

Estimation of the Factors influencing the Generation of the PM2.5 Contaminant in Temuco and Padre Las Casas, Chile

María Abarca, BSc¹, Carola Blazquez, PhD¹ y Karen Sauer-Brand, PhD¹

¹Universidad Andres Bello, Chile, m.abarcaabarca1@uandresbello.edu, cblazquez@unab.cl, k.sauerbrand@uandresbello.edu

Abstract- In the present study, we investigate different factors related to demography, land use, transportation and height that influence the generation of the PM2.5 contaminant due to woodburning in the conurbation of Temuco and Padre Las Casas, Chile. First, a global regression model was implemented to explore the relationships between the variables, and subsequently, spatial regression models were implemented (geographically weighted regression and spatial error model). Results suggest that height and infrastructure density have a strong association with the PM2.5 contaminant using the global regression and spatial error models. However, results indicate that residential density, distance to high vehicle congestion, and length of local roads within a vicinity of 100 m are significant when employing the geographically weighted regression. In addition, the latter type of regression model presents the best performance when compared to the other two models (global regression and spatial error). These results may help authorities in the allocation of resources to improve air quality in the studied conurbation. Future research should include other types of spatial regression models and other relevant variables that may influence the generation of PM2.5 due to woodburning.

Keywords- Spatial econometry, regression models, environmental contamination, Chile

Estimación de los Factores que Influyen en la Generación del Contaminante PM2.5 en Temuco y Padre Las Casas, Chile

María Abarca, BSc¹, Carola Blazquez, PhD¹ y Karen Sauer-Brand, PhD¹

¹Universidad Andres Bello, Chile, m.abarcaabarca1@uandresbello.edu, cblazquez@unab.cl, k.sauerbrand@uandresbello.edu

Resumen— *En el presente estudio, se investigan los diferentes factores demográficos y de uso de suelo, transporte y altura que influyen en la generación del contaminante PM2.5 producto de la quema de leña en la conurbación de Temuco y Padre Las Casas, Chile. Primero, se implementó un modelo de regresión global para explorar las relaciones entre las variables y posteriormente se utilizaron modelos de regresión espaciales (regresión ponderada geográficamente y modelo de error espacial). Los resultados sugieren que la altura y la densidad de la infraestructura posee una fuerte asociación con el contaminante PM2.5 usando los modelos de regresión global y de error espacial. En cambio, los resultados indican que la densidad residencial, la distancia a la congestión vehicular alta y el largo de calles locales en una vecindad de 100 m son significativos al utilizar la regresión ponderada geográficamente. Además, este último tipo de modelo de regresión presenta el mejor desempeño al compararlo con los otros dos modelos (regresión global y error espacial). Estos resultados pueden apoyar a las autoridades en la asignación de recursos para mejorar la calidad del medio ambiente en la conurbación estudiada. Futura investigación debiese incluir otros tipos de modelos de regresión espacial y otras variables relevantes que puedan incidir en la generación del PM2.5 por quema de leña.*

Palabras claves— *Econometría espacial, modelos de regresión, contaminación ambiental, Chile*

I. INTRODUCCIÓN

La contaminación ambiental es un problema que genera un impacto a nivel global ya que son muchos los países que cuentan con altos niveles de contaminación en el aire. Según un estudio realizado por científicos en Australia y China, revela que alrededor de la 99,82% de la superficie terrestre se encuentra expuesta a niveles de material particulado (PM) fuera de los límites recomendados por la Organización Mundial de la Salud, OMS [1]. El PM es emitido a la atmósfera debido a la quema de combustibles fósiles, la actividad industrial, el transporte, la agricultura, entre otros y puede ocasionar efectos adversos para la salud humana y el bienestar del medio ambiente en general. Según la OMS, en el año 2019, se estimó que alrededor de 4.2 millones de personas fallecieron a causa de la contaminación ambiental [2] El PM2.5 son partículas microscópicas con un diámetro de menos de 2.5 μm que al ser tan pequeñas pueden ingresar al torrente sanguíneo después de ser inhalado a través de los pulmones [3]. Los daños que causa el PM2.5 no solo depende de su concentración en el ambiente,

sino que también del tiempo que la persona esté expuesta. Mientras mayor sea el tiempo de exposición, mayor es el riesgo de producir un daño a la salud y al desarrollo de enfermedades [4].

Entre las ciudades más contaminadas del mundo en el año 2022, se encuentra Lahore (Pakistán), Hotan (China), Bhiwadi (India), Delhi (India) y Peshawar (Pakistán) [5]. Aunque no son las únicas ciudades que cuentan con el problema de la contaminación ambiental, países más desarrollados, como EEUU y algunos países europeos también sufren este problema de contaminación. Los gobiernos de los diferentes países del mundo han tomado medidas para abordar la contaminación del aire, entre ellas, se encuentra la reducción de las emisiones de los vehículos y de las industrias, la promoción de la utilización de tecnologías más limpias y amigables con el medioambiente, la educación de las personas para que tomen conciencia de esta problemática que afecta a todos [6]. Sin embargo, aún queda bastante por hacer para asegurar el bienestar de las personas y del medio ambiente.

América Latina con una población de 626 millones habitantes es la región que cuenta con mayor urbanización del mundo ya que cerca del 80% de la población vive en ciudades de 200.000 habitantes [7]. Por lo tanto, se genera una mayor producción de residuos y más emisiones de contaminantes en esta región lo que empeora la calidad del aire que se respira [8]. En Chile, la contaminación causada por el PM2.5 es un problema sumamente importante, especialmente en invierno en las ciudades del centro y sur de Chile debido a la quema de leña para calefacción. Lamentablemente, los niveles de contaminación han ido en aumento en el país en los últimos años. Por ejemplo, las ciudades en las regiones de Aysén, Los Lagos y La Araucanía en el sur de Chile presentaron los niveles más altos de PM2.5 entre los años 2017 y 2020, superando el valor de la norma (50 μm) con creces [9].

Temuco y Padre Las Casas es una conurbación en la región de La Araucanía que poseen un serio problema de contaminación ambiental generada por emisiones del PM2.5 debido a la quema de leña usada como combustible en cocinas y calefactores domésticos [10]. En este estudio, se pretende identificar los factores demográficos, de uso de suelo, de transporte, de meteorología y de altura que inciden en la contaminación por PM2.5 en Temuco y Padre Las Casas usando el modelo de regresión global de mínimos cuadrado ordinario y los modelos de regresión espaciales (regresión ponderada geográficamente y error espacial).

Digital Object Identifier: (only for full papers, inserted by LACCEI).

ISSN, ISBN: (to be inserted by LACCEI).

DO NOT REMOVE

II. REVISIÓN BIBLIOGRÁFICA

Existen distintos estudios que han realizado análisis de regresión espaciales para determinar los factores que inciden en la generación del contaminante PM_{2.5} a nivel nacional e internacional. En la Tabla I, se muestra una lista de los estudios más relevantes que se encuentra en la literatura en la ciudad de Temuco y en Chile [3, 11-16]. Cabe destacar que no encontré ningún estudio similar que se haya efectuado en Chile ni en Temuco y Padre las Casas. La Tabla II presenta una breve descripción de los estudios que se han llevado a cabo en diversos países del orbe [17-26].

TABLA I
ESTUDIOS SOBRE CONTAMINANTE PM_{2.5} EN TEMUCO Y CHILE

Autores y año	Lugar	Objetivo del estudio
Quinteros et al., 2023 [3]	Temuco	Crear modelos LUR espaciotemporales en Temuco para partículas finas y trazadores de leña levoglucosan y potasio soluble en agua (Ksol) evaluando su funcionamiento y aplicación y estimar exposiciones en una cohorte de 77 mujeres embarazadas utilizando posteriormente dicha información en estudios de salud.
Jorquera et al., 2018 [10]	Temuco	Caracterizar la contaminación del aire interior para evaluar la exposición a PM _{2.5} en Temuco.
Quinteros et al., 2023 [11]	Temuco	Investigar la distribución espacial y temporal de PM _{2.5} en Temuco como los trazadores de quema de biomasa de potasio soluble en agua (Ksol) y levoglucosan.
Blanco et al., 2022 [12]	Temuco	Caracterizar la distribución del material particulado (PM _{2.5}) a través de mediciones móviles durante las noches de invierno en Temuco.
Villalobos et al., 2017 [13]	Temuco	Identificar y estimar las fuentes de carbón orgánico (CO) ambiental y PM _{2.5} en Temuco mediante el uso de marcadores moleculares característicos de diferentes fuentes de emisión (vehículos diésel y gasolina, quema de madera, combustión de petróleo y carbón, residuos vegetales).
Jorquera, 2021 [14]	Chile	Evaluar la efectividad de las regulaciones ambientales de PM _{2.5} en Chile a través de la estimación de las tendencias meteorológicas ajustadas de PM _{2.5} ambiental.
Varela et al., 2018 [15]	Chile	Probar los patrones de diversidad de líquenes climáticos que podrían explicarse por los niveles de PM en varias áreas urbanas y determinar cuál es la mejor métrica basada en la biodiversidad para usar en estudios posteriores en áreas urbanas de América del Sur.
Yáñez, 2017 [16]	Chile	Evaluar el efecto de los datos meteorológicos sobre el material particulado en las principales ciudades del valle centro-sur de Chile durante la temporada invernal entre 2014 y 2016.

III. DATOS

A. Área de Estudio

Temuco pertenece a la provincia de Cautín y es la capital de la región de La Araucanía con una superficie de 464 km² y una población de 282.415 habitantes. Padre Las Casas está ubicada en la misma provincia y esta conurbada con Temuco y

cuenta con una superficie de 400,7 km² y con 76.216 habitantes [27]. La zona de estudio presenta temperaturas promedio de 12°C, veranos calurosos entre los meses de diciembre y marzo e inviernos fríos entre mayo y agosto. Además, existen precipitaciones de aproximadamente 1.000 mm durante todo el año. Temuco y Padre Las Casas corresponde a una de las zonas más pobres del país y una actividad económica centrada principalmente en la agricultura [28].

TABLA II
ESTUDIOS SOBRE CONTAMINANTE PM_{2.5} EN EL MUNDO

Autores y año	Lugar	Objetivo del estudio
Yang et al., 2021 [17]	China	Identificar de manera precisa las concentraciones de PM _{2.5} en Zibo, Shandong.
Quan et al., 2023 [18]	China	Establecer un modelo de regresión ponderada geográfica y temporal (GTWR) y estimar la concentración de PM _{2.5} entre 2015 y 2020 utilizando datos espaciales e imágenes satelitales.
Shao et al., 2020 [19]	China	Evaluar el nivel de uso intensivo de la tierra en el área de la provincia de Hebei y regresiones espacial para analizar la relación entre la concentración de PM _{2.5} y el uso intensivo de la tierra y la influencia de las variables del uso intensivo de la tierra sobre las PM _{2.5} .
Liu et al., 2023 [20]	China	Predicir la alta resolución de concentraciones de NO ₂ y PM _{2.5} en el suelo en el río Yangtze y revelar los mecanismos de influencia entre NO ₂ y PM _{2.5} junto con precursores y factores meteorológicos.
Berrocal et al., 2020 [21]	Estados Unidos	Evaluar diferentes enfoques de modelamiento de exposición del PM _{2.5} usando máquinas de aprendizaje.
Chang et al., 2020 [22]	Taiwan	Extraer las características espacio-temporales de las concentraciones de PM _{2.5} .
Chen et al., 2018 [23]	China	Desarrollar un método de estimación que combina regresiones lineales y entropía bayesiana.
Li et al., 2018 [24]	Hong Kong	Medir las concentraciones de partículas finas (PM _{2.5}) en las carreteras en una zona urbana densamente poblada de Hong Kong entre el invierno y verano del 2015 utilizando una plataforma de seguimiento móvil de tranvías.
Hankey & Marshall, 2015 [25]	Estados Unidos	Medir las concentraciones de partículas contaminantes del aire en horas peak de la mañana y tarde a través de una plataforma de monitoreo móvil basada en bicicletas.
Jun & Gu, 2023 [26]	Corea	Identificar las relaciones causales entre episodios de alto PM _{2.5} en Corea.

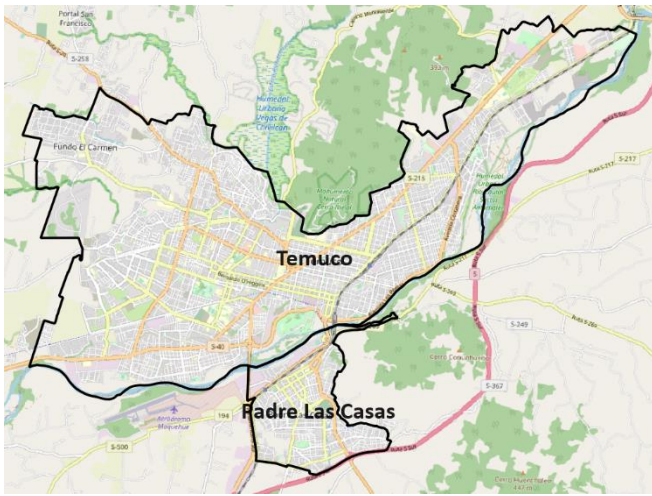


Fig. 1 Temuco y Padre Las Casas.

La Fig. 1 presenta la conurbación estudiada: Temuco y Padre Las Casas. Esta zona urbana presenta episodios de contaminación ambiental particularmente en la época invernal debido a las pobres condiciones de dispersión que facilitan la acumulación de contaminantes en el ambiente. La contaminación ambiental proviene principalmente de la quema de leña residencial junto a otras fuentes en menor medida que son emitidas de la industria y el tráfico vehicular [28].

B. Preparación y agregación de los datos

Entre los meses de mayo y julio del 2016, se realizó una campaña de monitoreo móvil del contaminante PM_{2.5} que representa una época de invierno con bajas temperaturas y precipitaciones. Durante esta campaña móvil, se realizaron viajes en un vehículo, divididos en cuatro rutas distintas entre las 20 y 2 h con velocidades entre 40 y 60 km/h. Como resultado de esta campaña, se obtuvo más de 300.000 mediciones con equipos de medición del PM_{2.5} cada un segundo. Debido a la gran cantidad de mediciones y la redundancia de los datos, se decidió agregar las mediciones del PM_{2.5} a nivel de segmentos de 50 m de la red vial y en celdas de una grilla de 100 m. Cabe mencionar que se asignó el promedio de las mediciones PM_{2.5} que se ubican a lo largo de cada segmento y dentro de cada celda. Para más detalle de esta campaña de monitoreo, revisar [12].

La Tabla III muestra los diferentes niveles de concentración PM_{2.5} y estados para cinco rangos de concentración de PM_{2.5}, según la normativa chilena de la calidad de aire. La simbología presentada en dicha tabla es usada en las Fig. 2a) y 2b) con las mediciones de PM_{2.5} agregadas a nivel de segmentos de 50 m (n = 5.523) y grilla de 100 m (n = 2.318), respectivamente. En ambas figuras, se aprecia una variabilidad espacial en la concentración del contaminante PM_{2.5}, sugiriendo valores promedio más bajo de este contaminante con estados bueno y regular hacia el centro de Temuco y aumentando hacia el este y el oeste de esta ciudad llegando a niveles sobre 170 $\mu\text{g}/\text{m}^3$ que corresponden a estado de emergencia.

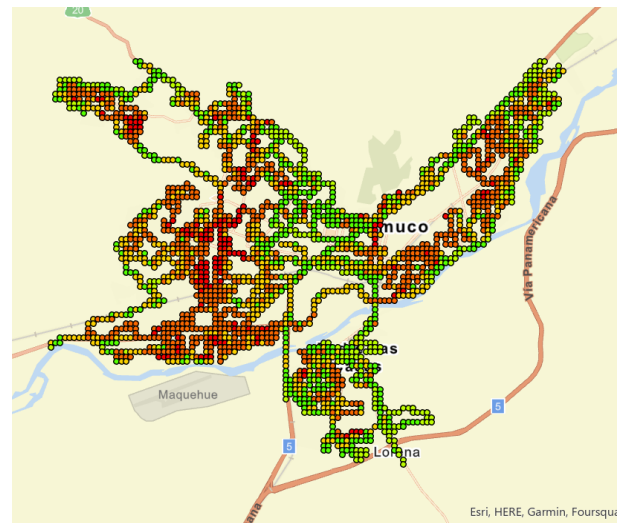
En la Fig. 3a), se observa las estadísticas básicas del PM_{2.5} agregado a nivel de segmentos de 50 m usando el diagrama de caja y bigotes. Esta figura indica que los datos fluctúan entre 6,50 y 555,65 $\mu\text{g}/\text{m}^3$ y la mediana de las mediciones PM_{2.5} es igual a 97,50 $\mu\text{g}/\text{m}^3$, lo sugiere que la mitad de las mediciones tiene un nivel del PM_{2.5} que indican un estado de alerta en ambas ciudades (i.e., 97,50 $\mu\text{g}/\text{m}^3$ pertenece al rango entre 80 y 109 $\mu\text{g}/\text{m}^3$). Estas estadísticas básicas no varían significativamente al agregar las mediciones del PM_{2.5} a nivel de grilla de 100 m (Ver Fig. 3b), pues la mediana es igual a 97,96 $\mu\text{g}/\text{m}^3$ y el valor mínimo y máximo son de 6,77 y 498,75 $\mu\text{g}/\text{m}^3$, respectivamente. Por lo tanto, se mantiene la mediana del PM_{2.5} en un estado de alerta para la zona de estudio.

TABLA III
NIVELES DE PM_{2.5} SEGÚN LA NORMATIVA CHILENA (BCN, S.F.).

Simbología	Estado	Concentración PM _{2.5} ($\frac{\mu\text{g}}{\text{m}^3}$)
	Bueno	0 – 50
	Regular	51 – 79
	Alerta	80 – 109
	Preemergencia	110 – 169
	Emergencia	≥ 170



a) agregación a nivel de segmentos de 50 m



b) agregación a nivel de grilla de 100 m

Fig. 2 Mediciones de contaminante PM_{2.5}.

IV. METODOLOGÍA

Primero, se exploró la relación entre las variables explicativas o independientes (factores) y la variable dependiente (PM2.5) usando el análisis de regresión global que minimiza la suma de los cuadrados de los residuales (OLS por sus siglas en inglés, Ordinary Least Squares), dada por (1).

$$y_i = \beta_0 + x_i\beta + \varepsilon_i \quad (1)$$

donde y_i representa la estimación de la medición del contaminante PM2.5 en la ubicación i , x_i es el vector de variables independientes (factores), β es el vector de coeficientes de regresión global, β_0 es el intercepto y ε_i es el error [29-31].

OLS asume que las observaciones son independientes y constantes en el área de estudio, y además los términos relacionados con el error no deben estar correlacionados [32]. Por el contrario, la regresión ponderada geográficamente (GWR, por sus siglas en inglés, Geographically Weighted Regression) permite las variables sean no estacionarias y modela las relaciones locales entre los predictores y la variable dependiente. La ecuación para el modelo GWR está dada por (2).

$$y_i = \beta_{i0} + \sum_{k=1}^m \beta_{ik} x_{ik} + \varepsilon_i \quad (2)$$

donde y_i es la estimación de la medición del contaminante PM2.5 en la ubicación i , x_{ik} es la variable independiente k -ésima en la ubicación i , m es el número de variables independientes, β_{i0} indica el intercepto en la ubicación i , β_{ik} es el coeficiente de regresión local para la k -ésima variable independiente en la ubicación i , y ε_i es el error aleatorio en la ubicación i [29].

Posteriormente, se implementó el modelo de error espacial (SER, por su sigla en inglés, Spatial Error Regression) que asume dependencia espacial en el error del OLS y descompone el error en la expresión (1) en dos términos. El primer término indica el componente espacial del error ($\lambda w_i \xi_i$) y el segundo término satisface el supuesto de regresión normal (ε_i), como se muestra en (3) [33].

$$y_i = \beta_0 + x_i\beta + \lambda w_i \xi_i + \varepsilon_i \quad (3)$$

donde se utilizan los mismos parámetros de OLS y además ξ_i indica la componente espacial del error, w_i es el vector de pesos espaciales, λ representa el nivel de correlación entre los componentes y ε_i es el término de error no correlacionado espacialmente. Para evaluar y comparar el desempeño de los modelos, se utilizaron distintos indicadores estadísticos tales como R^2 y AIC.

V. RESULTADOS

En esta sección, se presentan los resultados para cada uno de los modelos de regresión usando la agregación de los datos a nivel de segmentos de 50 m y grilla de 100 m. Para ello, se realizaron distintas iteraciones con los softwares ArcGIS Pro

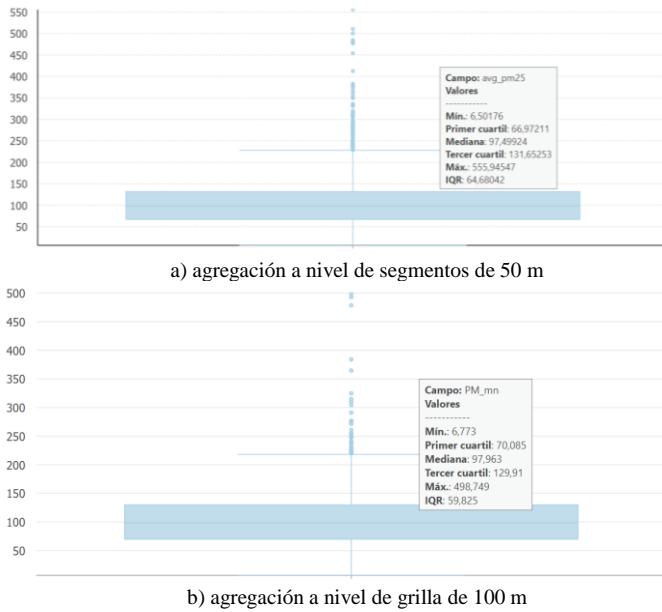


Fig. 3 Diagrama de caja y bigotes de las mediciones del contaminante PM2.5.

Las variables explicativas corresponden al área demográfica, uso de suelo, transporte, meteorología y altura, tal como lo se describe en la Tabla IV. Además, esta tabla indica el tipo de análisis espacial que se realizó para obtener un valor en cada centroide de segmentos de 50 m y grilla de 100 m.

TABLA IV
VARIABLES EXPLICATIVAS Y SUS RESPECTIVOS ANÁLISIS ESPACIALES.

Variable	Tipo de análisis espacial
Demográfica	
Población	Número dentro de buffers y valor de celda de la densidad de Kernel
Viviendas	
Mujeres	
Hombres	
Hogares con leña	
Uso de Suelo	
Áreas verdes	Área dentro de buffers y valor de celda de la densidad de Kernel
Residencial	
Infraestructura	
Equipamiento	
Actividades productivas	
Otros usos de suelo	
Transporte	
Transporte público	Largo dentro de buffers y distancia a segmento de vía más próximo
Carreteras	
Avenidas	
Calles locales	
Ferrovionario	
Congestión alta	
Congestión baja	
Meteorología	
Velocidad viento	Valor a estación de monitoreo más cercana
Humedad relativa	
Temperatura	
Precipitación	
Altura	Valor de la celda en modelo digital de elevación

and GeoDa hasta obtener el mejor modelo para cada regresión usando indicadores estadísticos.

A. OLS

En cada ejecución del modelo OLS, se verificó la significancia estadística de cada variable y así eliminar aquellas variables que no fuesen significativas ($p\text{-value} > 0,05$), se incluyeron las variables explicativas que con un VIF inferior a 7,5 para evitar la redundancia entre las variables y se mantuvieron aquellas variables independientes que tengan el sentido correcto con la variable dependiente. Además, se calculó el indicador Koenker para identificar la existencia de heterocedasticidad (no estacionario) en los modelos, el estadístico Jarque-Bera para determinar si los residuales (diferencia entre los valores observados y los predichos) se distribuyen normalmente y los indicadores R^2 y AIC para comparar el desempeño de los diferentes modelos.

Después de seis ejecuciones del OLS, se obtuvo los resultados mostrados en la Tabla V para los datos agregados a nivel de segmentos de 50 m. Estos resultados que sólo cinco variables son capaces de explicar la generación del contaminante PM2.5, entre ellas, la altura, congestión vehicular alta, las calles locales dentro de un buffer de 100 m y la densidad residencial y otros usos de suelo. La altura presenta el valor del coeficiente más alto indicando su impacto en el contaminante PM2.5. Esta tabla muestra que existe heterocedasticidad en el modelo OLS (Koenker es significativo) y que los residuales no se distribuyen normalmente (Jarque-Bera es significativo). Por lo tanto, hay una necesidad de analizar las variables con los modelos GWR.

Similarmente, se ejecutó cinco veces el OLS para los datos agrupados a nivel de grilla de 100 m, obteniéndose los resultados mostrados en la Tabla VI. Esta tabla indica que la densidad de infraestructura tiene el mayor impacto en la generación del contaminante PM2.5 mientras que la altura tiene una relación negativa significativa con el PM2.5. Nuevamente, este modelo OLS presenta heterocedasticidad y los residuales no son normalmente distribuidos. Al comparar los modelos OLS para ambos niveles de agregación, el AIC es inferior para los resultados del OLS con datos agrupados a nivel de grilla de 100 m y, además, se tiene un R^2 mayor para dicho modelo.

Un supuesto importante del modelo OLS es la independencia espacial de los residuales. Para ello, se calculó la autocorrelación espacial de los residuales usando el estadístico I de Moran con el software ArcGIS Pro. Este indicador detecta los patrones espaciales a nivel global (i.e., agrupación, dispersión o aleatoriedad). Un valor del indicador I de Moran cercano a 1 indica alta autocorrelación positiva (agrupación espacial) entre los datos, mientras que un valor cercano a -1 sugiere una alta autocorrelación negativa (dispersión espacial). Además, se calcula la puntuación z (z-score) y el p-value para determinar la significancia estadística [34], [35].

TABLA V
RESULTADOS DEL MODELO OLS A NIVEL DE SEGMENTOS DE 50 M.

Variable	Coficiente	VIF
Calles locales (buffer 100 m)	0,03*	1,88
Residencia (densidad Kernel)	0,01*	1,70
Otros usos (densidad Kernel)	0,19*	1,21
Congestión alta (distancia)	-0,05*	1,14
Altura	-0,48*	1,22
AIC	59254,97	
R^2	0,16	
Koenker	67,34*	
Jarque-Bera	19921,17*	

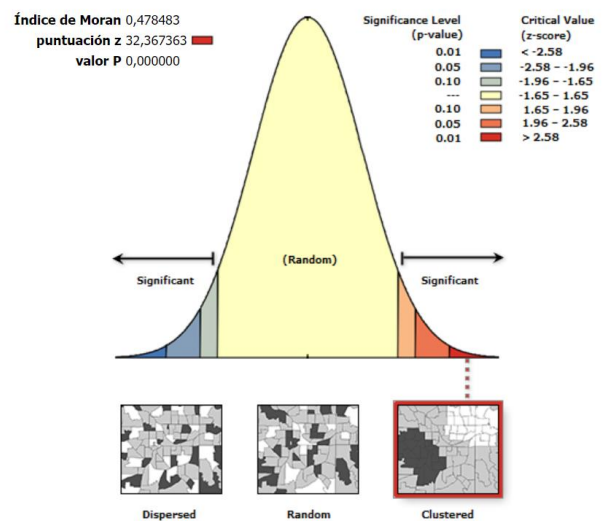
*p-value < 0,5

TABLA VI
RESULTADOS DEL MODELO OLS A NIVEL DE GRILLA DE 100 M.

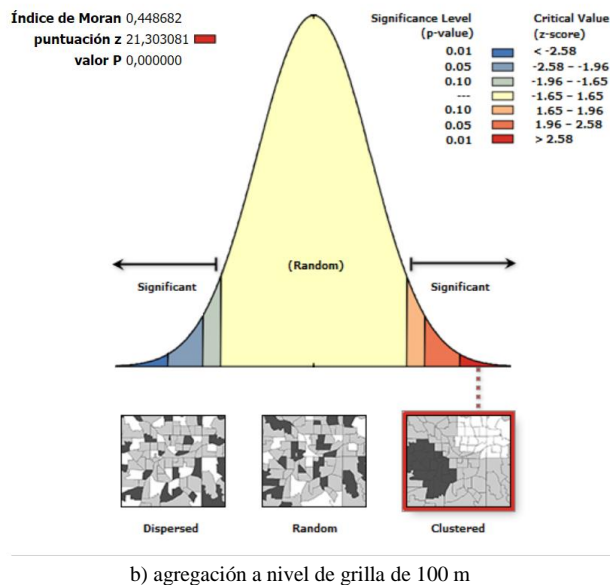
Variable	Coficiente	VIF
Áreas Verdes (buffer 100 m)	0,003*	1,03
Residencia (densidad Kernel)	0,02*	1,02
Infraestructura (densidad Kernel)	2,53*	1,10
Congestión alta (distancia)	-0,04*	1,07
Altura	-0,61*	1,11
AIC	24123,16	
R^2	0,19	
Koenker	17,38*	
Jarque-Bera	11.693,25*	

*p-value < 0,5

Las Fig. 4a) y 4b) presentan los resultados de la autocorrelación espacial de los residuales para ambos niveles de agregación de los datos. Como se puede apreciar de estas figuras, los residuales están agrupados, por lo tanto, es necesario implementar el modelo SER para poder abordar la dependencia espacial de los errores que aparecen en el modelo OLS.



a) agregación a nivel de segmentos de 50 m



b) agregación a nivel de grilla de 100 m
 Fig. 4 Autocorrelación espacial de los residuales usando el indicador I de Moran.

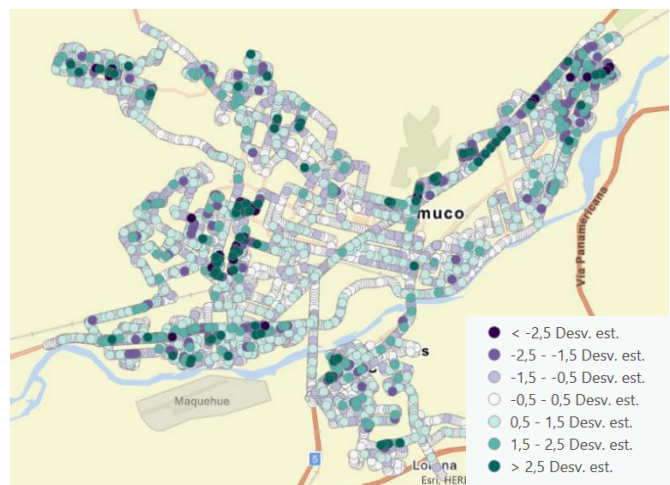
B. GWR

A continuación, se empleó el modelo GWR para abordar la heterocedasticidad en los modelos OLS usando las variables significativas mostradas en los resultados de la Tabla V a nivel de segmentos de 50 m y la Tabla VI a nivel de grilla de 100 m. Después de probar diferentes combinaciones de variables, se obtuvo los mejores resultados para ambos niveles de agregación de datos, tal como muestra en la Tabla VII. Esta tabla indica que la densidad residencial y la distancia a la alta congestión son variables que inciden en la generación del contaminante PM2.5 en ambos modelos. Además, se muestra un R^2 más alto al agregar los datos a nivel de segmentos mientras que el AIC es más bajo para el nivel de grilla de 100 m.

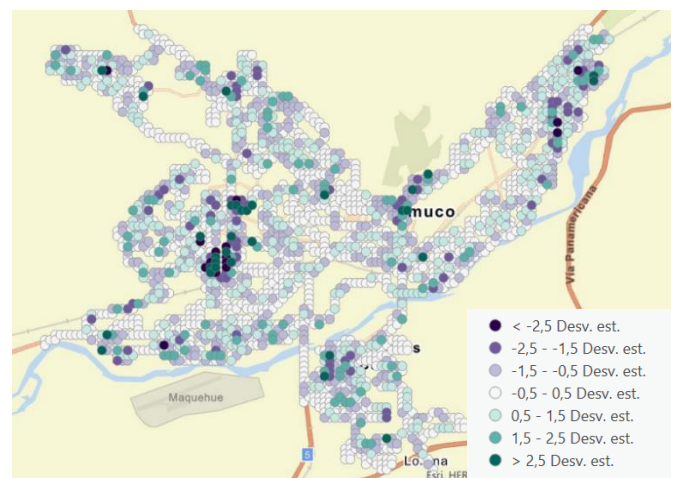
Las Fig. 5a) and b) muestran los residuales estandarizados para los resultados obtenidos en los modelos GWR usando datos agregados a nivel de segmentos de 50 m y grilla de 100 m, respectivamente. Estas figuras indican el nivel de ajuste de los modelos. Los residuales estandarizados positivos (negativos) corresponden a ubicaciones con valores observados más grandes (pequeños) que los valores estimados por el modelo. En general, las figuras muestran que los valores estimados con los modelos GWR se ajustan bastante a los valores observados del contaminante PM2.5.

TABLA VII
 RESULTADOS DEL MODELO GWR.

Nivel de agregación	Variables seleccionadas	AIC	R^2
Segmento de 50 m	Calles locales (buffer 100 m), Residencia (densidad Kernel), Congestión alta (distancia)	56694,78	0,64
Grilla de 100 m	Residencia (densidad Kernel), Congestión alta (distancia)	23249,27	0,61



a) agregación a nivel de segmentos de 50 m



b) agregación a nivel de grilla de 100 m

Fig. 5 Residuales estandarizados para los resultados de los modelos GWR.

C. SER

El modelo SER se utiliza para abordar la dependencia espacial en el error obtenido de los modelos OLS. En las Tablas VIII y IX, se muestran los resultados del modelo SER para datos agregados a nivel de segmentos de 50 m y de grilla de 100 m, respectivamente. La Tabla VIII indica que la variable altura tiene el mayor impacto en el contaminante PM2.5, seguido por la densidad de otros usos de suelo, mientras que en la Tabla IX se observa que la densidad de la infraestructura tiene una influencia significativa en la generación del contaminante PM2.5.

D. Comparación entre los modelos de regresión

En esta sección, se comparan los resultados de los modelos de regresión para los dos niveles de agregación de datos usando los indicadores estadísticos AIC y R^2 . En la Tabla X, se observa que el modelo GWR presenta el mejor desempeño con los valores más altos de R^2 y más bajos de AIC para ambos niveles de agregación de los datos. Además, se puede notar que el

modelo SER indica mejor ajuste que el modelo OLS tanto para el R² como el AIC. Por lo tanto, estos resultados sugieren que es importante abordar la heterocedasticidad con el modelo GWR y la dependencia espacial con el modelo SER para así mejorar la estimación de los factores que indican en la generación del PM2.5.

TABLA VIII
RESULTADOS DEL MODELO SER A NIVEL DE SEGMENTOS DE 50 M.

Variable	Coefficiente
Calles locales (buffer 100 m)	0,03*
Residencia (densidad Kernel)	0,01*
Otros usos (densidad Kernel)	0,19*
Congestión alta (distancia)	-0,04*
Altura	-0,48*
AIC	57697,00
R ²	0,42
Koenker	304,20*
Jarque-Bera	1555,92*

*p-value <0,0,5

TABLA IX
RESULTADOS DEL MODELO SER A NIVEL DE GRILLA DE 100 M.

Variable	Coefficiente
Áreas verdes (buffer 100 m)	-0,001*
Residencia (densidad Kernel)	0,02*
Infraestructura (densidad Kernel)	3,43*
Congestión alta (distancia)	-0,03*
Altura	-0,63*
AIC	23492,12
R ²	0,43
Koenker	128,78*
Jarque-Bera	628,98*

*p-value <0,0,5

TABLA X
RESULTADOS DE LA COMPARACIÓN ENTRE LOS MODELOS DE REGRESIÓN PARA AMBOS NIVELES DE AGREGACIÓN.

Indicador estadístico	OLS	GWR	SER
Segmentos 50 m			
AIC	59254,97	56697,78	57697,00
R ²	0,16	0,64	0,42
Grilla 100 m			
AIC	24123,16	23249,27	23492,12
R ²	0,19	0,61	0,43

VI. CONCLUSIONES

En el presente estudio, se utilizaron tres métodos de regresión espacial para estimar los factores que inciden en la generación del contaminante PM2.5 en la conurbación de Temuco y Padre Las Casas, Chile. En general, se aprecia que los resultados de los modelos pueden variar según el nivel de agregación de los datos que se utiliza en el análisis de regresión espacial. De esta manera, pueden existir diferentes variables significativas para el mismo tipo de modelo de regresión cuando se agregan los datos a los centroides de segmentos o celdas de una grilla.

Además, se puede apreciar que el desempeño entre los diferentes modelos varía al compararlos con los indicadores R² y AIC. Cada modelo identifica distintas variables significativas, obteniéndose el mejor ajuste entre las variables con el modelo GWR dándole énfasis a las variables residenciales, calles locales y congestión alta de vehículos. Estos resultados son

similares a los resultados presentados por los autores de los estudios en [3] y [11]. En particular, el factor relacionado con la densidad residencial es significativo en dichos estudios ya que la quema de leña en zonas residenciales es usada ampliamente por la población en la conurbación estudiada, principalmente por su bajo costo y contribuye significativamente a las emisiones de PM2.5.

Los resultados de este estudio permitirán adoptar medidas resolutivas para reducir el problema de la contaminación ambiental por PM2.5 a través, por ejemplo, de la implementación de políticas públicas. Algunas de las políticas públicas que se pueden implementar incluye el uso de calefacción sustentable más eficientes y menos contaminantes; incentivos para que la población reemplace calefacción a leña por alternativas más limpias; educación para concientizar a la población sobre los efectos negativos de la contaminación por quema de leña; y finalmente el aumento en la fiscalización para el cumplimiento de las normas de emisión y así desincentivar el uso de quema de leña [36]. Estas políticas deben ser rigurosas y contar con el respaldo gubernamental o municipal, garantizando así que toda la población afectada pueda cumplir con las medidas establecidas y que los cumplimientos de estas sean supervisados por las autoridades correspondientes. En futuro trabajo, se desea incluir otros tipos de modelos de regresión espacial y otras variables relevantes que puedan incidir en la generación del PM2.5 por quema de leña.

REFERENCIAS

- [1] Diario Financiero, "Estudio revela que menos del 1% de la Tierra tiene niveles seguros de contaminación del aire," 2023 [Online]. <https://www.df.cl/df-lab/cambio-climatico/estudio-revela-que-menos-del-1-de-la-tierra-tiene-niveles-seguros-de>
- [2] Organización Mundial de la Salud, OMS, "Contaminación del aire ambiente (exterior)," 2022 [Online]. [https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- [3] M. E. Quinteros et al., "Development of Spatio-Temporal Land Use Regression Models for Fine Particulate Matter and Wood-Burning Tracers in Temuco, Chile," *Environmental Science and Technology*, vol. 57, no. 48, pp. 19473–19486, 2023.
- [4] E. Purio, "¿Cómo se miden los niveles de PM 2.5?," Purio - Purificadores y Filtros de Aire, 2020 [Online]. <https://www.purioperu.com/como-se-miden-los-niveles-de-pm-2-5>.
- [5] IQAir, "Ciudades más contaminadas del mundo en 2021: Ranking PM2.5 AirVisual," 2022 [Online]. <https://www.iqair.com/es/world-most-polluted-cities>
- [6] United Nations Environment Assembly. UNEA, "Por un planeta sin contaminación," 2017 [Online]. https://wedocs.unep.org/bitstream/handle/20.500.11822/21839/UNEA2017_GovernmentToolkit_FINAL_SPANISH_FINAL.pdf?sequence=1&isAllowed=y
- [7] A. A. Schultz, J. J. Schauer, and K. MC. Malecki, "Allergic disease associations with regional and localized estimates of air pollution," *Environmental Research*, vol. 155, pp. 77–85, May 2017.
- [8] I. Romieu et al., "Multicity study of air pollution and mortality in Latin America (ESCALA study)," *Research Report (Health Effects Institute)*, no. 171, pp. 5–86, Oct. 2012.
- [9] Sistema Nacional de Información Ambiental, SINIA, "Capítulo 14 calidad del aire." 2021 [Online]. <https://sinia.mma.gob.cl/wp-content/uploads/2021/04/14-calidad-del-aire.pdf>
- [10] H. Jorquera et al., "Indoor PM2.5 in an urban zone with heavy wood smoke pollution: The case of Temuco, Chile," *Environmental Pollution*, vol. 236, pp. 477–487, May 2018.
- [11] M. E. Quinteros et al., "Spatio-temporal distribution of particulate matter and wood-smoke tracers in Temuco, Chile: A city heavily impacted by

- residential wood-burning,” *Atmospheric Environment*, vol. 294, no. 119529, Feb. 2023.
- [12] E. Blanco et al., “Spatial distribution of particulate matter on winter nights in Temuco, Chile: Studying the impact of residential wood-burning using mobile monitoring,” *Atmospheric Environment*, vol. 286, p. 119255, Oct. 2022.
- [13] A. M. Villalobos, F. Barraza, H. Jorquera, and J. J. Schauer, “Wood burning pollution in southern Chile: PM 2.5 source apportionment using CMB and molecular markers,” *Environmental Pollution*, vol. 225, pp. 514–523, Jun. 2017.
- [14] H. Jorquera, “Air quality management in Chile: Effectiveness of PM2.5 regulations,” *Urban Climate*, vol. 35, p. 100764, Jan. 2021.
- [15] Z. Varela et al., “Changes in epiphytic lichen diversity are associated with air particulate matter levels: The case study of urban areas in Chile,” *Ecological Indicators*, vol. 91, pp. 307–314, Aug. 2018.
- [16] M. A. Yáñez et al., “Urban airborne matter in central and southern Chile: Effects of meteorological conditions on fine and coarse particulate matter,” *Atmospheric Environment*, vol. 161, pp. 221–234, Jul. 2017.
- [17] J. Yang et al., “Estimation of PM2.5 concentration considering meteorological factors, policy, and interregional atmospheric transport,” *International Journal of Environmental Science and Technology*, vol. 19, no. 9, pp. 8479–8490, Oct. 2021.
- [18] W. Quan et al., “PM2.5 concentration assessment based on geographical and temporal weighted regression model and MCD19A2 from 2015 to 2020 in Xinjiang, China,” *PLOS ONE*, vol. 18, no. 5, pp. e0285610–e0285610, May 2023.
- [19] J. Shao et al., “Study on the relationship between PM2.5 concentration and intensive land use in Hebei Province based on a spatial regression model,” *PLOS ONE*, vol. 15, no. 9, p. e0238547, Sep. 2020.
- [20] M. Liu et al., “Synergistic analysis of atmospheric pollutants NO2 and PM2.5 based on land use regression models: a case study of the Yangtze River Delta, China,” *Environmental Monitoring and Assessment*, vol. 195, no. 9, Aug. 2023.
- [21] V. J. Berrocal et al., “A comparison of statistical and machine learning methods for creating national daily maps of ambient PM2.5 concentration,” *Atmospheric Environment*, vol. 222, pp. 117130–117130, Feb. 2020.
- [22] F.-J. Chang et al., “Explore spatio-temporal PM2.5 features in northern Taiwan using machine learning techniques,” *Science of The Total Environment*, vol. 736, p. 139656, Sep. 2020.
- [23] L. Chen et al., “Spatiotemporal modeling of PM 2.5 concentrations at the national scale combining land use regression and Bayesian maximum entropy in China,” *Environment International*, vol. 116, pp. 300–307, Jul. 2018.
- [24] Z. Li, J. Fung, and A. Lau, “High spatiotemporal characterization of on-road PM2.5 concentrations in high-density urban areas using mobile monitoring,” *Building and Environment*, vol. 143, pp. 196–205, Oct. 2018.
- [25] S. Hankey, and J. Marshall, “On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM 2.5, and particle size,” *Atmospheric Environment*, vol. 122, pp. 65–73, Dec. 2015.
- [26] M.-J. Jun, and Y. Gu, “Effects of transboundary PM2.5 transported from China on the regional PM2.5 concentrations in South Korea: A spatial panel-data analysis,” *PLOS ONE*, vol. 18, no. 4, p. e0281988, Apr. 2023.
- [27] Biblioteca del Congreso Nacional de Chile, BCN, “Reportes Comunal de Temuco,” 2023 [Online]. https://www.bcn.cl/siit/reportescomunales/comunas_v.html?anno=2017&idcom=9101
- [28] M. E. Quinteros et al., “Use of data imputation tools to reconstruct incomplete air quality datasets: A case-study in Temuco, Chile,” *Atmospheric Environment*, vol. 200, pp. 40–49, 2019.
- [29] A. Fotheringham, and T. Oshan, “Geographically weighted regression and multicollinearity: dispelling the myth,” *Journal of Geographical Systems*, vol. 18, no. 4, pp. 303–329, 2016.
- [30] M. Iturrieta, C. Blazquez, and K. Sauer-Brand, “Desigualdad socioeconómica y su asociación con accidentes de tránsito en el Gran Santiago, Chile,” Proceedings of the 21st LACCEI International Multi-Conference for Engineering, Education and Technology, 2023 [Online]. https://laccei.org/LACCEI2023-BuenosAires/papers/Contribution_207_a.pdf
- [31] N. Hamann and C. A. Blazquez, “Spatial modeling of confirmed cases of COVID-19 in Chile,” Proceedings of the 21th LACCEI International Multi-Conference for Engineering, Education and Technology, 2023 [Online]. https://laccei.org/LACCEI2023-BuenosAires/papers/Contribution_174_a.pdf
- [32] L. Anselin and D. Arribas-Bel, “Spatial fixed effects and spatial dependence in a single cross-section,” *Papers in Regional Science*, vol. 92, no. 1, pp. 3–17, Mar. 2013.
- [33] M. Ward and K. Gleditsch, “Spatial Regression Models,” Semantic Scholar, 2008. <https://www.semanticscholar.org/paper/Spatial-Regression-Models-Ward-Gleditsch/e08d6a024caccb1f2a7062266b09c902d5c4aa09e>
- [34] C. Blazquez, I. Puelma, and G. Khan, “Spatial Analysis of Bicycle Crashes in Chile,” Proceedings of the 18th International Conference on Intelligent Transportation Systems, pp. 2745–2750, Oct. 2015.
- [35] C. A. Blazquez et al., “Spatial autocorrelation analysis of cargo trucks on highway crashes in Chile,” *Accident Analysis & Prevention*, vol. 120, pp. 195–210, Nov. 2018.
- [36] ChileAtiende, “Medidas para mejorar la calidad del aire,” 2023 [Online] Gobierno de Chile, <https://www.chileatiende.gob.cl/fichas/35415-medidas-para-mejorar-la-calidad-del-aire>