
Modelación espacio-tiempo del material particulado en Santiago de Chile

Spatio-temporal modeling of particulate matter in Santiago of Chile

Enner Mendoza^a
ennerm@gmail.com

Eduardo Puraivan^b
puraivan@gmail.com

Resumen

En este trabajo se considera un modelo bayesiano jerárquico de espacio-tiempo para la concentración del material particulado 2.5 en la ciudad de Santiago de Chile propuesto por (Cameletti et al. 2013), tomando en consideración las covariables: temperatura, humedad relativa, velocidad del viento promedio diario y las coordenadas geográficas espaciales. Los datos utilizados fueron medidos en once estaciones de monitoreo de la calidad del aire de la ciudad de Santiago, registrados por hora durante 2012, 2013 y 2014. Con la información recopilada se construyeron mapas de predicción del material particulado 2.5 y de probabilidades de exceder el umbral permitido. Los cálculos fueron realizados en el *software* R específicamente con la librería INLA.

Palabras clave: contaminación del aire, material particulado 2.5 ($MP_{2.5}$), modelo bayesiano jerárquico espacio-tiempo.

Abstract

In this work, we consider a Bayesian hierarchical space-time model for concentration of particulate matter 2.5 in the City of Santiago de Chile proposed by (Cameletti et al. 2013). Taking into consideration the covariates: daily mean temperature, daily mean relative humidity, daily mean wind speed and spatial geographic coordinates. The data were measured in 11 air quality monitoring stations in the City of Santiago, recorded by hour during the years 2012, 2013 and 2014. The recopilatly information, we build prediction maps of the particulate material 2.5 and probability maps of exceeding the allowed threshold. The calculations were performed in the software the R-library INLA.

Keywords: air pollution; Particulate matter 2.5 ($PM_{2.5}$); Bayesian hierarchical space-time model.

^aEstudiante de Doctorado de la Universidad de Valparaíso, Chile

^bProfesor de la Universidad Viña del Mar, Chile.

1. Introducción

La Organización Mundial de la Salud (OMS) en el estudio realizado por *Global Urban Ambient Air Pollution Database* ((WHO 2016)) comparó la calidad del aire de casi tres mil ciudades en 103 países. En el informe se especifican las mediciones del material particulado 2.5 ($MP_{2.5}$) y el material particulado 10 (MP_{10}) por $\mu g/m^3$. Si bien, América Latina no es una de las regiones más contaminadas, en los diez primeros lugares aparecen algunas ciudades Chilenas, tales como: Coyhaique, con el mayor índice de contaminación registrando $64 \mu g/m^3$ de $MP_{2.5}$ y $75 \mu g/m^3$ de MP_{10} , la ciudad de Osorno, con $64 \mu g/m^3$ de $MP_{2.5}$, Rancagua, con $75 \mu g/m^3$ de MP_{10} , y, finalmente, Padre las Casas y Santiago con $64 \mu g/m^3$ de MP_{10} .

El material particulado se clasifica según su diámetro. El Ministerio del Medio Ambiente de Chile ((MMA 2016)) indica que existen dos tipos de medidas utilizadas para clasificar el material particulado: las menores a 10 micrones (MP_{10}) y las menores a 2.5 micrones ($MP_{2.5}$). El $MP_{2.5}$ es el más dañino para la salud humana debido a que su minúsculo diámetro le permite ingresar fácilmente al organismo, provocando enfermedades respiratorias y cardíacas, según la OMS. Ante esta realidad, la OMS señala los valores de referencia en la clasificación del grado permitido del material particulado en las ciudades, los cuales son los siguientes: el índice medio anual del $MP_{2.5}$ es $10 \mu g/m^3$ y para el MP_{10} es $20 \mu g/m^3$, asimismo, el índice medio 24 hrs del $MP_{2.5}$ es $25 \mu g/m^3$ y en el MP_{10} es $50 \mu g/m^3$, ((OMS 2017)).

Según (Vargas 2011), el $MP_{2.5}$ se produce por las emisiones directas de los procesos de combustión de fósiles, de la condensación de los gases de reacciones químicas en la atmósfera, de gases precursores como el dióxido de azufre, óxido de nitrógeno, compuestos orgánicos volátiles, amoníaco y otros. En Chile, Santiago es una de las ciudades que, históricamente, ha presentado alertas ambientales, por el índice medio del material particulado 24 hrs, esto es motivado por su geografía, la densidad poblacional y los números de agentes contaminantes. En las cifras del Instituto Nacional de Estadística (INE), la Región Metropolitana tiene 7,007,020 habitantes, lo que representa el 40% del total de población del país, además, la densidad es $454.9 \text{ hab}/\text{km}^2$ ((INE 2014)), también el parque automotriz en el 2011 era aproximadamente de 1,513,678 unidades ((INE 2011)). Cabe resaltar que la población de vehículos es el 41.4% del parque vehicular del país, adicionalmente, existen diversas industrias que generan contaminación del aire, esto lleva al Gobierno a tener sistema de alerta temprana en la calidad del aire. En 2012, la legislación en Chile modificó la norma de calidad del aire para el $MP_{2.5}$. La norma está en vigencia desde el 1 de enero del 2012, y como consecuencia generó cambios en el monitoreo de la calidad del aire en el transporte urbano, autos, buses y camiones, también en los proyectos de impacto ambiental. La norma establece que el índice medio anual del $MP_{2.5}$ es $20 \mu g/m^3$ y el índice medio 24 hrs del $MP_{2.5}$ es $25 \mu g/m^3$ ((BCN 2003)).

En la literatura más reciente, el estudio de contaminación del aire en la región metropolitana de Santiago de Chile ha sido tratado con diferentes metodologías,

entre ellos se tiene: (Silva et al. 2003) compara los resultados obtenidos utilizando análisis discriminante no-paramétrico, redes neuronales, regresión lineal múltiple y modelos MARS (*multivariate adaptive regression splines*). La utilización geoes-tadística aplicada por medio del método Kriging ordinario permite crear superficie de valores estimados de exposición del MP_{10} (Fuenzalida-Díaz 2015). Desde otro punto de vista, se tiene el artículo de (Álvarez E et al. 1999), donde se toma el tema utilizando la metodología de valoración contingente (VC), se estima la disposición a pagar (DAP) de la población de Santiago por un programa que reduzca 50% la contaminación atmosférica que la afecta. Las estimaciones realizadas permiten inferir que la DAP promedio individual es de entre 10 y 17 dólares por año.

Recientemente, los modelos espacio-tiempo permiten mostrar la distribución geográfica del riesgo y cómo estos evolucionan a lo largo del tiempo. Por este motivo, el uso de modelos bayesianos jerárquicos espacio-temporales se presenta como una herramienta muy útil para la vigilancia del $MP_{2.5}$. Con base en ello, el objetivo del estudio es aplicar el modelo bayesiano jerárquico espacio-temporales para identificar patrones espaciales, por medio de los mapas de riesgo del $MP_{2.5}$ en la ciudad Santiago de Chile, propuesto por (Cameletti et al. 2013). En el trabajo se utiliza el paquete *Integrated Nested Laplace Approximation* (INLA) propuesto por (Rue et al. 2009), que se encuentra en R project library (R 2011).

El trabajo esta organizado como se muestra a continuación. En la sección 2, materiales y métodos, hay una breve descripción de las ecuaciones, funcionamiento del programa INLA y su estructuración en la estimación de los parámetros. En la sección 3, resultados, se describe la notación usada, la región de estudio, análisis descriptivo del $MP_{2.5}$ para conocer las características de los datos o errores existentes. En la sección 4, estimación del modelo, se detalla la construcción de triangulación de región y la estimación de los parámetros. En la sección 5, predicción espacial, se explica cómo se realiza la predicción y las construcciones de los mapas. En la sección 6, discusión, se refleja la importancia de la metodología y facilidad de la predicción del $MP_{2.5}$ en las regiones a través de los mapas.

2. Materiales y métodos

Los procesos estocásticos en el espacio y tiempo se presentan de la siguiente manera:

$$\{Z(\mathbf{s}; t), \mathbf{s} \in \mathcal{D} \subset \mathbb{R}^2; t \in I \subset \mathbb{R}\},$$

Para cada par de localizaciones $(\mathbf{s}_1; t_1), (\mathbf{s}_2; t_2) \in \mathcal{D} \times I$, donde $\mathcal{D} \subset \mathbb{R}^2, I \subset \mathbb{R}$, la covarianza entre las correspondientes variables aleatorias:

$$C((\mathbf{s}_1; t_1), (\mathbf{s}_2; t_2)) = \text{Cov}[Z(\mathbf{s}_1; t_1); Z(\mathbf{s}_2; t_2)].$$

Estos procesos suelen descomponerse en una parte no estocástica que recoge la variación a gran escala y otro componente que recoge el comportamiento local, o variación a pequeña escala

$$Y(\mathbf{s}; t) = \mu(\mathbf{s}; t) + Z(\mathbf{s}; t), \quad \mathbf{s}, t \in \mathcal{D} \times \mathbb{R}.$$

De este modo, el comportamiento global está determinado por $\mu(\mathbf{s}; t)$, aunque sea desconocido, y se puede estimar mediante un modelo de regresión adecuado. Una vez que se ha ajustado un modelo a los datos, el interés es hacer predicciones en localizaciones no observadas, generalmente el propósito es construir un mapa completo del fenómeno en la región de estudio. Hay varios enfoques para esta construcción de los mapas, pero el método que es usado en este trabajo es el planteado por (Cameletti et al. 2013).

Según (Banerjee et al. 2014) en los modelos espacio-temporal básicamente el objetivo es descomponer la variable respuesta en tres partes. Una, debido a la tendencia con o sin covariables, la otra parte, al espacio-temporal, y por último, a un error aleatorio. La notación $y(\mathbf{s}_i, t)$ corresponde a las realizaciones de un proceso espacio-temporal $Y(\cdot, \cdot)$: en este caso, representan la concentración promedio diaria del $MP_{2.5}$ en la estación ubicada en \mathbf{s}_i , con $i = 1, 2, \dots, d$ y en el tiempo $t = 1, \dots, T$. El modelo propuesto por (Cameletti et al. 2013) establece lo siguiente:

$$y(\mathbf{s}_i, t) = z(\mathbf{s}_i, t)\boldsymbol{\beta} + \xi(\mathbf{s}_i, t) + \epsilon(\mathbf{s}_i, t) \quad (1)$$

Donde

1. $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ son los parámetros.
2. $z(\mathbf{s}_i, t)$ es un vector de covariables asociada a $y(\cdot, \cdot)$.
3. $\epsilon(\mathbf{s}_i, t) \sim N(0, \sigma_\epsilon^2)$ es el error de medición (un ruido blanco).
4. $\xi(\mathbf{s}_i, t)$ es la realización del proceso de estado, es decir, el verdadero valor no observado de la contaminación, que tiene una estructura de un proceso auto regresivo:

$$\xi(\mathbf{s}_i, t) = a\xi(\mathbf{s}_i, t-1) + \omega(\mathbf{s}_i, t)$$

con $t = 2, \dots, T$, $|a| < 1$, $\xi(\mathbf{s}_i, 1) \sim N\left(0, \frac{\sigma_\omega^2}{1-a^2}\right)$ y $\omega(\mathbf{s}_i, t)$ tiene distribución gaussiana de parámetros con media cero y función de covarianza que solo depende del espacio, en otros términos, para $i \neq j$

$$\text{Cov}[\omega(\mathbf{s}_i, t), \omega(\mathbf{s}_j, t')] = \begin{cases} 0 & \text{si } t \neq t' \\ \sigma_\omega^2 C(h) & \text{si } t = t' \end{cases}$$

Donde $\text{Var}[\omega(\mathbf{s}_i, t)] = \sigma_\omega^2$ para cada \mathbf{s}_i, t y $h = \|\mathbf{s}_i - \mathbf{s}_j\|$,

$$C(h) = \frac{1}{\Gamma(\nu)2^{\nu-1}} (\kappa h)^\nu K_\nu(\kappa h)$$

Donde K_ν es la función de Bessel de segundo tipo modificada de orden $\nu > 0$; además, $\nu > 0$ es el parámetro suavizamiento, su valor entero determina la diferencia en media cuadrática y $\kappa > 0$ es el parámetro de escala ((Whittle 1963)).

Para t fijo, se reescribir la ec. (1) en forma vectorial:

$$\mathbf{y}_t = \mathbf{z}_t \boldsymbol{\beta} + \boldsymbol{\xi}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\xi}_t = a \boldsymbol{\xi}_{t-1} + \boldsymbol{\omega}_t \quad (2)$$

donde $\boldsymbol{\epsilon}_t \sim N(0, \sigma_\epsilon^2 I_d)$, $\boldsymbol{\omega}_t \sim N(0, \Sigma = \sigma_\omega^2 \tilde{\Sigma})$ con $\tilde{\Sigma} = (C(\|\mathbf{s}_i - \mathbf{s}_j\|))_{i,j=1}^d$,
 $\mathbf{y}_t = (y(\mathbf{s}_1, t), \dots, y(\mathbf{s}_d, t))^\top$, $\mathbf{z}_t = (\mathbf{z}(\mathbf{s}_1, t)^\top, \dots, \mathbf{z}(\mathbf{s}_d, t)^\top)^\top$,
 $\boldsymbol{\xi}_t = (\xi(\mathbf{s}_1, t), \dots, \xi(\mathbf{s}_d, t))^\top$ con $\boldsymbol{\xi}_1 \sim N(0, \frac{\Sigma}{1-a^2})$. Los parámetros a estimar son $\boldsymbol{\beta}, \sigma_\epsilon^2, a, \sigma_\omega^2, \kappa$.

El procedimiento se enfoca en las ecuaciones diferenciales parciales estocásticas (EDPE) junto con el algoritmo INLA para superar los problemas computacionales relacionados con el modelamiento de campo gaussiano (CG). Además, se usa campo aleatorio de markov Gaussiano (CAMG) para el modelado del proceso espacial en la dependencia de los datos observados por unidades de área con respecto a las cuadrículas regulares, la estructura reticular o las regiones geográficas. Por otro lado, la inferencia bayesiana, se realiza a través de las técnicas de la cadena de Markov Monte Carlo (MCMC), esto es un reto debido a los problemas de convergencia y grandes cargas computacionales. Sin embargo, en el marco bayesiano, el tratamiento común para hacer inferencia en un modelo (es decir, estimación de parámetros y predicción espacial) es el muestreo en MCMC, esta propuesta consiste en representar un CG con función de covarianza Matérn como un CAMG mediante las EDPE. Básicamente, las EDPE usan una representación de elementos finitos para definir el campo Matérn en combinación lineal con las funciones básicas definidas en un dominio \mathcal{D} triangulado. Esto es subdividir \mathcal{D} en un conjunto de triángulos que no se intercepten entre sí, se encuentren en un borde y esquina común.

Utilizando el programa INLA, lo primero es construir la malla de la región, resulta ser una malla con pequeños triángulos en el dominio de interés, y triángulos más grandes en la región exterior. Así pues, la malla se construye tomando los puntos iniciales corresponden a las estaciones de monitoreo, entonces los puntos donde comienza la triangulación, permite definir un dominio que es un polígono convexo y que contiene a la región de interés.

Con respecto a la predicción espacial, vale la pena señalar que el algoritmo INLA proporciona la distribución condicional *a posteriori* de $\boldsymbol{\xi}$ para todos los n vértices de la triangulación. Una vez que se da $\boldsymbol{\xi}$, inmediatamente se obtiene una predicción para \mathbf{y}_t que se mapea el dominio triangulado. Esta es una ventaja considerable en términos de tiempo de cálculo con respecto a los métodos de MCMC que requieren primero obtener la distribución condicional completa de los parámetros, y luego simular a partir de la distribución predictiva posterior de $y(\mathbf{s}_0, t)$ para cada $\mathbf{s}_0 \in \mathcal{D}$.

3. Resultados

Los datos fueron medidos en once estaciones de monitoreo, y se obtuvieron del Ministerio del Medio Ambiente de Chile. Cada registro es el promedio del día desde el 1 de enero de 2012 hasta el 31 de enero de 2014 para distinguir las estaciones se utiliza la notación mostrada en Tabla 1, se usan letras mayúsculas para hacer distinción de las estaciones y facilitar las notaciones. Los registros de la

Tabla 1: *Estaciones de monitoreo.*

Notación	Nombre de la estación
F	Independencia
L	La Florida
M	Las Condes
N	Santiago(P. O'Higgins)
O	Pudahuel
P	Cerrillos
Q	El Bosque
R	Cerro Navia
S	Puente Alto
T	Talagante
V	Quilicura

base de datos están diferenciados de la forma siguiente: estación, fecha, longitud, latitud, promedio de la temperatura (TEM), promedio de la humedad relativa (HR), promedio de la velocidad del viento (VV) y el promedio del $MP_{2.5}$ por cada día, todas estas mediciones se realizaron por estación de monitoreo, esto equivale a una matriz de ocho columnas en donde las filas representan las estaciones por día.

3.1. La región de estudio

El área de estudio es la Región Metropolitana de Santiago de Chile, la cual está ubicada en la zona central y es la capital del país. Según información del (ODEPA 2015), la superficie de la región es de $15,403.2 \text{ km}^2$, con una extensión entre los $32^\circ 55'$ y $34^\circ 19'$ de latitud, y entre los $69^\circ 47'$ y $71^\circ 43'$ longitud. Las estaciones de monitoreo de la calidad de aire están ubicadas por latitud y longitud como se muestra en Figura 1.

3.2. Análisis descriptivo

La geoestadística tiene como objetivo proporcionar descripciones cuantitativas de las variables naturales distribuidas en el espacio o en el espacio tiempo (Chiles & Delfiner 2009); para ello, se generan los mapas de la región o área de interés, los cuales llevan a caracterizar el riesgo o valorar el ambiente, así son usados como herramienta de predicción. En esta sección se realizará un análisis preliminar de la base de los datos que nos permitirá controlar la presencia de posibles errores en la fase de introducción de estos, es decir, se detectarán los valores fuera de rango, o la presencia de valores perdidos, luego se realizará su respectiva corrección.

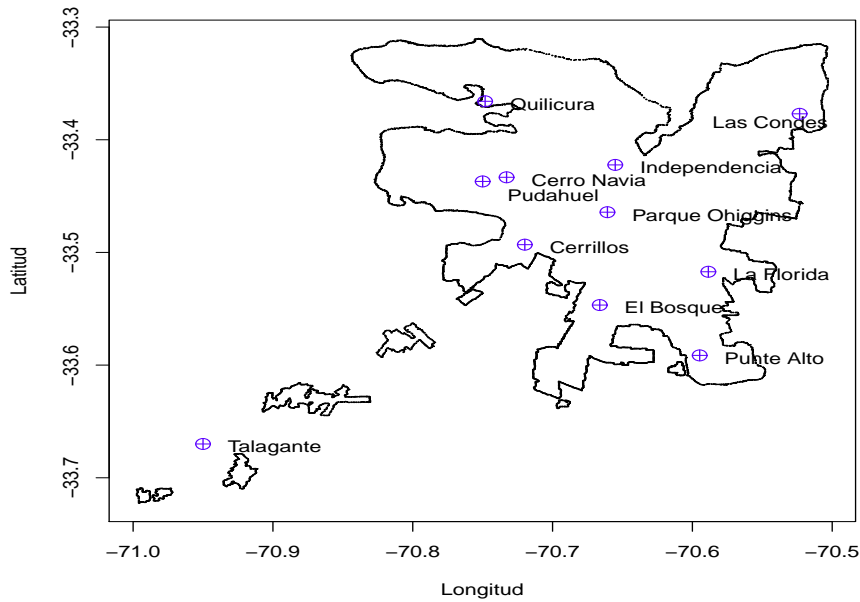


Figura 1: Región de estudio y las ubicaciones de las estaciones de monitoreo.

En la Tabla 2, el resumen de los promedios y las desviaciones estándar del $MP_{2.5}$, por estación de monitoreo según año. Se observa la variabilidad de las mediciones

Tabla 2: Promedio (\bar{x}) y desviación (σ) del $MP_{2.5}$ por estaciones.

Año		F	L	M	N	O	P	Q	R	S	T	V
2012	\bar{x}	23.76	26.15	19.99	26.59	26.95	25.39	28.26	29.24	28.26	21.74	24.78
	σ	11.85	11.63	8.1	14.09	19.82	14.72	15.65	21.07	11.64	14.56	13.73
2013	\bar{x}	24.41	25.86	21.53	27.82	24.70	26.71	29.58	27.28	35.29	21.06	25.07
	σ	11.36	10.95	7.57	12.25	15.08	13.40	15.02	17.62	15.11	12.14	11.74
2014	\bar{x}	28.20	29.61	23.25	24.46	29.53	31.33	32.98	33.14	27.01	19.17	28.83
	σ	16.20	18.55	13.08	12.60	24.48	20.84	22.49	22.49	15.83	12.28	19.76

por estación del $MP_{2.5}$, en particular la estación R presenta el mayor promedio en el 2012 y en el 2014. La estación S tiene el mayor promedio en el 2013. Las estaciones con los menores promedios: M en el 2012, y M, T en el 2013 y 2014. Las estaciones F, M, P, Q y V tienen un comportamiento creciente en cada año. En la Figura 2, el gráfico de barras describe los promedios anuales de la concentración del $MP_{2.5}$ por estación y la recta horizontal a la altura de $20 \mu g/m^3$, siendo esta línea correspondiente al índice medio anual según la legislación vigente en Chile permitida. En la figura, se aprecia que casi todas las estaciones monitoreo superan la norma anual salvo la estación T en el 2014. En la Tabla 3, se detallan

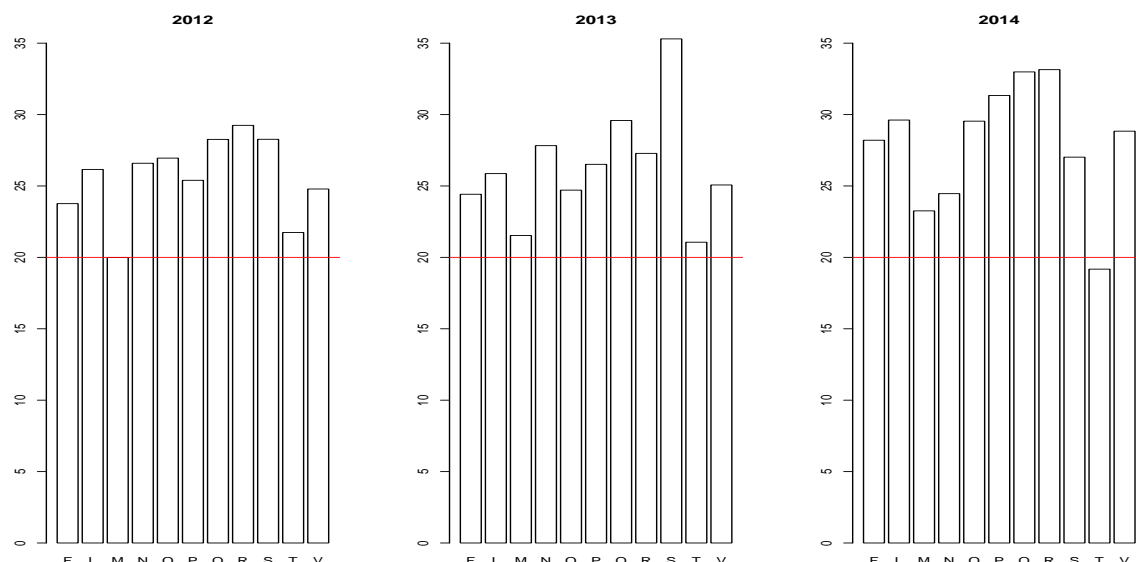


Figura 2: Nivel concentración promedio del $MP_{2.5}$ anual por cada estación de monitoreo.

los números de los días por año de las estaciones que superan el umbral de la Ley de Chile del índice medio anual del $MP_{2.5}$. En el 2012, las estaciones N, O y R presentan más de 30 días por arriba del índice permitido, asimismo, en el 2013 las estaciones Q, R y S tienen más de 30 días, también en el 2014 las estaciones F, L, O, P, Q, R, S y V tienen más de 30 días. Se puede apreciar que existe un aumento importante del total de días en que se superan el umbral de la norma en el 2014 respecto a los años anteriores. En la Figura 3, la estación T tiene la mediana me-

Tabla 3: Números de los días que superan el umbral del $20 \mu\text{g}/\text{m}^3$ por estación y por año.

Año	F	L	M	N	O	P	Q	R	S	T	V	Total de días
2012	13	13	1	33	41	25	43	57	17	18	24	285
2013	8	15	0	25	29	21	36	49	57	7	18	258
2014	46	57	17	16	58	49	69	71	39	12	54	488

nor por cada año, al contrario que la estación S tiene la mayor mediana en 2012 y 2013, asimismo, en el 2014, P y Q tienen la mayor mediana. Se puede apreciar que existen mediciones de valores extremos en la mayoría de las estaciones y en general los datos son levemente asimétricos. En los meses junio-septiembre correspondiendo a la temporada de invierno en la ciudad de Santiago, se observa un incremento significativo en la mediana del $MP_{2.5}$ (Figura 4). Por otro lado, se realiza el gráfico de densidad del $MP_{2.5}$, donde se observa que no existe simetría. Si aplicamos la

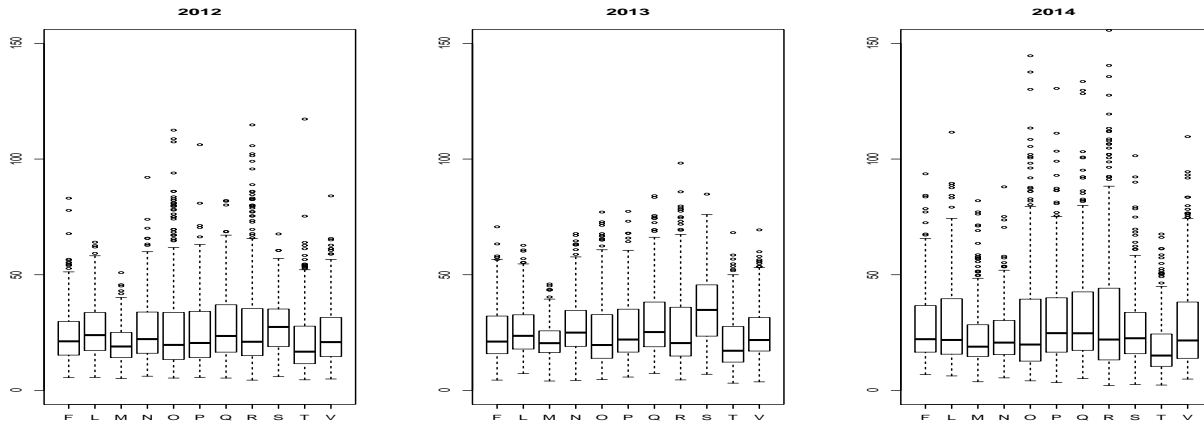


Figura 3: *Box-plot del $MP_{2.5}$ por estación según el año.*

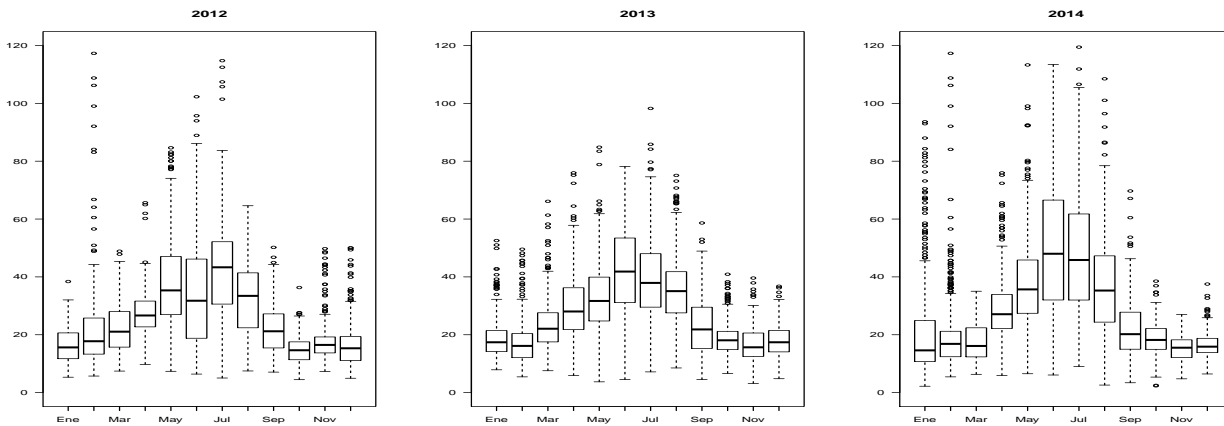


Figura 4: *Box-plot nivel del $MP_{2.5}$ por mes según el año.*

transformación logaritmo al $MP_{2.5}$, entonces el gráfico se puede apreciar que su soporte esta alrededor del intervalo de 1,5 a 4,5 y el centro es aproximadamente en 3 en escala logarítmica, de modo que esto es muy útil para el programa INLA y la estimar del modelo (Figura 5).

En la base de los datos existen valores perdidos en las variables TEM: 465 datos, HR: 2 datos, VV: 2 datos y el $MP_{2.5}$: 141 datos. Esta perdida se debe a causas desconocidas, pero se sabe que las mediciones se hacen con aparatos electrónicos que de alguna manera u otra dejan de realizar los registros, algunas veces se debe a las averías. En este caso, se hizo imputación de los valores perdidos con un programa de la librería de R Project llamado VIM (*Visualization and Imputation of Missing Values*), la cual usa un método de regresión (Templ et al. 2011a).

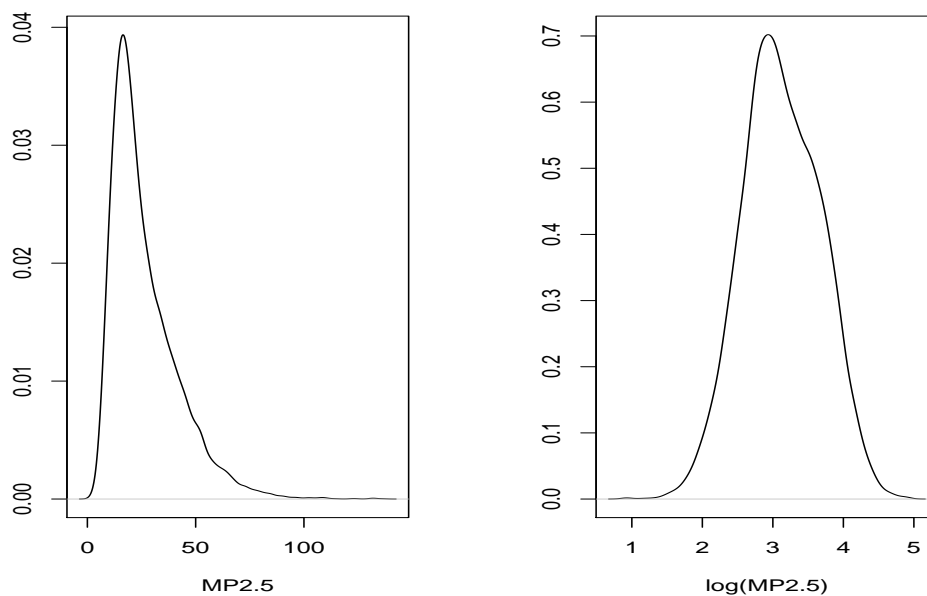


Figura 5: Densidad del $MP_{2.5}$ y la transformación.

4. Estimación del modelo

Utilizando el programa INLA, lo primero es construir la malla de la región, esto resulta ser una malla con pequeños triángulos en el dominio de interés, y triángulos más grandes en la región exterior. Así pues, el número de los vértices es 120, y los triángulos 211, (Figura 6). Los puntos azules son las estaciones validas, es decir, sus valores son de referencia para la estimación de los parámetros del modelo, junto con los puntos rojos que son las otras estaciones.

4.1. Parámetros estimados del modelo

En esta sección, se describe cómo hallar los parámetros del modelo; así, después que se construye la malla, se crea una matriz de covarianza espacial de Matérn de parámetro y coeficiente de suavizamiento $\nu = 2$, con la malla construida se obtienen los nuevos valores en los nodos (los vértices de los triángulos), en el cual se estima $\mathbf{y}(\cdot, \cdot)$ con el programa INLA. Los valores de los parámetros de la ecuación (2) que se muestran en la Tabla 4. Resultando que el valor del intercepto es aproximadamente 3.159, es decir, el valor real es $\exp(3.159) \approx 23.54704$, ya que se transformaron los valores del $MP_{2.5}$ (Figura 5), así que, el intercepto del modelo es el valor más explicativo y muy significativa para el nivel del $MP_{2.5}$

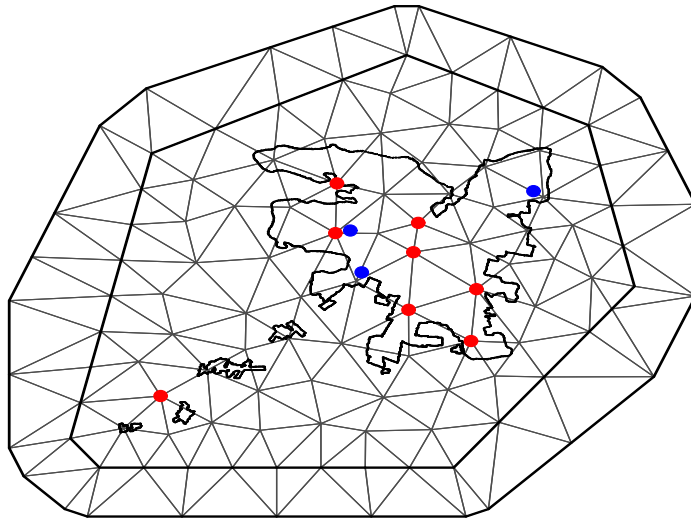


Figura 6: Triangulación de la región en estudio.

con las covariables. Igualmente, la desviación estándar 9.97635618 indica una alta variabilidad del modelo.

Tabla 4: Parámetros estimados del modelo de la ec. (2), desviación estándar y cuartiles.

		SD	Q_1	Q_2	Q_3
Intercepto	3.1592187	9.97635618	-18.3223555	3.2851794	23.8997833
Longitud	-0.8210709	0.0257748	-0.8716835	-0.821069	-0.7705155
Latitud	-0.3321282	0.02302713	-0.3773363	-0.3321298	-0.2869523
TEM	-0.7150994	0.03822415	-0.7903873	-0.7150201	-0.6403191
HR	-0.3173094	0.02975235	-0.375842	-0.3172705	-0.2590455
VV	0.6141912	0.02736371	0.5605062	0.6141778	0.6679012

Por otro lado, se obtienen los hiperparámetros del modelo, y todos los cálculos de los posteriores (media, desviación estándar y cuartiles), se muestra en la Tabla 5. Se observa una mayor variación en el término espacial en lugar del error de medición; además, el valor del $AR(1)$, los coeficientes de correlación temporal ratifican una persistencia de un modelo a corto plazo para el $MP_{2.5}$.

Tabla 5: *Estimación del posterior de la ec. (2), desviación estándar y cuartiles..*

		SD	Q_1	Q_2	Q_3
σ_ϵ^2	3.0452565	4.607415 E-02	2.9565503	3.0444838	3.137408
σ_w^2	8.1791904	5.972116 E-00	1.8639261	6.5285297	24.197699
ρ	1.8947079	5.213733 E-01	1.0814515	1.8219587	3.115985
a	0.9999612	5.025567 E-05	0.9998319	0.9999763	0.999997

5. Predicción espacial

Para el desarrollo de las predicciones espaciales del $MP_{2.5}$ en un día determinado: primero, se considera una malla que es dividida en 71x80 puntos en cuadrado de área 1.25x1.25 km que van desde la latitud 33.8 a 33.5 en la dirección norte y desde la longitud 71 a 70.4 en dirección este; segundo, para cada día (1 a 1095), se construye una malla por cada registro de TEM, HR y VV, es decir, que las mallas por cada registro son matrices tridimensionales de la superficie (latitud, longitud, día); posteriormente, con los parámetros estimados por el programa INLA tanto del modelo como del posterior, se presenta el resultado para la concentración del $MP_{2.5}$ el día 19 de julio de 2014 y la probabilidad de exceder el umbral de $25\mu\text{g}/\text{m}^3$ (Figura 7).

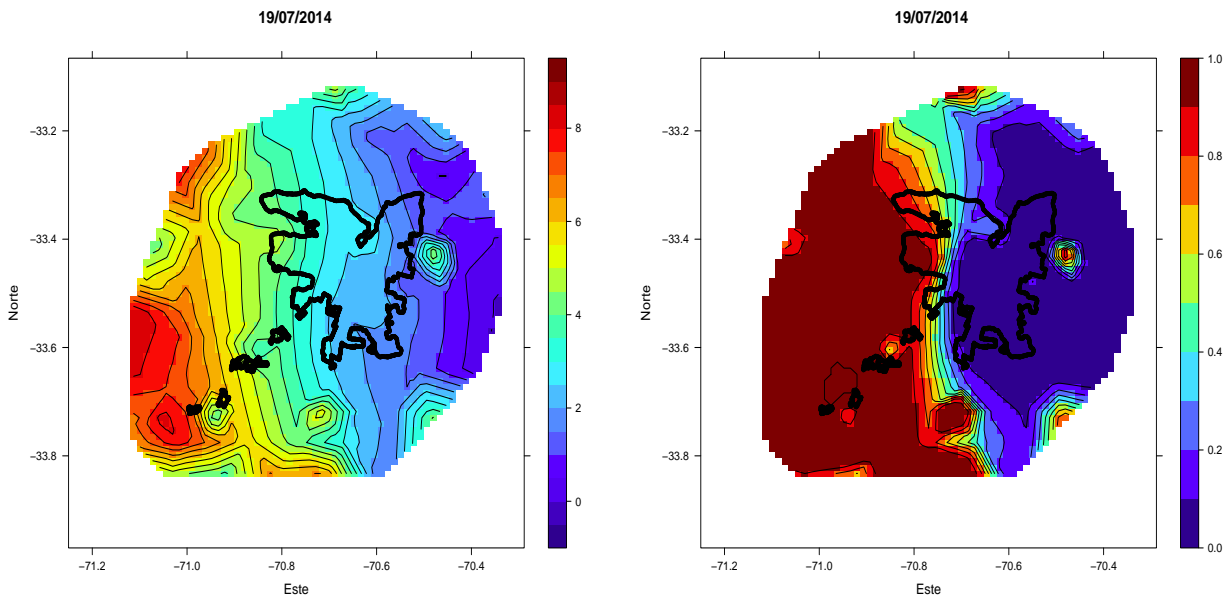


Figura 7: *Mapa de predicción $MP_{2.5}$ y mapa de probabilidad de exceder el umbral $25\mu\text{g}/\text{m}^3$ en escala logarítmica, para el día 19 de julio de 2014..*

6. Discusión

El programa INLA muestra ser una gran herramienta para la predicción espacial del $MP_{2.5}$ usando los mapas. Basta tener el modelo estimado y luego introduce un día determinado, de esta forma, se conoce el mapa de predicción y el mapa de probabilidad que excede el umbral $25\mu g/m^3$. Es muy eficaz, útil para crear seguimientos de las alertas tempranas de la contaminación del aire en la Región Metropolitana Santiago de Chile. Recapitulando, este procedimiento podría ser aplicado en otras zonas de Chile u otros países que poseen una alta tasa de contaminación. En síntesis, se muestra una predicción de la fecha 19 julio de 2014 como una evidencia del mapa de contaminación del aire y su mapa de probabilidad de exceder el umbral permitido en términos del $MP_{2.5}$ en escala logarítmica (Figura 7). En el mapa, se detallan las regiones de alto grado de contaminación del aire, y las zonas bajas de contaminación del aire por la presencia del $MP_{2.5}$. Las zonas con valores superiores a 3.2 se consideran que exceden la norma del $25\mu g/m^3$, por el contrario, se puede apreciar en la Figura 7 que las zonas que se encuentran en azul índigo son aquellas que no exceden la norma. Finalmente, el modelo detecta las zonas críticas que generan contaminantes como se especifica en la longitud 70.5 y latitud 33.4, allí hay una zona industrial y las zonas de color rojo oscuro el programa INLA estima correctamente, ya que, no se tienen mediciones.

Recibido: 16 de agosto de 2017

Aceptado: 2 de abril de 2018

Referencias

- Álvarez E, R., Figueroa B, E. & de F, S. V. (1999), 'Beneficios económicos de una reducción de la contaminación atmosférica en santiago de chile', *Investigación Económica* pp. 143–169.
- Banerjee, S., Carlin, B. P. & Gelfand, A. E. (2014), *Hierarchical modeling and analysis for spatial data*, Crc Press.
- BCN (2003), 'Biblioteca del congreso nacional de chile, decreto 110', Consulta: 09 de junio de 2017.
*<http://www.leychile.cl/Consulta/listaMasSolicitudesxmat?agr=1021sub=514tipCat=1>
- Cameletti, M., Lindgren, F., Simpson, D. & Rue, H. (2013), 'Spatio-temporal modeling of particulate matter concentration through the spde approach', *ASTA Advances in Statistical Analysis* pp. 1–23.
- Chiles, J.-P. & Delfiner, P. (2009), *Geostatistics: modeling spatial uncertainty*, Vol. 497, John Wiley and Sons.
- Fuenzalida-Díaz, M. (2015), 'Evaluación de modelos geoestadísticos aplicados a la exposición al contaminante atmosférico pm10 en chile', *Ciencias Espaciales* 8(1), 441–457.

- INE (2011), ‘Instituto nacional de estadística de chile’, Anuario 2011, Consulta: 09 de junio de 2017.
*<http://www.ine.cl/estadisticas/economicas/transporte-y-comunicaciones>
- INE (2014), ‘Instituto nacional de estadística de chile’, Anuario 2014, Consulta: 09 de junio de 2017.
*<http://www.ine.cl/estadisticas/demograficas-y-vitales>
- MMA (2016), ‘Ministerio del medio ambiente, gobierno de chile’, Consulta: 09 de junio de 2017.
*http://www.mma.gob.cl/1304/articles-52016_capitulo1.pdf
- ODEPA (2015), ‘Oficina de estudios y políticas agrarias’, Consulta: 09 de junio de 2017.
*http://www.odepa.cl/wp-content/files_mf/1437485361Metropolitanajulio.pdf
- OMS (2017), ‘Organización mundial de la salud’, Consulta: 09 de junio de 2017.
*<http://www.who.int/mediacentre/factsheets/fs313/es/>
- R (2011), ‘R core team: R a language and environment for statistical computing. vienna, austria: R foundation for statistical computing’.
*<http://www.R-project.org>
- Rue, H., Martino, S. & Chopin, N. (2009), ‘Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations’, *Journal of the royal statistical society: Series b (statistical methodology)* **71**(2), 319–392.
- Silva, C., Alvarado, S., Montaña, R. & Pérez, P. (2003), ‘Modelamiento de la contaminación atmosférica por partículas: comparación de cuatro procedimientos predictivos en santiago’, *Chile. Biomatemática* **13**, 113–27.
- Templ, M., Alfons, A., Kowarik, A. & Prantner, B. (2011a), ‘Vim: Visualization and imputation of missing values’, URL <http://CRAN.R-project.org/package=VIM>. R package version **3**(0).
- Vargas, C. (2011), ‘Efectos de la fracción gruesa (pm10-2.5) del material particulado sobre la salud humana’, *Revisión Bibliográfica MINSAL* .
- Whittle, P. (1963), ‘Stochastic-processes in several dimensions’, *Bulletin of the International Statistical Institute* **40**(2), 974–994.
- WHO (2016), ‘World health organization’, Consulta: 09 de junio de 2017.
*http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/