

Predicciones geospaciales del contaminante PM2.5 de estaciones distribuidas en la Ciudad de México.

Lara, A, Peralta, J

¹Departamento de Matemáticas, ESFM-IPN, México D.F., México

Teléfono (55) 5729-6000 Ext. 55017 Fax (55) 5729-55015 E-mail: jperaltag1500@esfm.ipn.mx

Resumen — Se obtuvieron mediciones realizadas en estaciones distribuidas en las diferentes alcaldías de la Ciudad de México del contaminante PM2.5 limpiando los datos. Con ayuda del modelo de interpolación de Kriging ordinario se realizaron las predicciones, en el cuál se demostró la isotropía de las mediciones realizadas, las correlaciones de las estaciones respecto a las distancias que las separa a partir del semivariograma, para la aplicación del mismo. Se obtuvieron predicciones del contaminante de radios de 10km desde las estaciones y se obtuvo el mapa de calor del contaminante.

Palabras Clave — PM2.5, Kriging ordinario, isotropía, semivariograma.

Abstract — Measurements of the pollutant PM2.5 were obtained from data bases in stations distributed in different cities of Mexico City, selecting days with the minimum value of null measurements. With the ordinary Kriging interpolation model, the predictions were made as well as the isotropy of the measurements made was demonstrated, the correlations of the stations with respect to the distances that separate them from the semivariogram, in order to construct the model. Pollutant predictions of 10km radii were obtained from the stations and the heat map of the pollutant was obtained.

Keywords — PM2.5, Ordinary Kriging, isotropy, semivariogram.

I. INTRODUCCIÓN

A. Geoestadística.

La Geoestadística es una rama de la estadística encargada en el análisis y modelación de datos referidos a localizaciones específicas mediante coordenadas [1]. A diferencia de la estadística, la geoestadística se centra en el análisis y modelización de variables asociadas a información espacial en base a métodos que cuyo objetivo es estimar valores espaciales desconocidos a partir de otros ya conocidos.

La Geoestadística hereda propiedades de la estadística, por ejemplo podemos definir las siguientes propiedades:

Varianza: Sea s cualquier localización geoespacial, la varianza del campo aleatorio Y se define como:

$$\text{var}[Y(s)] = E[Y(s) - \mu(s)]^2 \quad (1)$$

Covarianza: Sean s_1, s_2 dos localizaciones geoespaciales, la covarianza se define como

$$C(s_1, s_2) = E[(Y(s_1) - \mu(s_1))(Y(s_2) - \mu(s_2))] \quad (2)$$

Variograma: El variograma entre dos ubicaciones s_1, s_2 está definido como:

$$\begin{aligned} 2\gamma(s_1, s_2) &= \text{var}[Y(s_1) - Y(s_2)] \quad (3) \\ &= E[(Y(s_1) - Y(s_2)) \\ &\quad - (\mu(s_1) - \mu(s_2))]^2 \end{aligned}$$

Donde γ es el semivariograma y representa la autocorrelación especial de las mediciones hechas.

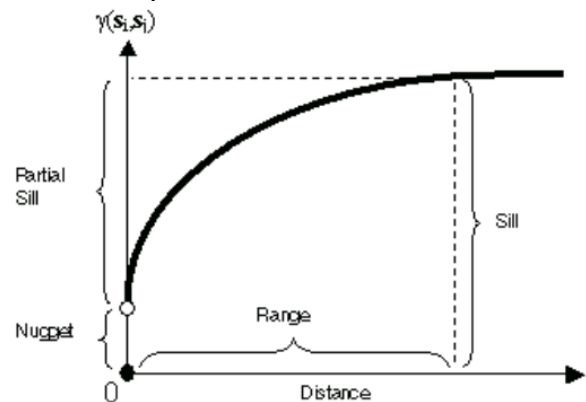


Fig. 1: Ejemplo de Semivariograma, donde el eje x representa la distancia, el eje y la correlación entre la ubicaciones de las mediciones, el valor nugget y la altura del semivariograma [1].

El semivariograma incrementa con la distancia por lo que se podría ver como una función disimilar [2], mientras que la covarianza decrementa con la distancia por lo que se puede ver como una función de similitud.

B. Variogramas.

Variogramas de puntos basados en múltiples o básicos retrasos (lag) y con una tolerancia absoluta de retraso.

En estos tipos variogramas, se manipula tanto la dirección de tolerancia, los ángulos azimutales y se busca que el área de búsqueda así como su separación de distancia entre mediciones incrementa a una distancia máxima del ancho de banda de una dirección fija de vector [3]; de esta forma, la búsqueda del área comienza como un cono circular o elíptico para convertirse en un cono circular o cilíndrico como se

observa en la siguiente figura:

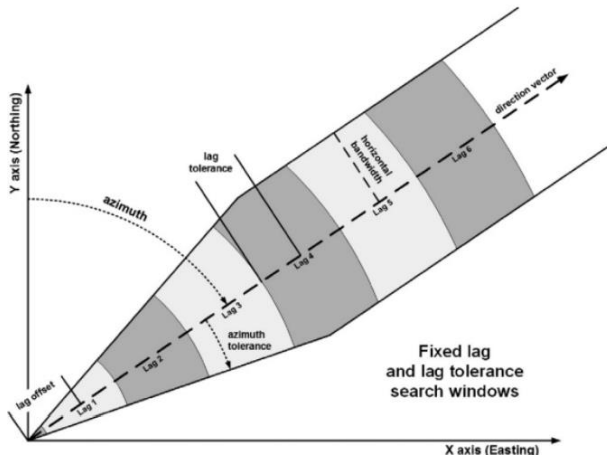


Fig. 2: Ejemplo de variograma con retrasos, donde se muestra el como este predice valores a partir de un valor de ángulo azimutal con un específico valor de retraso de tolerancia[3].

Variogramas de puntos basados en búsqueda de bloques

Este método está basado en bloques ortogonales que forman un modelo 3D centrado en el origen las coordenadas polares variografías. Dicho bloque extiende el control del retraso (lag) y del ángulo azimutal pero de forma diferente al método anterior [8]. Cada uno de los pares es comparado contra cada uno de los otros bloques con una muestra del modelo original; si la segunda muestra de pares falla, significa que el centroide de un bloque particular está cerca. Al final, cuando todos los pares se les asigna a bloques particulares, los valores del variograma bloque son calculados y guardados en el modelo, así como se observa en la siguiente figura:

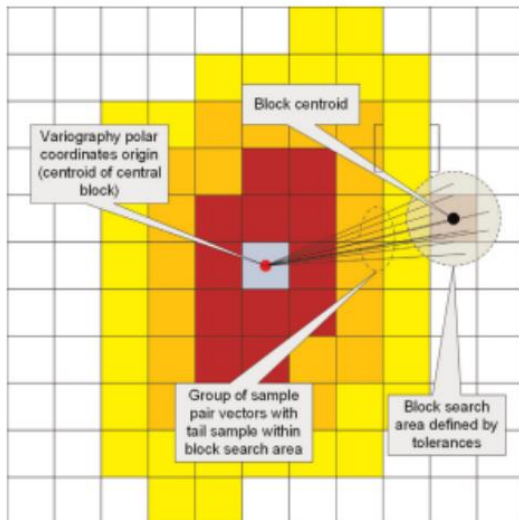


Fig. 3: Ejemplo de variograma de bloques, donde la búsqueda de bloques define las tolerancias así como el agrupamiento de la muestra respecto al centroide de los datos[4].

Variograma con retraso variable basado en el agrupamiento de k-means

Los métodos antes mencionados requieren que por prueba y error uno pueda conseguir la configuración del retraso óptimo del variograma. Sin embargo, la información para alcanzar una configuración aceptable del retraso está disponible en la muestra de pares para cualesquier dirección, aún para muestras de patrones irregulares.

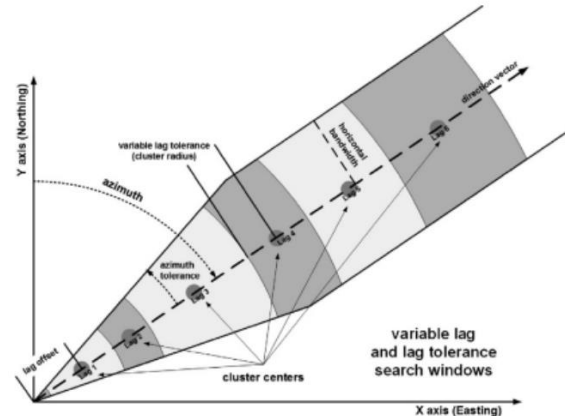


Fig. 4: Ejemplo de variograma con clusters en un ángulo azimutal fijo y con un retraso variable[3].

Con este método los grupos de pares pueden ser identificados por un procedimiento automatizado de aprendizaje no supervisado.

C. Clusterización por K-means.

El agrupamiento por k-means es un método originalmente utilizado en procesamiento de señales, utilizado para el minado de datos. Está técnica agrupa n observaciones en k grupos, cada observación se le asigna un agrupamiento con la media más cercana [11]. Es un algoritmo iterativo en el que antes de cualquier iteración, los grupos son centrados inicialmente en un número igual de observaciones, es asignado a un grupo en particular cuya media esté entre al menos el grupo de la suma de cuadrados. El segundo paso corresponde al cálculo de las nuevas medias a los centroides de las observaciones de los nuevos agrupamientos, para finalmente converger cuando no existe cambio en las tareas.

Algorímicamente se tiene lo siguiente:

1. Se inicializa el agrupamiento de la selección centrada. Se selecciona el agrupamiento inicial con un NC (número de agrupamientos requeridos) los cuales los primeros valores son seleccionados como agrupamientos centrales y los casos sobrantes son reprocesados como sigue:
 - i) Si $\min_i d(x_k, M_i) > 0$ y $d(x_k, M_m) > d(x_k, M_n)$ entonces x_k sustituye M_n . Si $\min_i d(x_k, M_i) > d_{mn}$ y $d(x_k, M_n) < d(x_k, M_m)$, el x_k sustituye M_m ; esto es, si la distancia entre x_k y la media del agrupamiento cercano es mayor que la distancia entre los agrupamientos cercanos (M_n, M_m), entonces x_k

- sustituye ya sea a M_m o a M_n , cualesquiera que esté cerca de x_k .
- ii) Si x_k no sustituye a la media de agrupamiento en i), un segundo intento se realiza:
1. Sea M_q el agrupamiento promedio más cercano a x_k .
 2. Sea M_p el segundo agrupamiento promedio cercano a x_k .
 3. Si $d(x_k, M_p) > \min_i d(M_q, M_i)$ entonces $M_q = x_k$

Donde NC es el número de agrupamientos solicitados, M_i la media de los i-ésimos agrupamientos, x_k el vector con la k-ésima observación, $d(x_i, x_j)$ es la distancia euclidiana de los vectores x_i, x_j , d_{mn} es la distancia entre las dos medias cercanas $\min_{i,j} d(M_i, M_j)$.

2. Inicializar la actualización de los agrupamientos centrales. Comenzando con el primer caso, cada caso es asignado al agrupamiento más cercano, y los agrupamientos promedios son actualizados. El agrupamiento inicial central es incluido en este promedio. La actualización del agrupamiento promedio son considerados como una clasificación a los agrupamientos centrales.
3. Inicializar la actualización de los agrupamientos centrales. Comenzando con el primer caso, cada caso es asignado al agrupamiento más cercano, y los agrupamientos promedios son actualizados. El agrupamiento inicial central es incluido en este promedio. La actualización del agrupamiento promedio son considerados como una clasificación a los agrupamientos centrales.

D. Estacionariedad.

La estacionariedad nos dice el como las campos aleatorias Y se comportan en una región geográfica.

Estacionariedad estricta

Una campo aleatoria Y se dice estrictamente estacionaria si para cualesquier vector h las distribuciones de $Y(s_1), \dots, Y(s_n), Y(s_1 + h), \dots, Y(s_n + h)$ son idénticos para un arbitrario n . Se dice que la campo aleatoria es invariante ante traslación. Puesto que esta definición conlleva a que los momentos de cualquier orden no deben de depender de la ubicación, y usualmente en aplicaciones geoestadísticas basta con los primeros momentos para el análisis y modelaje.

Estacionariedad de segundo orden

Se dice que una campo aleatoria es estacionaria de segundo orden si: (a) La esperanza existe y es independiente de la ubicación (b) La covarianza existe y depende solamente del vector h que separa a dos ubicaciones:

$$\mu(s) = E[Y(s)] = \mu \quad (4)$$

$$C(s + h, s) = C(s + h - s) = C(h) \quad (5)$$

Donde $C(h)$ se le llama covariograma (análoga a la función de autovarianza en series de tiempo). Cuando $C(0) >$

0, el correlograma $\rho(h)$, la correlación entre dos puntos separados por un vector h está definido como:

$$\rho(h) = C(h)/C(0) \quad (6)$$

Si un campo aleatorio es estacionario de segundo orden entonces el semivariograma se puede escribir como sigue:

$$\gamma(h) = C(0) - C(h) \quad (7)$$

Estacionariedad intrínseca

Se dice que una campo aleatoria tiene una estacionariedad intrínseca si (a) Su esperanza existe y es independiente de la ubicación. (b) Para cualesquier dos ubicaciones separadas por un vector h , la varianza de la diferencia $[Y(s) - Y(s + h)]$ existe y es función de h :

$$\text{var}[Y(s) - Y(s + h)] = 2\gamma(h) \quad (8)$$

E. Anisotropía e Isotropía.

La covarianza $C(h)$ es una función de h , si la covarianza exhibe diferentes comportamientos en diferentes direcciones se dice que el campo es anisotrópico. Por otro lado cuando $C(h)$ depende solamente de su longitud h , denotada por $|h|$, el campo se dice isotrópico.

En algunos casos la función $C(h)$, cuando es anisotrópica, puede ser representada por una función con covarianza isotrópica $f(|h_1|)$ transformandola linealmente el vector h a h_1 , esta tipo de anisotropía se le llama anisotropía geométrica.

F. Modelos de correlación isotrópicos.

De la definición del semivariogramas isotrópicos $\gamma(h) = C(0) - C(h)$, supondríamos que $C(0) = 0$ sin embargo, en ciertos casos puede haber una discontinuidad llamado el efecto nugget que refleja la variabilidad de su campo aleatorio. Cuando el semivariograma en el límite de la distancia h converge a un valor límite, a este valor se le llama "sill" o umbral. Existen diferentes tipos de modelos de correlación, entre los usuales se observan los siguientes:

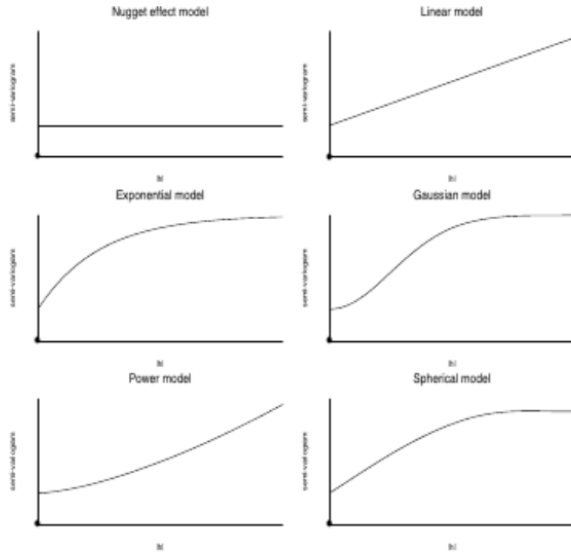


Fig. 5: Ejemplos de modelos de correlación isotrópicos, de izquierda a derecha y de arriba abajo se observan con efecto nugget, modelo lineal, modelo exponencial, modelo gaussiano, modelo potencia y modelo esférico[8].

Para cada uno de estas gráficas se definen $\gamma(h)$.

G. Modelos de correlación para procesos no estacionarios.

Cuando la suposición de covarianza homogénea no es suficiente para los campos aleatorios entonces se hace uso de distintas técnicas para obtener modelos de correlación entre ellos se encuentran: The Sampson-Guttorp The Higdon, Swall, and Kern The Fuentes [6].

H. Predicciones espaciales

Kriging

Este método de predicción supone que los datos están homogéneamente distribuidos espacialmente. Este método consiste en que dadas ubicaciones s_1, \dots, s_n con valor $Y(s_i), \forall i$. El objetivo es estimar la concentración $Y(s_0)$ en la ubicación s_0 . El interpolador de Kriging, una combinación lineal ponderada de las variables observadas $\{Y(s_i)\}$ tiene coeficientes que la hacen no sesgado y con mínimos errores de predicción, conocido como varianza de kriging [4]; de esta manera la aplicación de este método consiste en encontrar los pesos ponderados óptimos para la predicción de las mediciones en las ubicaciones dadas, estos pesos son covarianzas del campo aleatorio de las ubicaciones s_1, \dots, s_n , cuando las estructuras de covarianza son isotrópicas, los pesos están en función de las ubicaciones $\left(\alpha \frac{1}{r}\right)$.

Kriging ordinario

Supongamos un campo aleatorio $Y(\cdot)$ es intrínsecamente estacionario, i.e para cualesquier ubicación s :

$$E[Y(s)] = \mu \quad (9)$$

$$\text{Var}[Y(s) - Y(s+h)] = 2\gamma(|h|) \quad (10)$$

La interpolación local (kriging) para predecir el valor en una ubicación $Y(s_0)$ es un promedio de los pesos de los

niveles de las áreas en los alrededores. Es un estimador lineal óptimo de la forma:

$$Y^*(s_0) = \sum_{i=1}^n \alpha_i Y(s_i) \quad (11)$$

Donde los pesos α_i son escogidos para hacer al estimador no sesgado y con un error mínimo de predicción, para lograr el promedio de este estimador de kriging [9], debe de ser idéntico al promedio del campo aleatorio en la ubicación s_0 . i.e

$$E[Y(s_0)] = E\left[\sum_{i=1}^n \alpha_i Y(s_i)\right] = \sum_{i=1}^n \alpha_i \mu \quad (12)$$

Así la estimación es no sesgada $E[Y^*(s_0)] = E[Y(s_0)]$ si los pesos suman la unidad $\sum_{i=1}^n \alpha_i = 1$.

Por otro lado el error de predicción se puede escribir en términos del semivariograma:

$$\begin{aligned} \sigma_{s_0}^2 &= E[Y^*(s_0) - Y(s_0)]^2 \\ &= E\left[\sum \alpha_i (Y(s_i) - Y(s_0))\right]^2 \\ &= E\left[\sum_i \sum_j \alpha_i \alpha_j (Y(s_i) - Y(s_0))(Y(s_j) - Y(s_0))\right] \\ &= \sum_i \sum_j \alpha_i \alpha_j \gamma(|h_{ij}|) - \sum_i \alpha_i \gamma(|h_{i0}|) \end{aligned} \quad (13)$$

De esta manera la última expresión es lo que hay que minimizar en función de los pesos α_i . Ahora los pesos kriging se pueden minimizar a través de la siguiente función objetivo [10]:

$$f(\alpha_1, \dots, \alpha_n, \lambda) = \sigma_{s_0}^2 + 2\lambda \left(\sum_i \alpha_i - 1\right) \quad (14)$$

Al realizar la derivada parcial respecto a los multiplicadores de Lagrange λ se encuentra que:

$$\sum_j \alpha_j \gamma(|h_{ij}|) + \lambda = \gamma(|h_{i0}|), i = 1, \dots, n \quad (15)$$

$$\sum_j \alpha_j = 1 \quad (16)$$

Así dado el semivariograma, obtenemos el mejor predictor lineal de kriging.

II. METODOLOGÍA

A. Recolección de datos

Los datos utilizados fueron mediciones meteorológicas realizadas en estaciones distribuidas en la CDMX, estas estaciones tienen un rango de confiabilidad de 3km equidistante. Todas estas mediciones fueron extraídos de datos previamente tratados y estructurados de la base de datos de SAPIENS del IPN, que a su vez los toma de una base de datos publicada por la secretaría del medio ambiente [7]

(SEDEMA: <https://www.sedema.cdmx.gob.mx/>) con datos desde enero de 1986. Se recolectaron datos desde principios del 2016 hasta finales del 2017 parámetros de partículas PM2.5, partículas de 2.5 micrómetros mezcla que puede incluir sustancias químicas orgánicas, polvo, hollín y metales que se encuentran en el medio ambiente, entre todos los datos recolectados el 66% resultaron datos nulos entre las 21 estaciones distribuidas entre las diferentes alcaldías de la CDMX, fue por esta razón que entre los diferentes días en los que se obtuvieron datos, se escogieron aquellos en los que hubo un 5% o menos de mediciones nulas en promedio de todas las estaciones; manteniendo de esta manera fijos los días y enfocandonos en el análisis de los días que cumplen dicho criterio.

TABLA I.
DIAS SELECCIONADOS CON EL CRITERIO

2016						
May	Jun	Jul	Ago	Sep	Nov	Dic
17,18 31	3-12 21-30	1-9 11-13 15,16 20-31	2-24	2-4 23	11,12 22,23 25-30	3-5 12-28 30,31
2017						
Ene	Feb	Jun	Jul	Aug		
1-4 6-10 12,13 23-27 29,30	24-28	6,30	1-4	11-14 19-22		
		16- 25				

B. Tratamiento de los datos.

A los datos de los días seleccionados de la tabla I, fueron filtrados entre los horarios de 6AM y 10PM (horas laborales), a cada uno de los parámetros de PM2.5 fueron normalizados pues su distribución resultó estar sesgada hacia la derecha, como se observa en la figura 6:

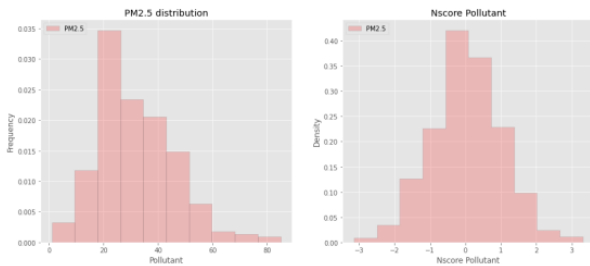


Fig. 6: Histograma del PM2.5 y normalización de los datos, la figura de la izquierda muestra el histograma de datos no normalizados y la figura derecha la de los normalizados.

Ahora, para calcular la distancia entre las estaciones se escribió una función de la fórmula de haversine que considera

las latitudes y longitudes de dos estaciones para regresar el valor de la distancia entre ellas.

$$d = 2r \arcsin \left(\sqrt{\sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos \phi_1 \cos \phi_2 \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right) \quad (17)$$

Donde r es el radio de la Tierra y ϕ, λ corresponde a la latitud y longitud expresadas en radianes de las coordenadas geográficas. Con la ecuación (17) se construyeron las locaciones geográficas equivalentes en un plano cartesiano procediendo a transformar las latitudes y longitudes en km, acomodando en el origen a la estación ubicada en el suroeste más lejano posible.

Para demostrar la isotropía de las mediciones, se procedió a construir el semivariograma definido en (3), aplicando el algoritmo de k-means para la detección del agrupamiento de las estaciones respecto a las mediciones realizadas por las estaciones de PM2.5.

Lo anterior, se aplico para cada uno de los días de la tabla I, para que de manera promediada se pudiera aplicar la interpolación por Kriging ordinario clusterizado por el algoritmo de k-means; y así predecir de manera geoespacial valores de contaminación intermedios entre las estaciones distribuidas en la Ciudad de México.

Todos los resultados fueron programados y obtenidos por el lenguaje de programación Python, junto con la librería GSLIB, librería basado en el software geoestadístico de dominio público creado por Clayton V. Deutsch y Andrei G. Journel por la Universidad de Alberta [5].

III. RESULTADOS

A. Isotropía de las mediciones.

En la construcción del semivariograma para la demostración de la isotropía de los datos, primeramente se realizó un barrido del ángulo azimutal en la que se distribuyen los datos, seleccionando el ángulo azimutal óptimo con el cual poder construir la interpolación de kriging ordinario. Los resultados obtenidos fueron los siguientes:

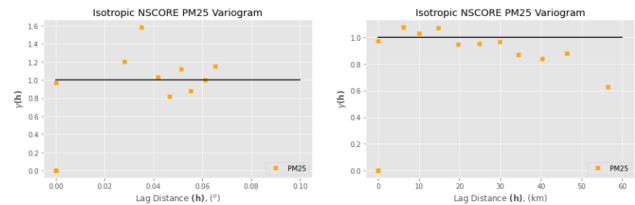


Fig. 7: Semivariogramas para $\phi = 90^\circ$, figura izquierda de las estaciones separadas por latitud y longitud, figura derecha de las estaciones separadas por su equivalente en kilómetros.

Se observa como al tratar los datos desde su longitud y latitud geográfica, el semivariograma con un ángulo azimutal de noventa grados (óptimo), se observa de manera aleatoria

la dependencia de la correlación entre las estaciones respecto a su distancia de separación. Sin embargo, al transformar las coordenadas en su equivalente a su separación en kilómetros se observa la similitud con el modelo de semivariograma esférico.

Puesto que al variar el ángulo azimutal en la que se distribuyen los datos no se observó cambio en el gráfico, se concluye que los datos son isotrópicos por lo que podemos aplicar el método de kriging o alguno de sus submodelos.

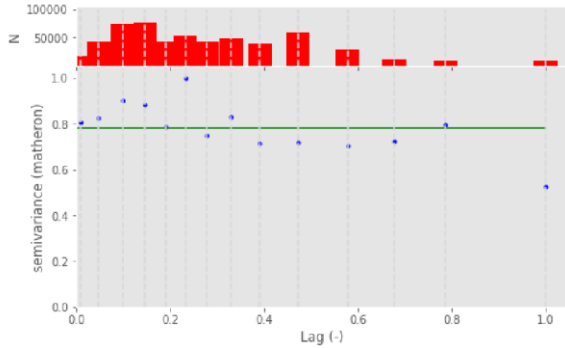


Fig. 8: Semivariograma con algoritmo de K-means, donde el eje x representa los retrasos configurados y el eje y el semivariograma o correlación entre las estaciones con mediciones respecto a su distancia.

B. Interpolación con Kriging ordinario.

Los valores de contaminación promediados de los días seleccionados de la tabla I por cada una de la estación se observan en la siguiente gráfica con su distribución geoespacial de la CDMX.

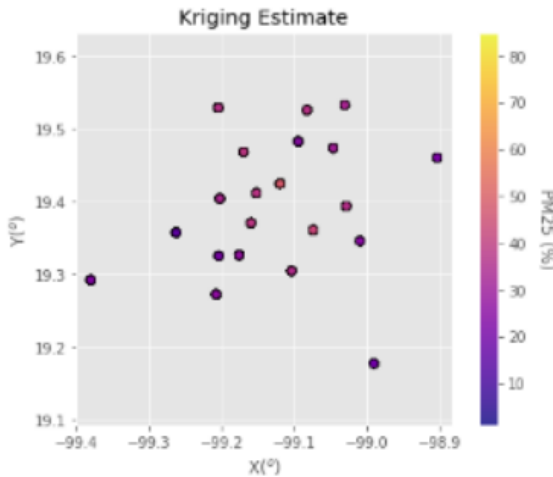


Fig. 9: Distribución de las estaciones de la CDMX, donde los ejes coordenados se encuentran en grados de latitud y longitud.

A partir de los resultados anteriores, se aplicó la interpolación con Kriging ordinario para la predicción de los valores geostatísticos de los valores promediados del PM2.5 en los días seleccionados de la tabla I con ayuda del algoritmo de agrupamiento de K-means.

Se obtuvo como valores óptimos:

Valores de silla: $3.44169127 \times 10^{-15}$

Rango: 0.4172698843221701

Nugget: 0.9432990955184472

(18)

Y con los valores de predicción que ofrece este modelo junto con sus incertidumbres se obtuvo la siguiente interpolación geoespacial de las estaciones de la Ciudad de México:

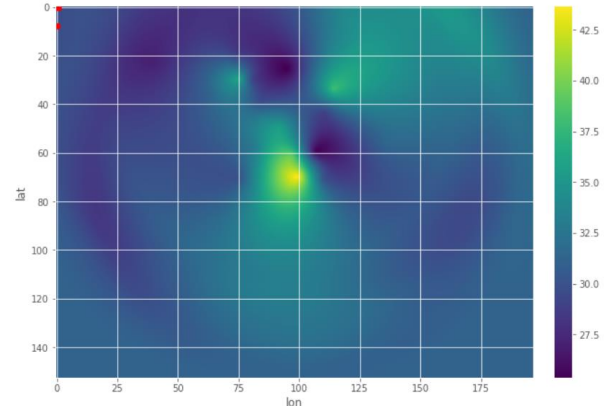


Figura 10: Predicción de valores de PM2.5 en la Ciudad de México, donde los ejes muestran la distancia de separación de las estaciones en kilómetros y la barra el grado de concentración de contaminación.

IV. DISCUSIÓN

El método de interpolación por Kriging ordinario nos ofrece una manera de predicción espacial cuando solo conocemos valores en ciertas coordenadas geográficas constantes. Este modelo tiene como hipótesis el uso de mediciones isotrópicas, para esto se necesita normalizar los valores de contaminación de PM2.5 que se tienen para la construcción del semivariograma y si el ángulo azimutal no varía la tendencia de los valores de autocorrelación se demuestra que las mediciones utilizadas son isotrópicas.

De la figura 8, observamos que la distancia es inversamente proporcional a la correlación de las estaciones, i.e. entre más lejos estén dos estaciones de medir un valor de contaminante de P M 2.5 menor será la correlación de este valor (Como se podría intuir con naturalidad). Y como la tendencia de correlación entre las estaciones es invariante ante cambios del ángulo azimutal, se justifica el uso de la interpolación de kriging ordinario, en el se empiezan a observar resultados en las predicciones de la distribución de estaciones de la figura 9 respecto al de la figura 10.

En la figura 10, se observa como la mayor concentración de contaminación se encuentra cerca de las estaciones con más altos valores de contaminación. Es de esta misma figura en la que se obtuvieron predicciones, en la que se pueden extraer valores de contaminación en las cercanías de las estaciones, pues si bien estas estaciones por la documentación de la SEDEMA, solo tienen un rango de confiabilidad de

3km, con estas predicciones tenemos la posibilidad de predecir más allá de esta restricción gracias a la correlación espacial que hay entre las estaciones.

V. CONCLUSIONES

El modelo de Kriging ordinario utilizado para la predicción de los valores de los contaminantes de PM_{2.5} a partir de las mediciones realizadas por las estaciones de SEDEMA de la Ciudad de México nos ofrece una mirada cuantitativa a la intuición que uno tendría en el comportamiento de los contaminantes, los resultados mostraron de manera concluyente como en las estaciones con mediciones altas de contaminación tienen un radio de más de 10 km en los que el contaminante está presente en la ciudad, estos resultados no hubieran sido capaz de realizarse con las limitaciones que las estaciones de SEDEMA tenían.

El semivariograma construido a partir del algoritmo de clusterización por k-means nos brindó la demostración de la hipótesis de que los datos se pueden trabajar como isotrópicos, es decir, invariantes ante cambios de dirección de los contaminantes, es por esta misma hipótesis por la que se tuvo la oportunidad de aplicar el modelo de Kriging ordinario. Sin embargo, este modelo tiene como hipótesis que las mediciones trabajadas fuesen independientes a valores externos; al realizar el estudio de los datos, se observó que las mediciones del contaminante PM_{2.5} en particular tiene correlaciones con O₂ y la temperatura. Es por esto que se tiene a futuro aplicar el modelo de CoKriging el cual es un método de interpolación que considera la dependencia de los parámetros, en particular sería de O₂ y de temperatura.

REFERENCIAS

- [1] Burrough, P. A. Principles of Geographical Information Systems for Land Resources Assessment. New York: Oxford University Press. 1986.
- [2] Heine, G. W. "A Controlled Study of Some Two-Dimensional Interpolation Methods". COGS Computer Contributions 3 (no. 2): 60–72. 1986.
- [3] McBratney, A. B. y R. Webster. "Choosing Functions for Semi-variograms of Soil Properties and Fitting Them to Sampling Estimates". Journal of Soil Science 37: 617–639. 1986.
- [4] Oliver, M. A. "Kriging: A Method of Interpolation for Geographical Information Systems". International Journal of Geographical Information Systems 4: 313–332. 1990.
- [5] Press, W. H., S. A. Teukolsky, W. T. Vetterling y B. P. Flannery. Numerical Recipes in C: The Art of Scientific Computing. New York: Cambridge University Press. 1988.
- [6] Royle, A. G., F. L. Clausen y P. Frederiksen. "Practical Universal Kriging and Automatic Contouring". Geoprocesamiento 1: 377–394. 1981.
- [7] Ambiente, S., 2022. Secretaría del Medio Ambiente. [online] Secretaría del Medio Ambiente. Available at: <<https://www.sedema.cdmx.gob.mx/>> [Accessed 24 June 2022].
- [8] Mark D. Fisk, Gary D. McCartor; Treatment of Correlation for Clustered Reference Data in Kriging Regional Seismic Discriminants. Bulletin of the Seismological Society of America 2008;; 98 (4): 1768–1780. doi: <https://doi.org/10.1785/0120070219>
- [9] Lloyd, Stuart P. "Least squares quantization in PCM." Information Theory, IEEE Transactions on 28.2 (1982): 129-137.
- [10] MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, pp. 281–297). California: University of California Press.
- [11] Steinley, D., & Brusco, M. J. (2007). Initializing k-means batch clustering: A critical evaluation of several techniques. Journal of Classification, 24(1), 99–121.