

Forecasting and Economic Impacts of PM 2.5 in Santiago, Chile

Stella Moisan^{a,*}, Rodrigo Herrera^a, Adam Clements^b

^a*Facultad de Economía y Negocios, Universidad de Talca, Chile*

^b*Business School, Queensland University of Technology, Australia*

Abstract

La importancia de predecir el nivel de contaminantes atmosféricos presentes en el aire recae tanto en los efectos que tienen en la salud humana como en las consecuencias económicas que esto genera. Actualmente, el pronóstico de calidad del aire en Chile se basa en el modelo Cassmassi desarrollado en 1999, pero su ajuste de predicción es muy bajo, registrando un 80% de falsas predicciones. En esta investigación, se presentan resultados del modelamiento y predicción de concentraciones horarias de PM2.5 para 11 estaciones en Santiago, Chile. Se muestra que un modelo de ecuaciones múltiples dinámico, como aproximación lineal, tiene el potencial de igualar o incluso superar la precisión del actual modelo utilizado (Cassmassi), otros modelos de pronóstico lineales (SARIMAX) y no lineales (Red Neuronal). Dentro de los resultados más importantes resalta la interpretación lineal de sus parámetros así como también su extrapolación a los impactos económicos directos provocados por las PM2.5.

Keywords: Air quality, Pollution, Linear Model, Economic Impacts.

1. Introducción

La OCDE en 2014 situó a Chile como el país más contaminado entre los 36 que conforman esta organización. Mientras que Santiago de Chile es la 4ta ciudad más contaminada del continente americano (WHO, 2011). Entre estos contaminantes se encuentra el PM10 y PM2.5 (Material Particulado con diámetro menor a 10 y 2.5 $\mu\text{g}/\text{m}^3$, respectivamente) los cuales son claves para medir la calidad del aire (OMS, 2006). En Chile, más de 10 millones de habitantes están expuestos a concentraciones de PM2.5 superiores a 15 $\mu\text{g}/\text{m}^3$, valor correspondiente a la norma anual vigente en EE.UU, mientras que el límite recomendado según la OMS es de 10 $\mu\text{g}/\text{m}^3$ (Cifuentes, 2010). Del mismo lado, 4000 personas apróx. mueren prematuramente al año por la exposición crónica de este contaminante (MMA, 2011).

Cassmassi (1999) desarrolló el modelo actual de pronóstico de calidad del aire utilizado en Chile. En Santiago se implementa en 11 estaciones, sin embargo, el desempeño del modelo en este lugar registra un 80% apróx. de falsas predicciones en preemergencias de PM10 (Perez and Reyes, 2006). La predicción e

*Corresponding author.

Email address: stellamoisanrodriguez@gmail.com (Stella Moisan)

impacto económico de PM2.5 es importante porque este contaminante tiene efectos más severos a la salud que partículas de tamaños más grandes. Además, Bohringer and Jochem (2007) exponen que la calidad ambiental, integrada al sentido económico y social es de importancia crítica para el desarrollo sostenible de las naciones. En este sentido, los beneficios económicos netos de normar de manera efectiva el PM2.5 en Chile se estiman en MM US\$7100 (SINIA, 2010).

En los últimos años, existe un interés creciente por el desarrollo de modelos predictivos de PM10 y PM2.5. Se distinguen dos grandes familias: estadísticos y deterministas, aunque diferentes autores sugieren que una combinación de ellos pudiese ser la mejor opción (Konovalov et al., 2009; Neal et al., 2014). Diversos autores exponen que los modelos estadísticos tienen en general un mayor poder predictivo en comparación con los modelos deterministas. Por ejemplo, Hrust et al. (2009) postula que estos modelos son más apropiados para capturar dependencias entre concentraciones de contaminantes atmosféricos y predictores potenciales.

Además, los modelos estadísticos se distinguen entre lineales y no lineales. En modelos lineales, se encuentran metodologías de Filtro de Kalman (Zolghadri and Cazaurang, 2006; Hoi et al., 2008) y modelos de Regresión Lineal Múltiple (RLM) (Stadlober et al., 2008; Genc et al., 2010). Por ejemplo, Perez and Reyes (2002); Perez and Gramsch (2016) desarrollan un modelo de RLM para predecir el promedio máximo diario de PM10 y PM2.5, respectivamente en Santiago, Chile. Exponen que la mejora en el modelo es posible mediante la inclusión de variables meteorológicas como predictores potenciales y que esto es más importante que el tipo de modelo utilizado. Análogamente, dentro de los modelos que utilizan FK destacan Zolghadri and Cazaurang (2006), los cuales utilizan una versión extendida para estimar la predicción de la media diaria de concentraciones de PM10 en Burdeos, Francia. Los resultados experimentales muestran que el modelo predice con exactitud la media diaria de PM10.

Por otro lado, la flexibilidad de los modelos no lineales ha inspirado diferentes metodologías. Por ejemplo, Kukkonen et al. (2003) proponen cinco modelos de Red Neuronal (RN) para la predicción de concentraciones de PM10 y NO2 en Helsinki, Finlandia, y lo comparan con un modelo lineal y otro determinístico. Sus resultados muestran que los modelos de RN entregan mejores resultados en términos de predicción. Hooyberghs et al. (2005) mejoran la calidad de predicción de concentración promedio diaria de PM10 utilizando información atmosférica como input dentro de una RN. Concluyen que las condiciones meteorológicas juegan un rol importante en las fluctuaciones diarias de PM10 en Bélgica.

En Chile, Perez and Gramsch (2016) utilizan RN con información de datos históricos de PM10 y PM2.5, algunas variables meteorológicas, y concentraciones de estaciones de monitoreo cercanas. Para las estimaciones utilizan información hasta las 19 hrs. y predicen hasta 15 hrs. después sólo durante períodos nocturnos en invierno. Exponen que este modelo puede ser considerado como una herramienta útil para anticipar episodios de crisis asociados a la calidad del aire. En resumen, la literatura muestra que, si bien los modelos no lineales logran un mejor ajuste en predicción (Sun et al., 2013; Stern et al., 2008; Shad et al., 2009), sus parámetros pierden interpretabilidad lineal (Zhou et al., 2000).

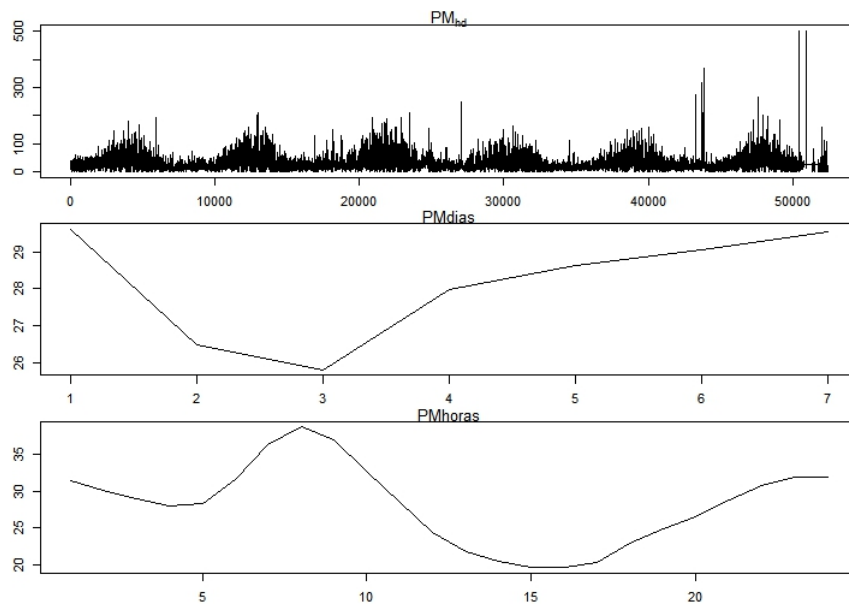


Figure 1: Desde panel superior a inferior: serie temporal horaria de PM_{2.5}, promedio por horas del día, y por días de la semana en un año, para Estación Cerrillos, Santiago, Chile desde Ene. 2008 a Dic. 2015.

El objetivo central de este trabajo consta de dos partes: (i) demostrar que el modelo de ecuaciones múltiples con coeficientes dinámicos tiene el potencial de lograr un pronóstico competitivo de PM_{2.5}, en relación a modelos planteados en la literatura, en la Región Metropolitana de Santiago, Chile; (ii) proveer un sentido económico a los coeficientes utilizados en el modelo para estudiar el impacto económico que puede ser alcanzado por un modelo preventivo de episodios críticos de contaminación atmosférica (alarma, preemergencia, y emergencia). Sección 2 analiza los factores estilizados de la serie de tiempo de PM_{2.5}. Sección 3 introduce la metodología propuesta, mientras que Sección 4 el primer modelo de comparación. Sección 5 muestra los resultados preliminares de estimación y predicción de los modelos. Finalmente, Sección 6 exhibe algunas conclusiones preliminares. Este resumen corresponde a un trabajo en desarrollo que tendrá resultados finales antes de que tome lugar SECHI 2016.

2. Descripción de los Datos

Para desarrollar el modelo se utilizan observaciones históricas horarias en 11 estaciones de monitoreo ubicadas en la Región Metropolitana de Santiago, Chile. Se trabaja con un periodo muestral desde el 01 Enero 2008 a 31 Diciembre 2014 extraídos desde SINCA (Sistema de Información Nacional de Calidad del Aire), organismo dependiente del Ministerio de Medio Ambiente de Chile.

Fig. 1 presenta la serie temporal horaria de PM_{2.5} para la muestra completa en Estación Cerrillos y la gráfica del promedio horario y diario del contaminante, respectivamente, al año 2008.

Al analizar las características de la serie temporal de PM2.5, se observan patrones estacionales diarios, mensuales y anuales. Por ejemplo, mensuales debido a que existe mayor contaminación en meses de otoño-invierno. De lo anterior, a partir del 01 de Abril al 31 de Agosto la Región Metropolitana emplea medidas de mitigación de PM2.5 a través de la Gestión de Episodios Críticos (GEC), distinguiendo niveles de Alerta/Preemergencia/Emergencia (MMA, 2010). Por otro lado, existe correlación intradiaria. La 1 muestra concentraciones según horas del día. La mayor concentración promedio horaria se observa entre 06:00 y 10:00, correspondiente al periodo en el que la población comienza su jornada laboral, por tanto es mayor la afluencia vehicular. Se observan patrones semanales y diarios, el mayor peak diario en la concentración del contaminante es Sábado, Domingo y Lunes, con diferencias significativas entre mitad de semana, ej. Miércoles, y fines de semana.

La Fig. 2 muestra la correlación entre PM2.5 y series de tiempo de Monóxido de Carbono (CO), Temperatura ($Temp$), Velocidad del Viento (VV) y Humedad Relativa (HR), como posibles variables a incluir.

3. Sistema de Ecuaciones Múltiples

Para capturar las características y variabilidad de las PM2.5, se propone un sistema de ecuaciones múltiples (SEM) que entrega 24 ecuaciones, una por cada hora h del día d . La importancia del modelo propuesto radica en que, al ser lineal en parámetros, sus variables continúan siendo fácilmente interpretables. Entonces, la ecuación 1 presenta el sistema de ecuaciones múltiples para la concentración de PM2.5, PM_{hd}

$$PM_{hd} = \theta_h^0 + \theta_{hd}^1 PM_{hd-1} + \theta_h^2(t) PM_{hd-7} + \theta_{hd}^3 PM_{h-1d} + \Phi_{h1} \varepsilon_{hd-1} + \Phi_{h2} \varepsilon_{hd-7} + \varepsilon_{hd} + \sum_{k=1}^K \delta_{hk} Z_{h-1d}^K \quad (1)$$

donde $h = 1, \dots, 24$ y ε_{hd} es el término de error. El proceso autorregresivo y de medias móviles del modelo cuentan con uno y siete días de retardo de PM2.5, PM_{hd-1} y PM_{hd-7} , y de los errores, ε_{hd-1} , ε_{hd-7} , respectivamente. Debido a que las ecuaciones horarias cuentan con retardos diarios, PM_{h-1d} se incluye como variable exógena, y explica que la concentración del contaminante depende del existente con una hora de retardo.

Luego, el coeficiente θ^1 incluye la estacionalidad semanal, mediante variables dicotómicas según días de la semana, y θ^2 la estacionalidad anual, según ciclos mediante una Serie de Fourier con 8760 observaciones (datos horarios por días en un año):

$$\theta_{hd}^1 = \sum_{p=1}^7 \eta_{hp} W_{dp} \quad (2)$$

$$\theta_h^2(t) = \tau_{h1} + \sum_{q=1}^m \left[\tau_{h2q} \sin \left(2q\pi \left(\frac{t}{8760} \right) \right) + \tau_{h3q} \cos \left(2q\pi \left(\frac{t}{8760} \right) \right) \right] \quad (3)$$

Además, siguiendo a autores como Hien et al. (2002) es posible utilizar algunas variables metereológicas (Humedad Relativa, “*HR*”, Velocidad del Viento, “*VV*”, Dirección del Viento, “*DV*” y Temperatura, “*Temp*”) y ambientales (PM10, SO2, NO2, NOX, NO, CO, O3) como predictores dentro del modelo. Acorde a la literatura, las covariables incluidas son *HR*, *Temp* y *CO* y $PM \max_{hd-1}$ como el valor horario máximo de la variable PM2.5 en el día anterior. Lo anterior, dado que autores como Saide et al. (2011) comprueban que la correlación entre CO y PM2.5 en Santiago es alta, incluso sobre 0.95 durante la noche, por tanto es relevante su inclusión en el modelo. Mientras que Zhou et al. (2014) exponen que las variables *HR* y *Temp* son buenos predictores de concentraciones de PM2.5. Dichas covariables se incluyen con una hora de retardo, donde $k = 1, \dots, K$, siendo K el número de covariables Z_{h-1d} que afectan al modelo.

4. Benchmark

El modelo propuesto se compara con un modelo *SARIMAX* como metodología lineal y una *RN* como metodología no lineal.

4.1. Modelo SARIMAX

Como modelo lineal se propone un modelo *ARIMA* (Box et al., 1994) que incluya un componente estacional (Taylor and McSharry, 2007) y variables exógenas, esto es, $SARIMAX(p, d, q) \times (P, D, Q)_S$ con la siguiente estructura:

$$\Phi^p(L)\Phi_S^P(L)\Delta^d\Delta^D PM_h = \delta_h + \Theta^q(L)\Theta_S^Q(L)\varepsilon_t \quad (4)$$

Donde, (p, d, q) representa los componentes del modelo *ARIMA* general y (P, D, Q) como componentes estacionales, con estacionalidad S (en este caso, 24 horas). En su forma general, Φ^p y Φ_S^P representa el componente autorregresivo $AR(p)$ y autorregresivo estacional $AR(P)$, respectivamente. Θ^q y Θ_S^Q son los componentes de medias móviles $MA(q)$ y $MA(Q)$ estacional. Finalmente, Δ^d y Δ^D es el componente de integración (I) no-estacional y estacional, respectivamente. δ_h se compone del intercepto del modelo y de un grupo de covariables exógenas incluidas en el modelo, como se muestra en la ecuación 5

$$\delta_h = \delta + \sum_{i=1}^{i=n} \sum_{k=1}^{k=m} \psi_{ik} Z_{i,h-k} \quad (5)$$

De esta forma, el modelo anterior se aplica según cada estación. Por ejemplo, en Estación Cerrillos se estudia la siguiente estructura

$$PM_h = \theta^0 + \beta_1 PM_{h-1} + \Phi_1 \varepsilon_{h-1} + \theta_h^1 PM_{h-24} + \Phi_2 \varepsilon_{h-24} + \theta_h^2 PM_{h-168} + PM \max_{h-24} + \gamma_1 CO_{h-24} + \gamma_2 Temp_{h-24} + \gamma_3 HR_{h-24} + \varepsilon_h \quad (6)$$

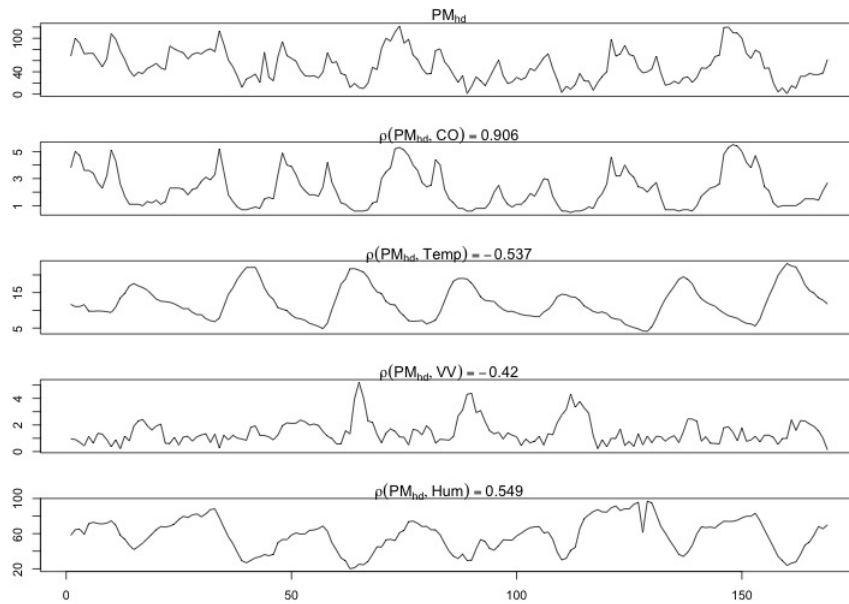


Figure 2: Correlación de PM2.5 con posibles variables a incluir en el modelo

Donde se incluye como variable exógena el retardo de 7 días del $PM_{2.5}$ (PM_{hd-168}) y las covariables, meteorológicas y ambientales, se incluyen con un día de retardo. Los coeficientes θ^1 y θ^2 se trabajan con las ecuaciones ya presentadas, de manera que sea comparable con el modelo propuesto.

4.2. Red Neuronal

La metodología propuesta se compara también con un modelo no lineal. Luego de explorar diferentes opciones, se encontró que resultados con mayor precisión se obtienen con la *RN* Feed Forward desarrollada por Perez and Gramsch (2016); Feng et al. (2015), que se compone de 13 inputs, 8 neuronas y 1 output, correspondiente a la concentración de $PM_{2.5}$ con cierta cantidad de horas de antelación.

Las variables de entrada son las seleccionadas en el modelo principal, de manera que sea comparable con el modelo propuesto (concentraciones ambientales y meteorológicas).

5. Predicción

En base al ajuste de modelos lineales revisados en la literatura, este trabajo utiliza un modelo lineal de ecuaciones múltiples a fin de pronosticar concentraciones horarias de $PM_{2.5}$ en las 11 estaciones de la Región Metropolitana. Se utiliza un periodo deslizante de 2 años como backtesting para evaluar la predicción desde el 01 de Abril al 31 de Agosto, periodo en el que el gobierno toma medidas de Gestión de Episodios Críticos (GEC). Del mismo lado, autores como Perez and Gramsch (2016) exponen que en este periodo existe un mayor número de alertas ambientales de $PM_{2.5}$ que en el resto del año.

La predicción se realiza por hora, reestimando el modelo cada 3 hrs., 6 hrs., 12 hrs. y 24 hrs. En este sentido, las covariables se incluyen con 24 horas de retardo según la estructura anteriormente explicada. Cada vez que el modelo es reestimado, se actualizan los valores horarios observados de PM2.5 según SINCA, mientras que dentro del intervalo de tiempo se utilizan los valores predichos. La predicción horaria termina a las 18 hrs. de cada día, dado que la Norma de Calidad Primaria de PM2.5 expone que es importante alertar un episodio crítico de contaminación atmosférica entre 20 hrs. y 21 hrs. del día anterior a la alerta (MMA, 2010).

5.1. Medidas de Ajuste y Predicción

Como medidas para evaluar el ajuste del pronóstico de la serie temporal PM2.5 se utiliza el error medio absoluto (MAE), la raíz del error cuadrático medio (RMSE) y el error medio absoluto escalado (MASE). En este sentido, el error de predicción simple se interpreta como $e_t = Y_t - F_t$, donde Y_t representa los valores observados y F_t los valores pronosticados.

Las ecuaciones 7, 8 y 9 presentan la estructura de las medidas usuales MAE, RMSE y el método MASE, respectivamente:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_t - F_t| \quad (7)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (Y_t - F_t)^2 \right]^{1/2} \quad (8)$$

$$MASE(q_t) = \frac{e_t}{\frac{1}{n-1} \sum_{i=2}^n |Y_t - Y_{t-1}|} \quad (9)$$

El MASE (Hyndman et al., 2006), presentado en la ecuación 9, utiliza el MAE y lo escala por el método “naïve”, generando pronósticos un periodo antes de cada dato en la muestra in-sample.

En este sentido, Hyndman et al. (2006) exponen que MAE asume una media estable en el tiempo (serie estacionaria), pero usualmente las observaciones en datos de alta frecuencia son susceptibles a estacionalidad y patrones comunes, características en las que el MASE es más adecuado.

6. Resultados Preliminares e Impacto Económico

Los resultados preliminares expuestos en este apartado se basan en la estimación de Estación Cerrillos. En cuanto al modelo en general, la Serie de Fourier captura el ciclo anual según la selección de sus coeficientes, m . Acorde a la literatura, autores como Clements et al. (2015); Zhou et al. (2000) utilizan 4 y 5 coeficientes de Fourier para capturar tendencia anual en datos de alta frecuencia. La Fig. 3 presenta la serie temporal PM2.5 con 4 coeficientes de Fourier. Se concluye que, efectivamente captura de manera adecuada el componente anual, siendo mayor el coeficiente en meses de invierno y en altas horas.

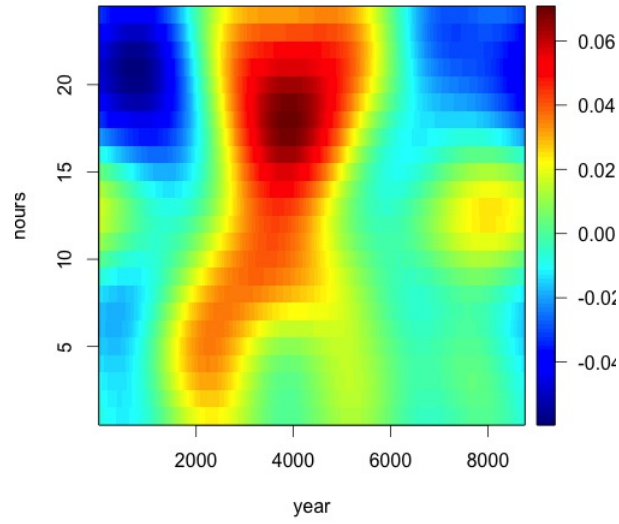


Figure 3: Estacionalidad Anual con 4 Coeficientes de Fourier.

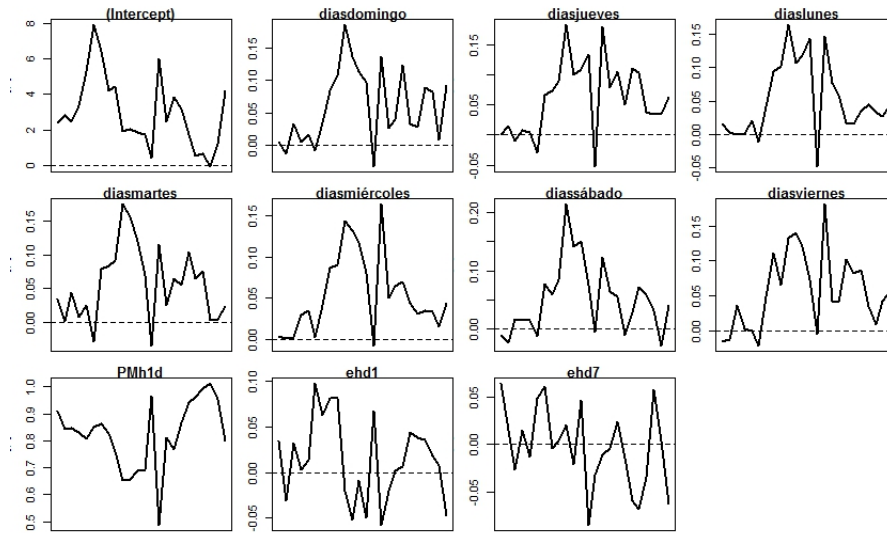


Figure 4: Coeficientes SEM Estación Cerrillos, Santiago, Chile.

Estación	SEM				SARIMA			
	RMSE	MAE	MASE	AIC	RMSE	MAE	MASE	AIC
Cerrillos	10.085	6.212	0.435	391983	10.407	6.281	0.440	394418
Condes	6.732	4.468	0.480	349607	6.745	4.472	0.481	348949
Pudahuel	9.312	5.522	0.372	383620	9.731	5.672	0.382	387372
Cerro Navia	9.499	5.693	0.349	385712	9.857	5.855	0.358	388724
El Bosque	10.129	6.589	0.448	392443	10.386	6.718	0.456	394201
Independencia	7.578	4.756	0.448	362023	7.645	4.795	0.452	362100
La Florida	8.764	5.449	0.459	377271	8.937	5.479	0.461	378451
Quilicura	10.089	5.783	0.453	392027	10.646	5.891	0.461	396790
Puente Alto	9.692	6.442	0.501	387816	9.778	6.444	0.501	387875
Parque O'higgins	9.729	5.546	0.440	388222	10.068	5.643	0.448	390940
Talagante	8.973	5.252	0.457	379733	9.074	5.255	0.457	380039

Table 1: Medidas de Ajuste para 11 estaciones de Santiago, Chile. Comparación MEM/SARIMA

Periodo	SEM			SARIMAX		
	RMSE	MAE	MASE	RMSE	MAE	MASE
24 hrs.	24.721	17.683	1.916	25.638	18.410	1.995
12 hrs.	23.857	16.702	1.810	24.937	17.515	1.898
6 hrs.	20.266	14.124	1.530	21.866	15.023	1.628
3 hrs.	16.852	11.822	1.281	22.016	15.032	1.629
1 hr.	12.467	8.976	0.972	17.134	11.968	1.297

Table 2: Medidas de Ajuste en Pronóstico, Comparación SEM / SARIMAX por periodos de reestimación.

Los coeficientes del modelo propuesto se presentan de manera gráfica en la Fig. 4, observando que en la mayor parte de las horas toman valores positivos e influyen estadísticamente en el modelo. En Tabla 1 se presentan algunas medidas de ajuste propuestas entre el SEM y el modelo lineal comparativo, SARIMAX. Se observa que, en las 3 medidas de error propuestas resulta ser mejor el SEM, lo que es un buen indicio del modelo aún cuando falta incluir covariables exógenas (Lluvia, Velocidad del Viento, Dirección del Viento). La Tabla 2 presenta los resultados de ajuste en predicción según periodos reestimados para la estación de Cerrillos. Se observa que el ajuste del pronóstico en el modelo propuesto es mejor a SARIMAX en los 5 periodos horarios evaluados.

Este estudio demuestra que una aproximación tradicional de ecuaciones múltiples, con un modelo por cada hora del día, logra ser una alternativa competitiva frente a otros modelos revisados en la literatura. Lo

anterior, obteniendo un RMSE de 24.721 en promedio de estimaciones cada 24 horas para la Estación Cerrillos, menor al modelo comparativo lineal SARIMAX, que en periodos de 24 hrs. logra un RMSE promedio de 25.638. Del mismo lado, a pesar que las medidas de ajuste son menores cuando el modelo se reestima cada una hora, es importante lograr reportes con anticipación diaria a fin de lograr estrategias oportunas frente a los ECCA. En resumen, se mejora en aproximadamente un tercio del error porcentual medio promedio de los pronósticos de 12 horas por delante reportados por la Comisión Nacional del Medio Ambiente.

En cuanto al impacto económico, es importante obtener un pronóstico asertivo de PM2.5 en cuanto a medidas de ajuste en predicción, a fin de tomar medidas estratégicas y regular restricciones temporales de emisión de este contaminante. Lo anterior, mediante el anuncio de Alertas Ambientales o Episodios Críticos de Contaminación Atmosférica (ECCA), debido a que sus consecuencias provocan un impacto económico de \$670 millones de dólares cada año.

Finalmente, es importante regular de manera efectiva el PM2.5, debido a que el daño en el bienestar social que significa un aumento en el riesgo de muerte puede ascender a \$1.900 millones de dólares. Lo anterior, asociado a costos médicos y pérdidas en la productividad laboral, debido al deterioro de la calidad del aire (MMA, 2013). Finalmente, el desafío para el trabajo futuro es evaluar el impacto económico de la calidad del aire en costos de salud de la población y en la productividad del trabajador, para empresas e industrias que abandonan sus labores los días que se decretan ECCA, usando esta nueva metodología para la predicción.

References

- Christoph Bohringer and Patrick EP Jochem. Measuring the immeasurable. a survey of sustainability indices. *Ecological economics*, 2007.
- GEP Box, GM Jenkins, and GC Reinsel. *Time Series Analysis- Forecasting and Control*. 1994.
- Joseph Cassmassi. Improvement of the forecast of air quality and of the knowledge of the local meteorological conditions in the metropolitan region. *Informe final, CONAMA RM*, 1999.
- Luis Cifuentes. Relación de la norma de calidad primaria mp2, 5 con la norma de calidad primaria de mp10. *Santiago, Chile. p158*, 2010.
- AE Clements, AS Hurn, and Z Li. Forecasting day-ahead electricity load using a multiple equation time series approach. *European Journal of Operational Research*, 2015.
- Xiao Feng, Qi Li, Yajie Zhu, Junxiong Hou, Lingyan Jin, and Jingjie Wang. Artificial neural networks forecasting of pm 2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, 107: 118–128, 2015.
- D Deniz Genc, Canan Yesilyurt, and Gurdal Tuncel. Air pollution forecasting in ankara, turkey using air pollution index and its relation to assimilative capacity of the atmosphere. *Environmental monitoring and assessment*, 166(1-4):11–27, 2010.
- PD Hien, VT Bac, HC Tham, DD Nhan, and LD Vinh. Influence of meteorological conditions on pm 2.5 and pm 2.5- 10 concentrations during the monsoon season in hanoi, vietnam. *Atmospheric Environment*, 36(21):3473–3484, 2002.
- KI Hoi, KV Yuen, and KM Mok. Kalman filter based prediction system for wintertime pm10 concentrations in macau. *Global NEST Journal*, 10(2):140–150, 2008.
- Jef Hooyberghs, Clemens Mensink, Gerwin Dumont, Frans Fierens, and Olivier Brasseur. A neural network forecast for daily average pm 10 concentrations in belgium. *Atmospheric Environment*, 39(18):3279–3289, 2005.
- Lovro Hrust, Zvezdana Bencetić Klaić, Josip Križan, Oleg Antonić, and Predrag Hercog. Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. *Atmospheric Environment*, 43(35):5588–5596, 2009.
- Rob J Hyndman et al. Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, 4(4):43–46, 2006.
- IB Konovalov, Matthias Beekmann, Frédéric Meleux, A Dutot, and Gilles Foret. Combining deterministic and statistical approaches for pm 10 forecasting in europe. *Atmospheric Environment*, 43(40):6425–6434, 2009.
- Jaakko Kukkonen, Leena Partanen, Ari Karppinen, Juhani Ruuskanen, Heikki Junninen, Mikko Kolehmainen, Harri Niska, Stephen Dorling, Tim Chatterton, Rob Foxall, et al. Extensive evaluation of neural network models for the prediction of no 2 and pm 10 concentrations, compared with a deterministic modelling system and measurements in central helsinki. *Atmospheric Environment*, 37(32):4539–4550, 2003.
- MMA. Análisis general del impacto económico y social de la norma de calidad primaria de material particulado 2.5. Technical report, AGIES Norma de Calidad Primaria para MP 2, 5. Ministerio del Medio Ambiente (MMA), 2010.
- MMA. Norma primaria de calidad ambiental para material particulado fino respirable mp 2,5. *Reporte Oficial de Ministerio del Medio Ambiente*, 2011.
- MMA. Primer reporte del estado del medio ambiente. Technical report, Ministerio del Medio Ambiente, 2013.
- LS Neal, P Agnew, S Moseley, C Ordóñez, NH Savage, and M Tilbee. Application of a statistical post-processing technique to a gridded, operational, air quality forecast. *Atmospheric Environment*, 98:385–393, 2014.
- OMS. Guías de calidad del aire de la oms, relativas al material particulado, el ozono, el dióxido de nitrógeno y el dióxido de azufre. Technical Report pp. 7-12, OMS (Organización Mundial de la Salud), pp. 7-12. Ginebra, Suiza., 2006.
- Patricio Perez and Ernesto Gramsch. Forecasting hourly pm2. 5 in santiago de chile with emphasis on night episodes. *Atmospheric Environment*, 124:22–27, 2016.

- Patricio Perez and Jorge Reyes. Prediction of maximum of 24-h average of pm10 concentrations 30h in advance in santiago, chile. *Atmospheric Environment*, 36(28):4555–4561, 2002.
- Patricio Perez and Jorge Reyes. An integrated neural network model for pm10 forecasting. *Atmospheric Environment*, 40(16):2845–2851, 2006.
- Pablo E Saide, Gregory R Carmichael, Scott N Spak, Laura Gallardo, Axel E Osses, Marcelo A Mena-Carrasco, and Mariusz Pagowski. Forecasting urban pm10 and pm2.5 pollution episodes in very stable nocturnal conditions and complex terrain using wrf-chem co tracer model. *Atmospheric Environment*, 45(16):2769–2780, 2011.
- Rouzbeh Shad, Mohammad Saadi Mesgari, Arefeh Shad, et al. Predicting air pollution using fuzzy genetic linear membership kriging in gis. *Computers, Environment and Urban Systems*, 33(6):472–481, 2009.
- SINIA. Analisis general del impacto económico y social de la norma de calidad primaria de material particulado 2.5. Technical report, Sistema Nacional de Información Ambiental, 2010.
- Ernst Stadlober, Siegfried Hörmann, and Brigitte Pfeiler. Quality and performance of a pm10 daily forecasting model. *Atmospheric Environment*, 42(6):1098–1109, 2008.
- R Stern, P Bultjes, M Schaap, R Timmermans, R Vautard, A Hodzic, M Memmesheimer, H Feldmann, E Renner, R Wolke, et al. A model inter-comparison study focussing on episodes with elevated pm10 concentrations. *Atmospheric Environment*, 42(19):4567–4588, 2008.
- Wei Sun, Hao Zhang, Ahmet Palazoglu, Angadh Singh, Weidong Zhang, and Shiwei Liu. Prediction of 24-hour-average pm2.5 concentrations using a hidden markov model with different emission distributions in northern california. *Science of the total environment*, 443:93–103, 2013.
- James W Taylor and Patrick E McSharry. Short-term load forecasting methods: An evaluation based on european data. *Power Systems, IEEE Transactions on*, 22(4):2213–2219, 2007.
- World Health Organization. WHO. Urban outdoor air pollution database. Technical report, Department of Public Health and Environment., 2011.
- Qingping Zhou, Haiyan Jiang, Jianzhou Wang, and Jianling Zhou. A hybrid model for pm2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of the Total Environment*, 496:264–274, 2014.
- Shuang Lin Zhou, Thomas Aquinas McMahon, Allan Walton, and Jane Lewis. Forecasting daily urban water demand: a case study of melbourne. *Journal of Hydrology*, 236(3):153–164, 2000.
- Ali Zolghadri and Franck Cazaurang. Adaptive nonlinear state-space modelling for the prediction of daily mean pm10 concentrations. *Environmental Modelling & Software*, 21(6):885–894, 2006.