

# *Evaluación de modelos geoestadísticos aplicados a la exposición al contaminante atmosférico $PM_{10}$ en Chile*

Manuel Fuenzalida-Díaz

## **Resumen:**

Se modela la exposición al contaminante atmosférico  $MP_{10}$  a partir de los datos registrados en 108 estaciones de monitoreo pertenecientes al Sistema Nacional de Calidad del Aire, utilizando geoestadística aplicada por medio del método Kriging ordinario, por su cualidad de representar fenómenos cuyos valores varían en función de la localización espacial de puntos de observación. El procedimiento se lleva a cabo en tres etapas: (1) Análisis exploratorio de los datos, para determinar si es necesario o no la transformación de los datos y la eliminación de tendencia en los mismos. (2) Análisis estructural para el cálculo del semivariograma experimental y el ajuste a seis modelos matemáticos teóricos seleccionados. (3) Predicciones para la obtención del Mapa predicción y mapa de errores estándar de la predicción.

Los resultados obtenidos han permitido crear una superficie de valores estimados de exposición a  $PM_{10}$ , que entre sus principales características, presenta una distribución territorial con un fuerte sesgo regional, concentrándose espacialmente en territorios del cluster minero y áreas metropolitanas. Los mayores errores de predicción se dan en sectores de baja densidad poblacional y de estaciones de monitoreo, en el desierto de Atacama y la cordillera de Los Andes.

**Palabras Clave:** Kriging, material particulado, exposición a contaminantes, sistemas de información geográfica.

## **Abstract:**

Is modeled exposure to the air pollutant particulate matter ( $PM_{10}$ ) starting from registered data of 108 monitoring stations belonging to the National Air Quality System, using applied geostatistics through ordinary Kriging method, for its attribute of

representing phenomena whose values vary based on spatial location of points of observation.

The procedure is conducted in three stages: (1) exploratory data analysis, to determine whether it is necessary or not the data transformation and removal of bias in them. (2) Structural analysis to calculate the experimental semivariogram and adjustment to six selected theoretical mathematical models. (3) Prediction for obtaining prediction map and map standard errors of prediction.

The obtained results have allowed the creation of an estimated exposure surface of PM<sub>10</sub> values, which among its main characteristics, presents a spatial distribution with strong regional bias, spatially concentrating in mining cluster territories and metropolitan areas. The largest prediction errors occur in low population density areas and, monitoring stations in the Atacama Desert and the Andes.

**Keywords:** Kriging, particulate matter, exposure to pollutants, geographical information systems.

---

**Manuel Fuenzalida-Díaz** (mfuenzal@uahurtado.cl). \* Departamento de Geografía, Universidad Alberto Hurtado, Chile. Trabajo realizado en el marco del proyecto FONDECYT N° 11121354, financiado por la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) de Chile.

## 1. INTRODUCCION

El contaminante atmosférico material particulado respirable de 10 micras ( $MP_{10}$ ) es actualmente reconocido como el principal agente de morbilidad y mortalidad para el sistema respiratorio (OMS, 2006). Los efectos por la exposición al contaminante  $MP_{10}$  en la población se manifiestan principalmente como: muertes prematuras, aumento en la frecuencia de cáncer pulmonar, síntomas respiratorios severos, irritación de vías respiratorias, bronco constricción, entre las más importantes (Pope et al. 1995; Bateson y Schwartz, 2004). Es un problema de salud que afecta indistintamente a toda la población, de países desarrollados o en vías de desarrollo, y la susceptibilidad puede variar dependiendo de la edad, siendo los niños y los ancianos los que más frecuentemente requieren de atención sanitaria de atención primaria o de alta complejidad (hospitalización).

A partir de una serie de consultas a expertos en salud pública, la Organización Mundial de la Salud (OMS) estableció en 2005 nuevos valores guía para las concentraciones de material particulado en el aire, fijando para  $MP_{10} = 20\mu\text{g}/\text{m}^3$  media anual y  $50\mu\text{g}/\text{m}^3$  de media en 24h., siendo el primer valor utilizado en estudios de exposición prolongada y el segundo en exposición breve. Estos valores son bastante más ambiciosos que las normas primarias de calidad del aire vigentes en Chile para  $MP_{10} = 50\mu\text{g}/\text{m}^3$  media anual y  $150\mu\text{g}/\text{m}^3$  de media en 24h. que guían la política pública concerniente a la declaración de Zonas Saturadas o Latente por contaminación de material particulado respirable  $MP_{10}$  y la definición de niveles de emergencia ambiental (1:alerta, 2:preemergencia y 3:emergencia) para ciertos días del año.

Existen tres tipos de fuentes emisoras de  $MP_{10}$ . La primera de ella corresponden a fijas y se dividen en areales (e.g. calefacción residencial) y puntuales (e.g. industrias). La segunda a móviles (e.g. locomoción). La tercera a fugitivas (e.g. polvo suspendido).

El acceso a la información de este contaminante ha mejorado notablemente durante la última década. Actualmente es posible consultar en línea información en el Sistema Nacional de Calidad del Aire (SINCA), dependiente del Ministerio del Medio Ambiente de Chile. La presente comunicación se enmarca en el Proyecto FONDECYT Iniciación N°11121354 de dos años de duración, y tiene por objetivo modelar espacialmente la exposición al contaminante  $MP_{10}$ , utilizando geoestadísti-

ca aplicada por medio del método Kriging ordinario, por su cualidad de representar fenómenos cuyos valores varían en función de la localización espacial de puntos de observación. Este será un insumo básico para en una ulterior investigación, contrastar si las altas concentraciones de este contaminante están asociadas con la morbi-mortalidad de la población por enfermedades respiratorias o cardiovasculares.

El desarrollo de la comunicación inicia con un apartado donde se tratan someramente algunas cuestiones teóricas sobre geoestadística, para inmediatamente explicitar la metodología adoptada en torno al modelamiento del kriging ordinario. Posteriormente se describen los resultados, para terminar con una sección de conclusiones.

## 2. CONTEXTO TEÓRICO Y ANTECEDENTES

Todo conjunto de datos que pueda ser referenciado a partir de un sistema de coordenadas con un punto de origen conocido (e.g. lat/long, UTM) se ven condicionados por la primera ley de geografía o principio de autocorrelación espacial, según el cual *todo se relaciona con todo, pero las cosas más próximas o cercanas se relacionan más que las distantes* (Tobler, 1970). La autocorrelación espacial viola un supuesto básico de muchos test estadísticos tradicionales, el cual supone que las observaciones sean independientes. La independencia en un conjunto de datos geográficos se daría si su localización no fuera significativamente distinta de una localización resultante de una asignación aleatoria para cada lugar  $i$  posible del mapa (Upton y Fingleton, 1985).

Como en la naturaleza esto rara vez ocurre, abordar la dependencia espacial de los datos llevó a desarrollar nuevos procedimientos que tuvieran en cuenta tal circunstancia. En el ámbito de la estadística, estos nuevos procedimientos se han englobado bajo un nuevo campo de conocimiento denominado geoestadística, donde el prefijo “geo” resalta el aspecto espacial de los problemas que estudia mediante la aplicación de métodos probabilísticos (Montero y Larraz, 2008:11).

Estos se aplican a la concentración de un elemento contaminante en el suelo, agua o atmósfera; la altitud topográfica en un punto del espacio geográfico; el número de árboles y su diámetro promedio en áreas de observación de un bosque; la ley de un mineral, entre otros, intentando conseguir una descripción matemática del fenómeno regionalizado (se extiende en el espacio geográfico y presenta una cierta

continuidad). Los Sistemas de Información Geográfica (SIG) actualmente presentan enormes ventajas para la aplicación de procedimientos de geoestadística, entre las cuales se encuentran herramientas para análisis exploratorio de datos espaciales y la generación de superficies continuas a partir de datos discretos, utilizando métodos de interpolación. En relación a este último punto, la geoestadística proporciona estimaciones probabilísticas de la calidad de la interpolación.

Numerosas son las investigaciones que han recurrido a la geoestadística para modelar la exposición a contaminantes atmosféricos, entre las cuales es menester citar a Briggs (2005), Chinery y Walker (2009), Setton, et al. (2010), Nadal, et al. (2011) las cuales concluyen que la variación espacial es predecible, y se pueden modelar a un alto grado de precisión, utilizando información sobre las fuentes de emisión.

Una cuestión importante a mencionar aquí, es que la mayoría de los métodos geoestadísticos sólo son óptimos si la variable de estudio sigue una distribución de probabilidad normal. Entre sus principales propiedades se encuentra que es continua, de modo que puede tomar un valor cualquiera (no sólo enteros); es una curva lisa en forma de campana, simétrica alrededor de la media de la distribución ( $\mu$ ); la desviación estándar de la distribución ( $\sigma$ ) es la distancia horizontal entre la media y el punto de inflexión de la curva (el punto donde la curva cambia de convexa a cóncava); el área bajo la curva es igual a uno y como es una distribución simétrica, la mitad del área está a la izquierda de la media y la otra mitad, a la derecha; el coeficiente de sesgo (asimetría horizontal) y de curtosis (grado de concentración que presentan los valores alrededor de la zona central de la distribución) es igual a cero (0).

Estos dos últimos parámetros son importantes, pues si la distribución de los datos no se asemeja a una distribución de probabilidad normal, esta deberá ser transformada. El tipo de transformación dependerá del grado de asimetría. De acuerdo con Webster y Oliver (2001) uno puede considerar una transformación de tipo raíz cuadrada si el coeficiente de sesgo (*skewness*) se encuentra entre 0,5 y 1, y de tipo logarítmico ( $\ln$  o  $\log$ ) si el coeficiente de sesgo excede a 1.

Una vez subsanado este requerimiento, Isaaks y Srivastava (1989) proponen que todo trabajo geoestadístico tiene que llevarse a cabo en tres etapas fundamentales:

1) *Análisis exploratorio de los datos*. Se estudian los datos muestrales sin tener en cuenta su distribución geográfica. Sería una etapa de aplicación de la estadística. Se comprueba la consistencia de los datos, eliminándose los erróneos e identificándose las distribuciones de las que provienen.

2) *Análisis estructural*. Estudio de la continuidad espacial de la variable. Se calcula el variograma, u otra función que explique la variabilidad espacial, y se ajusta al mismo un variograma teórico.

3) *Predicciones*. Estimaciones de la variable en los puntos no muestrales, considerando la estructura de correlación espacial seleccionada e integrando la información obtenida de forma directa en los puntos muestrales, así como la conseguida indirectamente en forma de tendencias conocidas.

### 3. MATERIALES Y MÉTODOS

#### 3.1. Fuentes de información

##### 3.1.1. Datos de PM<sub>10</sub>

Los datos oficiales de PM<sub>10</sub> fueron obtenidos del Sistema Nacional de Calidad del Aire (SINCA) -[www.sinca.mma.gob.cl](http://www.sinca.mma.gob.cl)-, dependiente del Ministerio del Medio Ambiente de Chile, que pone a disposición de la ciudadanía información de calidad del aire de todo el país en su parte continental, el cual se localiza entre los 17° 30' y 56° 30' de Latitud Sur, alineado en torno a los 70° de Longitud Oeste.

El contaminante PM<sub>10</sub> se registra en 155 estaciones de monitoreo que están distribuidas de forma desigual a lo largo del territorio nacional continental. Las zonas extremas (norte y sur) no presentan estaciones de monitoreo y en el resto de las regiones político administrativa el reparto es como sigue: Tarapacá: 12, Antofagasta: 25, Atacama: 12, Coquimbo: 13, Valparaíso: 28, Metropolitana: 11, O'Higgins: 11, Maule: 5, BioBio: 26, Araucanía: 4, Los Ríos: 3, Los Lagos: 4, Aysén:

1. La cuatro primeras regiones tienen mayoritariamente el origen de las emisiones contaminantes atribuibles a la actividad del cluster minero (faenas mineras, termoelectricas, fundiciones), las siguientes tres están asociadas a la actividad industrial y del transporte (privado y público), las cuatro últimas se imputan al uso residencial de

leña para cocinar y calefaccionar, con un fuerte arraigo cultural y precio accesible.

Producto de las distintas fechas de entrada en funcionamiento de las estaciones de monitoreo (desde 1992 hasta 2012), se procedió a inspeccionar el conjunto de estaciones con una clara meta: obtener el máximo de estaciones para un único año donde se registre  $\mu\text{g}/\text{m}^3$  de media anual, y dada la extensión territorial de Chile (4.400 Km. desde el límite Norte hasta el Estrecho de Magallanes), que no estén separados entre sí a una distancia superior a los 150 kilómetros. El resultado obtenido fueron 108 estaciones para el año 2008, distribuidas de la siguiente forma (ver figura n°1): Tarapacá: 9, Antofagasta: 20, Atacama: 8, Coquimbo: 11, Valparaíso: 25, Metropolitana: 11, O'higgins:8, Maule: 2, BioBio: 8, Araucanía: 2, Los Ríos: 2, Los Lagos: 2.

### 3.1.2. Otras fuentes

Junto a los datos de  $\text{PM}_{10}$  se ha utilizado cartografía digital sobre los límites políticos administrativos, ciudades y caminos principales procedentes del Sistema Integrado de Información Territorial (SIIT) de la Biblioteca del Congreso Nacional de Chile.

## 3.2. Técnicas de análisis

Las técnicas para la evaluación de modelos geoestadísticos aplicados a la exposición al contaminante atmosférico  $\text{PM}_{10}$  en Chile, se implementarán con la extensión *Geostatistical Analyst integrado en ArcMap*. Esta permite crear una superficie continua a partir de datos puntuales, de una capa raster o de polígonos. Para ello estudia en primer lugar la autocorrelación espacial de la variable a interpolar, luego calcula una función de interpolación que tiene en cuenta el grado y tipo de autocorrelación existente en la variable y finalmente utiliza el semivariograma para describir y elaborar modelos espaciales, prediciendo valores en puntos no muestrales y proporcionando una medida de valoración de los mismo (Cañada, 2006: 745-746).

### 3.2.1. Análisis exploratorio de los datos.

*Geostatistical Analyst > Explore Data > Histogram*. El propósito aquí es conocer si la distribución de los datos de  $\mu\text{g}/\text{m}^3$  de media anual de  $\text{PM}_{10}$  sigue una

distribución normal. Los parámetros que se obtienen corresponden a Count (número de datos), Min (dato menor), Max (dato mayor), Mean (Media), Std Dev (Desviación estándar), Skewness (Coeficiente de sesgo o asimetría), Kurtosis (curtosis), 1-st Quartile (primer cuartil), Median (Mediana) y 3-rd Quartile (tercer cuartil).

Para que la distribución sea normal o se aproxime, la media, y la mediana deben ser similares, se acepta una diferencia de una (1) unidad entre ella. El coeficiente de sesgo o asimetría debe presentar un valor entre 0 y 0,5. Si eso no se da se deben transformar los datos siguiendo la pauta de Webster y Oliver (2001): transformación de tipo raíz cuadrada si el coeficiente de sesgo se encuentra entre 0,5 y 1, y de tipo logarítmico (ln o log) si excede a 1.

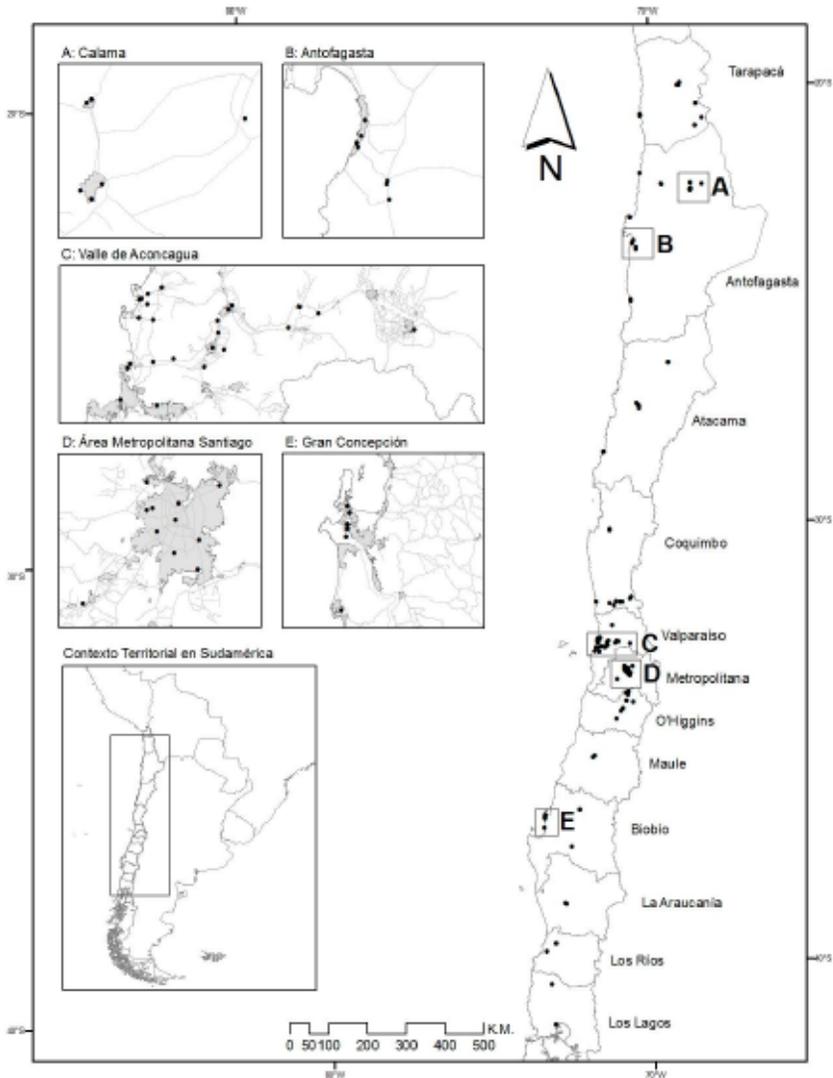
*Geostatistical Analyst > Explore Data > Trend Analysis.* La finalidad es descubrir la tendencia global que siguen los datos, es decir, si los datos pueden ser ajustados por funciones de primer (lineal, no tiene máximo ni mínimo), segundo (cuadrática, un máximo o un mínimo) o tercer grado (cúbica, dos máximos y un mínimo), para que luego en el posterior análisis estructural le indiquemos a la herramienta que sea removida. En términos representativos es un gráfico de dispersión visto en tres dimensiones, en donde los datos se proyectan en dos direcciones: N-S sobre un plano Y,Z y la dirección E-W sobre un plano X,Y.

### 3.2.2. Análisis estructural.

*Geostatistical Analyst > Geostatistical Wizard.* La meta es realizar un modelo geoestadístico con los datos.

*En Geostatistical methods seleccionaremos Ordinary Kriging > Prediction Map.* El kriging ordinario se basa en el modelo  $Z(s) = \mu + \varepsilon(s)$ , donde  $\mu$  es una constante desconocida. Se utiliza para datos que no tienen tendencia y cuya media es desconocida. El kriging ordinario usa semivariogramas o covarianzas (que son formas matemáticas de expresar la autocorrelación). También permite transformaciones de los datos, eliminación de tendencias y proporciona medidas de error (Cañada, 2006: 836-837).

Figura 1: Distribución territorial de estaciones de monitoreo del SINCA seleccionadas.



Fuente: Elaboración propia.

Para el modelado del semivariograma empírico se encuentran disponibles los modelos Circular (Circular), Spherical (Esférico), Tetraspherical (Tetraesferico), Pentaspherical (Pentaesférico), Exponential (Exponencial), Gaussian (Gaussiano), Rational Quadratic (covarianza cuadrática racional), Hole Effect (crecimiento no monótono), K-Bessel, J-Bessel y Stable. El modelo seleccionado influye en la pre-

dicción de los valores desconocidos, particularmente cuando la forma de la curva cerca del origen difiere significativamente. Cuanto más empinada la curva cerca del origen, más influencia a los vecinos más cercanos tendrán en la predicción.

Como resultado, la superficie de salida será menos suave. Cada modelo está diseñado para adaptarse a diferentes tipos de fenómenos con más precisión.

Nos referiremos brevemente a las características que se utilizan comúnmente para describir estos modelos. *Sill* (Meseta) corresponde al valor promedio de la varianza, *Range* (Rango) define la vecindad máxima sobre la cual los puntos de control deben ser seleccionados para estimar los puntos desconocidos de interés y finalmente *Nugget* (efecto pepita) corresponde al mínimo de varianza detectable.

### 3.2.3. Predicciones.

Cuando se predicen valores en lugares no muestrales, disponemos de Cross Validation (Validación Cruzada) para comprobar la validez del modelo que estamos utilizando. El modelo a seleccionar será aquel que mejor reproduzca los datos conocidos, por lo tanto cumplirá con las siguientes condiciones: *Root-Mean-Square -RMS-* (Raíz cuadrada del error medio): cuanto más pequeño sea, mejor serán las predicciones; *Average Standard Error -ASE-* (Error estándar promedio): pequeño, próximo a RMS, la variabilidad de la predicción se calcula correctamente y *Root-Mean-Square Standardized -RMSS-* (Raíz cuadrada del error medio): cerca de uno (1), los errores de la predicción son válidos.

Una vez que se presiona la tecla *Finish* (terminar) aparece el mapa de predicción. Seguido de esto, es posible obtener un mapa de errores. Para ello se debe hacer click con el botón derecho del ratón (mouse) sobre el mapa creado y elegimos la opción *Create Prediction Standard Error Map* (Crear mapa de predicción de error estándar) y se produce un mapa donde se observan los distintos niveles de error en la predicción. Las zonas más próximas a los puntos muestrales son las que puntúan un menor error estándar, sucediendo lo contrario con las más alejadas. Producto de lo que interesa conocer es el valor del espacio continuo alrededor cercano a la distribución de las estaciones de monitoreo, se calculará la confiabilidad del modelo como 100 menos el error mínimo.

Por lo tanto, para seleccionar el modelo que mejor modela nuestros datos,

es necesario aplicarles cada uno de ellos y escoger el que presente menor RMS, menor ASE, RMSS más cercano a uno y mayor porcentaje de confiabilidad.

## 4. ANÁLISIS DE RESULTADOS

### 4.1 Análisis exploratorio de los datos

Como se observa en el histograma de la figura 2a, los datos de  $\mu\text{g}/\text{m}^3$  de media anual de  $\text{PM}_{10}$  no siguen una distribución normal, pues estos presentan una asimetría positiva, donde la mayoría de las observaciones están a la izquierda de la proyección de la media. Los indicadores de normalidad *Mean* (Media) [53,344] y *Median* (Mediana) [47,94] no son iguales y en relación a *Skewness* (Coeficiente de sesgo o asimetría) [1,8583] y *Kurtosis* (curtosis) [8,5549] están bastante alejados de los valores ideales 0 y 3 respectivamente. Es por ello que se ha decidido transformar los datos. Siguiendo la pauta de Webster y Oliver (2001), la transformación ha sido del tipo logarítmico (ln o log) y los valores de normalidad se pueden apreciar en el histograma de la figura 2b, donde *Mean* [3,8878] y *Median* [3,8699] presentan una diferencia inferior a uno (1), que se considera aceptable y en relación a *Skewness* [-0,4337] y *Kurtosis* [5,416] están más cercanos a los valores ideales. Esta transformación se considerará al momento de realizar la predicción.

Para el análisis de tendencia, los gráficos que se observan en la figura 3 muestran la ausencia de tendencia en el comportamiento de los datos proyectados en el eje Y,Z. En particular, el trazado de la línea azul es una recta sin ninguna inclinación, por lo que se concluye que no hay tendencia de N-S. Los datos proyectados en el eje X,Z de dirección E-W, no presentan de forma clara las características típicas de las funciones de primer, segundo o tercer grado. Por lo tanto, se concluye que no es necesario modelar ni remover la tendencia en ninguna dirección.

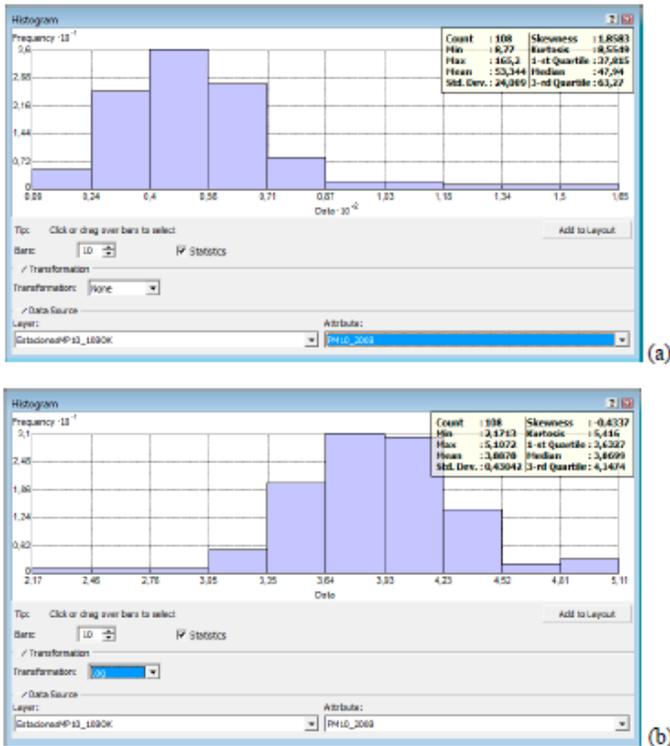
### 4.2 Análisis estructural y predicciones

El método geoestadístico utilizado corresponde al kriging ordinario, y la intención es crear un mapa de predicción. Los datos correspondientes a las 108 estaciones de monitoreo, al no seguir una distribución normal, se establece una transformación de tipo logarítmico y no es necesario remover o modelar tendencias.

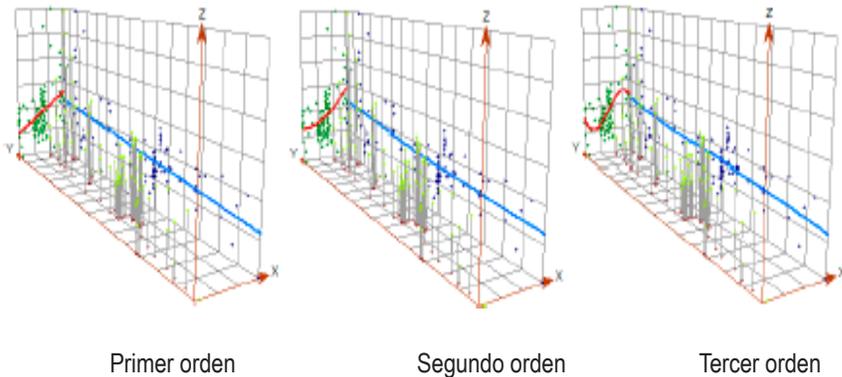
Los datos son representados en un semivariograma empírico y los modelos

matemáticos usados para su representación teórica (curva) corresponden a *Circular*, *Spherical*, *Tetraspherical*, *Pentaspherical*, *Exponential* y *Gaussian*, por ser lo que comúnmente se utilizan en temas de contaminación atmosférica. Se considera la misma variabilidad independiente de la dirección (fenómeno isotrópico).

Figura 2: Distribución de los datos de  $\mu\text{g}/\text{m}^3$  de media anual de  $\text{PM}_{10}$ .  
 (a) Distribución No-normal. (b) Distribución transformación logarítmica



Fuente: Elaboración propia a partir de datos SINCA.

Figura 3: Análisis de tendencia de los datos de  $\mu\text{g}/\text{m}^3$  de media anual de  $\text{PM}_{10}$ .

Fuente: Elaboración propia a partir de datos SINCA.

La elección del número de vecinos a utilizar corresponde a 15 y un mínimo de 5. La elipse dividida en cuatro sectores con ángulos de  $45^\circ$ . Los parámetros de los errores en la predicción se pueden observar en la tabla 1. De esta forma, seleccionar el modelo que mejor se adapta a nuestros datos corresponde al que cumple con los requerimientos de RMS pequeño, ASE pequeño próximo a RMS, RMSS cerca de 1 y un alto porcentaje de confiabilidad.

Tras un análisis de los valores, los modelos que presentan proximidad entre RMS y ASE, su RMSS cercano a uno, más exhiben un interesante grado de confiabilidad, corresponden a los modelos *Spherical* [0,78 – 1,024 – 84,15] y *Tetraspherical* [1,2 – 1,046 – 84,34]. Entre ellos, nos decantamos por el modelo *Spherical*, pues sacrificando apenas confiabilidad, los restantes parámetros entregan resultados bastante satisfactorios.

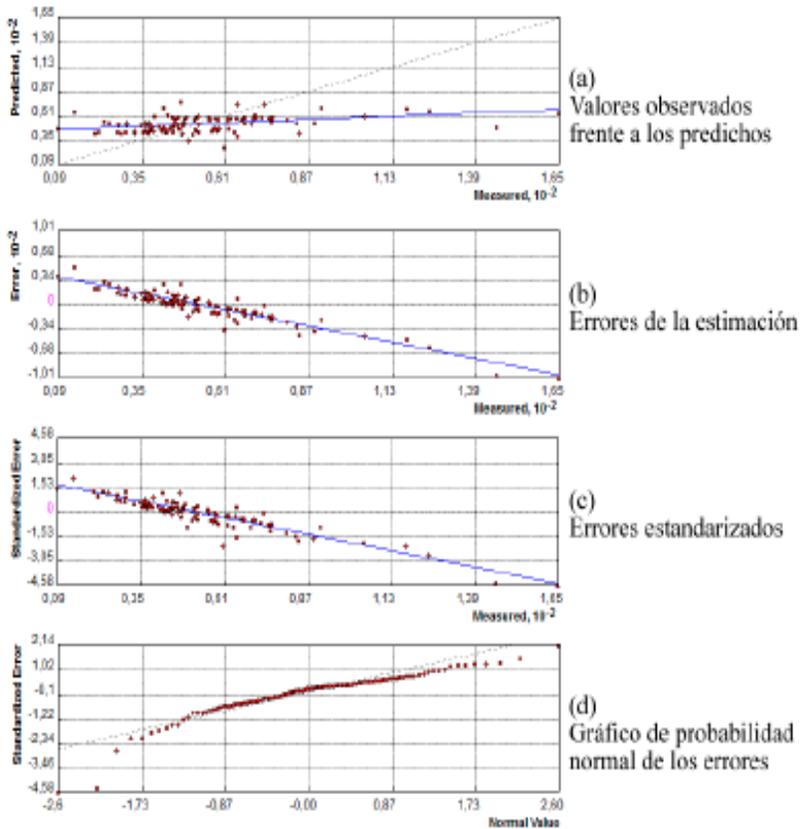
Tabla 1: Modelos utilizados y sus errores en la predicción

MODELO	RMS	ASE	RMSS	ERROR-m	CONFIABILIDAD
Circular	22,3	22,14	0,9945	16,19	83,81
Spherical	22,27	21,49	1,024	15,85	84,15
Tetraspherical	22,25	21,05	1,046	15,66	84,34
Pentaspheical	22,23	20,71	1,063	15,55	84,45
Exponential	21,99	19,62	1,112	14,78	85,22
Gaussian	22,97	21,28	1,08	14,19	85,81

Fuente: Elaboración propia.

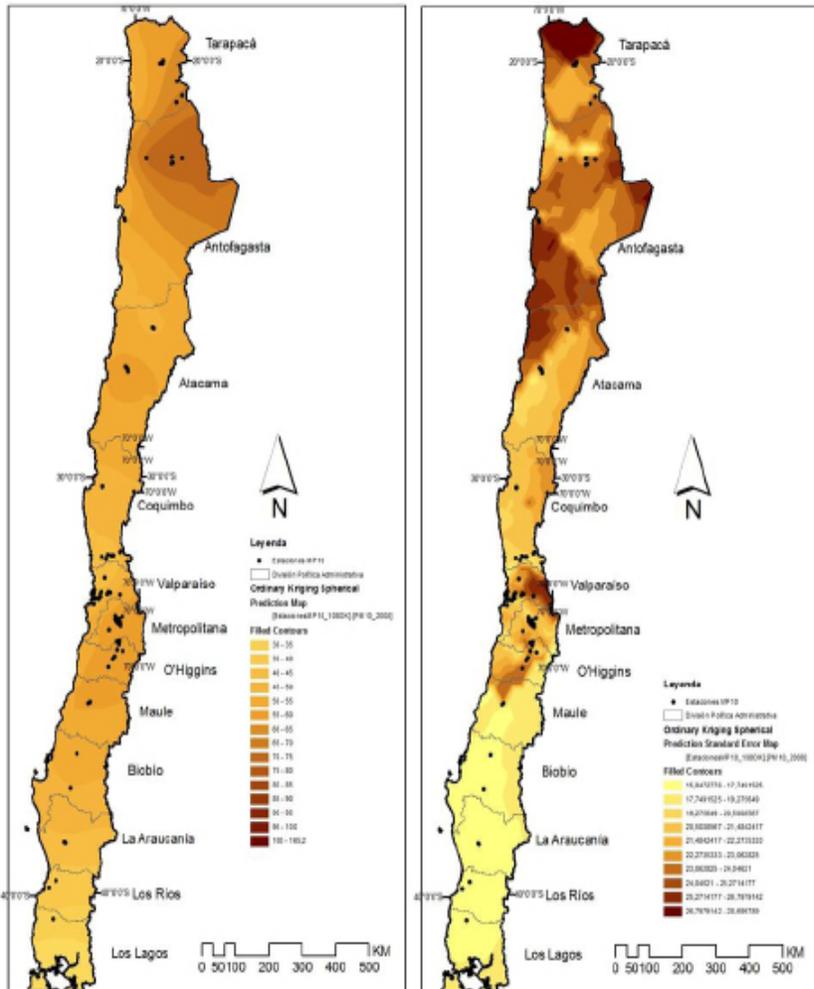
Los valores pronosticados se pueden valorar en los gráficos de la figura 4 y el mapa resultante de predicción (superficie generada con el kriging ordinario) más el mapa de errores estándar de la predicción se puede observar en la figura 5. En ellos es posible contemplar que los niveles altos de exposición a  $MP_{10}$  están presentes de norte a sur, con áreas peligrosas en el entorno de las regiones de Antofagasta, Atacama, Metropolitana y O’ Higgins. En relación a la pérdida de confiabilidad de la predicción, esta se presenta más grave en el norte (desierto de Atacama) y en la parte este de la región de Valparaíso (cordillera de Los Andes). En ambos sectores, la baja densidad de las estaciones de monitoreo, condicionan los resultados.

Figura 4: Gráficos de valores pronosticados, modelo Spherical.



Fuente: Elaboración propia.

Figura 5: Mapa predicción (superficie generada con el kriging ordinario) y mapa de errores estándar de la predicción.



Fuente: Elaboración propia.

## 5. DISCUSIÓN Y CONCLUSIONES

La aplicación del método kriging ordinario para la predicción de datos de  $\mu\text{g}/\text{m}^3$  de media anual de  $\text{PM}_{10}$  resultó útil para representar los niveles de exposición a contaminación en función de la variabilidad espacial (autocorrelación) asegurando la mínima varianza. Como paso previo, el análisis exploratorio de los datos

fue indispensable para conocer si estos cumplían o no con los requisitos necesarios para que el mapa de predicción sea válido y entregue información adecuada.

Al representar los datos en un semivariograma empírico y ajustarlo a uno de los seis modelos matemáticos usados para su representación teórica, fue posible comparar los resultados de error en la predicción. Para nuestro caso, el más satisfactorio corresponde al modelo Spherical.

Los resultados obtenidos han permitido crear una superficie de valores estimados de exposición a  $PM_{10}$ , que entre sus principales características, presenta una distribución territorial con un fuerte sesgo regional, concentrándose espacialmente en territorios del cluster minero y áreas metropolitanas. Los mayores errores de predicción se dan en sectores de baja densidad poblacional y de estaciones de monitoreo, en el desierto de Atacama y la cordillera de Los Andes.

## BIBLIOGRAFIA

- Bateson, T. y Schwartz, J. (2004): "Who Is Sensitive to the Effects of Particulate Air Pollution on Mortality? A Case-Crossover Analysis of Effect Modifiers", en *Epidemiology*, Vol. 15, No. 2, p. 143-149.
- Briggs, D. (2005) "The Role of Gis: Coping With Space (And Time) in Air Pollution Exposure Assessment", en *Journal of Toxicology and Environmental Health, Part A* 68:13-14, p. 1243-1261.
- Cañada, R. (2006): "Análisis exploratorio de datos espaciales: gráficos de distribución", en Moreno, A. (ed.): *Sistemas y análisis de la información geográfica*. Madrid, Ra-Ma, p. 745-780.
- Cañada, R. (2006): "Técnicas de interpolación geoestadísticas: Kriging ordinario", en Moreno, A. (ed.): *Sistemas y análisis de la información geográfica*. Madrid, Ra-Ma, p. 823-854.
- Chinery R. y Walker, R. (2009): "Development of Exposure Characterization Regions for Priority Ambient Air Pollutants", en *Human and Ecological Risk Assessment: An International Journal* 15:5, p. 876-889.

- Isaaks, E. H. and Srivastava, R. M. (1989): *An Introduction to Applied Geostatistics*. Oxford University Press, New York, USA.
- Nadal, M., Cadiach, O., Kumar, V., Poblet, P., Mari, M., Schuhmacher, M. y Domingo, J. (2011): "Health Risk Map of a Petrochemical Complex through GIS-Fuzzy Integration of Air Pollution Monitoring Data", en *Human and Ecological Risk Assessment: An International Journal* 17:4, p. 873-891.
- Montero, J.M. y Larraz, B. (2008): *Introducción a la geoestadística lineal*. Netbiblio, La Coruña. España.
- OMS (2006): *Guías de calidad del aire de la OMS relativas al material particulado, el ozono, el dióxido de nitrógeno y el dióxido de azufre. Actualización mundial 2005*. Ediciones de la OMS, Organización Mundial de la Salud, Ginebra, Suiza.
- Pope, CA.; Burnett, RT.; Thun, MJ.; Calle, EE.; Krewski, D.; Ito, K. y Thurston, GD. (2002): "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution", en *JAMA*, Mar 6;287(9), p. 1132-1141.
- Setton, E., Keller, P., Cloutier-Fisher, D. y Hystad, P. (2010): "Gender Differences in Chronic Exposure to Traffic-Related Air Pollution—A Simulation Study of Working Females and Males", en *The Professional Geographer* 62:1, p. 66-83.
- Upton, G.J. y Fingleton, B. (1985): *Spatial data analysis by example, volume 1: Point pattern and quantitative data*. Wiley, Toronto, Canadá.
- Webster, R. y Oliver, M. A. (2001): *Geostatistics for environmental science*. John Wiley and Sons, LTD. Toronto, Canadá.