
# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

## ANÁLISIS ESPACIO-TEMPORAL DE LA CONTAMINACIÓN ATMOSFÉRICA DEL VALLE DE MÉXICO

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRA EN CIENCIAS

PRESENTA:
KARLA VIANEY PALACIOS RAMÍREZ

DR. CARLOS DÍAZ AVALOS
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA IIMAS, UNAM

CDMX, ENERO DEL 2018

# AGRADECIMIENTOS

*A mi asesor de tesis, el Dr. Carlos Díaz Avalos por la ayuda, paciencia y orientación que me brindó al realizar este proyecto, por su apoyo y amistad que me permitieron aprender más sobre temas relacionados con estadística, por ser un excelente profesor.*

*A mis sinodales, por su apoyo, tiempo brindado y comentarios que ayudaron al desarrollo del proyecto.*

*A mi papá que siempre me ha entendido y cuidado, por su cariño, por todos sus sacrificios.*

*A mi mamá por ser mi guía para poder llegar a este punto de mi vida, que con su ejemplo y palabras siempre me apoyó en todas las decisiones importantes, por ser la mejor mamá, te amo.*

*A mis hermanas por siempre estar ahí apoyándome y alentándome a seguir adelante, gracias por todos los buenos momentos.*

*A mi abuela que siempre fue un apoyo y una gran amiga.*

*A mi abuelita Santa que fue una gran consejera y un gran apoyo en los momentos difíciles.*

*A mis amigos, que han estado conmigo en malos y buenos momentos, me han ayudado a seguir adelante y a nunca dejar mis sueños.*

*A mi abuelo que fue el apoyo incondicional en todos los momentos de mi vida, su amor y cariño me hicieron seguir en los momentos más difíciles.*

# Índice general

# Capítulo 1

# INTRODUCCIÓN

Conocer los mayores factores de riesgo en la salud es fundamental para poder generar estrategias efectivas orientadas a la prevención de daños a la salud, así como al mejoramiento de la calidad de vida de las personas. Según estimaciones de la OMS, en el 2012 las muertes provocadas por insalubridad del medio ambiente alcanzaron cerca de una cuarta parte de las muertes totales a nivel mundial, de las cuales, cerca de dos terceras partes son atribuibles a la contaminación del aire. Entre los principales problemas de salud asociados a la contaminación ambiental están las infecciones respiratorias, enfermedades cardiovasculares, distintos tipos de cáncer, enfermedad pulmonar obstructiva crónica, asma, así como una serie de consecuencias en las condiciones de recién nacidos (Prüss-Ustün, Wolf, Corvalán, Bos, & Neira, 2016).

La contaminación del aire es generada por emisiones de sustancias a la atmósfera que alteran su composición y originan desequilibrios en la misma (SEMARNAT, 2013). Estos contaminantes se clasifican según su origen en naturales (con una participación de aproximadamente el 20% en las emisiones) y antropogénicos, estos últimos derivados de las actividades impulsadas por el hombre, destinadas a proporcionar los bienes y servicios necesarios a la sociedad son responsables de la parte antropogénica de la contaminación del aire (INECC & SEMARNAT, 2013). Las emisiones de contaminación del aire ocurren en muchas etapas en los ciclos de vida de los productos y servicios, es decir, desde la materia prima, extracción de materiales, adquisición de energía, producción y fabricación, uso, reutilización, reciclaje, hasta la eliminación final. Las emisiones resultantes contribuyen a una amplia gama de impactos en la salud y el medio ambiente.

Existen distintos factores que influyen en la dispersión y concentración de contaminantes en el aire y por ello resulta necesario que en su estudio se consideren factores meteorológicos, geográficos y de emisión para un entendimiento integral de la dinámica de la contaminación atmosférica. Entre los factores naturales con mayor impacto en la dispersión y concentración de contaminantes en la atmósfera se encuentra la temperatura, el relieve, la humedad, las precipitaciones, la elevación, la presión del aire y la velocidad del viento. Dentro de los factores humanos que afectan la dispersión y concentración de contaminantes en el aire se encuentran principalmente la densidad de población, el número de vehículos, el PIB y el porcentaje de urbanización (Dongsheng, Mei-Po, Wenzhong , & Shaojian, 2017).

Con una tendencia creciente a la urbanización, las ciudades necesitan atención primordial en el estudio y tratamiento de la contaminación ambiental, ya que presentan condiciones de mayor riesgo (SEMARNAT, 2013). De acuerdo con datos de la ONU, la Ciudad de México es una de las 28 megaciudades del mundo que exceden los 10 millones de habitantes, concertándose como una de las zonas urbanas más densas del mundo (ONU, 2014). Respecto a la contaminación atmosférica, la Ciudad de México ha registrado niveles superiores a los sugeridos como óptimos (OMS, 2005) en términos de la concentración media anual de distintos contaminantes (OMS, 2006), por lo que los problemas de contaminación atmosférica en esta área son más sobresalientes que en otras regiones del país y resulta fundamental un estudio integral de la participación de estos contaminantes en los factores de riesgo para la salud de su población, con la finalidad de generar acciones de prevención y promoción de la salud en la región.

Los contaminantes atmosféricos son muy variables y de composición compleja (OMS, 2006), pudiéndose encontrar en forma gaseosa o en forma de partículas (ya sean líquidas o sólidas). Los principales indicadores de contaminación del aire utilizados por agencias ambientales a nivel mundial para la determinación de la calidad del aire son las concentraciones de material particulado, de ozono y de dióxido de nitrógeno por unidad geográfica (OMS, 2006). Por esta razón, casi la totalidad de los países tienen puntos de monitoreo distribuidos a lo largo de su extensión territorial que miden el nivel de concentración de estos contaminantes en el aire.

El PM o partículas suspendidas ha sido identificado en diversos estudios (An-

derson J., Thundiyil J. & Stolbach A., 2012) como el contaminante que afecta a más personas y que tiene mayores repercusiones en la salud. El material particulado (PM) está compuesto por distintos contaminantes y consiste en una mezcla compleja de partículas sólidas y líquidas de sustancias orgánicas e inorgánicas suspendidas en el aire, principalmente sulfatos, nitratos, amoníaco, cloruro de sodio, carbono, mineral, polvo y agua; es uno de los indicadores más utilizados en diversos estudios de contaminación en vista de que se pueden conocer las fuentes de emisión a partir de su composición (Zhang, R., et al, 2013).

En la mayor parte de los estudios epidemiológicos que consideran el PM se toman en cuenta únicamente tamaños pequeños del mismo (10 μm o menos), ya que tamaños más grandes son poco comunes y tienen una corta existencia en la atmósfera debido al efecto de la gravedad y corrientes de aire sobre ellos. Las partículas se identifican de acuerdo a sus diámetros aerodinámicos como PM10 o PM2.5 (diámetros aerodinámicos iguales o menores que 2.5 μm). La importancia de considerar tanto PM10 como PM2.5 radica en que este último puede ser un mejor indicador sobre la participación de fuentes antropogénicas en la generación de contaminantes en el aire. Esto se debe a que el papel de fuentes naturales de emisión de PM2.5 es menor que en PM10 y el primero podría ser un indicador guía primordial para generar políticas en materia ambiental (OMS, 2006).

El primer paso para instituir medidas de control para la contaminación del aire es tener una idea clara de la situación actual de la contaminación a través del monitoreo de la misma. La medición de la calidad del aire y la comprensión de sus impactos proporcionan una sólida base científica para su gestión y para el control de las fuentes de contaminación del aire. Por lo tanto, la calidad del aire o el control de la contaminación desempeña un papel vital en el desarrollo de políticas y estrategias, para medir el cumplimiento de los valores de referencia y el seguimiento del progreso hacia metas u objetivos ambientales. Se debe dedicar un esfuerzo considerable a la medición sistemática de niveles de contaminación del aire en diferentes escalas, desde locales hasta globales.
El objetivo final del monitoreo de la contaminación del aire es recopilar datos confiables que puedan ser utilizados por científicos, diseñadores de políticas y planificadores, a fin de permitirles contar con información de calidad para tomar decisiones informadas sobre la gestión y la mejora de la calidad general del medio ambiente. La selección del sitio es una parte muy importante de

cualquier programa de monitoreo de la contaminación del aire y la ubicación de una estación de monitoreo está directamente relacionada con el área que representarían los datos.
Identificar importantes fuentes de contaminación que contribuyen a las concentraciones ambientales de contaminantes es esencial para desarrollar un plan de gestión de calidad del aire eficaz. Los modelos de calidad del aire utilizan técnicas matemáticas y numéricas para simular los procesos físicos y químicos que afectan a los contaminantes del aire a medida que se dispersan y reaccionan en la atmósfera.
A pesar de esto, la variabilidad espacial y temporal inherente al fenómeno dificulta el análisis y las inferencias para predecir tendencias espaciales y temporales, por lo que es necesario postular modelos que consideren e incorporen dicha aleatoriedad. Además, teniendo en cuenta que las observaciones son a menudo escasas, los modelos pueden usarse para hacer inferencias en localidades donde no hay información. Combinado con la información sobre población, los modelos pueden ser fácilmente utilizados para estimar la exposición y, en última instancia, dar una idea sbre su posible efecto en la salud.

En esta tesis se presentan los resultados del análisis estadístico espacial y temporal de la contaminación por PM2.5 dentro de la Ciudad de México para los años 2002 a 2015. En el primer y segundo capítulo se presenta una breve introducción a ciertos conceptos generales de utilidad, así como una breve introducción de la importancia del modelo propuesto, también se presentará el problema a resolver y las bases necesarias para lograrlo.
En los capítulos III y IV se tratan y definen los conceptos importantes para este proyecto, se muestran aspectos como el tipo de investigación, las técnicas y procedimientos que fueron utilizados para llevar a cabo dicha investigación.
En el capítulo V se presentarán los resultados obtenidos del análisis de datos y de la metodología propuesta.
Por último, en el capítulo VI se presentan las conclusiones a las que se llegó durante el desarrollo de esta investigación.

# Capítulo 2

# FUNDAMENTO DEL PROBLEMA

La motivación de este proyecto fue el interés por tener un modelo para la dinámica espacio temporal de las partículas suspendidas, que sirva para evaluar los niveles de exposición y su variabilidad espacio temporal, como parte de un proyecto global que busca evaluar los factores de riesgo de enfermedades cardio pulmonares relacionadas a la contaminación atmosférica. Se utilizarán métodos geo-estadísticos para tratar de resolver el problema, usando un contexto espacio-temporal.

Los contaminantes sobre los cuales se tiene un valor observado son:

1. El monóxido de carbono (CO) que es un gas incoloro, inodoro, no irritante pero sumamente tóxico. Se produce naturalmente por una serie de procesos, sobre todo por la oxidación parcial del metano (CH4) que se forma en la descomposición de la materia orgánica por fermentación.
2. El ozono (O3) es un compuesto gaseoso incoloro, que posee la capacidad de oxidar materiales. El ozono es un contaminante secundario que se forma mediante la reacción química del dióxido de nitrógeno (NO2) y compuestos orgánicos volátiles (COV) en presencia de la luz solar.
3. Material particulado respirable (PM10). Son partículas de diámetro menor o igual a 10 micrones (un micrón es la milésima parte de un milímetro).
4. Material particulado respirable (PM2.5). Son partículas de diámetro menor o igual a 2.5 micrones.

Por su tamaño, el PM10 y PM2.5 son capaces de ingresar al sistema respiratorio del ser humano. Mientras menor sea el diámetro de estas partículas, mayor será

el potencial daño en la salud, así que nos concentraremos en el contaminante PM2.5 por ser el que parece afectar más al sistema respiratorio.
El modelo propuesto nos permitirá tener un valor del contaminante por Area Geo-Estadística Básica (AGEB), buscamos tener un indicador del contaminante en cada AGEB debido a que se tienen datos de casos respiratorios por entidad médica a esa escala y así se podrán ajustar modelos futuros para estimar los efectos de ciertos contaminantes en la salud.

# Capítulo 3

# MARCO TEÓRICO

## 3.1. Geoestadística

Considere el problema de predicción espacial, donde la región de interés es $D_s \subset \mathbb{R}^d$.
Sea $Y(\cdot) = \{Y(s) \ : s \in D_s\}$ los valores observados del fenómeno de interés en un conjunto de sitios $S_1, S_2, \ldots, S_m \in D$.
Ahora considere $Z(S_i) = Y(S_i) + \epsilon(S_i)$ para $i = 1, ..., m$, donde independientemente de $Y(\cdot)$, $\epsilon(\cdot) = \{\epsilon(s) \ : s \in D_s\}$ es un ruido blanco con media cero y varianza $\sigma^2_\epsilon \geq 0$.
Por simplicidad se asume que $Y(\cdot)$ y $\epsilon(\cdot)$ en $Z(\cdot) = Y(\cdot) + \epsilon(\cdot)$ son procesos gaussianos independientes, que pueden ser caracterizados mediante sus primeros dos momentos.
Suponiendo estacionariedad de segundo orden podemos decir que,

$$C_Y(h) \equiv Cov(Y(s), Y(s+h)) \quad \forall \ s, s+h \in D_s.$$

Definamos

$$\sigma^2_Y \equiv C_Y(0)$$

y

$$\lim_{h \to 0} \{C_Y(0) - C_Y(h)\} \equiv \sigma^2_0 \geq 0.$$

donde $\sigma^2_0 \leq \sigma^2_Y$.
Entonces, $\sigma^2_0$ representa la varianza del componente de micro escala del proceso $Y(\cdot)$.
Recordemos que como $Var(\epsilon(s)) = \sigma^2_\epsilon$, tenemos lo siguiente:

$$C_Z(h) \equiv Cov(Z(s), Z(s+h)) = \begin{cases} \sigma_Y^2 + \sigma_\epsilon^2 & si\, h = 0 \\ C_Y(h) & si\, h \neq 0 \end{cases}$$

lo cual implica que:

$$\lim_{h \to 0} \{C_Z(0) - C_Z(h)\} = \sigma_0^2 + \sigma_\epsilon^2 \equiv C_0 \geq 0$$

donde $C_0$ le llamaremos el efecto del nugget, el cual notemos que está compuesto por la suma de dos elementos no negativos.

### 3.1.1. Variograma estacionario

Los variogramas son un componente importante para la predicción lineal espacial óptima y miden la asociación espacial.

Sea $\{Y(s) \ : \ s \in D_s \subset \mathbb{R}^d\}$ un proceso espacial real valuado, definido en el dominio $D_s$ de un espacio euclideano $d$-dimensional $\mathbb{R}^d$.

Se dice que $Y(s)$ es un proceso estacionario intrínseco si:

$$E(Y(s+h) - Y(s)) = 0$$

$$Var(Y(s+h) - Y(s)) = 2\gamma_Y(h) \quad \forall\, s, s+h \in D_s$$

donde $2\gamma_Y(\cdot)$ es una función de la diferencia entre las localizaciones $s$ y $s+h$, la cual debe satisfacer lo siguiente:

$$\sum_{i=1}^{k} \sum_{j=1}^{k} \alpha_i \overline{\alpha_j} 2\gamma_Y(s_i - s_j) \leq 0$$

para todo entero positivo $k$, para todo conjunto de localizaciones $\{S_i \ : \ i = 1, ..., k\} \in D$ y cualquier conjunto de números complejos $\{\alpha_i \ : \ i = 1, ..., k\}$ tales que $\sum_{i=1}^{k} \alpha_i = 0$. Donde $\{\overline{\alpha_i} \ : \ i = 1, ..., k\}$ son los conjugados de $\{\alpha_i \ : \ i = 1, ..., k\}$.

La condición anterior asegura que la varianza sea no negativa, en cuyo caso diremos que $2\gamma_Y$ es un variograma válido.

La cantidad $2\gamma_Z(\cdot)$ es llamada variograma y $\gamma_Z(\cdot)$ semivariograma.

Cuando $2\gamma_Y(h)$ se puede escribir como función de $||h|| \equiv (h_1^2 + \cdots + h_d^2)^{1/2}$ para $h = (h_1, ..., h_d)' \in \mathbb{R}^d$ decimos que el variograma es isotrópico, en otro caso será anisotrópico.

El comportamiento del variograma cerca del origen es muy informativo acerca de las propiedades de continuidad del proceso aleatorio $Z(\cdot)$, y resultados similares se cumplen para la función de covarianza, algunos casos son los siguientes (Matheron (1971b, p.58)):

1. Si $2\gamma_Z(\cdot)$ es continuo en el origen, entonces $Z(\cdot)$ es $\mathcal{L}_2 - continuo$. Ya que $E\left(Z(s+h) - Z(s)\right)^2 \to 0$ si y sólo si $2\gamma_Z(h) \to 0$, conforme $||h|| \to 0$.

2. Si $2\gamma_Z(h)$ no se aproxima a cero conforme $h$ se aproxime al origen, entonces $Z(\cdot)$ no es $L_2 - continuo$ y es altamente irregular.

3. Si $2\gamma_Z(\cdot)$ es una constante positiva (excepto en el origen donde vale cero), entonces $Z(s_1)$ y $Z(s_2)$ son no correlacionados para cualquier $s_1 \neq s_2$, sin importar lo cercano que se encuentren; $Z(\cdot)$ es usualmente llamado ruido blanco.

Recordemos que un proceso aleatorio $Z(\cdot)$ se dice que es $\mathcal{L}_2 - diferenciable$ en $s$, si conforme $h_j \to 0$, $\{Z(s+h_j e_j) - Z(s)\}/h_j$ converge en $\mathcal{L}_2$, para $j = 1, ..., d$, donde $\{e_j : j = 1, ..., d\}$ es la base canónica en $\mathbb{R}^d$.
De este modo, si $\partial^2(2\gamma_Z(h))/\partial h_1^2, ..., \partial^2(2\gamma_Z(h))/\partial h_d^2$ existen y son finitas en $h = 0$, entonces $Z(\cdot)$ es $\mathcal{L}_2 - diferenciable$ para todo $s \in \mathbb{R}^d$.

#### 3.1.1.1. Modelos de variogramas isotrópicos (semivariogramas)

Características del semivariograma
En la Figura 3.1 se presenta el semivariorgama, mostrando la relación que existe entre los parámetros sill, nugget y rango.

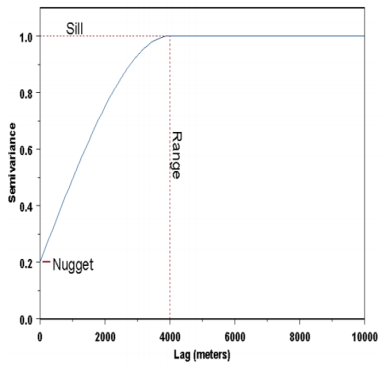


Figura 3.1: Semivariograma mostrando los parámetros sill, nugget y rango.

Rango: La distancia de rezago (lag) a la que el semivariograma (o componente de semivariograma) alcanza el valor del umbral. Es claro que, la autoco-

rrelación es esencialmente cero más allá del rango. El valor que el modelo de semivariograma obtiene en el rango (el valor en el eje y) se llama sill (umbral).

Sill (Umbral): Es el $\lim_{h\to\infty}\gamma(h)$ representando la varianza del campo aleatorio, es decir, el valor de semivarianza al que el variograma se nivela. También se usa para referirse a la "amplitud" de un cierto componente del semivariograma.

Nugget (Pepita): En teoría, el valor del semivariograma en el origen (lag=0) debe ser cero. Si es significativamente diferente de cero para rezagos muy cercanos a cero, entonces este valor del semivariograma se conoce como la pepita. Así, la pepita representa la variabilidad a distancias más pequeñas que el espaciado de muestra típico, incluido el error de medición. Por ejemplo, si el modelo de semivariograma intercepta el eje $y$ en 2, entonces el nugget es 2.

Usando $h$ para representar la distancia de rezago (lag), $a_{(\cdot)}, b_{(\cdot)}$ para representar el rango (práctico), $C_o$ para representar el nugget (pepita) y $C_{(\cdot)}$ para representar al umbral, algunos de los modelos más conocidos, son los siguientes:

- Modelo lineal (válido en $\mathbb{R}^d$, $d \geq 1$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + b_l||h|| & h \neq 0 \end{cases}$$

$\theta = (C_0, b_l')$, donde $C_0 \geq 0$ y $b_l \geq 0$. El valor de $C_0$ podría ser considerado el nugget (pepita). En la Figura 3.2 se presenta el modelo lineal.

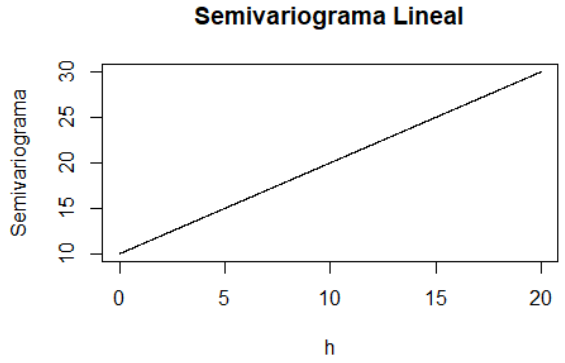


Figura 3.2: Modelo lineal

- Modelo esférico (válido en $\mathbb{R}^1, \mathbb{R}^2$ y $\mathbb{R}^3$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + C_s \left\{ \frac{3}{2} \left( ||h|| / a_s \right) - \frac{1}{2} \left( ||h|| / a_s \right)^3 \right\} & 0 < ||h|| \leq a_s \\ C_0 + C_s & h \geq a_s \end{cases}$$

$\theta = (C_0, C_s, a_S)'$, donde $C_0 \geq 0$, $C_s \geq 0$ y $a_s \geq 0$. En la Figura 3.3 se presenta el modelo esférico.

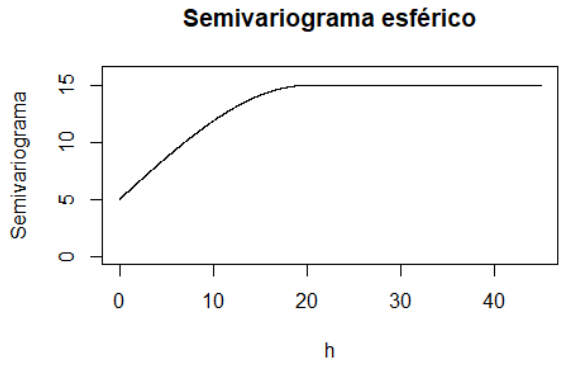


Figura 3.3: Modelo esférico

- Modelo exponencial (válido en $\mathbb{R}^d$, $d \geq 1$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + C_e \left\{ 1 - \exp\left( -||h|| / a_e \right) \right\} & h \neq 0 \end{cases}$$

$\theta = (C_0, C_e, a_e)'$, donde $C_0 \geq 0$, $C_e \geq 0$ y $a_e \geq 0$. En la Figura 3.4 se presenta el modelo exponencial.

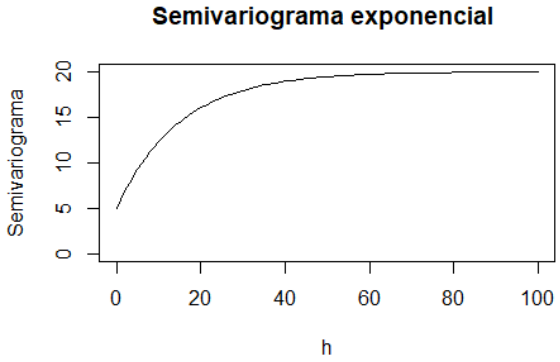


Figura 3.4: Modelo exponencial

- Modelo cuadrático racional (válido en $\mathbb{R}^d$, $d \geq 1$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + C_r||h||^2/\left(1 + ||h||^2/a_r\right) & h \neq 0 \end{cases}$$

$\theta = (C_0, C_r, a_r)'$, donde $C_0 \geq 0$, $C_r \geq 0$ y $a_r \geq 0$. En la Figura 3.5 se presenta el modelo cuadrático racional.

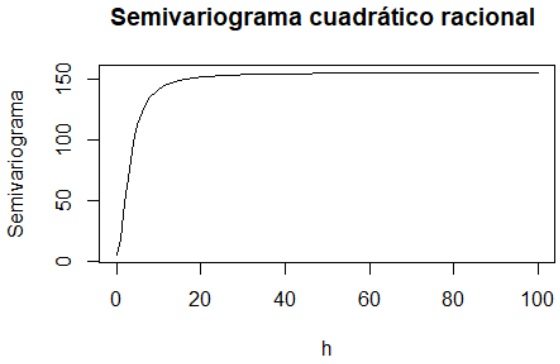


Figura 3.5: Modelo cuadrático racional

- Modelo de agujero (Wave model, válido en $\mathbb{R}^1, \mathbb{R}^2$ y $\mathbb{R}^3$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + C_w \left\{1 - a_w \sin\left(||h||/a_w\right)/||h||\right\} & h \neq 0 \end{cases}$$

$\theta = (C_0, C_w, a_w)'$, donde $C_0 \geq 0$, $C_w \geq 0$ y $a_w \geq 0$. En la Figura 3.6 se presenta el modelo de agujero.

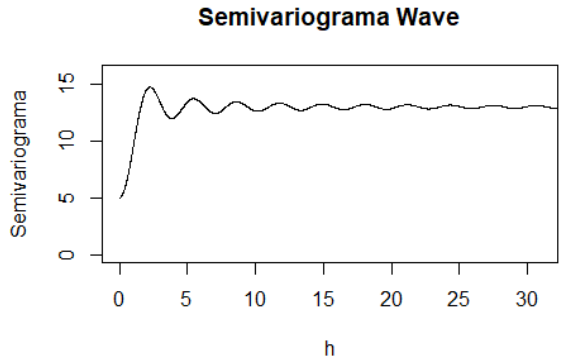


Figura 3.6: Wave model

Una condición más fuerte que el modelo de variograma debe satisfacer es:

$$2\gamma(h)/||h||^2 \to 0 \quad conforme \;\; ||h|| \to \infty$$

Así el siguiente modelo es un semivariograma válido:

- Modelo potencia (Power model, válido en $\mathbb{R}^d$, $d \geq 1$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + b_p||h||^\omega & h \neq 0 \end{cases}$$

$\theta = (C_0, b_p, \omega)'$, donde $C_0 \geq 0$, $b_p \geq 0$ y $0 \leq \omega < 2$.

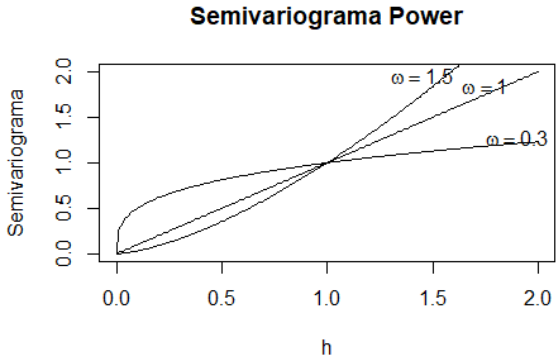


Figura 3.7: Power model

- Modelo Matérn (válido en $\mathbb{R}^d$, $d \geq 1$):

$$\gamma(h;\theta) = \begin{cases} 0 & h = 0 \\ C_0 + \sigma_0^2 I(||h|| = 0) + \sigma_1^2 \{2^{\theta_2 - 1}\Gamma(\theta_2)\}^{-1} \{||h||/\theta_1\}^{\theta_2} K_{\theta_2}(||h||/\theta_1) & h \neq 0 \end{cases}$$

$\theta = (\sigma_0^2, \sigma_1^2, \theta_1, \theta_2)'$, donde $\sigma_0^2 \geq 0$, $\sigma_0^2 \geq 0$, $\theta_1 > 0$, $\theta_2 > 0$ y $K_{\theta_2}$ es una función de Bessel modificada del segundo tipo de orden $\theta_2$.

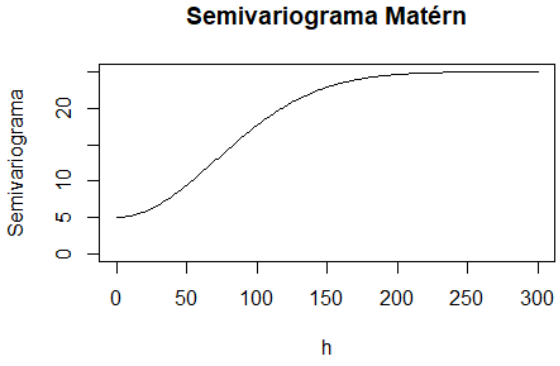


Figura 3.8: modelo Matérn

#### 3.1.1.2. Modelos de variogramas anisotrópicos (semivariogramas)

Decimos que el proceso $Z$ es anisotrópico cuando la dependencia entre $Z(s)$ y $Z(s+h)$ es función de la magnitud y dirección del vector $h$, de modo que el variograma no será función únicamente de las distancias entre localizaciones.
Muchas veces se puede corregir la anisotropía mediante una transformación lineal del vector $\text{lag}(h)$, entonces decimos que el variograma de $Z$ es anisotrópico geométricamente, es decir:

$$2\gamma(h) = 2\gamma^o(||Ah||) \quad h \in \mathbb{R}^d$$

donde $A_{d\times d}$ es una natriz y $2\gamma^o$ es una función real valuada.
En este caso, el espacio euclideano no es apropiado para medir distancias entre localizaciones, pero una transformación lineal sí lo es.

### 3.1.2. Función de covarianza estacionaria

Suponga que:

$$C_Y(h) = Cov(Y(s), Y(s+h)) \quad \forall\, s, s+h \in D_s$$

y además la media es constante, es decir,

$$E(Y(s)) = \mu \quad \forall\, s \in D_s.$$

Así, las condiciones anteriores definen las clases de procesos estacionarios de segundo orden en $D_s$, con función de covarianza estacionaria $C_Y(\cdot)$.
Donde $C_Y(\cdot)$ al ser una función de covarianza válida debe satisfacer:

$$\sum_{i=1}^{k}\sum_{j=1}^{k} \alpha_i \overline{\alpha_j} C_Y(s_i - s_j) \geq 0$$

para todo entero positivo $k$, para todo conjunto de localizaciones
$\{S_i \,:\, i = 1, ..., k\} \in D$ y cualquier conjunto de números complejos $\{\alpha_i \,:\, i = 1, ..., k\}$ (positiva definida). Donde $\{\overline{\alpha_i} \,:\, i = 1, ..., k\}$ son los conjugados de $\{\alpha_i \,:\, i = 1, ..., k\}$.
La condición anterior asegura que la varianza de combinaciones lineales del tipo $\sum \alpha_i Z(s_i)$ sea no negativa.

**Teorema de Bochner**

$C(\cdot)$ es una función positiva definida si y sólo si tiene representación espectral:

$$C(\cdot) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \cos(w'h) G(dw)$$

donde $G$ es una medida acotada positiva y simétrica; $G/C(0)$ es llamada función de distribución espectral. Si además $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |C(h)|dh < \infty$ podemos escribir $G(dw) = g(w)dw$ y así $g/C(0)$ es llamada densidad espectral.
Suponga $\{W(w) : w \in \mathbb{R}^d\}$ es un proceso estocástico $d$-dimensional, en los complejos con media cero y completamente aleatorios, donde:

$$\mathbb{E}(|W(dw)|^2) = G(dw)$$

y

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(dw) < \infty$$

Entonces el proceso:

$$Z(s) \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{iw's} W(dw)$$

tiene covariograma $C(h) = \mathbb{E}\left(Z(s+h) \cdot Z(s)\right)$, definido anteriormente.
El teorema nos asegura que cualquier función positiva definida $C(\cdot)$ corresponde al covariograma de un proceso estacionario de segundo orden.

### 3.1.3. Métodos de estimación del variograma

El variograma empírico proporciona una descripción de cómo los datos están relacionados (correlacionados) con la distancia. Dado que no es una función observable, es necesario estimar $C(h)$ a partir de los datos $Z(s_1), ..., Z(s_n)$. El estimador de momentos propuesto por Matheron (1962) es:

$$2\hat{\gamma}(h) = \frac{1}{|N(h)|} \sum_{N(h)} \left(Z(s_i) - Z(s_j)\right)^2, \quad h \in \mathbb{R}^d$$

donde

$$N(h) \equiv \{(s_i, s_j) \, : \, s_i - s_j = h; \; i, j = 1, ..., n\}$$

y $|N(h)|$ es el número de pares de observaciones separadas a una distancia $h$.
Notemos que: $2\hat{\gamma}(h) = 2\hat{\gamma}(-h)$.
Este estimador es válido independientemente de la media del proceso $Z(s)$.

## 3.2. Kriging (predicción espacial óptima)

El término kriging se refiere a hacer predicciones e inferencias sobre valores no observados del proceso aleatorio $Z(\cdot)$. Es una interpolación óptima basada en la

regresión contra los valores observados $z$ de los puntos de datos circundantes, ponderados según los valores de covarianza espacial.
Para ajustar un modelo kriging, necesitamos reemplazar el semivariograma empírico con un modelo de semivariograma aceptable. Debido a que el algoritmo de kriging necesitará acceder a valores de semivariograma para distancias de rezago (lags) distintas de las utilizadas en el semivariograma empírico. Más importante aún, los modelos de semivariograma utilizados en el proceso de kriging deben obedecer ciertas propiedades numéricas para que las ecuaciones de kriging puedan resolverse. (Técnicamente, el modelo de semivariograma necesita ser definido no negativo, para que el sistema de ecuaciones de kriging no sea singular.) Por lo tanto, los geoestadísticos eligen un modelo de semivariograma válido.

Ventajas del Kriging:

- Ayuda a compensar los efectos de la agrupación de datos, asignando a puntos individuales dentro de un cluster menos peso que a puntos de datos aislados (o, tratando los clusters más como puntos individuales).
- Da una estimación de la varianza del error de predicción (varianza del kriging), junto con la estimación de la variable $Z$, pero el mapa de error es básicamente una versión escalada de un mapa de distancia al punto de datos más cercano, por lo que no es único (Cressie 1993).
- La disponibilidad del error de predicción proporciona una base para la posible simulación estocástica de $Z(\cdot)$.

### 3.2.1. Kriging ordinario

La predicción espacial bajo las siguientes condiciones:

$$Z(s) = \mu + \delta(s), \quad s \in D,\ \mu \in \mathbb{R} \text{ y desconocida.}$$

donde:

$$Var(Z(s+h) - Z(s)) = 2\gamma_Z(h) \quad \forall\ s, s+h \in D_s$$

El predictor será de la forma:

$$p(Z;\ B) = \sum_{i=1}^{n} \lambda_i Z(s_i), \text{ sujeta a } \sum_{i=1}^{n} \lambda_i = 1$$

a continuación se supondrá $B = \{s_0\}$ y denotaremos al predictor como $\hat{Z}(s_0)$. La condición $\sum_{i=1}^{n} \lambda_i = 1$ es para evitar el sesgo en la predicción, es decir, el error de predicción $\sigma_e \equiv \mathbb{E}\left(Z(s) - \hat{Z}(s_0)\right)$ tiene que ser cero en promedio. Por

lo tanto, según Wackernagel (2003, p.29), tenemos que forzar los pesos $\lambda_i$ a sumar uno. Veamos que ocurre si la condición es cierta

$$\sigma_e \equiv \mathbb{E}\left(Z(s) - \sum_{i=1}^{n} \lambda_i Z(s_i)\right) = \mathbb{E}\left(Z(s)\right) - \sum_{i=1}^{n} \lambda_i \mathbb{E}Z(s_i) = \mu\left(1 - \sum_{i=1}^{n} \lambda_i\right) = 0$$

Existe una versión donde $\mu$ es conocida y no está la restricción de que los coeficientes del predictor lineal sumen uno, se conoce como Kriging Simple.
De modo que el predictor óptimo, será aquel que minimice la varianza del error de predicción:

$$\sigma_e^2 \equiv E\left(Z(B) - \hat{Z}(s_0)\right)^2$$

sobre la clase de predictores lineales $\sum_{i=1}^{n} \lambda_i Z(s_i)$, tales que $\sum_{i=1}^{n} \lambda_i = 1$.
Usando multiplicadores de Lagrange, buscaremos minimizar:

$$E\left(Z(s_0) - \sum_{i=1}^{n} \lambda_i Z(s_i)\right)^2 - 2m\left(\sum_{i=1}^{n} \lambda_i - 1\right)$$

respecto a $\lambda_1, ..., \lambda_n$ y $m$ el multiplicador de Lagrange.
De forma que los $\lambda_1, ..., \lambda_n$ están dados por:

$$\lambda' = \left(\gamma + 1\frac{(1 - 1'\Gamma^{-1}\gamma)}{1'\Gamma^{-1}1}\right)' \Gamma^{-1}$$

y

$$m = -\frac{(1 - 1'\Gamma^{-1}\gamma)}{1'\Gamma^{-1}1}$$

donde $\gamma = (\gamma(s_0 - s_1), ..., \gamma(s_0 - s_n))'$ y $\Gamma$ una matriz de $n \times n$, cuyo elemento $(i, j)$ es $\gamma(s_i - s_j)$.
La varianza del error de predicción minimizado usualmente es conocido como varianza del Kriging:

$$\sigma_k^2(s_0) = \gamma'\Gamma^{-1}\gamma - \frac{(1'\Gamma^{-1}\gamma - 1)^2}{1'\Gamma^{-1}1}$$

A partir de éste podremos construir intervalos de predicción donde, bajo el supuesto de que $Z(\cdot)$ es Gaussiano, tenemos que el intervalo al 95 de confianza es:

$$A \equiv \left(\hat{Z}(s_0) - 1.96\sigma_k(s_0),\ \hat{Z}(s_0) + 1.96\sigma_k(s_0)\right).$$

#### 3.2.1.1. Kriging en términos de la función de covarianza

Asumiremos el siguiente modelo:

$$Z(s) = \mu + \delta(s), \quad s \in D,\ \mu \in \mathbb{R} \text{ y desconocida.}$$

donde $\delta(\cdot)$ es un proceso estacionario de segundo orden, con media cero y covariograma $C_Z(h)$.
El predictor será de la forma:

$$p(Z;\, B) = \sum_{i=1}^{n} \lambda_i Z(s_i), \quad \sum_{i=1}^{n} \lambda_i = 1$$

De forma que los $\lambda_1, ..., \lambda_n$ están dados por:

$$\lambda' = \left( c + 1 \frac{(1 - 1'\Sigma^{-1}c)}{1'\Sigma^{-1}1} \right)' \Sigma^{-1}$$

y

$$m = -\frac{(1 - 1'\Sigma^{-1}c)}{1'\Sigma^{-1}1}$$

donde $c = (C(s_0 - s_1), ..., C(s_0 - s_n))'$ y $\Sigma$ una matriz de $n \times n$, cuyo elemento $(i, j)$ es $C(s_i - s_j)$.
De modo que $Z(\cdot)$ es un proceso estacionario de segundo orden.

### 3.2.2. Efectos de los parámetros del variograma en el Kriging

**Efecto Nugget** Recordemos que definimos al nugget como $C_0 = \lim_{||h|| \to 0} \gamma(h)$, donde $\gamma(\cdot)$ es el semivariograma de un proceso estacionario intrínseco $Z(\cdot)$ y además $\sigma_0^2 + \sigma_\epsilon^2 \equiv C_0$. En el kriging podemos considerarlo el valor distinto de cero que toma $\gamma(\cdot)$ cuando $h = 0$, producido por varias fuentes de error no explicado, por ejemplo, el error de medición.

**Rango** Puede ser interpretado como la distancia para la cual $Z(s)$ y $Z(s + a)$ son no correlacionados, es decir, las ubicaciones de muestra separadas por distancias más cercanas que el rango están autocorrelacionadas espacialmente, mientras que las ubicaciones más alejadas que el rango no lo están, lo cual por sí solo no nos permite determinar las vecindades del kriging.

**La relación Nugget : Sill (Pepita : Umbral)** Indica qué porcentaje de la varianza global se encuentra a una distancia más pequeña que el menor intervalo de retraso, y da una idea de la cantidad de varianza que considera o recupera el modelo.
Cuando los datos se distribuyen de manera independiente, es decir, carecen de estructura espacial, podemos esperar que haya poca diferencia en la varianza ($\gamma$) a cualquier lag. Sin embargo, cuando hay un patrón presente en la distribución,

podemos esperar que la varianza aumente con comparaciones de muestras cercanas, autocorrelacionadas, pero se nivelarán para formar un umbral cuando las muestras se vuelvan independientes. Para valores grandes de $h$, el variograma se nivela, lo que indica que ya no existe ninguna correlación entre los puntos de datos. El umbral debe ser igual a la varianza del conjunto de datos.

### 3.2.3. Kriging universal

La predicción espacial mediante Kriging Universal se hace suponiendo las siguientes condiciones:

$$Z(s) = \mathbb{E}(Z(s)) + \delta(s), \quad s \in D.$$

$$Var(Z(s+h) - Z(s)) = 2\gamma_Z(h) \quad \forall\, s, s+h \in D_s$$

donde $\mathbb{E}(Z(s))$ ya no es constante, sino una combinación lineal de funciones $\{f_o(s), ..., f_p(s)\}$ conocidas, además cada $f_j(s)$ puede ser escrita como función de localización de $s$.
Así, el kriging universal es una generalización del kriging ordinario, permitiendo que el valor medio del proceso sea una combinación lineal de funciones conocidas. Entonces, asumiremos el siguiente modelo:

$$Z(s) = \sum_{j=1}^{p+1} f_{j-1}(s)\beta_{j-1} + \delta(s) \quad s \in D.$$

donde $\beta = (\beta_0, ..., \beta_p)' \in \mathbb{R}^{p+1}$ es desconocido y $\delta(\cdot)$ es un proceso estacionario con variograma $2\gamma(\cdot)$ y $f_0(s) = 1$.
El predictor lineal óptimo, denotado por $\hat{p}(Z; B)$, es aquel que minimiza el error de predicción:

$$\sigma_e^2 = \mathbb{E}\left(Z(B) - p(Z; B)\right)^2$$

sobre $\lambda_1, \lambda_2, ..., \lambda_n$ sujeto a $\lambda' X = x'$.
Suponiendo que $B = \{s_0\}$, los pesos óptimos serán de la forma:

$$\lambda_U = \Gamma_U^{-1}\gamma_U$$

donde

$$\lambda_U \equiv (\lambda_1, ..., \lambda_n, m_0, .., m_p)'$$

con $m$ los multiplicadores de lagrange de la minimización con restricciones para un polinomio de orden $p$.

$$\gamma_U \equiv (\gamma(s_0 - s_1), ..., \gamma(s_0 - s_n), 1, f_1(s_0), ..., f_p(s_0))'$$

y $\Gamma_U$ es una matriz simétrica de $(n+p+1)\times(n+p+1)$ de la forma:

$$\Gamma_U \equiv \begin{cases} \gamma(s_i - s_j) & i = 1, ..., n \; j = 1, ..., n \\ f_{j-1-n}(s_i) & i = 1, ..., n \; j = n+1, ..., n+p+1 \\ 0 & i = n+1, ..., n+p+1 \\ & j = n+1, ..., n+p+1 \end{cases}$$

con $f_0(s) \equiv 1$.
Así los coeficientes de $\lambda$ están dados por:

$$\lambda' = \left\{\gamma + X(X'\Gamma^{-1}X)^{-1}(x - X'\Gamma^{-1}\gamma)\right\}'\Gamma^{-1}$$

y

$$m = -(x - X'\Gamma^{-1}\gamma)'(X'\Gamma^{-1}\gamma)^{-1}$$

donde

$$\gamma \equiv (\gamma(s_0 - s_1), ..., \gamma(s_0 - s_n))'$$

y $\Gamma_{n\times n}$ una matriz con elemento $\gamma(s_i - s_j)$ en la entrada $(i, j)$.
La varianza del Kriging será,

$$\lambda_U'\gamma_U(s_0)$$

suponiendo que el covariograma de $Z(\cdot)$ está bien definido, nos queda:

$$\sigma_k^2(s_0) = C(0) - 2\sum_{i=1}^{n}\lambda_i C(s_0 - s_i) + \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(s_i - s_j)$$

Los intervalos de predicción al 95 % de confianza ($\mathbb{P}(Z(s_0) \in A) = 0.95$), suponiendo que $Z(\cdot)$ es un proceso Gaussiano es:

$$A \equiv \left(\hat{Z}(s_0) - 1.96\sigma_k(s_0), \hat{Z}(s_0) + 1.96\sigma_k(s_0)\right)$$

En el caso de no suponer normalidad, usaremos la desigualdad de Chebyshev para construir los intervalos de confianza.

### Intervalos de confianza usando la desigualdad de Chebyshev

Usaremos una versión simplificada. Sea $X$ una variable aleatoria con media finita $\mu$ y varianza finita distinta de cero $\sigma^2$. Entonces, para cualquier número real $k > 0$, se cumple:

$$\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Notemos que cuando $k \leq 1$ se tiene $\frac{1}{k^2} \geq 1$ lo cual nos da una desigualdad trivial, pues siempre se cumple $\mathbb{P}(\cdot) \leq 1$.
Buscamos que $\frac{1}{k^2} = 0.10$, entonces $k = 3.16$.
Así el intervalo de confianza al 90 % de confianza será:

$$A \equiv \left(\hat{Z}(s_0) - 3.16\sigma_k(s_0), \hat{Z}(s_0) + 3.16\sigma_k(s_0)\right)$$

### 3.2.4. Block Kriging

Un problema común es tratar de predecir el promedio del proceso sobre un bloque $B$, cuya localización y geometría es conocido y tiene volumen $|B|$:

$$g(Z(\cdot)) = \int_B Z(s)ds/|B| \quad B \subset D$$

Para definir la integral correctamente, en la práctica aproximaremos la integral mediante sumas de Riemman; por ejemplo, en $\mathbb{R}^2$, sea $B$ el rectángulo $[a_1, a_2] \times [b_1, b_2]$ y considere

$$\sum_{i=1}^{l}\sum_{j=1}^{m} Z(u_i', v_j')(u_i - u_{i-1})(v_j - v_{j-1}),$$

donde $a_1 = u_0 < u_1 < \cdots < u_l = a_2$, $u_{i-1} \leq u_i' \leq u_i$, $b_1 = v_0 < v_1 < \cdots < v_m = b_2$, $v_{i-1} \leq v_i' \leq v_i$. Entonces, el límite en media cuadrática de la suma es la integral estocástica $\int_B Z(s)ds$.
El kriging universal, cuando se tiene $|B| > 0$ se convierte en:

$$\lambda_U = \Gamma_U^{-1}\gamma_U(B)$$

$$\sigma_k^2(B) = \lambda_U'\gamma_U(B) - \gamma(B, B)$$

donde

$$\gamma_U(B) \equiv (\gamma(B, s_1), ..., \gamma(B, s_n), 1, f_1(B), ..., f_p(B))',$$

$$\gamma(B, s_i) \equiv \int_B \gamma(u - s_i)du/|B| \quad i = 1, ..., n.$$

$$f_j(B) \equiv \int_B f_j(u)du/|B| \quad j = 1, ..., p$$

$$\gamma(B, B) \equiv \int_B \int_B \gamma(u - v)dudv/|B|^2$$

Aproximaremos la varianza por bloque $\gamma(B, B)$ mediante sumas de Riemman, de la siguiente forma:
Suponga que $r_j$ es el punto central en el bloque y considere $\{r_1, ....r_k\}$ como $k$ puntos dentro de cada bloque, se calculará el variograma en esas distancias, así la varianza por bloque queda estimada de la siguiente forma:

$$\overline{\gamma}(B_i, B_i) = \sum_{j=1}^{k} \frac{1}{k} \gamma(s_i, r_j)$$

## 3.3. Funciones de covarianza espacio-temporales

**Definición 3.3** Sea $\{f(u, v) : u, v \in D\}$ definida en $D \times D$, se dice que $f$ es no negativa definida, si para todo número complejo $\{a_i : i = 1, ..., m\}$, $\{u_i : i = 1, ..., m\}$ en $D$ y entero $m$, donde $\{\overline{a_i} : i = 1, ..., m\}$ son los conjugados de $\{a_i : i = 1, ..., m\}$, se tiene:

$$\sum_{i=1}^{m} \sum_{j=1}^{m} a_i \overline{a_j} f(u_i, u_j) \geq 0$$

Decimos que $f$ es una función de covarianza estacionaria espacio-temporal en $\mathbb{R}^d \times \mathbb{R}$, si satisface la Definición 3.3 y puede escribirse como:

$$f((s; t), (x; r)) = C(s - x;\ t - r) \qquad s, x \in \mathbb{R}^d\ t,\ r \in \mathbb{R}$$

Podemos considerar que la función de covarianza estacionaria es separable en espacio y tiempo.

- Estacionaria en el espacio:

$$cov\,(Y(s; t), Y(x; r)) \equiv C(s - x;\ t, r)$$

- Estacionaria en el tiempo:

$$cov\,(Y(s; t), Y(x; r)) \equiv C(s, x;\ t - r)$$

- Isotropía en el espacio:

$$cov\left(Y(s;t), Y(x;r)\right) \equiv C(||s-x||;\ t-r)$$

### 3.3.1. Separabilidad y simetría

**Definición 3.3.1** Un proceso aleatorio $Y(\cdot;\cdot)$ tiene una función de covarianza espacio-temporal separable si, para todo $s, x \in \mathbb{R}^d$, $t, r \in \mathbb{R}$ se tiene:

$$cov\left(Y(s;t), Y(x;r)\right) \equiv C^{(s)}(s,x) \cdot C^{(t)}(t,r)$$

donde $C^{(s)}$ es una función de covarianza espacial y $C^{(t)}$ una función de covarianza temporal.
Si ambas funciones de covarianza son estacionarias, se tiene:

$$cov\left(h;\tau\right) \equiv C^{(s)}(h) \cdot C^{(t)}(\tau)$$

**Definición 3.3.2** Un proceso aleatorio $Y(\cdot;\cdot)$ tiene una función de covarianza espacio-temporal simétrica si, para todo $s, x \in \mathbb{R}^d$, $t, r \in \mathbb{R}$ se tiene:

$$cov\left(Y(s;t), Y(x;r)\right) = cov\left(Y(s;r), Y(x;t)\right)$$

**Teorema 3.3.1 (Gneiting, Genton, y Guttorp (2007))**
Sea $\mu$ una medida finita, no negativa en el conjunto no vacio $\Theta$. Suponga que para cada $\theta \in \Theta$, $C^{(s_\theta)}$ y $C^{(t_\theta)}$ son funciones de covarianza estacionarias en $\mathbb{R}^d$ y en $\mathbb{R}$ respectivamente, también suponga que

$$\int_\Theta |C^{(s_\theta)}(0) \cdot C^{(t_\theta)}(0)| d\mu(\theta) < \infty$$

Entonces,

$$C(h;\tau) \equiv \int_\Theta C^{(s_\theta)}(h) \cdot C^{(t_\theta)}(\tau) d\mu(\theta) \qquad h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

es una función de covarianza estacionaria espacio-temporal.

### 3.3.2. Variograma espacio-temporal

El variograma espacio-temporal del proceso $Y(\cdot;\cdot)$ se define como:

$$Var(Y(s;t) - Y(x;r)) \equiv 2\gamma(s,x;\ t,r)$$

donde su versión estacionaria es $2\gamma(h;\tau)$; $h \in \mathbb{R}^d$,$\tau \in \mathbb{R}$.
Al igual que antes, a la expresión $\gamma$ se le llama semivariograma.

El proceso $Y(\cdot;\cdot)$ es estacionario intrínseco si tiene media constante y variograma estacionario. Su semivariograma estará dado por $\gamma(h;\tau) = C(0;0) - C(h;\tau)$.

El famoso teorema de Bochner (1955) establece que una función continua es definida positiva si y sólo si es la transformada de Fourier de una medida finita no negativa. Esto permite la siguiente caracterización de las funciones estacionarias de covarianza espacio-tiempo.

**Teorema de Bochner espacio-temporal**

Usaremos la representación espectral de $C(\cdot)$ en $\mathbb{R}^{d+1}$, enfatizando en espacio y tiempo de forma separada

$$C(h;\tau) = \int\int e^{ih'w+i\tau\xi} dF(w;\xi) \qquad h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

donde $F$ es una medida finita no negativa y simétrica en $\mathbb{R}^d \times \mathbb{R}$, llamada medida espectral. Si además $dF(w;\xi) = f(w;\xi)dwd\xi$ entonces, $f(\cdot;\cdot)$ es llamada densidad espectral y $\int f(w;\tau)dwd\tau = C(0;0)$. Esto ocurre si $C$ es integrable, es decir, $\int\int |C(h;\tau)|dhd\tau < \infty$ y así la transformada de Fourier de $C(\cdot)$ está dada por:

$$f(w;\xi) = (2\pi)^{-(d+1)} \int\int e^{-ih'w-i\tau\xi} C(h;\tau)dhd\tau$$

(Así, $C(\cdot)$ es la transformada de Fourier inversa de $f(\cdot)$ ).
Cualquiera de las dos funciones $f(\cdot)$ y $C(\cdot)$ nos sirven para representar la dependencia espacio-temporal del proceso.

La representación espectral nos permite construir funciones de covarianza espacio-temporales estacionaras válidas.

**Teorema 3.3.2 (Cressie and Huang (1999))**

Suponga que $C$ es una función simétrica, continua, acotada e integrable definida en $\mathbb{R}^d \times \mathbb{R}$. Entonces $C$ es una función de covarianza espacio-temporal estacionaria si y sólo si:

$$\alpha(w;\tau) = (2\pi)^{-d} \int e^{-ih'w} C(h;\tau)dh, \qquad \tau \in \mathbb{R}$$

es no negativa definida en $\tau$ (c.s), para toda $w \in \mathbb{R}^d$.
Ya que se tiene una familia definida para $\alpha(w;\tau)$, que cumplen ser no negativas definidas en $\tau$ y $\int |\alpha(w;\tau)|dw < \infty$, entonces dada $\tau$ la transformada inversa

de Fourier para $\alpha(\cdot;\tau)$:

$$C(h;\,\tau) \equiv \int e^{ih'w}\alpha(w;\tau)dw \qquad h \in \mathbb{R}^d, \tau \in \mathbb{R}$$

es una función de covarianza espacio-temporal. Generalmente, las familias construidas así no son separables, lo cual sólo ocurría cuando $\alpha(w;\tau) \equiv k(w)p(\tau)$.

## 3.4. Kriging espacio-temporal

Ya vimos anteriormente cómo se puede formular la predicción espacial lineal óptima (kriging) en función de las dependencias estadísticas espaciales de los procesos $Z(s,\omega)$. Estos se cuantifican típicamente a través de un variograma espacial o una función de covarianza espacial. En el contexto espacio-temporal, el variograma está definido de la siguiente forma:

$$2\gamma(s,x;\,t,r) \equiv var(Y(s;t) - Y(x;r)), \qquad s,x \in \mathbb{R}^d,\, t,r \in \mathbb{R}.$$

El objetivo de kriging es predecir $Y(s_0;t_0)$ en algún punto $s_0$ (o bien bloque $B_0$) y tiempo $t_0$ a partir de datos incompletos y ruidosos. Cualquier predictor lineal insesgado, $Y^*(s_0;t_0)$, de $Y(s_0;t_0)$, tiene la propiedad de que su error de predicción cuadrático medio, $\mathbb{E}\left(Y^*(s_0;t_0) - Y(s_0;t_0)\right)^2$ , se puede expresar en términos del variograma. De tal forma que el semivariograma, se puede obtener de la siguiente forma:

$$\gamma(s,x;\,t,r) = \frac{1}{2}\{C(s,s;\,t,t) + C((x,x;\,r,r)\} - C(s,x;\,t,r)$$

Anteriormente se proporcionaron modelos estacionarios válidos (es decir, no negativos-definidos) para $C(s,x;\,t,r)$. Esencialmente, la estacionariedad es una suposición de conveniencia para la estimación de parámetros y, en ocasiones, para la parte computacional; sin embargo, puede ser una suposición demasiado fuerte sobre grandes dominios espacio-temporales.

Desarrollaremos el kriging espacio-temporal basado en la especificación descriptiva general de la dependencia dada por la función de covarianza, $C(s,x;\,t,r)$. Lo hacemos para el caso de tiempo continuo, aunque es igualmente apropiado para tiempo discreto.

Suponga que los datos son,

$$Z(s_i;t_{ij}) = Y(s_i;t_{ij}) + \epsilon(s_i;t_{ij}), \qquad j = 1,...,T_i\ ,\ i = 1,...,m,$$

donde $\{\epsilon(s_i;t_{ij})\}$ es independiente de $Y(\cdot;\cdot)$ y representa el error de medición que, en adelante, se supone que es iid con media cero y varianza $\sigma^2_\epsilon$.

Sea $Z^{(i)} \equiv (Z(s_i;t_{ij}) : \ j = 1,...,T_i)'$; $i = 1,...,m$. Entonces buscamos predecir $Z(s_0;t_0)$ basados en la información dada por $Z^{(i)}$. Así, el predictor será:

$$Z^*(s_0;t_0) = \sum_{i=1}^{m}\sum_{j=1}^{T_i} \lambda_{ij} Z(s_i,t_{ij}) \equiv \lambda' Z$$

donde $Z \equiv \left(Z^{(1)\prime},...,Z^{(m)\prime}\right)'$ y $\lambda$ son optimizados de forma que minimicen el error de predicción cuadrático medio.

### 3.4.1. Kriging simple en el contexto espacio-temporal

El predictor lineal $Y^*(s_0;t_0)$, en el caso del Kriging simple (SK) toma la forma:

$$Y^*(s_0;t_0) = \sum_{i=1}^{m}\sum_{j=1}^{T_i} \lambda_{ij} Z(s_i,t_{ij}) + c \equiv \lambda' Z + c$$

donde $Z \equiv \left(Z^{(1)\prime},...,Z^{(m)\prime}\right)'$, $c$ y $\lambda$ son optimizados de forma que minimicen el error de predicción cuadrático medio.
Asumiremos en SK que la media del proceso $Y(\cdot,\cdot)$ es conocida,

$$\mu(s;t) \equiv \mathbb{E}(Y(s;t)), \qquad s \in D_s,\ t \in D.$$

Sea $C_Z \equiv var(Z)$, $c_0 \equiv cov(Y(s_0;t_0),Z)$ y $C_{0,0} \equiv var(Y(s_0;t_0))$, ahora si asumimos que $Y(\cdot,\cdot)$ y el error $\epsilon(\cdot,\cdot)$ son procesos gaussianos con matriz de covarianza $\Sigma_Y$ y $\Sigma_\epsilon$ respectivamente.
Se define $Z = Y + \epsilon$, entonces:

$$\begin{bmatrix} Y(s_0,t_0) \\ Z \end{bmatrix} \sim Gau\left(\begin{bmatrix} \mu(s_0;t_0) \\ \mu \end{bmatrix}, \begin{bmatrix} C_{0,0} & c_0' \\ c_0 & C_Z \end{bmatrix}\right) \qquad (3.4.1.1)$$

donde $\mu \equiv (\mu(s_i;t_{ij}) \ : \ j = 1,...,T_i \ , \ i = 1,...,m)'$ y $C_Z = \Sigma_Y + \Sigma_\epsilon$.
Así, la distribución condicional será,

$$Y(s_0,t_0)|Z \sim Gau\left(\mu(s_0;t_0) + c_0' C_z^{-1}(Z-\mu),\ C_{0,0} - c_0' C_Z^{-1} c_0\right)$$

Recordemos que el predictor lineal buscado es de la forma $Y^*(s_0;t_0) = \lambda' Z + c$, tal que minimice el error $\mathbb{E}\left(Y(s_0;t_0) - \lambda' Z - c\right)^2$, asumiendo la distribución Gaussiana conjunta en (3.4.1.1), es fácil ver que el predictor lineal óptimo será la media posterior al condicionar $Y(s_0,t_0)|Z$, entonces:

$$Y^*(s_0;t_0) = \mu(s_0;t_0) + c_0' C_Z^{-1}(Z-\mu) \qquad (3.4.1.2)$$

La varianza del kriging simple (SK) será simplemente la varianza del proceso al condicionar $Y(s_0, t_0)|Z$, dada por:

$$\sigma^2_{sk}(s_0, t_0) \equiv \mathbb{E}\left(Y^*(s_0; t_0) - Y(s_0; t_0)\right)^2 = C_{0,0} - c_0' C_Z^{-1} c_0 \quad (3.4.1.3)$$

Incluso si no asumieramos la distribución Gaussiana, podemos ver que los resultados adecuados son una generalización a los obtenidos en el kriging espacial. Es importante notar que se requiere el cálculo de la inversa de una matriz $C_Z = var(Z)$, que generalmente es complicado por la dimensión de los datos espacio-temporales, por lo cual es común asumir estructuras de covarianza separables en espacio-tiempo; sin embargo, los datos rara vez siguen este comportamiento, especialmente datos meteorológicos.

### 3.4.2. Kriging ordinario en el contexto espacio-temporal

Consideremos que el proceso $Y(\cdot, \cdot)$, tiene media constante y conocida $\mu$. Entonces $\mu = \mu\mathbf{1}$ donde $\mathbf{1}$ es un vector $T - dimensional$ de unos.
Recordando las ecuaciones del kriging ordinario (OK) en el caso espacial, el estimador generalizado de mínimos cuadrados es $\hat{\mu}_{gls} \equiv (1C_Z^{-1}1)1'C_Z^{-1}Z$, sustituyendo para $\mu(\cdot; \cdot) \equiv \mu$ en (3.4.1.2), se obtiene el predictor lineal óptimo para el kriging ordinario:

$$Y^*(s_0; t_0) \equiv \hat{\mu}_{gls} + c_0' C_Z^{-1}(Z - \hat{\mu}_{gls}1) \equiv \lambda' Z$$

donde $\lambda' \equiv \{c_0 + 1(1 - 1'C_Z^{-1}c_0)/(1C_Z^{-1}1)\}'C_Z^{-1}$. Claramente toma la forma del predictor en el kriging ordinario espacial.
La varianza será de la forma:

$$\sigma^2_{sk}(s_0, t_0) \equiv C_{0,0} - c_0' C_Z^{-1} c_0 + (1 - 1'C_Z^{-1}c_0)^2/(1C_Z^{-1}1)$$

### 3.4.3. Kriging universal en el contexto espacio-temporal

Consideremos el caso, $\mu(\cdot, \cdot)$ desconocida y de la forma:

$$\mu(s, t) = \sum_{h=1}^{p} a_h f_h(s, t)$$

,
donde $\{f_h(s, t),\ h = 1, ..., p\}$ son funciones conocidas y $a_h$ es un coeficiente constante.

De tal forma, las ecuaciones del kriging universal a resolver serán,

$$\left\{ \begin{array}{c} \sum_{j=1}^{n(s_0,t_0)} \lambda_j \gamma_\epsilon \left( (s_i,t_i) - (s_j,t_j) \right) + \sum_{k=1}^{p} \alpha_k f_k(s_i,t_i) = \gamma_\epsilon \left( (s_i,t_i) - (s_0,t_0) \right) \\ \forall i = 1, ..., n(s_0,t_0) \\ \sum_{i=1}^{n(s_0,t_0)} \lambda_i f_k(s_i,t_i) = f_k(s_0,t_0),\ \forall k = 1, ..., p \end{array} \right\}$$

Y la varianza de la predicción será:

$$\sigma^2_{sk}(s_0,t_0) \equiv \sum_{i=1}^{n(s_0,t_0)} \lambda_i \gamma_\epsilon \left( (s_i,t_i) - (s_0,t_0) \right) + \sum_{k=1}^{p} \alpha_k f_k(s_i,t_i)$$

# Capítulo 4

# ANÁLISIS

Como se mencionó en la introducción existe una asociación entre la incidencia de enfermedades cardio respiratorias con la contaminación del aire. Esto es particularmente notorio para la contaminación con partículas suspendidas. Aunque en la zona metropolitana de la Ciudad de México se han realizado diversos estudios al respecto, muchos de ellos no han considerado la variación espacial y temporal de las concentraciones de PM10 y PM2.5. El objetivo de este estudio es construir un modelo espacio temporal que permita explicar las variaciones espacio temporales de PM2.5 como un primer paso hacia la construcción de un modelo predictivo que permita evaluar los niveles de exposición a estos contaminantes de la población en la zona de estudio.

La información con que se cuenta y que será utilizada en este trabajo proviene de las estaciones de monitoreo que integran el Sistema de Monitoreo Atmosférico de la Ciudad de México, el cual es operado por la Secretaria del Medio Ambiente de la Ciudad de México. En esta red se reportan los datos diarios cada hora, para cada una de las estaciones de monitoreo que conforman la red, desde el año 2002 hasta la fecha. Cabe mencionar que no todas las estaciones han operado de manera continua y se han agregado y quitado estaciones a lo largo del periodo mencionado. La figura 4.1 muestra la localización geográfica de las estaciones de monitoreo y la tabla 4.1 presenta la lista de las estaciones junto con sus claves.

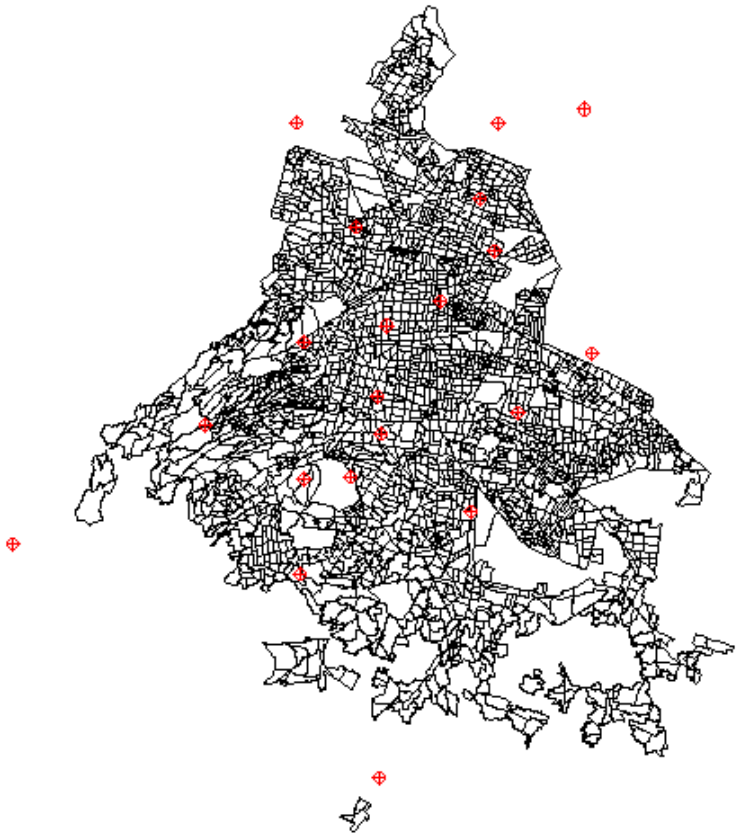

Figura 4.1: Estaciones de monitoreo y división en AGEBS de la CDMX

Las estaciones de monitorio son las siguientes:

| ESTACIÓN | DELEGACIÓN | LUGAR |
|---|---|---|
| ACO | Acolman | Estado de México |
| AJU | Tlalpan | CDMX |
| AJM | Tlalpan | CDMX |
| ATI | Atizapán de Zaragoza | Estado de México |
| BJU | Benito Juárez | CDMX |
| CAM | Azcapotzalco | CDMX |
| CCA | Coyoacán | CDMX |
| TEC | Gustavo A. Madero | CDMX |
| CHO | Chalco | Estado de México |
| COR | Xochimilco | CDMX |
| COY | Coyoacán | CDMX |
| CUA | Cuajimalpa de Morelos | CDMX |
| CUT | Tepotzotlán | Estado de México |
| DIC | Tlalpan | CDMX |
| EAJ | Tlalpan | CDMX |
| EDL | Cuajimalpa de Morelos | CDMX |
| FAC | Naucalpan de Juárez | Estado de México |
| GAM | Gustavo A. Madero | CDMX |
| HGM | Cuauhtémoc | CDMX |
| INN | Ocoyoacac | Estado de México |
| IZT | Iztacalco | CDMX |
| LPR | Tlalnepantla de Baz | Estado de México |
| LAA | Gustavo A. Madero | CDMX |
| IBM | Miguel Hidalgo | CDMX |
| LOM | Miguel Hidalgo | CDMX |
| LLA | Ecatepec de Morelos | Estado de México |
| MER | Venustiano Carranza | CDMX |
| MGH | Miguel Hidalgo | CDMX |
| MPA | Milpa Alta | CDMX |
| MON | Texcoco | Estado de México |
| MCM | Cuauhtémoc | CDMX |
| NEZ | Nezahualcóyotl | Estado de México |
| PED | Álvaro Obregón | CDMX |
| SAG | Ecatepec de Morelos | Estado de México |
| SJA | Gustavo A. Madero | CDMX |
| SNT | Magdalena Contreras | CDMX |
| SFE | Cuajimalpa de Morelos | CDMX |
| SHA | Miguel Hidalgo | CDMX |
| TAH | Xochimilco | CDMX |
| TLA | Tlalnepantla de Baz | Estado de México |
| TLI | Tultitlán | Estado de México |
| UIZ | Iztapalapa | CDMX |
| UAX | Coyoacán | CDMX |
| VIF | Coacalco de Berriozábal | Estado de México |

## 4.1. Análisis de los datos

El valor del contaminante PM2.5 se encuentra registrado en cada una de las estaciones de monitoreo desde el año 2002 hasta el 2017 con mediciones hechas cada hora. El primer paso es hacer promedios semanales de la información contenida en cada estación de monitoreo con el fin de poder visualizar el comportamiento espacio-temporal de manera más adecuada.
Se realizará un análisis descriptivo de los datos de forma espacial, para así comprender el comportamiento de la variable PM2.5. Después de esto procederemos a ajustar un modelo espacial a los datos, con el fin de comprobar una estructura separable en el espacio- tiempo.

### Análisis espacial de los datos

Se tomaron distintos intervalos de tiempo (mes de marzo, junio, septiembre, diciembre) y se realizó un análisis descriptivo (espacial) de los valores del contaminante, con el fin de ver si existe alguna relación en el tiempo cuando éste se mantiene constante.
Las figuras 4.2 a 4.11 muestran la distribución espacial de los valores de PM2.5 durante el mes de diciembre para los años 2005 a 2015, así como sus distribuciones empíricas marginales y su distribución con respecto a las coordenadas.

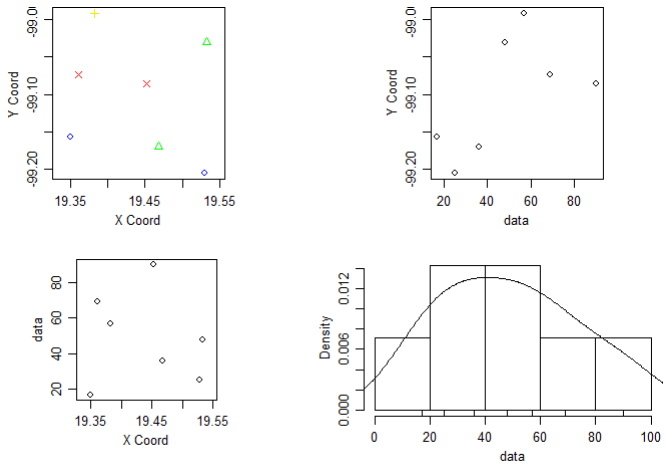


Figura 4.2: PM2.5 en las estaciones de monitoreo de la CDMX (2005)

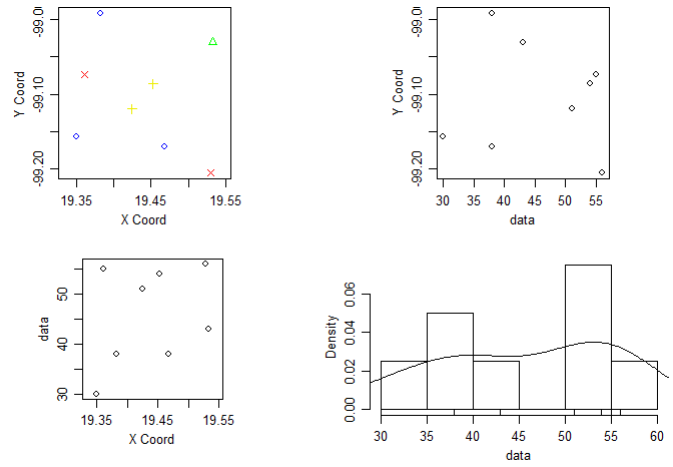


Figura 4.3: PM2.5 en las estaciones de monitoreo de la CDMX (2006)

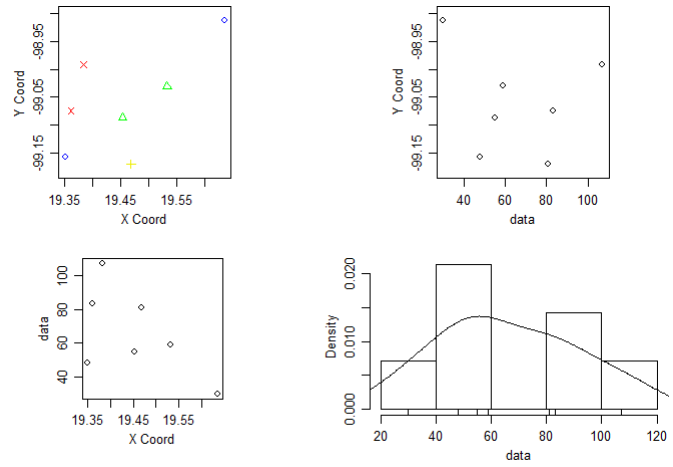


Figura 4.4: PM2.5 en las estaciones de monitoreo de la CDMX (2007)

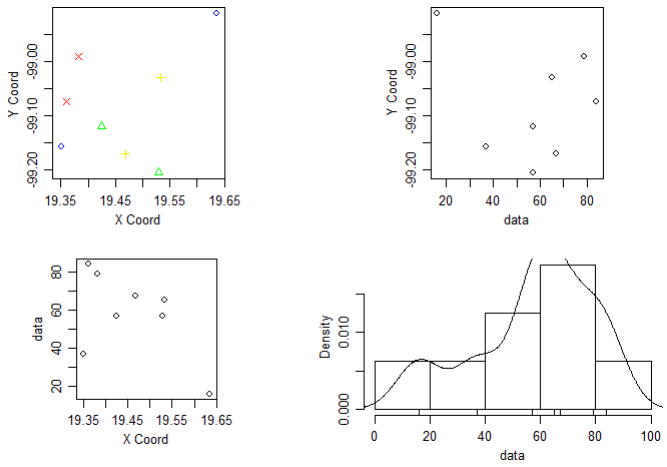


Figura 4.5: PM2.5 en las estaciones de monitoreo de la CDMX (2008)

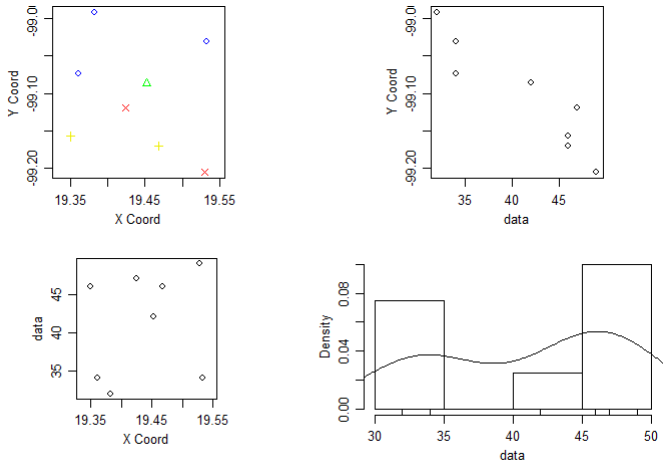


Figura 4.6: PM2.5 en las estaciones de monitoreo de la CDMX (2009)

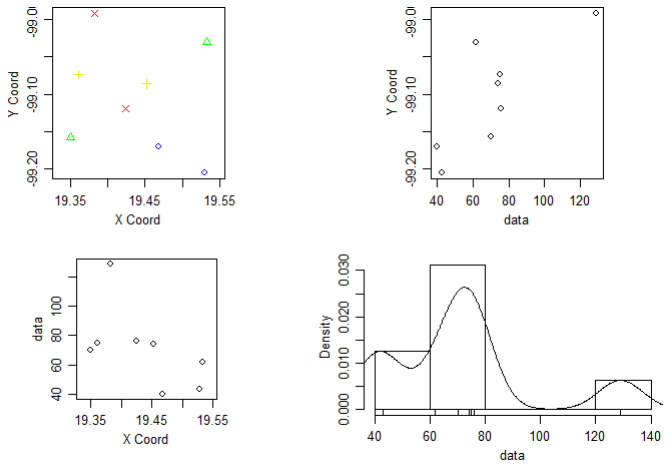


Figura 4.7: PM2.5 en las estaciones de monitoreo de la CDMX (2010)

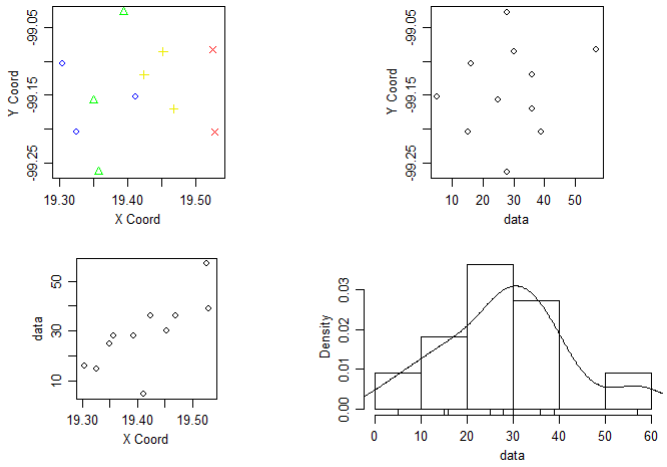


Figura 4.8: PM2.5 en las estaciones de monitoreo de la CDMX (2012)

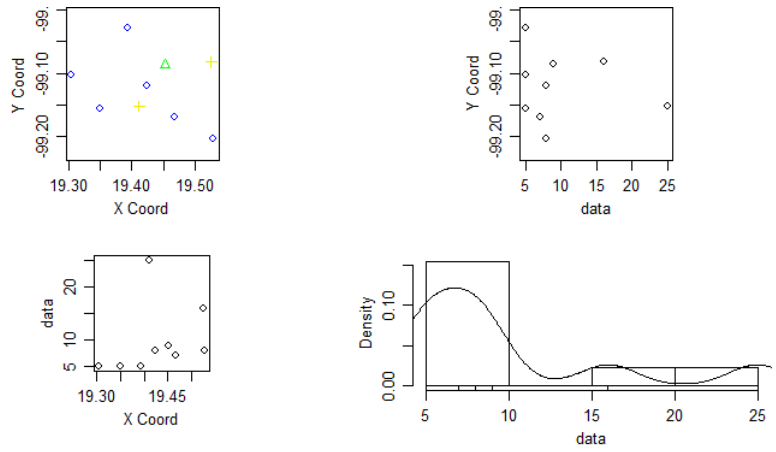


Figura 4.9: PM2.5 en las estaciones de monitoreo de la CDMX (2013)

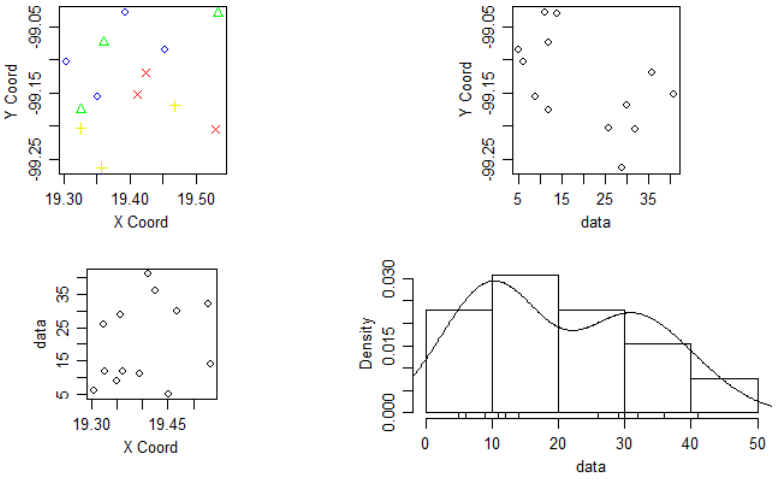


Figura 4.10: PM2.5 en las estaciones de monitoreo de la CDMX (2014)

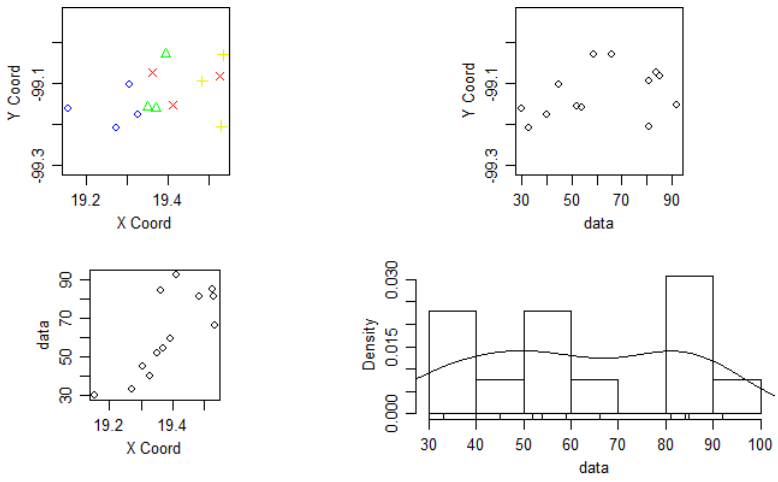


Figura 4.11: PM2.5 en las estaciones de monitoreo de la CDMX (2015)

Analizando el comportamiento del contaminante PM2.5 con el paso de los años, podemos observar un cambio evidente en los valores que dicho contaminante toma, dándonos visualmente una idea del cambio al paso de los años y cómo ha tomado valores cada vez más extremos. Aunque podríamos pensar que un año es un intervalo muy grande de tiempo para una variable que es medida cada hora y podría existir un comportamiento distinto a intervalos más cortos de tiempo, las gráficas de las figuras son de utilidad para darnos una idea de la existencia de posibles tendencias espaciales y temporales dentro de la región de estudio.
No tenemos evidencia clara de una estructura no separable en las variables espacio y tiempo, de modo que ajustaremos un modelo a una escala más pequeña para así poder concluir si se trata de un modelo separable o no separable.
En caso de ser separable, se modelará una estructura de dependencia en la variable tiempo y otra por separado de la variable espacio, es decir, una función de covarianza espacio-temporal separable de la forma

$$cov\left(Y(s;t), Y(x;r)\right) \equiv C^{(s)}(s,x) \cdot C^{(t)}(t,r)$$

para todo $s, x \in \mathbb{R}^d$, $t, r \in \mathbb{R}$, donde $C^{(s)}$ es una función de covarianza espacial y $C^{(t)}$ una función de covarianza temporal.
Debido a esto nos concentramos en ver qué ocurre en una escala más pequeña de tiempo, en este caso en un año, es decir, qué tanto se ve afectado el valor del contaminante al paso de los meses en un mismo año. Se toman cuatro meses representativos del año, elegidos de forma que cada uno ocurra en una estación

distinta del año, siguiendo este razonamiento tomamos el mes de marzo (primavera), junio (verano), septiembre (otoño), diciembre (invierno) de un año en específico, 2012-2013.

## 4.2. Ajuste del variograma espacial

El procedimiento siguiente se sigue para los cuatro meses mencionados antes, profundizaremos en el mes de diciembre para ejemplificar el método usado para ajustar el modelo de predicción.
Para el mes de diciembre del año 2012, su respectivo variograma es el siguiente:

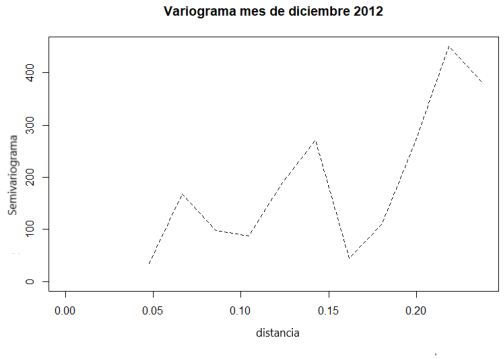


Figura 4.12: Variograma diciembre 2012

Se probaron varios modelos de variograma, para modelar el comportamiento del contaminante.
Entre los modelos probados (con ayuda de la libreria eyefit en R) , se encuentran:
Modelo gaussiano con los siguientes parámetros:

| PRUEBA | Modelo | Sill | $\phi$ | $\tau$ | Rango |
|---|---|---|---|---|---|
| 1) | Gaussian | 596.6 | 0.3 | 73.05 | 0.5192 |

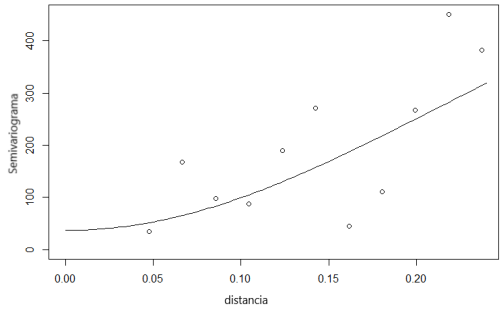


Figura 4.13: Ajuste modelo variograma Gaussiano

Modelo Wave con los siguientes parámetros:

| PRUEBA | Modelo | Sill | $\phi$ | $\tau$ | Rango |
|---|---|---|---|---|---|
| 2) | Wave | 389.62 | 0.1 | 85.23 | 1.4869 |

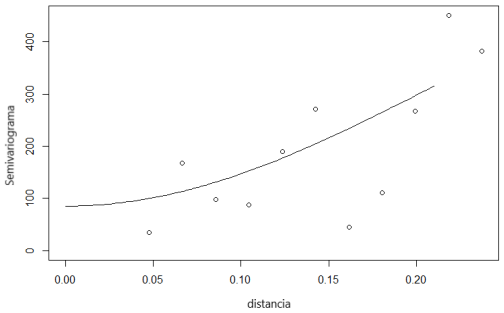


Figura 4.14: Ajuste modelo variograma Wave

En el siguiente modelo, se buscó tratar de estimar mejor a distancias pequeñas, por eso nos enfocamos en ver el ajuste a los primeros puntos de la gráfica.
Modelo Wave con los siguientes parámetros:

| PRUEBA | Modelo | Sill | $\phi$ | $\tau$ | Rango |
|---|---|---|---|---|---|
| 3) | Wave | 109.58 | 0.07 | 85.23 | 0.6272 |

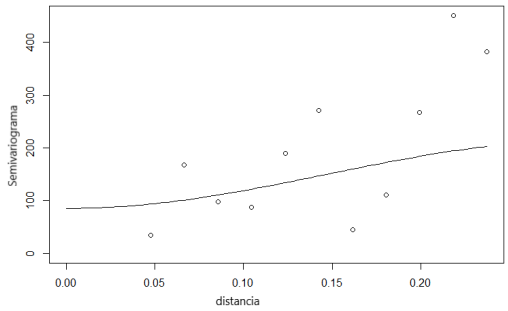


Figura 4.15: Ajuste modelo variograma Wave

## 4.3. Modelo de predicción (Kriging espacial)

La metodología a utilizar para realizar la predicción del valor del contaminate, es una modificación del Block Kriging para polígonos $B_{(i=1,...,2432)}$, que consideraremos como las agebs de la Ciudad de México, con un total de 2432, de modo que los bloques serán los siguientes:

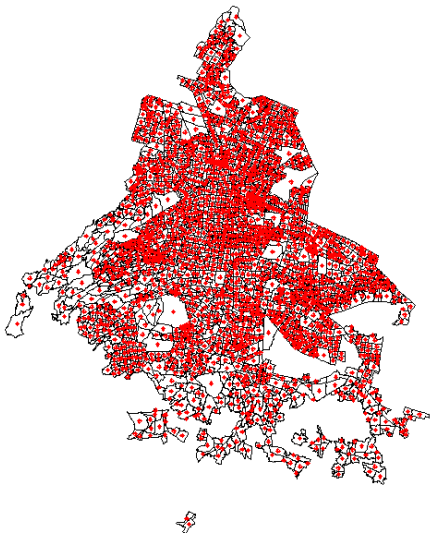

Figura 4.16: Localización de los centros de las AGEBS de acuerdo a datos de la INEGI

Se toma como observación puntual, el centro que cada una de las agebs, las cuales denotaremos por $\{s_1, s_2, ..., s_n\}$ para $n = 2432$.(Figura 4.16).

Recordemos que el modelo kriging para el bloque $|B| > 0$ es de la forma:

$$\lambda_U = \Gamma_U^{-1}\gamma_U(B)$$

$$\sigma_k^2(B) = \lambda_U'\gamma_U(B) - \gamma(B,B)$$

donde

$$\gamma_U(B) \equiv (\gamma(B,s_1),...,\gamma(B,s_n),1,f_1(B),...,f_p(B))'$$

$$\gamma_U(B) \equiv \int_B \gamma(u-s_i)du/|B|, \quad i = 1,...,n$$

$$f_j(B) \equiv \int_B f_j(u)du/|B|, \quad j = 1,...,p$$

$$\gamma(B,B) \equiv \int_B\int_B \gamma(u-v)dudv/|B|^2$$

Las ecuaciones del kriging varían respecto al kriging universal por la resta de $\gamma(B,B)$ en la varianza, así que por simplicidad de los métodos computacionales se estimará el kriging universal haciendo esta corrección en la varianza.
Por el momento nos enfocaremos en predecir $(\gamma(B_i,B_i))$, el variograma del proceso sobre un bloque $B_i$, cuya localización y geometría es conocido, de la siguiente forma:
Se simulan $k$ puntos dentro cada ageb, digamos $\{r_1,....r_k\}$ de modo que el variograma por bloque, definido anteriormente como la doble integral, será estimado por:

$$\overline{\gamma}(B_i,B_i) = \sum_{j=1}^{k}\frac{1}{k}\gamma(s_i,r_j)$$

El promedio del variograma (que se asumirá estacionario) evaluado en las distancias de los puntos simulados a los centros de las agebs. Los puntos simulados se observan en la Figura 4.17.

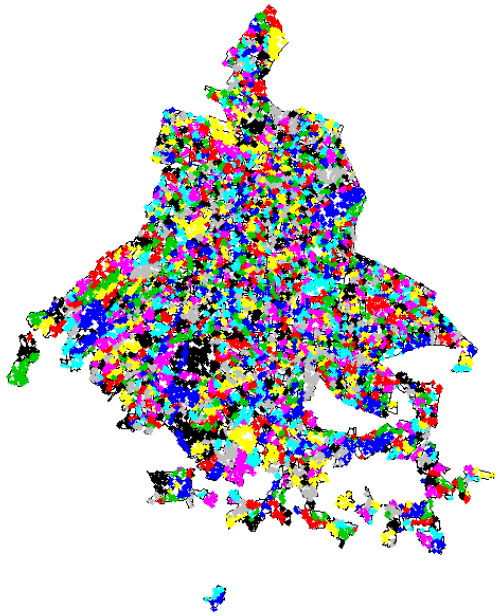

Figura 4.17: Puntos simulados dentro de las AGEBS

Por ejemplo, para la primera ageb se tienen los siguientes datos simulados (coordenadas UTM):

| | X | Y |
|---|---|---|
| 1 | 2780360 | 818108.7 |
| 2 | 2779756 | 817001.3 |
| 3 | 2779970 | 817835.8 |
| 4 | 2779522 | 816922.5 |
| 5 | 2779798 | 818209.6 |
| 6 | 2779752 | 817217.2 |
| 7 | 2779874 | 816820.7 |
| 8 | 2779940 | 818046.2 |
| 9 | 2779811 | 817073.6 |
| 10 | 2780458 | 817460.1 |

En la Figura 4.18 se observan los datos simulados y el punto rojo $s_1$ será el centro de la AGEB, con coordenadas $X = 2780141$, $Y = 817399.7$ a partir del

cual calcularemos la distancia.

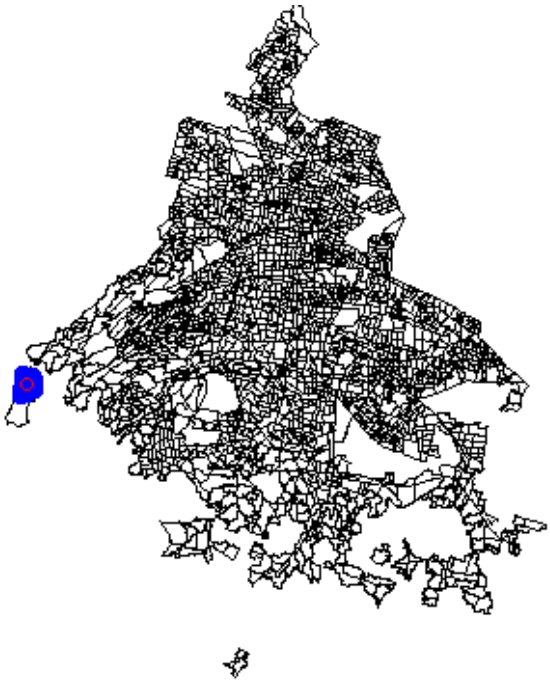

Figura 4.18: Puntos simulados dentro de la AGEB 1

Las distancias se muestran en la siguiente tabla.

| $k$ | Distancias $d(s_1, r_k)$ |
|---|---|
| 1 | 742.127 |
| 2 | 553.809 |
| 3 | 468.196 |
| 4 | 781.501 |
| 5 | 879.373 |
| 6 | 429.616 |
| 7 | 637.689 |
| 8 | 677.036 |
| 9 | 463.726 |
| 10 | 322.890 |

Se realizó el mismo procedimiento para cada una de las 2432 AGEBS. En la

siguiente tabla se muestran las distancias $h = \{h_1, ..., h_k\}$ calculadas desde los puntos centrales $s_i$ a los puntos simulados $\{r_1, ..., r_k\}$, no se muestran todos los bloques por simplicidad pues son 2432. En la Figura 4.18 se muestran gráficamente los puntos a partir de los cuales se calcularon las distancias, donde los círculos representan los puntos centrales $\{s_1, s_2, ..., s_n\}$ para $n = 2432$.

| $k$ | $d(s_1, r_k)$ | $d(s_2, r_k)$ | $d(s_3, r_k)$ | $d(s_4, r_k)$ | $d(s_4, r_k)$ |
|---|---|---|---|---|---|
| 1 | 742.127 | 74.711 | 750.471 | 147.513 | 284.997 |
| 2 | 553.809 | 260.452 | 957.466 | 102.438 | 115.090 |
| 3 | 468.196 | 156.952 | 1221.019 | 341.928 | 121.848 |
| 4 | 781.501 | 451.931 | 2213.170 | 339.886 | 83.138 |
| 5 | 879.373 | 394.096 | 713.792 | 131.892 | 80.728 |
| 6 | 429.616 | 245.956 | 1177.603 | 117.817 | 165.622 |
| 7 | 637.689 | 153.582 | 528.420 | 255.290 | 112.825 |
| 8 | 677.036 | 182.556 | 633.985 | 147.990 | 138.665 |
| 9 | 463.726 | 316.143 | 445.144 | 39.616 | 188.548 |
| 10 | 322.890 | 190.301 | 1187.265 | 144.692 | 52.825 |

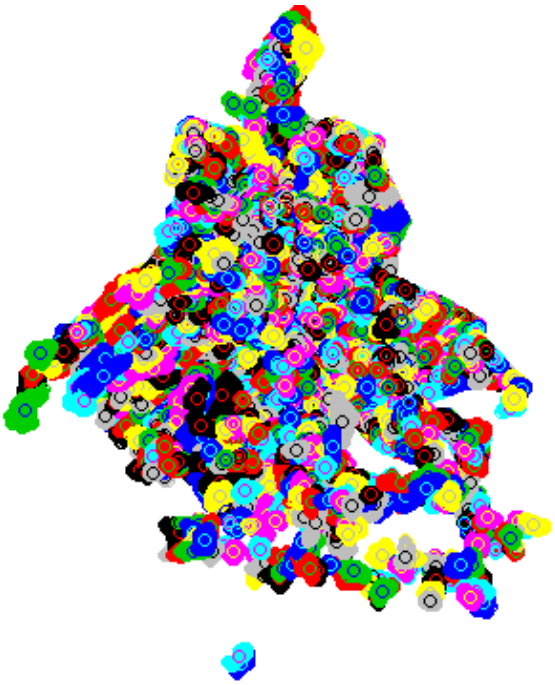

Figura 4.19: Puntos para calcular las distancias a los puntos centrales de las AGEBS

Como se mencionó el la sección 4.1, se probaron varios modelos de variograma para el mes de diciembre del año 2012, el primer modelo es un covariograma gaussiano,

$$C(h) = \sigma^2 \rho(h)$$

donde

$$\rho(h) = \exp(-(h/\phi)^2)$$

con parámetros $\sigma^2 = 596.6$ y $\phi = 0.3$.
Se calculó el valor del variograma en las distancias $h = \{h_1, ..., h_k\}$, para cada una de las 2432 agebs y después se estimó el valor del variograma por bloque mediante el promedio, es decir, $\overline{\gamma}(B_i, B_i) = \sum_{j=1}^{k} \frac{1}{k} \gamma(s_i, r_j)$.
Los resultados para las primeras 10 AGEBS se muestran en la siguiente tabla:

| $AGEB_i$ | $\overline{\gamma}(B_i, B_i)$ |
|---|---|
| 1 | 462.67 |
| 2 | 462.62 |
| 3 | 462.65 |
| 4 | 462.65 |
| 5 | 462.68 |
| 6 | 462.61 |
| 7 | 462.63 |
| 8 | 462.64 |
| 9 | 462.67 |
| 10 | 462.68 |

El segundo modelo propuesto es un variograma suave (Wave), el cual recordemos es de la forma:

$$\gamma(h) = \begin{cases} C_s \left( \frac{\sin(h)}{h} \right) & h > 0 \\ C_0 + C_s & h = 0 \end{cases}$$

donde $C_s$ es la meseta. O bien su covariograma dado de la siguiente forma:

$$C(h) = \sigma^2 \rho(h)$$

donde

$$\rho(h) = \left( \frac{\phi}{h} \right) \frac{\sin(h)}{\phi}$$

con parámetros $\sigma^2 = 389.62$ y $rango = 0.1$.
Se calcula el valor del variograma en las distancias $h = \{h_1, ..., h_k\}$, para cada una de las 2432 agebs y después se estima el valor del variograma por bloque mediante el promedio, es decir, $\overline{\gamma}(B_i, B_i) = \sum_{j=1}^{k} \frac{1}{k} \gamma(s_i, r_j)$.
Los resultados para las primeras 10 AGEBS se muestran en la siguiente tabla:

| $AGEB_i$ | $\overline{\gamma}(B_i, B_i)$ |
|---|---|
| 1 | 474.8520 |
| 2 | 474.8295 |
| 3 | 474.8495 |
| 4 | 474.8409 |
| 5 | 474.8833 |
| 6 | 474.9276 |
| 7 | 474.8583 |
| 8 | 474.8406 |
| 9 | 474.8583 |
| 10 | 474.8351 |

El tercer modelo probado es también un covariograma suave (Wave), pero con parámetros $\sigma^2 = 109.58$, $rango = 0.07$.
Se calcula el valor del variograma en las distancias $h = \{h_1, ..., h_k\}$, para cada una de las 2432 agebs y después se estima el valor del variograma por bloque mediante el promedio, es decir, $\overline{\gamma}(B_i, B_i) = \sum_{j=1}^{k} \frac{1}{k}\gamma(s_i, r_j)$.
Los resultados para las primeras 10 AGEBS se muestran en la siguiente tabla:

| $AGEB_i$ | $\overline{\gamma}(B_i, B_i)$ |
|---|---|
| 1 | 194.8112 |
| 2 | 194.8061 |
| 3 | 194.8092 |
| 4 | 194.8160 |
| 5 | 194.8227 |
| 6 | 194.8198 |
| 7 | 194.7948 |
| 8 | 194.8053 |
| 9 | 194.8106 |
| 10 | 194.8138 |

Utilizando validación cruzada sobre los distintos valores del covariograma, se ajustará el covariograma suave (Wave) con parámetros
$\sigma^2 = 109.58$, $rango = 0.07$.

Para el modelo Block Kriging que se busca ajustar, ya tenemos la constante $\overline{\gamma}(B_i, B_i)$ por cada bloque (AGEB) en la forma de la varianza

$\sigma_k^2(B) = \lambda_U' \gamma_U(B) - \gamma(B,B)$, basta calcular el valor de los coeficientes $\lambda_U$ en el modelo.

Para calcular los valores de $\lambda_U$ ajustaremos un kriging universal, (rutina en R) pues recordemos que el Block Kriging es una generalización de este modelo cuando $B \neq \{S_0\}$, pero sólo se realiza un ajuste en la varianza, quitando la constante obtenida anteriormente.

Así, los pesos óptimos serán de la forma:

$$\lambda_U = \Gamma_U^{-1} \gamma_U$$

donde

$$\lambda_U \equiv (\lambda_1, ..., \lambda_n, m_0, .., m_p)'$$

$$\gamma_U \equiv (\gamma(s_0 - s_1), ..., \gamma(s_0 - s_n), 1, f_1(s_0), ..., f_p(s_0))'$$

y $\Gamma_U$ es una matriz simétrica de $(n+p+1) \times (n+p+1)$ de la forma:

$$\Gamma_U \equiv \begin{cases} \gamma(s_i - s_j) & i = 1, ..., n \; j = 1, ...., n \\ f_{j-1-n}(s_i) & i = 1, ..., n \; j = n+1, ...., n+p+1 \\ 0 & i = n+1, ...., n+p+1 \\ & j = n+1, ...., n+p+1 \end{cases}$$

con $f_0(s) \equiv 1$.

Los coeficientes de $\lambda$ están dados por:

$$\lambda' = \left\{\gamma + X(X'\Gamma^{-1}X)^{-1}(x - X'\Gamma^{-1}\gamma)\right\}' \Gamma^{-1}$$

y

$$m = -(x - X'\Gamma^{-1}\gamma)'(X'\Gamma^{-1}\gamma)^{-1}$$

donde

$$\gamma \equiv (\gamma(s_0 - s_1), ..., \gamma(s_0 - s_n))'$$

y $\Gamma_{n \times n}$ una matriz con elemento $\gamma(s_i - s_j)$ en la entrada $(i,j)$.

# Capítulo 5

# RESULTADOS

## 5.1. Predicciones (Kriging espacial)

Las predicciones realizadas por el modelo anterior son las siguientes:
Se toman algunos meses significativos del año, para contemplar el comportamiento del contaminante, además se agrega el intervalo de confianza al 90 % para cada una de las predicciones realizadas, para así ver el peor escenario posible.
Para el mes de diciembre del 2012:

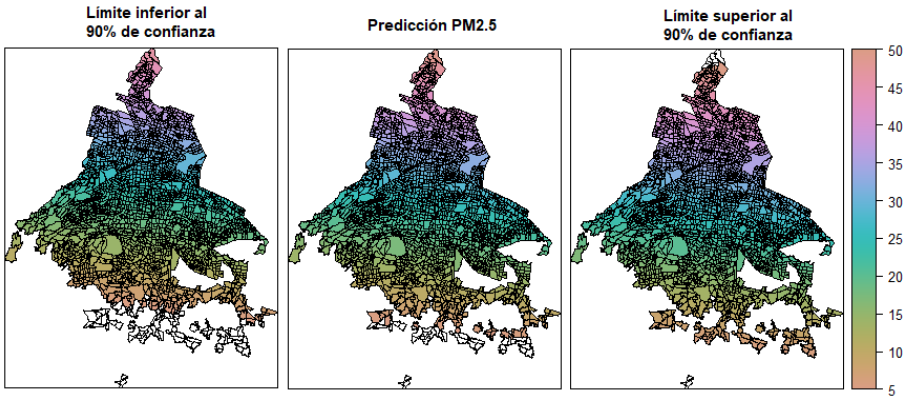


Figura 5.1: Predicciones de PM2.5 para el mes de diciembre del 2012

Se observa una mayor concentración de PM2.5 al norte de la ciudad en el mes de diciembre y va disminuyendo conforme se va al sur. Recordemos que al norte (Estado de México) se encuentra la zona industrial, lo cual tiene sentido pues

generalmente desprenden grandes cantidades de este contaminante.

Siguiendo el procedimiento anterior pero para el mes de marzo del 2013:

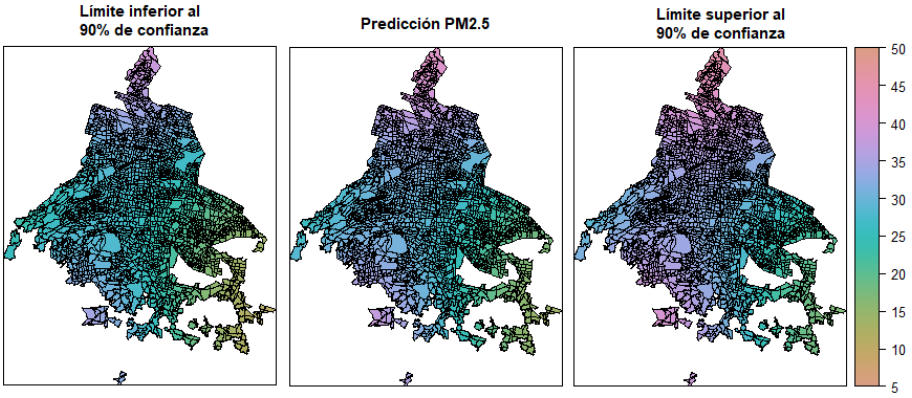


Figura 5.2: Predicciones de PM2.5 para el mes de marzo del 2013

Notamos un comportamiento parecido al mes de diciembre, es decir, se observa una estructura donde es más contaminada la zona norte y va disminuyendo hacia el sur. Hay presencia de valores extremos, o muy altos del contaminante (zona de color rosa).

Para el mes de junio del 2013:

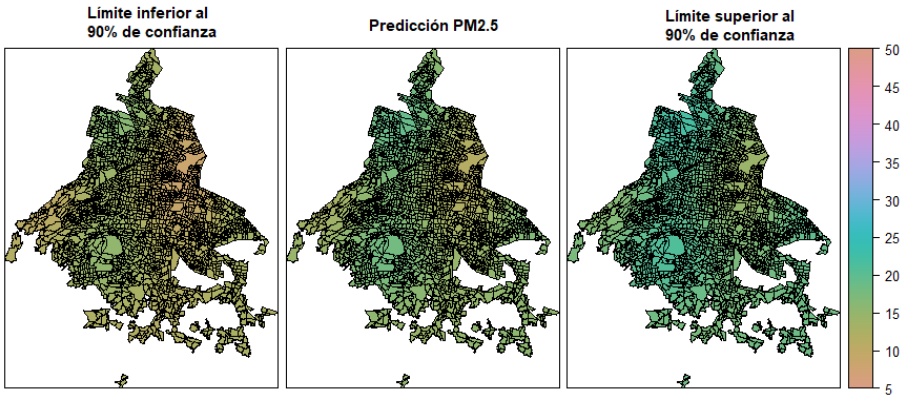


Figura 5.3: Predicciones de PM2.5 para el mes de junio del 2013

En el mes de junio, observamos un comportamiento completamente distinto pues no encontramos valores extremos del contaminante, en general se encuentra por debajo de los 25, parece una estructura uniforme. Podríamos pensar que el contaminante está disperso pues recordemos que junio es época de lluvias y vientos, lo cual nos da un sentido a los resultados.

Para el mes de septiembre del 2013:

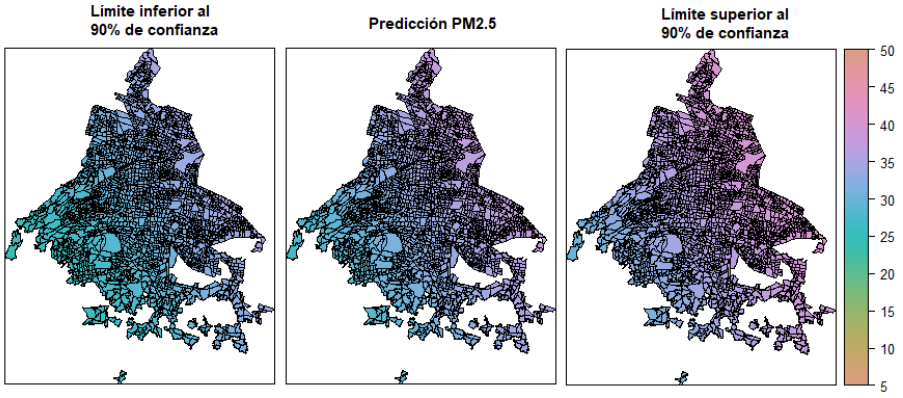


Figura 5.4: Predicciones de PM2.5 para el mes de septiembre del 2013

Para el mes de septiembre notamos de nuevo la presencia de valores extremos del contaminante (colores rosas), afectando a la zona noreste de la Ciudad de México y bajando los niveles conforme vamos al noroeste.

Podemos observar claramente un cambio en los valores del contaminante a lo largo del año, de modo que no tenemos una estructura separable en el espacio y el tiempo. No será lo mismo el comportamiento del contaminante PM2.5 en el mes de diciembre que en el mes de junio, donde hay otros componentes involucrados como la lluvia, el viento, la temperatura. Será necesario ajustar un covariograma espacio-temporal que logre capturar la dependencia en el espacio y en el tiempo de esta variable (PM2.5), buscaremos capturar los demás efectos aleatorios en un modelo más general. Un modelo de predicción espacio temporal.
Los intervalos al 90 % de confianza se calcularon usando la disigualdad de Chebyshev, de la siguiente forma:

$$A \equiv \left(\hat{Z}(B_i) - 3.16\sigma(B_i), \hat{Z}(B_i) + 3.16\sigma(B_i)\right)$$

donde,

$$\sigma(B_i) = \sqrt{\lambda'_U \gamma_U(B_i) - \gamma(B_i, B_i)}$$

## 5.2. Análisis de datos espacio-temporal

Para el ajuste espacio- temporal, consideraremos los mismos años que cuando se ajusta el modelo espacial (2012-2013) para así poder hacer una comparación entre ambos ajustes.
Recordemos que anteriormente se analizaban meses específicos, tomando en cuenta el promedio mensual como indicador del valor tomado por el contaminante en ese mes; ahora se toma la información diaria del valor tomado por el contaminante desde el 01/enero/2012 hasta el 31/dic/2013.
Al tener información diaria, la volatilidad de los datos aumenta considerablemente, por lo cual hubo la necesidad de usar una transformación que lograra el análisis correcto de los datos. Se utiliza el logaritmo de los valores observados, pues de otro modo la volatilidad no permitía el ajuste de ningún modelo, lo cual es razonable pues en datos de esta naturaleza tendremos variables que afecten a los valores del contaminante, por ejemplo: el viento, la temperatura, incluso fenómenos más grandes que a pesar de no afectar directamente, se puede observar un efecto en las partículas suspendidas, pues recordemos son aquellas con un diámetro muy pequeño.

En la figura 5.5 podemos observar cómo se ve el variograma del año 2012 si no hacemos la transformación logarítmica, donde claramente podemos observar que hay una variabilidad muy grande en intervalos de tiempo muy pequeños.

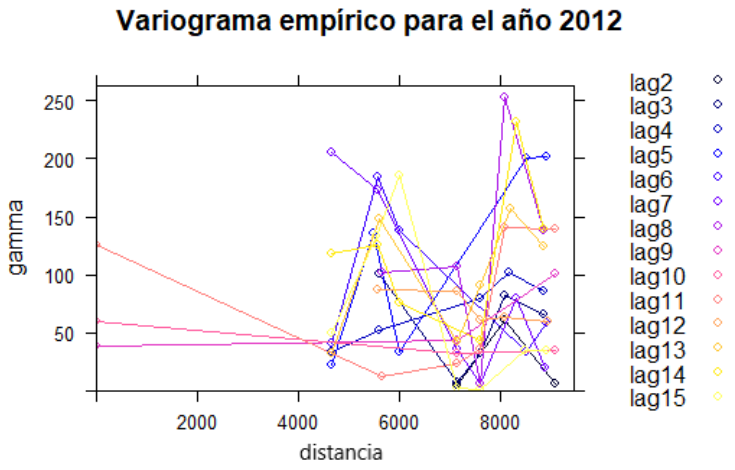


Figura 5.5: Variograma empírico de PM2.5 para el año 2012

A continuación, con el fin de ver la diferencia en las escalas al aplicar la transformación, se muestran los datos del año 2012, en su escala original y la transformación logarítmica de los datos para el año 2013, tomando algunos meses específicos:

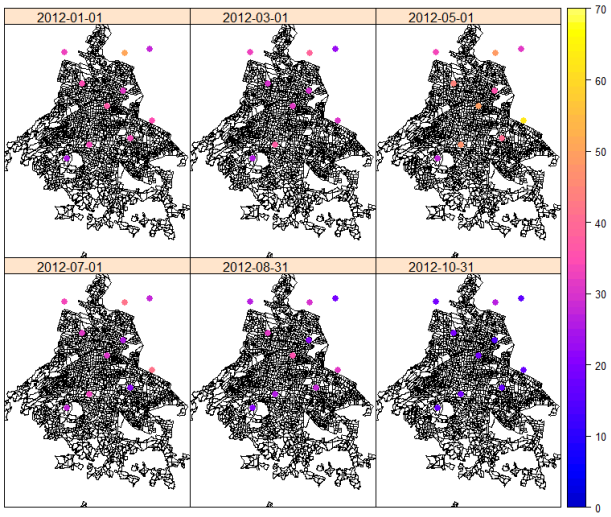


Figura 5.6: Valores de PM2.5 en escala original para el año 2012

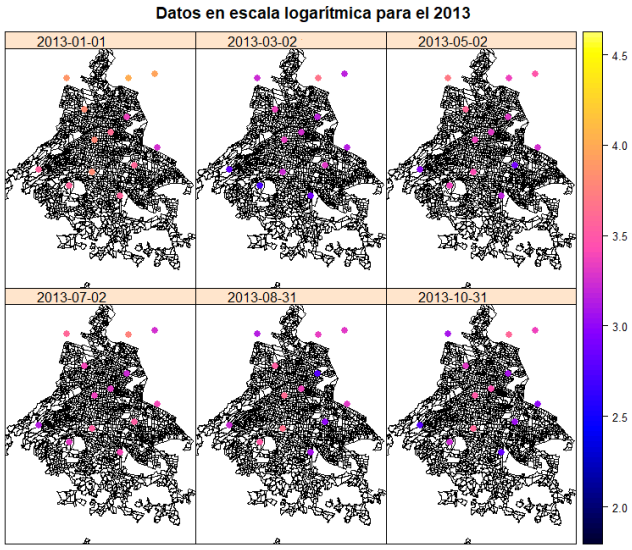


Figura 5.7: Valores de PM2.5 en escala logarítmica para el año 2013

## 5.3. Ajuste variograma espacio-temporal

Los variogramas empíricos para los años 2012 y 2013 son los siguientes, ajustando un variograma por año.

### Año 2012

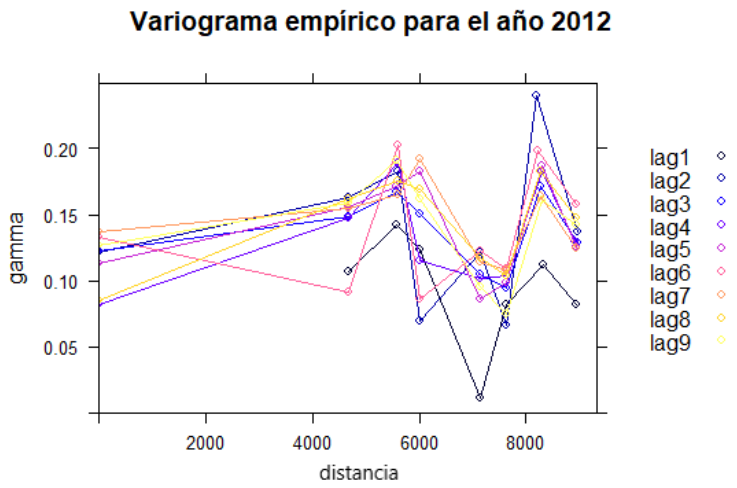


Figura 5.8: Variograma de PM2.5 en escala logarítmica para el año 2012

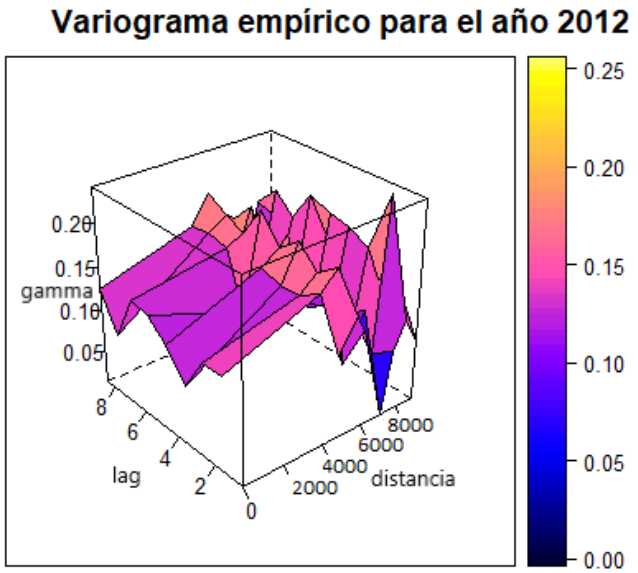


Figura 5.9: Variograma de PM2.5 en escala logarítmica para el año 2012

Se prueban distintos modelos de variogramas espacio-temporales.
Una vez elegido el modelo espacio-temporal (separable, suma producto, métrico, etc), se realizaron combinaciones de los siguientes variogramas tanto en espacio y tiempo:

⊕Modelo exponencial
⊕Modelo esférico
⊕Modelo cuadrático
⊕Modelo matérn

La combinación que mejor modelaba la estructura de covarianza fue cuando se eligió el modelo matérn tanto para la parte espacial como para la parte temporal, claro con los parámetros distintos para cada modelo espacio-temporal.

### Modelo Separable

Se asume que el componente espacio y tiempo son separables, el modelo asumido es de la forma:

$$C_{sep}(h,u) = C_s(h)C_t(u)$$

Este modelo es relativamente parsimonioso y fácil de interpretar, sin embargo existen muchos fenómenos físicos que no satisfacen la separabilidad. Muchos procesos ambientales, por ejemplo, no satisfacen el supuesto de la separabilidad, cómo lo mencionamos anteriormente al ajuste un modelo puramente espacial. Esto significa que este modelo debe usarse con cuidado.
Veamos como ajusta a nuestros datos:

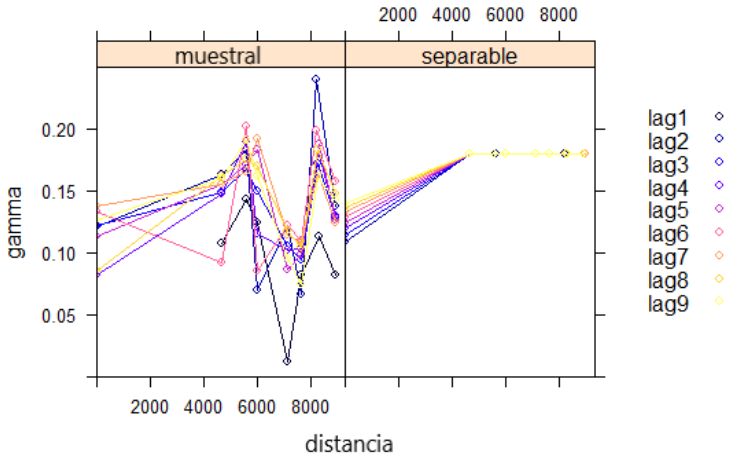


Figura 5.10: Ajuste de variograma separable para el año 2012

Claramente se observa una estructura muy distinta, entre las observaciones y cómo ajusta el modelo de variograma separable. Al calcular el MSE (mean squared error) nos da prácticamente el valor del parámetro sill. Lo cual quiere decir que prácticamente está ajustando al valor inicial dado como parámetro.

### Modelo Suma producto

El modelo asumido es de la forma:

$$C_{sp}(h,u) = kC_s(h)C_t(u) + C_s(h) + C_t(u)$$

donde $k > 0$ es una constante tal que cumple la condición de covarianza positiva-definida.
Veamos cómo ajusta a nuestros datos:

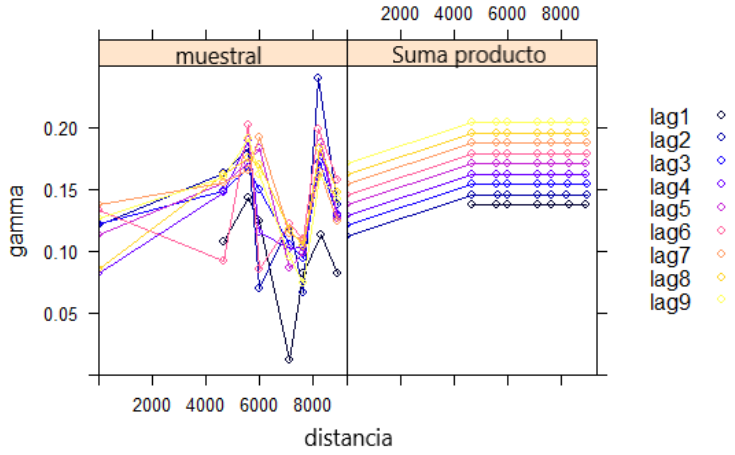


Figura 5.11: Ajuste de variograma suma producto para el año 2012

### Modelo métrico

El modelo asumido es de la forma:

$$C_m(h,u) = C_j\left(\sqrt{h^2 + (\kappa u)^2}\right)$$

Sigue la idea natural de extender el espacio geográfico bidimensional en un espacio espacio-temporal tridimensional. Para lograr un espacio isotrópico, el dominio temporal tiene que cambiarse de escala para que coincida con el espacial (corrección de anisotropía espacio-temporal $\kappa$). Todas las distancias espaciales, temporales y espacio-temporales son tratadas igualmente resultando en un modelo de covarianza conjunta.

La idea es la misma que en el caso espacial, agrupamiento de ubicaciones según su distancia de separación. En el caso espacio-temporal, las distancias son pares de distancia espacial y temporal que dan una superficie de variograma, no una sola línea.

Veamos cómo ajusta a nuestros datos:

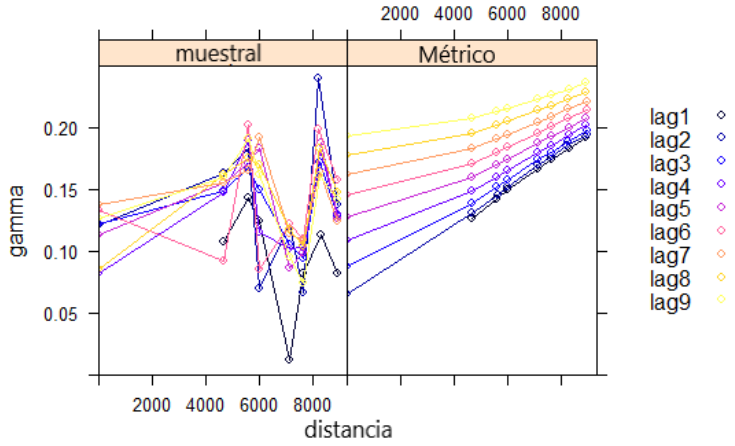


Figura 5.12: Ajuste de variograma métrico para el año 2012

**Modelo suma métrico**

El modelo asumido es de la forma:

$$C_{sm}(h,u) = C_s(h) + C_t(u) + C_j\left(\sqrt{h^2 + (\kappa u)}^2\right)$$

Originalmente, este modelo permite efectos de nugget espaciales, temporales y conjuntos, una versión simplificada puede permitir solo un nugget conjunto. Veamos cómo ajusta a nuestros datos, con el modelo original:

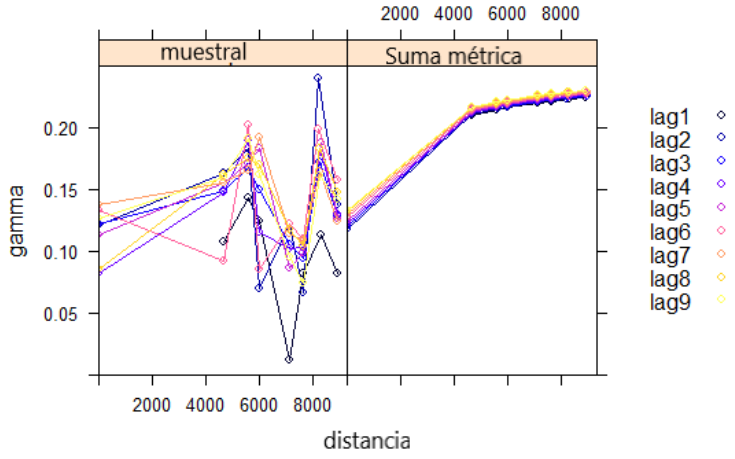


Figura 5.13: Ajuste de variograma suma métrica para el año 2012

Veamos cómo ajusta a nuestros datos, cuando sólo permitimos un nugget conjunto:

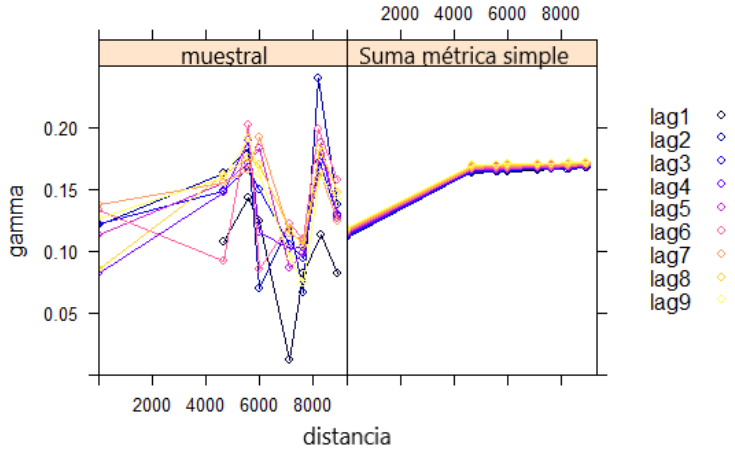


Figura 5.14: Ajuste de variograma suma métrica para el año 2012

**Comparación de los modelos de variogramas ajustados**

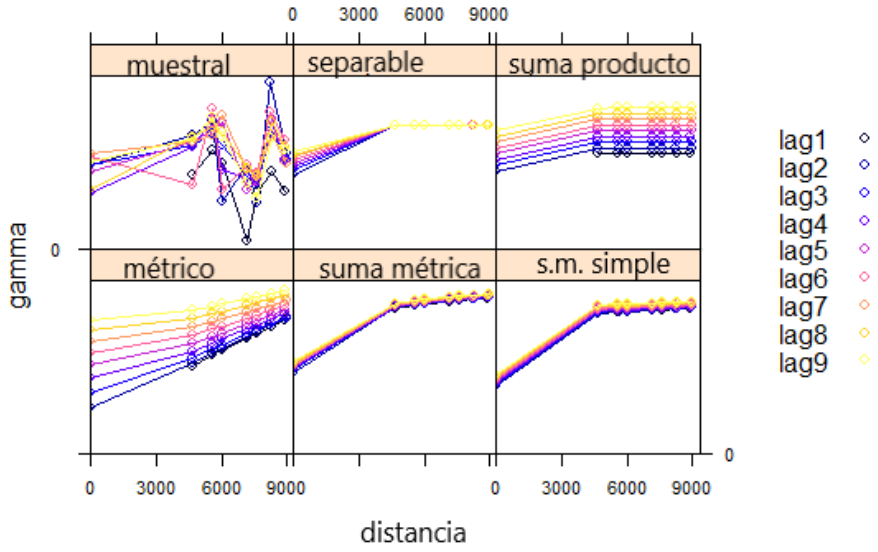


Figura 5.15: Variogramas ajustados para el año 2012

### Año 2013

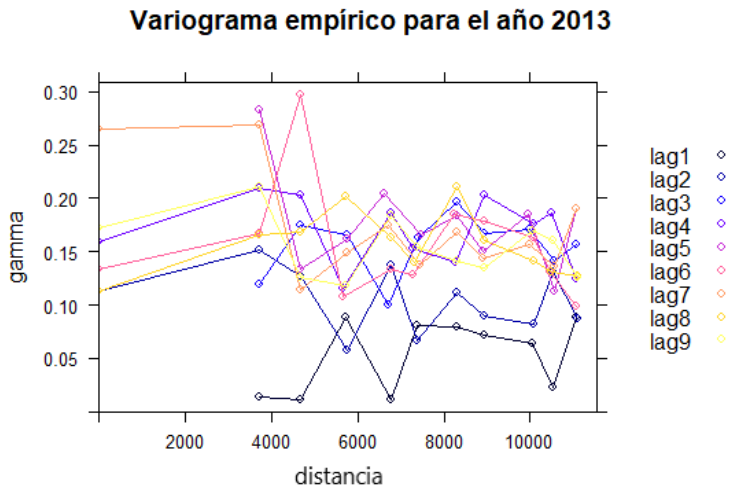


Figura 5.16: Variograma de PM2.5 en escala logarítmica para el año 2013

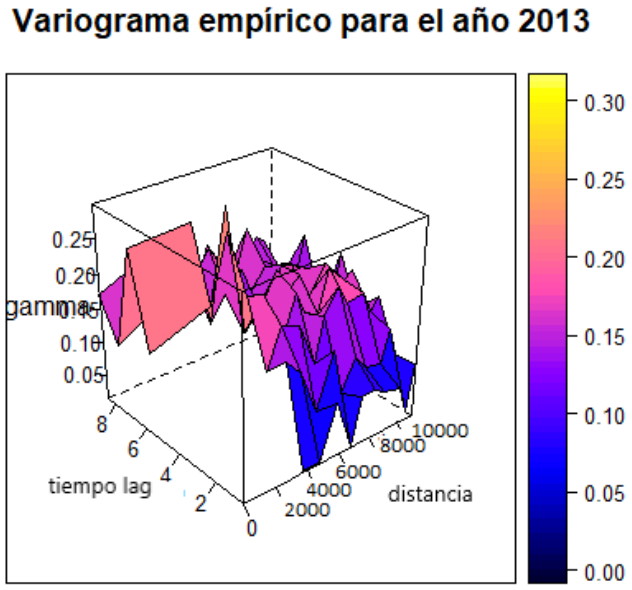


Figura 5.17: Variograma de PM2.5 en escala logarítmica para el año 2013

Se probaron distintos modelos de variogramas espacio-temporales.

**Modelo Separable**

$$C_{sep}(h,u) = C_s(h)C_t(u)$$

Veamos cómo ajusta a nuestros datos:

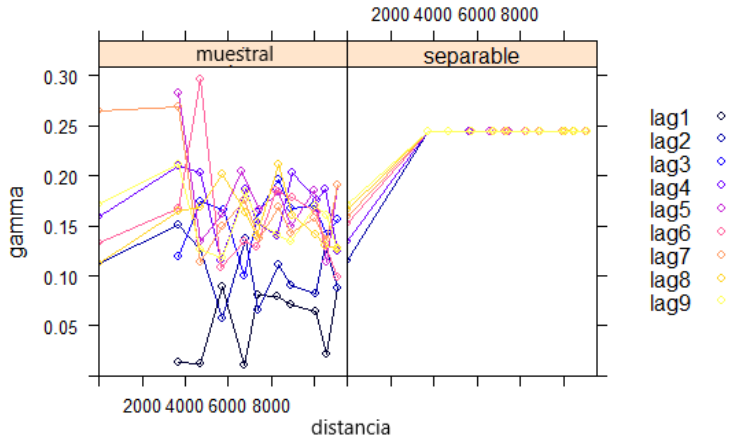


Figura 5.18: Ajuste de variograma separable para el año 2013

Se tiene un ajuste distinto a las observaciones.

**Modelo Suma producto**

$$C_{sp}(h,u) = kC_s(h)C_t(u) + C_s(h) + C_t(u)$$

donde $k > 0$ es una constante tal que cumple la condición de covarianza positiva-definida.
Veamos cómo ajusta a nuestros datos:

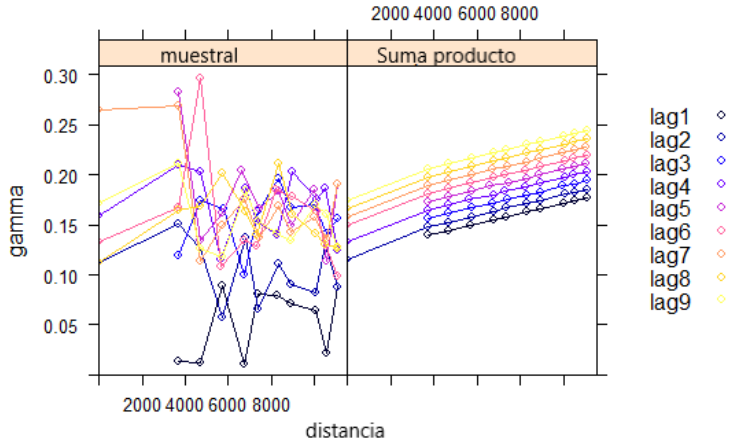


Figura 5.19: Ajuste de variograma suma producto para el año 2013

**Modelo métrico**

$$C_m(h,u) = C_j\left(\sqrt{h^2 + (\kappa u)^2}\right)$$

Veamos cómo ajusta a nuestros datos:

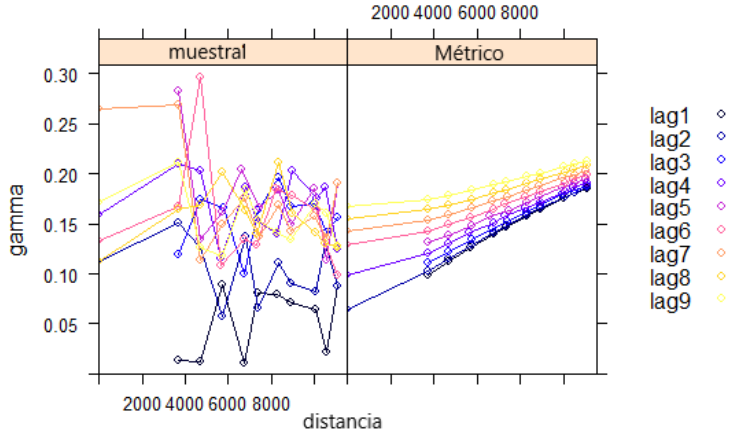


Figura 5.20: Ajuste de variograma métrico para el año 2013

**Modelo suma métrico**

$$C_{sm}(h,u) = C_s(h) + C_t(u) + C_j\left(\sqrt{h^2 + (\kappa u)^2}\right)$$

Veamos cómo ajusta a nuestros datos, con el modelo original:

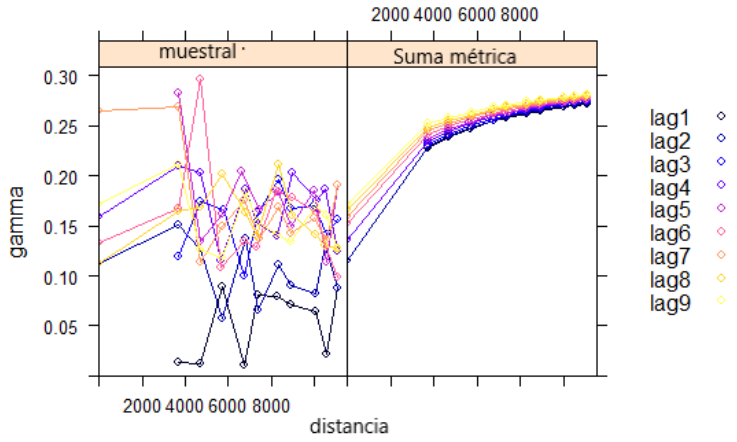


Figura 5.21: Ajuste de variograma suma métrica para el año 2013

Veamos cómo ajusta a nuestros datos, cuando sólo permitimos un nugget conjunto:

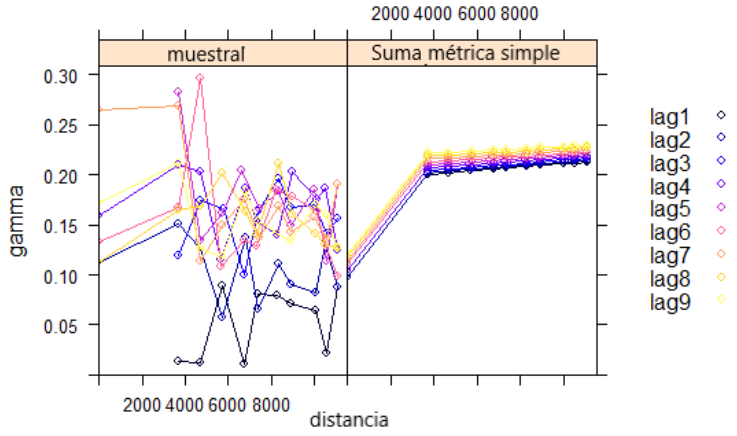


Figura 5.22: Ajuste de variograma suma métrica para el año 2013

### Comparación de los modelos de variogramas ajustados

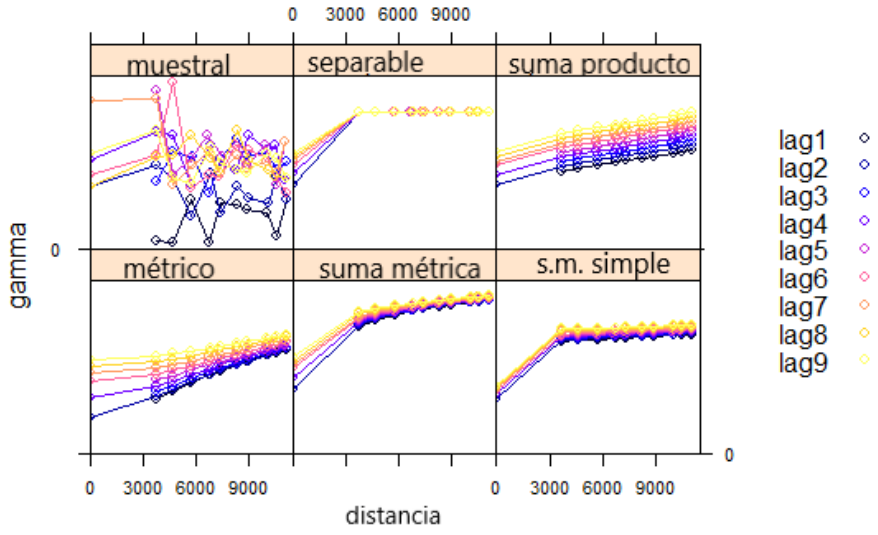


Figura 5.23: Variogramas ajustados para el año 2013

Para ambos años, el modelo que mejor ajusta a los datos es el variograma métrico de la forma $C_m(h,u) = C_j\left(\sqrt{h^2+(\kappa u)^2}\right)$, con $C_j(r,\theta)$ un modelo matérn de la forma:

$$C_j(r;\theta) = \sigma_0^2 I(||r|| = 0) + \sigma_1^2\{2^{\theta_2-1}\Gamma(\theta_2)\}^{-1}\{||r||/\theta_1\}^{\theta_2}K_{\theta_2}(||r||/\theta_1)$$

$\theta = (\sigma_0^2, \sigma_1^2, \theta_1, \theta_2)'$, donde $\sigma_0^2 \geq 0$, $\sigma_0^2 \geq 0$, $\theta_1 > 0$, $\theta_2 > 0$ y $K_{\theta_2}$ es una función de Bessel modificada del segundo tipo de orden $\theta_2$.
Con los siguientes parámetros para el año 2012:

| Umbral (Sill) | Pepita (nugget) | Rango | Anis |
|---|---|---|---|
| .363 | 0.289 | 150 | 1085 |

Con los siguientes parámetros para el año 2013:

| Umbral (Sill) | Pepita (nugget) | Rango | Anis |
|---|---|---|---|
| .290 | .0338 | 150 | 1080 |

Los variogramas fueron ajustados a la transformación logarítmica, por eso los valores tan pequeños de umbral y pepita.

## 5.4. Predicciones espacio-temporales de PM2.5

El modelo usado es el Block kriging, espacio temporal, de la forma:

$$\left\{\begin{array}{c} \sum_{j=1}^{n(B_0,t_0)} \lambda_j \gamma_\epsilon\left((B_i,t_i)-(s_j,t_j)\right) + \sum_{k=1}^{p}\alpha_k f_k(B_i,t_i) = \gamma_\epsilon\left((B_i,t_i)-(B_0,t_0)\right) \\ \forall i = 1,...,n(s_0,t_0) \\ \sum_{i=1}^{n(B_0,t_0)} \lambda_i f_k(B_i,t_i) = f_k(B_0,t_0),\ \forall k = 1,...,p \end{array}\right\}$$

Y la varianza de la predicción es corregida nuevamente de la siguiente forma:

$$\sigma_{sk}^2(B_0,t_0) \equiv \sum_{i=1}^{n(B_0,t_0)} \lambda_i \hat{\gamma}_\epsilon\left((B_i,t_i)-(B_0,t_0)\right) + \sum_{k=1}^{p}\alpha_k f_k(B_i,t_i) - \hat{\gamma}(B_i,B_i,t_i)$$

Tendremos que calcular la varianza por bloque espacio temporal, definida como:

$$\gamma(B,B,t) \equiv \int_B\int_B\int \gamma(u-v,t)dudvdt/|B|^2$$

Para poder predecir $(\gamma(B_i,B_i,t_i))$, el variograma del proceso sobre un bloque $B_i$, cuya localización y geometría es conocido, se hace lo siguiente:

Se simulan $k$ puntos dentro cada ageb, digamos $\{r_1, ..., r_k\}$ de modo que el variograma por bloque, definido como la triple integral, será estimado por:

$$\overline{\gamma}(B_i, B_i, t_i) = \sum_{j=1}^{k} \frac{1}{k} \gamma(s_i, r_j;\ t_i)$$

Es decir, tomamos la fecha a la cual se busca realizar la predicción, digamos $t_i$ y calculamos el promedio de los variogramas $\gamma(s_i, r_j;\ t_i)$, donde $r_j$ son los puntos simulados y $s_j$ es el centro de la ageb $j$. Donde usamos el modelo de variograma espacio temporal ajustado en la sección anterior.

Recordemos que las distancias de los puntos simulados al centro de las agebs ya había sido calculada en la parte del modelo espacial; así, basta tomar una coordenada más que será el tiempo a predecir para calcular el variograma promedio y así estimar la varianza por bloque.

Ahora los bloques son de la forma $(B_i, t_i)$, pues se realizan predicciones diarias así que se toma un solo tiempo, notemos que podríamos tener un bloque de la forma $(B_i, (t_1, t_2, ...t_7))$ en caso que de querer una predicción semanal (7 días) en un área específica $B_i$, o bien $(B_i, (t_1, t_2, ...t_k))$ una predicción que abarque $k$ días.

Se muestran las predicciones realizadas por el modelo block kriging espacio temporal, tomando algúnos días significativos del año:

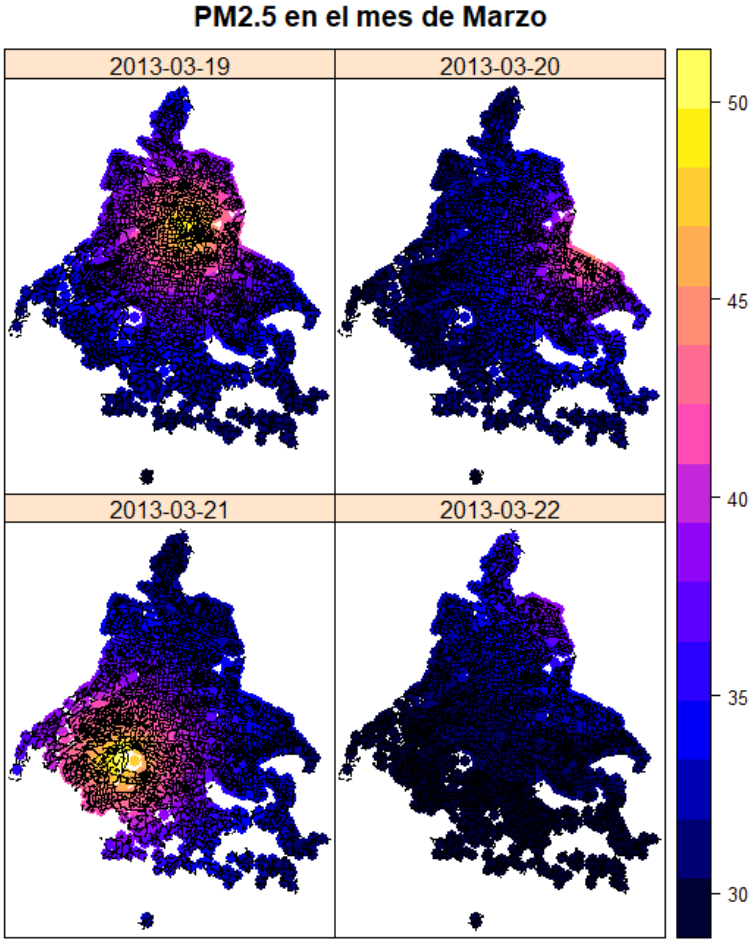


Figura 5.24: Predicciones de PM2.5

En el mes de marzo el contaminante toma valores desde $30\mu g/m^3$ hasta $50\mu g/m^3$, lo cual de acuerdo a la OMS ya es un valor crítico del contaminante. Podemos observar cómo el valor del contaminante fluctuá mucho, incluso de un día a otro, aunque parece siempre tener una estructura parecida, es decir, se observa la concentración de valores altos del contaminante en áreas específicas de la CDMX y conforme nos alejamos de ese punto crítico el valor del contaminante va disminuyendo, esto tiene sentido pues el viento juega un papel fundamental en la concentración y propagación del contaminante.

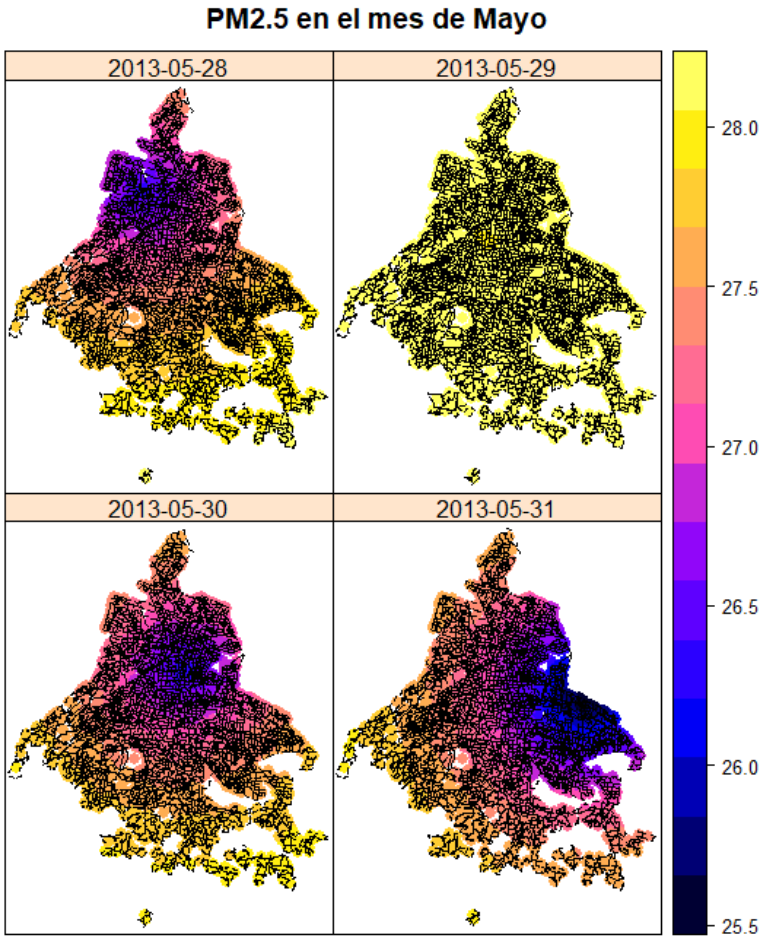


Figura 5.25: Predicciones de PM2.5

En el mes de mayo el comportamiento es completamente distinto, ahora la estructura de concentración es sobre los valores más pequeños que toma el contaminante, es decir, alrededor de 25-26 $\mu g/m^3$ y las partes más contaminadas con valores $28\mu g/m^3$ son aquellas que están en el sur y casi en la frontera de la CDMX, recordemos que esta fecha es próxima a la Semana Santa y podría ser un punto de influencia en cuanto al comportamiento del tránsito en la ciudad. Es importante notar que los valores más altos que tomó el contaminante en este mes no son valores críticos y están dentro de los estándares permitidos.

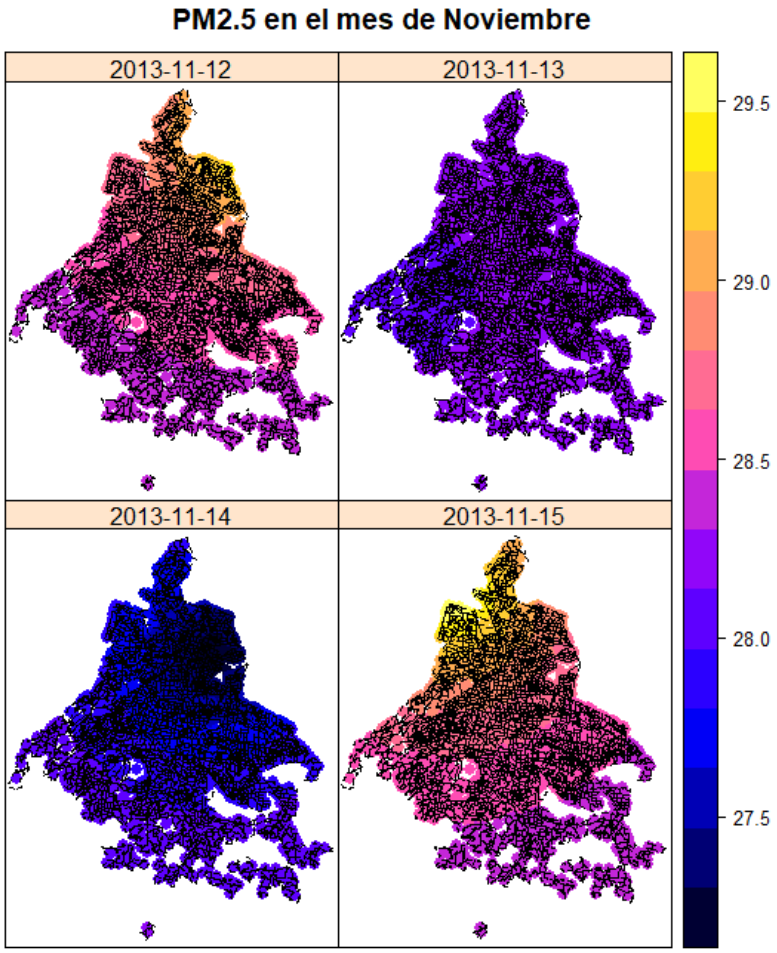


Figura 5.26: Predicciones de PM2.5

Para el mes de noviembre, se observa que el cambio día a día es desde los valores $27\mu g/m^3$ a $29\mu g/m^3$, en la gráfica pareciera que existe un cambio drástico de un día a otro pero pensemos en que la escala es muy pequeña y realmente lo que se observa es un cambio de valores altos a bajos, pero realmente quedando constante en un intervalo de 27 a $29\mu g/m^3$.

Notemos que las gráficas se muestran con escalas distintas, esto debido a que existen meses donde la variación es muy poca y otros en los cuales la variación es demasiada, esto nos prohíbe observar a detalle el comportamiento mensual en el caso de graficar con la misma escala.(Figura 5.27).

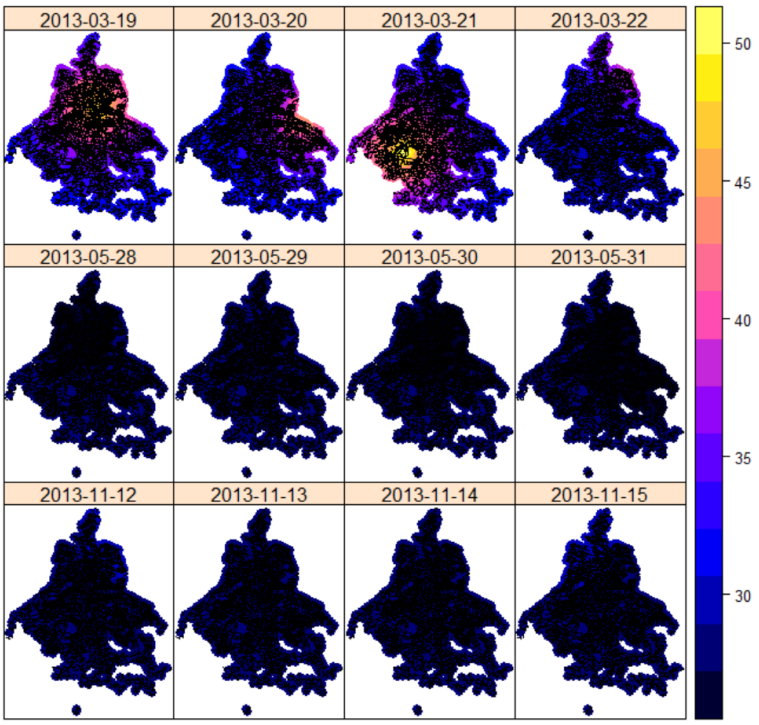


Figura 5.27: Predicciones de PM2.5

### Intervalos de confianza de la predicción

Los intervalos al 90 % de confianza se calculan usando la disigualdad de Chebyshev, de la siguiente forma:

$$A \equiv \left(\hat{Z}(B_0,t_0) - 3.16\sigma_{sk}(B_0,t_0), \hat{Z}(B_0,t_0) + 3.16\sigma_{sk}(B_0,t_0)\right)$$

donde la varianza es tal que,

$$\sigma^2_{sk}(B_0,t_0) \equiv \sum_{i=1}^{n(B_0,t_0)} \lambda_i \hat{\gamma}_\epsilon\left((B_i,t_i) - (B_0,t_0)\right) + \sum_{k=1}^{p} \alpha_k f_k(B_i,t_i) - \hat{\gamma}(B_i,B_i,t_i)$$

A continuación se muestran los intervalos de confianza de las predicciones realizadas por el modelo, al 90 % de confianza.

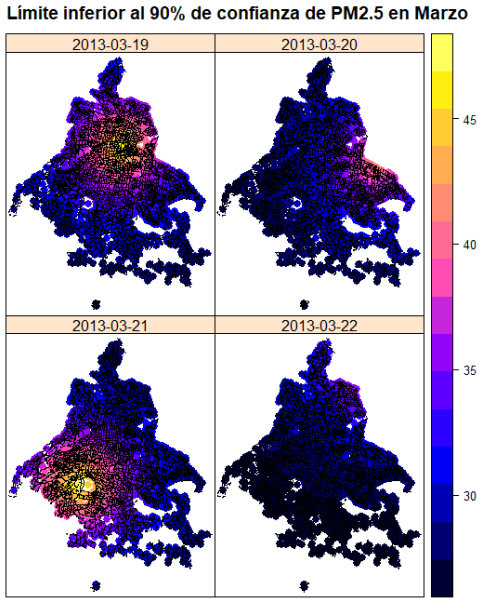


Figura 5.28: Intervalos de confianza PM2.5

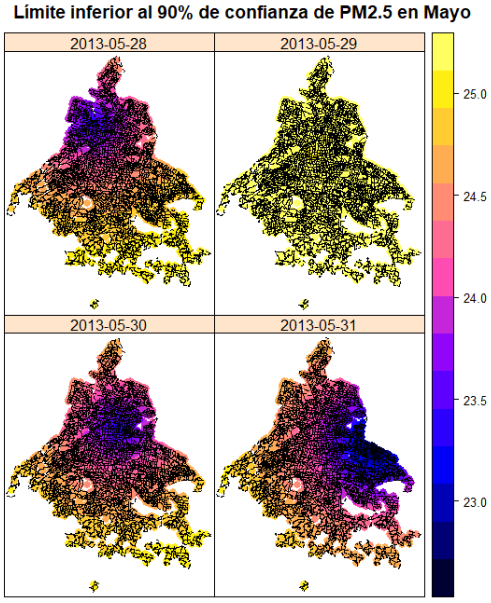


Figura 5.29: Intervalos de confianza PM2.5

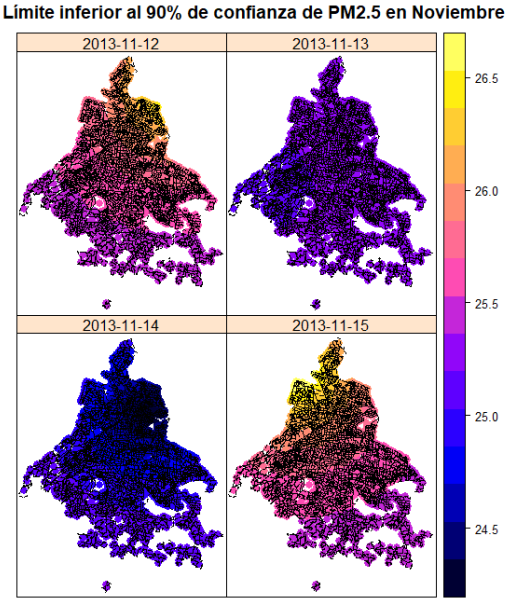


Figura 5.30: Intervalos de confianza PM2.5

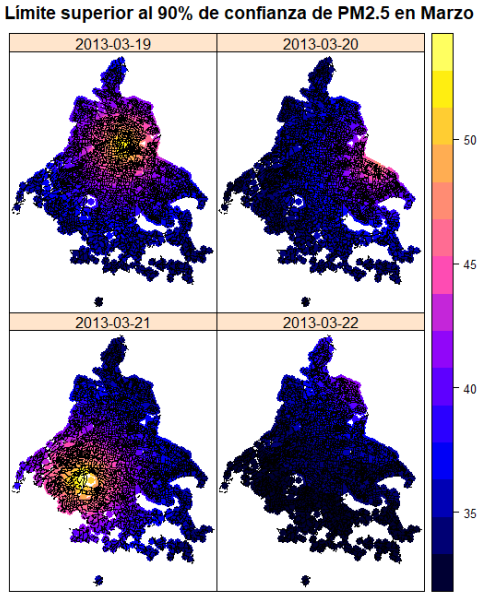


Figura 5.31: Intervalos de confianza PM2.5

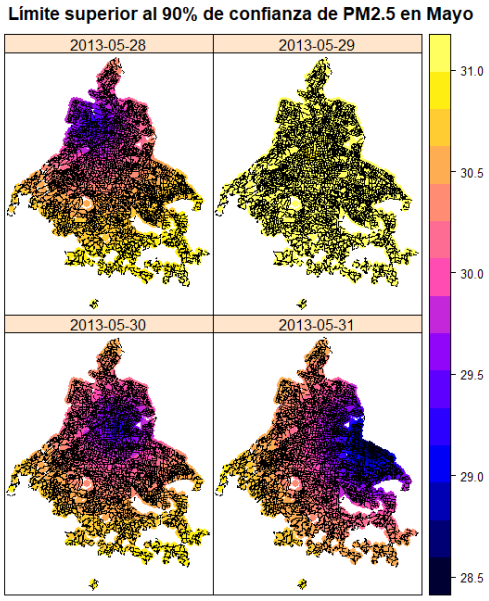


Figura 5.32: Intervalos de confianza PM2.5

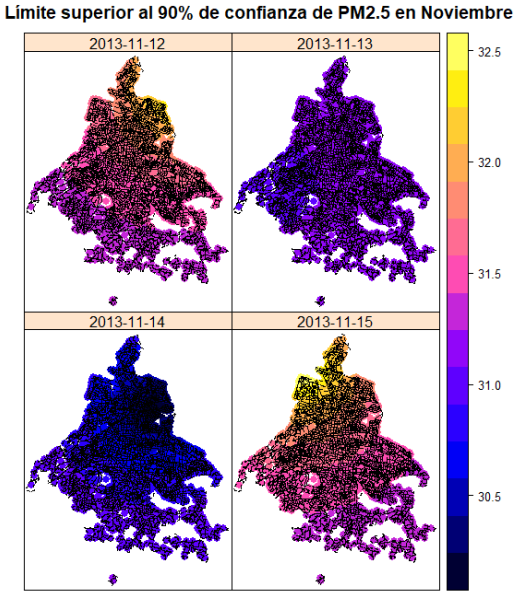


Figura 5.33: Intervalos de confianza PM2.5

# Capítulo 6

# CONCLUSIONES

La manipulación y el análisis de los datos espacio-temporales a menudo se complica por el tamaño y la complejidad de estos datos. Además, los datos pueden venir en muchas formas diferentes, pueden ser ricos en tiempo, ricos en espacio, y además vienen como conjuntos de puntos de espacio-tiempo o como trayectorias. Sobre la base de la infraestructura existente para datos espaciales y temporales, hemos implementado con éxito un conjunto de modelos para la predicción del contaminante PM2.5 en la CDMX, aunque nos encontramos con distintos retos pues la estructura de la base de datos era muy inestable al tener datos faltantes y estaciones de monitoreo cerradas por temporadas.

Las clases y métodos presentados en este documento son un primer intento de cubrir la necesidad del monitoreo constante de los contaminantes, pues como sabemos son parte importante de la salud y otros fenómenos naturales.
Nuestro estudio sobre un conjunto de datos reales de gran dimensión ha demostrado que el paso del análisis de datos puramente espacial a espacial-temporal es altamente no trivial, la exploración visual de datos se vuelve intrincada y la dependencia a lo largo del tiempo puede ser asimétrica y ocurrir mediante varios tipos (función de covarianza separable o no separable, modelo dinámico, ..., etc).
El enorme tamaño de los conjuntos de datos requiere algoritmos implementados eficientemente para almacenar, manipular y explorar datos. Los modelos deben ser capaces de capturar las características específicas de los datos dependientes del tiempo a través de estructuras adecuadas de covarianza espacio-tiempo, mientras permanecen susceptibles de una estimación y predicción eficiente y relativamente rápida a través de kriging. Las funciones de covarianza están ligadas

a la condición de ser no negativa, de modo que la construcción de los modelos admisibles y las funciones de distancia en el plano o en la esfera para datos globales o de gran superficie es complicada.
En el trabajo anterior realizamos el análisis de una sola variable PM2.5, mediante un modelo que ofrece funcionalidad para el análisis y la predicción de datos espacio-temporales, pero sería interesante el desarrollo de un modelo multivariantes espacio-temporal, para así analizar simultáneamente otros contaminantes como PM10, O3, entre otros. Es claro que hay una interacción interesante entre los contaminantes y podrían incluso estarse afectando unos a otros.

# BIBLIOGRAFÍA


1. Cressie Noel, K.Wikle Christopher, *Statistics for Spatio-Temporal Data*, revised edition. John Wiley & Sons, New York, 2010.

2. Cressie Noel, *Statistics for Spatial Data*, revised edition. John Wiley & Sons, New York, 1993.

3. Cressie N, Chan N; *Spatial Modeling of Regional Variables*, Journal of the American Statistical Association, 1989.

4. Haslett J, Raftery AE. *Space-time Modelling with Long-memory Dependence*, Applied Statistics, 1989.

5. Cressie, N. and Johannesson, *G.Fixed rank kriging for very large spatial data sets*, J. of the Royal Statist. Society, Series B, 2008.

6. Fuentes M, *Testing for separability of spatial-temporal covariance functions*, Journal of Statistical Planning and Inference, 2005.

7. Gneiting T, *Nonseparable, stationary covariance functions for space-time data*, Journal of the American Statistical Association,590–600,2002.

8. Gneiting T, Marc G, Peter Guttorp,*Geostatistical Space-Time Models, Stationarity, Separability and Full Symmetry,* 2007.

9. Dongsheng, Z., Mei-Po, K., Wenzhong , Z., & Shaojian, W. (8 de Diciembre de 2017). *Spatiotemporal Variations and Driving Factors of Air Pollution in China.* International Journal of Environmental Research and Public Health(14).

10. Instituto Nacional de Ecología y Cambio Climático & Secretaría de Medio Ambiente y Recursos Naturales. (2013). *Informe Nacional de Calidad del Aire 2013, México.* Coordinación General de Contaminación y Salud Ambiental. Ciudad de Mèxico: INECC, SEMARNAT.

11. OMS. (2005). *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. Ginebra: OMS.

12. OMS. (2006). *Air Quality Guidelines*. Global Update 2005. Alemania: Druckpartner Moser.

13. ONU. (10 de junio de 2014). Naciones Unidas. Recuperado el 18 de diciembre de 2017, de http://www.un.org/es/development/desa/news/population/world-urbanization-prospects-2014.html

14. Prüss-Ustün, A., Wolf, J., Corvalán, C., Bos, R., & Neira, M. (2016). *Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks*. Francia.

15. Secretaría de Medio Ambiente y Recursos Naturales SEMARNAT. (2013). *Calidad del aire: una práctica de vida. Ciudad de Mèxico: SEMARNAT.*

16. Anderson J., Thundiyil J., Stolbach A. *Clearing the air: a review of the effects of particulate matter air pollution on human health*. J Med Toxicol 2012; 8:166-175

17. Zhang, R., Jing, J., Tao, J., Hsu, S.-C., Wang, G., Cao, J., Lee, C. S. L., Zhu, L., Chen, Z., Zhao, Y., and Shen, Z.: *Chemical characterization and source apportionment of PM2.5 in Beijing: seasonal perspective*, Atmos. Chem. Phys., 13, 7053-7074, https://doi.org/10.5194/acp-13-7053-2013, 2013.