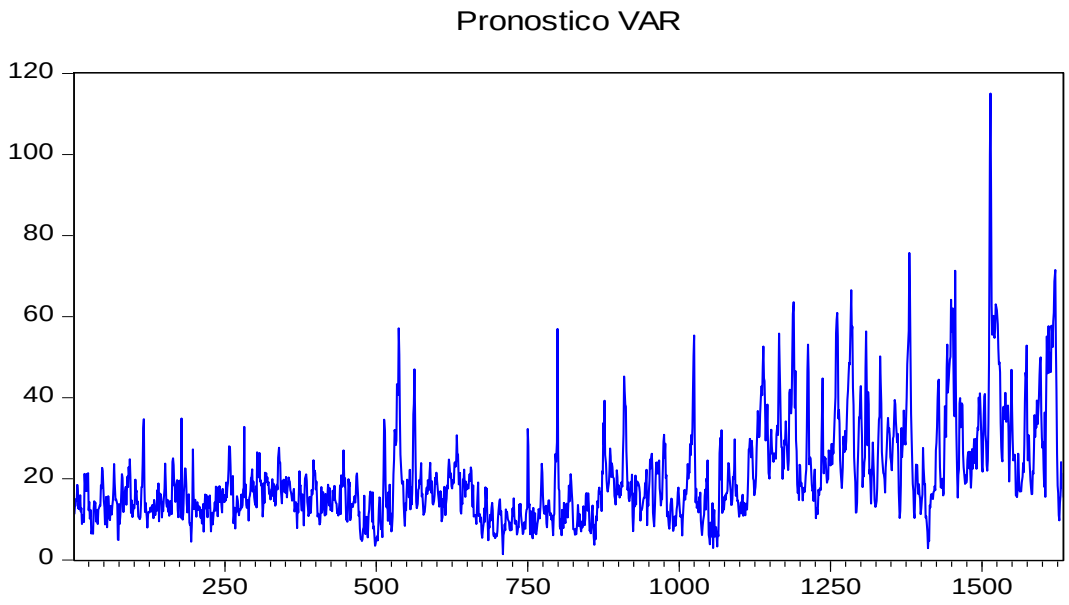

Ilustración 23: Pronósticos Modelo econométrico VAR por hora. Fuente: Elaboración propia.

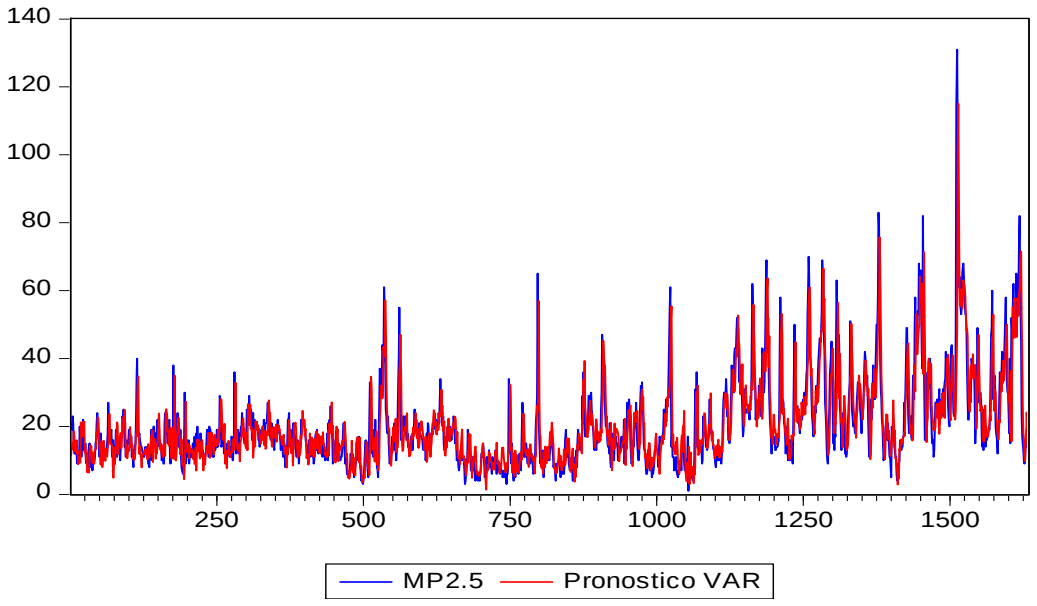

Ilustración 24: MP2.5 y Pronósticos VAR para MP2.5 por hora. Fuente: Elaboración propia.

La ilustración 24, muestra tanto el MP2.5 real diario como el pronóstico entregado por el modelo VAR, se puede visualizar que el pronóstico se asemeja bastante a los datos reales.

5.1.2 Modelo Econométrico Por Día

Del mismo modo que para el caso del modelo por hora, se realizaron las la prueba de Dickey-Fuller Aumentada para comprobar la estacionariedad de las variables, se obtuvo que tanto como el MP2.5 como la temperatura, humedad relativa, velocidad del viento y dirección del viento eran series estacionarias, por lo que no se modificaron las variables.

De acuerdo a la prueba de causalidad de Granger se obtuvo que todas las variables eran necesarias para ser incluidas en el modelo acorde con lo observado en estudios acerca del tema de que los factores meteorológicos de gran importancia a la hora de predecir el nivel de material particulado.

Posteriormente se procedió a calcular la cantidad de rezagos necesarios para el correcto pronóstico del modelo, esto realizando la prueba de AKAIKE. Esta entrego como conclusión que la cantidad adecuada era de 6 rezagos.

Luego de esto se generó el modelo y los resultados entregados por el modelo econométrico para el MP2.5 por día entrego un MSE de 64,5807 y un MAPE de 0,4193, el pronóstico se puede apreciar claramente en la ilustración 25.

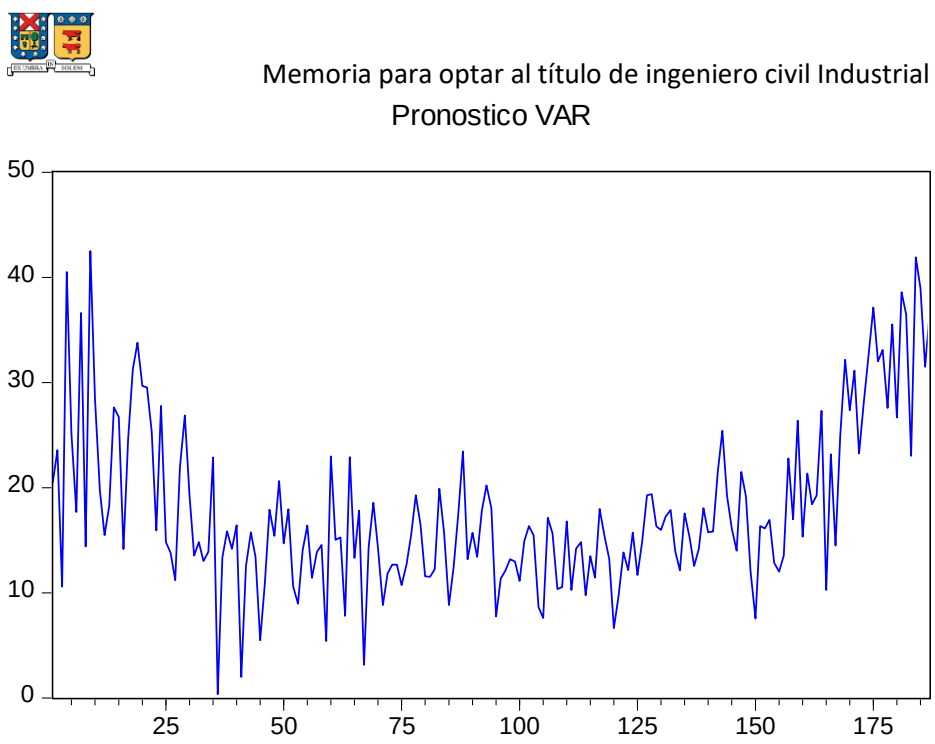

Ilustración 25: Pronósticos MP2.5 por día. Fuente: Elaboración propia.

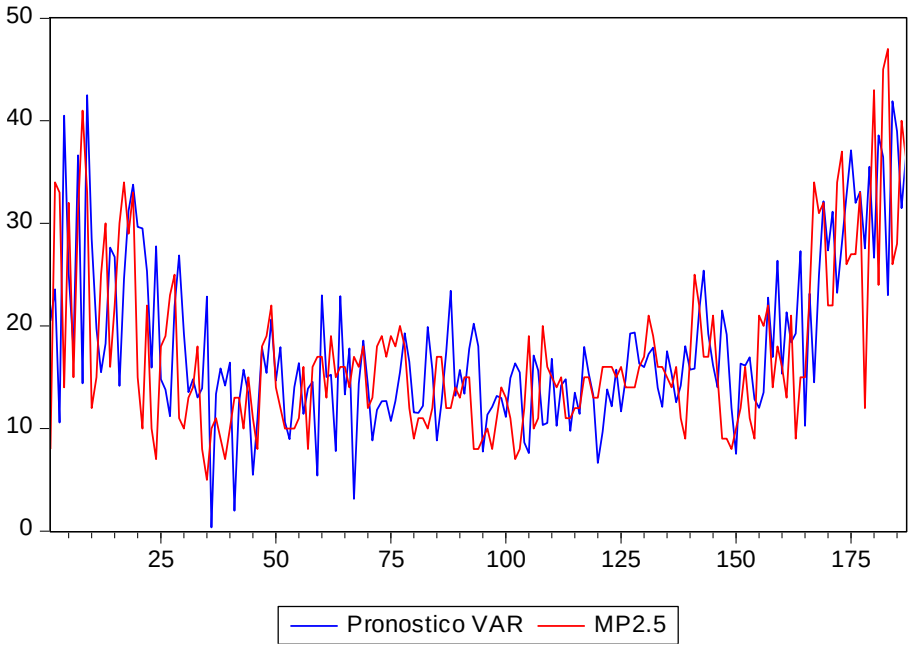

Ilustración 26: MP2.5 y Pronósticos VAR para MP2.5 por día. Fuente: Elaboración Propia.

En la figura 26, están expuestos tanto el MP2.5 real por día como el pronóstico del modelo econométrico, se puede ver de forma más clara el desfase que se produce en los pronósticos debido al gran peso que tiene el valor del día anterior.

5.2 Modelo híbrido

5.2.1 Modelo híbrido por hora

Se hicieron en total 366 modelos híbridos, existiendo una gran cantidad de modelos que entregaron malos indicadores; los mejores modelos están descritos en la tabla 1, y es posible observar que corresponden solamente a modelos que mezclan los pronósticos entregados por el modelo VAR y los datos autorregresivos del MP2.5.

Tabla 1: Mejores Modelos Híbridos para Pronóstico por hora. Fuente: Elaboración propia.

Tipo de Modelo	Autorregresivo	Ventana Móvil	Capas	Neuronas	MSE	MAPE
VAR + Autorregresivo	2	50	2	15	40,7449438	0,231265904
				20	39,8348053	0,230614201
				25	41,636458	0,236132972
			3	15	43,2431252	0,236024575
	5	100	2	25	40,6256962	0,232817973
	10	50	3	25	48,5820295	0,256827511
		100	3	15	48,0258537	0,251885909
				25	44,4175691	0,246989771

El mejor modelo híbrido encontrado corresponde a aquel con autorregresivo 2, ventana móvil de 50, 2 capas y 20 neuronas, los pronósticos pueden verse en la ilustración 27, este modelo entrega un MSE de 39,8348 y un MAPE de 0,2306.

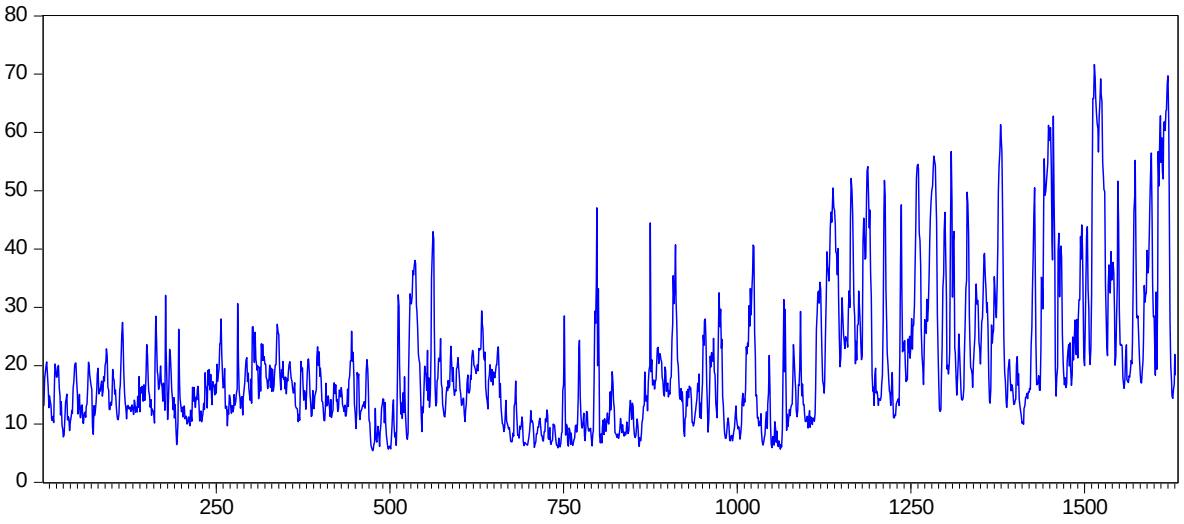

Ilustración 27: Pronóstico Modelo Híbrido autorregresivo 2, ventana móvil 50, 2 capas, 20 neuronas. Fuente: Elaboracion Propia.

En la ilustración 28, se puede apreciar tanto los valores de MP2.5 real, así como los pronósticos realizados por el mejor modelo. Es posible apreciar como el modelo presenta un buen ajuste a los datos.

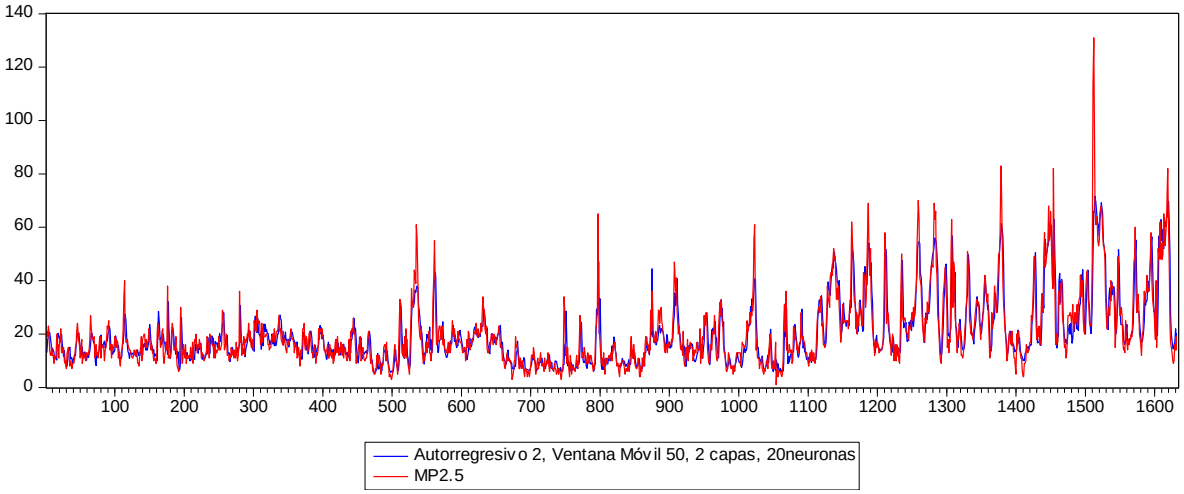

Ilustración 28: MP2.5 real por hora y Pronóstico Modelo Híbrido autorregresivo 2, ventana móvil 50, 2 capas, 20 neuronas. Fuente: Elaboración Propia.

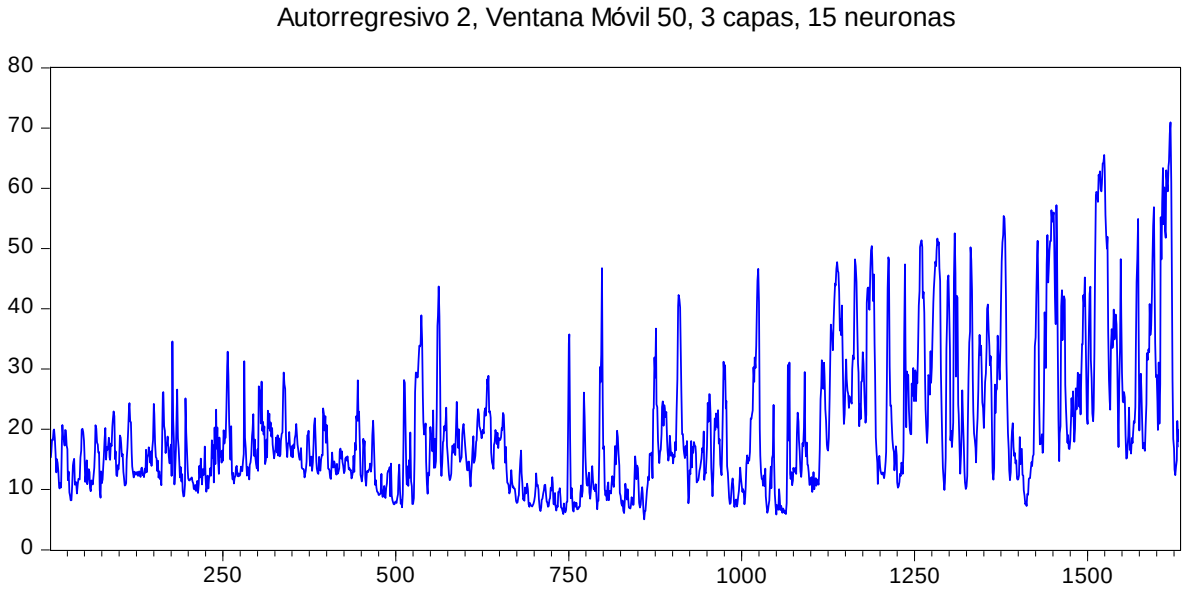

Ilustración 29: Pronóstico Modelo Híbrido autorregresivo 2, ventana móvil 50, 3 capas, 15 neuronas. Fuente: Elaboración Propia.

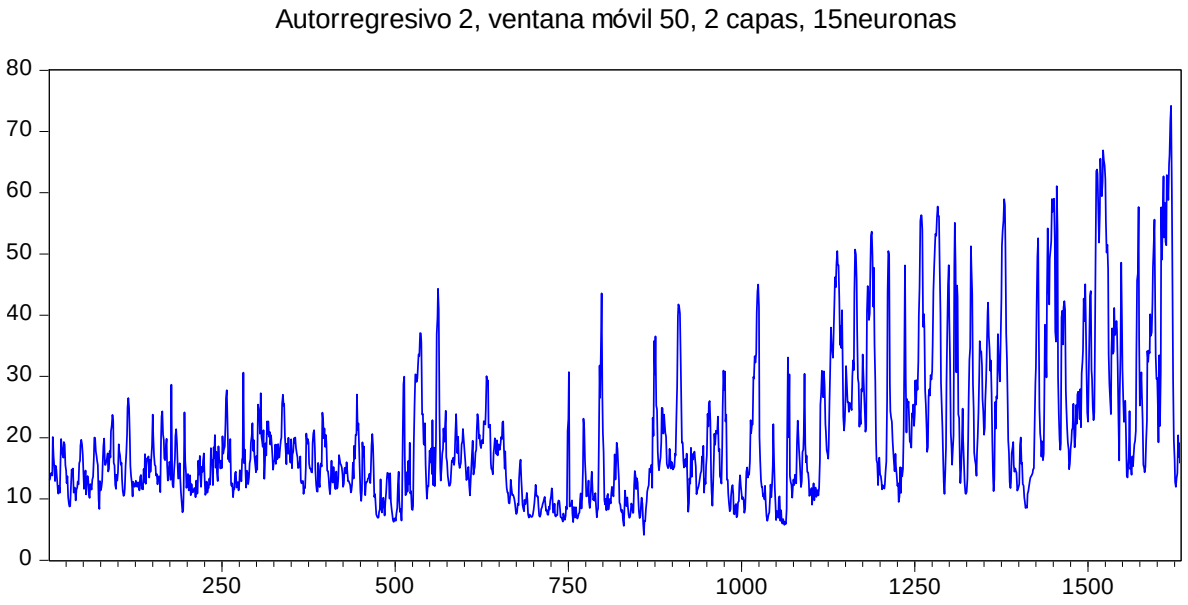

Ilustración 30: Pronóstico Modelo Híbrido autorregresivo 2, ventana móvil 50, 2 capas, 15 neuronas. Fuente: Elaboración Propia.

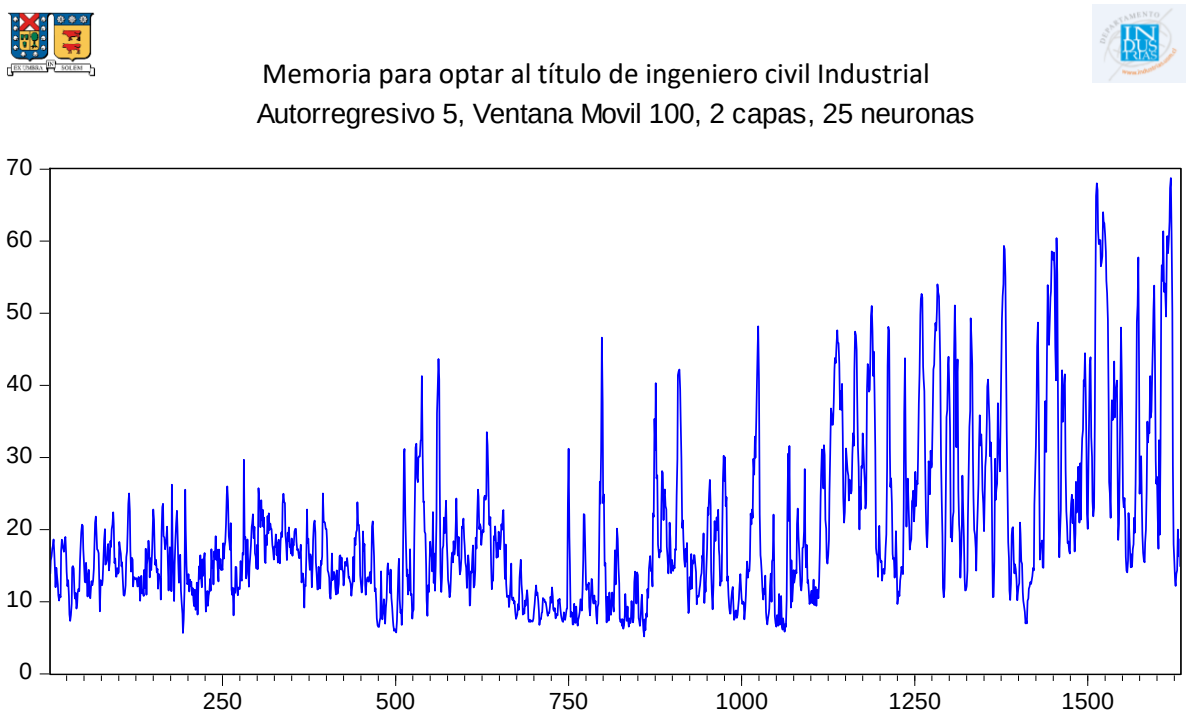

Ilustración 31: Pronóstico Modelo Híbrido autorregresivo 5, ventana móvil 100, 2 capas, 25 neuronas. Fuente: Elaboración Propia.

En las figuras 29, 30 y 31, es posible observar los pronósticos generados por otros tres modelos que entregan valor de MSE y MAPE cercanos al mejor modelo. Es importante destacar que el mejor modelo y 3 de los 4 con mejores resultados presentan 2 capas ocultas.

Se puede apreciar como existe una disminución del MSE y MAPE en el caso de todos los modelos híbridos presentados, por sobre el modelo econométrico, además en cuanto al mejor modelo horario se tiene que el MSE presenta una mejor de 65,7512 a 39,8348, esto es cerca del 39,41%.

5.2.2 Modelo híbrido por día

Se realizaron un total de 161 modelos híbridos por día, variando cantidad de autorregresivos, ventana móvil, capas, neuronas e inclusión de variables meteorológicas como se detalló en la sección 4.3. En el caso del modelo hibrido por día, la cantidad de

modelos que entregaron un buen rendimiento es considerablemente mayor que en el caso por hora, y estos pueden ser apreciados en las tablas 2, 3 y 4.

Tabla 2: Mejores Modelos Híbridos para Pronóstico por día 1. Fuente: Elaboración propia.

Tipo de modelo	Autorregresivo	Ventana Móvil	Capas	Neuronas	MSE	MAPE
Completo	1	20	1	15	46,69477388	0,33847933
			2	15	49,08209038	0,339217725
				20	48,42551154	0,347726038
				25	50,32500577	0,349784381
			3	10	49,27313061	0,348798471
				15	47,57103276	0,347755338
				25	48,63299194	0,354444215
			5	5	46,52646828	0,344357389
				15	47,57996932	0,353127438
				25	48,78322014	0,353748211

Tabla 3: Mejores Modelos Híbridos para Pronóstico por día 2. Fuente: Elaboración propia.

Tipo de modelo	Autorregresivo	Ventana Móvil	Capas	Neuronas	MSE	MAPE
Completo	6	20	1	5	47,1178342	0,342073464
				10	49,6319767	0,352082521
			2	15	46,5332461	0,344533897
				20	48,2606946	0,349406969
			3	5	46,9519246	0,340421707
				10	48,4399711	0,35166266
				15	49,7257208	0,350293854
				20	47,4173721	0,348443492
				25	48,7283177	0,352868111
			5	5	49,5969098	0,350992569
				10	48,134786	0,349128362
				20	48,217501	0,353394257
				25	48,7972872	0,35073611
		50	1	10	51,8463817	0,383092823

Tabla 4: Mejores Modelos Híbridos para Pronóstico por día 3. Fuente: Elaboración propia.

Tipo de modelo	Autorregresivo	Ventana Móvil	Capas	Neuronas	MSE	MAPE
Clima + VAR	0	20	1	10	50,49137113	0,345890414
			2	5	48,92471514	0,348998865
				15	47,04900481	0,347362439
				20	47,47751238	0,348875453
			3	5	48,51005787	0,35576931
				10	48,80263478	0,353463526
				15	49,31634455	0,351121165
				20	46,68646313	0,339546534
			5	20	47,13746543	0,346073732
				25	49,66737749	0,348124014

Es importante destacar que la mayoría de los modelos que entregaron los mejores resultados son aquellos que mezclan tanto el pronóstico entregado por el modelo VAR junto a las variables climatológicas y los autorregresivos del MP2.5, a diferencia con el pronóstico horario, los modelos que mezclan Autorregresivos más VAR para el pronóstico diario no entregan buen pronóstico, además se observa que la cantidad de capas y neuronas es variada en general en los modelos, no permitiendo ver superioridad de alguna sobre otra.

En las ilustraciones 32, 33, 35 ,37 ,38 y 39 se pueden observar los pronósticos entregados por los mejores modelos para el nivel de MP2.5 por día.

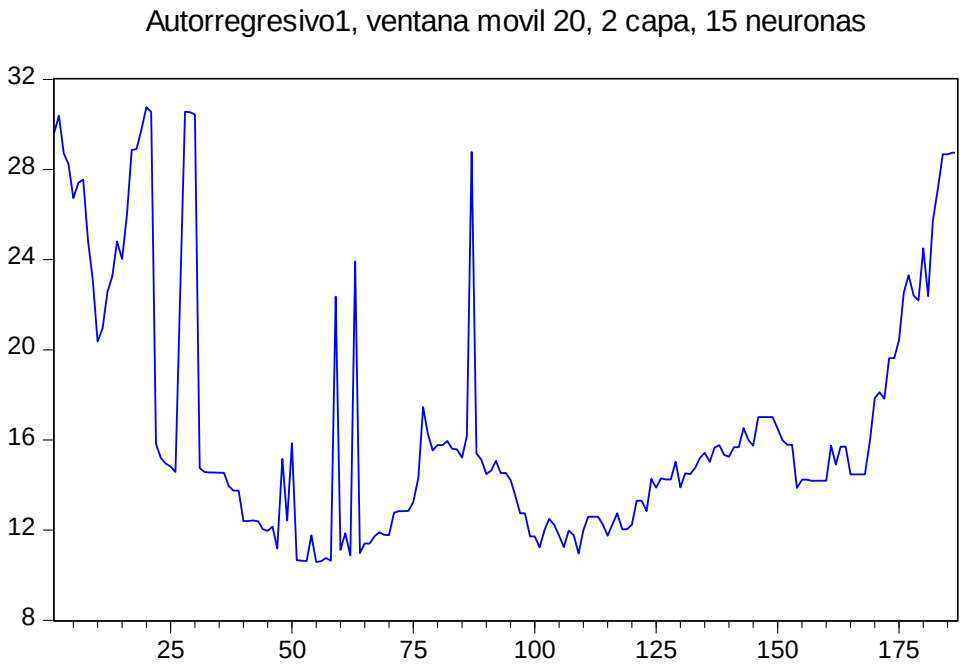

Ilustración 32: Pronóstico Modelo Híbrido autorregresivo 1, ventana móvil 20, 2 capas, 15 neuronas. Fuente: Elaboración Propia.

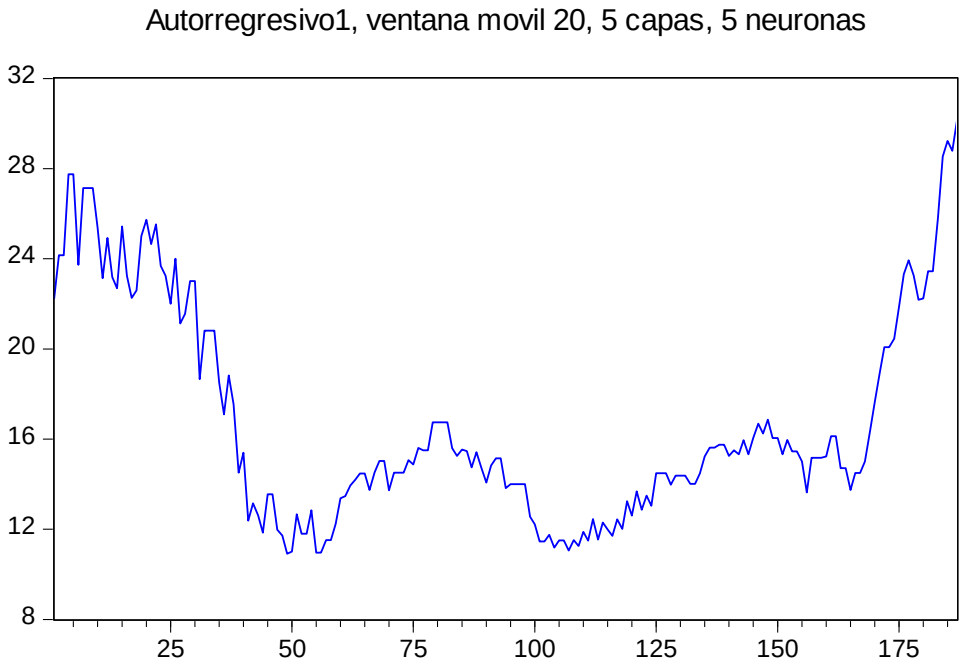

Ilustración 33: Pronóstico Modelo Híbrido autorregresivo 1, ventana móvil 20, 5 capas, 5 neuronas. Fuente: Elaboración Propia.

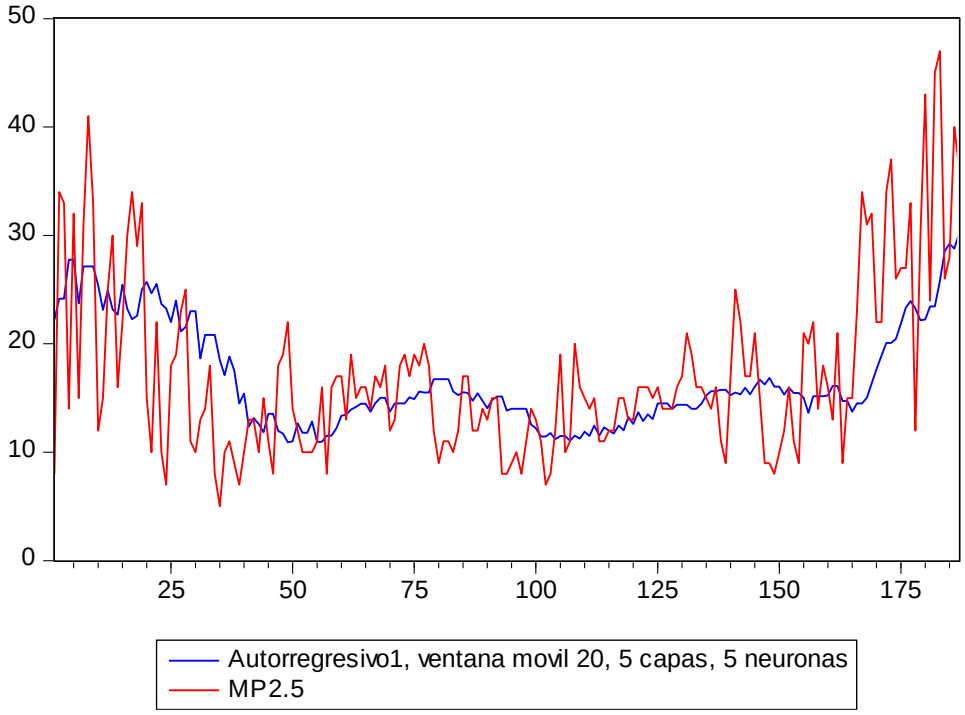

Ilustración 34: MP2.5 por Día y Pronóstico modelo híbrido autorregresivo 1, ventana móvil 20, 5 capas y 5 neuronas. Fuente: Elaboración propia.

En la ilustración 34 podemos observar el comportamiento del modelo híbrido con autorregresivo 1, ventana móvil 20, 5 capas y 5 neuronas, comparado con el nivel real de MP2.5 por día, este modelo entrega un MSE de 46,5264 y un MAPE de 0,3443, es uno de los mejores modelos presentados.

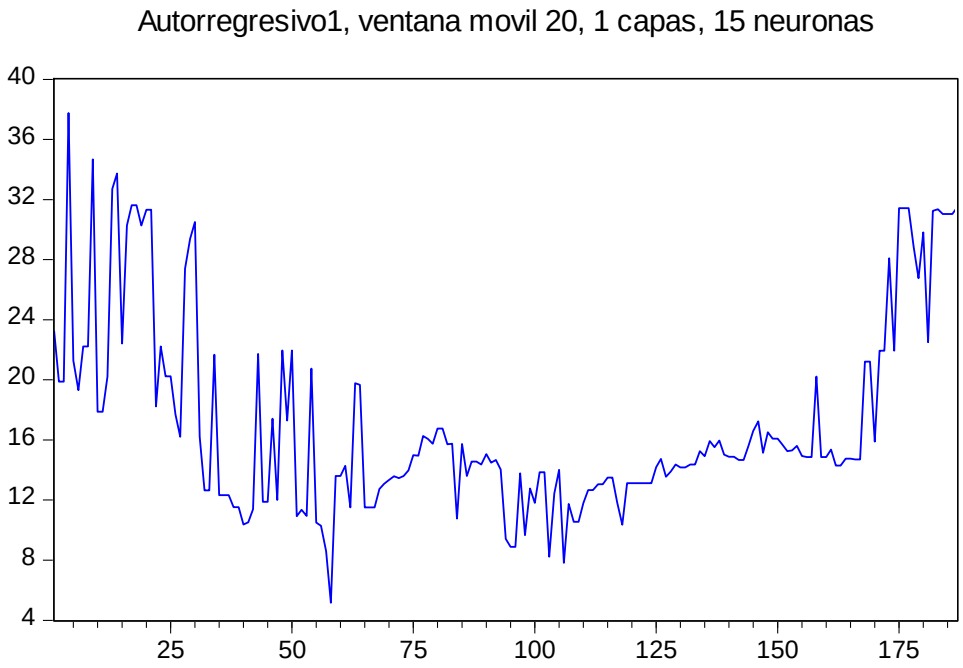

Ilustración 35: Pronóstico Modelo Híbrido autorregresivo 1, ventana móvil 20, 1 capas, 15 neuronas. Fuente: Elaboración Propia.

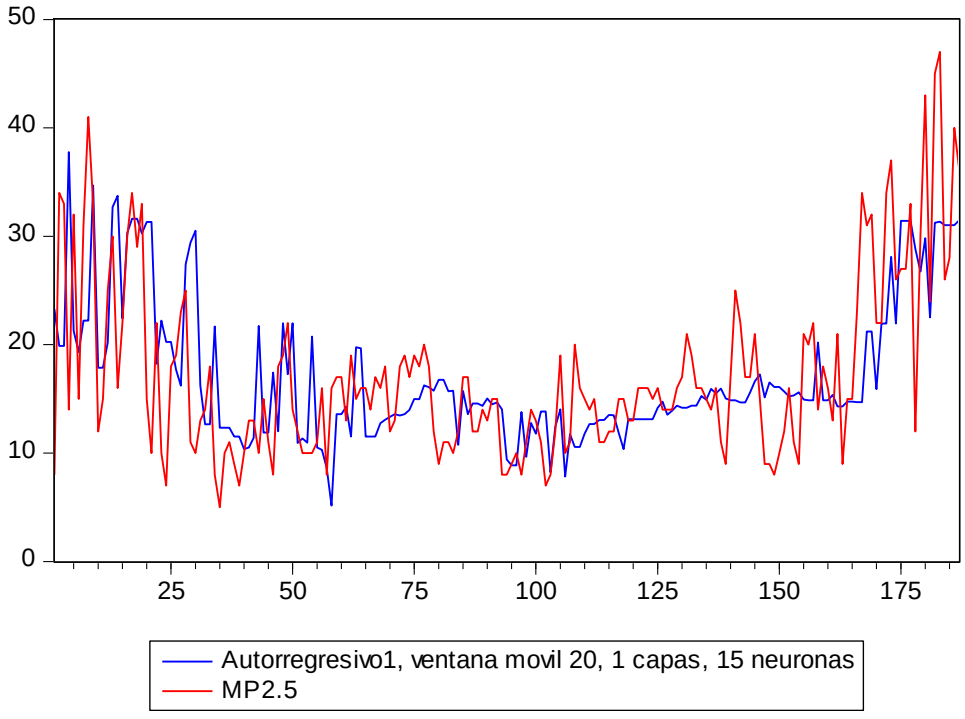

Ilustración 36: MP2.5 por Día y Pronóstico modelo híbrido autorregresivo 1, ventana móvil 20, 1 capas y 15 neuronas. Fuente: Elaboración propia.

En la figura 36 podemos observar otro de los modelos con mejores resultados, el modelo hibrido con autorregresivo 1, ventana móvil 20, 1 capa y 15 neuronas contrastado al valor real de MP2.5 por día, este modelo a diferencia del mostrado en la figura 34 tiene un comportamiento más variable, presentando valores por muy debajo de la media o muy superiores.

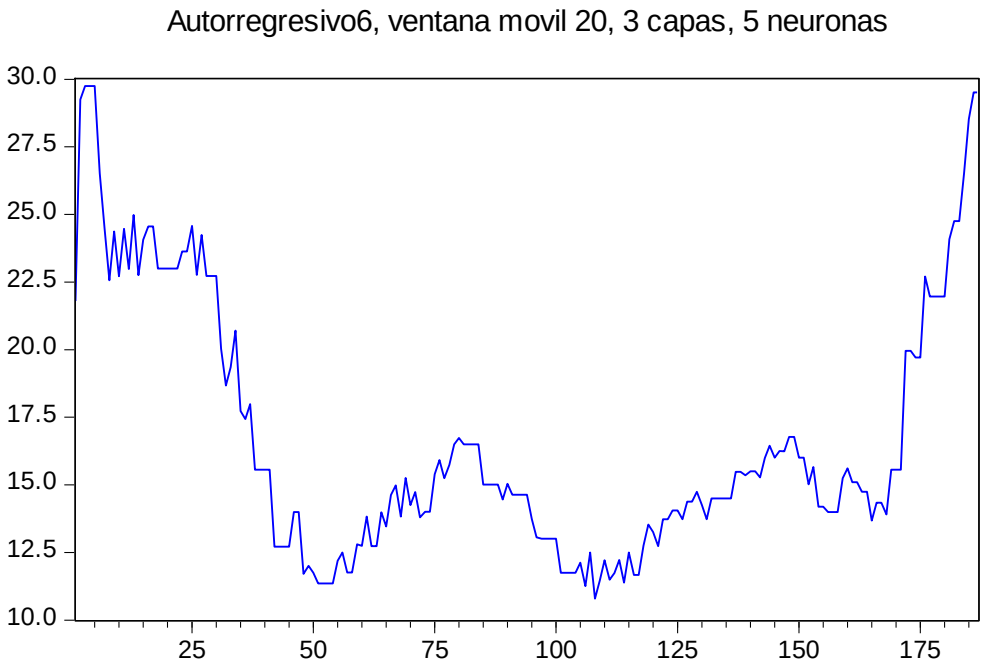

Ilustración 37: Pronóstico Modelo Híbrido autorregresivo 6, ventana móvil 20, 3 capas, 5 neuronas. Fuente: Elaboración Propia.

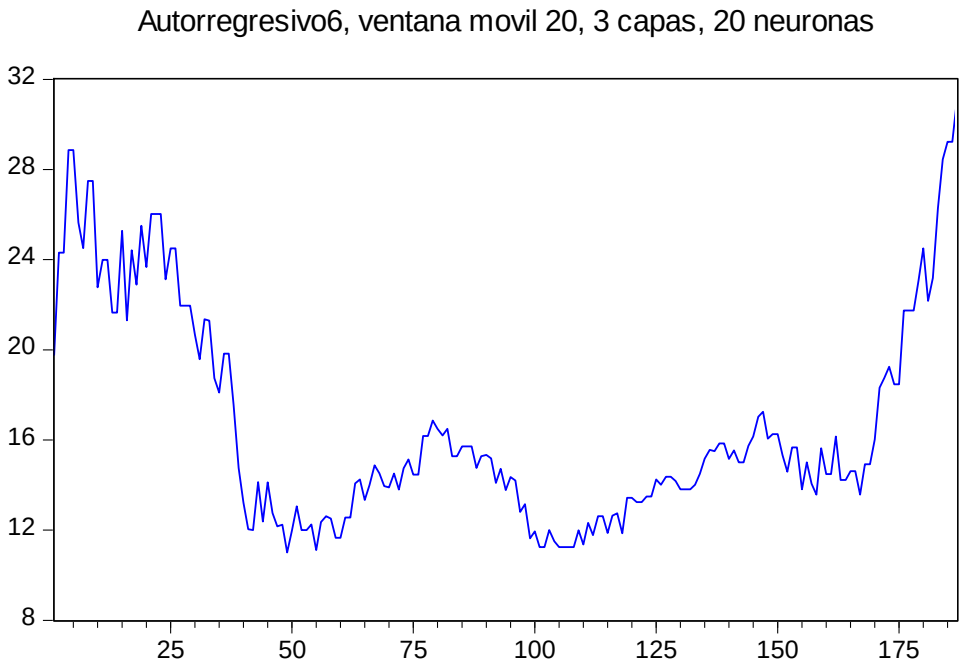

Ilustración 38: Pronóstico Modelo Híbrido autorregresivo 6, ventana móvil 20, 3 capas, 20 neuronas. Fuente: Elaboración Propia.

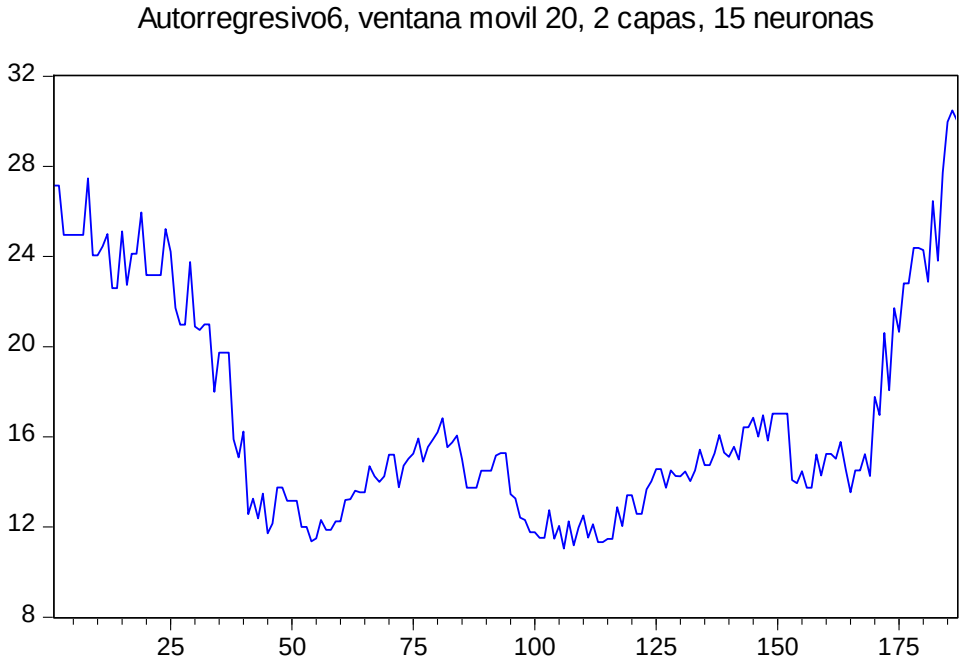

Ilustración 39: Pronóstico Modelo Híbrido autorregresivo 6, ventana móvil 20, 2 capas, 15 neuronas. Fuente: Elaboración Propia.

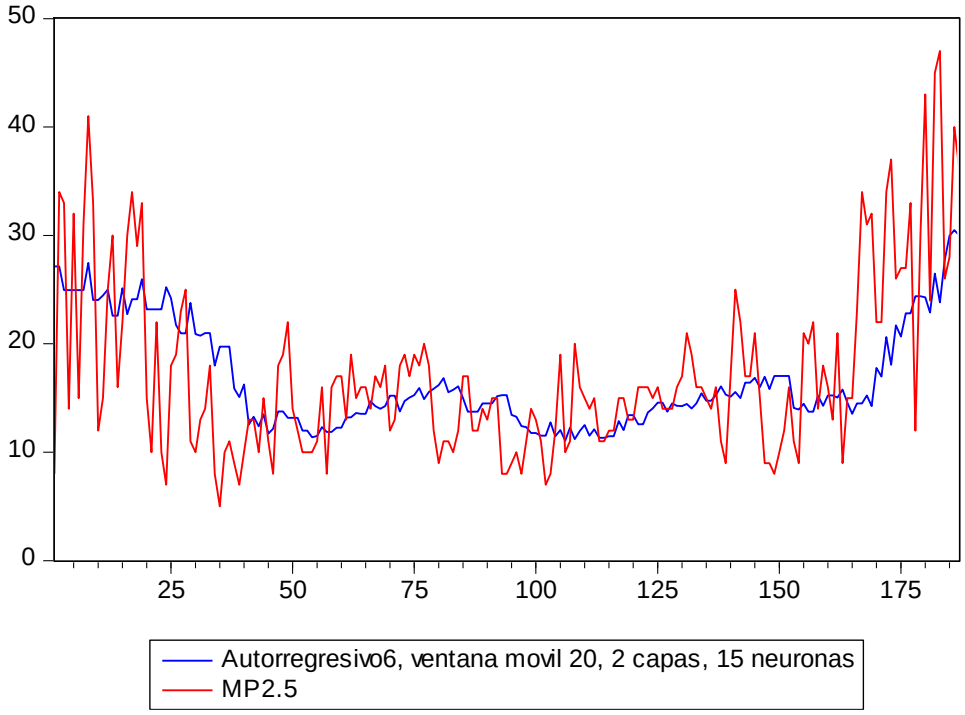

Ilustración 40: MP2.5 por Día y Pronóstico modelo híbrido autorregresivo 6, ventana móvil 20, 2capas y 15 neuronas. Fuente: Elaboración propia.

En las figuras 40, se puede apreciar el tercer mejor modelo, al igual que el modelo presentado en la figura 34, este modelo no presenta valores extremos, presenta menos variabilidad en el pronóstico.

Claramente los modelos híbridos presentados entregan mejores indicadores que el modelo econométrico, con una mejora del MSE de alrededor del 11% y del MAPE de aproximadamente un 17%.

6. Conclusiones

La contaminación medioambiental es un tema sumamente importante en la actualidad debido a la gran cantidad de repercusiones que tiene no solo en la salud de las personas, es por tal motivo que cada vez se crean más planes para prevenir y controlarlo.

El material particulado menor a 5 µm específicamente el MP2.5 es especialmente dañino ya que se deposita en los órganos y es difícil de eliminar, por esta razón se convierte en un objetivo de monitoreo constante de las centrales. En Santiago el nivel de contaminación es bastante alto producto del transporte, industrias y sus características topográficas que dificultan la correcta ventilación de la ciudad.

De acá nace la necesidad de crear modelos que permitan anticiparse a los distintos escenarios y crear medidas preventivas. Tras la realización de los modelos se obtuvo un muy buen resultado con el modelo econométrico el cual se ajustó de muy buena forma a los datos, aun así varios de los modelos híbridos presentaron mejor resultado que los modelos econométricos, probando de esta manera la eficacia de los modelos híbridos en pronósticos de diversos tipos.

La gran cantidad de variables incluidas en la red neuronal en vez de mejorar el pronóstico parecían entorpecerlo alejándolo del objetivo, esto sucedía cuando se incluían una gran cantidad de valores autorregresivos, esto puede deberse a que una gran cantidad de componentes aumentan la complejidad del modelo. Además se puede observar que la entrada de la variable hora no mejora el pronóstico, ya que el modelo que presento menor MSE y MAPE de los híbridos no contenía la variable hora como entrada.

En base a los resultados obtenidos por los modelos se puede apreciar que los mejores resultados son en aquellos modelos que utilizan una ventana móvil menor, al aumentar el tamaño de la ventana móvil no se vio ninguna mejora. Tanto en el modelo por hora como en el modelo por día ocurre de esta manera. El número de capas y neuronas presento resultados variables no pudiendo determinarse alguna como mejor que otra.

Además se puede mencionar que los modelos que incluían solo como entrada el modelo VAR más los valores autorregresivos presentaron mejor rendimiento que los modelos que incluían solamente las variables climatológicas y el pronóstico del modelo VAR para el caso del modelo horario, sin embargo los mejores modelos para pronosticar de forma diaria corresponden a aquellos que contienen tanto el pronóstico generado por el modelo VAR más los valores autorregresivos y las variables meteorológicas, no obstante es importante destacar que el nivel autorregresivo es bajo, ya que aquellos modelos con mayor nivel de datos autorregresivos tenían un rendimiento menor.

Luego de estos resultados se puede determinar que efectivamente y conforme a lo señalado en diversos estudios, los modelos híbridos permiten mejorar el pronóstico presentando por los modelos econométricos, siendo de gran utilidad y permitiendo pronósticos más acertados.

El pronóstico del material particulado puede ser especialmente útil, más aun en una ciudad tan contaminada y poblada como Santiago, además el realizar un pronóstico por día, permite tener una mirada global de lo que podría ocurrir y de esta forma anticiparse a este escenario, para posteriormente utilizando el pronóstico horario ir alertando a la población para que esta pueda tomar medidas de resguardo que protejan su salud.

Bibliografía

Alvarado Zuñiga, G. M. (2010). *Estudio integrado de factores que influyen sobre la contaminación atmosférica por material particulado respirable de pudahuel.* Santiago.

Biancofiore, F., Busilacchio, M., Verdecchia, M., Tomassetti, B., Aruffo, E., Bianco, S., . . . Di Carlo, P. (2017). Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Elsevier*, 652-659.

Cabrera , A., & Ortiz, F. (2012). Pronóstico del rendimiento del IPC (Índice de Precios y Cotizaciones) mediante el uso de redes neuronales diferenciales. *Contaduría y Administración, 57*(2), 63-81.

CENTRO DE ANÁLISIS DE POLÍTICAS PÚBLICAS. (2016). *Estado del medio ambiente en Chile.*

Chile, U. d. (2016). *ESTADO DEL MEDIO AMBIENTE EN CHILE.*

Delgadillo-Ruiz, O., Ramírez-Moreno, P., Leos-Rodríguez,, J., Salas González, J., & Valdez-Cepeda,, R. (2016). Pronósticos y series de tiempo de rendimientos de granos básicos en México. *Acta Universitaria, 26*(3), 23-32.

Elangasinghe, M., Singhal, N., Kim N, D., & Salmond, J. (2014). Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmospheric pollution research*, 696-708.

Feng, X., Li, Q., Zhu, Y., Hou,, J., Jin, L., & Wang, J. (2015). Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment, 107*, 118-128.

Gujarati, D. (2010). *Econometria.* Mexico: Mcgraw-hill.

Hidalgo, M. A. (2014). *Vectores Autorregresivos.*

Jian, L., Zhao, Y., Zhu, Y.-P., Zhang, M.-B., & Bertolatti, D. (2012). An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China. *Elsevier*, 336-345.

Kristjanpoller, W., & Minutolo, M. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model. *Expert Systems with Applications, 42*, 7245-7251.

Lopez, I., Figueroa , A., & Corrales, J. (2015). Un mapeo sistemático sobre predicción de calidad del agua mediante técnicas de inteligencia computacional. *Revista Ingenierías Universidad de Medellín*.

Lv, B., Cobourn, W., & Bai, Y. (2016). Development of nonlinear empirical models to forecast daily PM2.5 and ozone levels in three large Chinese cities. *Elsevier*, 209-223.

Matich, D. J. (2001). Redes neuronales: Conceptos basicos y aplicaciones.

Mercado Polo, D., Pedraza Caballero, L., & Martínez Gómez, E. (2015). Comparación de Redes Neuronales aplicadas a la predicción de Series de Tiempo. *PERSPECTIVA*, *13*(2), 88-95.

MMA, D. d. (2017). *Tercer Reporte del Estado del Medio Ambiente.*

Niu, M., Gan, K., Sun, S., & Li, F. (2017). Application of decomposition-ensemble learning paradigm with phase space reconstruction for day-ahead PM2.5 concentration forecasting. *Elsevier*, 110-118.

Niu, M., Wang, Y., Sun, S., & Li, Y. (2016). A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting. *Elsevier*, 168-180.

Perez, P., & Gramsch, E. (2016). Forecasting hourly PM2.5 in Santiago de Chile with emphasis on night episodes. *Elsevier*, 22-27.

Rios, G. (2008). *Series de tiempo.*

Romero, M., Diego, F., & Álvarez, M. (2006). La contaminación del aire: su repercusión como problema de salud. *Revista Cubana de Higiene y Epidemiología*, *44*(2), 1-14.

Salini Calderon, G. A. (2009). *Desarrollo de un modelo para pronosticar concentraciones extremas de PM2.5 en Santiago.* Santiago.

Salini, G., & Perez, P. (2006). ESTUDIO DE SERIES TEMPORALES DE CONTAMINACIÓN AMBIENTAL MEDIANTE TÉCNICAS DE REDES NEURONALES ARTIFICIALES. *Ingeniare*, *14*(3), 284-290.

Silva, C., Alvarado, S., Montaño, R., & Pérez, P. (2003). Modelamiento de la contaminación atmosférica por particulas: Comparación de cuatro procedimientos predictivos en Santiago,Chile. *Biomatemática XIII*, 113-127.

Valencia, M., Vanegas, J., Correa, J., & Restrepo, J. (2017). Comparación de pronósticos para la dinámica del. *Lecturas de Economia*(86), 199-230.

Anexos

Anexo1: Prueba Dickey-fuller MP2.5

Null Hypothesis: MP2_5__UG_M3__PRELIMINAR has a unit root
Exogenous: Constant
Lag Length: 15 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.856043	0.0000
Test critical values:	1% level	-3.431156	
	5% level	-2.861781	
	10% level	-2.566940	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(MP2_5__UG_M3__PRELIMINAR)
Method: Least Squares
Date: 05/30/18 Time: 08:54
Sample (adjusted): 17 6661
Included observations: 6645 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MP2_5__UG_M3__PRELIMINAR(-1)	-0.048379	0.005463	-8.856043	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-1))	0.032076	0.012823	2.501494	0.0124
D(MP2_5__UG_M3__PRELIMINAR(-2))	-0.026180	0.012760	-2.051748	0.0402
D(MP2_5__UG_M3__PRELIMINAR(-3))	-0.045858	0.012696	-3.612080	0.0003
D(MP2_5__UG_M3__PRELIMINAR(-4))	-0.013020	0.012658	-1.028591	0.3037
D(MP2_5__UG_M3__PRELIMINAR(-5))	-0.034702	0.012594	-2.755532	0.0059
D(MP2_5__UG_M3__PRELIMINAR(-6))	-0.064820	0.012556	-5.162682	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-7))	-0.057596	0.012495	-4.609374	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-8))	-0.099698	0.012382	-8.051747	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-9))	-0.071066	0.012375	-5.742844	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-10))	-0.037414	0.012326	-3.035304	0.0024
D(MP2_5__UG_M3__PRELIMINAR(-11))	-0.057128	0.012292	-4.647477	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-12))	-0.041735	0.012290	-3.395968	0.0007
D(MP2_5__UG_M3__PRELIMINAR(-13))	-0.060051	0.012251	-4.901605	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-14))	-0.059635	0.012240	-4.872157	0.0000
D(MP2_5__UG_M3__PRELIMINAR(-15))	-0.068131	0.012265	-5.554996	0.0000
C	1.296570	0.182144	7.118389	0.0000

Anexo2: Prueba Dickey-fuller Humedad Relativa

Null Hypothesis: HUMEDAD_RELATIVA____ has a unit root
Exogenous: Constant
Lag Length: 15 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.521533	0.0000
Test critical values:	1% level	-3.431156	
	5% level	-2.861781	
	10% level	-2.566940	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(HUMEDAD_RELATIVA____)
Method: Least Squares
Date: 05/30/18 Time: 09:01
Sample (adjusted): 17 6661
Included observations: 6645 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
HUMEDAD_RELATIVA____(-1)	-0.021949	0.002918	-7.521533	0.0000
D(HUMEDAD_RELATIVA____(-1))	0.531998	0.012234	43.48627	0.0000
D(HUMEDAD_RELATIVA____(-2))	-0.024023	0.013824	-1.737776	0.0823
D(HUMEDAD_RELATIVA____(-3))	0.021492	0.013792	1.558267	0.1192
D(HUMEDAD_RELATIVA____(-4))	-0.037767	0.013765	-2.743768	0.0061
D(HUMEDAD_RELATIVA____(-5))	-0.021065	0.013759	-1.531021	0.1258
D(HUMEDAD_RELATIVA____(-6))	-0.072326	0.013734	-5.265999	0.0000
D(HUMEDAD_RELATIVA____(-7))	-0.057842	0.013738	-4.210354	0.0000
D(HUMEDAD_RELATIVA____(-8))	-0.054620	0.013727	-3.979023	0.0001
D(HUMEDAD_RELATIVA____(-9))	-0.046254	0.013710	-3.373874	0.0007
D(HUMEDAD_RELATIVA____(-10))	-0.045032	0.013671	-3.293889	0.0010
D(HUMEDAD_RELATIVA____(-11))	-0.028358	0.013672	-2.074142	0.0381
D(HUMEDAD_RELATIVA____(-12))	-0.051681	0.013657	-3.784154	0.0002
D(HUMEDAD_RELATIVA____(-13))	-0.057909	0.013672	-4.235588	0.0000
D(HUMEDAD_RELATIVA____(-14))	-0.072022	0.013679	-5.265258	0.0000
D(HUMEDAD_RELATIVA____(-15))	-0.123011	0.012190	-10.09111	0.0000
C	1.401014	0.191815	7.303982	0.0000

Anexo3: Prueba Dickey-fuller Direccion del viento

Null Hypothesis: DIRECCION_DEL_VIENTO____ has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on SIC, maxlag=34)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-38.19985	0.0000
Test critical values:	1% level	-3.431154	
	5% level	-2.861780	
	10% level	-2.566940	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DIRECCION_DEL_VIENTO____)
Method: Least Squares
Date: 05/30/18 Time: 09:04
Sample (adjusted): 4 6661
Included observations: 6658 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DIRECCION_DEL_VIENTO____(-1)	-0.629998	0.016492	-38.19985	0.0000
D(DIRECCION_DEL_VIENTO____(-1))	-0.070367	0.014928	-4.713674	0.0000
D(DIRECCION_DEL_VIENTO____(-2))	-0.051009	0.012243	-4.166481	0.0000
C	121.2768	3.282647	36.94482	0.0000

R-squared	0.346372	Mean dependent var	0.000381
Adjusted R-squared	0.346077	S.D. dependent var	84.14745
S.E. of regression	68.04624	Akaike info criterion	11.27885
Sum squared resid	30809957	Schwarz criterion	11.28294
Log likelihood	-37543.30	Hannan-Quinn criter.	11.28026
F-statistic	1175.367	Durbin-Watson stat	2.002720
Prob(F-statistic)	0.000000		

Anexo4: Prueba Dickey-fuller Velocidad del viento

Null Hypothesis: VELOCIDAD_DEL_VIENTO__M_ has a unit root
Exogenous: Constant
Lag Length: 15 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.39671	0.0000
Test critical values:	1% level	-3.431156	
	5% level	-2.861781	
	10% level	-2.566940	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(VELOCIDAD_DEL_VIENTO__M_)
Method: Least Squares
Date: 05/30/18 Time: 09:05
Sample (adjusted): 17 6661
Included observations: 6645 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
VELOCIDAD_DEL_VIENTO__M_(-1)	-0.127415	0.009511	-13.39671	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-1))	0.084738	0.014024	6.042340	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-2))	-0.018465	0.013791	-1.338915	0.1806
D(VELOCIDAD_DEL_VIENTO__M_(-3))	0.032699	0.013487	2.424410	0.0154
D(VELOCIDAD_DEL_VIENTO__M_(-4))	0.042444	0.013250	3.203322	0.0014
D(VELOCIDAD_DEL_VIENTO__M_(-5))	0.052678	0.013048	4.037158	0.0001
D(VELOCIDAD_DEL_VIENTO__M_(-6))	-0.017603	0.012912	-1.363272	0.1728
D(VELOCIDAD_DEL_VIENTO__M_(-7))	-0.026390	0.012706	-2.076973	0.0378
D(VELOCIDAD_DEL_VIENTO__M_(-8))	-0.064730	0.012486	-5.184305	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-9))	-0.060824	0.012369	-4.917385	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-10))	-0.035283	0.012264	-2.876828	0.0040
D(VELOCIDAD_DEL_VIENTO__M_(-11))	-0.058538	0.012238	-4.783245	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-12))	-0.079412	0.012213	-6.502182	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-13))	-0.107397	0.012201	-8.802232	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-14))	-0.096964	0.012141	-7.986295	0.0000
D(VELOCIDAD_DEL_VIENTO__M_(-15))	-0.112734	0.012187	-9.250245	0.0000
C	0.189318	0.015303	12.37093	0.0000

Anexo5: Prueba Dickey-fuller Temperatura

Null Hypothesis: T___C_ has a unit root
Exogenous: Constant
Lag Length: 15 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.639971	0.0001
Test critical values:	1% level	-3.431156	
	5% level	-2.861781	
	10% level	-2.566940	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(T___C_)
Method: Least Squares
Date: 05/30/18 Time: 09:08
Sample (adjusted): 17 6661
Included observations: 6645 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
T___C_(-1)	-0.007815	0.001684	-4.639971	0.0000
D(T___C_(-1))	0.594493	0.012194	48.75225	0.0000
D(T___C_(-2))	0.007998	0.014138	0.565681	0.5716
D(T___C_(-3))	-0.020995	0.014127	-1.486158	0.1373
D(T___C_(-4))	-0.075364	0.014119	-5.337585	0.0000
D(T___C_(-5))	-0.035915	0.014137	-2.540600	0.0111
D(T___C_(-6))	-0.071504	0.014124	-5.062501	0.0000
D(T___C_(-7))	-0.058097	0.014134	-4.110567	0.0000
D(T___C_(-8))	-0.045683	0.014139	-3.231026	0.0012
D(T___C_(-9))	-0.052826	0.014127	-3.739334	0.0002
D(T___C_(-10))	-0.054613	0.014108	-3.871172	0.0001
D(T___C_(-11))	-0.036722	0.014113	-2.601996	0.0093
D(T___C_(-12))	-0.033378	0.014083	-2.370108	0.0178
D(T___C_(-13))	-0.038398	0.014085	-2.726206	0.0064
D(T___C_(-14))	-0.111358	0.014097	-7.899211	0.0000
D(T___C_(-15))	-0.125352	0.012195	-10.27898	0.0000
C	0.140912	0.031568	4.463750	0.0000