



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA QUÍMICA, BIOTECNOLOGÍA Y
MATERIALES

MODELO DE DISPERSIÓN DE COMPUESTOS ORGÁNICOS VOLÁTILES PRIMARIOS EN TIEMPO REAL MEDIANTE SENSORES DE BAJO COSTO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL QUÍMICO

KEVIN BHAYRON VALENZUELA CONTRERAS

PROFESOR GUÍA:
FELIPE ANDRÉS DÍAZ ALVARADO

MIEMBROS DE LA COMISIÓN:
YEANICE ANDREA VÁSQUEZ PINO
JOSÉ CRISTIAN SALGADO HERRERA

Este trabajo ha sido parcialmente financiado por:
Airflux SpA

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL QUÍMICO
POR: KEVIN BHAYRON VALENZUELA CONTRERAS
FECHA: 2022
PROF. GUÍA: FELIPE ANDRÉS DÍAZ ALVARADO

MODELO DE DISPERSIÓN DE COMPUESTOS ORGÁNICOS VOLÁTILES PRIMARIOS EN TIEMPO REAL MEDIANTE SENSORES DE BAJO COSTO

La contaminación es una problemática a nivel mundial, solo en Chile 12.000 decesos son producto de este problema. Esta situación se agudiza en ambientes con múltiples fuentes de emisión, ya que los sistemas de monitoreo presentan una baja cobertura espacial y actualización cada una hora de los niveles de contaminación. En consecuencia, estos sistemas no permiten reconocer situaciones de riesgo en tiempo real, generando situaciones como los más de 1500 casos de intoxicación ocurridos el 2018 en Quintero y Puchuncaví.

Motivado por lo anterior, el proyecto de memoria desarrolla un modelo de dispersión en tiempo real, mediante el uso de sensores de bajo costo. Ya que estos dispositivos han permitido aumentar la cobertura espacial y temporal del seguimiento ambiental; además, los modelos de dispersión permiten aumentar el conocimiento de la distribución espacial de los contaminantes. Aplicando la metodología de interpolación de los vecinos naturales se desarrolla un algoritmo de rápida ejecución, que estima los niveles de concentración en un área definida, a partir de valores medidos en ubicaciones establecidas. Si bien el caso de estudio se centra en los compuestos orgánicos volátiles, se plantea una metodología aplicable a cualquier contaminante. Su ejecución permite actualizar los niveles de contaminación cada 2 minutos, por lo que puede ser utilizado como un sistema de alerta temprana de calidad del aire, al identificar rápidamente zonas con mayores niveles de contaminación. En consecuencia, el modelo tiene un potencial uso como sistema de alerta temprana de calidad del aire.

El desempeño del modelo se analiza mediante el mecanismo de validación cruzada *K-Fold* y *Leave-One-Out*, utilizando información validada de $MP_{2,5}$ de la red del SINCA. Se obtienen porcentajes de error entre las concentraciones estimadas y observadas del 30 y 40 % en las temporadas de invierno y verano respectivamente. Además, las estimaciones logran representar adecuadamente el ciclo diario de contaminantes, presentando un índice de correlación promedio entre las estimaciones y los valores observados de 0,8 en la temporada de verano y de 0,86 en la temporada de invierno.

La principal conclusión corresponde a que la complejidad del ambiente condiciona el número de equipos de monitoreo que aseguren estimaciones confiables. Al ser un método basado en interpolaciones su calidad depende significativamente de la ubicación de los equipos de monitoreo, especialmente su posición respecto a las zonas de máxima y mínima concentración en el plano. Las mejoras sugeridas tienen relación con la influencia del viento en el cálculo de los coeficientes de correlación, además del potencial acople de sensores móviles a la red de monitoreo, que permitirían responder a cambios dinámicos en el ambiente, al ubicarse en las zonas de máxima y mínima concentración.

Tabla de Contenido

1. Introducción	1
1.1. Antecedentes generales	1
1.2. Objetivos	3
1.2.1. Objetivo General	3
1.2.2. Objetivos específicos	4
1.3. Alcances	4
1.4. Metodología	4
2. Antecedentes sobre la atmósfera y la dispersión de COV	7
2.1. Atmósfera	7
2.1.1. La tropósfera	10
2.1.2. La estratósfera.....	14
2.1.3. La mesósfera y la termósfera.....	14
2.2. Química Atmosférica.....	15
2.2.1. Química de los COV en la tropósfera	17
3. Antecedentes sobre el monitoreo ambiental	20
4. Antecedentes sobre modelos de Dispersión	23
4.1. Modelos de dispersión.....	24
4.1.1. Modelos de dispersión físicos.....	26
4.1.2. Modelos de dispersión matemáticos.....	27
4.1.2.1. Métodos estocásticos	27
4.1.2.2. Métodos deterministas.....	29
4.1.2.2.1 Interpolación de distancia Inversa Ponderada.....	29
4.1.2.2.2 Interpolación de Vecinos Naturales	30
4.2. Software tradicionales.....	36
4.3. Validación cruzada.....	40
4.4. Estado del Arte Modelación Atmosférica	42
4.4.1. Proyecto Deo Nai, Vietnam.....	46
4.4.2. Proyecto NaturalKriging.....	48
4.4.3. Proyecto Valencia, España	49
5. Modelo propuesto para la dispersión en tiempo real de COV atmosféricos utilizando sensores de bajo costo	52
5.1. Algoritmo Original.....	55
5.2. Algoritmo modificado.....	57
5.3. Modelo	61

6. Resultados y validación del modelo	67
6.1. Test 1	72
6.2. Test 2	86
6.3. Test 3	92
7. Conclusión	106
Bibliografía	111
Anexos	118
Anexo A. Diagrama de Voronoi	118
Anexo B. Sistema Nacional de Información de Calidad del Aire SINCA	120
B.1. Estaciones de Monitoreo	120
B.2. Contaminantes Monitoreados SINCA	121
Anexo C. Diagramas de Voronoi Estaciones Sinca	124
Anexo D. Modelos de dispersión	133
Anexo E. Software	138
Anexo F. PID	140
Anexo G. Resultados	142
G.0.1. Test 1	142
G.0.2. Test 2	146

Índice de Tablas

4.1.	Comparación Modelos de dispersión. Elaboración propia.	35
4.2.	Comparación <i>Software</i> utilizados en la modelación atmosférica. Elaboración propia.	39
4.3.	Resultados validación cruzada para MP 2,5 [72].	47
4.4.	Resultados validación cruzada para CO [72].	47
4.5.	Resultados validación cruzada para CO ₂ [72].	48
4.6.	Comparación del estado del Arte de la modelación atmosférica en tiempo real. Elaboración propia.	50
6.1.	Concentraciones medias registradas en estaciones de monitoreo utilizadas para interpolación. Elaboración propia.	72
6.2.	Concentraciones medias registradas en estaciones de monitoreo utilizadas para validación. Elaboración propia.	73
6.3.	Concentraciones medias de las estaciones utilizadas para interpolación del primer set de datos. Temporada Invierno. Elaboración propia.	77
6.4.	Concentraciones medias estaciones utilizadas para la validación. Temporada Invierno. Elaboración propia.	77
6.5.	Resultados métricas de error media obtenidas en la temporada de verano, entre las concentraciones estimadas y observadas en estaciones de validación, para el método de los vecinos naturales y el caso A, B y C del algoritmo modificado en la temporada de verano.	81
6.6.	Resultados métricas de error media obtenidas en la temporada de invierno, entre las concentración estimadas y observadas en las estaciones de validación, para el método de los vecinos naturales y el caso A, B y C del algoritmo modificado.	81
6.7.	Comparación tiempo de calculo de Coeficientes de Correlación y peso de los datos generados en el calculo de los Coeficientes de Correlación.	83
6.8.	Comparación errores de interpolación desconectando la información de una ubicación de monitoreo. Temporada Invierno.	84
6.9.	Comparación errores de interpolación desconectando la información de una ubicación de monitoreo. Temporada Verano.	85
6.10.	Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.	86
6.11.	Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.	87
6.12.	Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.	89
6.13.	Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.	89
6.14.	Resultados métricas de error media obtenidas en la temporada de verano, entre las concentraciones estimadas y observadas en estaciones de validación, para el algoritmo de vecinos naturales VN y el caso A y C del algoritmo modificado. Test 2.	91

6.15.	Resultados métricas de error media obtenidas en la temporada de invierno, entre las concentraciones estimadas y observadas en estaciones de validación, para el algoritmo de vecinos naturales VN y el caso A y C del algoritmo modificado. Test 1	91
6.16.	Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.....	92
6.17.	Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.....	97
6.18.	<i>leave-one-out</i> . Comparación errores de interpolación desconectando la información de una ubicación, periodo Verano.	102
6.19.	Comparación errores de interpolación desconectando la información de una ubicación, periodo Invierno.	103
B.1.	Estaciones de monitoreo Región Metropolitana. Elaboración propia.....	120
B.2.	Niveles de calidad del aire de norma primaria de MP _{2,5} . Elaboración propia.....	122
B.3.	Ejemplo de tablas con colores de filas.	123

Índice de Ilustraciones

2.1.	Capas de la atmósfera y el perfil de temperatura en función de la altura. Tomado de [16].	8
2.2.	Escalas espaciales y temporales de variabilidad de los elementos atmosféricos. Adaptado y traducido de [16].	9
2.3.	Esquema del primitivo modelo de circulación global de Hadley". Tomado de [18].	11
2.4.	Sistema de circulación global [20].....	12
2.5.	Diagrama del cambio de la temperatura del aire con la altitud y difusión de contaminantes en una situación normal y de inversión térmica. Tomado de [21].	13
2.6.	Composición Química de la atmósfera terrestre Tomado de “El Cambio Climático”, por M. Molina, 2017, p. 34 [23].	15
2.7.	Perfiles diurnos medios de (NO_2), Humedad relativa (HR), temperatura, presión, concentraciones de O_3 , O_x , NO_2 , NO , HONO , CO observados, especies orgánicas volátiles antropogénicas (AVOC) e isopreno en el sitio SYY durante el EP1, EP2 y EP3, respectivamente. Tomado de [33].	19
3.1.	Visualización niveles de $\text{MP}_{2,5}$ en plataforma del estado [4].	21
3.2.	Montaje en paralelo de Equipos DOAS y sensores de bajo costo PID para su calibración.	22
4.1.	Mapas de distribución de las concentraciones ($\mu\text{g}/\text{m}^3$) de (A) BTEX y (B) NO_2 en Pamplona. Tomado de [39].	23
4.2.	Esquema localización de la fuente contaminante mediante sensores estáticos. Tomado de [47].	25
4.3.	Variable regionalizada en un espacio unidimensional. Por ejemplo, podría representar la concentración de un contaminante medida a lo largo del tiempo en una estación de monitoreo. Tomado de [55].	28
4.4.	Variograma [57].	29
4.5.	Ejemplo Interpolación por distancia inversa ponderada. Tomada de [59].	30
4.6.	Polígono de Voronoi en punto de interpolación sobre el diagrama de Voronoi inicial de los puntos de monitoreo. Tomado de [62].	31
4.7.	Diagrama de Voronoi dos puntos. Elaboración propia	32
4.8.	Diagrama de Voronoi incluyendo el punto de interpolación. Elaboración propia.	32
4.9.	Intersección celda de Voronoi en DV inicial. Elaboración propia.	33
4.10.	Software recomendados por diversas guías de modelación (a). Tomado de [66].	36
4.11.	Software recomendados por diversas guías de modelación (b). Tomado de [66].	37
4.12.	Mapa ubicaciones de medición estudio Kolkata [76].	42
4.13.	Sensor de bajo costo utilizado en estudio Kolkata.	43
4.14.	Configuración red de sensores Caso de estudio Lyon. Tomado de [5].	44

4.15.	Resultados de la simulación para las iteraciones 5, 15 y 25, respectivamente. (a) Distribución de contaminantes, (b) Diagrama de Voronoi de ubicación de los nodos, (c) reconstrucción del modelo. Tomado de [81].	46
4.16.	Dron DJI Inspire 2 montado con sensores de bajo costo para el monitoreo de calidad del aire. Tomado de [72].	47
4.17.	Comparación de RMSE de los 4 métodos de interpolación analizados. Tomado de [82].	49
5.1.	Esquema solapamiento áreas de influencia muestras [83]	53
5.2.	Diagrama de Voronoi a partir de cuatro nodos. Ejemplo interpolación Vecinos Naturales. Elaboración propia.	55
5.3.	Ejemplo Diagrama de Voronoi a partir de cinco puntos (a). Método de los Vecinos Naturales. Elaboración propia.	56
5.4.	Ejemplo Diagrama de Voronoi a partir de cinco puntos (b). Método de los Vecinos Naturales. Elaboración propia.	56
5.5.	Ejemplo Diagrama de Voronoi a partir de cuatro ubicaciones. Método iterativo de los Vecinos Naturales. Elaboración propia.	58
5.6.	Ejemplo Diagrama de Voronoi a partir de cinco nodos. Interpolación iterativa. Elaboración propia.	59
5.7.	Ejemplo DV Interpolación Iterativa, Interpolación 2do punto. Elaboración propia.	60
5.8.	Ejemplo DV de cinco puntos en Interpolación Iterativa, aplicando un cambio de orden de interpolación del 2do punto. El punto C_5 corresponde al punto C_6 en $(x_6, y_6) = (4.35, 2)$. Elaboración propia.	61
5.9.	Diagrama funcionamiento modelo en tiempo real de dispersión de contaminantes. Elaboración propia.	62
5.10.	Diagrama algoritmo modulo que realiza las particiones de Voronoi necesarias para implementar el método. Elaboración propia.	64
5.11.	Diagrama algoritmo módulo que calcula los coeficientes de correlación para cada punto de interpolación definido en la grilla de puntos. Elaboración propia.	65
5.12.	Diagrama algoritmo módulo que calcula las interpolaciones de los puntos definidos en la grilla de puntos. Elaboración propia.	66
6.1.	Mapa estaciones de monitoreo de calidad del Aire Región Metropolitana. En rojo estaciones utilizadas para la interpolación y en azul estaciones utilizadas para validación. Elaboración propia.	69
6.2.	Diagrama de Voronoi de ubicaciones de estaciones del SINCA utilizadas para la interpolación. Elaboración propia.	69
6.3.	Mapa de estaciones de monitoreo de calidad del Aire Región Metropolitana con información en línea de $MP_{2,5}$. Elaboración propia.	70
6.4.	Diagrama de Voronoi de ubicaciones de estaciones del SINCA utilizadas para la interpolación en test 2. Elaboración propia.	70
6.5.	Mapa de dispersión de concentración de $MP_{2,5}$ en Región la Metropolitana periodo Invierno	71
6.6.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia IVN.	73
6.7.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia. Caso A.	73
6.8.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia Caso B.	74

6.9.	Comparación Concentraciones reales e interpoladas en ubicación de estación Cerro Navia. Temporada Invierno, Caso C.....	74
6.10.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación de VN. Temporada Verano.....	75
6.11.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto IVN.....	76
6.12.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia IVN.....	77
6.13.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia All.....	78
6.14.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia Caso B.....	78
6.15.	Comparación Concentraciones reales e interpoladas en ubicación de estación Cerro Navia. Temporada Invierno, Caso C.....	78
6.16.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. VN. Invierno.....	79
6.17.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto IVN.....	80
6.18.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. VN. Temporada Verano.....	87
6.19.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IA. Temporada Verano.....	87
6.20.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IC. Temporada Verano.....	88
6.21.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. VN. Temporada Verano.....	88
6.22.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. VN. Temporada Invierno.....	90
6.23.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. VN. Temporada Invierno.....	90
6.24.	Comparación Concentraciones reales e interpoladas en ubicación estación Pudahuel para un mes de datos. Temporada Verano 2021.....	93
6.25.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. Temporada Verano 2021.....	93
6.26.	Comparación Concentraciones reales e interpoladas en ubicación estación Quilicura para un mes de datos. Temporada Verano 2021.....	94
6.27.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. Temporada Verano 2021.....	94
6.28.	Comparación Concentraciones reales e interpoladas en ubicación estación Parque O'higgins para un mes de datos. Temporada Verano 2021.....	94
6.29.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto para un mes de datos. Temporada Verano 2021.....	95
6.30.	Comparación Concentraciones reales e interpoladas en ubicación estación La Florida para un mes de datos. Temporada Verano 2021.....	96
6.31.	Comparación Concentraciones reales e interpoladas en ubicación estación Las Condes para un mes de datos. Temporada Verano 2021.....	96

6.32.	Comparación Concentraciones reales e interpoladas en ubicación estación Pudahuel para un mes de datos. Temporada Invierno 2021.	97
6.33.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. Temporada Invierno 2021.	98
6.34.	Comparación Concentraciones reales e interpoladas en ubicación estación Quilicura para un mes de datos. Temporada Invierno 2021.	98
6.35.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. Temporada Invierno 2021.	99
6.36.	Comparación Concentraciones reales e interpoladas en ubicación estación Parque O'higgins para un mes de datos. Temporada Invierno 2021.	99
6.37.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto para un mes de datos. Temporada Invierno 2021.	100
6.38.	Comparación Concentraciones reales e interpoladas en ubicación estación La Florida para un mes de datos. Temporada Invierno 2021.	101
6.39.	Comparación Concentraciones reales e interpoladas en ubicación estación Las Condes para un mes de datos. Temporada Invierno 2021.	101
A.1.	Ejemplo de Diagrama de Voronoi de puntos 1, ..., 6 [85].	118
B.1.	Técnicas de medición estación Las Condes. [4]	121
C.1.	Diagrama de Voronoi excluyendo estación de Quilicura para validación. Temporada Verano 2021.	124
C.2.	Diagrama de Voronoi considerando la estación de Quilicura para validación. Temporada Verano 2021.	125
C.3.	Diagrama de Voronoi excluyendo estación de Parque O'higgins para validación. Temporada Verano 2021.	125
C.4.	Diagrama de Voronoi considerando la estación de Parque O'higgins para validación. Temporada Verano 2021.	126
C.5.	Diagrama de Voronoi excluyendo estación de Pudahuel para validación. Temporada Verano 2021.	126
C.6.	Diagrama de Voronoi considerando la estación de Pudahuel para validación. Temporada Verano 2021.	127
C.7.	Diagrama de Voronoi excluyendo estación de Cerro Navia para validación. Temporada Verano 2021.	127
C.8.	Diagrama de Voronoi considerando la estación de Cerro Navia para validación. Temporada Verano 2021.	128
C.9.	Diagrama de Voronoi excluyendo estación de El Bosque para validación. Temporada Verano 2021.	128
C.10.	Diagrama de Voronoi considerando la estación de El Bosque para validación. Temporada Verano 2021.	129
C.11.	Diagrama de Voronoi excluyendo estación de Puente Alto para validación. Temporada Verano 2021.	129
C.12.	Diagrama de Voronoi considerando la estación de Puente Alto para validación. Temporada Verano 2021.	130
C.13.	Diagrama de Voronoi excluyendo estación de La Florida para validación. Temporada Verano 2021.	130
C.14.	Diagrama de Voronoi considerando la estación de La Florida para validación. Temporada Verano 2021.	131

C.15.	Diagrama de Voronoi excluyendo estación de Las Condes ara validación. Temporada Verano 2021	131
C.16.	Diagrama de Voronoi considerando la estación de Las Condes para validación. Temporada Verano 2021	132
D.1.	Ejemplo modelo de caja [87].....	134
D.2.	Ejemplo modelo de gauss [87].....	135
D.3.	Ejemplo variograma [87].....	137
E.1.	Esquema de procesamiento de datos sistema AERMOD [69].....	138
E.2.	Tabla parámetros de entrada sistema AERMOD [69].....	139
F.1.	Esquema de funcionamiento sensor de fotoionización.....	140
F.2.	Esquema funcionamiento interno sensor PID <i>alphasense</i>	141
G.1.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa A. Temporada Verano.	142
G.2.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa B. Temporada Verano.	143
G.3.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa C. Temporada Verano.	143
G.4.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IA. Temporada Verano.	143
G.5.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IB. Temporada Verano.	144
G.6.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IC. Temporada Verano.	144
G.7.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. IA. Temporada Invierno.	144
G.8.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. IB. Temporada Invierno.....	145
G.9.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1 Orden.....	145
G.10.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto All	145
G.11.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto All	146
G.12.	Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto Orden.....	146
G.13.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IA. Temporada Verano	146
G.14.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IC. Temporada Verano.	147
G.15.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IA. Temporada Invierno.	147
G.16.	Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IC. Temporada Invierno.....	147
G.17.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IA. Temporada Invierno.	148
G.18.	Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IC. Temporada Invierno.	148

Lista de Abreviaturas

Abreviaturas	Significado
Ar	Argon
AVOC	Compuestos Orgánicos Volátiles Antropogénicos
BTEX	Benceno, Tolueno, Etilbenceno y Xileno
CH ₄	Metano
Cl ⁻	Ion Cloruro
CO	Monóxido de Carbono
CO ₂	Dióxido de Carbono
COV	Compuestos Orgánicos Volátiles
DOAS	Espectroscopía de Absorción Óptica Diferencial
DV	Diagrama de Voronoi
EPA	Agencia de Protección ambiental de Estados Unidos
IVN	Interpolación de Vecinos Naturales
IDW	Interpolación de Distancia Inversa Ponderada
IOA	Índice de Correlación
MAE	Error Medio Absoluto
MAPE	Porcentaje de Error Absoluto
ME	Error Medio
MP _{2,5}	Material Particulado diámetro ≤ 2,5 micras
MP ₁₀	Material Particulado diámetro ≤ 10 micras
N ₂	Nitrógeno
NO	Óxido de Nitrógeno
NO ₂	Dióxido de Nitrógeno
NO ₃ ⁺	Ion Nitrato
NO _x	Óxidos de Nitrógeno
SINCA	Sistema Nacional de Información de Calidad del Aire
SO ₂	Dióxido de azufre
O ₂	Oxígeno
O ₃	Ozono
SEIA	Sistema de Evaluación de Impacto Ambiental
SEA	Servicio de Evaluación Ambiental
OH	Ion Hidroxilo
NH ₃	Amoniaco
RMSE	Error Cuadrático Medio
PE	Porcentaje de Error
VR	Variable Regionalizada

Capítulo 1

Introducción

1.1. Antecedentes generales

La contaminación ambiental es una problemática en Chile y el mundo. Según la OMS, a nivel mundial, la polución sería más letal que el tabaco, causando alrededor de siete millones de muertes anuales [1], y en Chile un 12 % de las muertes son generadas por ambientes contaminados, lo que equivale a 11.300 decesos al año [2]. Debido a lo anterior, en la Conferencia de la ONU sobre el Cambio Climático de 2017 (COP23), se elaboró el informe “Hacia un planeta libre de contaminación”, en el que se sentaban las bases de 50 estrategias a nivel global para personas, empresas y gobiernos, para luchar contra la contaminación y el cumplimiento de los Objetivos de Desarrollo Sostenible, y de las cuales 12 de esas estrategias están relacionadas con la contaminación del aire.

Los orígenes de los contaminantes son variados, sin embargo, se pueden clasificar según la naturaleza de la fuente emisora en biogénicas o antropogénicas. Las fuentes biogénicas son aquellas producidas por fenómenos propios de la naturaleza, se reconocen la erosión, los incendios forestales, las erupciones volcánicas, la descomposición de la vegetación, tormentas de polvo, entre otras. Por su parte, las fuentes antropogénicas son aquellas producidas por actividades o intervenciones que realizan las personas, siendo la principal causa la combustión de materiales, estas además se pueden clasificar en fuentes fijas, móviles y fugitivas, ejemplo de ellas pueden ser la chimenea de una industria, un vehículo en movimiento y la suspensión de partículas generada por un fenómeno atmosférico respectivamente.

Además, se pueden clasificar en contaminantes primarios y secundarios. Los primarios son aquellos emitidos directamente a la atmósfera, como el material particulado primario (MP), los Compuestos Nitrogenados (NOx), algunos Compuestos Orgánicos Volátiles (COV), etc. Los secundarios, son aquellos que se generan a raíz de reacciones en las que participan contaminantes primarios, como el ozono troposférico, hidrocarburos oxidados, Material particulado secundario o ultrafino, entre otros [3]. De esta forma, conocer y cuantificar los contaminantes primarios es el primer paso para comprender la dinámica de formación de contaminantes secundarios, la capacidad oxidativa de la atmósfera y el riesgo a la salud producido por estos.

En Chile, la gestión de la calidad del aire se inició con el Decreto 144 del Ministerio de Salud en 1961, que estableció normas para evitar emanaciones o contaminantes atmosféricos

de cualquier naturaleza. Sin embargo, el decreto no estableció límites de concentraciones, dejando esta potestad al Ministerio de Salud. Posteriormente, en 1978, mediante la Resolución 1.215 del Ministerio de Salud, se definieron concentraciones máximas para algunos contaminantes, como las partículas en suspensión, el anhídrido sulfuroso, el monóxido de carbono, algunos oxidantes fotoquímicos y dióxido de nitrógeno; originándose así la primera norma de calidad del aire en el país. En 1988, el Ministerio de Salud creó la Red de Monitoreo Automático de la Calidad del Aire y Variables Meteorológicas en la Región Metropolitana, llamada Red-MACAM [3].

Posteriormente, con la publicación del Reglamento del Sistema de Evaluación de Impacto Ambiental (SEIA) en 1997, aumentó el número de estaciones de monitoreo para el seguimiento y control de las emisiones, especialmente desde fuentes fijas industriales. Luego, en 2005, comenzó a implementarse el Sistema de Información Nacional de calidad de aire (SINCA), que concentra toda la información disponible de las estaciones de monitoreo de calidad del aire en Chile. Actualmente, la red de monitoreo del SINCA cuenta con información en línea de 204 estaciones de monitoreo, distribuidas a lo largo de todo el país y, de las cuales 57 son del Ministerio del Medioambiente [4]. Pese a lo anterior, los actuales sistemas de monitoreo son limitados, ya que no permiten una gran visualización espacial y temporal de los niveles de contaminación, reduciendo el monitoreo a ciertos puntos disgregados con actualización de la información de calidad del aire cada una hora [4].

Los actuales sistemas de monitoreo cuentan con metodologías de referencia, que utilizan tecnologías que tienen un alto valor, por lo cual su aplicación a gran resolución no siempre es la mejor alternativa [5]. Una alternativa viable o complementaria es el uso de sensores inalámbricos [6]-[7], por ejemplo, el sensor electroquímico, que es más pequeño, económico y genera mediciones de calidad del aire confiables posterior a una etapa de calibración y validación científica. Por lo tanto, se posicionan como una alternativa viable para el seguimiento de la calidad del aire [8]. El uso de sensores de inalámbricos en el monitoreo de calidad del aire en general tiene dos objetivos [9], el primero, de regular la contaminación atmosférica mediante el muestreo y mapeo de contaminantes; y el segundo, de detectar cuando se sobrepasan los umbrales de alerta de algún contaminante, como lo realizado por Boubrima en [10] y [11].

Por otro lado, el monitoreo y fiscalización se realiza solo en torno a contaminantes normados, dejando fuera del monitoreo gran cantidad de especies. Este es el caso de los COV, los que según organismos del estado podrían ser los principales precursores de los mas de 1000 casos de intoxicación ocurridos en Quintero y Puchuncaví en 2018 [12]. Por si fuera poco, aún no se cuenta con una norma específica para ellos. Actualmente, el Ministerio del Medio Ambiente se encuentra trabajando en la elaboración del anteproyecto de Norma Primaria de Calidad del Aire para COV y en el rediseño de la red de monitoreo de calidad del aire de la zona [13]. En Anexo B.2 se puede observar un detalle de los contaminantes actualmente medidos por las estaciones de monitoreo de la plataforma del SINCA.

Los COV son todos aquellos hidrocarburos que se presentan en estado gaseoso en condiciones normales, poseen propiedades volátiles, liposolubles, tóxicas e inflamables; algunos COV son el butano, el propano, el xileno, el cloro benceno, el tolueno, etc. Los COV tienen efectos directos en la salud y promueven la formación de contaminantes secundarios,

como el Ozono troposférico, que afecta de manera negativa en la fotosíntesis de las plantas y provoca enfermedades respiratorias y neurológicas en seres humanos [14]. En consecuencia, la importancia de estudiar su concentración en la atmósfera radica en prever los efectos directos a la salud de las especies, así como su rol en la formación de contaminantes secundarios de igual o mayor nocividad.

La deficiente normativa y calidad de la información enfatizan la problemática en zonas con alta actividad industrial, lo que pone en riesgo enormemente la vida humana, generando situaciones como la ocurrida en Quintero y Puchuncaví el 2018, zona que cuenta con 13 estaciones de monitoreo de calidad del aire y no ha podido plantear medidas efectivas para el control de la contaminación. Esta problemática es relatada en detalle por la Fundación TERRAM, en el informe “La negligente realidad de la Bahía de Quintero” [15], que analizó la precaria normativa y fiscalización existente en la zona. El estudio plantea la necesidad de monitorear COV como el tolueno, el benceno, los Compuestos Orgánicos Persistentes (COP), sustentado en la poca información pública disponible para poder evaluar la situación que afecta a la Bahía de Quintero.

Actualmente, Airflux se encuentran trabajando en una propuesta de norma primaria para compuestos orgánicos volátiles junto al Ministerio del Medio del Ambiente, para lo cual, busca generar y recopilar antecedentes técnicos que permitan elaborar una norma de calidad del aire de uno o más COV que presenten impactos negativos a la salud. Como resultado se espera contar con los antecedentes técnicos necesarios para elaborar una norma primaria de COV, donde se indique qué compuesto(s) es (son) viables de normar, su concentración límite y sistema de monitoreo para dar seguimiento a esta normativa. Debido a esto, surge la necesidad de utilizar las nuevas tecnologías presentes en el mercado como son los sensores de bajo costo, para responder a la necesidad de sesgo o error existente, monitorear con resolución espacial y temporal necesarias de una norma de calidad de aire sin costos elevados de instalación y operación.

Por la problemática anterior, el proyecto de título busca generar mayor conocimiento de la distribución espacial de contaminantes, mediante un Modelo de dispersión de Compuestos Orgánicos Volátiles primarios en la ciudad de Santiago, con actualización en tiempo real de los niveles de contaminación, mediante el uso de sensores de bajo costo alphasense modelo PID-A12. En anexo F se encuentra un detalle del funcionamiento del sensor. El modelo busca predecir en tiempo real la concentración de contaminantes en puntos donde no existe monitoreo ambiental, de manera de aumentar la visualización espacial y temporal del monitoreo de calidad del aire. De esta forma, el modelo servirá como un sistema de alerta temprana ante elevados índices de contaminación, identificando claramente las zonas con los mayores y menores niveles de contaminación.

1.2. Objetivos

1.2.1. Objetivo General

Modelar la dispersión de compuestos orgánicos volátiles primarios en la atmósfera con adquisición de datos en tiempo real generados por una red de sensores de bajo costo

1.2.2. Objetivos específicos

1. Analizar los fenómenos fisicoquímicos que afectan a los COV en la atmósfera.
2. Comparar modelos de dispersión de contaminantes atmosféricos.
3. Comparar los Software utilizados para la modelación atmosférica.
4. Formular y resolver un modelo de dispersión de COV primarios con adquisición de datos en tiempo real.

1.3. Alcances

Los alcances de este trabajo parten desde la adquisición de datos en tiempo real de la plataforma ThingSpeak. El desarrollo de los dispositivos de medición y su electrónica son ejecutados por la empresa Airflux. Del mismo modo, la calibración de los sensores y conexión a la plataforma ThingSpeak, es desarrollada por otras personas del equipo de Airflux.

En concreto, el trabajo realizado se centra en el análisis de los modelos de dispersión existentes y los software utilizados para la modelación atmosférica, para luego seleccionar los más apropiados para la implementación de un modelo de dispersión de COV en tiempo real. Junto a lo anterior, en el desarrollo del modelo seleccionado en la ciudad de Santiago, a partir del uso de cinco sensores PID de COV. Por otro lado, en la validación del modelo mediante información real presente en la plataforma del SINCA de todas las estaciones de monitoreo de la Región Metropolitana.

1.4. Metodología

La metodología de trabajo consistirá en primera instancia en una investigación bibliográfica respecto a los principales fenómenos fisicoquímicos que afectan a los COV en la atmósfera, junto a ello, se revisan los principales modelos de dispersión atmosféricos desarrollados y los principales *software* utilizados en su modelación. Luego, se selecciona el modelo a implementar basado en el menor costo computacional de su implementación, así mismo, se escoge el software que presenta mayor flexibilidad para implementar el modelo seleccionado con adquisición de datos en tiempo real. Finalmente, se implementa un modelo computacional para la dispersión de compuestos orgánicos volátiles en tiempo real, a partir de la adquisición de datos de concentración en el espacio generados por una red de sensores. La evaluación del método se realiza mediante una metodología de validación cruzada *K-Fold* y *Leave- One-Out*.

Para lo anterior, el estudio incluyó una revisión bibliográfica en revistas técnicas, analizando principalmente fuentes de información primaria y secundaria. Para el análisis de los modelos aplicados en tiempo real se incluyen publicaciones principalmente de la última década, para la profundización de modelos, fenómenos físicos y químicos, se priorizó recurrir a fuentes primarias, principalmente libros. Para el análisis de los *software* utilizados, se recurrió a publicaciones en revistas científicas, páginas de gobierno e instituciones medioambientales.

Se utiliza la herramienta computacional “Matlab” para el desarrollo del modelo. Para la adquisición de datos en tiempo real se utilizó la nube de Matlab “*ThingSpeak*”, especialmente destinada para proyectos que utilizan sensores. Para la adquisición de datos en tiempo real, se utilizaron cinco sensores de bajo costo Alphasense PID-A12 de compuestos orgánicos volátiles.

A continuación, se presenta la metodología específica para lograr el cumplimiento de los objetivos planteados.

- Objetivo 1: Analizar los fenómenos fisicoquímicos que afectan a los COV en la atmósfera

Este objetivo busca identificar las características que dominan los fenómenos físicos y químicos que afectan a las especies en la atmósfera. En particular, los que afectan a los COV una vez emitidos, considerando los principales fenómenos y reacciones que ocurren en la atmósfera. Para lo anterior, se prosigue de la siguiente forma:

1. Se analiza el sistema atmósfera, identificando las principales características atmosféricas y geográficas que afectan en el transporte y reacción de los compuestos orgánicos volátiles.
2. Se Identifican los principales COV presentes en la atmósfera.
3. Se analizan las principales reacciones a las que se ven expuestos los COV identificados.

- Objetivo 2: Comparar modelos de dispersión de contaminantes atmosféricos.

Este objetivo busca identificar los modelos de dispersión existentes y caracterizarlos, para posteriormente evaluar cuál de ellos es el más apropiado para modelar la dispersión de contaminantes con adquisición de datos en tiempo real. Para lo anterior, se prosigue de la siguiente forma:

1. Se Identifican los principales modelos de dispersión atmosféricos utilizados y sus requerimientos.
2. Se Identifican los principales usos y alcances de cada modelo.
3. Se comparan los modelos identificados.

- Objetivo 3: Comparar los *Software* utilizados para la modelación atmosférica

Este objetivo busca identificar los principales software utilizados en el modelamiento atmosférico. Posteriormente, se evalúa cuál de ellos es el más apropiado para modelar la dispersión de contaminantes con adquisición de datos en tiempo real. Para lo anterior, se prosigue de la siguiente forma:

1. Se Identifican las principales herramientas computacionales utilizadas actualmente para el modelamiento de contaminantes atmosféricos.
2. Se identifican los principales usos y alcances de cada herramienta computacional.
3. Se comparan las herramientas identificadas.

- Objetivo 4: Formular y resolver un modelo de dispersión de COV primarios con adquisición de datos en tiempo real.

Este objetivo busca formular un modelo de dispersión de contaminantes rápido y eficaz, con adquisición de datos en tiempo real de una red de sensores de bajo costo. Para lo anterior, se prosigue de la siguiente forma:

1. Se selecciona un modelo de dispersión.
2. Se selecciona una herramienta computacional.
3. Se implementa el modelo de dispersión en tiempo real.

Capítulo 2

Antecedentes sobre la atmósfera y la dispersión de COV

Para comprender el comportamiento de estos compuestos, es importante estudiar el sistema en el cual se emiten, transportan y reaccionan. Conocer las características del sistema, sus fenómenos fisicoquímicos y las variables externas que los afectan, permiten prever el comportamiento de los COV y anteponerse a situaciones de riesgo a la salud de las personas y el medio ambiente. Por lo anterior, en este capítulo, se estudiará la atmósfera terrestre y sus componentes, posteriormente, se presentarán las principales características de la química troposférica en la que se encuentran involucrados los COV.

2.1. Atmósfera

La atmósfera corresponde a la parte gaseosa que rodea el planeta tierra, según algunos autores tiene un espesor de aproximado de 500 km de altitud [16], altura a la cual la densidad de la atmósfera es prácticamente nula, de hecho, el 75 % de su masa se concentra en los primeros 11 km de altura [16]. Como sistema termodinámico, la atmósfera está caracterizada por tres variables fundamentales: presión, temperatura y composición. La variación de estas componentes genera que la atmósfera no sea uniforme en todo su volumen, de hecho, la variación de su perfil de temperatura con la altura es la base para distinguir varias capas atmosféricas, con distinta composición y propiedades físicas.

Por su parte, la presión atmosférica en la superficie de un planeta se define como el peso del total de la masa de la atmósfera dividida por el área de la superficie del planeta. Como resultado de la competencia entre la aceleración de gravedad y la cinética de traslación de las masas de aire, la presión atmosférica disminuye con la altura [17]. La presión en el Sistema Internacional de Unidades (SI) se define por el Pascal (Pa), en estas unidades de medida la presión atmosférica en la superficie de la Tierra es de $1,01325 \times 10^5$ [Pa] y varía hasta en seis ordenes de magnitud desde el suelo hasta su parte más alta [16].

La variación de temperatura con la altura no es tan evidente como en el caso de la presión, de hecho, el gradiente de temperatura cambia de signo varias veces a lo largo de atmósfera, lo que permite dividirla en cuatro regiones. Estas regiones se distribuyen desde el suelo y al aumentar en la altura en el siguiente orden: Tropósfera, Estratósfera, Mesósfera y Termósfera

como se muestra en la figura 2.1. La temperatura media en la superficie de la Tierra es de 288 [°K] (15°C) [16].

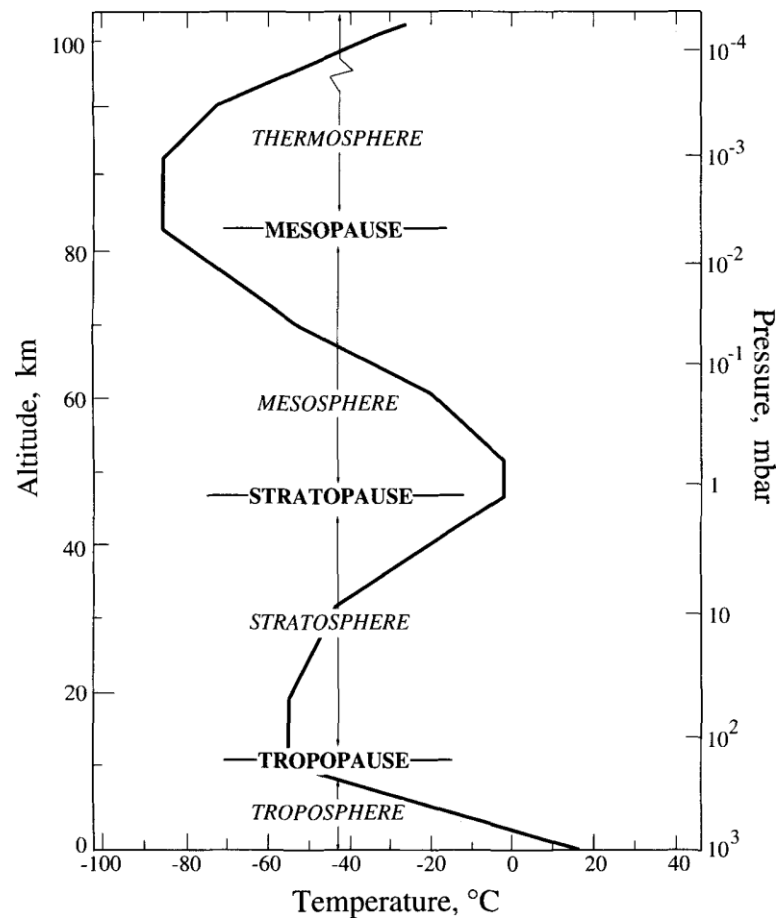


Figura 2.1: Capas de la atmósfera y el perfil de temperatura en función de la altura. Tomado de [16].

La composición de las distintas capas de la atmósfera puede verse alterada por diferentes causas, como la emisión de contaminantes o las reacciones químicas entre las especies que se encuentran en ella. Estas reacciones, ocurren principalmente por vía fotoquímica e influyen en los perfiles de temperatura y concentración de las especies presentes en las diferentes capas. Producto de estas reacciones, no todas las especies poseen los mismos tiempos de residencia atmosféricos, en la figura 2.2 se observan las escalas temporales y espaciales de las principales especies en la atmósfera.

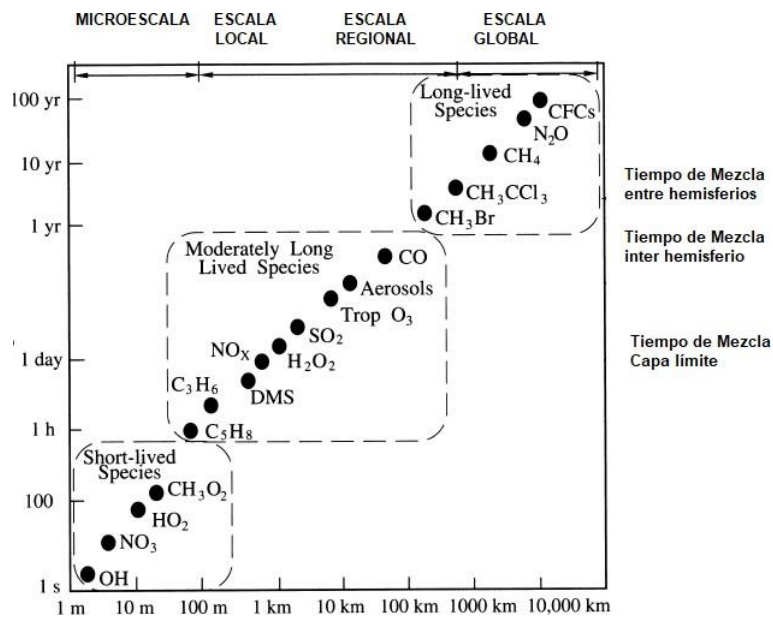


Figura 2.2: Escalas espaciales y temporales de variabilidad de los elementos atmosféricos. Adaptado y traducido de [16].

Como se puede observar, los contaminantes y especies químicas en la atmósfera presentan múltiples escalas espaciales y temporales de hasta 8 órdenes de magnitud [16]. De esta forma, se identifican 4 escalas de movimiento atmosféricas: la microescala, la mesoescala, la escala sinóptica y la escala global [16]. A continuación se detalla cada una de estas.

1. Microescala: Los fenómenos de movimiento se producen del orden de 0-100 [m], como la dispersión de contaminantes de una chimenea.
2. Mesoescala: Los fenómenos de movimiento se producen del orden de decenas a cientos de kilómetros, como los vientos que se producen en zonas costeras.
3. Escala sinóptica: Los fenómenos de movimiento se producen del orden de cientos a miles de kilómetros, como los ciclones y anticiclones.
4. Escala global: Los fenómenos de movimiento se producen en ordenes superiores a los 5×10^3 kilómetros, como los vientos alisios u otras corrientes atmosféricas de escala global

Muchos fenómenos atmosféricos se solapan en las diversas escalas. Por ejemplo, la contaminación atmosférica urbana y regional (1,2), los fenómenos asociados a la interacción aerosoles/clima (2,4), el aumento de gases de efecto invernadero (2,3,4) y la capacidad oxidativa de la atmósfera (1,4). Además de la escala espacial, es importante considerar los tiempos de residencia atmosféricos de las especies, que corresponde a la vida media de la especie una vez emitida a la atmósfera, los cuales van desde los segundos para los radicales libres, hasta muchos años en especies más estables. Por ejemplo, el radical hidroxilo (OH),

tiene un tiempo de residencia atmosférica del orden de las fracciones de segundo, mientras que el metano (CH_4), presenta un tiempo de residencia en la atmósfera de 10 años [16].

La escala temporal y espacial es de gran importancia en los problemas de modelación atmosférica, dependiendo de los fenómenos que se quieren considerar se pueden o no realizar varias simplificaciones. Por ejemplo, en un modelo de transporte atmosférico de escala global no tiene mucha utilidad considerar el transporte de especies con bajos tiempos de residencia atmosféricos, las cuales no lograrían visualizarse en esta escala. Sin embargo, si podría ser de interés en un modelo de escala global, en el cual se analice la variación de la capacidad oxidativa de la atmósfera, considerar especies con bajos tiempos de residencia atmosféricos, como el radical hidroxilo, principal agente oxidante de la atmósfera.

2.1.1. La tropósfera

La tropósfera es la capa que se encuentra justo en contacto con la superficie de la Tierra, su estudio es de gran relevancia, ya que es el lugar donde se desarrollan los seres vivos. Además, es la región de la atmósfera donde ocurren todos los fenómenos meteorológicos que afectan el clima, prácticamente todo el vapor de agua, nubes y precipitaciones ocurren en esta zona. El espesor medio de esta capa varía entre 10 km y 15 km de altitud, dependiendo la latitud y época del año [16].

En la tropósfera, la temperatura desciende con la altitud a razón de 5 a 10 °C por km de altura y su descenso es función del porcentaje de agua contenido en la parcela de aire, por ejemplo, en presencia de aire seco la temperatura desciende a razón de 10°C por cada kilómetro de altitud [16], por otro lado, la temperatura varía horizontalmente de unas zonas a otras del planeta. Debido a estos gradientes de temperatura, la tropósfera está muy sometida a perturbaciones y procesos turbulentos, que mantienen el aire en constante agitación, dando lugar a movimientos horizontales y verticales de las masas de aire. Estos movimientos de las masas de aire permiten transportar y diluir la contaminación hasta lugares muy alejados de sus fuentes de emisión.

La tropósfera se divide en dos subcapas:

1. Capa límite: Se extiende desde la superficie hasta unos 2 km de altura aproximadamente. Es una región turbulenta y de mezcla rápida. La mezcla turbulenta es causada por el incremento de la fricción entre las masas de aire con la superficie terrestre y puede servir localmente para concentrar y transportar contaminantes
2. Tropósfera libre: La circulación atmosférica no se encuentra impedida por la superficie, pero si afectada en menor medida. Como consecuencia de la disminución de la temperatura con la altitud, la mezcla convectiva vertical es relativamente rápida, del orden de las horas, en comparación a la mezcla convectiva horizontal en todo el hemisferio, del orden de uno o dos meses [16]. Por lo anterior, solo especies con tiempos de vida atmosféricos relativamente largos, como el metano CH_4 y el dióxido de nitrógeno NO_2 muestran una distribución homogénea dentro del hemisferio en esta parte de la tropósfera, mientras que las especies con tiempos de vida atmosféricos

cortos presentan variaciones entre las diferentes masas de aire, dependiendo de la región, factores meteorológicos y geográficos.

Uno de los fenómenos más importantes que sucede en la tropósfera es el sistema de circulación global, que determina en gran medida el clima de la superficie. Se genera producto de las masas de aire que redistribuyen el calor que llega del Sol, desde las zonas más calientes, hacia las más frías. Este fenómeno se ilustra en la figura 2.3.

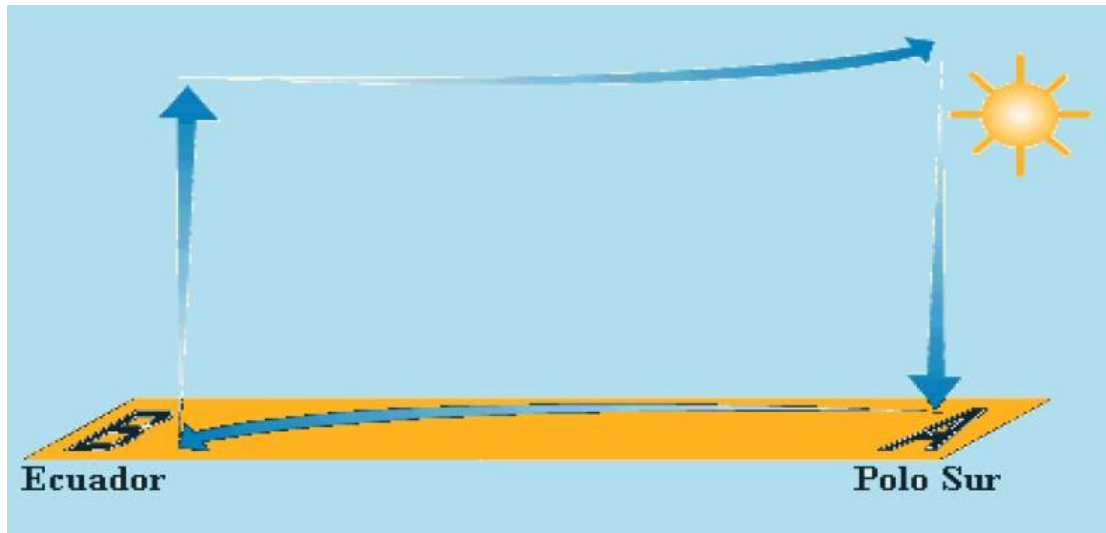


Figura 2.3: Esquema del primitivo modelo de circulación global de Hadley". Tomado de [18].

El ciclo inicia en el ecuador con el calentamiento del suelo, que a su vez provoca el calentamiento de las masas de aire en contacto con la superficie y su ascenso a zonas más frías. Durante el ascenso, este aire se expande y el vapor de agua que arrastra se enfría y finalmente, se condensa y produce lluvias. Conforme las masas de aire se desplazan hacia los polos, se enfrían y retornan a la superficie, desde donde inician el retorno hacia la zona ecuatorial [19].

La figura 2.3 muestra una simplificación del sistema de circulación global. En realidad el sistema de circulación esta constituido por varias celdas de convección. En la figura 2.4 se puede observar un diagrama con las distintas corrientes de aire que circulan el globo, conocidas como células de convección.

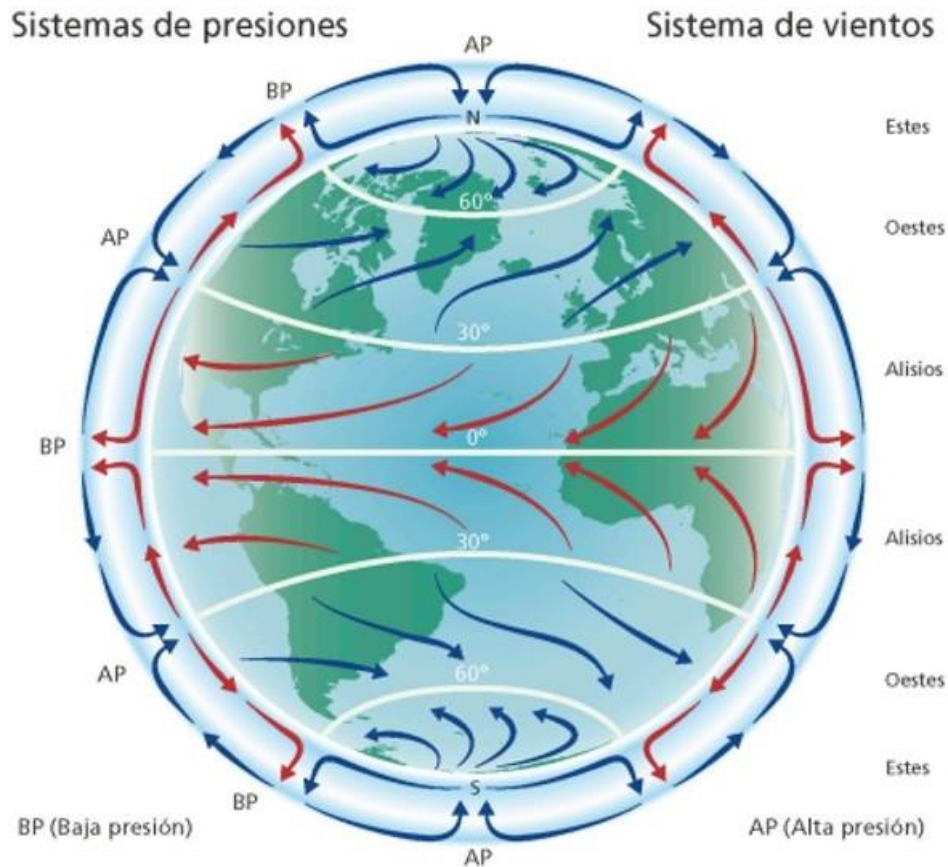


Figura 2.4: Sistema de circulación global [20].

En la figura 2.4 se pueden observar los sistemas de presiones generados por los movimientos de las masas de aire. En primer lugar, al calentarse las masas de aire en el suelo (pto 1), estas se vuelven menos densas y se desplazan a partes superiores de la atmósfera (pto 2), lo que genera un sistema de baja presión en el punto de ascenso (pto 1), al generarse pseudo vacío en la zona donde las masas de aire se levantaron. Las masas de aire al elevarse van desplazando horizontalmente las masas de aire circundantes, hacia un punto de acumulación en la altura (pto 3), esta acumulación genera un sistema de alta presión en la altura, el cual genera el retorno de las masas de aire al suelo (pto 4). Posterior a esto, las masas de aire que retornan al suelo (pto 4), desplazan horizontalmente las masas de aire circundantes hacia el punto inicial (pto 1).

Además del sistema de circulación global de aire troposférico, se deben considerar los fenómenos locales de movimientos de aire influenciados por la topografía del lugar. Estos movimientos son generados por el ciclo diario solar, que calienta diferenciadamente la superficie del lugar, por ejemplo, en sectores costeros el sol calienta más rápido la superficie del suelo que la del mar, generando un ascenso de las masas de aire por encima de la superficie del suelo y un desplazamiento de las masas de aire por encima de la superficie del agua, induciendo de esa forma un viento en dirección mar-tierra. Así mismo, durante la noche, la superficie del suelo se enfría más rápido que la superficie del agua, por lo tanto, las masas de aire por encima del agua son las cálidas, provocando un ascenso de las masas de aire por encima de del agua, induciendo un desplazamiento del aire en dirección tierra mar. Por otro lado, las

islas de calor urbana son fenómenos generados por la capacidad calórica de los materiales de construcción. El calentamiento de la ciudad puede provocar una circulación que induce el intercambio de masas de aire con las zonas aledañas a esta [19].

En ocasiones, los fenómenos meteorológicos pueden hacer que el gradiente negativo de temperatura se acentúe, disminuya e incluso se interrumpa a una altitud determinada. El gradiente de temperatura regula los movimientos verticales de las masas de aire en la troposfera y, por lo tanto, la dispersión de la contaminación. Cuando el gradiente vertical de temperatura se interrumpe, la temperatura del aire aumenta con la altura, lo que se conoce como inversión térmica, el cual se puede observar en la Figura 2.5.

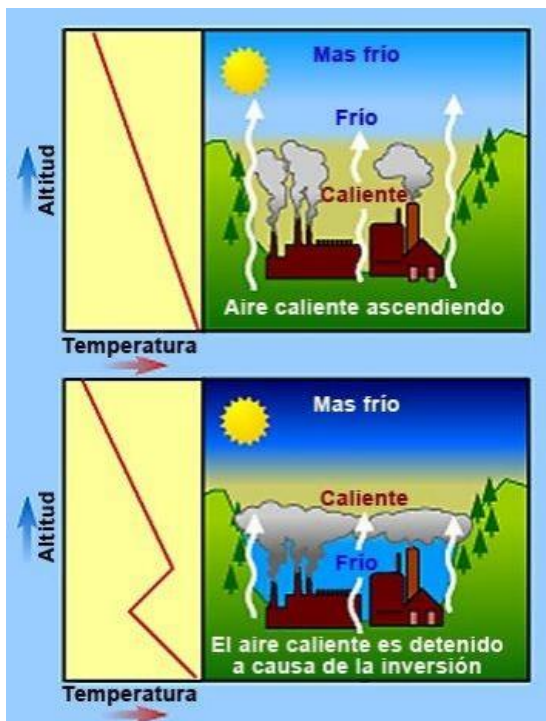


Figura 2.5: Diagrama del cambio de la temperatura del aire con la altitud y difusión de contaminantes en una situación normal y de inversión térmica. Tomado de [21].

Las masas de aire por debajo de la inversión, más frías y densas, no pueden elevarse por encima del aire más cálido de la zona de inversión. Como consecuencia, se produce el aislamiento de las masas de aire que quedan por debajo y por encima de esta inversión. Este fenómeno de inversión térmica juega papel importante en la dispersión de contaminantes en la tropósfera, ya que pueden provocar que el volumen disponible para que se diluyan los gases emitidos sea menor, esto puede generar episodios de contaminación críticos.

Este es el caso de Santiago, que presenta un serio problema de contaminación atmosférica, de hecho, en 1996 de acuerdo con el decreto supremo N°131 fue declarada zona saturada por Ozono, material particulado, entre otros y, además, fue declarada zona latente por dióxido de nitrógeno, registrando los mayores niveles de contaminación del país [22]. Entre de las variables que generan esta situación de contaminación crítica en la ciudad de Santiago se

encuentran las condiciones meteorológicas, las múltiples emisiones y la topografía. De los aspectos meteorológicos más relevantes se encuentra la capa inversión térmica de subsidencia, provocada por el anticiclón subtropical del Pacífico [22], en invierno ésta se sitúa a muy baja altura y en verano se eleva por efectos del calentamiento atmosférico. La inversión térmica tiene una altura que es variable dependiendo de la temperatura. Aunque en promedio se ubica a unos 600-800 metros sobre el nivel del mar y a unos 200-300 metros sobre Santiago.

Además del efecto generado por la meteorología y las emisiones, se encuentra el inducido por la topografía. En Santiago, los factores topográficos más relevantes son la cordillera de los andes al este de la ciudad, el cordón de Chacabuco al norte, los cerros de la angostura de Paine al sur y la cordillera de la costa al oeste. Estos provocan condiciones desfavorables para la dispersión de contaminantes al formar la conocida cuenca de Santiago [22]. Estas adversas condiciones para la dispersión de contaminantes fuera de la ciudad, sumada a la mezcla turbulenta en la baja atmósfera, sustenta el estado de saturación para varios contaminantes en la ciudad de Santiago. Por lo tanto, es poco probable la presencia de pronunciados perfiles de concentración espacial y temporal.

2.1.2. La estratósfera

La estratosfera es la capa de la atmósfera comprendida entre los 11 km y los 50 km de altura aproximadamente [16]. Recibe ese nombre porque el aire se mueve en estratos horizontales y el intercambio de gases verticalmente entre los estratos es escaso. En esta capa la temperatura aumenta con la altura hasta alcanzar los 0 [°C], producto de las reacciones fotoquímicas producidas por la absorción de la radiación ultravioleta por parte del ozono. Este aumento de temperatura con altura impide el ascenso de las masas de aire más frías y densas por encima de las más calientes y menos densas. Por este motivo, en esta sección no se producen fenómenos meteorológicos, ya que, según lo visto en la sección anterior sobre la tropósfera, son los desplazamientos verticales de las masas de aire las que inducen la precipitación.

Entre los rasgos más importantes de esta capa del punto de vista de su composición, se encuentra la presencia prácticamente nula de vapor de agua. Además, en esta zona, el ozono es un componente que aglomera cerca del 80 % del O₃ de la atmósfera, en lo que se conoce como capa de ozono. Otros compuestos presentes en la estratosfera, minoritarios, pero importantes desde el punto de vista de las reacciones químicas que se producen a esta altitud, son el metano CH₄, los óxidos de nitrógeno NO_x o los compuestos halogenados. En esta sección toman relevancia los compuestos que participan en los procesos de destrucción de la capa de ozono, un COV que forma parte de este grupo es el clorofluorocarbono (CFC), estos son compuestos muy estables, que presentan altos tiempos de permanencia atmosféricos, que le permiten alcanzar la estratósfera y por medio de reacciones de fotólisis con la radiación UV procede a destruir la capa de ozono [17].

2.1.3. La mesósfera y la termósfera

Estas son las capas más externas, en ellas la presión atmosférica es muy baja y es allí donde son absorbidos los rayos más energéticos provenientes del Sol. La mesósfera se extiende hasta los 90 km de altura aproximadamente [16] y contiene sólo el 0,1 % de la masa de la atmósfera.

En esta capa, la temperatura vuelve a descender con la altura, hasta los 90 °C bajo cero. Por encima de esta capa se sitúa la termósfera.

En la termósfera, el gradiente de temperatura es positivo, producto de la absorción de radiación de onda corta por las moléculas de O₂ y de N₂ [16]. Esta absorción da paso a reacciones de foto ionización, razón por la cual la zona entre los 50 km y los 300 km de altitud también se denomina ionosfera, debido a la presencia de iones producidos por la interacción entre la radiación solar y los elementos presentes. La ionósfera contribuye a la reflexión de las ondas de radio emitidas desde la superficie terrestre, lo que favorece que estas señales puedan viajar grandes distancias.

2.2. Química Atmosférica

La atmósfera terrestre está compuesta fundamentalmente por N₂ (78,08 %), O₂ (20,95 %) y Ar (0,93 %), en la figura 2.6 se observan los principales componentes de la atmósfera. El vapor de agua es una de las especies más abundantes, se encuentra principalmente en la baja atmósfera y su concentración es muy variable, alcanzando concentraciones del 4 % [17], el resto de los constituyentes de la atmósfera se encuentran en cantidades del orden de los ppt, abarcando menos del 1 % del total. Aquellos componentes menores al 0,0001 %, se les conoce como componentes traza, estos juegan un papel fundamental en el balance radiativo de la atmósfera y pueden participar procesos químicos que ocurren en ella, como la destrucción de la capa de ozono, el calentamiento global o episodios de contaminación a escala local [16].



Figura 2.6: Composición Química de la atmósfera terrestre Tomado de “El Cambio Climático”, por M. Molina, 2017, p. 34 [23].

Esta mezcla de especies químicas, en condiciones de presión, temperatura y exposición a diferentes niveles y tipos de radiación, pueden generar una serie de radicales. Estos radicales constituyen la base de la química atmosférica, determinando la capacidad oxidativa de la atmósfera terrestre, la cual determina los tiempos de residencia atmosféricos de las distintas especies que la componen. Muchas de estas especies son consideradas contaminantes, por sus efectos adversos en las especies y ecosistemas.

Diversas son las definiciones de contaminación, según la EPA, contaminación es: “Cualquier sustancia en el agua, el suelo o el aire que degrade la calidad natural del medio ambiente, ofenda los sentidos de la vista, el gusto o el olfato, o cause un peligro para la salud. La utilidad del recurso natural suele verse afectada por la presencia de contaminantes.” [24]. En Chile, mediante la Ley 19300 de bases generales del medioambiente se tipifica contaminación como: “la presencia en el ambiente de sustancias, elementos, energía o combinación de ellos, en concentraciones o concentraciones y permanencia superiores o inferiores, según corresponda, a las establecidas en la legislación vigente”. Sin embargo, la problemática de esta normativa es que solo mandata la fiscalización de algunos contaminantes normados en la ley, ver anexo B.3. Por lo tanto, el riesgo a incidentes ambientales que afectan negativamente la salud de las personas y el medioambiente es mayor, producto de la falta de monitoreo de contaminantes.

Los contaminantes se pueden clasificar según el origen de su emisión, en naturales y antropogénicas. Las naturales son aquellas fuentes de emisión producidas por factores biológicos o fenómenos naturales, por ejemplo, la emisión de CO₂ producto de la respiración celular o erupciones volcánicas. Por otro lado, las antropogénicas, son aquellas fuentes de emisión que se intensifican o producen por la acción humana, por ejemplo, la emisión de contaminantes por la actividad industrial y el transporte.

Otra forma de clasificar los contaminantes es en primarios y secundarios:

- Contaminantes Primarios: Son aquellos contaminantes emitidos directamente a la atmósfera por fuentes de emisión identificables. En este grupo se encuentra el monóxido de carbono (CO), el dióxido de carbono CO₂, el dióxido de nitrógeno (NO₂), dióxido de azufre (SO₂), Compuestos orgánicos volátiles (COV), el amoníaco (NH₃), etc.
- Contaminantes secundarios: Son las especies que se forman como consecuencia de los procesos químicos en los que intervienen los contaminantes primarios. Estos contaminantes ambientales pueden ser más nocivos que los emitidos directamente a la atmósfera, como el ozono troposférico (O₃) o el material particulado ultrafino (PM_{2,5-1,0}).

La lista de contaminantes atmosféricos es sumamente variada, solo en la familia de los COV se han identificado más de 1000 especies distintas en la atmósfera [17]. Esta amplia gama de contaminantes corresponde a todos los hidrocarburos que se presentan en estado gaseoso en condiciones normales de presión y temperatura o, que son muy volátiles a estas temperaturas [25]. En esta familia se incluye el (CO), el CO₂, hidrocarburos de fórmula general C_nH_m e hidrocarburos que contienen heteroátomos (Oxígeno, nitrógeno, halógenos o azufre) [26]. Su peligrosidad depende del tipo de COV, existen unos extremadamente peligrosos para la salud, como el caso del benceno, el cual en concentraciones superiores a 10000 ppm puede causar la muerte entre 5 y 10 minutos de exposición [27], por otro lado, desde la perspectiva ambiental, las preocupaciones hacia los COV se centra en su rol como precursor de ozono troposférico, el que es dañino para la salud y es el principal responsable del smog fotoquímico [17].

Dentro de los COV, el más abundante es el metano (CH₄), por su potencial de calentamiento global suele analizarse por separado. El resto de los hidrocarburos, mucho

menos presentes en la atmósfera se agrupan en lo que se conoce como hidrocarburos no metánicos o NMHCs por su acrónimo en inglés (Non Methanic Hydrocarbons) [28]. Un subgrupo de gran interés de estos COV es el de los BTEX, por las iniciales de los compuestos Benceno, Tolueno, Etilbenceno y Xileno.

La relevancia del estudio de los BTEX radica en su abundancia y sus múltiples fuentes de emisión. Los BTEX pueden corresponder hasta el 60 % de los COV no metánicos en ambientes urbanos [29]. Por otro lado, sus fuentes de emisión son variadas, son generados por el uso de combustibles fósiles o sus derivados, donde destacan la industria química con el uso de solventes, gasolineras y el tráfico vehicular, de hecho, los BTEX pueden constituir hasta el 10 % de las gasolinas sin plomo [22].

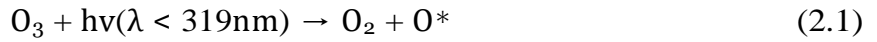
Los COV en la tropósfera participan en los procesos de producción de Ozono y otros oxidantes que forman parte del denominado smog fotoquímico. El smog fotoquímico puede definirse como la acumulación de gases y aerosoles como resultado de reacciones entre los óxidos de nitrógeno (NO_x), algunos COV y el oxígeno bajo la influencia de la radiación solar. Se observó por primera vez en los ángeles, en 1944, por su impacto a la vegetación [30]. El ozono es uno de los principales componentes del smog fotoquímico estudiado por la comunidad científica, debido a sus grandes daños asociados [31].

En la estratósfera, la emisión de COV también puede afectar la capa de ozono, ya que los COV poco reactivos pueden alcanzar las capas superiores y afectar el balance foto estacionario que regula la concentración de Ozono. Por ejemplo, el caso de los CFC, usados antiguamente en la industria de la refrigeración, son compuestos muy estables, que presentan altos tiempos de permanencia atmosféricos. Producto de lo anterior, logran alcanzar la estratosfera y por medio de reacciones de fotólisis con la radiación UV procede a destruir la capa de ozono [17]. Actualmente, el uso de este tipo de compuestos está totalmente prohibido, tras la firma del protocolo de Montreal (1987) y han sido reemplazados por aditivos como los hidroclorofluorocarbonos (HCFCs), hidrofluorocarbonos (HFCs), entre otros, los cuales poseen tiempos de permanencia atmosféricos mucho menores, por lo que no logran alcanzar la capa de ozono [17].

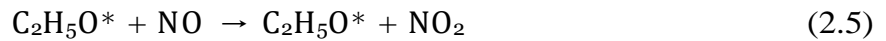
2.2.1. Química de los COV en la tropósfera

La química en la tropósfera es extremadamente compleja, ya que en ella existen una serie de especies trazas de alta reactividad, las cuales pueden participar en la foto-oxidación de las especies emitidas al ambiente. Las especies más relevantes son el radical OH, el radical nitrato NO_3 , el Cl y la molécula de (O_3) [17]. Estas corresponden a las principales especies oxidantes de la atmósfera, donde el radical OH es el principal agente oxidante durante el día, el radical NO_3 es el principal agente oxidante durante la noche, mientras que el Ozono actúa durante el día y la noche, pero con una menor reactividad que el OH y el NO_3 [32].

Los COV, inician su reacción en la atmósfera por la oxidación con el radical OH. Luego en presencia de óxidos de nitrógeno y luz solar dan paso a la formación de ozono. El radical OH se forma en la troposfera en la reacción del vapor de agua con el oxígeno excitado procedente de la fotólisis del ozono, esta reacción se observa en las ecuaciones 2.1 y 2.2.



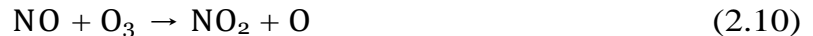
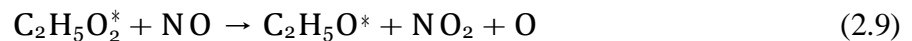
La reacción de los COV con los radicales OH genera un radical alquilo (R^*), que en presencia de oxígeno molecular (O_2), evoluciona a un radical peróxido (ROO^*). Estos radicales peróxido se transforman en radicales alcoxilo (RO^*) a través de su reacción con NO que se convierte en NO_2 . Por ejemplo, a continuación, se observa el caso del etano en las ecuaciones 2.3, 2.4 y 2.5.



Finalmente, es la fotólisis del dióxido de nitrógeno la principal responsable de la formación de ozono troposférico, mediante las ecuaciones 2.6 y 2.7 [17].



En la ecuación anterior, M es una molécula que interviene en las colisiones necesarias para excitar las moléculas, pero que no participa directamente en la formación de productos. El NO generado puede seguir diversas vías de reacción, que pueden llevar su consumo a la acumulación de NO_2 o consumo de O_3 , mediante las reacciones de las ecuaciones 2.8, 2.9 y 2.10.



Por lo tanto, el NO puede reaccionar propiciando el consumo de O_3 o la formación de NO_2 y con ello, nuevamente la generación de O_3 . Por lo tanto, en función de la composición atmosférica, el NO puede ser más afín a la reacción con el radical peróxido proveniente de la degradación de COV, favoreciendo la producción de NO_2 y con ello la formación de O_3 .

Esta relación entre la química de los COV y el NO_2 en la formación de ozono puede generar claros patrones diarios, producto de sus reacciones fotoquímicas, controladas por

el ciclo diario de radiación solar. A continuación, en la figura 2.7 se observan resultados de un estudio realizado en la ciudad de Shenzhen, China, sobre el ozono y sus precursores fotoquímicos. En la figura, se pueden observar los ciclos diarios del O_3 , NO_2 , NO y compuestos orgánicos volátiles antropogénicos (AVOC).

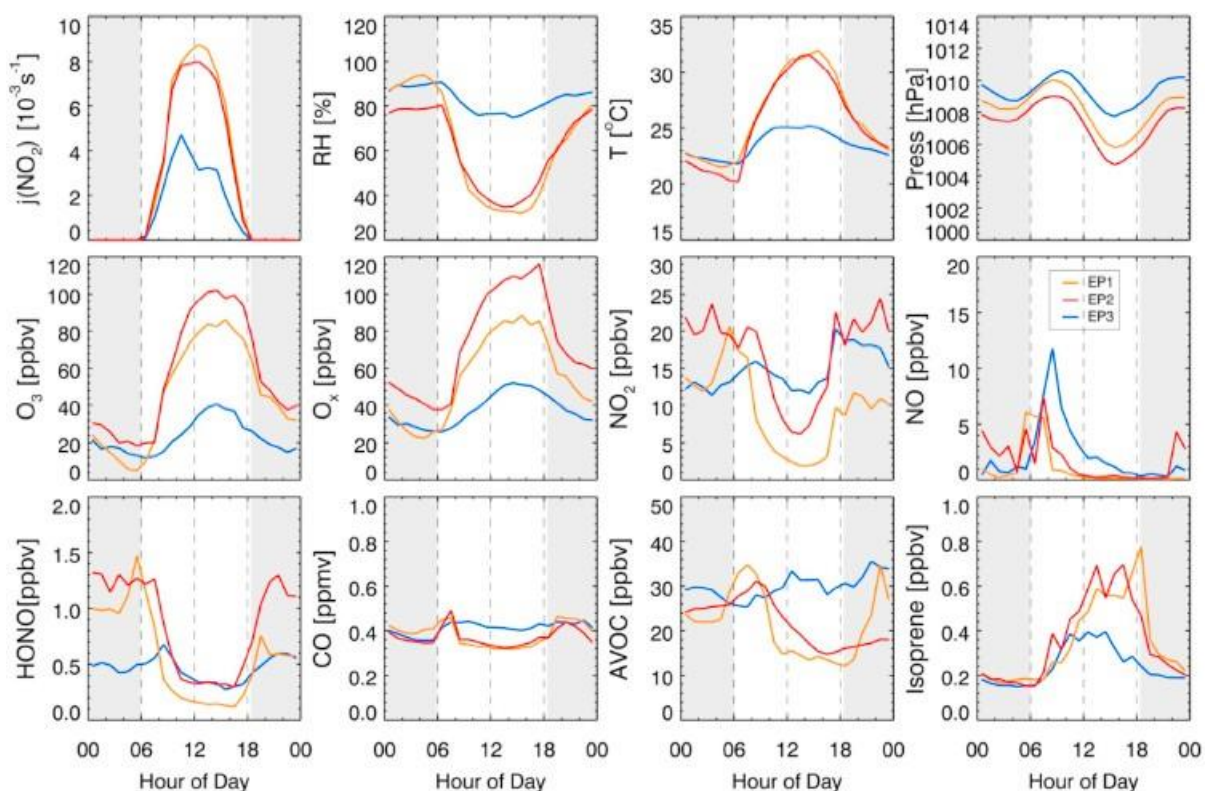


Figura 2.7: Perfiles diurnos medios de (NO_2), Humedad relativa (HR), temperatura, presión, concentraciones de O_3 , O_x , NO_2 , NO , HONO, CO observados, especies orgánicas volátiles antropogénicas (AVOC) e isopreno en el sitio SYJ durante el EP1, EP2 y EP3, respectivamente. Tomado de [33].

La figura 2.7 relata las mediciones de tres episodios, en el caso de una atmósfera contaminada con presencia importante de hidrocarburos. Se observa que los NO_x inician su disociación por efecto de la radiación solar desde el amanecer, incrementando los niveles de NO y de O . Estos últimos, posteriormente, por su combinación con el O_2 dan paso a la formación de O_3 . Conforme avanza el día, los radicales generados a partir de los COV, reaccionan con el NO , lo que induce el incremento en las concentraciones de NO_2 y de O_3 . En las horas centrales del día los niveles de NO son bajos, debido a las reacciones con los COV que provocan que los niveles de ozono se incrementen rápidamente. A medida que avanza la tarde y que los hidrocarburos se consumen, la formación de NO_2 disponible para la fotólisis es menor, lo que combinado con la disminución de la radiación solar, provoca una reducción de la concentración de ozono hasta los valores mínimos nocturnos [33].

Capítulo 3

Antecedentes sobre el monitoreo ambiental

En Chile, el seguimiento de la calidad del aire se realiza por medio de “Estaciones de Monitoreo”, estas son estructuras compuestas por diversas tecnologías, capaces de medir la concentración de múltiples contaminantes en la atmósfera. Por ejemplo, la estación de monitoreo “Parque O’Higgins” es capaz de medir concentraciones de contaminantes en tiempo real de: MP_{10} , $MP_{2,5}$, O_3 , NO_x , SO_2 y otras especies contaminantes. Estas especies, son monitoreadas mediante técnicas de medición basadas en atenuación de rayos beta, fotometría UV, quimioluminiscencia en fase gaseosa y fluorescencia pulsante respectivamente [34].

Estas técnicas de medición se conocen como metodologías de referencia, ya que aseguran mediciones confiables de calidad del aire, además, cuentan con estándares establecidos en el decreto supremo N° 61/08 del Ministerio de Salud, en cuanto a los chequeos, mantenciones, calibraciones de los equipos y al procedimiento de validación de los datos que el operador de la estación debe realizar [35]. Los sistemas de monitoreo que cuentan con metodologías de referencia para el monitoreo continuo de contaminantes tienen un alto costo de instalación, por su gran tamaño y valor de los equipos que la constituyen, por lo tanto, su aplicación en masa en múltiples puntos de monitoreo no siempre es la mejor opción [5]. Según un estudio realizado por el Centro Mario Molina a cinco redes de monitoreo de calidad del aire, para un total de 36 estaciones de monitoreo, relata que el costo de inversión y operación a 10 años es de aproximadamente 880 mil dólares [36].

Una alternativa viable o complementaria para aumentar los puntos de monitoreo de contaminantes es el uso de sensores inalámbricos de bajo costo [6]-[7], que corresponden a tecnologías portátiles y más económicas en comparación a las referenciales. De hecho, el costo de una estación de monitoreo compuesta por sensores de bajo costo es alrededor de 42 veces menor que el de una tecnología de referencia [37]. Estos sensores de bajo costo se posicionan como una alternativa viable para el seguimiento de la calidad del aire [8]. Sin embargo, el uso de sensores de bajo costo requiere de una etapa de ajuste, en la cual, las mediciones del sensor de bajo costo se deben contrastar con las de una metodología de referencia y reducir el error entre las mediciones de la metodología de bajo costo con la de referencia.

Junto a la baja resolución espacial de los sistemas de monitoreo, que no permite conocer la calidad del aire en todo el territorio, se presenta la baja resolución temporal de la información

del SINCA, que presenta actualización cada una hora de los niveles de concentración de contaminantes en los puntos de monitoreo. Por otro lado, la visualización de la información es compleja y poco amigable a la ciudadanía, presentando una serie de tiempo de los niveles de contaminación en los puntos de monitoreo, como se puede observar en la figura 3.1. Este formato de información y visualización no permite tener un conocimiento global de la calidad del aire, reduciendo la información a un solo punto, por lo cual, no es una herramienta muy eficaz para alertar y prever situaciones de riesgo.

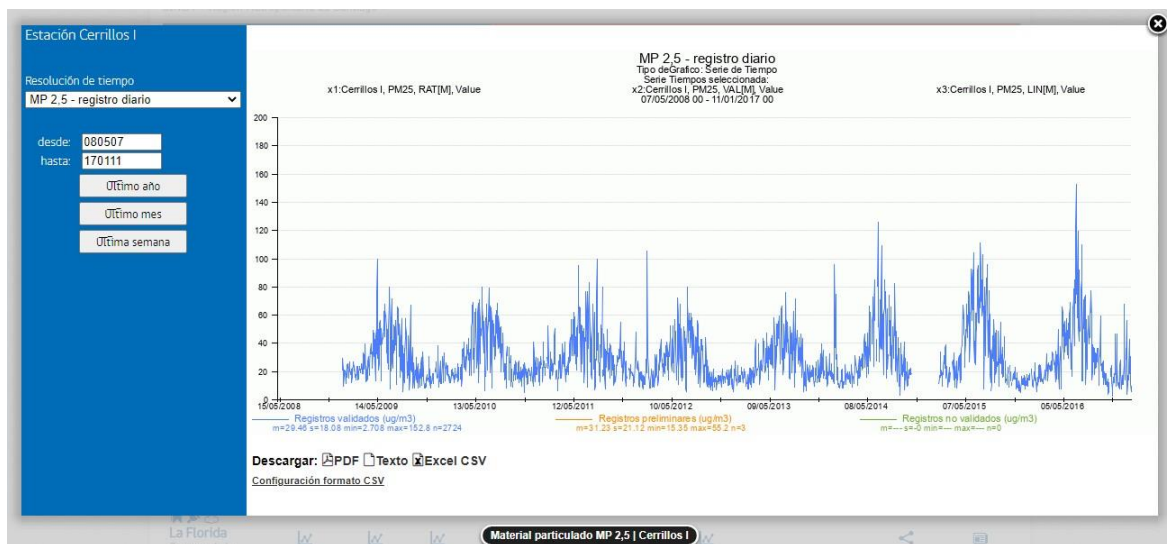


Figura 3.1: Visualización niveles de $MP_{2.5}$ en plataforma del estado [4].

Los sensores de bajo costo utilizados en este estudio corresponden a sensores de fotoionización (PID), tienen un costo aproximado entre 800 y 900 dolares. Su funcionamiento se basa en la ionización de un gas por medio de una luz ultravioleta y el par electrón-cación formado genera una diferencia de potencial proporcional a la concentración del gas en la atmósfera. En el anexo F se presenta un detalle de estos sensores.

Los sensores PID son calibrados con un equipo basado en la Espectroscopia de absorción óptica diferencial "DOAS", por sus siglas en ingles *Differential Optical Absorption Spectroscopy*. Para la calibración, los sensores PID deben ser ubicados en paralelo al equipo DOAS para la adquisición de datos por al menos 30 días, tal como se puede ver en la figura 3.2 que presenta el montaje de los equipos utilizados en este estudio.



Figura 3.2: Montaje en paralelo de Equipos DOAS y sensores de bajo costo PID para su calibración.

El DOAS es una tecnología de medición óptica que cuantifica las concentraciones de compuestos gaseosos, midiendo la absorción de la luz UV por parte de los compuestos químicos del aire, aplicando la ley de Beer-Lambert. Un equipo DOAS tiene un costo aproximado entre 60.000 y 200.000 dólares [38]. Un sistema DOAS típico consta de una fuente de luz, un receptor, un espectrómetro y, dependiendo de la configuración del sistema, un retroreflector. Su funcionamiento consiste en la emisión de una luz desde la fuente, la luz se enfoca y se dirige a la atmósfera por medio de un telescopio transmisor. Un telescopio receptor recupera y enfoca el haz de luz atenuado, y el espectrómetro mide el cambio de absorbencia causada por la luz ultravioleta. La señal digital del espectrómetro es recogida por un sistema informático y comparada con espectros de referencia.

Capítulo 4

Antecedentes sobre modelos de Dispersión

Una forma de mejorar la comunicación y completar la información espacial de calidad del aire es mediante modelos de dispersión, que permiten inferir la distribución de contaminantes en la atmósfera a partir de una serie de datos y parámetros de entrada, como se puede observar en la figura 4.1. En esta sección se discuten algunos antecedentes generales para comprender los objetivos y alcances de la modelación atmosférica. Seguidamente, se exponen ciertos de los modelos más utilizados en el análisis de la dispersión atmosférica y luego se exponen algunos de los software mas utilizados para desarrollarlos.

Posteriormente, se presenta el mecanismo de validación cruzada, que es el mecanismo más utilizado para el análisis de error entre las estimaciones y los valores observados en la modelación atmosférica. Finalmente, se presenta el estado del arte de la modelación en tiempo real, con el trabajo realizado por diversos autores.

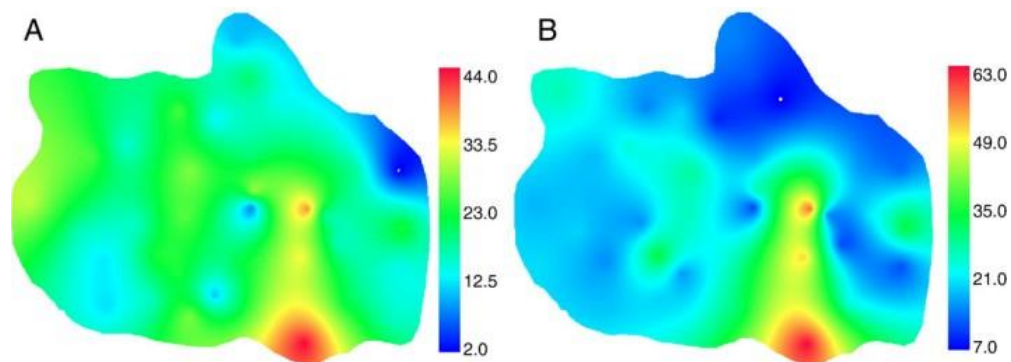


Figura 4.1: Mapas de distribución de las concentraciones ($\mu\text{g}/\text{m}^3$) de (A) BTEX y (B) NO_2 en Pamplona. Tomado de [39].

4.1. Modelos de dispersión

Los modelos de dispersión sirven para estudiar y predecir la distribución de concentración espacial de gases y/o partículas contaminantes emanados por las chimeneas industriales u otras fuentes, sean naturales o antropogénicas. Los modelos de dispersión de contaminantes son ampliamente utilizados con fines regulatorios y tienen diversos alcances, los que permiten identificar áreas de mayor impacto ambiental, identificar fuentes de emisión, predecir episodios de contaminación, elección de sitios óptimos para la instalación de redes de monitoreo y otros usos. Un buen modelo deberá predecir las concentraciones de los contaminantes en estudio para diversas condiciones atmosféricas y de emisión en la zona estudiada.

Los modelos utilizados con fines regulatorios corresponden a modelos descriptivos, que modelan los fenómenos físicos y químicos de los contaminantes en la atmósfera, lo que permite analizar el alcance de los contaminantes emitidos por diversas fuentes y en distintas condiciones meteorológicas. Por lo anterior, son ampliamente utilizados en el análisis de impacto ambiental en la instalación de grandes fuentes de emisión. Por ejemplo, en Chile, los proyectos que deben someterse a estudios de impacto ambiental, tienen que realizar un modelo de dispersión para estimar el impacto de sus emisiones, para esto el servicio de evaluación ambiental (SEA) elaboro la guía para el uso de modelos de calidad del aire en el SEIA, que reconoce cuatro tipos de modelos para estos fines, los modelos gaussianos, eulerianos, lagrangeanos y modelos tipo puff, estos modelos requieren gran cantidad de datos meteorológicos, topográficos y de fuentes de emisión [40].

Cada modelo tiene diversos requerimientos y características según la escala del sistema, que puede considerar variables como temperatura, dirección y velocidad de viento, condiciones meteorológicas, topografía, etc [41]. Producto de las múltiples escalas espaciales y temporales de la atmósfera, los modelos atmosféricos pueden clasificarse en estas escalas de movimiento atmosférico. Esto es importante, ya que dependiendo de la escala meteorológica a estudiar, se puede llegar a despreciar el efecto generado por la difusión, la turbulencia o la advección [42].

Por las dimensiones del problema, los modelos de escala global presentan órdenes temporales de horas o días, por lo tanto, no suelen considerarse especies con bajos tiempos de residencia atmosféricos. Por esta razón, en general, este tipo de modelos son más utilizados para predecir el cambio de las condiciones meteorológicas, que para el transporte de contaminantes. Sin embargo, actualmente es reconocido que especies químicas con bajos tiempos de residencia atmosféricos, particularmente aquellas asociadas a la formación de aerosoles, pueden afectar el clima, cambiando el balance energético terrestre e incidiendo en la formación de nubes [43]. Por otro lado, los de menor escala, en general buscan representar el comportamiento local de la atmósfera, permitiendo predecir la evolución de contaminantes o especies químicas presentes en las masas de aire de un lugar determinado y se conocen típicamente como modelos de calidad del aire.

Los modelos de dispersión en general responden a dos objetivos: el primero, corresponde a los que buscan determinar la ubicación y características de una fuente contaminante (muchas veces, desconocida); el segundo, corresponde a los modelos que buscan representar

la evolución de contaminantes en la atmósfera, predecir y conocer la distribución de concentración en el espacio. Estos últimos, en general basados en entradas de datos meteorológicos e información de fuentes contaminantes, como las tasas de emisión y la altura de la chimenea, están diseñados para caracterizar los contaminantes primarios que se emiten directamente a la atmósfera y, en algunos casos, los contaminantes secundarios que se forman como resultado de reacciones químicas en la [44].

Los modelos utilizados para determinar la ubicación de una fuente de emisión corresponden a problemas de modelado inverso, en los cuales se calculan los parámetros de la fuente de emisión a partir de la medición de sensores. El problema de modelado inverso es altamente no lineal [45] y sujeto a datos de entrada que suelen ser esporádicos, ruidosos y escasos [46]. Tradicionalmente, se utiliza una red de sensores estáticos en el suelo para estimar la ubicación y características de la fuente, como se ilustra en la Figura 4.2.

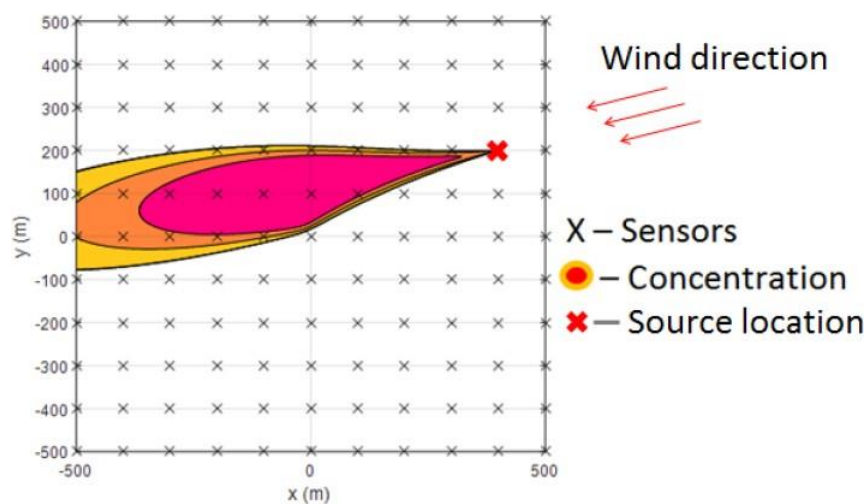


Figura 4.2: Esquema localización de la fuente contaminante mediante sensores estáticos. Tomado de [47].

Estos modelos requieren gran cantidad de datos meteorológicos y su desarrollo tiene un alto costo computacional, ya que debe dar solución de manera iterativa a las ecuaciones que describen su comportamiento. El desafío de estos modelos, es evolucionar de modelos descriptivos de la dispersión de contaminantes ante ciertos escenarios, a modelos que permitan representar la dispersión espacial de contaminantes sincrónicamente.

Otra forma de categorizar los modelos existentes es según su formulación, permitiendo dividir en dos grandes grupos. El primero, corresponde a los modelos de dispersión físicos, que responden a la fenomenología a la que se ven expuestos las especies en el ambiente, mediante la resolución de las ecuaciones de continuidad y dinámica de fluidos. El segundo, corresponde a los modelos matemáticos o numéricos, que analizan la correlación existente entre los datos registrados para interpolar una distribución de concentraciones. A continuación, se profundizará en este tipo de métodos.

4.1.1. Modelos de dispersión físicos

Estos modelos buscan relatar los fenómenos físicos y químicos en los que participan las partículas o gases en la atmósfera. Es decir, las reacciones y los fenómenos de transporte en los que están involucrados, controlados por la acción de la difusión molecular, turbulencia y advección. Para lo anterior, los modelos físicos dan resolución al balance de masa de una especie en una unidad de volumen del espacio mediante la ecuación 4.1, que relata la acción de los fenómenos involucrados en el transporte, consumo y producción de un contaminante en la atmósfera [16].

$$\frac{\partial C_i}{\partial t} + \frac{\partial}{\partial x_j} (u_j C_j) = \frac{\partial^2}{\partial x_j^2} + (R_i C_i, \dots, C_n, T) + S(x, t) \quad (4.1)$$

En la ecuación 4.1, el primer término representa la variación de concentración del contaminante i en el tiempo; el segundo, la entrada neta de contaminante asociada al flujo atmosférico (advección y turbulencia); el tercero, la variación de la concentración de contaminante debido a la difusión molecular; el cuarto, la velocidad de desaparición del contaminante por reacción química; y el quinto, la aparición de contaminante en el entorno (generación in situ). El problema es que la resolución analítica de estas ecuaciones no es posible, puesto que no se conoce la forma de la función $(u_j C_j)$ ni de la difusión turbulenta asociada. Además, hay que tener en cuenta que es necesario resolver las N ecuaciones diferenciales simultáneamente. Producto de esta complejidad, actualmente, existen diversos modelos ampliamente desarrollados y validados por la comunidad científica que presentan ciertas simplificaciones a la ecuación anterior y logran representar el comportamiento de los agentes atmosféricos adecuadamente.

Los modelos de dispersión físicos de contaminantes se pueden dividir en cinco tipos principales: los modelos de caja [48], modelos de pluma gaussiana [49], modelos lagrangianos [50], modelos de dispersión euleriana [51] y modelos de gas denso [52]. En el anexo D se encuentra un detalle de los más relevantes. Cada uno de estos modelos presenta distintas funcionalidades, simplificaciones o alternativas para desarrollar las ecuaciones que representan la fenomenología fisicoquímica. La elección sobre qué modelo utilizar depende de la escala y características del sistema atmosférico a modelar, los fenómenos fisicoquímicos, disponibilidad de datos, entre otros.

Por ejemplo, los modelos de caja son utilizados cuando se modela un sistema con una fuente difusa de contaminación y se realiza un balance de masa sobre una caja que englobe todo el sistema, por otro lado, el modelo de pluma gaussiana, que es el más utilizado con fines regulatorios sirve para modelar sistemas con fuentes puntuales de emisión y plantea simplificaciones a términos como el factor que representa la difusión molecular. Por lo tanto, para la correcta elección del modelo se debe tener claro el problema que se espera resolver, si se desea identificar una fuente puntual de emisión, o evaluar el impacto que genera, o hacia donde se transportaran los contaminantes generados por una industria, el modelo de pluma gaussiana sería el más apropiado, por otro lado, si se quiere modelar una fuentes de emisión difusa, cuando la emisión se genera por un área o volumen, entonces el modelo de caja sería una alternativa viable.

La principal ventaja de estos modelos es la amplia validación científica en torno a sus

resultados, de hecho, los modelos de dispersión recomendados por el SEA corresponden a modelos de dispersión físicos. Su desventaja es el alto costo computacional para desarrollarlos, producto de la gran cantidad de datos requeridos para implementarlos y la resolución múltiples sistemas de ecuaciones. Por lo tanto, la implementación de un modelo de dispersión físico en tiempo real requeriría gran capacidad de cómputo, múltiples estaciones de monitoreo de contaminantes y meteorológicas desplegadas.

4.1.2. Modelos de dispersión matemáticos

Los modelos de dispersión matemáticos corresponden a métodos basados en interpolaciones, se sustentan en la correlación existente entre los datos medidos para efectuar la interpolación de otros puntos. El factor que representa esto es el coeficiente de correlación W_{pq} e indica el peso de interpolación o correlación existente entre los puntos p y q. Existen varios tipos de técnicas de interpolación, todas diferentes en cuanto a los puntos que se consideran y cómo se determinan los pesos de interpolación o coeficientes de correlación.

En general, los puntos cercanos obtienen más peso que los puntos lejanos, tal y como apoya la primera ley de geografía de Tobler, que plantea que "Todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las lejanas" [53]. Es decir, la concentración o el valor de un parámetro en un punto $P_{x,y,z}$, debe ser muy similar al valor de la concentración o parámetro en un punto $P_{x+dx,y+dy,z+dz}$ muy cercano al punto inicial.

El costo computacional de estos métodos es menor que aquellos involucran la resolución de sus ecuaciones fundamentales de transporte, sin embargo, la calidad de los métodos de interpolación está condicionados por el número y distribución espacial de los puntos de muestreo. Los métodos de interpolación se pueden clasificar según la metodología de cálculo de los coeficientes de correlación, que puede ser de manera estocástica o determinista. A continuación, se presenta un detalle de los métodos de interpolación comúnmente utilizados en el mapeo de la dispersión atmosférica.

4.1.2.1. Métodos estocásticos

Se basan en los principios de la geoestadística, que es la aplicación de la teoría de las variables regionalizadas. De manera general, diremos que un fenómeno es regionalizado cuando se desplaza en el espacio manifestando una cierta estructura. Si $f(x)$ designa el valor en el punto "x" de una característica f de este fenómeno, diremos que $f(x)$ es una variable regionalizada (V.R) [54]. Entonces, del punto de vista matemático, una V.R es simplemente una función $f(x)$ del punto "x", que en general es una función muy irregular, por ejemplo, la ley de un mineral en una veta, la porosidad y permeabilidad de la roca en un reservorio, etc. Se caracteriza por presentar cierta continuidad espacial y temporal y, por variar irregularmente, como se observa en la figura 4.3.

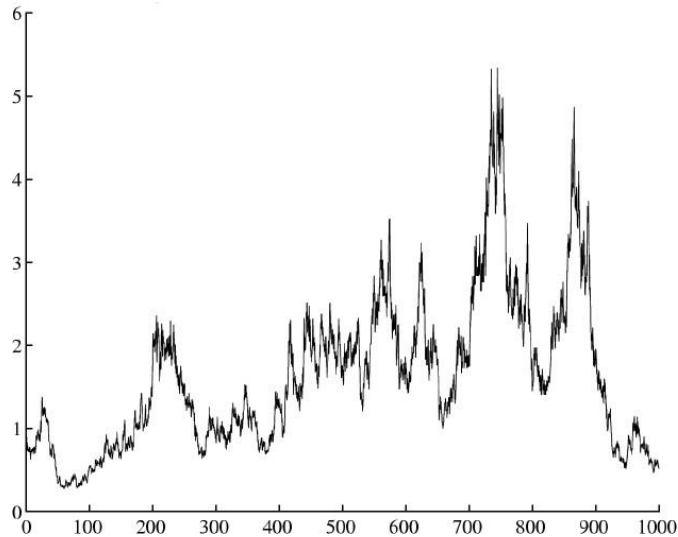


Figura 4.3: Variable regionalizada en un espacio unidimensional. Por ejemplo, podría representar la concentración de un contaminante medida a lo largo del tiempo en una estación de monitoreo. Tomado de [55].

El kriging es el método estocástico por excelencia, fue desarrollada por primera vez por el matemático Georges Matheron en 1951. El método de interpolación recibió el nombre de kriging, en honor al estudiante Danie Krige que elaboró el método durante su investigación de maestría [56]. La idea del kriging es calcular la disimilitud entre pares de puntos de datos y utilizar esta relación para estimar los pesos de interpolación. Las disimilitudes entre los puntos se utilizan para cuantificar la dependencia espacial de los puntos en un llamado variograma.

Podemos definir el variograma como la media de los cuadrados de las diferencias entre pares de las muestras separadas por una distancia h . A continuación, en la ecuación 4.2 se observa la formula del variograma. En el anexo 4 se profundiza en esta metodología.

$$Y = \frac{1}{2N} \sum (Z(x) - Z(x + h))^2 \quad (4.2)$$

En la figura D.3 se observa un ejemplo de variograma, este informa que tanto varia la influencia de un punto a medida que me alejo de él. En el eje “y” se observa la covarianza entre la estimación y la medición, el eje “x” corresponde a la distancia respecto al punto de interpolación. Nos permite conocer la distancia a la cual ya no existe correlación entre los datos, este valor corresponde al rango “*range*” y se alcanza cuando la curva del variograma alcanza el silo “*Sill*”, punto en el que se alcanza la mínima covarianza. El kriging crea un variograma teórico y experimental, los cuales se utilizan para predecir los pesos en lugares no observados, de forma que se minimice la varianza de la estimación [54].

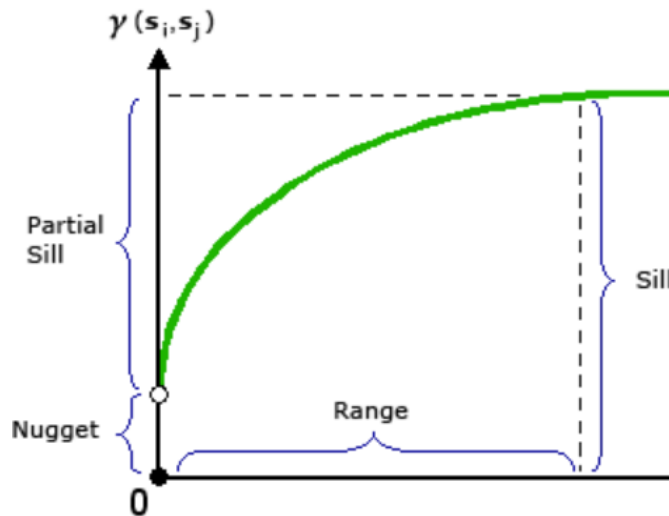


Figura 4.4: Variograma [57].

El kriging realiza una combinación lineal de todos los puntos muestreados, ponderados por los coeficientes de correlación que son calculados mediante la minimización de la varianza entre los puntos muestreados y los interpolados. De esta forma, el factor de interpolación es función del variograma, a diferencia de los métodos deterministas donde la interpolación es función de características geométricas, como la distancia entre los puntos en el método de interpolación de distancia inversas. En el kriging, la forma de variación dependerá de la forma del variograma, si el variograma es muy suave, entonces la variabilidad espacial es muy suave, por lo tanto, el mapa tendrá una forma muy distinta de cuando la variabilidad espacial es muy abrupta.

El cálculo de las ponderaciones de kriging puede ser un proceso muy exigente desde el punto de vista informático cuando se trata de grandes conjuntos de datos. La creación del variograma y la inversión de grandes matrices durante el kriging pueden ser procesos que consumen mucho tiempo. La forma habitual de superar este problema es incluir localmente sólo un subconjunto de puntos cercanos en el proceso de kriging [58].

4.1.2.2. Métodos deterministas

Los métodos deterministas crean superficies a partir de los puntos medidos y basándose en la extensión de estas superficies se calculan los coeficientes de correlación. La interpolación por distancia inversa y de vecinos naturales corresponden a este tipo de métodos. A continuación, se relatan las principales características de estos métodos de interpolación.

4.1.2.2.1. Interpolación de distancia Inversa Ponderada

Es uno de los métodos deterministas más simples, asume que los valores más cercanos están más relacionados que los lejanos. El valor de una variable Z_p en un punto “p” desconocido se calcula en función de la distancia entre los puntos p y q, mediante la ecuación 4.3. Es bueno cuando sus datos son densos y están espaciados uniformemente.

$$Z_p = \frac{\sum_{i=1}^n (z_q W_{q,p})}{\sum_{i=1}^n (W_{q,p})} \quad (4.3)$$

$$W_{q,p} = \frac{1}{d_{q,p}} \quad (4.4)$$

En la ecuación 4.3, z_q representa el valor de una variable en un punto q conocido, $W_{q,p}$ representa el coeficiente de correlación entre los puntos (q,p) y $d_{q,p}$, representa la distancia entre los puntos (q,p) . Por ejemplo, en la figura 4.5 se observa un ejemplo, en el que se cuenta con un set de valores conocidos de una variables en determinados puntos y se desconoce el valor de la variable en el punto “?” desconocido.

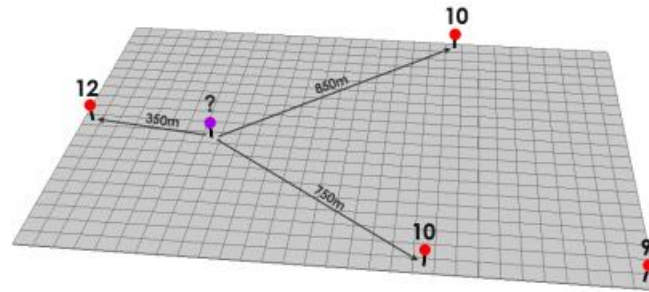


Figura 4.5: Ejemplo Interpolación por distancia inversa ponderada. Tomada de [59].

Aplicando la ecuación 4.3 para el set de puntos conocidos $q = 1, 2, 3$, con valores z_q iguales a 12, 10 y 10 unidades respectivamente. Las distancias al punto de interpolación con valores $d_{q,p}$ iguales a 350, 750 y 850 metros respectivamente. Se obtiene que el valor de la variable en el punto “?” es igual a 11,1 unidades, como se puede observar en la ecuación 4.5.

$$Z_p = \frac{\frac{12}{350} + \frac{10}{750} + \frac{10}{850}}{\frac{1}{350} + \frac{1}{750} + \frac{1}{850}} = 11,1 \quad (4.5)$$

4.1.2.2.2. Interpolación de Vecinos Naturales

A diferencia del método de interpolación de distancia inversa, el método de vecinos naturales (V.N) realiza la interpolación en función de un área de representación asociada a cada punto de muestreo. Se sustenta en una división geométrica del plano conocida como teselación o diagrama de Voronoi. Los diagramas de Voronoi han sido utilizados en diversas ramas de la ciencia, el primero en utilizarlas en estudios en torno a la interpolación espacial fue Thiessen (1991) que utilizó las regiones de Voronoi como ayuda para calcular estimaciones más precisas de los promedios regionales de reincidencia de lluvias [60].

El diagrama de Voronoi divide el plano en N celdas, una para cada nodo, que corresponden a las ubicaciones (x_i, y_j) de los puntos de monitoreo. Las celdas se construye en función de la distancia euclidiana entre los nodos y los demás puntos del plano. La distancia euclidiana de un punto a su nodo generador es menor a la distancia euclidiana a cualquier otro nodo del plano. De esta forma, los puntos de una celda son los mas cercanos a la ubicación del nodo

generador. En el anexo A, se presenta un detalle de las propiedades del diagrama de Voronoi y sus aplicaciones.

El método de interpolación de los vecinos naturales consta de dos pasos. En primer lugar, se crea un diagrama de Voronoi de los puntos observados, generando una celda para cada ubicación de monitoreo, estas celdas de Voronoi definen todas las ubicaciones más cercanas a ese punto de monitoreo. En segundo lugar, se crea una nueva partición de Voronoi, considerando los puntos de monitoreo más el punto de estimación. Se crea una nueva partición por cada punto de estimación establecido. El efecto de este nuevo punto en la partición original de los puntos de monitoreo, se utiliza para calcular los coeficientes de correlación entre los puntos de monitoreo y el punto de estimación.

El cálculo de los coeficientes de correlación para el punto de interpolación q , se realiza observando todas las celdas de Voronoi V_p de los puntos de observación que se ven alteradas la introducir el nuevo punto de estimación. El peso de interpolación de cada punto de monitoreo, es el área que se intersecta de su celda de Voronoi dividida por el área total de la celda de Voronoi generada por el punto que se desea interpolar [61], como se observa en la figura 4.6.

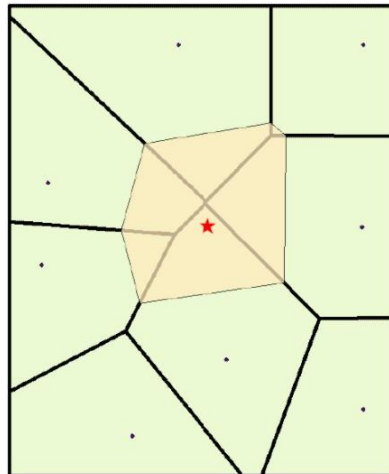


Figura 4.6: Polígono de Voronoi en punto de interpolación sobre el diagrama de Voronoi inicial de los puntos de monitoreo. Tomado de [62].

De acuerdo con este método, inicialmente, para un conjunto dado de puntos de muestreo se construye el diagrama de Voronoi original; posteriormente, el punto de interpolación objetivo “ q ” se inserta en el diagrama, alterando el diagrama de Voronoi original. El punto “ q ” en el nuevo diagrama de Voronoi está asociado con la región V_q y los pesos de interpolación o coeficiente de correlación $W_{p,q}$ se calculan en función de las áreas de las regiones generadas [63]. A continuación, se presenta la ecuación 4.6 para el calculo de los coeficientes de correlación en la interpolación de los vecinos naturales y la ecuación 4.7 para el calculo del valor en el punto de interpolación q .

$$W_{p,q} = \frac{\text{Área}(V_q \cap V_p)}{\text{Área}(V_q)} \quad (4.6)$$

$$C_q = \sum_{p=1}^n W_{p,q} * C_p \quad (4.7)$$

Por ejemplo, si se cuenta con 2 sensores de gases en las ubicaciones C_1 y C_2 , con coordenadas en $(x_1, y_1) = (1, 2)$ y $(x_2, y_2) = (4, 2)$. Primero, se genera un diagrama de Voronoi para las ubicaciones de los puntos de monitoreo C_1 y C_2 . La partición de Voronoi generada se observa en la figura 4.7, en la figura se observa claramente la principal propiedad de los D.V, que indica que todos los puntos de una celda son más cercanos a su origen y solo los puntos en las aristas que separan dos celdas son equidistantes a dos nodos (C_1 y C_2).

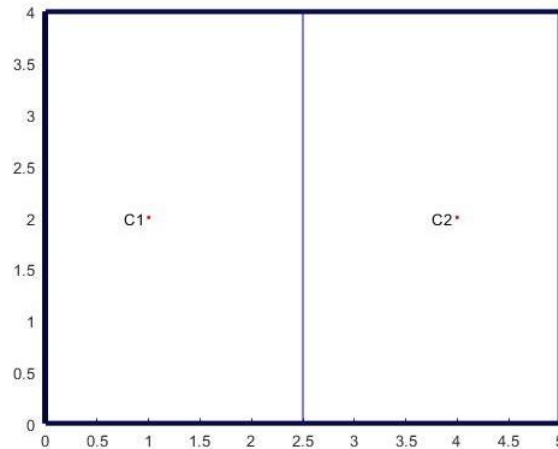


Figura 4.7: Diagrama de Voronoi dos puntos. Elaboración propia.

Luego, si se interpola un nuevo punto en la ubicación $(x_3, y_3) = (3, 3)$, se debe generar un nuevo diagrama de Voronoi incluyendo dicho punto, como se observa en la figura 4.8

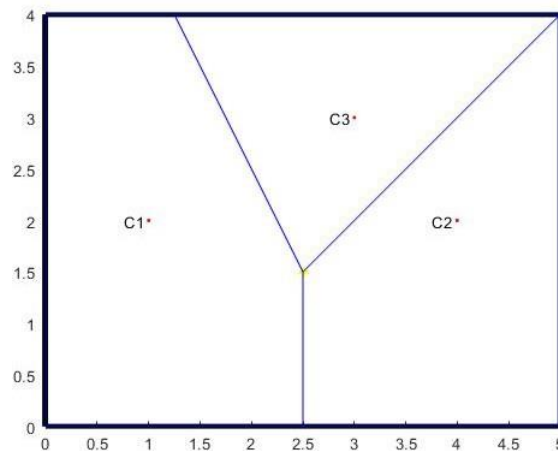


Figura 4.8: Diagrama de Voronoi incluyendo el punto de interpolación. Elaboración propia.

Luego el valor el punto C_3 , depende de los coeficientes de correlación $W_{1,3}$ y $W_{2,3}$, los cuales se calculan mediante la fracción del área de la nueva celda de Voronoi generada por el punto C_3 sobre las celdas C_1 y C_2 , como se observa a continuación en la figura 4.9.

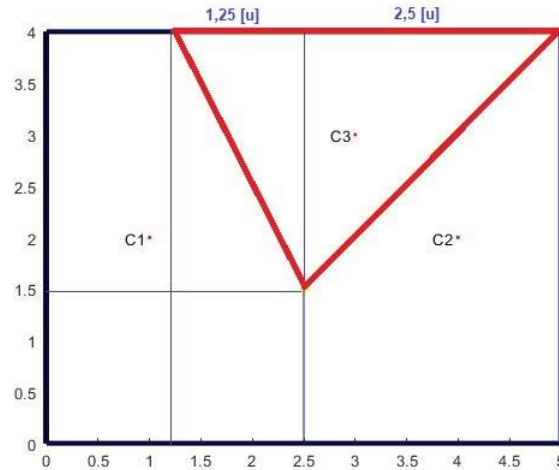


Figura 4.9: Intersección celda de Voronoi en DV inicial. Elaboración propia.

En la figura 4.9, se puede observar que el área de la celda de Voronoi generada por C_3 , corresponde al área de un triángulo de base 3,75 [unidades] y altura 2,5 [unidades] aproximadamente, por lo tanto, el área de la celda C_3 es aproximadamente 4,69 [unidades²], como se observa en la ecuación 4.8.

$$\text{Área}C_3 = \frac{3,75 * 2,5}{2} = 4,687 \quad (4.8)$$

Luego, se analiza la fracción del área de C_3 sobre las celdas iniciales C_1 y C_2 , mediante las ecuaciones 4.9 y 4.10 respectivamente. Se observa que se generan 2 nuevos triángulos rectángulos, uno de base 1,25 [unidades] sobre la celda C_1 y otro, de base 2,5 [unidades] sobre la celda C_2 . El triángulo sobre la celda C_1 posee un área aproximada de 1,56 [unidades] y el triángulo sobre la celda C_2 de 3,125 [unidades], sumando la totalidad del área de la celda C_3 .

$$\text{Área}(C_3 \cap C_1) = \frac{1,25 * 2,5}{2} = 1,56 \quad (4.9)$$

$$\text{Área}(C_3 \cap C_2) = \frac{2,75 * 2,5}{2} = 3,125 \quad (4.10)$$

Luego, como se ya menciono, los coeficientes de correlación $W_{1,3}$ y $W_{2,3}$ de los puntos 1 y 2 para el punto de interpolación 3 respectivamente. Se calculan mediante la proporción del área de C_3 sobre C_1 y C_2 , como se observa en las ecuaciones 4.11 y 4.12 respectivamente. Es claro que los coeficientes solo pueden tomar valores en el intervalo [0,1].

$$W_{1,3} = \frac{1,56}{4,69} = 0,34 \quad (4.11)$$

$$W_{2,3} = \frac{3,125}{4,69} = 0,66 \quad (4.12)$$

De esta forma, se calculan los coeficientes de correlación mediante el método de los vecinos naturales, cabe destacar, que los coeficientes dependen solo de las ubicaciones iniciales de los equipos de monitoreo. Luego para realizar la interpolación, solo se debe conocer los valores de las mediciones en los puntos C_1 y C_2 y la concentración en el punto C_3 queda determinada por la ecuación 4.13.

$$C_3 = 0,34 * C_1 + 0,66 * C_2 \quad (4.13)$$

Una propiedad favorable de la interpolación del vecino natural es que se adapta a la anisotropía de la geometría de los datos, es decir, datos que varían según la dirección en que son examinadas. Algo anisótropo podrá presentar diferentes características según la dirección, incluso para este tipo de condiciones se obtienen distribuciones de pesos realistas. Por lo tanto, la interpolación del vecino natural funciona bien para diversas distribuciones geométricas de los datos de entrada y no se ve afectada en mayor medida por no considerar los efectos de la dirección en los puntos de interpolación. Este aspecto es de gran importancia en el modelamiento atmosférico, ya que este es dominado por la magnitud y dirección del viento.

Pese a lo anterior, una desventaja del método es que sigue siendo determinista, por lo tanto, no logra representar condiciones mayores o menores a los datos muestreados. Los resultados pueden ser buenos en términos de distribución geométrica de pesos, pero la relación espacial real dentro de los datos no se capta ni se tiene en cuenta durante la interpolación [58]. Por lo tanto, en condiciones en que las variables en estudio cambien abruptamente en el espacio y/o tiempo, será necesario un mayor número de puntos de muestreo para lograr representar el cambio de las variables de estudio en el espacio.

El uso del método de interpolación de vecinos naturales tiene muchas características beneficiosas, se puede utilizar tanto para interpolación como para la extrapolación y funciona bien para un conjunto de ubicaciones aleatorias de monitoreo [64]. Presenta un algoritmo de interpolación rápido y es un buen método para interpolar grandes conjuntos de datos, porque sólo se utiliza un conjunto de puntos vecinos para el cálculo de los pesos. Tiene las desventajas de que la función de interpolación se limita a la zona acotada por el conjunto de puntos monitoreo y la superficie resultante no puede ser aceptada en muchos campos como la geología o la hidrología [64].

En la tabla 4.2, se observa una comparación de las técnicas de modelación analizadas en este capítulo.

Tabla 4.1: Comparación Modelos de dispersión. Elaboración propia.

Tipo de Modelo	Técnica de Calculo	Principales aplicaciones	Ventajas	Desventajas
Dispersión de Gauss	Resuelve el balance de masa del componente y plantea una simplificación al factor de difusión molecular	Usos regulatorios de impacto ambiental de fuentes de emisión fijas	Amplia validación científica	Alto costo computacional y N ^o de datos de entrada
Kriging	Interpola valores ajustando un variograma teórico y experimental mediante la minimización de la varianza entre valores conocidos y estimados	Calculo de ley de mineral en depósitos mineros	Amplia validación científica	Inversión de múltiples matrices en la optimización
I.D.W	Interpola proporcionalmente a la distancia al punto	Interpolar todo tipo de datos ubícales en un SIG	Método rápido y simple	Poco robusto y costoso cuando considera todos los puntos del plano
V.N	Interpola proporcionalmente, al área de Voronoi generada por el punto de interés sobre un DV inicial	Interpolación espacial, análisis de áreas de influencia, vecinos mas cercanos a un punto.	Método rápido y simple	Requiere gran cantidad de puntos de muestreo

4.2. Software tradicionales

Actualmente, existen diversas herramientas computacionales que permiten modelar una amplia variedad de sistemas. Entre ellas se pueden encontrar programas específicos, que presentan funcionalidades o módulos predefinidos que facilitan la labor de las y los desarrolladores. Por otro lado, se pueden encontrar otros programas no tan desarrollados, que requieren mayor programación por parte del desarrollador para generar tales módulos atmosféricos de transporte, fotoquímicos, entre otros. Además, muchas de estas herramientas implican altos costos monetarios para acceder a sus licencias. En esta sección se presentarán algunas de las principales herramientas utilizadas en el desarrollo de modelos de dispersión de contaminantes.

De acuerdo con la EPA [65], dependiendo del nivel de sofisticación de la modelación se realiza la selección de los modelos. Se definen dos niveles de sofisticación, el primero de ellos consiste en una técnica que de manera simple evalúa el peor escenario meteorológico para identificar las concentraciones máximas posibles, esta técnica es conocida como modelos *screening*. Si las concentraciones determinadas por la técnica *screening* exceden las concentraciones de las normas de calidad del aire, se pasa al siguiente nivel de sofisticación en donde se debe hacer uso de modelos de calidad del aire refinados.

En las figuras 4.10 y 4.11, se observa un detalle de los principales software recomendados por diversos organismos para la modelación de calidad del aire. Se observa que los más populares corresponden al software AERMOD, CALPUFF y SCREEN3, estos programas se sustentan en modelos de dispersión físicos.

Ref	Screen	Refinado	Fuentes móviles	Fumigación costera	Grandes distancias	Fotoquímico	Entidad	País	Año
1	CTSCREEN	AERMOD	CALINE3	OC	CALPUFF	CMAQ CAMx REMSAD UAM-V	US Environmental Protection Agency	USA	2015
2	AERMOD, ISC3, SCREEN3, CALPUFF, CALSCREEN						Environmental Health Department	Albuquerque, New Mexico, USA	2008
3	CDM/CDMQC, AQDM, TCM, RAM, HIWAY						ADE	Arizona, USA	1979
5	SCREEN3 CTSCREEN VISCREEN	ISCST3 CTDMPPLUS	AERMOD, CALPUFF				Idaho Department of Environmental Quality	Idaho, USA	2002
6	AERMOD						Department of Natural Resources	Iowa, USA	2008
7	AERSCREEN	AERMOD, CALPUFF					Montana Department of Environmental Quality	Montana, USA	2007
8	SCREEN3 AERSCREEN CTSCREEN OZIPR CAL3QHC/MOBILE6 VISCREEN TSCREEN	AERMOD ISC3	CAL3QHC/MOBILE6	SDM OC	CALPUFF FLAG	CMAQ UVM	New York State Department of Environmental Conservation	New York, USA	2006
9	SCREEN3	AERMOD, CALPUFF					Oklahoma Department of Environmental Quality	Oklahoma, USA	2006
10	AERMOD, SCREEN3, ISCST3, ISCST3-PRIME						Utah Division of Air Quality	Utah, USA	2008

Figura 4.10: Software recomendados por diversas guías de modelación (a). Tomado de [66].

Ref	Screen	Refinado	Fuentes móviles	Fumigación costera	Grandes distancias	Fotoquímico	Entidad	País	Año
11	SCREEN3 ISC-PRIME AERMOD CALPUFF	ISC PRIME ISC_OLM AERMOD RTDM CTDPLUS CALPUFF					Alberta Environment	Alberta, Canada	2003
12	NO DEFINE MODELOS						Ministry of Environment	British Columbia, Canada	2006
13	SCREEN3	AERMOD ISC3/ISCPRIME	CAL3QHCR	SDM	CALPUFF		Ontario Ministry of the Environment	Ontario, Canada	2005
14			CALINE4 ISCST3	FDM			Environmental Department	Protection Hong Kong	2006
15	CUENTAN CON UNA BASE DE DATOS DE 125 MODELOS						Forum for Air Quality Modelling in Europe	Europa	2006
16	AUSPLUME, ISCST3, AERMOD, CTDPLUS, CALPUFF						Ministry for the Environment	Nueva Zelanda	2004
17	AUSPLUME, AERMOD, CALPUFF, ISCST3						Department of Environment	Australia	2006
18				OCD			Comisión Nacional de Energía Atómica	Argentina	1997
19	SCREEN	ISCLT					Instituto colombiano de normas técnicas y certificación - ICOTEC	Colombia	1998

Figura 4.11: Software recomendados por diversas guías de modelación (b).
Tomado de [66].

AERMOD es un programa desarrollado por la Agencia de Protección Ambiental de los Estados Unidos. Resuelve un modelo físico de pluma gaussiano de estado estacionario, que asume que las concentraciones en todas las distancias están gobernadas por la meteorología promedio de una hora. Las ventajas de AERMOD radican en que permite ingresar archivos de terreno y datos meteorológicos completos, los cuales influyen considerablemente en la dispersión de contaminantes. La desventaja es que representan una predicción de la realidad, por lo que no serán totalmente exactos y si los datos de entrada son escasos o deficientes, entonces los resultados no serán totalmente representativos.

CALPUFF es un sistema avanzado de modelización meteorológica y de calidad del aire en estado no estable, es un programa desarrollado por científicos de Exponent. Resuelve un modelo de dispersión físico lagrangeano. La ventaja de CALPUFF es que se presenta como una excelente alternativa para modelar sistemas complejos de terreno o turbulencia, así como condiciones atmosféricas con bajas de velocidades de viento, donde los modelos gaussianos no se comportan tan bien.

AERMOD y CALPUFF son modelos bastante similares, que pueden ofrecer varios niveles de complejidad, pero solo son aplicables a casos particulares, ya que no permite al desarrollador acceder al código y realizar cambios según sus requisitos. Un problema particular al aplicar uno de estos modelos es que no permiten responder rápidamente al aumento de la concentración de contaminantes en la atmósfera, en particular en situaciones accidentales, es necesario que los modelos de difusión sean parte integral de los sistemas de medición automatizados que se instalan en la zona de influencia, para que pueda proporcionar campos de concentración en tiempo real [67]. Por lo tanto, estas herramientas no se posicionan como la mejor alternativa para el análisis de la dispersión de contaminantes en tiempo real, el desafío de estos programas es la integración en tiempo real, de manera de observar las variaciones de la dispersión de contaminantes a medida que varían los flujos de viento y las emisiones de contaminantes.

Airviro es un sistema basado en la web de última generación para la gestión de la calidad

del aire. Es un sistema integrado para el manejo de datos de series de tiempo, inventarios de emisiones y modelado de dispersión. Se ha desarrollado continuamente desde 1990 con miles de usuarios en todo el mundo [68]. Airviro puede controlar los principales aspectos en la evaluación y gestión de la calidad del aire: Monitoreo, Inventarios de emisiones y Modelado de dispersión.

Airviro es adecuado para manejar grandes cantidades de datos y usuarios. Satisface las altas exigencias de seguridad, rendimiento y estabilidad. Airviro utiliza un diseño modular que le permite adaptar el sistema a las necesidades del usuario, además, ofrece alojamiento de la información en sus servidores y se encarga de las copias de seguridad, administración de sistema y actualizaciones del servidor. Por lo tanto, es un sistema que no sobrecarga los recursos computacionales de sus usuarios.

Actualmente, muchas autoridades, consultores ambientales e industrias de todo el mundo utilizan Airviro para mejorar su trabajo en gestión ambiental, ejemplo de ello es la Red automática urbana y rural del Reino Unido (AURN) y la plataforma del SINCA, que permitió incorporar los inventarios de emisiones existentes realizados por MMA y el monitoreo en tiempo real de sus estaciones. Su falencia es que corresponde a un sistema con licencia pagada.

Otros programas utilizados son herramientas como Python y R, ambos lenguajes de código abierto y entregan funcionalidades mucho más diversas y flexibles para el usuario, sin embargo, requieren mayor trabajo por parte del programador al no tener las funcionalidades preestablecidas de los software más específicos. Las ventajas de estos lenguajes de programación es la capacidad de formular modelos específicos a los requerimientos establecidos por el usuario, evitando tener que utilizar algoritmos predefinidos a casos generales y desarrollar modelos mucho más eficientes. Por ejemplo, Teggi desarrolla un nuevo modelo gaussiano [69], que se codificó en lenguaje Python y se basa en gran medida en una formulación simplificada del modelo AERMOD y la comparación de sus resultados muestra que el modelo reproduce satisfactoriamente las simulaciones de AERMOD y es considerablemente más rápido.

Por otro lado, también destaca la herramienta de Matlab, la cual presenta tantas o más funcionalidades que Python o R y su información se encuentra totalmente documentada. Es una herramienta ampliamente utilizada y validada por la comunidad científica, sin embargo, esta es una herramienta licenciada, por lo que su utilización implica mayores costos. Lo atractivo de esta herramienta son su lenguaje intuitivo y la excelente documentación de sus funcionalidades, además, Matlab cuenta con una nube ThingSpeak, especialmente destinada al desarrollo de proyectos Iot, que permite conectar sensores, transmitir, almacenar y analizar datos en la nube.

Los trabajos reportados en la literatura en su gran mayoría presentan modelaciones tipo *offline*, debido a su bajo costos computacional y facilidad de operación [66]. En la modelación *offline* la meteorología y la química atmosférica son tratadas por separado dentro del modelo de calidad del aire. Sin embargo, los modelos analizados corresponden principalmente a modelos de dispersión física y su complejidad constituye un obstáculo para la implementación *online*.

Tabla 4.2: Comparación *Software* utilizados en la modelación atmosférica.
Elaboración propia.

<i>Software</i>	<i>Software Libre</i>	Modelo que Resuelve	Ventajas	Desventajas
Aermod	Si	Físico de Gauss	Amplia validación científica	No aplicable en tiempo real y deficiente con bajas magnitudes de viento
Calpuff	Si	Físico Lagrangeano	Amplia validación científica y se comporta bien sistemas complejos o turbulentos	No aplicable en tiempo real
Airviro	No	Físico editable por el programador	Cuenta con su propia nube por lo que no sobrecarga los recursos del usuario	Costoso
Matlab	No	Completa decisión del programador	Flexible a los intereses del usuario y alta documentación de sus funciones	Requiere mayor conocimiento del programador
Python	Si	Completa decisión del programador	Flexible a los intereses del usuario	Requiere mayor conocimiento del programador y menor documentación de sus funciones

4.3. Validación cruzada

Un mapa interpolado sin una descripción de su calidad suele considerarse un resultado incompleto. Para describir la calidad y desempeño de un determinado modelo de dispersión se utiliza una técnica denominada validación cruzada. Esta se trata de una validación habitual en la investigación medioambiental [70] y se pueden distinguir dos tipos de validación cruzada comúnmente utilizados, el leave-one-out y la validación cruzada k-fold [71].

En la validación cruzada "leave –one – out" la información de una de las ubicaciones de monitoreo conocidas, se deja fuera del conjunto de datos de entrada y todas las demás localizaciones se utilizan para calcular el valor en la estación excluida. Esto se repite hasta que todos los puntos conocidos se han dejado fuera una vez. Los valores interpolados y los valores observados en todas las ubicaciones de monitoreo se comparan y se calculan diferentes métricas de error [71].

Por otro lado, *K-fold* es otro método de validación cruzada, en el cual el conjunto de datos se divide aleatoriamente en dos subconjuntos. Uno de estos subconjuntos de datos se utilizan como lugar de validación y el otro, como condiciones iniciales para interpolar los valores en las ubicaciones del conjunto excluido. Se repite hasta generar los k subconjuntos de igual tamaño posibles e interpolar los valores de concentración en las ubicaciones de los puntos de los subconjuntos de validación. De hecho, el método *Leave-One-Out*, es un caso particular del método *K-Fold*, donde el subgrupo de validación son los datos de solo una ubicación y en el método *K-Fold* podrían ser los datos de 2,3 o más ubicaciones.

Dicho de otra forma, la idea general de la validación cruzada es subdividir un conjunto de datos en varias partes, utilizando una parte de los datos como lugares de entrada para la interpolación de todos los puntos del plano y otra parte, como lugares de validación. Las ubicaciones de entrada se utilizan ahora para interpolar los valores en las ubicaciones de validación. Los valores reales en estos lugares se conocen y pueden comparar con los valores estimados. Las diferencias entre los valores reales y los valores estimados se utilizan para determinar diversas medidas de error entre las estimación y el valor real.

En este estudio, en la validación cruzada de las interpolaciones realizadas, se calculan las métricas del error medio (ME), el error medio absoluto (MAE) y el error cuadrático medio (RMSE), como lo realizado por N. Long [72]. El ME indica la ausencia de sesgo en la predicción, mientras que el MAE proporciona una medida de hasta qué punto la estimación puede estar en error, ignorando el signo [73]. El RMSE proporciona una medida sensible a los valores atípicos [74]. El ME, el MAE y el RMSE se calculan mediante la ecuación 4.14, 4.15 y 4.16 respectivamente.

$$ME = \frac{1}{N} \sum_{i=1}^N (P_i - P_i^*) \quad (4.14)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |P_i - P_i^*| \quad (4.15)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - P_i^*)^2} \quad (4.16)$$

Otra métrica relevante en la validación cruzada es el porcentaje de error (PE). Este es independiente de la escala y genera un método de previsión cuantitativo, ya que produce una medida del ajuste global relativo. Se utiliza sobre todo para comparar la precisión entre dos métodos cualquiera en porcentaje. Es fácil de interpretar y se define según la ecuación 4.17.

$$PE(\%) = \frac{RMSE}{\frac{1}{N} \sum_{i=1}^N P_i^*} * 100(\%) \quad (4.17)$$

En las ecuaciones 4.14, 4.15, 4.16 y 4.17, (P_i^*) corresponde al valor observado, (P_i) al valor predicho y N es el número de observaciones.

Otra métrica ampliamente utilizada corresponde al porcentaje de error absoluto (MAPE), al igual que el PE, esta métrica es independiente de la escala. En la ecuación 4.18 [70].

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \frac{|O_i - E_i|}{O_i} \right) * 100 \quad (4.18)$$

El índice de correlación (IOA) se utiliza para determinar la relación o concordancia entre los datos reales y los interpolados. Es un parámetro adimensional y su valor oscila entre 0 y 1. El IOA puede expresarse como en ecuación 4.19

$$IOA = 1 - \frac{\sum_{i=1}^n (O_i - E_i)^2}{\sum_{i=1}^n (|O_i - \mu| + |E_i - \mu|)^2} \quad (4.19)$$

En la ecuación 4.19, μ representa la media de los valores observados. El IOA se utiliza sobre todo en los estudios de investigación social en los que existe una gran necesidad de un índice de acuerdo para cada hecho [75]. En los estudios sobre la contaminación atmosférica, se pueden comparar dos algoritmos de interpolación espacial mediante un procedimiento de validación cruzada utilizando el índice de correlación.

El estudio realizado por Khazaz en [?], concluye que para un área de estudio no existe un esquema de interpolación óptimo predeterminado, depende de la disposición espacial de las redes de vigilancia. Por lo tanto, para identificar la mejor técnica de interpolación en un área determinada, se deben estudiar y comparar las distintas técnicas mediante métricas estadísticas de error.

4.4. Estado del Arte Modelación Atmosférica

A continuación, se presentan una serie de estudios y/o proyectos realizados sobre el modelamiento en tiempo real de la calidad del aire por medio de sensores de bajo costo.

Proyecto Kolkata, India

En la fase II del Proyecto de Implementación de Vehículos Eléctricos en Kolkata, se realiza un estudio de la calidad del aire por IDEATION TECHNOLOGY SOLUTIONS en [76]. Se implementa una red de sensores portátiles de bajo costo para medir datos de contaminación atmosférica en puntos localizados. La red utilizada consta de 19 ubicaciones de la región de Kolkata, ocho sensores de bajo costo *Cleair* y once estaciones de monitoreo gubernamentales. El perímetro de monitoreo es de aproximadamente 70 km, un detalle de las de las 19 ubicaciones se observa en la figura 4.12, la estación gubernamental situada en la Universidad Rabindra Bharti se utiliza para calibrar los sensores de bajo costo y en la figura 4.13 se observa una imagen del sensor de bajo costo utilizado en Kolkata.



Figura 4.12: Mapa ubicaciones de medición estudio Kolkata [76].



Figura 4.13: Sensor de bajo costo utilizado en estudio Kolkata.

El modelo de dispersión implementado en este estudio se sustenta en el diagrama de Voronoi, donde los sensores de bajo costo y las estaciones gubernamentales se utilizan para formar células Voronoi en la región de Kolkata. Cada celda de Voronoi representa una estación y el nivel de contaminación dentro de la celda. A diferencia del método de interpolación de vecinos naturales visto anteriormente, en este estudio, las celdas de Voronoi son la entrada a un modelo de dispersión físico, que para cada celda asigna las lecturas del sensor a una ecuación subyacente de difusión y advección, que representa la propagación de las partículas en cada celda.

El modelo de dispersión que se aplica muestra un enfoque para predecir la contaminación en tiempo real. El modelo produce resultados en estado estacionario del perfil de partículas conectado suavemente con las lecturas obtenidas en las ubicaciones de monitoreo. Este modelo se pondrá en funcionamiento en cuanto se recojan suficientes datos de la red de sensores y se validará utilizando sensores móviles para registrar las desviaciones de las estimaciones del modelo.

Proyecto Lyon, Francia

A. Boubrima, plantea un nuevo modelo de dispersión en tiempo real basado en la interpolación de distancia inversa ponderada [5]. El modelo propuesto considera la mejora de incluir el efecto de la velocidad y magnitud del viento en el cálculo del coeficiente de correlación entre dos puntos. La formulación propuesta es lo suficientemente general para adaptarse a otros métodos de dispersión deterministas, como el de los vecinos naturales.

Además, este estudio busca determinar las mejores posiciones de los sensores, para lo cual plantea un algoritmo de optimización que busca minimizar el costo de implementación de la red, asegurando la calidad de cobertura requerida. Junto a ello, plantea un segundo modelo, que busca optimizar la calidad de la cobertura sin exceder un presupuesto de implementación

determinado. Los valores de los coeficientes de correlación se pueden considerar como una variable de entrada o utilizar modelos de dispersión físicos para obtener valores de concentración simuladas y a partir de ellas, determinar los coeficientes de correlación Z_{pq} entre distintos puntos (p,q) y determinar a partir de esa información la ubicación óptima de los sensores [77].

En este estudio, A. Boubrima, utiliza una red de cinco sensores de NO₂ en el distrito lionés de La-Part-Dieu, en la figura 4.14 se observa un diseño de la red. Se utiliza un software de dispersión atmosférica mejorado llamado SIRANE [78], que está diseñado para zonas urbanas y tiene en cuenta el impacto de los cañones de las calles en la dispersión de la contaminación.

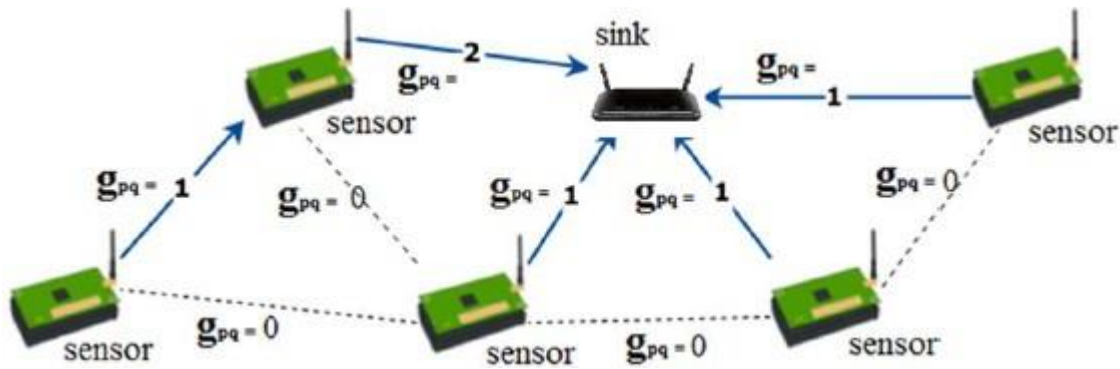


Figura 4.14: Configuración red de sensores Caso de estudio Lyon. Tomado de [5].

Boubrima plantea mejoras a la definición determinista de los coeficientes de correlación, integrando mas variables como un factor de atenuación a la distancia, velocidad, dirección viento, entre otros. Lo anterior debido a que el calculo de coeficientes de correlación mediante distancia inversa ponderada o vecinos naturales, supone que la interpolación de $p \rightarrow q$ o de $q \rightarrow p$ tienen el mismo valor. Para superar estas limitaciones, se puede utilizar la ecuación 4.20 para calcular los parámetros de los coeficientes de correlación W_{pq} . Esta formulación, es lo suficientemente general como para ajustarse a otros métodos de interpolación determinista, principalmente la interpolación de vecino natural, pero también a otros como la interpolación de vecino más cercano [9]. La interpolación del vecino natural se puede obtener configurando el factor de atenuación a un valor suficientemente alto para que sólo los vecinos más cercanos se tengan en cuenta en la fórmula de interpolación.

$$W_{pq} = B_{pq} * \frac{1}{D(p,q)} * |\vec{u}| \frac{(1 + \cos(\angle(\vec{p}q, \vec{u})))}{2} \quad (4.20)$$

En la ecuación 4.20, \vec{u} denota el vector viento y $|\vec{u}|$ la velocidad normalizada del viento en el rango [0, 1] y $\cos(\angle(\vec{p}q, \vec{u}))$ es el coseno del ángulo entre los vectores $\vec{p}q$ y \vec{u} . Gracias al uso de la función coseno, los valores de correlación se reducen cuando la dirección del viento es opuesta a $\vec{p}q$. Para tener en cuenta el impacto de la topografía urbana la

topografía urbana en la dispersión de los contaminantes, D puede definirse como la distancia más corta a lo

largo de la red de carreteras. Esto nos permite asignar valores de correlación pequeños a los puntos separados por edificios, aunque estén cerca con respecto a la distancia euclidiana [78].

El problema del despliegue de las redes de sensores inalámbricos se ha abordado ampliamente en la literatura, los principales aspectos que se abordan son la cobertura, la conectividad y la vida útil de la red. Los enfoques de despliegue existentes son consientes de los eventos [79] o de la correlación [80]. En el primer caso, se supone que un sensor tiene un rango de detección normalmente circular, dentro del cual el sensor es capaz de detectar cualquier evento que pueda ocurrir. La segunda clase de enfoques de despliegue se basa en la correlación que pueden presentar las mediciones de los sensores y con ello, seleccionar el número mínimo de nodos de detección.

De los resultados obtenidos por Boubrina en la evaluación de óptimo despliegue. Se observa que en cuanto al valor de la función objetivo, el costo de despliegue óptimo aumenta con la variabilidad de la contaminación; es decir, se necesitan más sensores cuando la variabilidad de la contaminación es mayor.

Proyecto Wang, China

Wang [81], presenta un enfoque de modelamiento de la distribución de los contaminantes utilizando una red de sensores móviles. En el estudio Wang utiliza 25 sensores móviles. En el enfoque propuesto, se presenta un modelo de detección autónomo y un sistema de control de movimiento basado en la mínima energía consumida por los dispositivos.

Existen dos formas de monitoreo basadas en sensores móviles, una sin modelos y otra basada en modelos. En el primer método los nodos sensores se desplazan directamente a la fuente de contaminación, sin modelar la distribución de los contaminantes y los nodos se distribuyen de la misma manera que la distribución del contaminante. En los basados en modelos, los nodos se mueven hacia la fuente de contaminación y, mientras tanto, construyen el modelo de distribución según las mediciones de los sensores. En este estudio, Wang desarrolla un monitoreo en tiempo real mediante el uso de sensores móviles, aplicando un método basado en modelos. Desarrolla un modelo basado en la regresión de vectores de soporte (SVR) y utiliza la partición de Voronoi para acotar el desplazamiento de cada nodo a una celda Voronoi, mediante una serie de iteraciones el nodo móvil se desplaza por su celda de Voronoi alimentando el modelo de regresión hasta reconstruir la distribución de contaminantes.

Un amplio conjunto de simulaciones demuestra que el modelo desarrollado puede modelar la distribución de contaminantes con menos tiempos de iteración y mayor precisión. Incluso para un campo de concentración relativamente complejo, la similitud entre el modelo reconstruido y la distribución de contaminantes puede alcanzar el 95 % mediante unas 20 iteraciones. La figura 4.15 muestra el procedimiento de simulación de 30 iteraciones, donde la primera fila es el estado inicial y las siguientes tres filas son los resultados en las iteraciones 5, 15 y 25, respectivamente. Se puede observar que el modelo reconstruido es cada vez más preciso a medida que los nodos sensores se dispersan gradualmente desde sus posiciones originales. Esto se debe a que los nodos sensores obtienen suficientes conocimientos sobre la distribución de los contaminantes durante el movimiento. Cada nodo sensor se

mueve constantemente hacia la fuente de contaminación, y el proceso de movimiento se guía por la energía residual de los nodos y un criterio de distancia entre ellos.

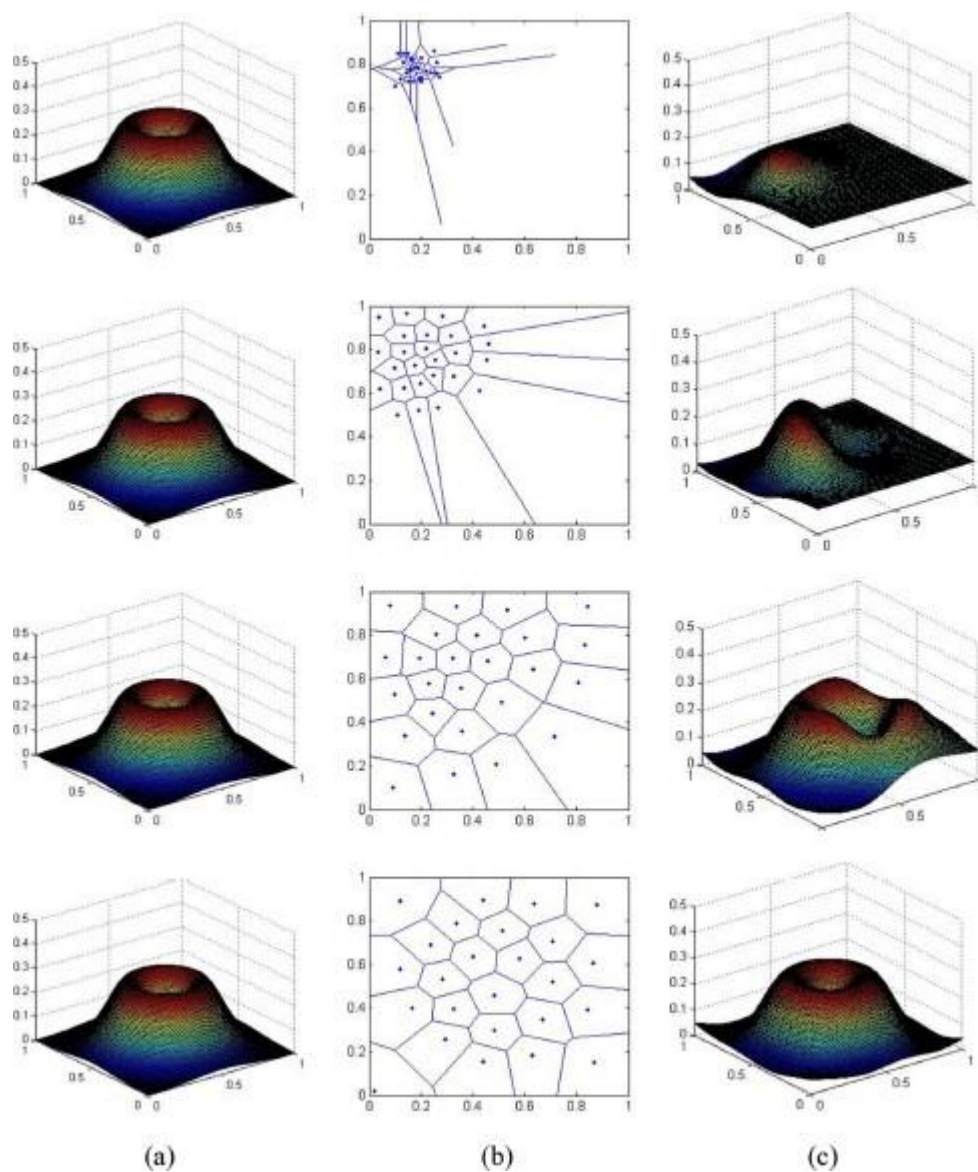


Figura 4.15: Resultados de la simulación para las iteraciones 5, 15 y 25, respectivamente. (a) Distribución de contaminantes, (b) Diagrama de Voronoi de ubicación de los nodos, (c) reconstrucción del modelo. Tomado de [81].

4.4.1. Proyecto Deo Nai, Vietnam

Long genera un modelo de dispersión mediante el uso de un sensor móvil [72]. Este corresponde a un dron DJI Inspire 2, que se presenta en la figura 4.16. El DJI Inspire 2 puede despegar con un objeto de 810 g, esto permite que el dron vuele con seguridad con los sensores de bajo costo montados en él. Los sensores de bajo costo utilizados corresponden a sensores electroquímicos y un sensor de material particulado. El estudio se realizó en la mina de carbón a cielo abierto de Deo Nai, situada en el norte de Vietnam.



Figura 4.16: Dron DJI Inspire 2 montado con sensores de bajo costo para el monitoreo de calidad del aire. Tomado de [72].

Este estudio plantea un enfoque para estimar con precisión la concentración 3D de los contaminantes monitoreados en la atmósfera de la mina de carbón. Para lograr este objetivo, se probaron diferentes métodos de interpolación con los datos adquiridos por el sensor móvil. Se eligió el método más adecuado para construir la distribución 3D de la calidad del aire basándose en la validación cruzada realizada con cuatro métricas de error, incluyendo el error medio (ME), el error medio absoluto (MAE), el error cuadrático medio (RMSE) y el error porcentual (PE). Respecto a los resultados obtenidos, que se presentan en las tablas 4.3, 4.4 y 4.5, se observa que la interpolación trilineal obtuvo el resultado más preciso, seguido del vecino natural y el vecino más cercano [72].

Tabla 4.3: Resultados validación cruzada para MP 2,5 [72].

Modelo	ME	MAE	RMSE	PE(%)
Interpolación Trilineal	-0,116	4,7	5,8	15,2
Vecino mas cercano	0,25	5,75	7,1	18,4
Vecino Natural	-0,5	5	6,2	16,2

Tabla 4.4: Resultados validación cruzada para CO [72].

Modelo	ME	MAE	RMSE	PE(%)
Interpolación Trilineal	0,058	0,102	0,192	10,3
Vecino mas cercano	0,04	0,145	0,295	15,2
Vecino Natural	0,066	0,118	0,210	11,3

Tabla 4.5: Resultados validación cruzada para CO_2 [72].

Modelo	ME	MAE	RMSE	PE(%)
Interpolación Trilineal	4.428	14,086	19,189	3,7
Vecino mas cercano	7,208	18,458	26,894	5,2
Vecino Natural	5,631	15,396	19,702	3,8

Los resultados muestran que el sistema de monitorización de la calidad del aire funcionó de forma eficiente y segura en las condiciones de las minas de carbón a cielo abierto. Para cada método de interpolación, se empleó el mecanismo de validación cruzada *K-Fold*, que utilizó el 80 % de los datos observados para calcular la concentración de los componentes del aire en las ubicaciones del 20 % de los datos excluidos, estos últimos fueron utilizados para el análisis de las métricas de error entre la estimación y la observación.

De los resultados obtenidos por Long, se puede concluir que distintos contaminantes presentan diversas distribuciones espaciales y por lo tanto, los coeficientes de correlación de cada método pueden representar de mejor manera algunos contaminantes. Estas diferentes distribuciones espaciales pueden tener su origen en factores como la cantidad de puntos de emisión en el área de cada contaminantes, las reacciones químicas, entre otras. Por lo anterior es que se presentan errores porcentuales mayores en el caso del $MP_{2,5}$, en comparación al CO y al CO_2 . Esto se debe, a que el estudio se realizó en un área minera, donde producto de los procesos la mayor cantidad de emisiones que se generan corresponde a material particulado, por lo tanto presenta una mayor variabilidad espacial que el CO y al CO_2 .

4.4.2. Proyecto NaturalKriging

Van del Graaf integra las metodologías de interpolación estocásticas y deterministas del Kriging y Vecinos naturales en un solo método [58]. Sustentado en que el kriging requiere la inversión de matrices y, por lo tanto, puede ser un proceso computacionalmente ineficiente cuando se trata de grandes conjuntos de datos. Una forma de superar este problema, es el uso de una vecindad de interpolación acotada, un subconjunto de puntos seleccionados que están cerca del punto de estimación son utilizados para implementar el algoritmo del Kriging. Este subconjunto de puntos queda determinado por los vecinos naturales al punto de interpolación, reduciendo considerablemente el tamaño de la matriz de covarianza.

El método de kriging de los vecinos naturales es analizado con los “Datos de prueba de Davis”. Este conjunto de datos de altura fue creado por Davis en 1973, se utiliza habitualmente como conjunto de datos de prueba para interpolaciones. El conjunto de datos consta de 52 valores de altura en un sistema de coordenadas local de tres dimensiones.

Se comparan los resultados de interpolación para los tres métodos de interpolación y se estiman los errores medios de predicción utilizando los dos tipos de validación cruzada. Además se consideran diferentes modelos de variograma para kriging de vecinos naturales para obtener una comprensión general del método.

El RMSE medio para para la interpolación de vecinos naturales es de 798 m^2 , para el

kriging ordinario de 636 m² y para el kriging de vecinos naturales de 614 m². La desviación estándar del MSE de estos métodos es de 125 m² para la interpolación de vecinos naturales, 64 m² para el kriging ordinario y 61 m² para el kriging natural. El kriging de vecino natural tiene el menor RMSE medio y la desviación estándar más baja.

4.4.3. Proyecto Valencia, España

Lidia Contreras realiza un estudio en España, en la ciudad de Valencia, para integrar el efecto de la dirección y magnitud del viento en un modelo de dispersión de contaminantes en tiempo real basado en interpolaciones espaciales [82]. Se comparan varios modelos de regresión capaces de predecir los niveles de cuatro contaminantes (NO, NO₂, SO₂ y O₃) en seis lugares diferentes de la ciudad. Se analizan diversas técnicas de regresión para integrar el efecto de la dirección y magnitud de viento en la dispersión de contaminantes.

Se analiza el método de interpolación que tiene en cuenta la dirección y magnitud del viento, se compara con otros métodos de interpolación espacial, como la ponderación de distancia inversa (IDW), el Kriging, una variante de la ponderación de distancia inversa en función de la distancia euclidiana (LIDW) y el método propuesto de IDW considerando el efecto del viento. La metodología de validación validación cruzada muestra que este nuevo método obtiene mejores resultados en comparación con los demás métodos conocidos. En la figura 4.17 se observan los resultados para la validación cruzada, el mejor modelo de predicción se destaca en negrita.

	A)3 Known - 3 Unknown 20it					B)4 Known - 2 Unknown 15it					C)5 Known - 1 Unknown 5it				
	Mean	LIDW	IDW	Wind	Krig.	Mean	LIDW	IDW	Wind	Krig.	Mean	LIDW	IDW	Wind	Krig.
NO	22.15	19.55	17.67	16.92	19.94	18.13	17.76	17.89	14.99	19.89	17.13	16.72	14.76	14.71	22.98
NO ₂	21.49	22.06	21.46	20.02	19.99	19.04	22.19	20.80	17.31	20.56	17.90	19.20	16.31	16.09	24.18
O ₃	15.37	14.54	17.54	17.72	13.16	12.32	20.44	14.44	15.13	12.66	12.03	13.74	13.87	13.90	18.66
SO ₂	3.05	3.23	2.95	2.91	2.90	3.19	2.96	3.05	2.47	3.42	2.97	2.90	2.37	2.47	3.42

Figura 4.17: Comparación de RMSE de los 4 métodos de interpolación analizados. Tomado de [82].

Resumen Estado del Arte

A continuación, en la tabla 4.6 se presenta un resumen de las principales características de los estudios analizados sobre el estado del arte de la modelación atmosférica en tiempo real.

Tabla 4.6: Comparación del estado del Arte de la modelación atmosférica en tiempo real. Elaboración propia.

Proyecto	N° de equipos	Modelo que Resuelve	Beneficios para la Investigación
Kolkata	19	Físico acotado por diagramas de Voronoi	Integra el uso de diagramas de Voronoi con modelos físicos
Francia	5	Matemático IDW que incluye la variable viento	Integra la variable viento en métodos matemáticos
China	25	Físico acotado por diagramas de Voronoi	Integra el uso de diagramas de Voronoi con modelos físicos
Vietnam	1	Matemático Interpolación Trilineal, Vecino mas Cercano y Vecino Natural	Uso de sensores móviles que buscan fuentes de emisión
Natural Kriging	50	Matemático Kriging acotado por Vecinos Naturales	Logra integrar los métodos matemáticos del Kriging y Vecinos Naturales
España	No informa	Matemático IDW que considera la variable viento	Integra la variable viento en un método de dispersión matemático

Se observa que los modelos de dispersión matemáticos basados en interpolaciones son ampliamente utilizados en el desarrollo de modelos de dispersión con adquisición de datos en tiempo real. Estos métodos han demostrado excelentes resultados en la modelación en tiempo real de contaminantes en la atmósfera, sin embargo, todos los autores que desarrollan este tipo de modelos plantean mejoras en las variables consideradas en el calculo de los coeficientes de correlación. En general, se incluyen factores de atenuación al efecto generado por el viento o la topografía del lugar, lo cual ha permitido obtener excelentes resultados.

Otro aspecto relevante a notar, es el amplio uso de los diagramas de Voronoi en la modelación atmosférica no solo para realizar interpolaciones. Muchos autores lo utilizan para acotar los puntos considerados en un problema, a un set mas pequeño y con mayor representatividad o correlación. La propiedad de dividir el espacio en celdas que contienen los vecinos mas cercanos a un punto “x”, ha permitido su integración con variados modelos

físicos y químicos, acotando el área de modelación a la celda de Voronoi o vecinos mas cercanos del punto de interés. Ha demostrado ser lo suficientemente robusto para adecuarse con otros métodos de dispersión físicos y matemáticos, generando excelentes resultado. Por lo tanto, el uso de diagramas de Voronoi se plantea como una excelente alternativa para el mapeo de contaminantes o definir áreas de representación.

De los resultados obtenidos de Long y Contreras, se observa que para un mismo método, distintos contaminantes presentan diferentes resultados. Además, Contreras demostró mejoras significativas en los resultados obtenidos para los contaminantes relacionados directamente a procesos de combustión. También concluyen mejoras significativas al aumentar el numero de puntos de monitoreo para realizar las interpolaciones.

Capítulo 5

Modelo propuesto para la dispersión en tiempo real de COV atmosféricos utilizando sensores de bajo costo

En este capítulo, se presenta el método seleccionado para el modelo de dispersión en tiempo real y una modificación a su algoritmo de cálculo, que permite integrar el efecto de puntos de monitoreo más lejanos. En primer lugar, se presentan los requerimientos del modelo de dispersión en tiempo real que justifican la selección del método y herramientas computacionales. Posteriormente, se presenta el cuestionamiento al método original que sustenta el análisis del algoritmo modificado.

Seguidamente, se expone en detalle el funcionamiento del método original y el modificado. Para cada método, se expone un ejemplo de cálculo para la estimación de dos puntos a partir de cuatro ubicaciones de monitoreo en una área definida. Finalmente, se presenta el funcionamiento global del modelo, desde el envío de datos, hasta el funcionamiento de los módulos del algoritmo de dispersión en tiempo real.

El modelo de dispersión de contaminantes en tiempo real busca desarrollar una alternativa de modelación de bajo costo computacional, que pueda ser aplicada como un mecanismo de alerta temprana de calidad del aire. Esta, debe dar una respuesta rápida, que permita prever tempranamente situaciones de riesgo ambiental. De esta forma, se puede sacrificar la precisión de los resultados, siempre y cuando, estos conserven la verosimilitud.

Por lo anterior, se desarrolla un modelo basado en la interpolación de vecinos naturales, debido a sus menores requerimientos de implementación. Estos, permiten desarrollar un modelo simple y rápido. Además, según la bibliografía analizada en la sección 4.1.2.2.2, es un método que permite representar la distribución de contaminantes atmosféricos de manera fidedigna, posicionándose como una excelente alternativa para su aplicación en tiempo en real.

Cabe mencionar, que el método de los vecinos naturales requiere datos de concentración en ubicaciones predefinidas, para interpolar la concentración en otros puntos del espacio. Por lo tanto, la precisión de sus resultados está condicionada a la disponibilidad de datos reales para interpolar los demás puntos. Por este motivo, para la adquisición de datos, se utilizaron

cinco sensores de foto-ionización de bajo costo alphasense. Un detalle del funcionamiento de estos equipos se presenta en el anexo F.

Como la ejecución del modelo de dispersión requiere la adquisición de datos sincrónicos, se utiliza la herramienta de Matlab para implementarlo, ya que es un lenguaje flexible, que facilita el desarrollo de un modelo con adquisición de datos en tiempo real. Junto a lo anterior, Matlab, cuenta con una nube especialmente destinada para la conexión de sensores, envío y almacenamiento de información, a la cual se puede acceder en todo momento desde la plataforma de Matlab. Por lo tanto, se posiciona como una excelente alternativa, que facilita la adquisición de datos y permite la flexibilidad necesaria para plantear un modelo de dispersión en tiempo real.

Por otro lado, un aspecto fundamental en la geoestadística es el alcance de la correlación espacial de los datos muestreados. Este aspecto el Kriging lo integra mediante el variograma, que es la función de covarianza de los puntos de un plano respecto de la distancia entre ellos. El alcance hace referencia a la distancia espacial, a la cual el valor de una VR, como la concentración en un punto del plano, presenta covarianza nula con el valor de la VR en otro punto del plano. Es decir, una alta covarianza entre “X” e “Y”, indica que si aumenta “X”, entonces “Y” aumenta.

De esta forma, el alcance se define como la distancia a la cual el valor de una VR presenta covarianza nula entre dos puntos. Este es el caso de covarianza cero o sin solapamiento, que se observa en la parte inferior de la figura 5.1, donde la distancia que separa a las circunferencias es igual alcance del sistema. Este aspecto se observa, ya que no se genera un solapamiento de los círculos y por lo tanto, no existe correlación entre dichos puntos.

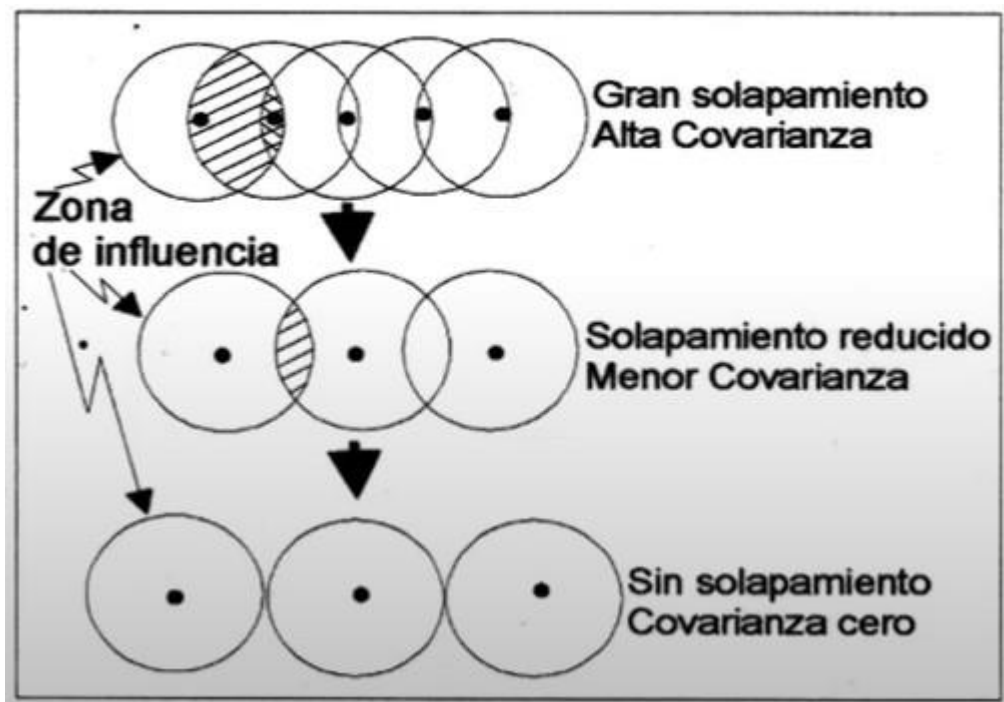


Figura 5.1: Esquema solapamiento áreas de influencia muestras [83]

Por otro lado, en la parte superior de la figura 5.1, se observa el área de influencia de dos puntos contiguos, mediante círculos superpuestos. En este caso, se observa que el área de influencia de los círculos generados por dos puntos contiguos, es mucho mayor, que el área de influencia de los círculos de la parte central de la figura 5.1, que se encuentran más alejados entre sí. Por lo tanto, esta mayor área de solapamiento indica que existe una mayor covarianza entre los puntos.

Una posible falencia del método es que la interpolación de los vecinos naturales solo considera los puntos más cercanos para realizar la interpolación. Por lo tanto, se podrían estar excluyendo puntos del plano, con los cuales si podría existir correlación matemática. Esto se podría apreciar en el caso del solapamiento reducido, ya que, el área de influencia del primer círculo se superpone sobre el segundo y del mismo modo, el área del segundo círculo, se superpone sobre el tercero. Además, se observa, que el área del primer círculo no presenta un solapamiento con el tercero, por lo tanto, hay covarianza cero entre el primer y tercer punto de la parte central de la figura 5.1. Sin embargo, la lógica plantea que si el punto 1 presenta correlación con el punto 2 y, el punto 2 presenta correlación con el punto 3, entonces el punto 1 debería presentar correlación con el punto 3.

Este efecto podría ser relevante, ya que en la realidad, no todos los puntos del plano presentan una correlación constante. Esto, debido a que características meteorológicas o geográficas condicionan que algunos puntos presenten mayor correlación que otros. Por ejemplo, producto que el viento sople en una dirección o la presencia de montes y/o edificios que perturben la dispersión de contaminantes en una dirección. Además, otros factores antropogénicos, como fuentes de emisión de contaminantes pueden inducir reacciones químicas de manera diferenciada en el área de estudio, dependiendo la magnitud y tipo de emisión.

Con la intención de plantear un algoritmo que logre representar de mejor forma la distribución de contaminantes en el espacio, se desarrolla el algoritmo de interpolación de vecinos naturales como se presenta en el capítulo 4.7 y, una modificación iterativa, que busca modificar la correlación espacial planteada por el método original e integrar la posible correlación espacial de puntos mas lejanos. El método iterativo se sustenta en que si la ubicación 1 esta relacionada con 2 y la ubicación 2 con la 3, entonces 1 presenta correlación con 3. De esta forma, si se interpola la concentración C_2 en un punto 2, a partir de la concentración C_1 en un punto 1 y luego, se interpola la concentración C_3 en punto 3 a partir de C_2 , entonces el cálculo de C_3 considerará intrínsecamente el efecto del valor de C_1 , por medio de C_2 . De esta forma, en el cálculo del punto 3, se considera la correlación espacial existente entre el punto 1 y 2, y además, la existente entre los puntos 2 y 3.

A continuación se profundiza en el mecanismo de interpolación de cada uno de estos métodos. Se presenta un ejemplo de interpolación del algoritmo original y modificado, a partir de la ubicación de 4 puntos de monitoreo en una área definida.

5.1. Algoritmo Original

Para comprender de mejor manera el mecanismo de resolución de este algoritmo, se plantea un ejemplo de estimación de dos puntos desconocidos. Se suponen 4 sensores de contaminantes en las ubicaciones C_1 , C_2 , C_3 , C_4 . Primero, se genera un diagrama de Voronoi para las ubicaciones de los puntos de monitoreo $(x_1, y_1) = (1, 1)$, $(x_2, y_2) = (1, 3)$, $(x_3, y_3) = (4, 1)$, $(x_4, y_4) = (4, 3)$. La partición de Voronoi generada se observa en la figura 5.2.

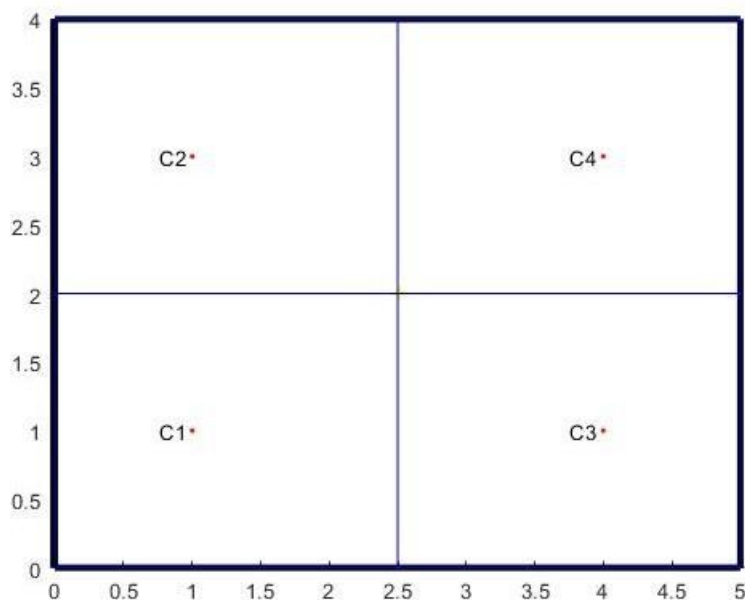


Figura 5.2: Diagrama de Voronoi a partir de cuatro nodos. Ejemplo interpolación Vecinos Naturales. Elaboración propia.

Luego, si se quisieran interpolar los puntos en las ubicaciones $(x_5, y_5) = (2, 5, 2)$ y $(x_6, y_6) = (4, 35, 2)$, se generarán dos nuevos diagramas de Voronoi. Cada nuevo diagrama, incluye la ubicación del punto de interpolación y las ubicaciones de los sensores. De esta forma, se genera un DV que contiene los puntos $a = (1, 2, 3, 4, 5)$ y otro que contiene los puntos $b = (1, 2, 3, 4, 6)$. Luego se analiza el efecto generado por la nueva celda ingresada en el diagrama de Voronoi original, tal como se presentó en el ejemplo del capítulo 4.1.2.2.2. Los nuevos diagramas de Voronoi se pueden observar en las figuras 5.3 y 5.4 respectivamente.

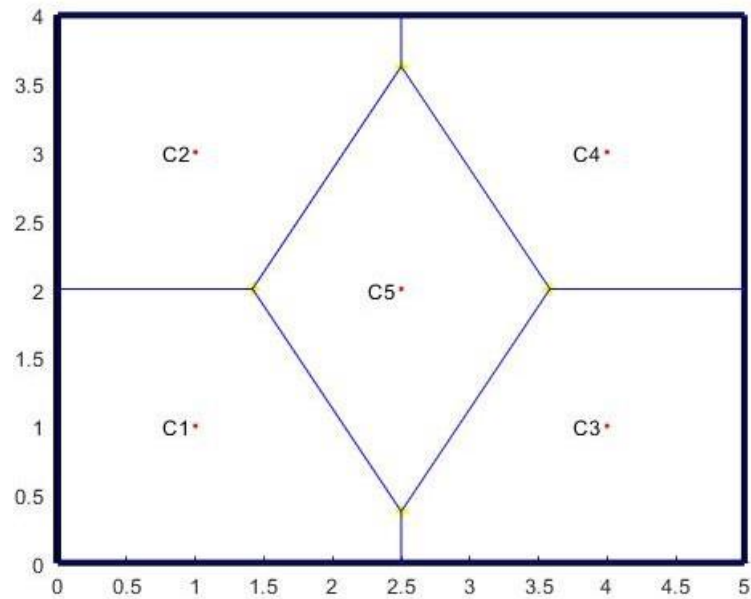


Figura 5.3: Ejemplo Diagrama de Voronoi a partir de cinco puntos (a). Método de los Vecinos Naturales. Elaboración propia.

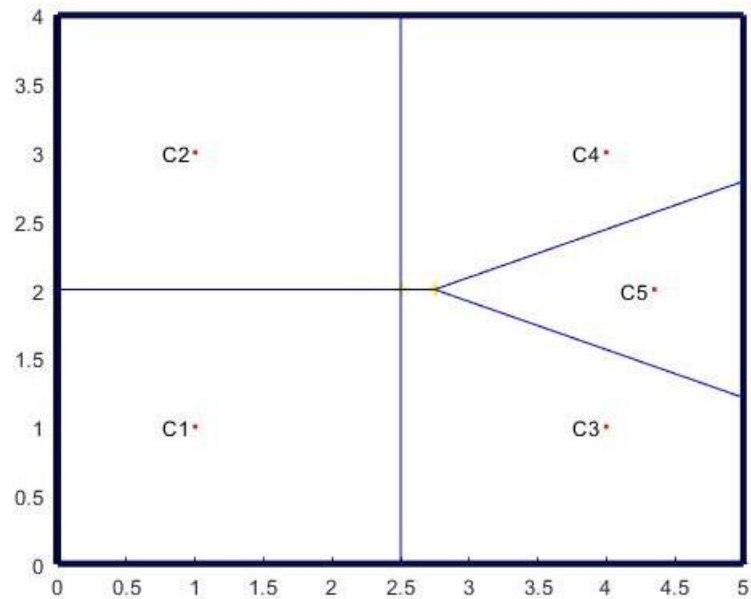


Figura 5.4: Ejemplo Diagrama de Voronoi a partir de cinco puntos (b). Método de los Vecinos Naturales. Elaboración propia.

En este método, solo se considera el efecto de los vecinos naturales del punto de estimación para calcular el coeficiente de correlación. Se observa en la figura 5.3 del caso (a), que los vecinos naturales de las coordenadas $(x_5, y_5) = (2.5, 2)$, se posicionan sobre las antiguas celdas C_1, C_2, C_3, C_4 . Por lo tanto, el valor en el punto C_5 depende de las concentraciones en $C_1,$

C_2 , C_3 y C_4 . Del mismo modo, en la figura 5.4 se observa que en el caso (b), los vecinos naturales de las coordenadas $(x_6, y_6) = (4.35, 2)$, se posicionan sobre las antiguas celdas C_3 y C_4 , por lo que su cálculo solo depende de los valores de C_3 y C_4

De esta forma, el cálculo de C_5 depende de cuatro coeficientes de correlación, se observa que el área de la celda C_5 se intersecta equitativamente sobre las celdas iniciales C_1 - C_4 , por lo tanto todos sus coeficientes de correlación son equivalente.

$$W_{1,5} = \frac{1}{4} = 0,25 \quad (5.1)$$

$$W_{2,5} = \frac{1}{4} = 0,25 \quad (5.2)$$

$$W_{3,5} = \frac{1}{4} = 0,25 \quad (5.3)$$

$$W_{4,5} = \frac{1}{4} = 0,25 \quad (5.4)$$

Por otro lado, los coeficientes de correlación de C_6 dependen de C_3 y C_4 . De esta forma, el coeficiente de correlación $W_{3,6}$ se calcula mediante el cociente entre área de la celda C_6 superpuesta sobre el área de la celda C_3 , dividida en el área total de la celda C_6 . Del mismo modo, el coeficiente de correlación $W_{4,6}$ se calcula mediante el cociente entre área de la celda C_6 superpuesta sobre el área de la celda C_4 , dividida en el área total de la celda C_6 . En las ecuaciones 5.5 y 5.6 se presenta su cálculo.

$$W_{3,6} = \frac{(0,75 * 2,25)/2}{(1,5 * 2,25)/2} = 0,5 \quad (5.5)$$

$$W_{4,6} = \frac{(0,75 * 2,25)/2}{(1,5 * 2,25)/2} = 0,5 \quad (5.6)$$

De esta forma, las concentraciones C_5 y C_6 quedan definidas por las ecuaciones 5.8 y ?? respectivamente. :

$$C_5 = W_{1,5} * C_1 + W_{2,5} * C_2 + W_{3,5} * C_3 + W_{4,5} * C_4 \quad (5.7)$$

$$C_6 = W_{3,6} * C_3 + W_{4,6} * C_4 \quad (5.8)$$

Evaluando los coeficientes de correlación calculados, se obtiene:

$$C_5 = 0,25 * C_1 + 0,25 * C_2 + 0,25 * C_3 + 0,25 * C_4 \quad (5.9)$$

$$C_6 = 0,5 * C_3 + 0,5 * C_4 \quad (5.10)$$

5.2. Algoritmo modificado

La idea de este mecanismo es generar un mayor solapamiento o correlación al momento de realizar la interpolación. Para esto, una vez interpolado un punto mediante el método de vecinos naturales, este valor se incluye como un valor valido en la lista de puntos iniciales. Luego, cuando se interpola un segundo punto, la partición de Voronoi original a considerar es

la realizada para interpolar el punto anterior, considerando el valor calculado en la iteración anterior, como un valor valido en el cálculo actual.

Siguiendo con el ejemplo anterior, se suponen las mismas ubicaciones para 4 sensores de contaminantes en las ubicaciones C_1 , C_2 , C_3 y C_4 . Primero, se genera un diagrama de Voronoi para las ubicaciones de los puntos de monitoreo $(x_1, y_1) = (1, 1)$, $(x_2, y_2) = (1, 3)$, $(x_3, y_3) = (4, 1)$, $(x_4, y_4) = (4, 3)$. La partición de Voronoi generada se observa en la figura 5.5.

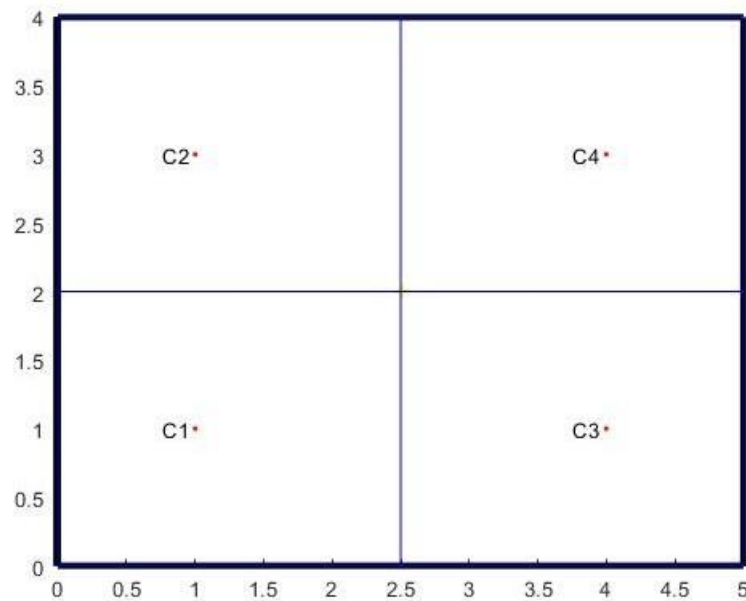


Figura 5.5: Ejemplo Diagrama de Voronoi a partir de cuatro ubicaciones. Método iterativo de los Vecinos Naturales. Elaboración propia.

Luego, si se quisieran interpolar los puntos en las ubicaciones $(x_5, y_5) = (2.5, 2)$ y $(x_6, y_6) = (4.35, 2)$, primero se debe seleccionar que ubicación se interpolará en primer lugar. En este caso, se sigue el mismo orden del ejemplo anterior con la ubicación $(x_5, y_5) = (2.5, 2)$. El diagrama de Voronoi resultantes es:

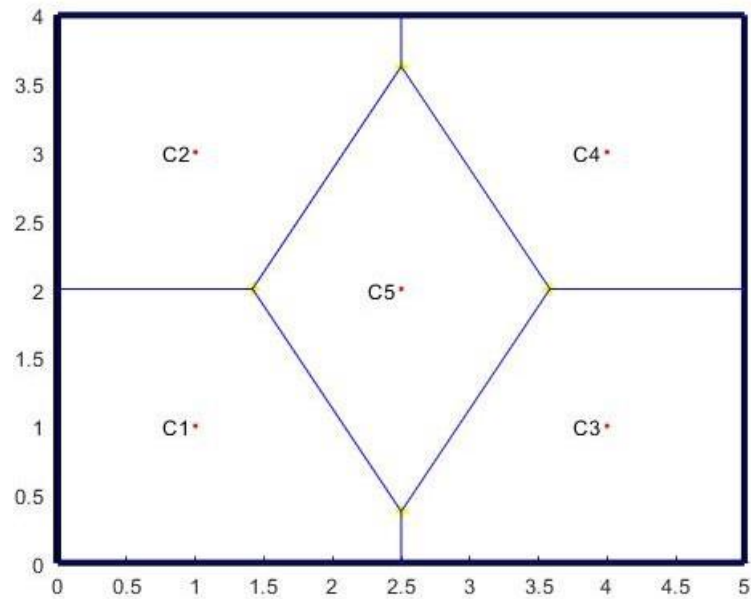


Figura 5.6: Ejemplo Diagrama de Voronoi a partir de cinco nodos. Interpolación iterativa. Elaboración propia.

Al igual a lo obtenido en el caso anterior, se observa claramente que la celda generada por C_5 intersecta las celdas C_1 , C_2 , C_3 y C_4 , luego el valor en el punto C_5 depende de las concentraciones en C_1 , C_2 , C_3 y C_4 . La diferencia se observa al interpolar el 2do punto en $(x_6, y_6) = (4.35, 2)$. En este caso, se debe mantener el diagrama anterior antes de agregar el nuevo punto de interpolación. El diagrama de Voronoi resultante se presenta en la figura 5.7.

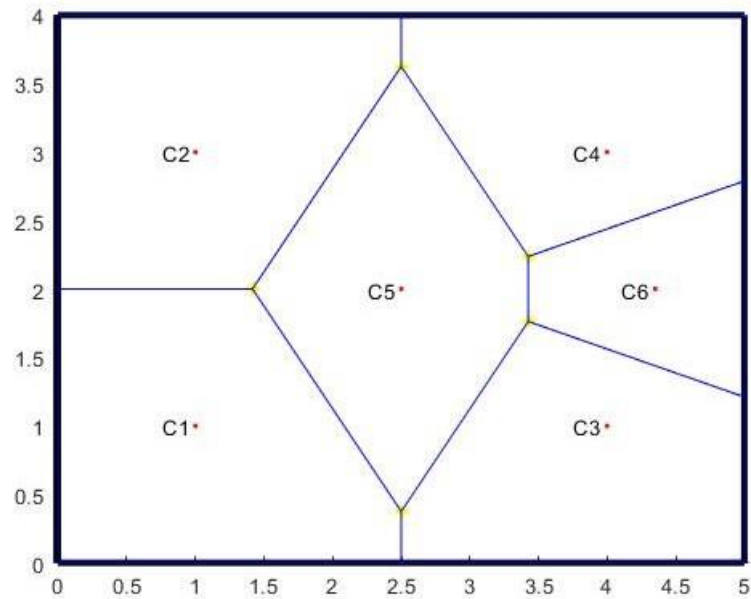


Figura 5.7: Ejemplo DV Interpolación Iterativa, Interpolación 2do punto.
Elaboración propia.

En este caso, se observa que la interpolación del 2do punto depende de los valores de las celdas C_3 , C_4 y C_5 . El aspecto importante de notar es que en este caso el valor de la celda C_6 se solapa sobre la celda C_5 , por lo tanto, su valor depende intrínsecamente de los valores en C_1 y C_2 con los que fue calculada la celda C_5 . Por el contrario, si se interpola en primera instancia la coordenada del punto C_6 , $(x_6, y_6) = (4.35, 2)$, sin antes haber interpolado el punto en las coordenadas $(x_5, y_5) = (2.5, 2)$, el valor de C_6 dependería solo de C_3 y C_4 , tal como se presentó en el caso original y como se observa en la figura 5.8:

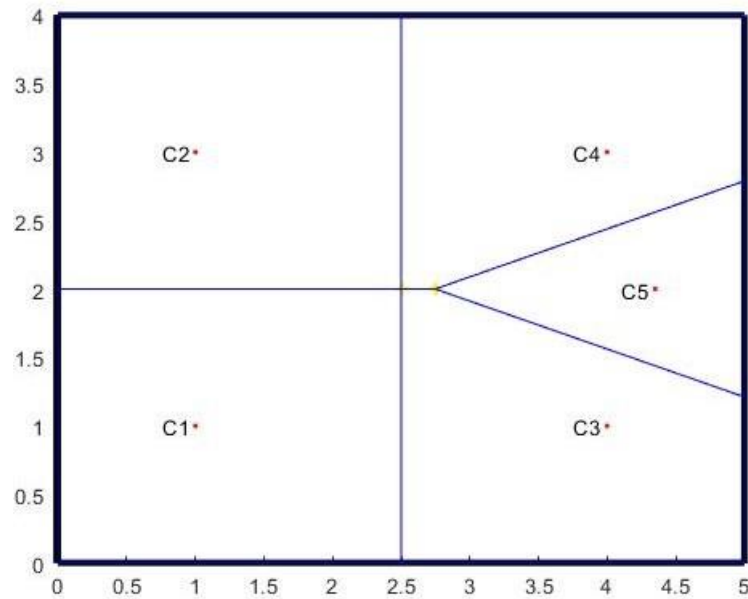


Figura 5.8: Ejemplo DV de cinco puntos en Interpolación Iterativa, aplicando un cambio de orden de interpolación del 2do punto. El punto C_5 corresponde al punto C_6 en $(x_6, y_6) = (4.35, 2)$. Elaboración propia.

Este mecanismo permite variar la configuración de correlación espacial representada con el método original, induciendo un efecto producto del orden de interpolación de los puntos. En el ejemplo presentado, se induce la correlación entre de los puntos C_1 y C_2 con el punto $(x_6, y_6) = (4.35, 2)$ por medio del punto central $(x_5, y_5) = (2.5, 2)$. Esto en la realidad podría darse por una serie de factores que generen una mayor correlación de entre los puntos C_1 y C_2 con el punto $(x_6, y_6) = (4.35, 2)$, como por ejemplo que el viento sople desde esas direcciones, generando una mayor correlación espacial en la dirección del viento, que el método de los vecinos naturales podría no representar en algunas circunstancias.

5.3. Modelo

En esta sección se presenta el modelo de dispersión en tiempo real desarrollado mediante la interpolación de los vecinos naturales y tres variantes iterativas del método. En primera instancia se presentan los módulos que componen el modelo, su función, entradas y salidas. Posteriormente, se presentan los datos utilizados y los mecanismos para evaluar el desempeño del modelo.

A continuación, en la figura 5.9 se observa la configuración que sigue el flujo de datos generados por una red de sensores inalámbricos. En primer lugar, la información generada por la red de sensores es enviada a la nube de ThingSpeak, que almacena la información enviada por los dispositivos de bajo costo. Posteriormente, desde un computador conectado a internet se puede acceder a los datos almacenados en la nube ThingSpeak, desde la plataforma de Matlab, en la cual son manipulados para la integración en un modelo de dispersión de contaminantes basado en interpolaciones. Finalmente, se obtiene un mapa de distribución de

concentraciones de contaminantes en el espacio.

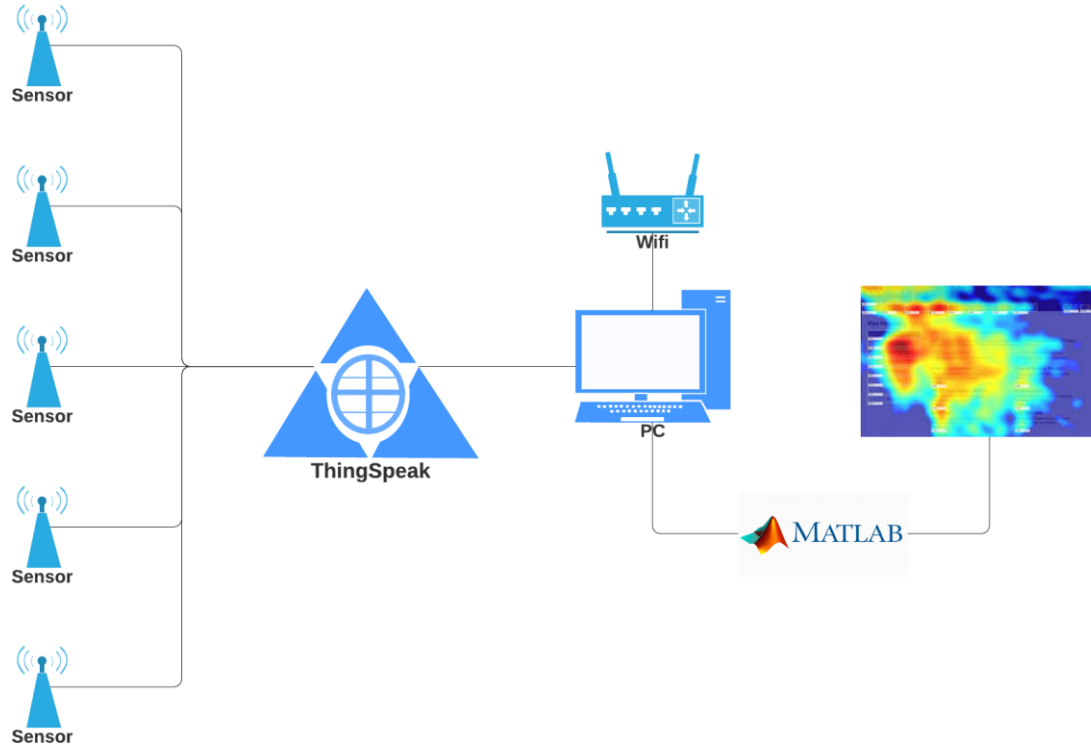


Figura 5.9: Diagrama funcionamiento modelo en tiempo real de dispersión de contaminantes. Elaboración propia.

El modelo de dispersión está constituido por tres módulos principales. El primero encargado de realizar los diagramas de Voronoi requeridos para interpolar cada punto del área de interés. El segundo, encargado del cálculo de los coeficientes de correlación para interpolar cada uno de los puntos, y el tercero, es el que ejecuta las interpolaciones de todos los puntos definidos en el plano. A continuación, se presenta cada de uno de estos módulos en mayor detalle.

1. Particiones de Voronoi: Este módulo realiza las $N+1$ particiones de Voronoi del plano, requeridas para el cálculo de los coeficientes de correlación para interpolar N puntos. El módulo genera una partición de Voronoi para las S ubicaciones monitoreo y una que por cada ubicación definida en una grilla de puntos, que representan las ubicaciones a estimar en el área de modelamiento. Con el objetivo de reducir el tiempo de ejecución del modelo en tiempo real, este módulo se ejecuta de manera offline, ya que, solo depende de variables fijas, como las ubicaciones de los equipos de monitoreo, el área de interpolación y el número de puntos a interpolar en el área definida.

Este módulo se sustenta en el programa “clippedVoronoi” realizado por Kantaros [84], que utiliza la función `Voronoin` de MATLAB y traza el diagrama de Voronoi derivado de los nodos con coordenadas $S = [Xnodes\ Ynodes]$, mientras delimita el resultado en

el área descrita por los vértices $\text{área} = [X_{\text{area}} Y_{\text{area}}]$. Esta función permite acceder a las coordenadas de las celdas de Voronoi generadas en el área definida. Posteriormente, una segunda función utiliza el algoritmo “clippedVoronoi” para generar los DV para cada uno de los puntos a interpolar, que se encuentran definidos a partir de una grilla equispaciada en el área de modelamiento.

La resolución de vecinos naturales *versus* la resolución iterativa, entre otras, se diferencian en que la primera desarrolla un diagrama de Voronoi de S celdas y N diagramas de Voronoi con S + 1 celdas, lo que corresponde a incluir cada punto de la grilla, por si solo en el diagrama de Voronoi de las ubicaciones iniciales en S. Por otro lado, en la resolución iterativa, por cada punto interpolado, el módulo genera un diagrama de Voronoi con una celda extra, de manera que la ultima iteración genera un diagrama de Voronoi de S + N celdas, una celda para ubicación interpolada, más las S celdas de las ubicaciones de monitoreo.

De esta forma, el módulo genera una matriz de múltiples páginas, en cada página guarda las coordenadas de las celdas de Voronoi generadas en cada iteración. El método de VN e iterativo se diferencian en esta sección, en que el primero, cada página de la matriz contiene la información de un diagrama de Voronoi de S+1 puntos, que corresponden a las coordenadas iniciales de monitoreo más el punto de interpolación. Por otro lado, el método iterativo, en cada iteración conserva el punto calculado en la iteración anterior y sobre el diagrama de la iteración anterior interpola el nuevo punto. De esta forma, en cada interacción el diagrama de Voronoi generado contiene una celda extra al generado en la iteración anterior. En la figura 5.10 se presenta un diagrama de funcionamiento de este algoritmo.

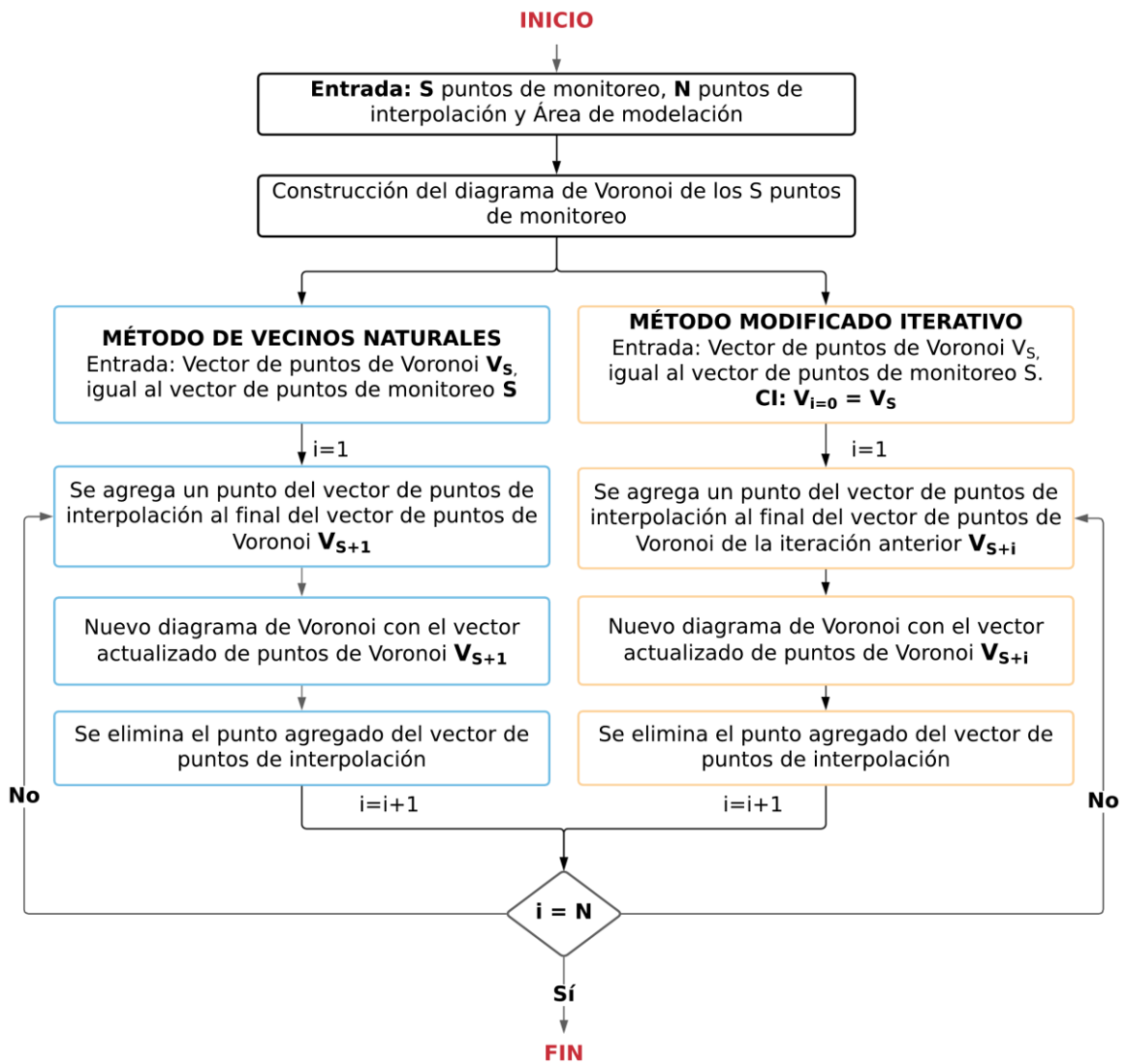


Figura 5.10: Diagrama algoritmo modulo que realiza las particiones de Voronoi necesarias para implementar el método. Elaboración propia.

2. Cálculo coeficientes de interpolación: Para aumentar la rapidez del modelo, este módulo también se realiza de manera offline, ya que al igual que en el módulo anterior, los coeficientes solo dependen de las ubicaciones de los puntos de monitoreo en un área predefinida. Ubicaciones, que en este caso se fijan y no son móviles.

Esta función recorre la matriz generada anteriormente que contiene la información de las celdas de Voronoi de los DV de todos los puntos de interpolación. En el caso de los vecinos naturales, interseca la ultima celda de Voronoi de cada pagina de la matriz generada por el punto de interés, con todas las celdas del diagrama de Voronoi de los puntos de monitoreo iniciales. Por otro lado, en la resolución iterativa, se interseca la ultima celda de Voronoi de cada pagina de la matriz con todas las celdas del diagrama de Voronoi de la iteración o pagina anterior. De esta forma, se calculan los coeficientes

de correlación de cada uno de los puntos de interpolación, los cuales dependen solo de las ubicaciones de los equipos de monitoreo en el caso de los vecinos naturales y, en la resolución iterativa, además, dependen del orden de interpolación de los puntos. En la figura 5.11, se presenta un esquema del algoritmo de este módulo.

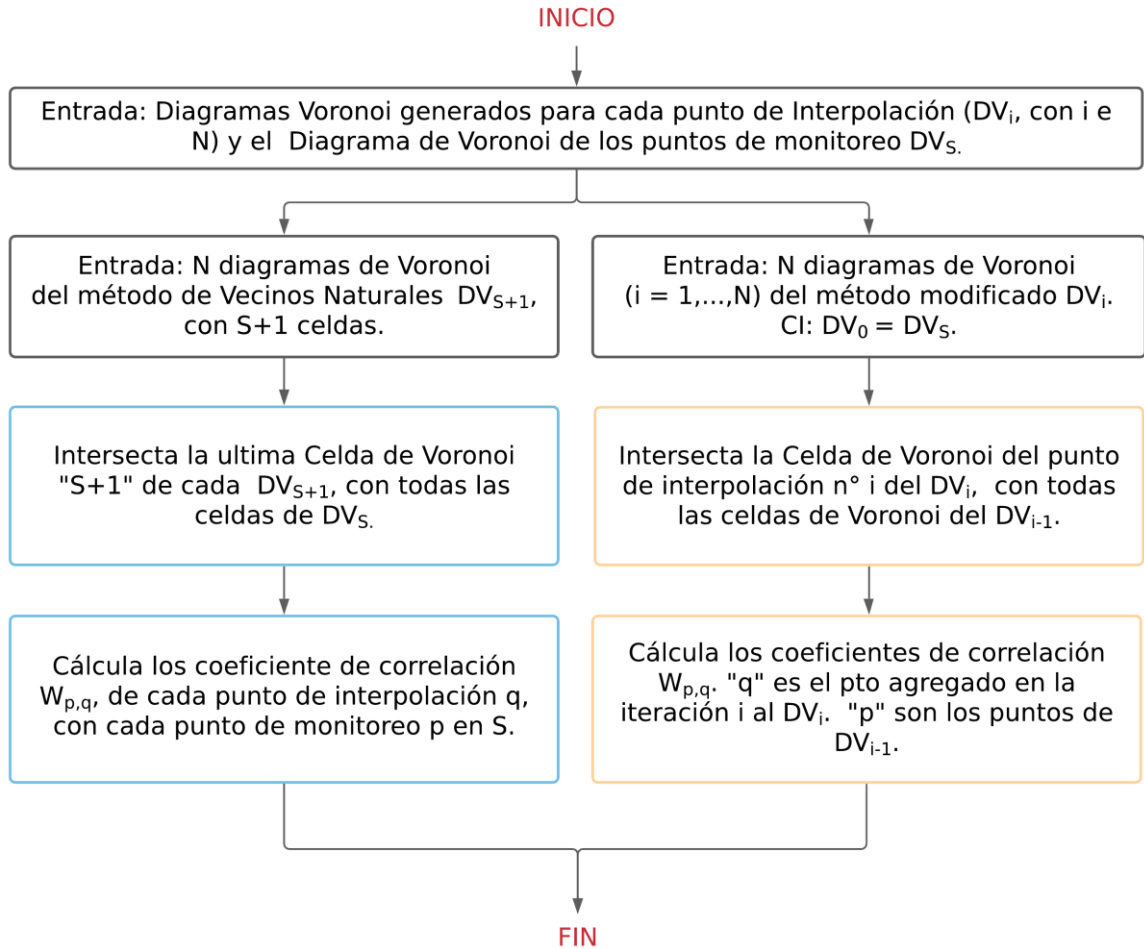


Figura 5.11: Diagrama algoritmo módulo que calcula los coeficientes de correlación para cada punto de interpolación definido en la grilla de puntos. Elaboración propia.

3. Interpolación en tiempo real: Este módulo accede a la información enviada por los sensores a la nube ThingSpeak. A partir de la media de las concentraciones registradas por cada sensor en los últimos 5 minutos, se interpolan todos los puntos del plano definidos anteriormente, accediendo a los coeficientes calculados en el modulo anterior.

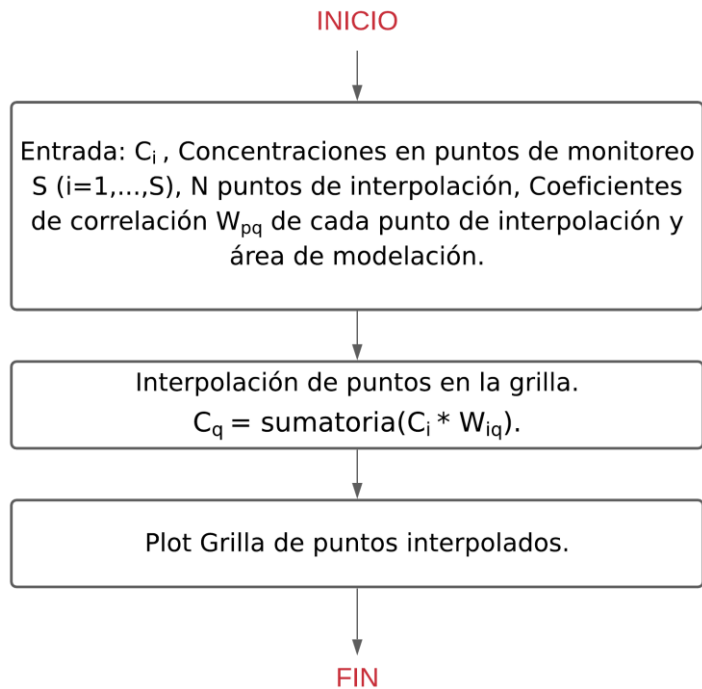


Figura 5.12: Diagrama algoritmo módulo que calcula las interpolaciones de los puntos definidos en la grilla de puntos. Elaboración propia.

Capítulo 6

Resultados y validación del modelo

En esta sección, se presentan los resultados obtenidos en la implementación del modelo de dispersión en tiempo real. Cabe mencionar, que los sensores de bajo costo utilizados en este estudio no se encontraban calibrados al momento de la validación, razón por la cual, el desempeño del modelo se analizó con información validada de las estaciones de monitoreo del Estado de la Región Metropolitana. Los sensores de bajo costo solo fueron utilizados para analizar la implementación del modelo con adquisición de datos en tiempo real.

En primera instancia, se exponen los datos utilizados y objetivos de los test realizados para la validación del modelo. Luego, se presentan los mapas de las ubicaciones de las estaciones utilizadas en ambos set de datos, junto a los diagramas de Voronoi de las ubicaciones de monitoreo. Posteriormente se presentan las principales gráficas obtenidas y resultados numéricos de los errores entre las estimaciones y los valores observados.

Para evaluar el desempeño del modelo se utilizan dos set de datos de las estaciones de monitoreo del Estado disponibles en la página del SINCA. El primer set de datos corresponde a la información de material particulado de dos días de la temporada de verano y dos días de invierno del año 2015, año que presenta mayor número de estaciones de monitoreo con información de $MP_{2.5}$ en la Región Metropolitana. El segundo set de datos corresponde a la información de material particulado de un mes de la temporada de verano y un mes de la temporada de invierno del año 2021.

Mediante estos set de datos se realizan tres test, que evalúan el modelo mediante los mecanismos de validación cruzada *K-Fold* y *Leave-One-Out*. El primer test tiene múltiples objetivos, en primer lugar, analizar la representación del ciclo diario de un contaminante, observando la semejanza de las concentraciones estimadas y observadas. Esta semejanza se analiza mediante el cálculo de las métricas de error entre las concentraciones reales y estimadas en los puntos de validación y, además, mediante la superposición de las curvas reales y estimadas en los puntos de validación.

Junto a lo anterior, el primer test tiene por objetivo analizar el costo computacional del método implementado y sus variaciones, ya que una condición de diseño es su bajo costo computacional para una rápida respuesta. Además, en este test, se analiza la resiliencia del modelo ante la pérdida de información de un equipo de monitoreo, producto de alguna falla en la medición, envío y recepción de datos.

El segundo test, tiene por objetivo, comparar con mayor robustez las diferencias entre las estimaciones y los valores observados en las temporadas de invierno y verano, para el método de los vecinos naturales y sus variaciones. Para lo anterior, se utiliza un mayor conjunto de datos, de manera de observar diferencias significativas entre las temporadas de invierno y verano, reduciendo el sesgo de sus resultados. Junto a lo anterior, tiene el objetivo de identificar con mayor precisión los puntos donde el modelo no logra representar adecuadamente la realidad, además, de identificar con mayor certeza la tendencia a la sub o sobre estimación del método.

Por último, el tercer test, tiene por objetivo principal analizar el desempeño del modelo en situaciones de interpolación y extrapolación. Para lo anterior, en este test se aplica el mecanismo de validación cruzada *Leave-One-Out*, en el que se analiza el error del modelo al interpolar todos los puntos del plano excluyendo los datos solo de una estación de monitoreo, la cual será utilizada para el cálculo de las métricas de error. Se realiza este procedimiento hasta excluir todas las estaciones por sí sola y estimar las concentraciones en su ubicación, cuando esta fue excluida. De esta forma, se podrá comparar el error entre la estimación y los datos observados en cada punto de muestreo, cuando su valor fue estimado a partir de la información en los demás puntos de monitoreo.

Puntos de Monitoreo y Validación

A continuación, en la figura 6.1 se presenta la ubicación de las estaciones de monitoreo correspondientes al primer set de datos. Los puntos en rojo corresponden a las estaciones utilizadas para la interpolación de todos los puntos del plano y los puntos en azul corresponden a las estaciones utilizadas para analizar el error entre la estimación y la observación. Luego, en la figura 6.2, se presenta el diagrama de Voronoi inicial a partir de las ubicaciones de las estaciones monitoreo utilizadas para interpolar la grilla de puntos.



Figura 6.1: Mapa estaciones de monitoreo de calidad del Aire Región Metropolitana. En rojo estaciones utilizadas para la interpolación y en azul estaciones utilizadas para validación. Elaboración propia.

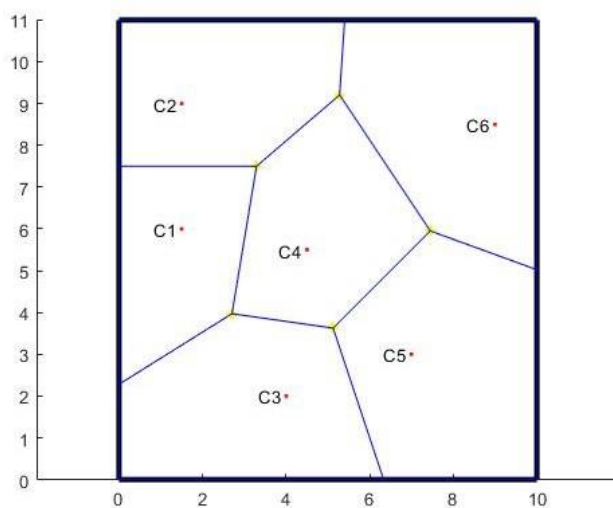


Figura 6.2: Diagrama de Voronoi de ubicaciones de estaciones del SINCA utilizadas para la interpolación. Elaboración propia.

En la figura 6.3 se presenta la ubicación de las estaciones de monitoreo correspondientes al segundo set de datos. Los puntos en rojo corresponden a las estaciones utilizadas para la interpolación de todos los puntos del plano y los puntos en azul corresponden a las estaciones utilizadas para analizar el error entre la estimación y la observación. Luego, en la figura 6.4 se presenta el diagrama de Voronoi inicial a partir de las ubicaciones de estaciones monitoreo utilizadas para interpolar la grilla de puntos en el segundo set de datos.

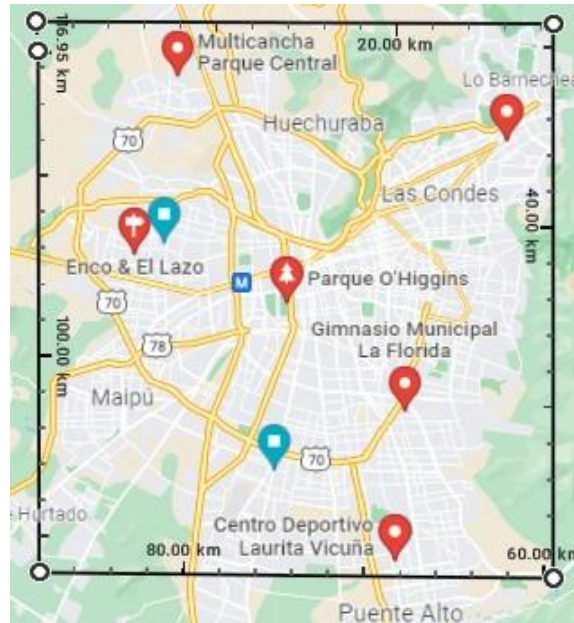


Figura 6.3: Mapa de estaciones de monitoreo de calidad del Aire Región Metropolitana con información en línea de $MP_{2.5}$. Elaboración propia.

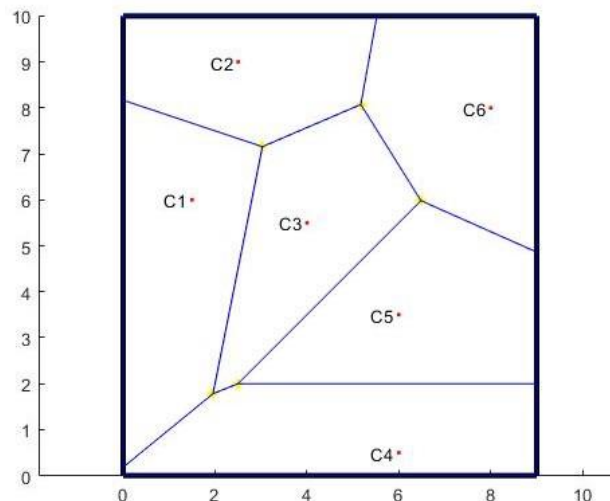


Figura 6.4: Diagrama de Voronoi de ubicaciones de estaciones del SINCA utilizadas para la interpolación en test 2. Elaboración propia.

Visualización modelo dispersión

A continuación, se presentan una serie de gráficas de los resultados obtenidos de la interpolación de los datos del periodo 2015. En primer lugar, en la figura 6.5 se presentan los resultados obtenidos de la interpolación de todos los puntos del plano definidos en la grilla a partir de la información obtenida en seis ubicaciones en la figura 6.1. Los resultados se presentan mediante el mapa de calor de la concentración de material particulado en la Región Metropolitana del día 2 de Julio del 2015 a las 23:00 hrs. Esta visualización presenta una notoria mejoría respecto a la información de calidad del aire presente en el SINCA, que corresponde a una serie de tiempo de la concentración en una ubicación puntual, como se puede observar en la figura 3.1. En este sentido, el modelo se posiciona como una excelente alternativa para aumentar la información espacial de la calidad del aire, pero además, presentar los resultados de manera simple y amigable, permitiendo reconocer fácilmente sectores con altos niveles de contaminación.

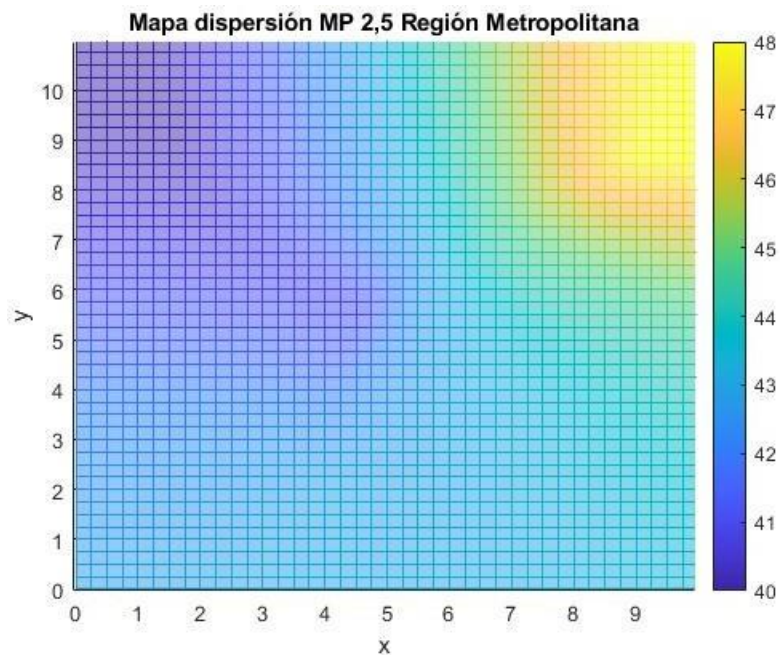


Figura 6.5: Mapa de dispersión de concentración de $MP_{2,5}$ en Región la Metropolitana periodo Invierno

6.1. Test 1

A continuación, se presentan los resultados obtenidos del test realizado con el primer set de datos. En primer lugar, se presentan los objetivos de este test, luego se presentan las concentraciones medias registradas por cada estación en el intervalo de medición. Posteriormente, se presentan las series de tiempo de las concentraciones estimadas y observadas para cada una de las tres estaciones de monitoreo utilizados en la validación cruzada de los resultados. Finalmente, se presentan las métricas de error promedio calculadas al comparar las diferencias entre los valores estimados y observados de las tres estaciones de validación.

En este test, se utiliza el primer set de datos de material particulado de nueve estaciones de monitoreo en la temporada de invierno y verano. Para la temporada de invierno se utiliza la información comprendida entre el 01/07/2015 y el 02/07/2015, por otro lado, para la temporada de verano se utiliza la información comprendida entre el 01/01/2015 y el 02/01/2015. Se aplica el mecanismo de validación cruzada *K-Fold*, que utiliza los datos de seis estaciones para realizar las interpolaciones y la información de las tres restantes para el análisis de error entre la predicción y el valor observado, ver figura 6.1. Posteriormente, se realiza un análisis de resiliencia del modelo, en casos donde los equipos de monitoreo fallen y dejen de enviar información desde su ubicación de monitoreo, para lo que se analizan las métricas de error obtenidas al desprestigiar las mediciones de una de las seis estaciones utilizadas para la interpolación y se comparan las métricas de error de cuando los valores excluidos si fueron considerados.

En este test, se comparan los resultados obtenidos por la interpolación de los vecinos naturales y tres casos (A, B, C) del algoritmo modificado, variando el orden en que se interpolan los puntos de la grilla. En los casos A y B se interpolan los puntos aleatoriamente y en el C, en un orden predefinido, interpolando los puntos en orden, recorriendo fila a fila la grilla de puntos. El algoritmo modificado busca observar alguna configuración óptima, que reduzca el error entre la estimación y el valor observado.

Resultados Verano Test 1

En la tabla 6.1, se presentan los promedios de las concentraciones registradas por cada estación de interpolación en el intervalo de medición de la temporada de verano. En la tabla 6.2 se presentan los promedios de concentración registrados en las estaciones de validación. Se observa que los máximos y mínimos registros medios se presentaron en las estaciones de validación de Cerrillos y Cerro Navia respectivamente. Las medias corresponden a los datos de las estaciones de la figura 6.1.

Tabla 6.1: Concentraciones medias registradas en estaciones de monitoreo utilizadas para interpolación. Elaboración propia.

Estaciones	Pudahuel	Quilicura	El Bosque	Parque O.	La Florida	Las Condes
\bar{C} (ug/m ³)	22,3	20,8	22,8	23,4	17,2	19,0

Tabla 6.2: Concentraciones medias registradas en estaciones de monitoreo utilizadas para validación. Elaboración propia.

Estaciones Validación	Cerro Navia	Cerrillos	Puente Alto
\bar{C} ($\mu\text{g}/\text{m}^3$)	14,3	28,1	21,4

A continuación, se presentan las series de tiempo de las concentraciones estimadas y observadas en los puntos de validación. Se observan las gráficas de las curvas estimadas y observadas para cada estación y método de interpolación. En primer lugar, se presentan los resultados del método de vecinos naturales y luego, las tres versiones A, B y C del algoritmo iterativo. En las gráficas, en azul se observan las curvas de datos observados, mientras que en naranja se observan las curvas de datos estimados.

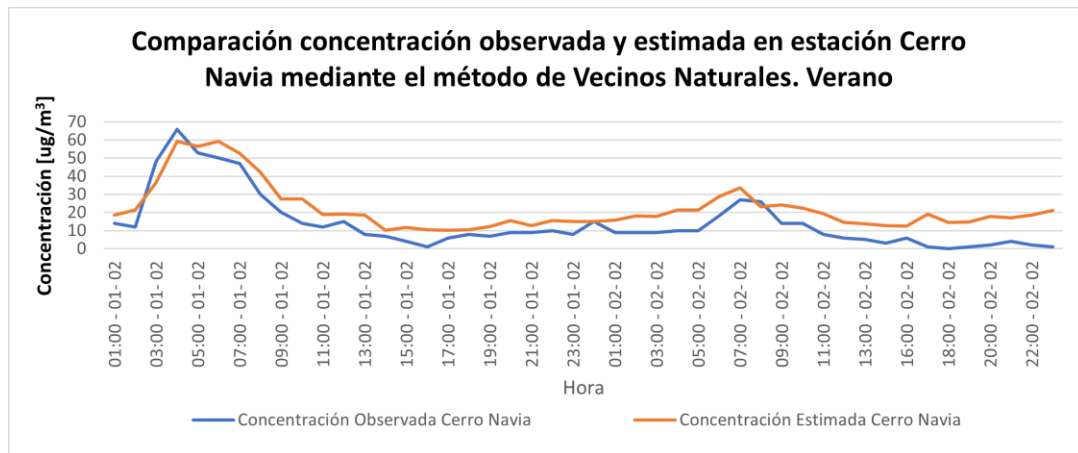


Figura 6.6: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia IVN.

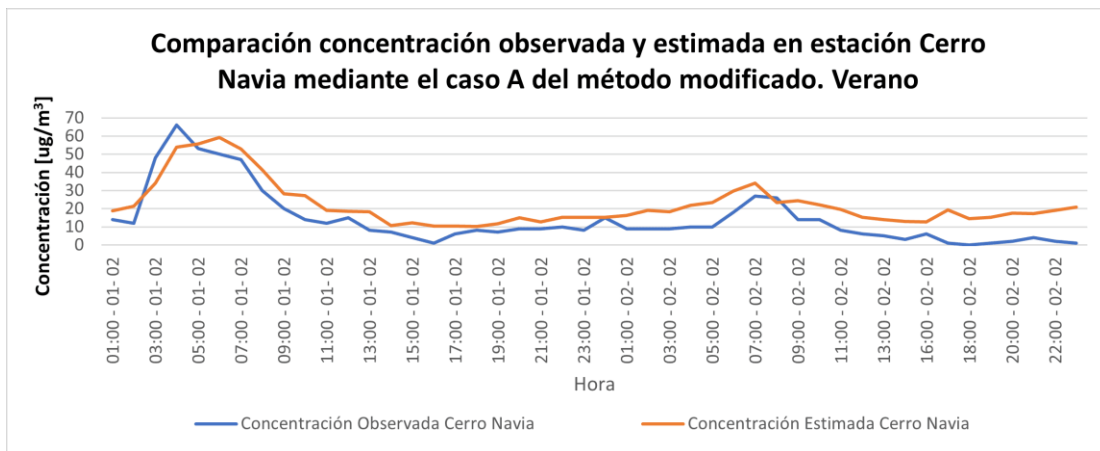


Figura 6.7: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia. Caso A.

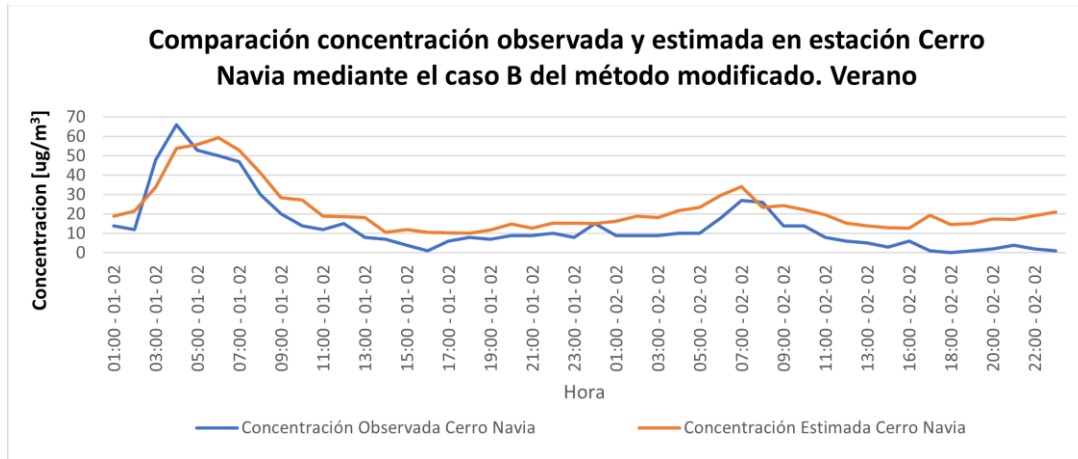


Figura 6.8: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia Caso B.

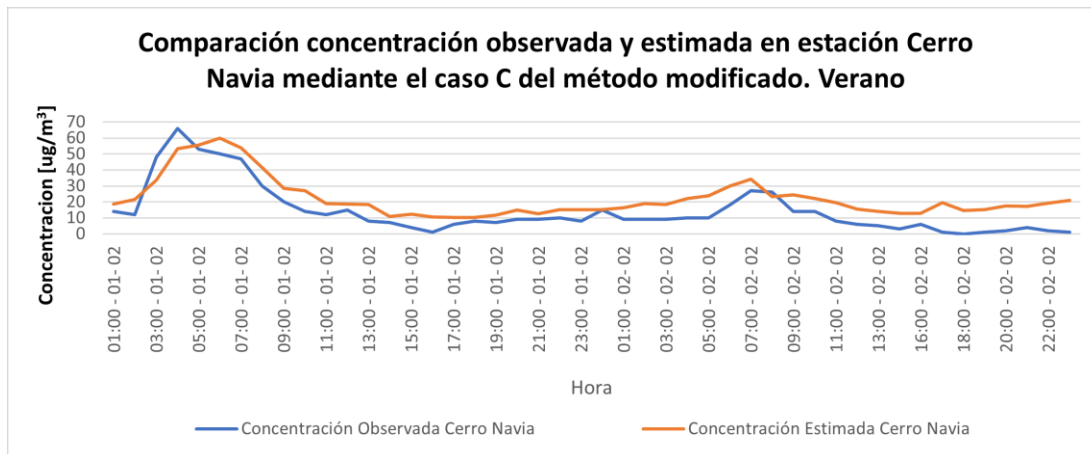


Figura 6.9: Comparación Concentraciones reales e interpoladas en ubicación de estación Cerro Navia. Temporada Invierno, Caso C.

En las figuras 6.6, 6.7, 6.8 y 6.9, se observa que visualmente los resultados obtenidos por el método de Vecinos naturales e iterativo (A, B y C) presentan resultados bastante similares entre sí, ya que no se observan diferencias significativas entre los métodos. Producto de lo anterior, en adelante solo se presentan las gráficas del método de vecinos naturales para cada estación de validación. El resto de las gráficas se presentan en el anexo G.

Por otro lado, se observa que la curva de datos estimados representa el comportamiento y tendencias diarias de las concentraciones observadas. Es claro como la curva de datos estimados presenta sus máximos y mínimos en los mismos momentos que la curva de datos observados.

Lo anterior, se debe a que las estimaciones se realizan con una imagen real de los niveles de contaminación registrada por las estaciones de monitoreo. Estos niveles de concentración se encuentran condicionados por diferentes variables fisicoquímicas que dominan los niveles diarios de concentración, por ejemplo, la radiación solar en caso de reacciones fotoquímicas, que puede determinar momentos de máxima o mínima concentración

producto de la generación o consumo del contaminante, o el viento, que induce momentos de mayor dispersión y dilución de contaminantes en la atmósfera. Estas variables, presentan diversos patrones diarios, con niveles máximos y mínimos durante el día. Por lo tanto, los contaminantes regidos por estas variables, también presentaran momentos de máxima y mínima concentración, por ejemplo, el ciclo diario de contaminantes de la ciudad Shenzhen en China, que se presenta en la figura 2.7, donde se pueden observar claramente los perfiles diarios de varios contaminantes, condicionados por el efecto conjunto de estas variables.

Pese a que el modelo logra representar el ciclo diario de contaminantes, las estimaciones en la estación de Cerro Navia se encuentran sobre estimadas. La sobre estimación de los valores observados ocurre porque los valores reales en el punto de monitoreo se encuentran fuera del intervalo de datos utilizados en la interpolación, ya que en la estación de Cerro Navia se presentaron los menores registros medios. Por lo tanto, es congruente que los valores estimados en esta ubicación sean mayores a los observados, ya que fueron interpolados con valores medios mayores a los del punto de validación en Cerro Navia.

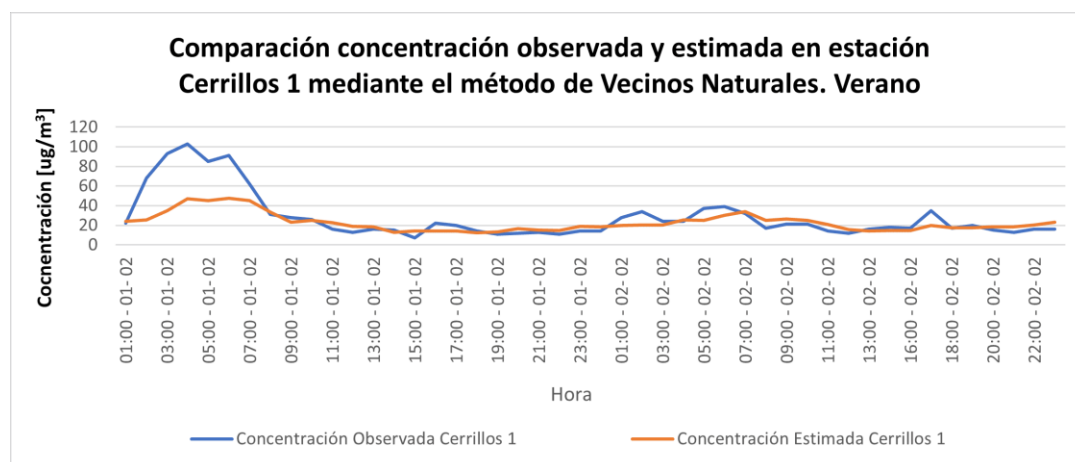


Figura 6.10: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación de VN. Temporada Verano.

Respecto a los resultados obtenidos en la estación de Cerrillos presentes en las figuras 6.10, G.1, G.2 y G.3, nuevamente se observa que los resultados obtenidos por el método de VN e iterativo (A, B y C) no presentan diferencias significativas. Junto a lo anterior, nuevamente se observa que las estimaciones representan adecuadamente el comportamiento y tendencias diarias de las concentraciones de $MP_{2,5}$ observadas en prácticamente todo el dominio temporal, a excepción de los puntos iniciales, en los cuales los valores observados son notoriamente mayores a las demás concentraciones registradas en la estación Cerrillos.

Se puede observar que los niveles de $MP_{2,5}$ en la sección inicial son del orden de los 100[ug/m³], muy por encima de las medias registradas por las estaciones utilizadas en la interpolación del orden de los 20 [ug/m³], como se puede observar en la tabla 6.1. Además, se observa que la estación de Cerrillos presenta la máxima media de las estaciones de monitoreo, en consecuencia, como el método se sustenta en interpolaciones, los registros de las estaciones utilizadas en la interpolación son incapaces de estimar valores por fuera de su intervalo de monitoreo.

Por otro lado, si bien los mayores registros medios se presentaron en la estación de Cerrillos, estos se deben a las emisiones puntuales al inicio del intervalo de monitoreo, ya que posterior a los puntos iniciales, la curva de datos estimado representa verosímilmente la curva de datos observados. Este aspecto denota la importancia de contar una red distribuida, con registros en las ubicaciones donde se puedan presentar las mayores o menores niveles de concentración del contaminante de interés, ya que si se hubiera utilizado la información de la estación de Cerrillos, los valores interpolados se hubieran encontrado dentro del intervalo de medición.

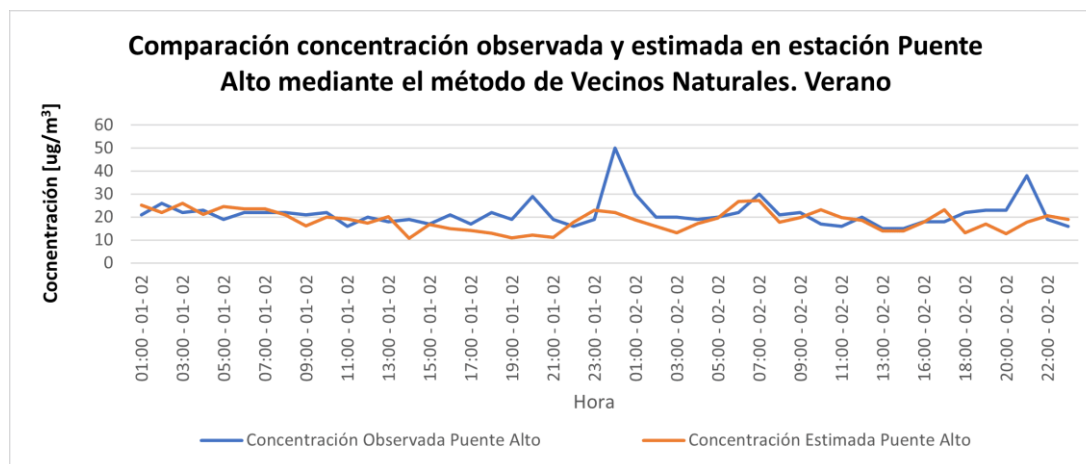


Figura 6.11: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto IVN.

En la figura 6.11 nuevamente se observa que las estimaciones representan adecuadamente el comportamiento y tendencias diarias de las concentraciones observadas. Sin embargo, nuevamente se observa que el modelo no logra representar los picos máximos de la curva de datos reales. Esto, al igual que lo observado anteriormente, se debe a que dichos valores se encuentran fuera del rango de medición de las estaciones utilizadas en la interpolación.

Respecto al análisis de las series de tiempo entre la estimación y el valor real observado, se concluye que en la temporada de verano en general los valores estimados representan el comportamiento diario de los datos observados en la estación de monitoreo utilizada para la validación, identificando los periodos de máxima y mínima concentración. Sin embargo, no se comporta adecuadamente al cuantificar dichos niveles máximos o mínimos, cuando los valores se presentan por fuera de los rangos de medición, como lo ocurrido en la estación de Cerrillos. Además, se observa que en la estación de Cerro Navia, la predicción representa de mejor manera la tendencia de los valores reales, esto se debe a que la estación de Cerro Navia se encuentra dentro del intervalo de monitoreo y, por tanto, sus predicciones relatan de mejor manera la variación de la dinámica espacial, al ser calculada un mayor número de estaciones colindantes.

Resultados Invierno Test 1

A continuación, se presentan los resultados obtenidos para la temporada de invierno. En la tabla 6.3 se presentan los promedios de las concentraciones medidas en cada estación monitoreo utilizada en la interpolación y en la tabla 6.4, se presentan los promedios de concentración registrados en las estaciones de validación para el periodo Invierno. En este caso, se observa que el mayor y menor registro medio se presentó en las estaciones de interpolación de Pudahuel y las Condes respectivamente. Las medias de las estaciones presentes en las tablas 6.3 y 6.4, corresponden a datos registrados en las estaciones de monitoreo de la figura 6.1.

Tabla 6.3: Concentraciones medias de las estaciones utilizadas para interpolación del primer set de datos. Temporada Invierno. Elaboración propia.

Estaciones	Pudahuel	Quilicura	El Bosque	Parque O.	La Florida	Las Condes
\bar{C} ($\mu\text{g}/\text{m}^3$)	95,4	83,9	78,0	82,4	76,8	55,3

Tabla 6.4: Concentraciones medias estaciones utilizadas para la validación. Temporada Invierno. Elaboración propia.

Estaciones Validación	Cerro Navia	Cerrillos	Puente Alto
\bar{C} ($\mu\text{g}/\text{m}^3$)	91,9	91,2	64,5

A continuación, se presentan las series de tiempo de las concentraciones observadas y estimadas en los puntos de validación para la temporada de invierno. Al igual que en la temporada de verano, en azul se presentan las curvas de datos observadas y naranja las curvas de concentraciones estimadas. Se presentan los gráficos por estación de cada método analizado.

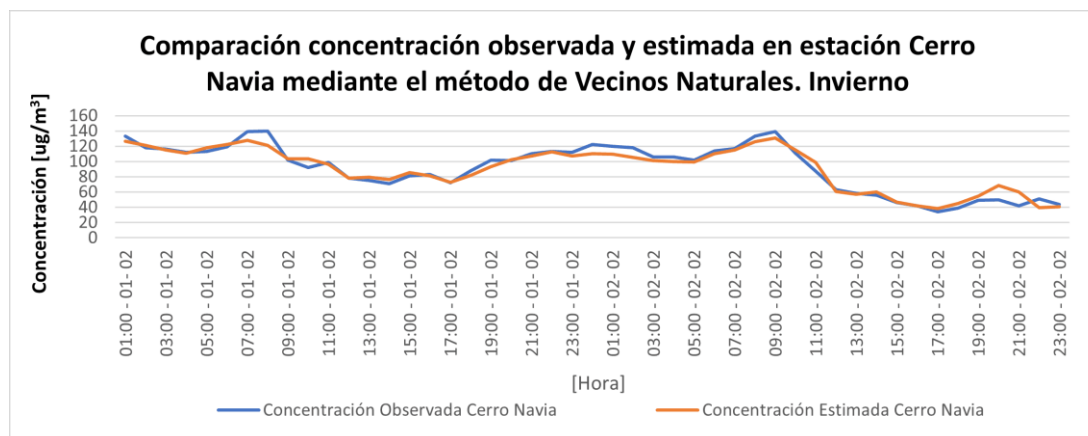


Figura 6.12: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia IVN.

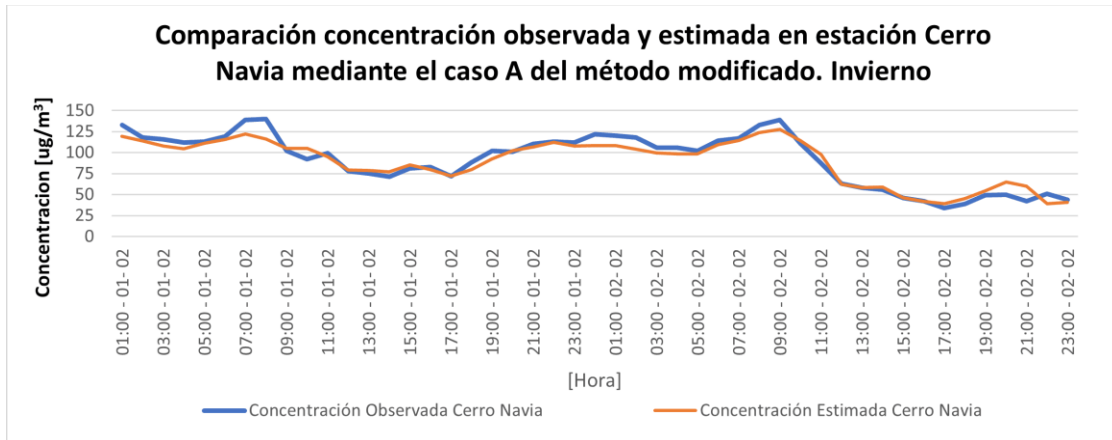


Figura 6.13: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia Al1.

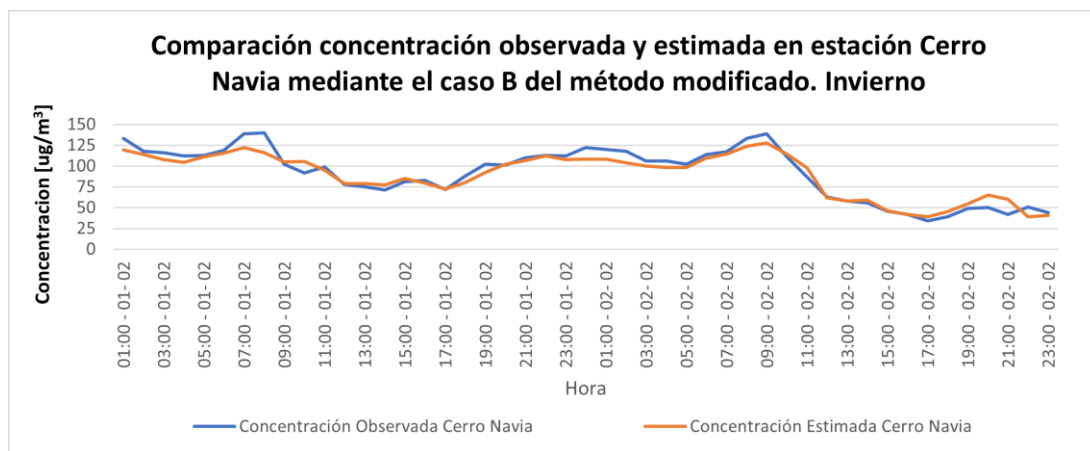


Figura 6.14: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia Caso B.

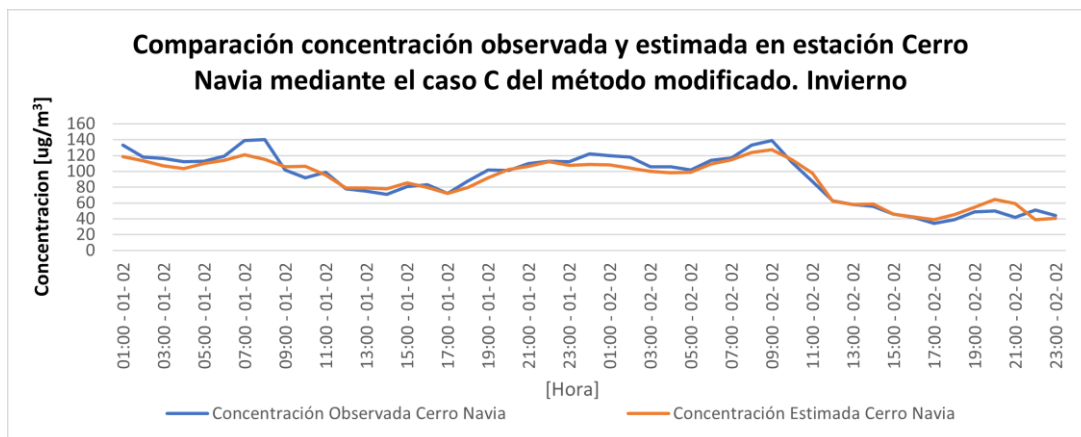


Figura 6.15: Comparación Concentraciones reales e interpoladas en ubicación de estación Cerro Navia. Temporada Invierno, Caso C.

En las figuras 6.12, 6.13, 6.14 y 6.15, nuevamente se observa que visualmente los resultados

obtenidos por el método de V.N e iterativo (A, B y C) presentan resultados bastante similares entre sí. Producto de lo anterior, en adelante, solo se presentaran las gráficas del método de vecinos naturales para cada estación de validación. En el anexo G se presentan la totalidad de las gráficas por estación.

En las figuras 6.12, 6.13, 6.14 y 6.15, nuevamente se observa que las estimaciones representan adecuadamente el ciclo diario de material particulado, identificando con bastante verosimilitud los momentos en que presentan los picos y valles de la curva de datos observados. Finalmente, no se observa una clara tendencia a la sobre o sub estimación de las concentraciones, esto se debe a que el punto de estimación de Cerro Navia se encuentra dentro del área de monitoreo y además, sus registros medios se encuentran dentro del intervalo de medición de las estaciones circundantes. Por lo tanto, su estimación corresponde a un problema de interpolación en el que el método de vecinos naturales se comporta adecuadamente.

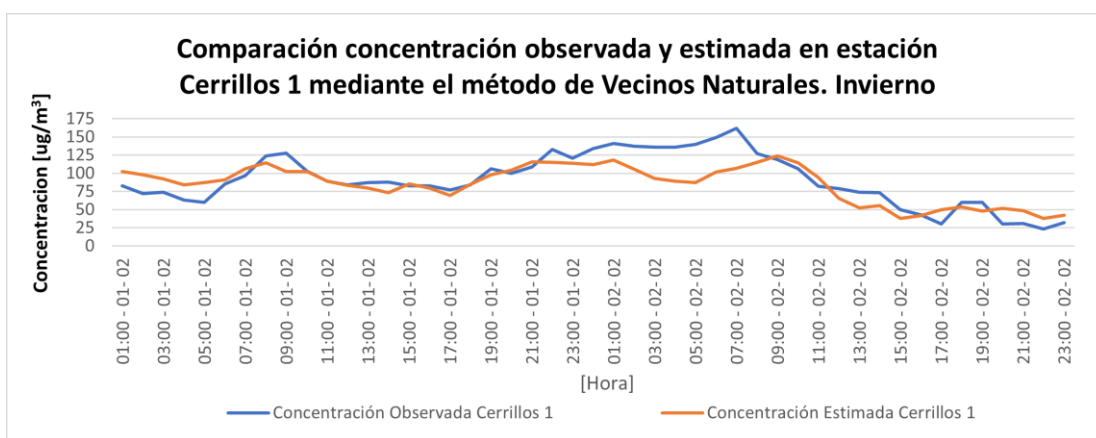


Figura 6.16: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. VN. Invierno.

En este caso, nuevamente se observa que el modelo no logra representar la zona donde se presenta el mayor peak de medición, que presenta un valor del orden de los 150[ug/m³]. Se presume que esto debe tener su origen en los altos registros medios presentados en la estación de Cerrillos, con un valor de 91,2 [ug/m³], solamente superados por los 95,4 [ug/m³] en la estación de Pudahuel. Por lo tanto, en el momento que se presentan los máximos registros observados en el punto de estimación, las estaciones circundantes de interpolación debieron presentar menores concentraciones que la registrada en la estación de Cerrillos 1. Por lo tanto, los niveles estimados en ese momento se encuentran sub estimados.

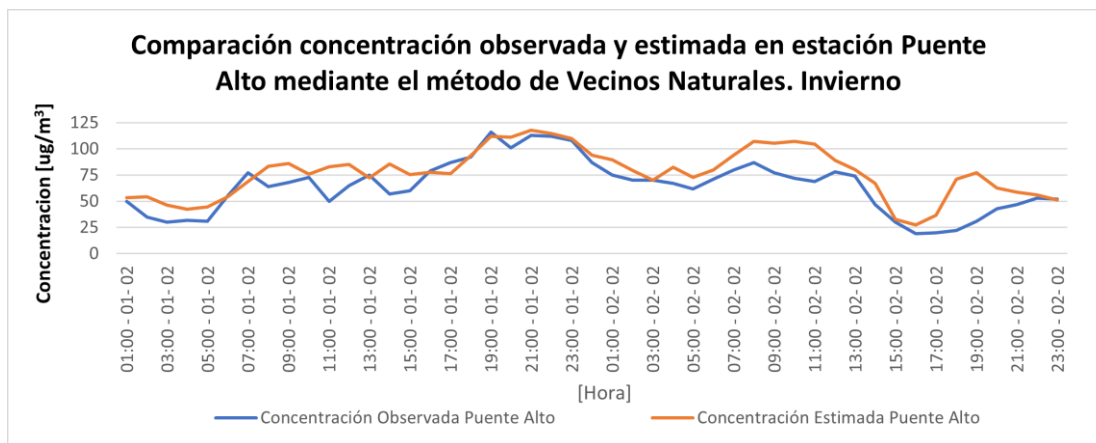


Figura 6.17: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto IVN.

En la figura 6.17, se observa que en este caso las estimaciones no representan la totalidad de picos y valles del ciclo diario de datos observados. Junto a lo anterior, se observa que las estimaciones en la estación Puente Alto se encuentran sobre estimadas con respecto a las concentraciones observadas. Esta sobre estimación se debe a la baja concentración media en la estación de Puente Alto, de hecho, solo la estación de las Condes presenta un menor registro medio. Sumado a lo anterior, la estación de Puente Alto se encuentra en la periferia del área de monitoreo, por lo tanto, su estimación implica un problema de extrapolación y en estos casos, el método de los vecinos naturales no se comporta tan bien, ya que no cuenta con estaciones colindantes para representar adecuadamente la variación espacial de contaminantes.

A diferencia de la temporada de verano, en la temporada de invierno se observa la presencia de menor cantidad de grandes peak en las series de tiempo, producto de esto, las estimaciones presentan mejores resultados visuales. Este aspecto, se puede observar en la superposición de las rectas de valores observados y estimados, las que presentan una mayor suavidad a la observada en la temporada de verano. Por otro lado, en este caso, nuevamente la estación de Cerro Navia presenta los mejores resultados visuales, producto que esta estación se encuentra dentro del área de monitoreo y en consecuencia, su estimación considera el efecto de mayor numero de estaciones colindantes que permiten representar de mejor manera la variación espacial de contaminantes.

Resultados Numéricos Temporada Invierno y Verano Test 1

A continuación se analiza numéricamente el desempeño y diferencias de los métodos desarrollados, mediante el calculo del ME, MAE, RMSE y PE. Las métricas de error fueron calculadas contrastando los valores observados y estimados de cada hora en cada estación de validación. En la tablas 6.5 y 6.6, se presentan las métricas de error promedio de las tres estaciones de validación para el periodo de verano e invierno respectivamente.

Tabla 6.5: Resultados métricas de error media obtenidas en la temporada de verano, entre las concentraciones estimadas y observadas en estaciones de validación, para el método de los vecinos naturales y el caso A, B y C del algoritmo modificado en la temporada de verano.

Modelo	ME	MAE	RMSE	PE(%)
I. Vecinos Naturales	-0,17	7,92	12,16	57,1
A) Iterativo aleatorio 1	-0,19	7,95	12,19	57,3
B) Iterativo aleatorio 2	-0,19	7,95	12,19	57,3
C) Iterativo Orden grilla	-0,08	7,96	12,28	57,7

Tabla 6.6: Resultados métricas de error media obtenidas en la temporada de invierno, entre las concentración estimadas y observadas en las estaciones de validación, para el método de los vecinos naturales y el caso A, B y C del algoritmo modificado.

Modelo	ME	MAE	RMSE	PE(%)
I. Vecinos Naturales	1,98	11,9	16,74	20,3
A) Iterativo aleatorio 1	1,55	12,1	16,59	20,1
B) Iterativo aleatorio 2	1,55	12,1	16,59	20,1
C) Iterativo Orden grilla	1,36	12,0	16,46	19,9

Del calculo de las métricas de error entre las concentraciones estimadas y observadas en las estaciones de Cerro Navia, Cerrillos y Puente alto, para la temporada de verano que se presentan en la tabla 6.5, se puede concluir que los resultados generados por el algoritmo de vecinos naturales y los tres casos de la resolución iterativa no presentan una diferencia significativa entre ellos. En términos numéricos, presentan una diferencia del 0,6 % entre el método de los vecinos naturales y el caso C del algoritmo iterativo. La mayor diferencia de error se presento en el error medio, donde el caso C del método modificado presento un valor de -0,08 [$\mu\text{g}/\text{m}^3$], prácticamente la mitad del error medio presentado por los demás casos. Las demás métricas de error presentaron su menor valor en la interpolación de vecinos naturales, sin embargo, su diferencia con respecto al valor obtenido en los demás casos es menor al 1 %, por lo cual, se puede decir que no existen diferencias significativas entre las variaciones del método analizadas.

De las métricas de error de la temporada de invierno, calculadas de la diferencia entre las concentraciones estimadas y observadas en las estaciones de Cerro Navia, Cerrillos y Puente alto, que se presentan en la tabla 6.6, se observa al igual que en la temporada de verano, no se presentan diferencias significativas entre los casos analizados. Además, nuevamente la mayor diferencia de error se presento en el error medio, donde el caso C de la interpolación iterativa presento nuevamente los mejores resultados, con un valor de 1,35 [$\mu\text{g}/\text{m}^3$], aproximadamente un 30 % menor al mayor error medio presentado en el caso de la interpolación de vecinos naturales. Las demás métricas de error presentan una diferencia del orden del 2 % entre los cuatro casos analizados, por lo tanto, se puede concluir que no existen diferencias significativas entre los métodos estudiados.

Al comparar las métricas de error obtenidas para la temporada de invierno y verano de las tablas 6.6 y 6.5, se observa que aunque el error medio (ME), el error medio absoluto (MAE) y el error cuadrático medio (RMSE) presentaran mejores resultados en la temporada de verano, el modelo presentó un mejor desempeño en la temporada de invierno. Lo anterior, porque en invierno, el error porcentual es del orden del 20 %, muy inferior a los errores del 57 % obtenidos en la época de verano. Esto se debe a que las concentraciones medias registradas en la temporada de invierno son mucho mayores a las concentraciones medias de la temporada de verano, ya que en invierno la media de los registros se presenta entre 55,3 y 95,4 [$\mu\text{g}/\text{m}^3$], mientras que en verano la media de los registros se presenta entre 14,3 y 28,1 [$\mu\text{g}/\text{m}^3$], tal como se puede observar en las tablas 6.3 y 6.2 respectivamente. Por lo tanto, un mismo error porcentual en invierno y verano, generará mayores errores absolutos en la temporada de verano.

El mayor error porcentual obtenido de las estimaciones realizadas en el periodo de verano debe tener su origen en la formulación del método basado en interpolaciones, ya que este no es capaz de representar valores por fuera del intervalo de medición. En este caso, al analizar las tablas 6.1 y 6.2, que presentan las concentraciones medias de la temporada de verano en las estaciones utilizadas para realizar las interpolaciones y las utilizadas para el cálculo del error respectivamente. Se observa que las estaciones de validación de Cerrillos y Cerro Navia presentan el mayor y menor registro medio respectivamente. Por lo tanto, en la interpolación se excluyeron los registros máximos y mínimos del área de monitoreo, en consecuencia, es esperable que las estimaciones no logren representar adecuadamente los valores reales observados en dichos puntos.

Por otro lado, al analizar las tablas 6.3 y 6.4, que presentan las concentraciones medias de la temporada de invierno, se observa que en este caso el máximo y mínimo se presentan en las estaciones de Pudahuel y las Condes respectivamente, ambas estaciones utilizadas para la interpolación de los puntos. Por lo tanto, en este caso, los registros utilizados para realizar las interpolaciones cubren los máximos y mínimos del área de monitoreo y por ello, es esperable que presenten menores errores. Este aspecto muestra la importancia de considerar adecuadamente los puntos de monitoreo, de manera de abarcar no solo los puntos del plano donde se generen las mayores emisiones, sino también, aquellos que presentan las menores concentraciones del contaminante de interés.

Además, denota como los niveles de calidad del aire no son constantes entre las temporadas de invierno y verano. Ya que como se observa, en invierno las máximas y mínimas se presentaron en las estaciones de interpolación y en verano, dichos máximas y mínimos se presentaron en los puntos de validación, es decir, se presentaron en diferentes ubicaciones geográficas. Por lo tanto, para un modelo robusto, es necesario considerar puntos de monitoreo donde se produzcan los máximos y mínimos niveles de concentración en las temporadas de invierno y verano o en su defecto, modificar la disposición de equipos entre una temporada y otra. Una solución óptima a esta problemática, es el uso de sensores móviles, ya que permite que los dispositivos de monitoreo se ubiquen en los puntos de máxima y mínima emisión, asegurando que los puntos interpolados se encuentren dentro del intervalo de medición.

Otro aspecto interesante de notar es el signo del error medio obtenido en la temporada de invierno y verano, que presenta valores positivos y negativos respectivamente. Estos signos

indican si los valores se encuentran sobre o sub estimados, ya que corresponden al promedio de la resta entre el valor observado y estimado, de esta forma, los valores negativos indican una tendencia media a la sobre estimación y por el contrario, los valores positivos indican una tendencia a la sub estimación. Este aspecto es más claro en la temporada de invierno, ya que presenta el mayor error medio. Si se observan las tablas 6.3 y 6.4, los altos registros medios de las estaciones de validación de Cerro Navia y Cerrillos solo son superados por los registros medios en la estación de interpolación de Pudahuel. Por lo tanto, es consistente que sus resultados se encuentren sub estimados y por consiguiente, errores medios positivos.

Se presume que los menores errores obtenidos en la temporada de invierno pueden tener su origen en las dinámicas de circulación atmosférica. Estos fenómenos son dominados por la energía del sol, que induce los movimientos de las masas de aire por convección y por ello, la menor turbulencia atmosférica en la temporada de invierno. Esta menor turbulencia atmosférica genera deficientes condiciones de ventilación de los contaminantes emitidos.

Las deficientes condiciones de ventilación ocasionadas por la menor turbulencia atmosférica en la temporada de invierno se ven enfatizadas con el aumento de las emisiones de $MP_{2,5}$, producto del mayor uso de calefacción y de vehículos para el transporte. Estos factores conducen a que en invierno se alcancen condiciones de saturación para el $MP_{2,5}$, por lo tanto, es poco probable la presencia de marcados perfiles de concentración en el espacio/tiempo, los que si pueden presentarse en verano con condiciones de mayor turbulencia atmosférica. Estos marcados perfiles de concentración que se pueden presentar por la mayor turbulencia atmosférica en Verano, generan una mayor posibilidad de no lograr cubrir los puntos del plano que presentan los mayores o menores niveles de contaminación y, en consecuencia, es más probable que los puntos de monitoreo no se posicionen en los puntos de máxima y mínima concentración.

Costo Computacional

El interés de la investigación es desarrollar un modelo rápido y de bajo costo computacional. Por ello, por un lado se analiza el tiempo de ejecución del modulo que calcula los coeficientes de correlación, ya que este, es el algoritmo que realiza mayor cantidad de cálculos. Se analiza mediante el tiempo de ejecución del algoritmo y del peso del archivo generado que contiene todas las variables calculadas en la ejecución del cada uno de los algoritmos estudiados. En la tabla 6.7 se presentan los resultados obtenidos para cada método analizado.

Tabla 6.7: Comparación tiempo de calculo de Coeficientes de Correlación y peso de los datos generados en el calculo de los Coeficientes de Correlación.

Modelo	Tiempo [min]	Peso [KB]
I. Vecinos Naturales	1.8	367
A) Iterativo aleatorio 1	197	17780
B) Iterativo aleatorio 2	203	17780
C) Iterativo Orden grilla	195	18085

Tal y como se esperaba, en la tabla 6.7 se observa que el método de los vecinos naturales es significativamente más rápido que el algoritmo modificado, ya que el primero calcula N diagramas de Voronoi de $S + 1$ celdas, con S igual al número de puntos de monitoreo. Por otro lado, el mecanismo modificado solo en su primera iteración genera un diagrama con $S + 1$ celdas y en cada iteración posterior, genera un diagrama con una celda extra al generado en la iteración anterior. De esta manera, el último diagrama de Voronoi tiene $N + S$ celdas, con N igual al número de puntos de interpolación. Por este mismo motivo, el tiempo de cálculo y peso de los resultados obtenidos por el algoritmo modificado son mayores, ya que en cada iteración deben realizar un mayor número de cálculos a los realizados en la iteración anterior.

Resiliencia del modelo

En esta sección se analiza la resiliencia del modelo, para lo cual se plantean situaciones donde un equipo de monitoreo deja enviar información de calidad del aire desde su ubicación, por algún motivo que imposibilite una solución rápida o a larga distancia. Esta falta de información implica la imposibilidad de interpolar la grilla de puntos a partir de los coeficientes de correlación calculados, ya que el método requiere valores de concentración en todos los puntos de monitoreo definidos. Por este motivo, se plantean dos posibles soluciones para ejecutar la interpolación en ausencia de la información de la estación de la Florida.

La primera alternativa, es conservar los coeficientes de correlación calculados a partir de las seis ubicaciones iniciales de monitoreo y la concentración en el punto faltante, será interpolada a partir de las cinco estaciones restantes. Posteriormente, se interpolan los puntos desconocidos a partir de los coeficientes de correlación obtenidos a partir de las seis ubicaciones de monitoreo y las seis concentraciones obtenidas posterior a estimar la concentración en el punto faltante. Por lo tanto, en este caso se utilizan los coeficientes de correlación de seis puntos del plano, pese a que se cuenta con información en solo cinco de ellos.

La segunda alternativa, es calcular todos los coeficientes de correlación a partir de las cinco ubicaciones de las que se dispone de información y luego, interpolar a partir de los nuevos coeficientes todos los puntos del plano. Dado que el objetivo del algoritmo es ser utilizado como un sistema de alerta de temprana, el método de los vecinos naturales es el único que responde a los requerimientos temporales para recalculer rápidamente los coeficientes de correlación. Por este motivo, en este análisis las alternativas planteadas solo se analizarán mediante el método de los vecinos naturales. Se comparan los resultados obtenidos mediante la validación cruzada *K-Fold*, en los mismos puntos de validación anteriores, ver figura 6.1. Los resultados para la temporada de invierno y verano se presentan en las tablas 6.8 y 6.9 respectivamente.

Tabla 6.8: Comparación errores de interpolación desconectando la información de una ubicación de monitoreo. Temporada Invierno.

Modelo	ME	MAE	RMSE	PE(%)
Conservación de Coeficientes interpolando el 6to punto	-3,49	11,8	16,1	19,5
Recalculer Coeficientes a partir de 5 puntos	2,0	13,2	17,9	21,7

Tabla 6.9: Comparación errores de interpolación desconectando la información de una ubicación de monitoreo. Temporada Verano.

Modelo	ME	MAE	RMSE	PE(%)
Conservación de Coeficientes interpolando el 6to punto	0,29	7,7	12,1	57,1
Recalcular Coeficientes a partir de 5 puntos	1,19	8,16	12,3	57,9

En las tablas 6.8 y 6.9, se observa que para la temporada de invierno y verano recalcular los coeficiente de correlación presenta mayor error, que conservar los coeficientes de los puntos iniciales y estimar el sexto punto faltante. Otro punto interesante, es que en este caso los errores de interpolación son menores a los obtenidos en la interpolación de vecinos naturales a partir de seis puntos, que se presentan en la tabla 6.5 de la sección anterior.

Esto podría indicar que la variabilidad espacial de ese punto ya se encontraba representada con los demás puntos de monitoreo, e incluir esa ubicación solo genera una sobre representación de su valor y no entrega más información de la variación espacial de contaminantes, al menos para las ubicaciones de Cerrillos y Cerro Navia utilizadas en la validación cruzada. Es más, como se puede observar en la figura 6.1, la estación de la Florida se encuentra por el sur oriente, al lado derecho del mapa y, los puntos de validación de Cerrillos y Cerro Navia se encuentran por el lado izquierdo, entre varios puntos de monitoreo. Por lo tanto, incluir el efecto de la estación puede inducir un mayor error en la estimación de dichos puntos. Este aspecto no ocurriría con una red densamente distribuida, donde no existen vecinos naturales tan lejanos a un punto.

Sin embargo, no se puede concluir que la alternativa de conservar los coeficientes es mejor, ya que la exclusión de los datos de la Florida podría inducir mejoras en las estimaciones de los puntos de validación que se encuentran lejanos a ella. Pero dado la baja distribución de puntos de monitoreo en el sector sur oriente de la Región Metropolitana, se desconoce el desempeño del modelo en los puntos cercanos a la estación de la Florida, donde su exclusión debe inducir un aumento del error en las estimaciones del sector.

De esta forma, de los resultados obtenidos en la temporada invierno y verano, se presenta como una mejor alternativa conservar los coeficientes de correlación del número original de equipos de monitoreo y la concentración en el punto faltante estimarla a partir de los nodos restantes. Sin embargo, es necesario analizar el error respecto a otros puntos del plano, ya que extraer la estación de la Florida podría mejorar los resultados de las interpolación en los puntos de validación del lado izquierdo, pero podría presentar mejoras en las estimaciones del sector sur oriente de la Región Metropolitana. Por lo tanto, es necesario un mayor numero de puntos de monitoreo en la región para concluir cual alternativa es más viable, incluso, esta podría cambiar dependiendo el punto de monitoreo que deje de enviar información en tiempo real.

Estos resultados son interesantes desde la perspectiva que plantean la posibilidad de mejorar los resultados de interpolación a partir del mismo número de equipos de monitoreo S, pero interpolando los puntos de la grilla considerando una mayor cantidad de ubicaciones iniciales. Para analizar esto, se deben calcular los coeficientes de correlación a partir de M ubicaciones, con M igual a S + T y donde T representa las ubicaciones de los puntos de monitoreo ficticios, cuyos valores serán interpolados a partir de los S puntos de monitoreo

reales.

6.2. Test 2

A continuación, se presentan los resultados obtenidos del test realizado con el segundo set de datos. En primer lugar, se presentan los objetivos de este test, luego se presentan las concentraciones medias registradas por cada estación en el intervalo de medición. Posteriormente, se presentan las series de tiempo de las concentraciones estimadas y observadas para cada una de las dos estaciones de monitoreo utilizados en la validación cruzada de los resultados. Finalmente, se presentan las métricas de error promedio calculadas al comparar las diferencias entre los valores estimados y observados de las dos estaciones de validación.

En este test, para mayor robustez de los resultados, se utiliza información de un mes de material particulado de ocho estaciones de monitoreo del SINCA, correspondientes al segundo set de datos de material particulado de ocho estaciones de monitoreo en la temporada de invierno y verano. Para la temporada de invierno se utiliza la información comprendida entre el 01/07/2021 y el 01/08/2021, por otro lado, para la temporada de verano se utiliza la información comprendida entre el 01/01/2021 y el 01/02/2021. Se aplica el mecanismo de validación cruzada *K-Fold*, que utiliza los datos de seis estaciones para realizar las interpolaciones y la información de las dos restantes para el análisis de error entre la predicción y el valor observado, ver figura 6.3. En este caso, se analizan dos nuevas métricas de error, el MAPE y el IOA.

En este test, se comparan los resultados obtenidos por la interpolación de Vecinos Naturales y dos casos (A y C) del algoritmo modificado, variando el orden en que se interpolan los puntos de la grilla, se excluye el caso B por la similitud de los resultados con el caso A. En el caso A se interpolan los puntos aleatoriamente y en el C, en un orden predefinido, recorriendo fila a fila la grilla de puntos. Lo anterior busca analizar la existencia de una configuración óptima, que reduzca el error entre la estimación y el valor observado.

Resultados Verano Test 2

En las tablas 6.10 y 6.11 se presentan los registros medios de la temporada de verano en las estaciones utilizadas en la interpolación y en la validación respectivamente. Se observa que el máximo y mínimo registro medio se presentó en las estaciones de interpolación de Puente Alto y Pudahuel respectivamente. Por lo tanto, el intervalo de medición de las estaciones de validación se encuentran dentro del intervalo de monitoreo de las estaciones de interpolación.

Tabla 6.10: Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.

Estaciones	Pudahuel	Quilicura	Parque O.	Puente Alto	La Florida	Las Condes
\bar{C} (ug/m ³)	11,6	13,0	12,2	17,0	12,1	11,8

Tabla 6.11: Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.

Estaciones Validación	Cerro Navia	El Bosque
\bar{C} ($\mu\text{g}/\text{m}^3$)	14,7	14,7

A continuación, se presentan las series de tiempo de las concentraciones observadas y estimadas en los puntos de validación para la temporada de verano. La curva azul representa las concentraciones observadas y la curva naranja representa las concentraciones estimadas. En este test, nuevamente no se presentaron diferencias significativas entre el método de los vecinos naturales y el caso A y C del algoritmo modificado, por lo tanto, solo presentaran las tres gráficas de la estación de Cerro Navia, la totalidad de las gráficas se presentan en el anexo G.

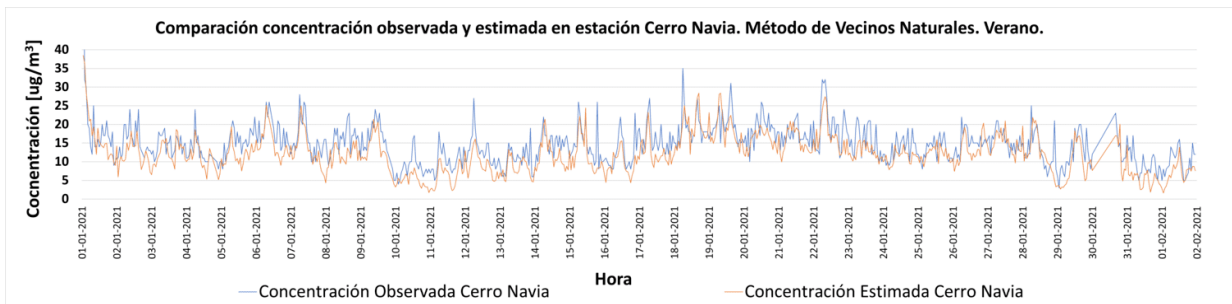


Figura 6.18: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. VN. Temporada Verano.

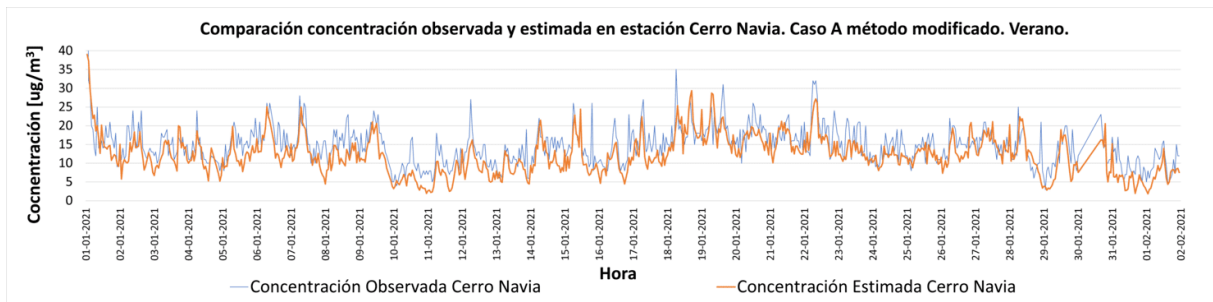


Figura 6.19: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IA. Temporada Verano.

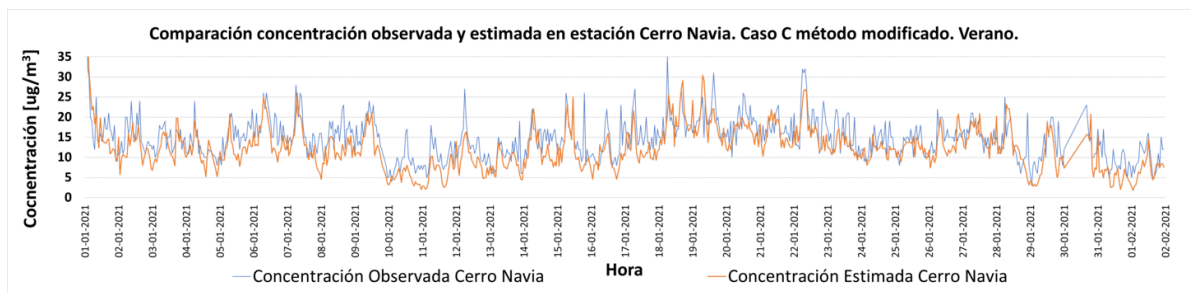


Figura 6.20: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IC. Temporada Verano.

En las figuras 6.18, 6.19 y 6.20, se observa el mismo comportamiento obtenido en el test 1, en el que las concentraciones estimadas logran representar el ciclo diario de $MP_{2,5}$, al identificar los puntos donde se presentan los máximos y mínimos de la curva de datos reales. Este aspecto es de gran relevancia, ya que implica que el mapa de distribución interpolado logra representar la variación del ciclo diario de contaminantes, identificando los momentos de mayor o menor concentración. Sin embargo, pese al solapamiento de las curvas de datos reales y estimados, se observa que las concentraciones se encuentran sub estimadas, por lo tanto, pese a identificar zonas y momentos de mayor concentración de contaminantes, los niveles de contaminación podrían ser mayores a los estimados.

Esta sub estimación de los niveles de $MP_{2,5}$ se debe a que a la media de los registros en la estación de Cerro Navia solo es superada por los registros de la estación Puente Alto, la cual se encuentra en uno de los extremos del plano. Debido a esto, al momento de interpolar las concentraciones en la ubicación de la estación de Cerro Navia, su calculo se realiza con los valores de las estaciones más cercanas a ella, que corresponden a las estaciones de Pudahuel, Parque O'higgins y Quilicura. Estas estaciones presentan medias entre los 11,6 y 13,0 [ug/m^3], muy por debajo de la media de 14,7 [ug/m^3] presentada en la estación de Cerro Navia. Por lo tanto, los valores reales se encuentran por debajo del intervalo de monitoreo y por ello, las estimaciones en este punto se encuentran sub estimadas. Nuevamente se observa la importancia de contar con una red ampliamente distribuida que se posicione en los puntos con los mayores y menores niveles de contaminación. De esta forma, al momento de realizar las interpolaciones se reduce el error de sub o sobre estimación.

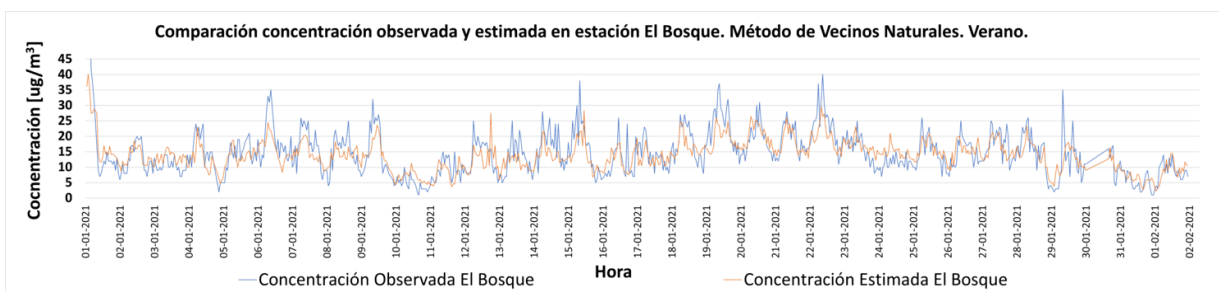


Figura 6.21: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. VN. Temporada Verano.

En la figura 6.21, nuevamente se observa que las concentraciones estimadas logran representar el ciclo diario de $MP_{2,5}$, identificando los puntos donde se presentan los máximos y mínimos de la curva de datos reales en la estación del Bosque. En este caso, no se observa un comportamiento tan claro a la sobre o sub estimación de los valores modelados, esto se debe a que el registro medio en la estación del Bosque se encuentra dentro del intervalo de medición de las estaciones de interpolación. La estación de validación del Bosque presenta un registro medio de 14,7 [ug/m^3] y es colindante a las estaciones de Puente Alto, la Florida y Parque O'higgins, las que presentan registros medios entre los 12,1 [ug/m^3] y los 17,0 [ug/m^3] en las estaciones de la Florida y Puente alto respectivamente. Por lo tanto, a diferencia de lo ocurrido en la estación de Cerro Navia, en este caso los valores interpolados en la estación del Bosque se encuentran dentro del rango de monitoreo y en consecuencia, los valores estimados no presentan una tendencia a la sobre o sub estimación.

Por otro lado, nuevamente se observa que los valores interpolados no logran representar los puntos máximos de la curva de datos reales que se encuentran muy por encima de los valores medios. En tales casos, el máximo de la curva estimada alcanza una altura menor que el máximo de la curva de datos reales. Esto se debe, a que si bien las medias se encuentran dentro del intervalo de medición, esto no asegura que en todos los registros horarios las concentraciones de $MP_{2,5}$ en la estación del Bosque se encuentren dentro del rango de monitoreo de sus estaciones vecinas.

Esta variabilidad, torna compleja la problemática sobre cuales son las ubicaciones óptimas de monitoreo que logren representar la variación espacial de los niveles de contaminación. Esto se debe principalmente a la incertidumbre que presenta la variación de los niveles de emisión en el plano, que se puede generar por el aumento del trafico vehicular en una autopista, o por la falla de los sistemas de filtración de una industria, entre otros. Una solución a esta problemática, es considerar dentro de los puntos de monitoreo los lugares más propensos a la variabilidad de sus índices de emisión, por ejemplo, autopistas o industrias.

Resultados Invierno Test 2

A continuación, en las tablas 6.12 y 6.13 se presentan los registros medios de las estaciones en la temporada de invierno utilizadas en la interpolación y validación respectivamente. En este caso, se observa que el máximo y mínimo registro medio se presenta en las estaciones de Cerro Navia y las Condes respectivamente.

Tabla 6.12: Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.

Estaciones	Pudahuel	Quilicura	Parque O.	Puente Alto	La Florida	Las Condes
\bar{C} (ug/m^3)	70,4	57,5	50,6	41,3	48,2	27,9

Tabla 6.13: Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.

Estaciones Validación	Cerro Navia	El Bosque
\bar{C} (ug/m^3)	72,8	68,7

A continuación se presentan las series de tiempo de las concentraciones estimadas y observadas en los puntos de validación de Cerro Navia y El Bosque para la temporada de invierno. La Curva azul representa las concentraciones observadas y la curva naranja las concentraciones estimadas en los puntos de validación. Solo se presentan los resultados del método de los vecinos naturales, ya que por la similitud de los resultados, las demás gráficas no entregan mayor información. Un detalle de la totalidad de las gráficas del método modificado se presenta en el anexo G.

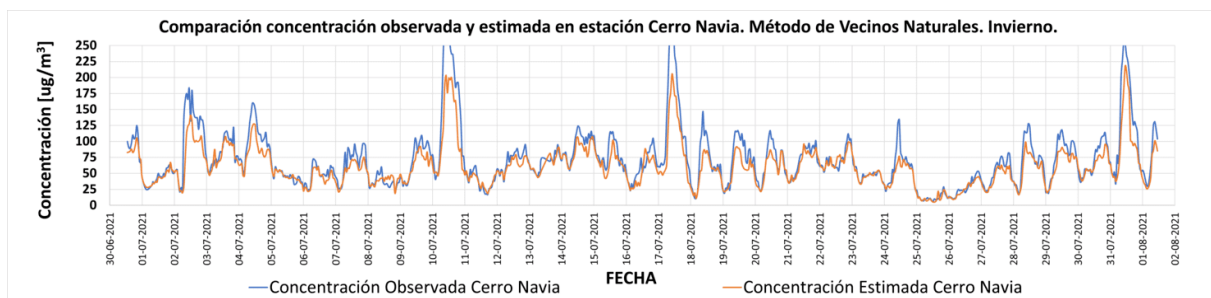


Figura 6.22: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. VN. Temporada Invierno.

En la figura 6.22, se observa una pequeña sub estimación de los valores modelados en comparación a los observados. Sin embargo, en este caso es mucho menor a lo ocurrido en el periodo de verano y las diferencias solo son notorias en los picos de las curvas. Lo anterior, se debe a que en este caso, la concentración estimada en el punto de Cerro Navia se encuentra influenciada en gran parte por la estación de Pudahuel, que se encuentra colindante a Cerro Navia. Además, la media en Pudahuel, es ligeramente menor a la registrada en Cerro Navia, con valores de 70,4 y 72,8 [$\mu\text{g}/\text{m}^3$] respectivamente. Por lo tanto, la estimación en este punto se ve bien representada por sus estaciones colindantes y, nuevamente se concluye que el método no se comporta adecuadamente en los picos máximos de la curva de datos observados.

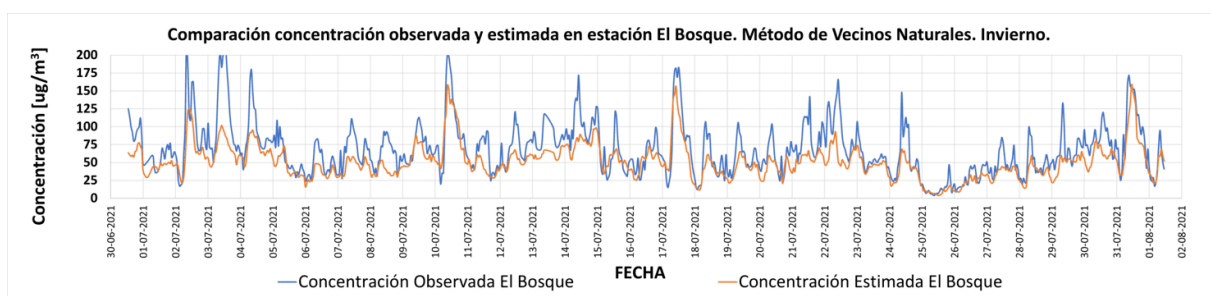


Figura 6.23: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. VN. Temporada Invierno.

En la figura 6.23, nuevamente se observa una sub estimación de los valores modelados en la estación del Bosque. En este caso, la sub estimación es mucho más clara que lo observado en la estación de Cerro Navia, ya que la media de la estación del Bosque se encuentra fuera del rango de medición de sus estaciones circundantes. Además, el registro medio en la

estación del bosque de 68,7 [ug/m³], presenta diferencias significativamente superiores a las medias de registradas por sus estaciones circundantes, que corresponden a las estaciones de Puente Alto, la Florida y Parque O'higgins, con registros medios de 41,3, 48,2 y 50,6 [ug/m³] respectivamente.

Resultados Numéricos Temporada Invierno y Verano Test 2

A continuación, se analiza numéricamente el desempeño y diferencias de los métodos desarrollados, mediante el calculo del ME, MAE, RMSE, PE, MAPE e IOA, las cuales se presentaron en la sección 4.3. Las métricas de error fueron calculadas contrastando los valores observados y estimados de cada hora en las estaciones de Cerro Navia y El Bosque. En la tablas 6.14 y 6.15 se presentan las métricas de error promedio en las estaciones de validación para el periodo de verano e invierno respectivamente.

Tabla 6.14: Resultados métricas de error media obtenidas en la temporada de verano, entre las concentraciones estimadas y observadas en estaciones de validación, para el algoritmo de vecinos naturales VN y el caso A y C del algoritmo modificado. Test 2

Método	ME	MAE	RMSE	PE(%)	MAPE(%)	IOA
VN	-1,62	3,51	4,69	31,91	27,7	0,83
IA	-1,81	3,49	4,65	31,63	26,21	0,84
IC	-1,82	3,52	4,69	31,93	26,38	0,84

Tabla 6.15: Resultados métricas de error media obtenidas en la temporada de invierno, entre las concentraciones estimadas y observadas en estaciones de validación, para el algoritmo de vecinos naturales VN y el caso A y C del algoritmo modificado. Test 1

Método	ME	MAE	RMSE	PE(%)	MAPE(%)	IOA
VN	-14,37	16,45	24,92	35,22	21,09	0,88
IA	-15,55	17,58	26,27	37,12	22,35	0,86
IC	-15,79	18,07	27,05	38,23	22,9	0,85

Respecto a las métricas de error calculadas, se concluye al igual que lo observado en el test 1, que no se presentan diferencias significativas entre los métodos analizados. Por otro lado, en este test, se observan menores diferencias entre los resultados de Invierno y Verano, esto se debe al mayor conjunto de datos que permite reducir el sesgo inducido por grandes errores locales. En este caso, nuevamente los mejores resultados se presentaron en la temporada de invierno, ya que presenta menores porcentajes de error absoluto (MAPE) y mayores índices de correlación (IOA). Sin embargo, a diferencia de lo observado en el test 1, en este caso, en invierno se presentaron mayores porcentajes de error en la temporada invierno, se presume que esto se debe, a que los valores en las estaciones de validación en la temporada de verano son significativamente mayores a los de las estaciones de interpolación, por lo tanto, los valores estimados presentan mayores errores porcentuales. Nuevamente se observa, como la distribución espacial de contaminantes varía entre una temporada y otra, por lo tanto, para

buenos resultados, se debe considerar el cambio de los puntos de máxima y mínima emisión de una temporada a otra.

6.3. Test 3

A continuación, se presentan los resultados obtenidos del tercer test, realizado con el segundo set de datos. En primer lugar, se presentan los objetivos de este test, luego se presentan las concentraciones medias registradas por cada estación en el intervalo de medición. Posteriormente, se presentan las series de tiempo de las concentraciones estimadas y observadas para cada una de las estaciones de monitoreo, cuando la estación fue utilizada para la validación cruzada de los resultados. Finalmente, se presentan las métricas de error promedio calculadas al comparar las diferencias entre los valores estimados y observados para cada una de las estaciones en la temporada de verano e invierno respectivamente.

En este test, se utiliza el mismo set de datos de un mes de material particulado utilizado en el caso anterior, con información de los periodos del 01/07/2021 al 01/08/2021 y del 01/01/2021-01/02/2021, para la temporada de invierno y verano respectivamente. En este caso, se analiza particularmente el método de los vecinos naturales y se aplica el mecanismo de validación cruzada *leave-one-out*, que excluye la información de una estación de monitoreo del grupo de datos de interpolación y la información en este punto es utilizada para el análisis de error entre la estimación y la observación. De esta forma, se analizan las métricas de error en cada estación de monitoreo, cuando dicha estación no fue considerada para realizar las interpolaciones del plano. En este caso, se busca observar el efecto generado al extraer cada punto de monitoreo y analizar el desempeño del modelo en problemas de interpolación y extrapolación.

Resultados Verano Test 3

En la tabla 6.16, se presentan las medias de cada estación de monitoreo de la temporada de verano. Se observa que el máximo y mínimo registro se presentan en las estaciones de Puente Alto y Pudahuel respectivamente.

Tabla 6.16: Concentraciones medias en puntos de monitoreo. Temporada Verano 2021.

Estaciones	Pudahuel	Cerro Navia	Quilicura	El Bosque	Parque O.	Puente Alto	La Florida	Las Condes
\bar{C} (ug/m ³)	11,6	14,7	13,0	14,7	12,2	17,0	12,1	11,8

A continuación, se presentan las series de tiempo de las concentraciones reales y estimadas en cada estación de monitoreo, cuando dicha estación se excluyó del set de datos de interpolación y fue utilizada para el análisis de error. En este test, solo se analiza el método de los vecinos naturales, ya que por los resultados obtenidos en los test 1 y 2, no se presentaron diferencias significativas entre los algoritmos. Además, por simplicidad, ya que en este test se aplica el método *leave-one-out*, que implica calcular 8 set de coeficientes de correlación distintos, cada uno en el cual es excluida una estación de monitoreo. Por lo tanto, realizar ese cálculo es demasiado costoso computacionalmente para el método modificado.

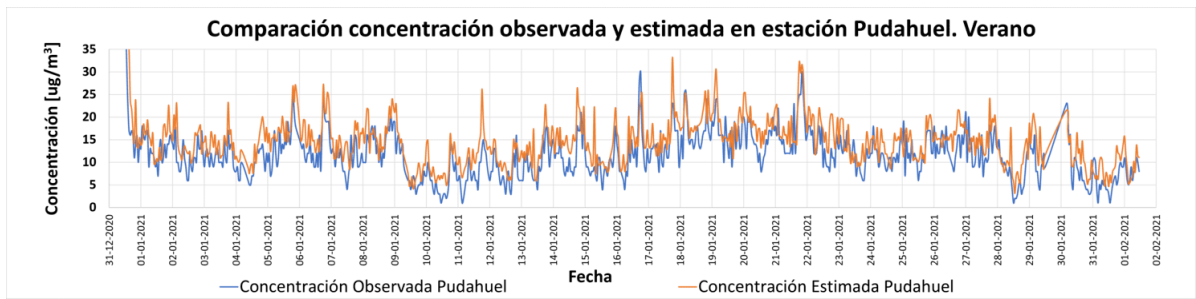


Figura 6.24: Comparación Concentraciones reales e interpoladas en ubicación estación Pudahuel para un mes de datos. Temporada Verano 2021.

En la figura 6.24, se observa que los valores estimados para la estación de Pudahuel se encuentran sobre estimados. Esto se debe a que la media en la estación de Pudahuel de $11,6 \mu\text{g}/\text{m}^3$ es menor a la media de sus estaciones colindantes, que corresponden a las estaciones de Cerro Navia, Quilicura, El Bosque y Parque O'higgins, que presentan registros medios iguales a $14,7$, $13,0$, $14,7$ y $12,2 \mu\text{g}/\text{m}^3$ respectivamente. Por lo tanto, su interpolación quedara acotada a las mediciones de sus estaciones colindantes y en consecuencia, esta será sobre estimada.

Además, ya que la estación de Pudahuel se encuentra en la periferia del área de monitoreo, los valores estimados para la estación de Pudahuel se ven representados en mayor medida por los datos de la estación de Cerro Navia y no se ven afectados significativamente por los registros de las demás estaciones. Por lo tanto, el valor en la estación de Pudahuel se calcula casi en su totalidad con los valores de la estación de Cerro Navia. Esto se puede observar en mayor detalle en el anexo A, en las figuras C.5 y C.6, que corresponden a los diagramas de Voronoi de cuando se excluye y de cuando se considera la Estación de Pudahuel en el en el área de modelamiento, en ellas se observa que la celda de Pudahuel C_8 se posiciona mayoritariamente sobre la celda de la estación de Cerro Navia.

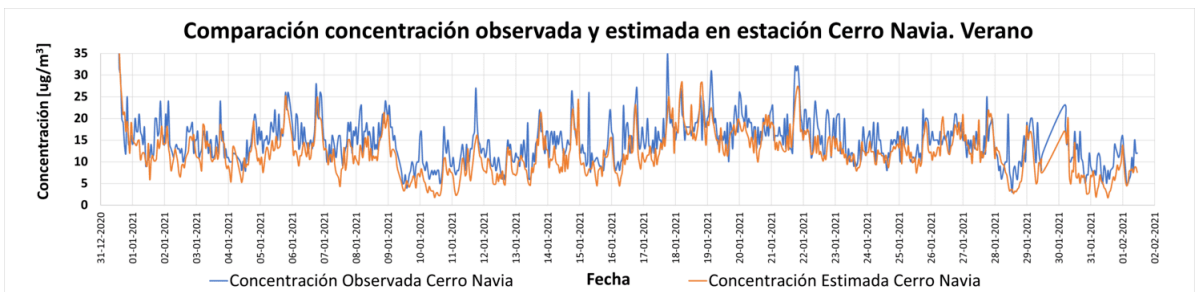


Figura 6.25: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. Temporada Verano 2021.

A diferencia de lo obtenido para la estación de Pudahuel, en este caso se observa que las estimaciones se encuentran sub estimadas, ya que ahora es la estación de Cerro Navia la que presenta la mayor media entre sus estaciones colindantes. Por lo tanto, los niveles de concentración en ese punto son calculados con estaciones que presentan en promedio menores valores de concentración. En consecuencia, los valores estimados se acotaran al intervalo de

datos de interpolación y por ello, los resultados obtenidos se encuentran sub estimados.

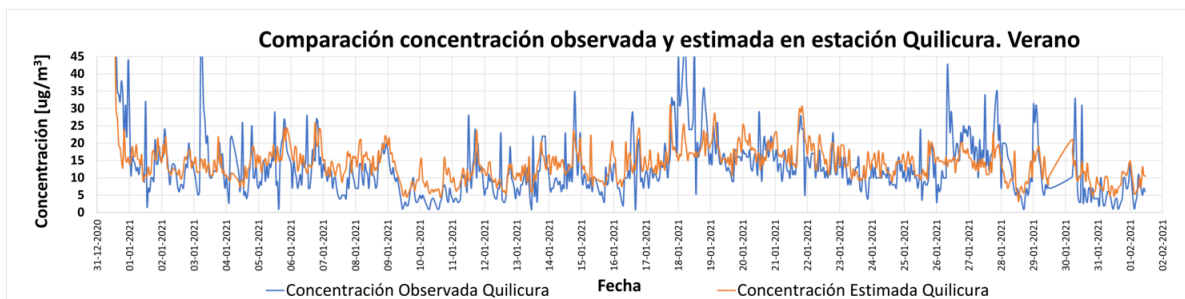


Figura 6.26: Comparación Concentraciones reales e interpoladas en ubicación estación Quilicura para un mes de datos. Temporada Verano 2021.

Respecto a los resultados obtenidos en la estación de Quilicura, presentados en la figura 6.26, se observa que en este caso no se presenta una clara tendencia a la sub o sobre estimación. Esto se debe a que su registro medio se encuentra dentro del intervalo de medición de sus estaciones colindantes, que corresponden principalmente a las estaciones de Pudahuel y Cerro Navia. Estas presentan un valor de 11,6 y 14,7 $[\mu\text{g}/\text{m}^3]$ respectivamente, por lo tanto, la media de la estación de Quilicura de 13.0 $[\mu\text{g}/\text{m}^3]$ se encuentra dentro del rango de monitoreo y por ello, los valores estimados no presentan una tendencia a la sobre o sub estimación.

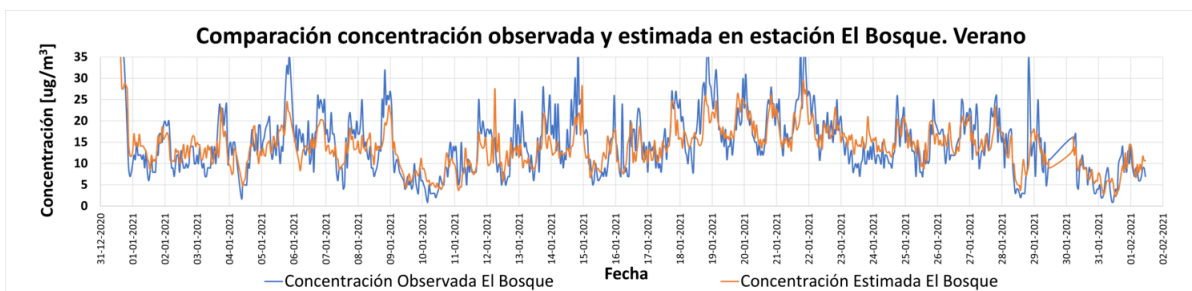


Figura 6.27: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. Temporada Verano 2021.

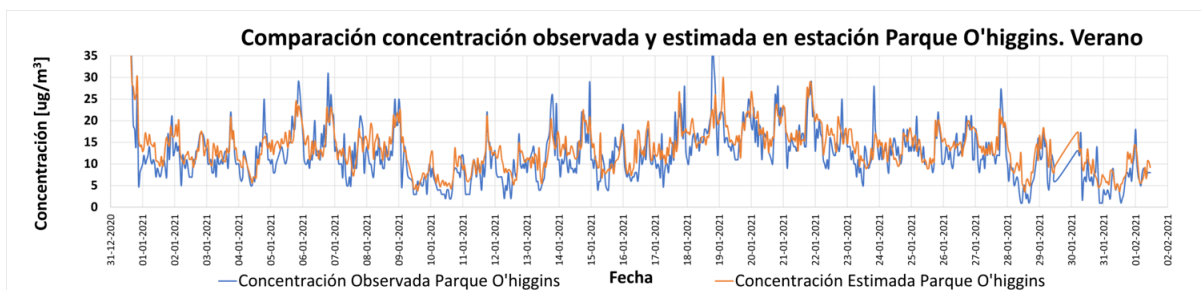


Figura 6.28: Comparación Concentraciones reales e interpoladas en ubicación estación Parque O'higgins para un mes de datos. Temporada Verano 2021.

Respecto de las estimaciones en las estaciones del Bosque y Parque O'higgins, que se presentan en las figuras 6.27 y 6.28 respectivamente, se observa que estas no presentan una clara tendencia a la sobre o sub estimación. Esto se debe, a que estas estaciones se encuentran en el centro del área de monitoreo, por lo tanto, en estos casos se presenta un problema de interpolación, en los cuales el método de vecinos naturales se comporta adecuadamente. Además, los valores medios registrados en las estaciones del Bosque y Parque O'higgins se encuentran dentro del rango de mediciones de las estaciones colindantes con las que se calcula su valor, ya que el máximo se presenta en la estación de Puente Alto y los mínimos en las de Pudahuel y las Condes.

El método de vecinos naturales se comporta adecuadamente en casos de interpolación porque como se observa, en este caso, los puntos centrales son rodeados por múltiples puntos de monitoreo, que permiten identificar la variación espacial en todo el contorno del punto interpolado. A diferencia de los puntos en los extremos del área de monitoreo, en los que su estimación no se realiza considerando puntos de todo su contorno, sino, que solo con una fracción de puntos que se encuentren por dentro del área de monitoreo. Por lo tanto, una buena estimación se realiza cuando el punto de interpolación se encuentra rodeado de puntos de monitoreo, pero además, cuando en este punto no se presenta un máximo o mínimo local.

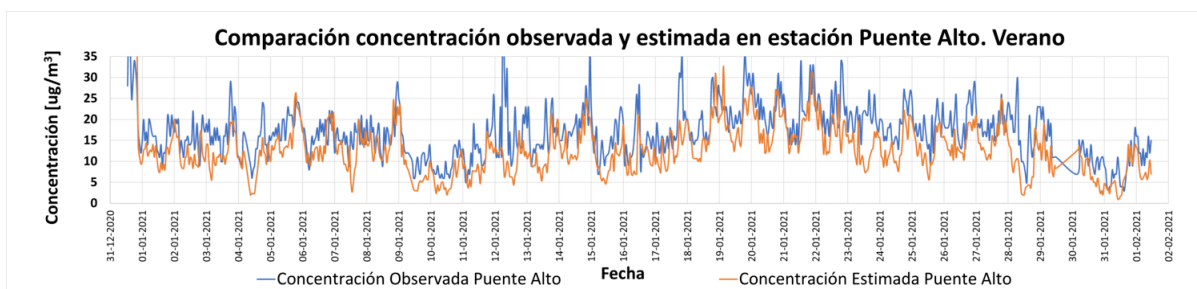


Figura 6.29: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto para un mes de datos. Temporada Verano 2021.

En la figura 6.29, se presentan los resultados obtenidos en estación de Puente Alto, se observa que esta presenta una clara sub estimación de las concentraciones, con diferencias significativas entre las estimaciones y observaciones. Este aspecto es consecuente con lo discutido anteriormente, ya que la estación de Puente Alto presenta los mayores registros medios, además, se encuentra en los extremos del área de modelación. Por lo tanto, las demás estaciones son incapaces de estimar un valor fuera del rango de monitoreo.

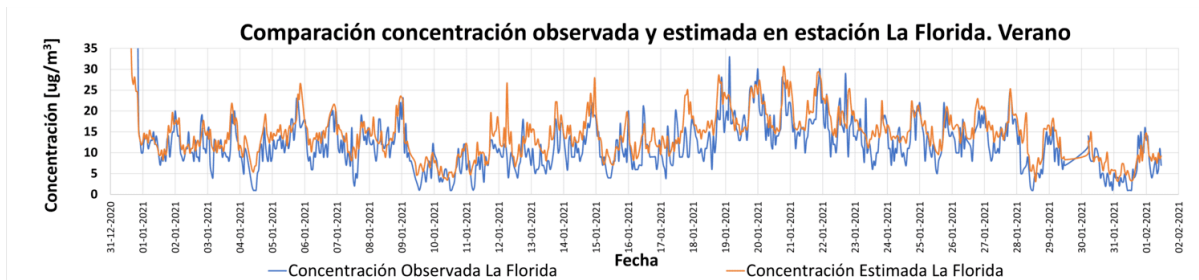


Figura 6.30: Comparación Concentraciones reales e interpoladas en ubicación estación La Florida para un mes de datos. Temporada Verano 2021.

En la figura 6.30, se presentan los resultados obtenidos en la estación de la Florida. En este caso se observa una tendencia a la sobre estimación de las concentraciones estimadas. Esto se debe a que la media de 12,1 [ug/m³] en la estación de la Florida, es menor a los registros medios de las estaciones colindantes, como se observa en la tabla 6.16 solo la estación de las Condes presenta un valor menor.

Además, como se puede observar en el anexo A, en las figuras C.13 y C.14, que presentan los diagramas de Voronoi excluyendo y considerando la celda de Voronoi de la estación la Florida, las estimaciones en este punto se realizan a con los registros de las estaciones de Parque O'higgins, El Bosque, Puente Alto y las Condes. Por lo tanto, los menores niveles de concentración en la estación de la Florida no podrán ser representados adecuadamente por sus estaciones circundantes y en consecuencia, se genera una sobre estimación.

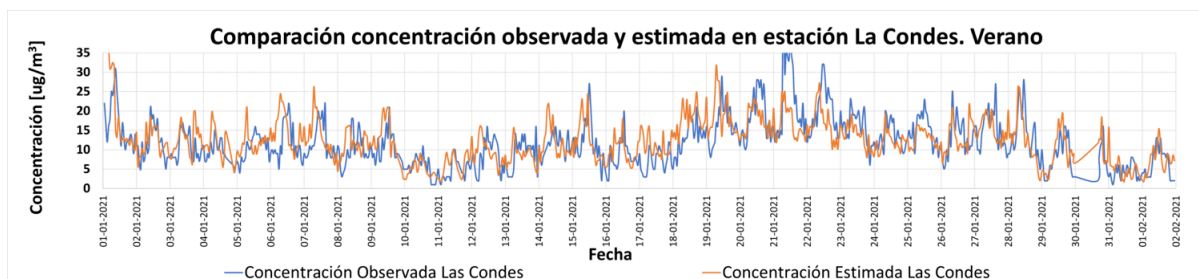


Figura 6.31: Comparación Concentraciones reales e interpoladas en ubicación estación Las Condes para un mes de datos. Temporada Verano 2021.

Finalmente, las concentraciones estimadas para la estación de las Condes no presentan una clara tendencia a la sobre o sub estimación. Esto se debe a que su valor es estimado mayoritariamente con los registros de las estaciones de Quilicura, Parque O'higgins y la Florida, que presentan medias de 13,0 [ug/m³], 12,2 [ug/m³] y 12,1 [ug/m³] respectivamente. Por lo tanto, si bien presentan medias mayores, los valores no se encuentran tan alejados del registro medio en las Condes de 11,8 [ug/m³].

Resultados Invierno Test 3

En la tabla 6.17, se presentan las medias de cada estación de monitoreo de la temporada de verano. Se observa que el máximo y mínimo registro se presenta en las estaciones de Cerro Navia y las Condes respectivamente. Recordar que en la figura 6.3, se presenta el mapa de las ubicaciones de estas estaciones en la Región Metropolitana.

Tabla 6.17: Concentraciones medias en puntos de monitoreo. Temporada Invierno 2021.

Estaciones	Pudahuel	Cerro Navia	Quilicura	El Bosque	Parque O.	Puente Alto	La Florida	Las Condes
\bar{C} ($\mu\text{g}/\text{m}^3$)	70,4	72,8	57,5	68,7	50,6	41,3	48,2	27,9

A continuación, se presentan las series de tiempo de las concentraciones observadas y estimadas en la temporada invierno. La curva azul representa los datos observados y la naranja los estimados en la estación de monitoreo.

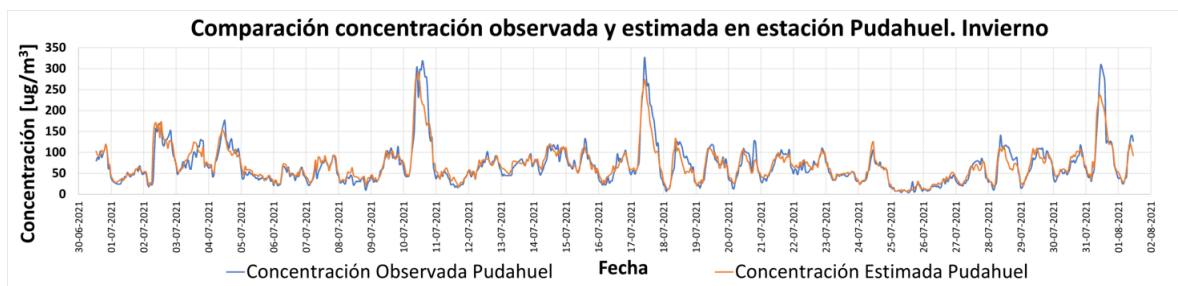


Figura 6.32: Comparación Concentraciones reales e interpoladas en ubicación estación Pudahuel para un mes de datos. Temporada Invierno 2021.

En la figura 6.32 se presentan las concentraciones estimadas y observadas en la estación de validación de Pudahuel. Se observa que la concentración estimada representa adecuadamente la concentración observada, la curva de datos estimados se superpone prácticamente en la totalidad de picos y valles de la curva de datos reales. Esto se debe, a que el valor en este punto se ve muy influenciado por la medición de la estación de Cerro Navia, que se encuentran significativamente cerca en comparación a las demás estaciones en el área de monitoreo y en consecuencia, presentan una mayor correlación.

Esta mayor correlación puede observar en las figuras C.5 y C.6, que presentan los diagramas de Voronoi del plano en las situaciones en que se considera y excluye la ubicación de la estación de Pudahuel. En los diagramas se observa claramente que la celda de Voronoi generada por la estación de Pudahuel se intersecta prácticamente en su totalidad con la celda de Voronoi de la estación de Cerro Navia. Por lo tanto, la estimación en el punto estará condicionada en gran medida al valor registrado en la estación de Cerro Navia y como se puede observar en la tabla 6.17, las estaciones de Pudahuel y Cerro Navia presentaron registros medios muy cercanos, con valores de 70,4 y 72,8 [$\mu\text{g}/\text{m}^3$] respectivamente. Por lo tanto, es esperable que las estimaciones en este punto presenten buenos resultados.



Figura 6.33: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. Temporada Invierno 2021.

En la figura 6.33 se presentan las concentraciones estimadas y observadas en la estación de validación de Cerro Navia. Se observa al igual que en la estación de Pudahuel, que las curvas de datos estimados y observados se solapan prácticamente en todo el dominio temporal. Sin embargo, nuevamente se observa que las estimaciones no alcanzan los picos máximos de la curva de concentraciones observadas, esto se debe a que al momento de presentarse los peak que las estimaciones no logran representar, las estaciones colindantes presentaron menores registros que los observados en la estación de Cerro Navia. De hecho, en la tabla 6.17, se observa que la estación de Cerro Navia presenta el máximo registro medio, por lo tanto es esperable que en algunos puntos exista una pequeña sub estimación.

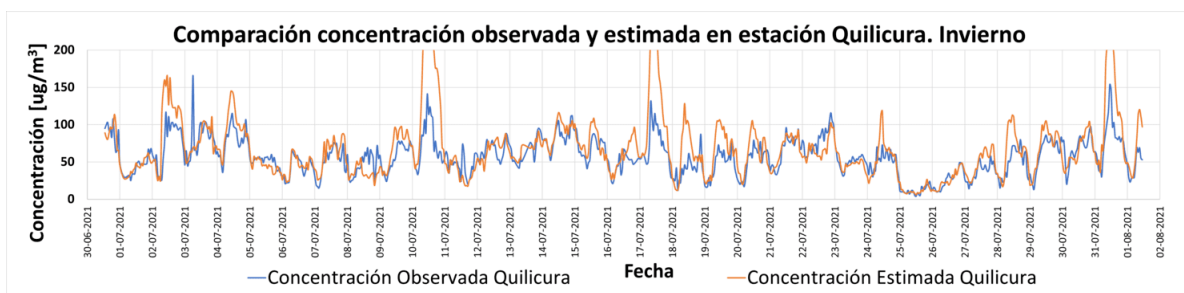


Figura 6.34: Comparación Concentraciones reales e interpoladas en ubicación estación Quilicura para un mes de datos. Temporada Invierno 2021.

En la figura 6.34 se presentan las concentraciones estimadas y observadas en la estación de validación de Quilicura. En este caso se observa que las concentraciones calculadas se encuentran sobre estimadas, es más, se puede observar como la curva de datos estimados presenta peak notoriamente superiores a los observados. Esto se debe a que las estaciones de Pudahuel y Cerro Navia, colindantes a la estación de Quilicura, presentan mayores registros medios, como se observa en la tabla 6.17. Por lo tanto, es esperable una sobre estimación.

Además, al observar las figuras C.1 y C.2, que presentan los diagramas de Voronoi del plano en situación con y sin la celda de Voronoi de la estación de Quilicura, se observa que la celda de la estación de Quilicura se intersecta principalmente sobre la celda de Cerro Navia y en menor medida, en las de Pudahuel y las Condes. Por lo tanto, su valor estimado se ve condicionado principalmente por los registros de la estación de Cerro Navia, que presenta los mayores registros medios. En consecuencia, los valores estimados en la ubicación de la

estación de Quilicura se sobre estiman.

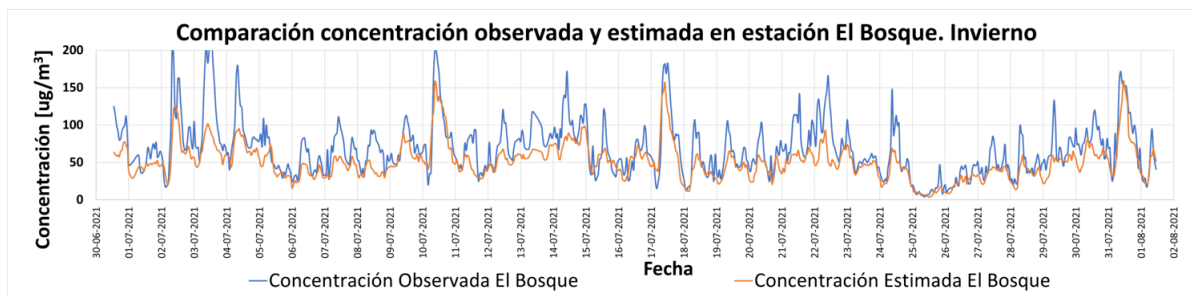


Figura 6.35: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. Temporada Invierno 2021.

En la figura 6.35 se presentan las concentraciones estimadas y observadas en la estación de validación del Bosque. En este caso, a diferencia de observado en la temporada de verano en este punto, se observa una clara sub estimación de las concentraciones estimadas. Esta problemática se debe a que la celda de Voronoi de la ubicación del Bosque se intersecta principalmente con la celda de Voronoi de Puente Alto y en menor medida, con las de Pudahuel, Parque O'higgins y la Florida. Por lo tanto, el valor estimado para la ubicación del Bosque esta condicionado en gran medida a los registros de la estación de Puente Alto, los que presentan una media de 41,3 [ug/m³], muy por debajo de la media observada en la ubicación del Bosque, de 68,7 [ug/m³]. En consecuencia de anterior, los niveles de concentración son sub estimados en este punto.

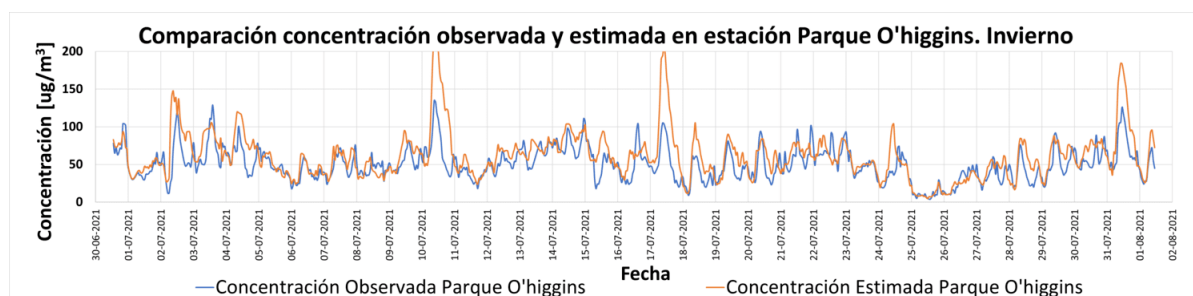


Figura 6.36: Comparación Concentraciones reales e interpoladas en ubicación estación Parque O'higgins para un mes de datos. Temporada Invierno 2021.

En la figura 6.36 se presentan las concentraciones estimadas y observadas en la estación de validación de Parque O'higgins. En este caso, al igual que en la temporada de verano, no se observa una tendencia clara a la sub o sobre estimación. Esto se debe, a que la estación de Parque O'higgins se encuentra en el centro del área de monitoreo. Por lo tanto, el problema presentado es una interpolación y su valor se calcula con un mayor número de estaciones colindantes.

En las figuras C.3 y C.4, que presenta los diagramas de Voronoi en las situaciones sin y con la presencia de la ubicación de Parque O'higgins. Se observa que sus valores se calculan

mediante la estación de Cerro Navia, Quilicura, las Condes, la Florida y el Bosque. Además, los registros medios en la estación de Parque O’higgins presentan un valor de 50,6 [ug/m³], el que se encuentra dentro del intervalo de monitoreo de sus estaciones colindantes, que va desde los 27,9 a los 72,8 [ug/m³] en las estaciones de las Condes y Cerro Navia respectivamente. Por lo tanto, el valor real se encuentra dentro del intervalo de monitoreo de su estaciones colindantes.

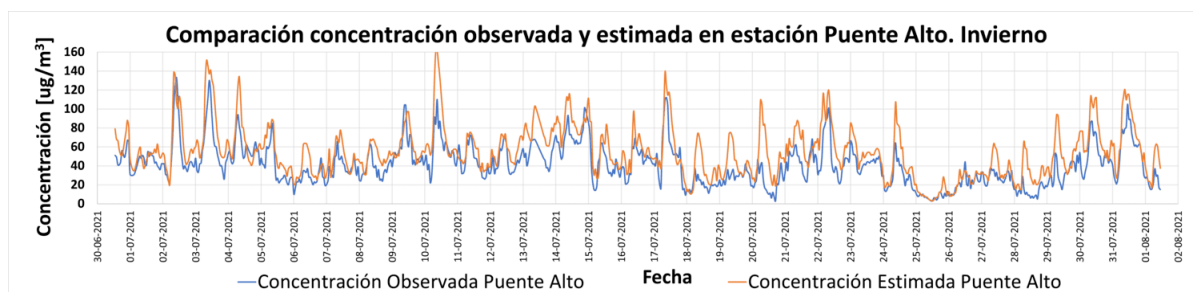


Figura 6.37: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto para un mes de datos. Temporada Invierno 2021.

En la figura 6.37 se presentan las concentraciones estimadas y observadas en la estación de validación de Puente Alto. En este caso se observa una sobre estimación de las concentraciones estimadas en este punto. Pese a lo anterior, la curva de datos estimados representa de buena manera la serie de tiempo de las concentraciones observadas, identificando los puntos de máxima y mínima concentración. Este aspecto es importante, ya que en este caso el problema se trata de una extrapolación, donde la estación de Puente de Alto se encuentra en los extremos del área de monitoreo.

La sub estimación, se debe a que los registros medios de la estación de Puente Alto se encuentran por debajo del intervalo de monitoreo de sus estaciones colindantes. Ya que, la estimación en este punto se realiza mediante la información de las estaciones del Bosque y la Florida, las cuales presentan registros medios iguales a 68,7 y 48,2 [ug/m³] respectivamente, superior a la media registrada en la estación de Puente Alto de 41,3 [ug/m³]. En consecuencia, los valores predichos en este punto presentan una sobre estimación.

Por otro lado, se presume que el comportamiento menos errático presentado por la curva estimada en la estación de Puente Alto en la temporada de invierno, en comparación al observado en la temporada de Verano, debe tener su origen en la menor turbulencia atmosférica en la época de invierno. Esta menor turbulencia, genera condiciones desfavorables a la dispersión y dilución de contaminantes, por lo tanto, es menos probable la presencia de grandes cambios en los perfiles de concentración espaciales y temporales, generando mayores condiciones de saturación para el MP_{2,5}. Producto de estas condiciones de saturación, los cambios en los niveles de concentración en el plano son menores en invierno y por ello, el menor error obtenido al extrapolar un punto en el extremo del plano.

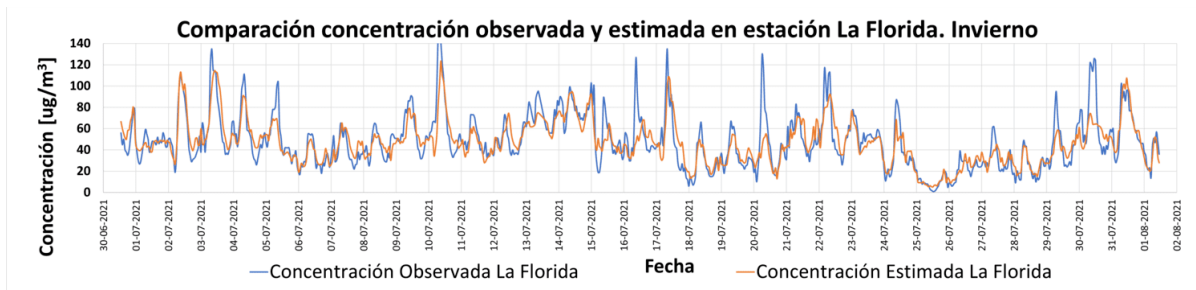


Figura 6.38: Comparación Concentraciones reales e interpoladas en ubicación estación La Florida para un mes de datos. Temporada Invierno 2021.

En la figura 6.38 se presentan las concentraciones estimadas y observadas en la estación de validación de la Florida. En este caso no se observa una tendencia a la sobre o sub estimación, esto se debe a que la estación de la Florida se encuentra en punto de monitoreo con varias estaciones colindantes, es decir, se trata de un problema de interpolación. Como se puede observar en las figuras C.13 y C.14, la concentración estimada en la estación la florida se calcula con los datos de Parque O'higgins, El Bosque, Puente Alto y las Condes, que presentan registros medios de 50,6, 68,7, 41,3 y 27,9 [$\mu\text{g}/\text{m}^3$] respectivamente, por lo tanto, la estación de la Florida presenta un registro medio dentro del intervalo de monitoreo de 48,2 [$\mu\text{g}/\text{m}^3$] y en consecuencia, sus buenos resultados.

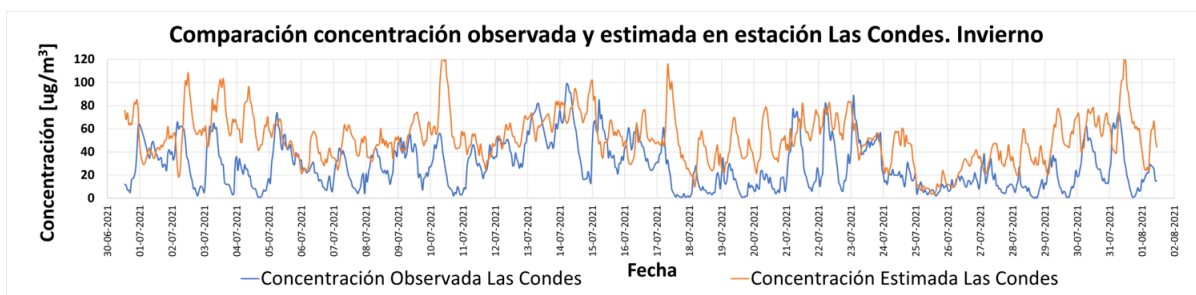


Figura 6.39: Comparación Concentraciones reales e interpoladas en ubicación estación Las Condes para un mes de datos. Temporada Invierno 2021.

Finalmente, en la figura 6.39 se presentan las concentraciones estimadas y observadas en la estación de validación de las Condes. En este caso, se observa una clara sobre estimación de las concentraciones estimadas en este punto. Esta sobre estimación se debe a que la estación de las Condes presenta el mínimo registro medio de las estaciones de monitoreo, con un valor de 27,9 [$\mu\text{g}/\text{m}^3$]. Por lo tanto, es imposible que las demás estaciones estimen un valor debajo de su rango de monitoreo.

Además, la estación de las Condes se encuentra en los extremos del área de monitoreo, por lo que en este caso se trata de un problema de extrapolación y se puede observar un comportamiento bastante errático, similar a lo ocurrido en la estación de Puente Alto en la temporada de verano. Esto se debe, a que la estación de las Condes además de encontrarse en los extremos del área de monitoreo, su ubicación se encuentra en altura en comparación

a los demás puntos de monitoreo de la Región Metropolitana. Esta problemática, no es tan relevante en la temporada de verano, donde la mayor turbulencia atmosférica induce una mayor dilución de los contaminantes y en consecuencia, el efecto de la altura no es tan relevante. De hecho, en las tablas 6.16 y 6.17, se puede observar que la diferencia de los registros medios de la estación las Condes con las demás estaciones de monitoreo, son notoriamente mayores en invierno que en verano.

Además, de acuerdo lo visto en el capítulo 2.5, en verano la capa de inversión térmica presenta una mayor altura a la presentada en invierno, por lo tanto los contaminante tienen a acumularse en las partes inferiores de la Región. Por otro lado, se presume, que la mayor humedad atmosférica en invierno favorece la deposición del $MP_{2,5}$ a las zonas más bajas. Por consiguiente, las mayores diferencias de concentración que presenta la estación de las Condes entre la temporada de verano e invierno. Por lo tanto, dado que el método solo considera dos ejes coordenados, se concluye que el método no se comporta adecuadamente cuando los puntos de monitoreo se ubican en diferentes alturas, en dichos casos, sería apropiado aplicar un método de dispersión en 3D, que considere el efecto de una tercera coordenada en la altura.

Resultados Numéricos Test 3

A continuación, se analiza numéricamente el desempeño de las estimaciones realizadas para cada punto de monitoreo anteriormente presentado, mediante el cálculo de las métricas del ME, MAE, RMSE, PE, MAPE e IOA, según las ecuaciones presentadas en la sección 4.3. Las métricas de error fueron calculadas contrastando los valores observados y estimados para cada hora del intervalo de medición, cuando se excluyó cada una de las estaciones de monitoreo del set de puntos de interpolación. En las tablas 6.18 y 6.19, se presentan las métricas de error promedio en cada estación para el periodo de verano e invierno respectivamente.

Tabla 6.18: *leave-one-out*. Comparación errores de interpolación desconectando la información de una ubicación, periodo Verano.

Estación de Validación	ME	MAE	RMSE	PE(%)	MAPE(%)	IOA
Pudahuel	2,96	3,63	4,54	38,9	43,2	0,81
Cerro Navia	-2,61	3,40	4,36	29,6	23,7	0,82
Quilicura	1,09	5,09	6,91	53,3	62,5	0,68
El Bosque	-0,64	3,62	5,00	34,1	30,8	0,84
Parque O'higgins	1,48	3,07	3,97	32,5	38,5	0,86
Puente Alto	-4,06	5,02	6,47	38,0	29,6	0,75
La Florida	2,30	3,56	4,71	38,9	47,5	0,85
Las Condes	0,57	4,04	5,57	47,18	47,5	0,75
Promedio	0,14	3,93	5,19	39,07	40,41	0,80

Tabla 6.19: Comparación errores de interpolación desconectando la información de una ubicación, periodo Invierno.

Estación de Validación	ME	MAE	RMSE	PE(%)	MAPE(%)	IOA
Pudahuel	0,96	11,2	17,7	25,1	18,4	0,96
Cerro Navia	-11,5	13,4	22,2	30,5	15,6	0,92
Quilicura	10,5	17,7	30,2	52,5	30,74	0,78
El Bosque	-17,3	19,5	27,4	39,9	26,7	0,81
Parque O'higgins	11,3	16,2	24,3	47,9	34,9	0,79
Puente Alto	13,7	15,0	19,6	47,5	47,7	0,85
La Florida	-1,12	9,49	13,8	28,5	23,9	0,90
Las Condes	23,8	27,9	33,6	120,74	inf	0,51
Promedio*	0,95	14,63	22,15	38,85	28,28	0,86

En las figuras 6.18 y 6.19 se puede observar que en la temporada de invierno las métricas de error porcentuales presentaron mejores resultados que los obtenidos en la temporada de verano. Sin embargo, cabe mencionar que en el cálculo de estas métricas, se excluyó el resultado de la estación de las condes, por presentar valores anómalos en la temporada de invierno. Debido a que esta se encuentra en altura respecto a las demás estaciones de monitoreo y en invierno los contaminantes tienden a depositarse en las zonas bajas de la Región.

Respecto al ME, MAE y RMSE, obtuvieron mejores resultados en la temporada de verano, ya que como se discutió anteriormente, los registros medios de la temporada de verano son significativamente menores a los registros medios de la temporada de invierno, como se puede observar en las tablas 6.16 y 6.17. Por lo tanto, un mismo error porcentual en invierno y verano, implica un mayor error absoluto en la temporada de invierno, sin que esto signifique un peor desempeño del modelo.

Sin embargo, el mayor error absoluto obtenido en invierno, si toma relevancia al analizar el modelo como un sistema de alerta temprana de calidad del aire. Ya que en invierno, es cuando se presentan las condiciones de calidad del aire de mayor riesgo, por lo tanto, un error absoluto muy grande, podría estar sub estimando los niveles de calidad del aire en situaciones que podrían ser de emergencia ambiental. Por lo tanto, estos errores se deben tener en consideración cuando los niveles de calidad del aire son cercanos a los límites que indican alerta ambiental, pre emergencia o emergencia ambiental, que en el caso del material particulado son iguales a 80, 110 y 170 [$\mu\text{g}/\text{m}^3$] respectivamente. En Chile, estos son definidos en el decreto N°12/2011 del ministerio del medio ambiente que “ESTABLECE NORMA PRIMARIA DE CALIDAD AMBIENTAL PARA MATERIAL PARTICULADO FINO RESPIRABLE $\text{MP}_{2,5}$ ”. Por lo tanto, desde un punto de vista de seguridad social, es preferible una configuración de monitoreo que sobre estime los niveles de contaminación, antes de generar condiciones de falsa seguridad producto de la sub estimación de los niveles de calidad del aire.

Por otro lado, los mayores registros de material particulado identificados en la Región Metropolitana en la temporada de invierno, tienen su origen en el aumento de los niveles de emisión por el mayor uso de combustibles, principalmente por el aumento del transporte

vehicular. Junto a ello, se presentan las deficientes condiciones de ventilación generadas por la menor turbulencia atmosférica, ya que como se presento en la figura 2.4, la dinámica atmosférica esta dominada por las diferencias de temperatura entre la atmósfera y el suelo. Por lo tanto, en épocas de invierno, producto de la menor radiación solar, las condiciones de turbulencia atmosférica son menores. De esta forma, la dispersión de contaminantes es menor y estos tienden a acumularse.

Sumado a las desfavorables condiciones de dispersión en la temporada de invierno. Como se presento en la sección 2.1.1, la cuenca de Santiago presenta condiciones topográficas desfavorables para la dilución de contaminantes y además, un descenso de la capa de inversión térmica. Por lo tanto, el aumento de emisiones, la menor turbulencia atmosférica, la topografía de la Región y en descenso de la capa de inversión térmica, son de los principales factores que favorecen la acumulación de contaminantes en la época de invierno.

Por el contrario, en la temporada de verano, se presume que los mayores errores y menor correlación obtenida, se deben a la mayor turbulencia atmosférica. Esta mayor turbulencia atmosférica genera mayor dispersión de los contaminantes y en consecuencia, la presencia de mayores perfiles de concentración en el espacio y tiempo. Producto de esto, en esta época, son más propensos los cambios abruptos en la distribución espacial de contaminantes, pudiendo cambiar con mayor rapidez los puntos donde se presentan las máximas y mínimas concentraciones del plano. En consecuencia, las interpolaciones son más propensas a los errores en que la estimación se encuentre fuera del intervalo de monitoreo, producto de la variación de los puntos de máxima y mínima concentración en todo el dominio temporal.

Otro aspecto interesante de notar, son los menores IOA que se presentaron en las estaciones de Quilicura, Puente Alto y Las Condes. Esta menor correlación en las estaciones, se debe a que estas ubicaciones se encuentran en los extremos del área de monitoreo. Por lo tanto, al excluirlas, se presenta un problema de extrapolación y en consecuencia, se demuestra que el método se comporta mejor en problemas de interpolación, estimando puntos dentro del área de monitoreo.

Del mismo modo, en la temporada de invierno, se observa que los menores IOA se presentan en las estaciones de Quilicura, Parque O y las Condes. En este caso, nuevamente el mínimo IOA se presenta en una estación del extremo del área de monitoreo, en la estación de Quilicura. Por lo tanto, nuevamente su estimación corresponde a resolver un problema de extrapolación.

Por otro lado, los deficientes resultados obtenidos en la estación de Parque O'higgins en la temporada de invierno, estación que se encuentra dentro del área de monitoreo, se explica porque esta presenta un valor por debajo de la media de sus estaciones colindantes, por lo tanto, su calculo no logra ser representando adecuadamente, pese a encontrarse dentro del área de monitoreo. Por consiguiente, no solo es relevante que los puntos de estimación se encuentren en zonas densamente monitoreadas, también es importante cubrir los puntos de máxima y mínima concentración del área de modelación.

Además, se observa que pese a que el modelo logra representar adecuadamente el comportamiento del ciclo diario de un contaminante, es deficiente en los puntos de máxima

y mínima concentración. Como se presentó, en ciertas circunstancias el modelo es sensible a la sobre o sub estimación excesiva de algunos puntos. Esto se debe, a que la selección de los puntos de monitoreo tienen un efecto considerable en los resultados del modelo si no se consideran los puntos de máxima y mínima concentración del área de modelación. Si esto ocurre, el modelo no es capaz de representar dichos niveles de contaminación fuera de su intervalo de monitoreo. De esta forma, los resultados generados en estos puntos se encontrarán sub estimados, si los puntos de máxima concentración no fueron considerados en el monitoreo, por el contrario, serán sobre estimados, si los puntos de mínima concentración no fueron considerados en el monitoreo.

Capítulo 7

Conclusión

El trabajo tuvo por objetivo general desarrollar un modelo de dispersión de contaminantes en tiempo real, mediante el uso de sensores de bajo costo, el cual actué como un sistema de alerta temprana de calidad del aire, identificando tempranamente los puntos de máxima y mínima concentración. Este objetivo se considera cumplido, ya que se implementa un modelo de dispersión de contaminantes con adquisición de datos sincrónicos, basado en la interpolación de los vecinos naturales producto de su rápida respuesta, que logra interpolar aproximadamente 1800 puntos del plano en 1,8 minutos. La ejecución del modelo para el caso del $MP_{2,5}$, presenta errores porcentuales del orden del 30 % y 40 % en las temporadas de invierno y verano respectivamente. Además, las estimaciones logran representar adecuadamente el ciclo diario de $MP_{2,5}$ y presentan índices de correlación superiores al 80 % en ambas temporadas. Por lo tanto, se implementa un método de dispersión que entrega una respuesta rápida y que a la vez, mantiene la verosimilitud de las estimaciones.

Junto a lo anterior, el trabajo tuvo por objetivo estudiar los fenómenos fisicoquímicos que afectan a las especies en la atmósfera. De esta forma, se concluye que la dispersión de contaminantes se encuentra dominada principalmente por la radiación solar, la cual regula el ciclo diario del viento, y a su vez, interfiere en las reacciones de muchas especies. Producto de lo anterior, en verano suelen presentarse mayores condiciones de turbulencia y dispersión atmosférica, que favorecen la dilución de los contaminantes, y en consecuencia, se presentan mejores niveles de calidad del aire que en invierno.

En base al estudio de los modelos utilizados en la modelación atmosférica, se identifica que los modelos físicos tienen una gran validación científica, sin embargo, estos tienen un alto costo computacional, por ese motivo, no se posicionan como la mejor alternativa para desarrollar un modelo en tiempo real. A diferencia de los modelos matemáticos basados en interpolaciones, los cuales por sus bajos requerimientos, presentan algoritmos simples y rápidos, por lo tanto, se posicionan como una excelente alternativa en la modelación en tiempo real. De esta forma, se cumple el objetivo específico número dos, identificando la mejor alternativa para la modelación con adquisición de datos sincrónicos.

Del mismo modo, en base al estudio de los software más utilizados en la modelación atmosférica, se concluye que los más populares corresponden a las herramientas de AERMOD y CALPUFF. Sin embargo, la falencia de estos software, es su imposibilidad de ejecutar un modelo con adquisición de datos sincrónicos, ya que corresponden a programas con módulos

predefinidos, que solo permiten cargar información histórica de calidad del aire y de esa forma, generar modelos estimados de la distribución de contaminantes en distintas condiciones meteorológicas. Por este motivo, las herramientas de Matlab y Python se presentan como excelentes alternativas para desarrollar el modelo de dispersión en tiempo real, ya que la flexibilidad que otorga al programador permite integrar la adquisición de datos sincrónicos, dentro de un módulo de dispersión.

El modelo desarrollado utiliza el método de los vecinos naturales, porque es un método sencillo y de bajo costo computacional, además, los diagramas de Voronoi con los cuales se desarrolla, han sido ampliamente utilizados para dividir el plano en zonas con características similares. Junto a lo anterior, se utiliza la herramienta de Matlab, por su flexibilidad y por la mayor documentación que presenta de sus funciones. Además, esta última cuenta con la nube *ThingSpeak*, especialmente destinada para la recopilación y análisis avanzado de datos generados por sensores inalámbricos.

En comparación a la información de calidad del aire de la plataforma del SINCA, se desarrolla una mejora sustancial en la presentación de la información de calidad del aire. Se desarrolla un algoritmo que presenta una gráfica de la distribución espacial de los contaminantes, identificando fácilmente las zonas con mejores y peores índices de calidad del aire. En contraste con la información del SINCA, que informa una serie de tiempo de los niveles de contaminación en un punto, la cual no permite reconocer los lugares de mayor o menor riesgo. Por lo tanto, el algoritmo podría ser utilizado como un sistema de alerta temprana de calidad del aire, permitiendo reconocer fácilmente zonas de riesgo y tomar las medidas preventivas correspondientes, como alejarse del lugar, uso de mascarillas, etc.

De los resultados obtenidos en el test 1, se concluye que el método logra representar adecuadamente los patrones diarios del ciclo de contaminantes, pese a que el modelo no considera el efecto de las variables físicas y químicas que dominan su comportamiento. Lo anterior, porque las interpolaciones son realizadas con información real de los niveles de contaminación en los puntos de monitoreo, por lo tanto, el nivel medido está condicionado intrínsecamente por las variables físicas y químicas que lo dominan, por lo tanto, si a las 12 del día se producen los máximos niveles de concentración, los resultados generados también representarían dicha situación. Este aspecto es de gran relevancia, ya que implica que el mapa de distribución estimado logra representar la variación del ciclo diario de contaminantes, identificando los momentos de mayor o menor concentración de la especie de interés.

Además, se identifica que el modelo es sensible a sobre o sub estimación de los resultados, cuando los equipos de monitoreo no se posicionan en los puntos de máxima y mínima concentración del área de modelamiento. Si esto ocurre, el modelo no es capaz de representar dichos niveles de contaminación fuera de su intervalo de monitoreo. De esta forma, los resultados generados en estos puntos se encontrarán sub estimados, si los puntos de máxima concentración no fueron considerados en el monitoreo, por el contrario, serán sobre estimados, si los puntos de mínima concentración no fueron considerados en el monitoreo.

Además, según lo observado en el test 1, los puntos de máxima y mínima concentración pueden variar durante el día de un punto a otro, producto del aumento de las emisiones u otros factores. Junto a ello, en los *test* se observa que estos puntos de máxima y mínima

concentración también varían de una temporada a otra. Esta variabilidad, torna compleja la problemática sobre cuales son las ubicaciones óptimas de monitoreo que logren representar la variación espacial de los niveles de contaminación. Una solución a esta problemática, es considerar dentro de los puntos de monitoreo los lugares más propensos a la variabilidad de sus índices de emisión, por ejemplo, autopistas o industrias.

Del análisis de resiliencia del test 1, se concluye que las estimaciones podrían presentar mejoras con la inclusión de puntos de monitoreo ficticios, los cuales son estimados a partir de los puntos de monitoreo reales. Sin embargo, falta realizar un mayor número de simulaciones en redes de monitoreo altamente distribuidas, que permitan analizar el error global al incluir puntos de monitoreo ficticios en las interpolaciones.

Por otro lado, de las variaciones realizadas al algoritmo de los vecinos naturales, analizadas en los test 1 y 2, resulta que estas no presentan diferencias significativas en las métricas de error calculadas. Por este motivo, se concluye que variar la correlación espacial de un punto para que tenga un mayor alcance, no presenta una mejora en las métricas de error dada la configuración utilizada, de hecho, solo generan que el modelo sea más costoso computacionalmente.

Del test 2 se concluye principalmente que una buena estimación de los niveles de concentración en un punto, ocurre cuando dicho punto se encuentra rodeado por múltiples ubicaciones de monitoreo, las que no deben estar muy alejadas entre si. Además, estas ubicaciones de monitoreo se deben ubicar estratégicamente, de manera de medir en los puntos del plano donde se presentan los mayores y menores niveles de contaminación durante el día. Como estos puntos varían constantemente, producto de las condiciones meteorológicas y niveles emisión, en instalaciones fijas sin capacidad de reubicación de los puntos de monitoreo, las ubicaciones se deben determinar considerando las máximas y mínimas de ambas temporadas. Por otro lado, en configuraciones que permitan una fácil reubicación de los equipos de monitoreo, bastaría con reubicar sensores en las ubicaciones donde se presenten las máximas y mínimas en cada temporada. Otra alternativa, es el uso de sensores móviles, los cuales se pueden desplazar para ubicarse en los puntos de máxima y mínima concentración, asegurando un monitoreo óptimo en todo momento.

Por otro lado, del test 3 se concluye que el método no se comporta tan eficientemente en problemas de extrapolación. Esto, ya que los menores coeficientes de correlación se presentaron en las estimaciones realizadas en las ubicaciones de los extremos del área de monitoreo. En la temporada de verano, se presentaron en las estaciones de Quilicura, Puente Alto y las Condes, con un IOA igual a 0,68, 0,75 y 0,75 respectivamente. Por otro lado, en invierno, los mínimos IOA se presentaron en las estaciones de Quilicura y las Condes, con valores de 0,78 y 0,51. Por lo tanto, en invierno y verano, los menores índices de correlación se presentaron en las estaciones de la periferia del área de modelación. La menor correlación se debe a que las estimaciones en estos puntos solo tienen información de los niveles de concentración de una fracción del plano, en comparación a los puntos del centro del área de monitoreo que tienen información de calidad del aire de todo su contorno.

Por este motivo, el modelo presenta mejores resultados en problemas de interpolación, al estimar valores de concentración en puntos que se encuentran dentro del área de monitoreo.

En el test 3, se observa que los mayores índices de correlación se obtienen en las estaciones de Parque O'higgins y el Bosque, las cuales se encuentran dentro del área de monitoreo, así mismo, en invierno, los mejores IOA se presentan en las estaciones de Pudahuel y Cerro Navia, que también se encuentran en el centro del área de monitoreo. Por este motivo, se recomiendan redes de monitoreo altamente distribuidas, que aseguren que los puntos a estimar se encuentren rodeados de puntos de monitoreo en todas las direcciones.

Del análisis de los resultados obtenidos en la temporada de invierno y verano, se observa que el modelo presenta mejores resultados en la época de invierno, con un MAPE del 28 % *versus* el 40 % obtenido en verano. Esto se debe a la menor turbulencia atmosférica en la época de invierno, que disminuye las condiciones de ventilación y producto de ello, los contaminantes tienden a acumularse y la variación espacial es menor. Producto de lo anterior, los contaminantes se distribuyen de manera más homogénea y es menos probable que se presenten cambios abruptos en los puntos de máxima y mínima concentración. La menor variación de los puntos de máxima y mínima concentración aseguran interpolaciones en las que los puntos de estimación se encuentren dentro del intervalo de medición, por el contrario, en verano, la mayor turbulencia atmosférica induce mayor cantidad de cambios abruptos en los perfiles de concentración, siendo más probable que cambien los puntos de máxima y mínima concentración, induciendo errores de sobre o sub estimación al no cubrir los puntos máximos y mínimos del plano.

Por lo anterior, para implementar el modelo en ambientes turbulentos, como zonas costeras o parques industriales, donde los perfiles de concentración pueden variar rápidamente, es necesario contar con una red de monitoreo ampliamente distribuida, que asegure cubrir los puntos de máxima y mínima concentración del plano en todo momento. De esta forma, se aseguran estimaciones que no se encuentren sobre o sub estimadas, producto de que la dinámica atmosférica haga variar los puntos donde se producen las máximas y mínimas concentraciones.

Por otro lado, pese que los mejores resultados porcentuales se presentaron en la temporada de invierno, en esta se presentaron los mayores errores absolutos. Lo anterior producto de los mayores niveles de contaminación que se presentan en invierno, por lo tanto, un mismo error porcentual en invierno y verano, implica un mayor error absoluto en invierno, sin que esto implique necesariamente un peor desempeño del modelo. Este factor toma gran relevancia al analizar el algoritmo como un sistema de alerta temprana. Ya que desde un punto de vista de seguridad social, es preferible una configuración de monitoreo que sobre estime los niveles de contaminación, antes de generar condiciones de falsa seguridad producto de la sub estimación de los niveles de calidad del aire.

Respecto a los posibles estudios de profundización, se identifica prometedor el análisis de la inclusión de puntos de monitoreo ficticios, interpolados a partir de las concentraciones medidas en los equipos desplegados. Estos puntos de monitoreo ficticios permitirían generar una mayor suavidad en el cambio de las concentraciones, integrando valores de concentración ficticias en puntos que presentan una alta correlación espacial con los puntos de monitoreo. Para esto es necesario realizar mayor número de simulaciones, en las cuales a partir de un número *S* de puntos de monitoreo generar puntos de monitoreo ficticios, estimando las concentraciones en algunas ubicaciones a partir de los puntos de monitoreo reales. La

hipótesis es que esto podría reducir el error de las interpolaciones, al incluir la correlación espacial de puntos intermedios ampliamente distribuidos.

Con base a estas conclusiones, dentro de las mejoras identificadas, se reconoce que integrar la variable del viento, como lo propuesto por Boubrima [5] o Contreras [82], presenta excelentes resultados, permitiendo reducir considerablemente el número de equipos de monitoreo. Junto a ello, el uso de sensores móviles, que permiten adaptar su posición constantemente a las ubicaciones de máxima y mínima concentración, se posiciona como excelente solución al problema de óptima ubicación de puntos de monitoreo, asegurando que los puntos de interpolación se encuentren en todo momento dentro del intervalo de medición.

Más aún, el uso de sensores móviles para solventar los cambios dinámicos de los perfiles de concentración en ambientes turbulentos, permitirán asegurar que estos siempre se ubiquen en las posiciones óptimas de máxima y mínima concentración del plano, y de esta forma, reducir errores de sub o sobre estimación. En consecuencia, se podría plantear un mecanismo de control *feedback*, donde algunos sensores móviles se utilicen como puntos de validación, que retro alimenten al modelo de errores y den la orden de movimiento a los sensores utilizados en las interpolaciones. De esta forma, se proponen una red de sensores móviles que aporten al modelo y una red (pequeña) de sensores que permitan hacer validación permanentemente, identificando los cambios de los perfiles de concentración y reubicando los sensores que aportan al modelo en las ubicaciones óptimas. Este mecanismo puede ser costoso en términos computacionales, en comparación a otros procesos y sistemas de control, pero considerando que estamos hablando de la atmósfera y la salud de las personas, el esfuerzo extra cobra sentido.

Bibliografía

- [1] Organización Mundial de la Salud, “Nueve de cada diez personas de todo el mundo respiran aire contaminado Sin embargo, cada vez hay más países que toman medidas,” 2018, <https://www.who.int/es/news/item/02-05-2018-9-out-of-10-people-worldwide-breathe-polluted-air-but-more-countries-are-taking-action> (visitado el 2020-11-02).
- [2] Álvarez C, H., “11.300 muertes por contaminación se producen en Chile cada año - IPSUSS,” 2016, <http://www.ipsuss.cl/ipsuss/analisis-y-estudios/11-300-muertes-por-contaminacion-se-producen-en-chile-cada-ano/2016-03-22/151633.html> (visitado el 2020-11-02).
- [3] MINISTERIO DEL MEDIO AMBIENTE, GUÍA DE CALIDAD DEL AIRE EDUCACIÓN AMBIENTAL, vol. 003. 2016, <https://mma.gob.cl/wp-content/uploads/2018/08/Guia-para-Docentes-Sobre-Calidad-del-Aire-003.pdf>.
- [4] SINCA, “Redes de monitoreo en línea,” 2020, <https://sinca.mma.gob.cl/index.php/redes> (visitado el 2020-12-01).
- [5] Boubrima, A. y Al., E., “On the deployment of wireless sensor networks for air quality mapping: Optimization models and algorithms,” *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1629–1642, 2019, doi:10.1109/TNET.2019.2923737.
- [6] Kumar, A., Kim, H., y Hancke, G. P., “Environmental monitoring systems: A review,” *IEEE Sensors Journal*, vol. 13, no. 4, pp. 1329–1339, 2013, doi:10.1109/JSEN.2012.2233469.
- [7] Yick, J., Mukherjee, B., y Ghosal, D., “Wireless sensor network survey,” *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008, doi:10.1016/j.comnet.2008.04.002.
- [8] Mead, M., Popoola, O., Stewart, G., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J. J., McLeod, M. W., Hodgson, T. F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J. R., y Jones, R. L., “The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks,” *Atmospheric Environment*, vol. 70, pp. 186–203, 2013, doi:10.1016/j.atmosenv.2012.11.060.
- [9] ETSI, “Electromagnetic compatibility and Radio spectrum Matters (ERM); System Reference Document (SRdoc): Spectrum Requirements for Short Range Device, Metropolitan Mesh Machine Networks (M3N) and Smart Metering (SM) Applications,” rep. tec., 2011.
- [10] Boubrima, A., Matigot, F., Bechkit, W., Rivano, H., y Ruas, A., “Optimal deployment of wireless sensor networks for air pollution monitoring,” *Proceedings - International Conference on Computer Communications and Networks, ICCCN*, vol. 2015-Octob, 2015, doi:10.1109/ICCCN.2015.7288443.

- [11] Boubrima, A., Bechkit, W., y Rivano, H., “Optimal WSN Deployment Models for Air Pollution Monitoring,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2723–2735, 2017, doi:10.1109/TWC.2017.2658601.
- [12] Fernández, O., “Superintendencia de Medio Ambiente reabre investigación por contaminación en Quintero,” 2020, [{O}](https://www.litoralpress.cl/sitio/Prensa_Texto?LPKey=o0p.Mbfq.Gdc.R.Rniu.Vj.E.U.Y.Pxm.A.Z.Ld.Strda.Xar.Tdvs7s.S0.) (visitado el 2020-11-02).
- [13] MMA, “A dos años de los episodios de contaminación en Quintero y Puchuncaví Seremi del Medio Ambiente reafirma importantes avances para descontaminar el territorio,” 2020, <https://mma.gob.cl/a-dos-anos-de-los-episodios-de-contaminacion-en-quintero-y-puchuncavi-seremi-del-medio-ambiente-reafirma-importantes-avances-para-descontaminar-el-territorio/> (visitado el 2020-11-02).
- [14] EPA, “Technical Overview of Volatile Organic Compounds.”, <https://www.epa.gov/indoor-air-quality-iaq/technical-overview-volatile-organic-compounds> (visitado el 2020-11-02).
- [15] Fundación TERRAM, “La negligente realidad de la Bahía de Quintero,” *Publicaciones Fundación Terram*, vol. 31, p. 27, 2018.
- [16] Seinfeld, J. H. y Pandis, S. N., *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, vol. 51. 2 ed., 1998, doi:10.1063/1.882420.
- [17] Moreno, A., *Química atmosférica de compuestos orgánicos volátiles oxigenados*. PhD thesis, Universidad de Castilla, 2016.
- [18] Inzunza, J., “Circulación de la atmósfera,” en *Meteorología Descriptiva*, cap. 8, p. 36, Santiago: Universitaria, 1 ed., 2019.
- [19] et al Gallego, A., *Contaminación atmosférica*, vol. 1. Madrid: UNED, 2012.
- [20] Larios, R., “El viento en movimiento. aprende en casa ii secundaria,” 2020, <https://www.unionjalisco.mx/2020/11/27/el-viento-en-movimiento-aprende-en-casa-ii-secundaria/>.
- [21] Carmona, V. y Díaz, C., *Una propuesta de material didáctico (juego de mesa) que favorece el proceso de enseñanza aprendizaje de la contaminación atmosférica y sus efectos en la salud humana*. PhD thesis, Universidad Del Valle, Santiago de Cali, 2013.
- [22] Vásquez, Y., *Contribución de las Distribuidoras de Combustible a los Niveles de Concentración de los Compuestos Orgánicos Volátiles en la Región Metropolitana*. PhD thesis, Universidad de Chile, 2011.
- [23] Molina, M., Sarukhán, J., y Carabias, J., *El Cambio Climático*, vol. 1. Fondo de Cultura Económica (FCE), 1 ed., 2017, http://www.nutricion.org/publicaciones/pdf/prejuicios_y_verdades_sobre_grasas.pdf https://www.colesterolfamiliar.org/formacion/guia.pdf https://www.colesterolfamiliar.org/wp-content/uploads/2015/05/guia.pdf.
- [24] United States Environmental Protection Agency, “Vocabulary Catalog.”, https://sor.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do;jsessionid=0cIINNkoJhu84Tma3-h154XAoA5dCpHiAnPVjhzrQM_SAEec_k7AL!13317628?details=\&vocabName=GreenhouseEmissionsGlossary\&filterTerm=poll\&checkedAcronym.

- [25] Ministerio para la Transición Ecológica y el Reto Demográfico, “De compuestos orgánicos volátiles.”, https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/emisiones/act-emis/compuestos_organicos_volatiles.aspx (visitado el 2021-10-28).
- [26] Brasseur, G. P., Orlando, J. J., y Tyndall, G. S., Atmospheric Chemistry and Global Change. New York: Oxford University Press, 1999.
- [27] ATSDR, “Resúmenes de Salud Pública - Benceno. Agencia para Sustancias Tóxicas y el Registro de Enfermedades,” 2016, https://www.atsdr.cdc.gov/es/phs/es_phs3.html (visitado el 2021-10-28).
- [28] Wayne, R. P., Chemistry of Atmospheres. No. 2, 3 ed., 2000, [doi:10.2307/2260295](https://doi.org/10.2307/2260295).
- [29] Lee, S., Chiu, M., Ho, K., Zou, S., y Wang, X., “Volatile organic compounds (VOCs) in urban atmosphere of hong kong,” Chemosphere, vol. 48, no. 3, pp. 375–382, 2002, [doi:10.1016/s0045-6535\(02\)00040-1](https://doi.org/10.1016/s0045-6535(02)00040-1).
- [30] NAPCA, Air Quality Criteria for Photochemical Oxidants. Washington: National Air Pollution Control Administration, 1970.
- [31] Carmona-Cabezas, R., Gómez-Gómez, J., Gutiérrez de Ravé, E., Sánchez-López, E., Serrano, J., y Jiménez-Hornero, F. J., “Improving graph-based detection of singular events for photochemical smog agents,” Chemosphere, vol. 253, p. 126660, 2020, [doi:10.1016/J.CHEMOSPHERE.2020.126660](https://doi.org/10.1016/J.CHEMOSPHERE.2020.126660).
- [32] Finlayson-Pitts, B. J., Chemistry of the Upper and Lower Atmosphere. Academic Press, 2000.
- [33] Yu, D., Tan, Z., Lu, K., Ma, X., Li, X., Chen, S., Zhu, B., Lin, L., Li, Y., Qiu, P., Yang, X., Liu, Y., Wang, H., He, L., Huang, X., y Zhang, Y., “An Explicit Study of Local Ozone Budget and NO_x-VOCs Sensitivity in Shenzhen China,” Atmospheric Environment, vol. 224, p. 117304, 2020, [doi:10.1016/J.ATMOENV.2020.117304](https://doi.org/10.1016/J.ATMOENV.2020.117304).
- [34] SINCA, “Estación Parque O’Higgins - Sistema de Información Nacional de Calidad del Aire.”, <https://sinca.mma.gob.cl/index.php/estacion/index/id/273> (visitado el 2021-11-01).
- [35] SINCA, “Validación.”, <https://sinca.mma.gob.cl/index.php/pagina/index/id/validacion> (visitado el 2021-11-29).
- [36] CMM, “EVALUACIÓN Y REDISEÑO DE LAS REDES DE MONITOREO DE CALIDAD DEL AIRE- II Informe,” rep. tec., Centro Mario Molina, Chile, 2017.
- [37] Ahumada, S., EVALUACIÓN DEL DESEMPEÑO DE LOS SENSORES ELECTROQUÍMICOS DE SO₂ Y NO₂ UTILIZANDO INSTRUMENTOS DE MONITOREO REGULATORIO. PhD thesis, Universidad de Chile, 2020.
- [38] EPA, EPA Handbook : Optical Remote Sensing for Measurement and Monitoring of Emissions Flux. No. September, 2011.
- [39] Parra, M. A., Elustondo, D., Bermejo, R., y Santamaría, J. M., “Ambient air levels of volatile organic compounds (VOC) and nitrogen dioxide (NO₂) in a medium size city in Northern Spain,” Science of The Total Environment, vol. 407, pp. 999–1009, 2009, [doi:10.1016/J.SCITOTENV.2008.10.032](https://doi.org/10.1016/J.SCITOTENV.2008.10.032).
- [40] SEA, GUÍA PARA EL USO DE MODELOS DE CALIDAD DEL AIRE EN EL SEIA.

Gobierno de Chile, 2012.

- [41] Perdiz, J., Modelado y simulación de la dispersión de contaminantes en la pluma de una chimenea por el método de dispersión Gaussiano. PhD thesis, UNIVERSIDAD NACIONAL DE TRUJILLO, 2016.
- [42] Dobbins, R. A., Atmospheric Motion and Air Pollution: An Introduction for Students of Engineering and Science. John Wiley & Sons Inc, 1979, <https://www.amazon.es/Atmospheric-Motion-Air-Pollution-Environmental/dp/0471216755>.
- [43] Grell, G. y Baklanov, A., “Integrated modeling for forecasting weather and air quality: A call for fully coupled approaches.” 2011.
- [44] EPA, “Air Quality Models.”, <https://www.epa.gov/scram/air-quality-models> (visitado el 2020-11-19).
- [45] Singh, S. K., Sharan, M., y Issartel, J. P., “Inverse modelling methods for identifying unknown releases in emergency scenarios: An overview,” International Journal of Environment and Pollution, vol. 57, no. 1-2, pp. 68–91, 2015, doi:10.1504/IJEP.2015.072121.
- [46] Singh, S. K. y Rani, R., “A least-squares inversion technique for identification of a point release: Application to Fusion Field Trials 2007,” Atmospheric Environment, vol. 92, pp. 104–117, 2014, doi:10.1016/j.atmosenv.2014.04.012.
- [47] Hutchinson, M., Oh, H., y Chen, W. H., “A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors.” 2017.
- [48] Ragland, K. W., “Multiple box model for dispersion of air pollutants from area sources,” Atmospheric Environment (1967), vol. 7, no. 11, pp. 1017–1032, 1973, doi:10.1016/0004-6981(73)90213-8.
- [49] Bosanquet, C. H. y Pearson, J. L., “The spread of smoke and gases from chimneys,” Trans. Faraday Soc., vol. 32, pp. 1249–1263, 1936, doi:10.1039/TF9363201249.
- [50] Pisso, I., Sollum, E., Grythe, H., Kristiansen, N., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R. L., Zwaafink, C. D. G., Evangelou, N., Sodemann, H., Haimberger, L., Henne, S., Brunner, D., Burkhardt, J. F., Fouilloux, A., Brioude, J., Philipp, A., Seibert, P., y Stohl, A., “The lagrangian particle dispersion model FLEXPART version 10.3,” 2019, doi:10.5194/gmd-2018-333.
- [51] Holmes, N. S. y Morawska, L., “A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available,” Atmospheric Environment, vol. 40, no. 30, pp. 5902–5928, 2006, doi:10.1016/j.atmosenv.2006.06.003.
- [52] Ermak, D., User’s manual for slab: An atmospheric dispersion model for denser-than-air-releases. 1990.
- [53] Tobler, W. R., “A Computer Movie Simulating Urban Growth in the Detroit Region,” Economic Geography, vol. 46, p. 240, 1970, doi:10.2307/143141.
- [54] Matheron, G., LA TEORÍA DE LAS VARIABLES REGIONALIZADAS Y SUS APLICACIONES. (M Alfaro, Trans). (Obra original publicada en1970), 2005.
- [55] Emery, X., Geoestadística. Santiago: Facultad de Ciencias Físicas y Matemáticas Universidad de Chile, 2013, <https://www.u-cursos.cl/usuario/f697544fb48474f415>

3ef2aafbe4c963/mi_blog/r/Geoestadistica.pdf.

- [56] Krige D.G, A statistical approach to some mine valuation and allied problems on the Witwatersrand. PhD thesis, Master's thesis of the University of Witwatersrand, 1951.
- [57] ArcGIS, "Cómo funciona kriging.", <https://desktop.arcgis.com/es/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm> (visitado el 2021-11-29).
- [58] Van Der Graaf, S., Natural Neighbour Kriging. PhD thesis, Delft University of Technology, 2016.
- [59] GISGeography, "Inverse Distance Weighting (IDW) Interpolation," 2021, <https://gisgeography.com/inverse-distance-weighting-idw-interpolation/> (visitado el 2021-11-08).
- [60] "Spatial tessellations: concepts and applications of voronoi diagrams," 2000.
- [61] Krige D.G, A Brief Description of Natural Neighbor Interpolation. PhD thesis, In: Barnett, V., Ed., Interpreting Multivariate Data, John Wiley & Sons, New York, 21-36, 1981.
- [62] ARCGIS, "Cómo funciona Vecino natural.", <https://desktop.arcgis.com/es/arcmap/10.4/manage-data/terrains/hidden-natural-neighbor-interpolation.htm> (visitado el 2020-12-28).
- [63] Quanfu Fan, A. E., Vladlen Koltun, Shankar Krishnan, S. V., "Hardware-assisted natural neighbor interpolation," , 2005.
- [64] Dumitru, P. D., Plopeanu, M., y Badea, D., "Comparative study regarding the methods of interpolation." 2013.
- [65] EPA, "40 CFR 51. Revision to the Guideline on Air Quality Models: Adoption of a Preferred General Purpose (Flat and Complex Terrain) Dispersion Model and Other Revisions," 2005, <https://www.federalregister.gov/documents/2005/11/09/05-21627/revision-to-the-guideline-on-air-quality-models-adoption-of-a-preferred-general-purpose-flat-and>.
- [66] UDC; y EPA; "Diseño del sistema inteligente de monitoreo de la calidad ambiental del distrito de cartagena convenio administrativo 0133-2015," vol. 4, p. 165, 2015, <http://observatorio.epacartagena.gov.co/ftp-uploads/dt-tomo-iv-protocolo-aire-20151412.pdf>.
- [67] Grsic, Z., Milutinovic, P., Arbutina, D., Milinčić, M., Dramlic, D., Velikic, Z., y Dramlic, S., "a Real Time Air Pollution Dispersion Modelling of Pc Nuclear Facilities of Serbia," Ecology of Urban Areas, no. October, pp. 97–104, 2013.
- [68] AIRVIRO, "Airviro.", <https://www.airviro.com/airviro/> (visitado el 2020-12-02).
- [69] Teggi, S., Costanzini, S., Ghermandi, G., Malagoli, C., y Vinceti, M., "A GIS-based atmospheric dispersion model for pollutants emitted by complex source areas," Science of the Total Environment, vol. 610-611, pp. 175–190, 2018, [doi:10.1016/j.scitotenv.2017.07.196](https://doi.org/10.1016/j.scitotenv.2017.07.196).
- [70] Priya, G., "Evaluation of Interpolation Techniques for Air Quality Monitoring using Statistical Error Metrics - A Review," INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICONNECT, vol. 06, 2018, www.ijert.org.
- [71] Geisser, S., Predictive Inference: An Introduction. Springer US, 1993, [doi:10.1007/978-1-4899-4467-2](https://doi.org/10.1007/978-1-4899-4467-2).

- [72] Long, N. Q., Cuong, C. X., y Canh, L. V. A. N., “3D Spatial Interpolation Methods for Open-Pit Mining Air Quality with Data Acquired by Small UAV Based Monitoring System,” pp. 263–274, 2020.
- [73] Hulme, M., Conway, D., Jones, P. D., Jiang, T., Barrow, E. M., y Turney, C., “Construction of a 1961–1990 European climatology for climate change modelling and impact applications,” *International Journal of Climatology*, vol. 15, no. 12, pp. 1333–1363, 1995, [doi:10.1002/joc.3370151204](https://doi.org/10.1002/joc.3370151204).
- [74] Nalder, I. A. y Wein, R. W., “Spatial interpolation of climatic Normals: Test of a new method in the Canadian boreal forest,” *Agricultural and Forest Meteorology*, vol. 92, no. 4, pp. 211–225, 1998, [doi:10.1016/S0168-1923\(98\)00102-6](https://doi.org/10.1016/S0168-1923(98)00102-6).
- [75] Robinson, W. S., “The statistical measurement of agreement,” *American Sociological Review*, vol. 22, p. 17, 1957, [doi:10.2307/2088760](https://doi.org/10.2307/2088760).
- [76] Ideation Technology Solutions, “Report On Air Quality Monitoring For Kolkata EV Implementation Project Phase II,” no. October, 2020.
- [77] Soulhac, L., Salizzoni, P., Mejean, P., Didier, D., y Rios, I., “The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study,” *Atmospheric Environment*, vol. 49, pp. 320–337, 2012, [doi:10.1016/j.atmosenv.2011.11.031](https://doi.org/10.1016/j.atmosenv.2011.11.031).
- [78] Tilloy, A., Mallet, V., Poulet, D., Pesin, C., y Brocheton, F., “BLUE-based NO₂ data assimilation at urban scale,” *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 4, pp. 2031–2040, 2013, [doi:10.1002/jgrd.50233](https://doi.org/10.1002/jgrd.50233).
- [79] Chakrabarty, K., Iyengar, S., Qi, H., y Cho, E., “Grid coverage for surveillance and target location in distributed sensor networks,” *IEEE Transactions on Computers*, vol. 51, no. 12, pp. 1448–1453, 2002, [doi:10.1109/TC.2002.1146711](https://doi.org/10.1109/TC.2002.1146711).
- [80] Ranieri, J., Chebira, A., y Vetterli, M., “Near-optimal sensor placement for linear inverse problems,” *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1135–1146, 2014, [doi:10.1109/TSP.2014.2299518](https://doi.org/10.1109/TSP.2014.2299518).
- [81] Wang, Y., Wang, Y., Zhang, X., Wang, D., y Yan, J., “Modeling of pollutant distribution based on mobile sensor networks,” *Environmental Science and Pollution Research*, vol. 27, no. 10, pp. 11413–11424, 2020, [doi:10.1007/s11356-020-07684-w](https://doi.org/10.1007/s11356-020-07684-w).
- [82] Contreras, L. y Ferri, C., “Wind-sensitive interpolation of urban air pollution forecasts,” *Procedia Computer Science*, vol. 80, pp. 313–323, 2016, [doi:10.1016/j.procs.2016.05.343](https://doi.org/10.1016/j.procs.2016.05.343).
- [83] FCNYM, “Geoestadística .”, <https://www.fcnym.unlp.edu.ar/catedras/geoeconomica/tp/tp13ppt.pdf>.
- [84] Kantaros, Y., “clippedVoronoi(nodes,area)”, 2021, https://www.mathworks.com/matlabcentral/fileexchange/59609-clippedvoronoi-nodes-area?s_tid=srchtitle.
- [85] Lindenbergh, R., *Limits of Voronoi Diagrams*. Universiteit Utrecht, 2020.
- [86] SINCA, “Normativa aplicable.”, <https://sinca.mma.gob.cl/index.php/pagina/index/id/norma> (visitado el 2021-11-29).
- [87] Martínez, B. U. P. d. O. y Merklings, P. U. P. d. O., “TEMA 2 "MODELOS DE CONCENTRACIÓN DE CONTAMINANTES ATMOSFÉRICOS”, rep. tec., 2008.

[88] EPA, “Meteorological Processors and Accessory Programs | Support Center for Regulatory Atmospheric Modeling (SCRAM).”, <https://www.epa.gov/scram/meteorological-processors-and-accessory-programs> (visitado el 2020-11-20).

[89] Alphasense, “Alphasense PhotoIonisation Detection (PID),” rep. tec., 2010.

Anexos

Anexo A

Diagrama de Voronoi

Un acercamiento a las propiedades del diagrama de voronoi es mediante el análisis del emplazamiento de los supermercados en una ciudad. Permite entre otras, responder a la pregunta, ¿Cuál es el supermercado más cercano a mi ubicación?, con esta premisa divide la ciudad en regiones, considerando el supermercado más cercano de los habitantes de la región. Una celda de Voronoi de un supermercado es la región del espacio donde el supermercado es el más cercano que cualquier otro. En la figura A.1 vemos algunos límites, por ejemplo el límite entre el supermercado 1 y el supermercado 3. Las personas que viven en este límite están a la misma distancia entre esos dos.

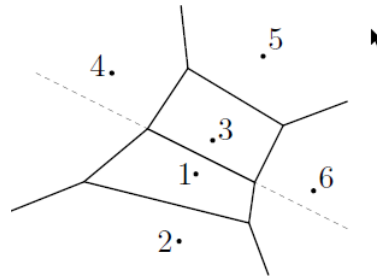


Figura A.1: Ejemplo de Diagrama de Voronoi de puntos 1,...,6 [85].

La línea punteada de la figura es la bisectriz de 1 y 3, porque a un lado de la línea la gente está más cerca del supermercado 1 y en el otro lado la gente está más cerca del supermercado 3. Por lo tanto, la bisectriz divide el plano en dos sectores o semiplanos. El sector que contiene supermercado 1 se llama semiplano de Voronoi $vh(1, 3)$ y el sector del otro lado de la bisectriz el semiplano de Voronoi $vh(3, 1)$. Se puede expresar el sector $V(1)$ del supermercado 1 en términos de estos semiplanos de Voronoi, según la ecuación A.1.

$$V(1) = vh(1, 2) \cap vh(1, 3) \cap vh(1, 4) \cap vh(1, 5) \cap vh(1, 6) \quad (A.1)$$

Dado un conjunto $S = \{p_1, \dots, p_n\}$ de n puntos distintos (x, y) en un plano, el diagrama de Voronoi $V(S)$ de N , es la familia de subconjuntos de \mathbb{R}^2 formado por las celdas de Voronoi y todas sus intersecciones. El límite de una célula de Voronoi está formada por aristas de Voronoi y vértices de Voronoi. Es una división del plano, que asigna a cada nodo una celda $V(p_i)$ o región del espacio cuyos puntos no son más cercanos a ningún otro nodo. Se define la celda de voronoi $V(p_i)$ de un punto $p_i \in S$ [85]:

Definición A.1 *Celda de Voronoi*

$$V(p_i) := \{q \in \mathbb{R}^2 : d(p_j, q) \geq d(p_i, q) \forall j \neq i\}$$

En la ecuación A.1, $d(p_i, q)$ es la distancia euclidiana ordinaria entre p_i y q . El diagrama de voronoi $V(S)$, donde S es la familia de puntos o nodos. La frontera de una celda de voronoi está formada por aristas y vértices, un punto $q \in \mathbb{R}^2$ pertenece a una arista que separa a las celdas $V(p_i)$ y $V(p_j)$ del diagrama de Voronoi si:

$$d(q, p_i) = d(q, p_j) \wedge d(q, p_k) \geq d(q, p_i) \forall i, j, k \neq i, j \quad (\text{A.2})$$

Un punto $q \in \mathbb{R}^2$ es un vértice de voronoi si esta presente en al menos dos aristas de voronoi. Un círculo C es un círculo vacío con respecto a S si no hay puntos de S dentro de del círculo. Para tres puntos cualesquiera, p_i, p_j y p_k , que no son colineales, existe una única circunferencia C_{ijk} que pasa por p_i, p_j y p_k . Un círculo C_{ijk} es un círculo de Voronoi si es un círculo vacío.

Lema A.1 $q \in \mathbb{R}^2$ es un vértice de Voronoi si y sólo si, q es el centro de un círculo de Voronoi.

La bisectriz $B(p_i, p_j)$ de dos puntos p_i y p_j es la recta que equidista de p_i y p_j . Es perpendicular al segmento de recta $p_i p_j$ que pasa por $\frac{1}{2}(p_i + p_j)$. Un punto q está en el semiplano de Voronoi $uh(p_i, p_j)$ si no está más cerca de p_j que de p_i :

$$uh(p_i, p_j) = \{q \in \mathbb{R}^2 \mid d(q, p_i) \leq d(q, p_j)\} \quad (\text{A.3})$$

En consecuencia, el semiplano de Voronoi $uh(p_i, p_j)$ está delimitado por la bisectriz $B(p_i, p_j)$. Toda célula de Voronoi es una intersección de semiplanos. Por lo tanto, cualquier célula de Voronoi $V(p_i)$ es convexa y está acotada o no.

Lema A.2 $V(p_i)$ no está limitada si y sólo si $p_i \in \delta CH(S)$, donde $\delta CH(S)$ denota la frontera del casco convexo $CH(S)$ [85].

$V(p_i)$ no está limitada si y sólo si $p_i \in \delta CH(S)$, donde $\delta CH(S)$ denota la frontera del casco convexo $CH(S)$.

Lema A.3 Sean p_i, p_j y p_k tres puntos distintos que no son colineales. Sea c el centro de la circunferencia que pasa por p_i, p_j y p_k [85]. Entonces:

$$c = B(p_i, p_j) \cap B(p_i, p_k) \quad (\text{A.4})$$

Por lo tanto, el punto c es equidistante a p_i, p_j y p_k .

Anexo B

Sistema Nacional de Información de Calidad del Aire SINCA

B.1. Estaciones de Monitoreo

A continuación, en la tabla se observa un detalle de las estaciones de monitoreo de la Región Metropolitana. A excepción de las estaciones de monitoreo de “Quilicura 1” y “Cerrillos 1”, que dejaron de emitir información de calidad del aire en 2016 y 2017 respectivamente, todas las demás estaciones se encuentran actualmente operativas y con actualización cada una hora en la plataforma del SINCA, consultar en <https://sinca.mma.gob.cl/index.php/region/index/id/M>. Por su lejanía, la estación de Talagante fue excluida de este estudio.

Tabla B.1: Estaciones de monitoreo Región Metropolitana. Elaboración propia.

Estaciones Sinca	Ubicación aproximada
Parque Ohiggins	Centro Parque
Quilicura 1	Estadio Municipal Quilicura
Quilicura	Multicancha Parque Central
Pudahuel	Av. El Lazo & Enco
Cerro Navia	Paula Jaraquemaca & Las Siemprevivas
Cerrillos 1	Jardín Infantil Los Cerrillos
El Bosque	Av. Padre Hurtado & Riquelme
La Florida	Gimnasio Municipal La Florida
Puente Alto	Centro Deportivo Laurita Vicuña
Talagante	Casa de la Cultura de Talagante
Las Condes	Fray Luis de León N° 11023

Las estaciones de monitoreo corresponden a instalaciones destinadas a medir y registrar regularmente diversas variables de calidad de aire y/o meteorológicas mediante equipos automáticos. Las normas de calidad de aire para efectos de evaluación de norma exige que éstas sean declaradas como Estación de monitoreo con representación poblacional (EMRP) y estación de monitoreo con representación poblacional por gases (EMRPG). Los requisitos que deben cumplir se encuentran en las normas (Ver Dto 59/1998, Dto 112/2002, Dto 104/2018).

Dto 114/2002 y Dto 115/2002 todos de MINSEGPRES).

Antes de que los datos de monitoreo sean aceptados en una base de datos final, los datos erróneos deben ser filtrados o extraídos. Este proceso de filtrado recibe el nombre de validación, y corresponde a la verificación de la exactitud, integridad y consistencia de la información generada, tomando en consideración criterios cuantitativos y cualitativos [35].

Los datos de monitoreo en línea, enviados por los operadores de las redes a SINCA, pueden presentar una serie de inexactitudes inherentes al proceso, como valores erráticos o inexistentes por cortes de energía, valores correspondientes a calibraciones y no a mediciones, valores no representativos debido a la presencia de fuentes ocasionales, etc. A estos datos se les denomina Registros no validados, ya que no han pasado por una validación por parte del operador de la red [35].

Una vez obtenidos los registros preliminares, estos deben pasar por un segundo proceso de validación, en el cual se verifica si la información es representativa y puede ser utilizada para evaluar las normas de calidad del aire primarias, estos datos tienen el nombre de Registros oficiales [35].

A continuación, en la figura B.1 se presenta el detalle de las técnicas de medición de la estación de monitoreo de las Condes.

Parámetro	Fecha primer registro	Fecha último registro	Técnica de medición	Gráficos
Dióxido de azufre (SO ₂ - µg/m ³)	08-05-1997	12-03-2009	FLUORESCENCIA PULSANTE - THERMO 43i	
Monóxido de nitrógeno (NO - ppb)	01-01-2000	21-11-2021	QUIMILUMINISCENCIA EN FASE GASEOSA - THERMO 42i	
Dióxido de nitrógeno (NO ₂ - ppb)	14-01-2005	2021-11-25	QUIMILUMINISCENCIA EN FASE GASEOSA - THERMO 42i	
Monóxido de carbono (CO - ppm)	05-04-1997	24-11-2021	FOTOMETRIA IR DE FILTRO DE CORRELACIÓN DE GAS-THERMO 48i	
Ozono- (O ₃ - ppb)	02-04-1997	2021-11-25	FOTOMETRIA UV - THERMO 49h	
Metano (CH ₄ - ppm)	20-03-1988	29-12-2010	No informado	
Óxidos de nitrógeno (NO _x - ppb)	01-01-2000	24-11-2021	QUIMILUMINISCENCIA EN FASE GASEOSA - THERMO 42i	
Carbono orgánico	01-01-2001	09-12-2004	No informado	
Carbono total	01-01-2001	09-12-2004	No informado	
Hidrocarburos no metánicos (HCNM - ppm)	16-12-2003	30-12-2010	No informado	
Material particulado MP 10 (MP 10 - µg/m ³)	24-04-1997	2021-11-25	ATENUACION BETA- MET ONE 1020	
Material particulado MP 2.5 (MP 2.5 - µg/m ³)	01-01-2000	2021-11-25	ATENUACION BETA- MET ONE 1020	

Figura B.1: Técnicas de medición estación Las Condes. [4]

B.2. Contaminantes Monitoreados SINCA

En Chile, la Ley sobre Bases Generales del Medio Ambiente (Ley N° 19.300, modificada por la Ley N° 20.417/2010, ambas del Ministerio Secretaría Regional de la Presidencia), establece en su artículo 32, la existencia de dos tipos de normas de calidad ambiental: primarias y secundarias. Las normas de calidad primarias, son aquellas normas de calidad ambiental que tienen como objetivo proteger la salud de la población humana dentro del territorio nacional;

las normas de calidad secundaria, tienen por objetivo proteger o conservar el medio ambiente o la naturaleza y son de carácter local y no necesariamente nacional. Una definición más precisa de norma primaria y secundaria es la que se establece en artículo 2º de la Ley:

Norma Primaria de Calidad Ambiental: aquella que establece los valores de las concentraciones y períodos, máximos o mínimos permisibles de elementos, compuestos, sustancias, derivados químicos o biológicos, energías, radiaciones, vibraciones, ruidos o combinación de ellos, cuya presencia o carencia en el ambiente pueda constituir un riesgo para la vida o la salud de la población.

Por ejemplo, la norma primaria de material particulado definido en el artículo 5 de la ley 19.300, que plantea lo siguiente:

Artículo 5: Defínase como niveles que originan situaciones de emergencia ambiental para material particulado respirable $MP_{2,5}$ aquéllos en que la concentración de 24 horas se encuentre dentro de los rangos que da cuenta la siguiente tabla:

Tabla B.2: Niveles de calidad del aire de norma primaria de $MP_{2,5}$.
Elaboración propia.

Nivel	Concentración 24 horas $MP_{2,5}$ ($\mu\text{g}/\text{m}^3$)
1 Alerta	80-109
2 Pre emergencia	110-169
3 Emergencia	170 o superior

Norma Secundaria de Calidad Ambiental aquella que establece los valores de las concentraciones y períodos, máximos o mínimos permisibles de sustancias, elementos, energía o combinación de ellos, cuya presencia o carencia en el ambiente pueda constituir un riesgo para la protección o la conservación del medio ambiente, o la preservación de la naturaleza.

A continuación se presentan las normas primarias de calidad del aire, actualmente vigentes [86] :

1. Norma de calidad del aire para $MP_{2,5}$ (D.S. N° 12/2010 del Ministerio de Medio Ambiente)
2. Norma de calidad del aire para MP_{10} (D.S. N° 59/1998, modificado por D.S. N° 45/2001, ambos del Ministerio Secretaría General de la Presidencia)
3. Norma de calidad del aire para SO_2 (D.S. N° 104/2018 del Ministerio de Medio Ambiente)
4. Norma de calidad del aire para NO_2 (D.S. N° 114/2002 del Ministerio Secretaría General de la Presidencia)
5. Norma de calidad del aire para CO (D.S. N° 115/2002 del Ministerio Secretaría General de la Presidencia)

6. Norma de calidad del aire para Plomo (D.S. N° 136/2000) del Ministerio Secretaría General de la Presidencia)
7. Norma de calidad del aire para O₃ (D.S. N° 112/2002 del Ministerio Secretaría General de la Presidencia)

En lo que respecta a las normas secundarias de calidad del aire, las siguientes son las vigentes en nuestro país:

1. Norma de calidad del aire para SO₂ (D.S. N° 22/2009 del Ministerio Secretaría General de la Presidencia)
2. Norma de calidad del aire para MPS en la cuenca del río Huasco, III Región (D. Exento N° 4/1992 del Ministerio de Agricultura)

Muchos proyectos privados están obligados a monitorear sus emisiones según lo establecido en su respectiva Resolución de Calificación Ambiental (RCA), por lo cual deben constituir sus propias estaciones de monitoreo, por ejemplo, las redes de Cemento Melón, Central Renca, Petrox, San Isidro-Nehuenco, entre otras. Desde el 1 de enero de 2012, el Ministerio del Medio Ambiente se ha hecho cargo de la administración de todas las redes públicas que existían en el país, las que antes de esa fecha pertenecían al Ministerio de Salud (Red SIVICA, que abarca desde la región de Coquimbo a Aysén, sin contar la Región Metropolitana) y a la Seremi de Salud de la Región Metropolitana (Red MACAM III). Actualmente, la red del SINCA concentra información de estaciones publicas y privadas.

A continuación, en la tabla B.3 se observa un detalle de los contaminantes monitoreados por las distintas estaciones de monitoreo de la Región Metropolitana, que se encuentran disponibles en la plataforma del SINCA. Los contaminantes monitoreados de izquierda a derecha corresponden al Material particulado 10, Material particulado 2,5, Dióxido de azufre, Dióxido de nitrógeno, Óxidos de nitrógeno, Óxido de nitrógeno, Monóxido de carbono, Ozono, Metano e Hidrocarburos no metánicos. Las "X" indican que el contaminante es medido por la estación de monitoreo.

Tabla B.3: Ejemplo de tablas con colores de filas.

Estación	MP 1,0	MP 2,5	SO₂	NO₂	NO_x	NO	CO	O₃	CH₄	HCNM
Parque O'higgins	X	X	X	X	X	X	X	X	X	X
Quilicura 1	X	X	X	X	X	X	X	X	X	X
Pudahuel	X	X	X	X	X	X	X	X	X	X
Cerro Navia	X	X	X	X	X	X	X	X	X	X
El Bosque	X	X	X	X	X	X	X	X	X	X
La Florida	X	X	X	X	X	X	X	X		
Puente Alto	X	X	X	X	X	X	X	X	X	X
Talagante	X	X	X	X	X	X	X	X	X	X
Las Condes	X	X	X	X	X	X	X	X	X	X

Anexo C

Diagramas de Voronoi Estaciones Sinca

A continuación se presentan los diagramas de voronoi de todas las estaciones cuando estas fueron excluidas y consideradas de la división de Voronoi. En los diagramas que consideran la ubicación de la estación, esta sera representada por la celda de Voronoi numero 8 (C8).

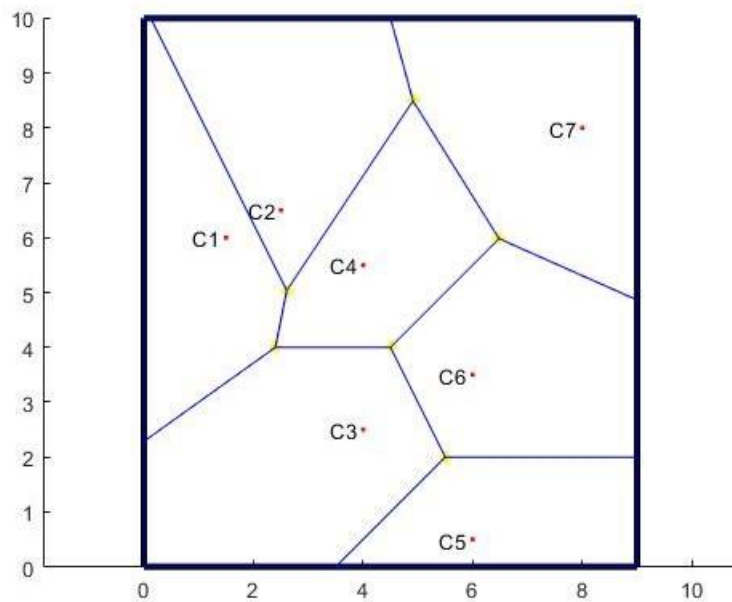


Figura C.1: Diagrama de Voronoi excluyendo estación de Quilicura para validación. Temporada Verano 2021

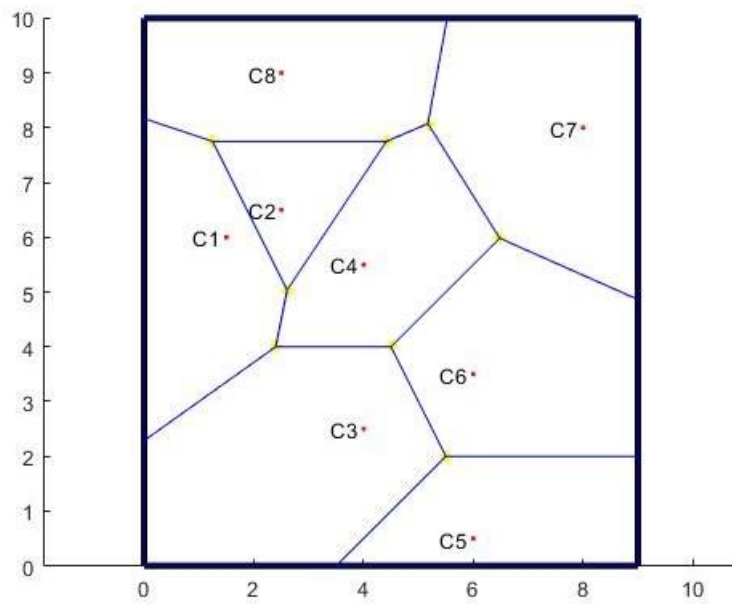


Figura C.2: Diagrama de Voronoi considerando la estación de Quilicura para validación. Temporada Verano 2021

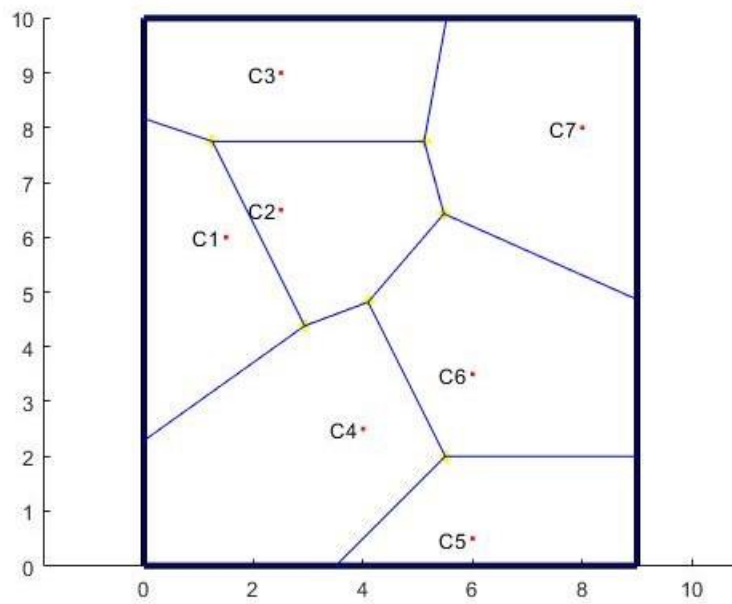


Figura C.3: Diagrama de Voronoi excluyendo estación de Parque O'higgins para validación. Temporada Verano 2021

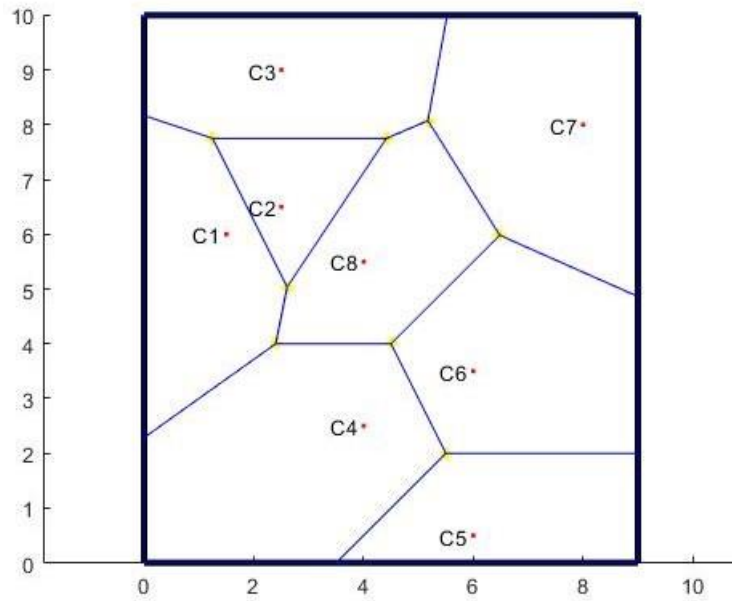


Figura C.4: Diagrama de Voronoi considerando la estación de Parque O'higgins para validación. Temporada Verano 2021

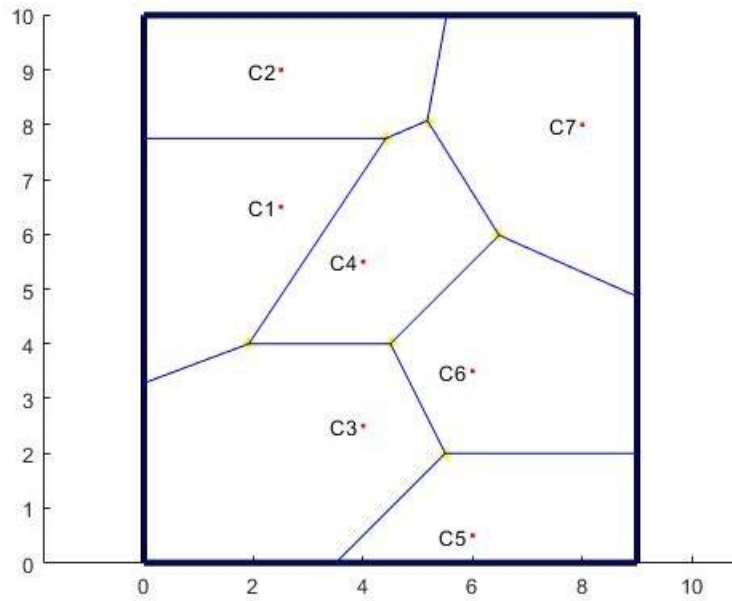


Figura C.5: Diagrama de Voronoi excluyendo estación de Pudahuel para validación. Temporada Verano 2021

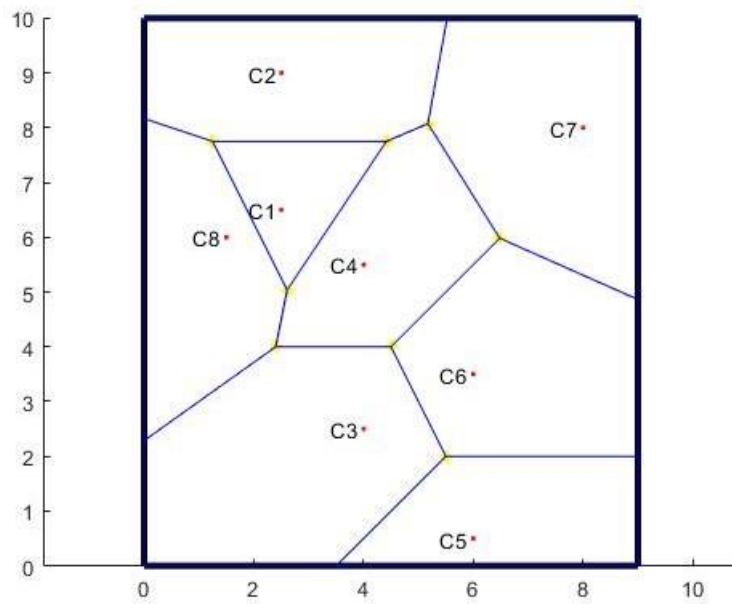


Figura C.6: Diagrama de Voronoi considerando la estación de Pudahuel para validación. Temporada Verano 2021

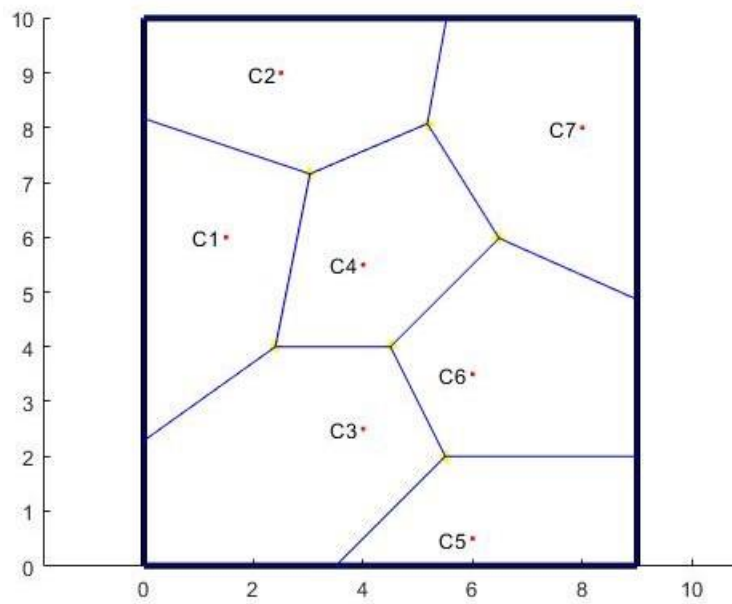


Figura C.7: Diagrama de Voronoi excluyendo estación de Cerro Navia para validación. Temporada Verano 2021

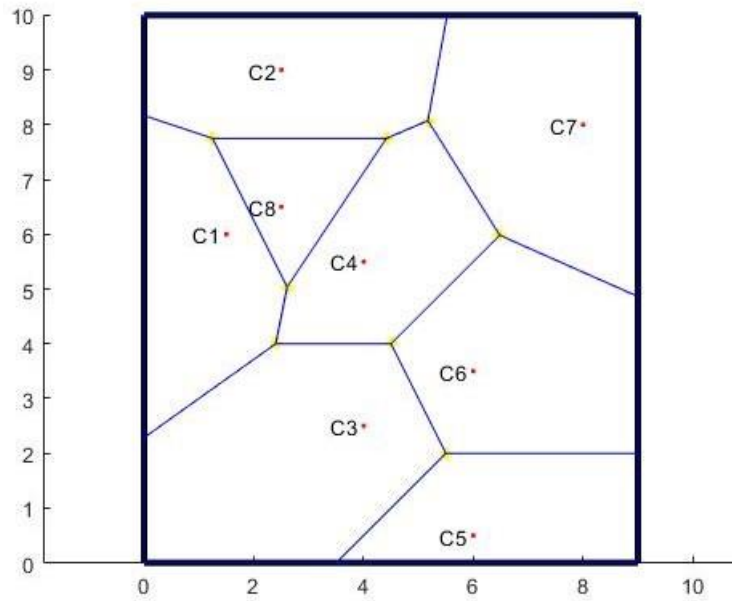


Figura C.8: Diagrama de Voronoi considerando la estación de Cerro Navia para validación. Temporada Verano 2021

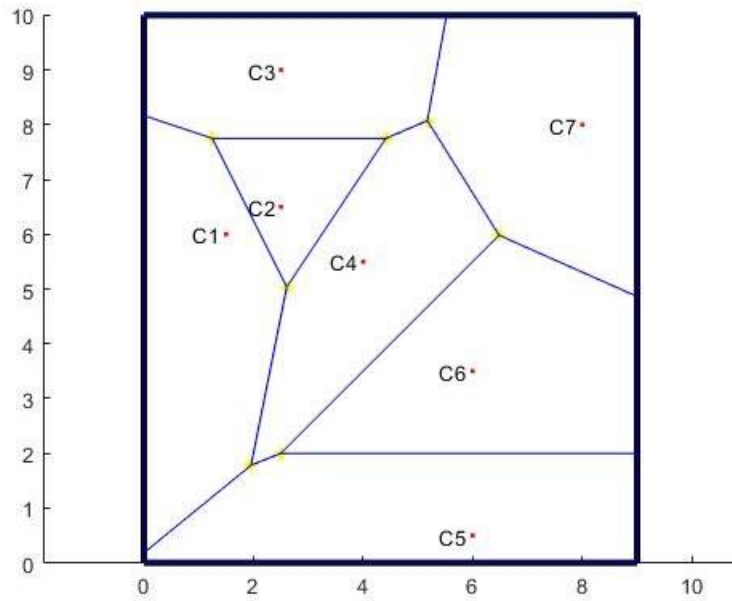


Figura C.9: Diagrama de Voronoi excluyendo estación de El Bosque para validación. Temporada Verano 2021

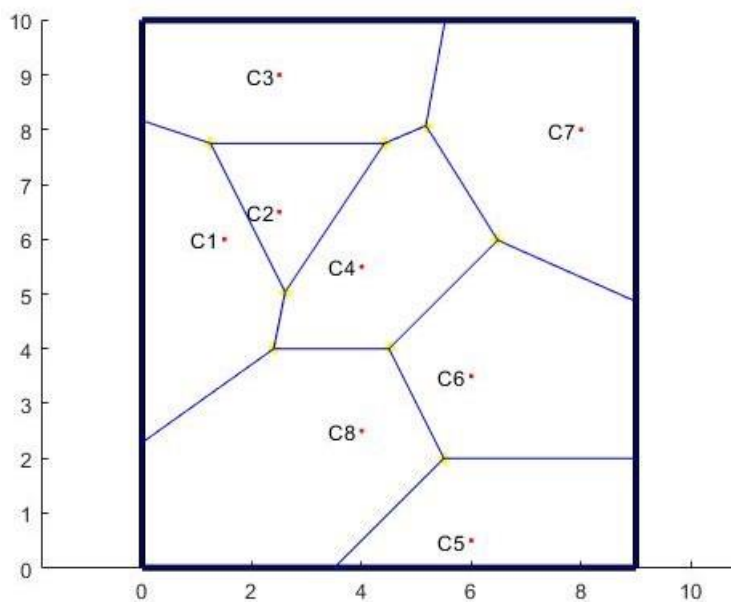


Figura C.10: Diagrama de Voronoi considerando la estación de El Bosque para validación. Temporada Verano 2021

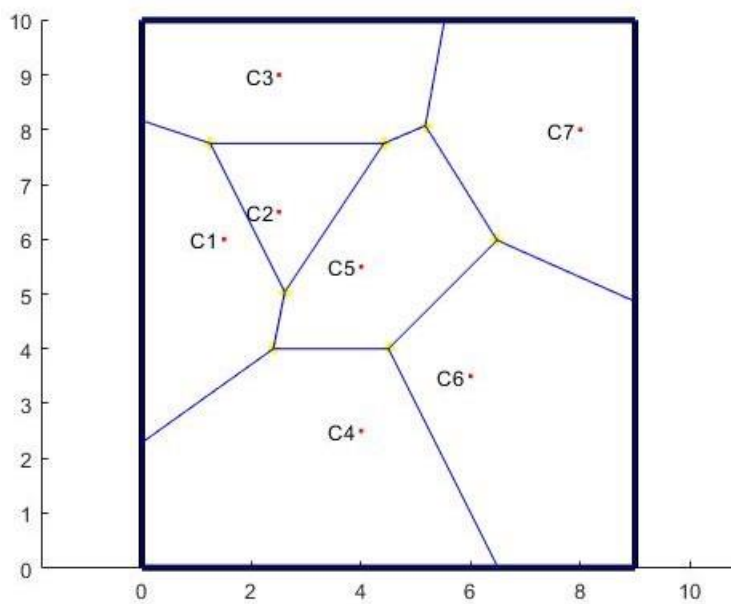


Figura C.11: Diagrama de Voronoi excluyendo estación de Puente Alto para validación. Temporada Verano 2021

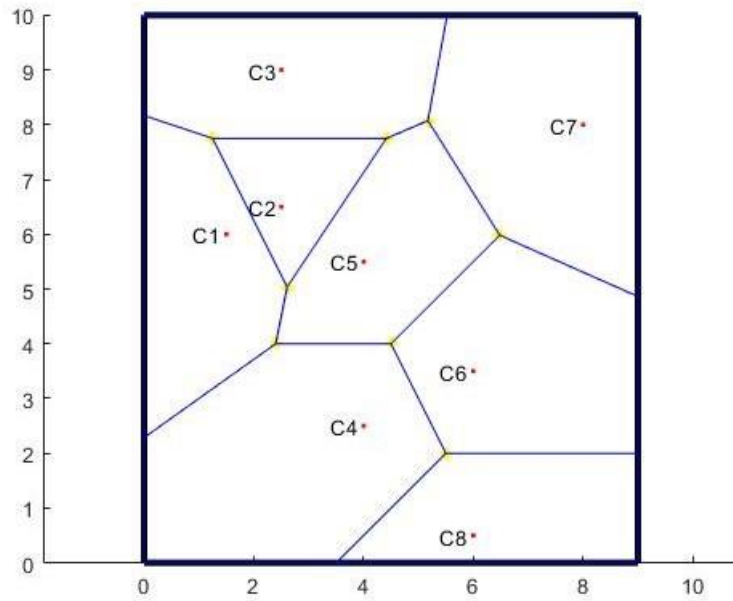


Figura C.12: Diagrama de Voronoi considerando la estación de Puente Alto para validación. Temporada Verano 2021

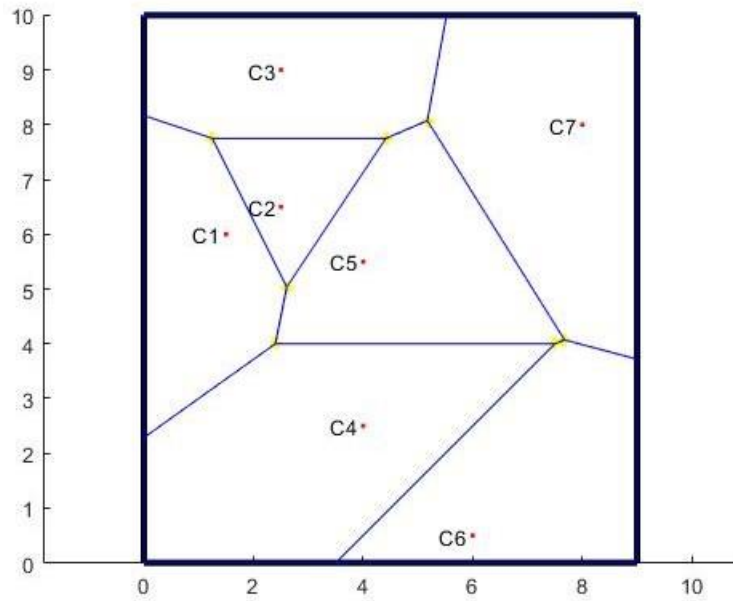


Figura C.13: Diagrama de Voronoi excluyendo estación de La Florida para validación. Temporada Verano 2021

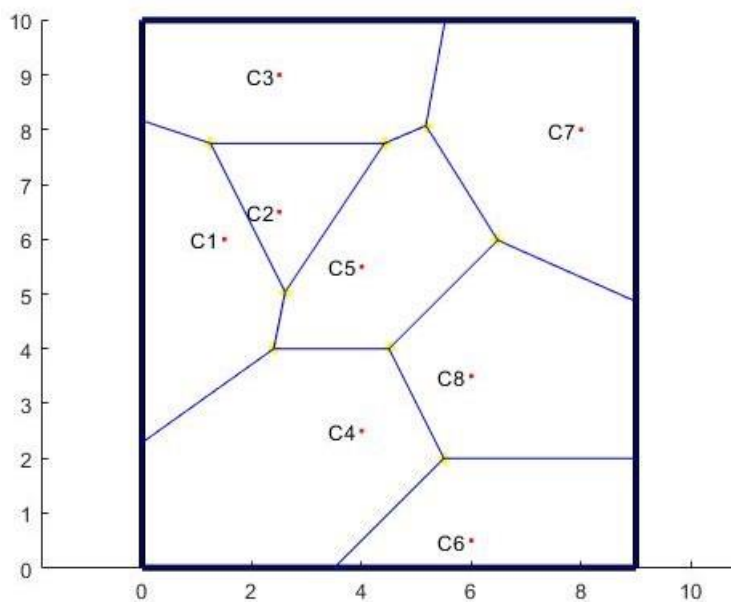


Figura C.14: Diagrama de Voronoi considerando la estación de La Florida para validación. Temporada Verano 2021

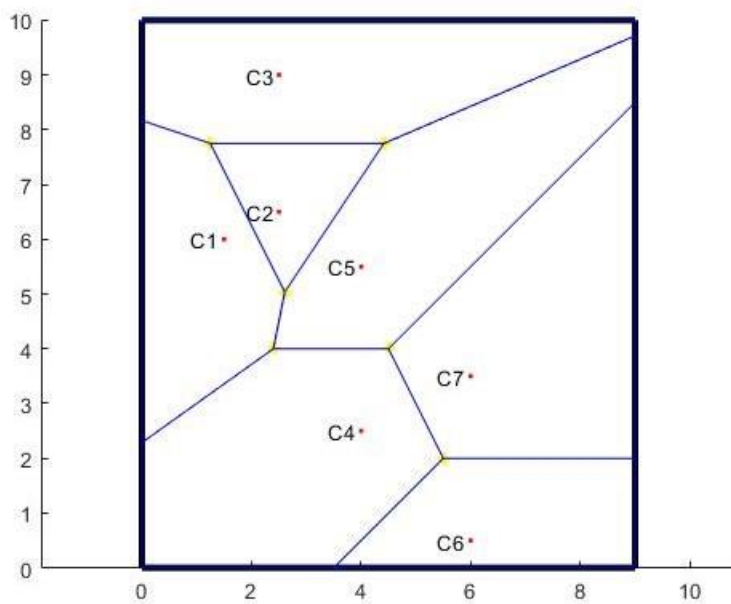


Figura C.15: Diagrama de Voronoi excluyendo estación de Las Condes ara validación. Temporada Verano 2021

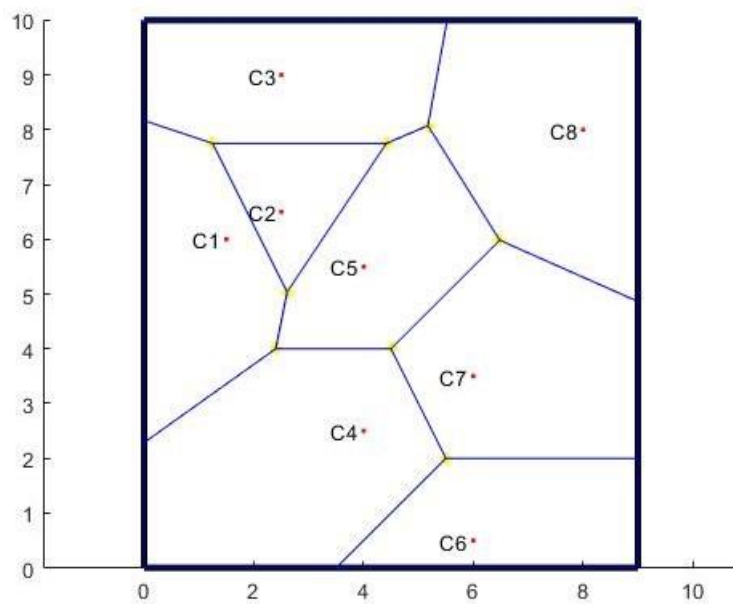


Figura C.16: Diagrama de Voronoi considerando la estación de Las Condes para validación. Temporada Verano 2021.

Anexo D

Modelos de dispersión

Los modelos de dispersión atmosférica son una de las soluciones más extendidas al problema del comportamiento de los contaminantes en la atmósfera. Su planteamiento general sigue el esquema de un modelo de procesos o fenómenos, es decir, consisten en aplicar un modelo matemático a cada fenómeno que experimentan los contaminantes en la atmósfera, y sumar las contribuciones de todos los modelos de proceso en una ecuación de balance o conservación para cada propiedad considerada. El balance de contaminante en la atmósfera aplicable de forma general es [16]:

$$\frac{dC_i}{dt} + \frac{d}{dx_j} (\mu_j C_j) = \frac{\partial^2 C_i}{\partial x_j^2} + R_i(c_1, \dots, c_n, T) + S(x, t) \quad (D.1)$$

1. Modelos de caja: Están basados en el principio de conservación de masa. El sitio de interés es tratado como una caja dentro de la cual se emiten contaminantes y sufren transformaciones químicas. Requieren de la entrada de datos simples de la meteorología del lugar y emisiones, el movimiento dentro y fuera de contaminantes en las fronteras de la caja es permitido. Dentro de la caja se hace la suposición de que todos los contaminantes están perfectamente mezclados. Una de las ventajas de este modelo es que la meteorología simplificada del mismo permite calcular complejas reacciones químicas, al igual que la dinámica de aerosoles. Su mayor desventaja es que no pueden determinar concentraciones locales[57]. Estos modelos son los más simples y aquellos que se utilizan en el caso de estudiar una emisión difusa o no puntual. El caso típico es el de una ciudad donde la fuente principal de contaminación es el tráfico rodado.

Los modelos de celda fija parten de asumir las siguientes hipótesis esenciales [48]:

- * La ciudad es un rectángulo de dimensiones W y L, con uno de sus lados paralelo a la dirección del viento (L).
- * La turbulencia atmosférica produce la mezcla completa de los contaminantes hasta la altura de mezclado (H), y no hay mezcla por encima de esa altura, con lo que se puede asumir que la concentración en la celda es homogénea en el espacio y de valor C.
- * El viento sopla con velocidad u en la dirección x, con fuerza independiente del tiempo y de la elevación sobre el suelo.
- * El viento sopla con velocidad u en la dirección x, con fuerza independiente del tiempo y de la elevación sobre el suelo.

concentración de contaminante que entra en la ciudad ($x=0$) es constante e igual a b (concentración de fondo).

- * El índice de emisiones por unidad de área es q (por ejemplo, en $gs^{-1}m^{-2}g$), con lo que la emisión total es $Q = q \cdot A$, siendo $A = W \cdot L$ el área de la ciudad.
- * El índice de emisiones no varía ni con el tiempo ni con el viento.
- * Ningún contaminante entra o sale por los lados perpendiculares a la dirección del viento.

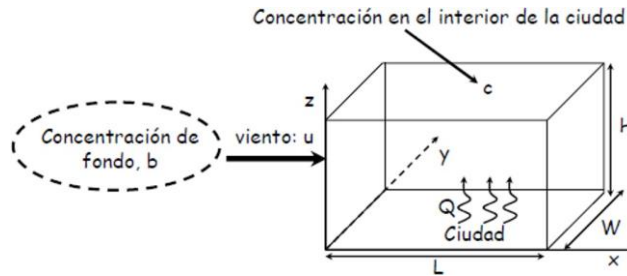


Figura D.1: Ejemplo modelo de caja [87]

$$u * b * W * H + q * W * L - \mu * C_e * W * H = W * L * H * \frac{\partial C}{\partial t} \quad (D.2)$$

2. Modelos Gaussianos. Este tipo de modelos son ampliamente utilizados en el modelado de dispersión de contaminantes y en particular con fines regulativos y a menudo son integrados dentro de modelos Euleriano y Lagrangianos. El principio de funcionamiento de los modelos Gaussianos está basado en distribuciones Gaussianas de plumas tanto en dirección vertical como horizontal bajo condiciones de estado permanente [57].

Los modelos gaussianos son los modelos que emplean los técnicos para evaluar el efecto de emisiones puntuales hasta distancias que no superen los 20 Km de distancia. No son pues adecuados para tratar problemas que impliquen largas distancias tales como la lluvia ácida. El problema típico que trata un modelo gaussiano es el de una chimenea de una fábrica que emite, de forma continua o esporádica, una cierta cantidad de contaminante.

Es importante señalar que, aunque el modelo gaussiano nos proporcione, en principio, la concentración de contaminante en cualquier punto del espacio (x,y,z), lo más común es trabajar para $z = 0$ (ras de suelo) y para $y = 0$ (en la dirección del viento). El caso de $z = 0$ es importante porque es en la superficie donde se encuentra la población y donde más interesa conocer las previsiones de contaminación. El modelo de la columna de humo gaussiana se basa en las siguientes hipótesis.

- a) La columna de humo emitida por la chimenea se eleva hasta cierta altura efectiva "H", que es la suma de la altura de la chimenea "h", más el ascenso "dh", debido a factores como la diferencia de temperaturas entre el gas saliente y el aire que le rodea.

- b) A partir de la altura efectiva “H” la columna de humo se mueve horizontalmente en la dirección del viento (dirección x) con velocidad u y se dispersa en las direcciones perpendiculares a éste, “y” y “z”. (y = horizontal, z = vertical). La dispersión en el plano yz se debe principalmente a la turbulencia atmosférica y puede calcularse como una distribución gaussiana.

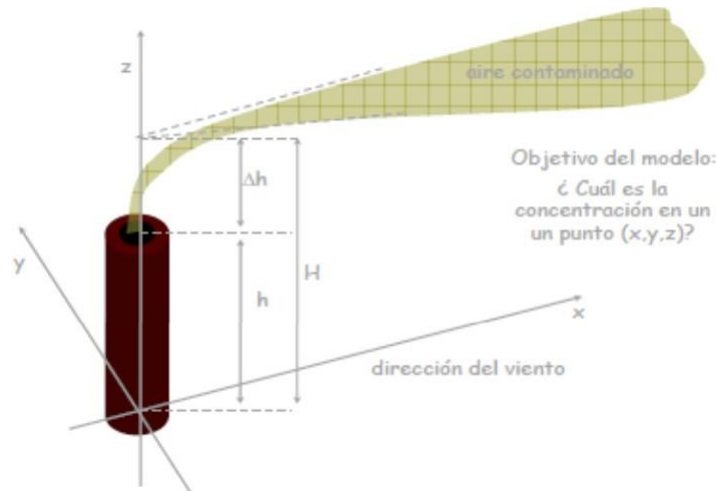


Figura D.2: Ejemplo modelo de gauss [87]

A continuación, se observa la solución del modelo de gauss en 2 dimensiones con $z=0$, equivalente a la altura de emisión.

$$C_{x,y,z} = \frac{Q}{2\pi\mu\sigma_y\sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2} + \frac{(z-H)^2}{2\sigma_z^2}\right) \quad (D.3)$$

Donde $C_{x,y,z}$ es la concentración Q es la cantidad de contaminante emitido por unidad de tiempo (gs^{-1}), μ es la velocidad del viento σ_y y σ_z son los coeficientes de dispersión turbulenta, que dependen de la clase de estabilidad y de la distancia al foco en la dirección del viento “x”.

3. Modelos eulerianos y lagrangeanos: Utilizan un sistema de referencia absoluto y tratan de calcular la difusión referida a este sistema mediante distintas soluciones de integración numérica. Una de ellas es la solución euleriana, según la cual se establece un sistema de referencia fijo (habitualmente, centrado en el foco emisor), y se tratan de resolver las ecuaciones de balance de materia, para cada contaminante, con diferentes grados de aproximación. La otra es la solución lagrangiana, que utiliza dos sistemas de coordenadas, uno fijo (el foco emisor) para el movimiento general de los contaminantes (advección) y otro móvil, que se desplaza con el penacho, para la descripción de la difusión turbulenta en el seno del penacho (Seinfeld, 2012). En ambas soluciones, se asume el conocimiento de las variables que definen el flujo atmosférico y condicionan la dispersión de los contaminantes.

4. Kriging: Métodos de geoestadística basados en interpolaciones espaciales, se reconoce el Kriging Simple o Universal. Estas metodologías analizan la correlación espacial de los puntos existentes, su importancia radica en que permite describir la continuidad espacial de las variables y estimar puntos cercanos, de los cuales se desconoce su valor. Por ejemplo, A. Patricia. Et al (2002) [59] realiza un estudio comparativo entre un modelo geoestadístico Kriging y un modelo gaussiano.

Los métodos kriging, además, proporcionan la herramienta del semi variograma, que permite tener una mayor comprensión de los datos, ya que en geoestadística es fundamental conocer la correlación espacial existente entre ellos, que consiste en estimar la dependencia espacial entre los datos de diversos puntos, la cual, se realiza mediante un análisis estructural del variograma. Los variogramas son estimadores de la varianza poblacional relacionados con la dirección y la distancia e indican como varia la dependencia espacial entre los puntos.

El método geoestadístico se basa en el supuesto de estacionalidad de la variable de interés y la existencia de una distribución de probabilidad. En caso que la variable de interés presente una distribución normal, entonces, el método será más exacto[59], por lo cual la eficacia de este método depende de la uniformidad del área de estudio. Si los datos no presentan tal distribución, se debe realizar un tratamiento preliminar para poder utilizar el método.

El semivariograma está determinado por 3 parámetros, la meseta, el rango y el Nugget. La meseta o silo es la distancia a la cual la semivariancia no aumentara más, el rango es la distancia entre el inicio del semivariograma y el comienzo de la meseta y al extrapolar la curva del semivariograma a la distancia cero se obtiene un valor no nulo denominado Nugget. En el método kriging comúnmente son utilizados 4 tipos de variogramas, el Lineal, Esférico, Exponencial y Gaussiano. Estos variogramas son función de la varianza, el nugget, el silo, la varianza asintótica y la distancia de separación.

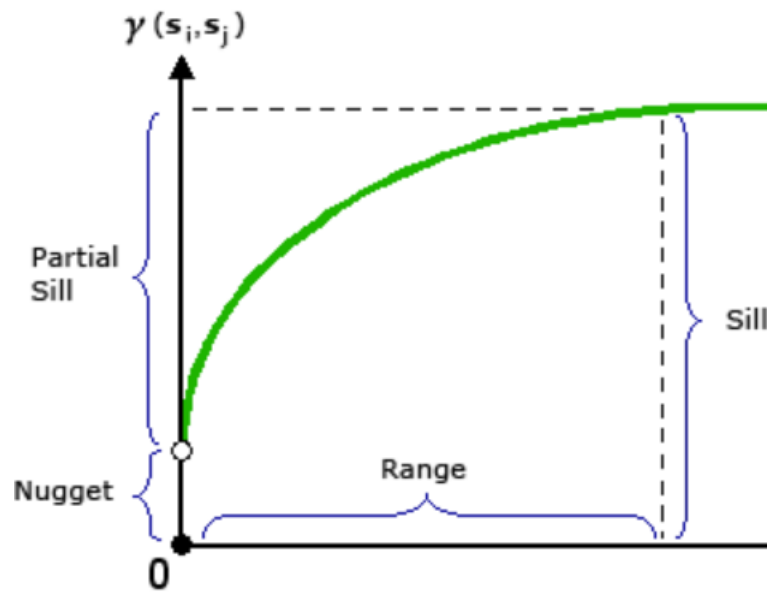


Figura D.3: Ejemplo variograma [87].

Anexo E

Software

La Figura 13 muestra el diagrama de flujo para el procesamiento de información en el modelo de dispersión AERMOD. Este consiste en un módulo principal (AERMOD) y dos preprocesadores (AERMET y AERMAP). Los datos de la superficie y del aire en altitud, proporcionados por NWS, son entradas importantes para los modelos de calidad del aire. Antes de que estos datos se utilicen en algunos de los modelos de dispersión de la EPA, se utilizan procesadores meteorológicos para manipular los datos. Estos procesadores utilizan los datos de observación del NWS para calcular las entradas importantes para los modelos de dispersión, uno de ellos es el procesador AERMET [88].

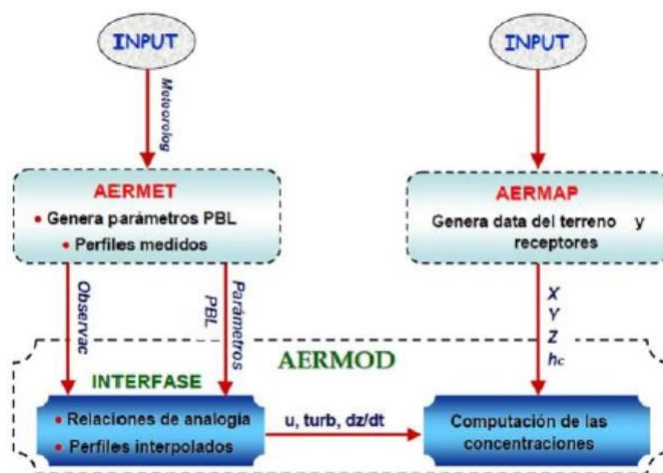


Figura E.1: Esquema de procesamiento de datos sistema AERMOD [69].

AERMET usa mediciones meteorológicas representativas del dominio de modelación para calcular parámetros de la capa límite. La INTERFASE interna de AERMOD usa estos parámetros calculados para generar perfiles de viento, temperatura y turbulencia. La profundidad de la capa límite y la dispersión de los contaminantes dentro de ésta están influenciadas a escala local por las características superficiales: rugosidad, albedo y radio de Bowen. Estas variables de superficie, junto con observaciones meteorológicas básicas (velocidad de viento, dirección de viento, temperatura y cobertura de nubes) son datos requeridos por AERMET para calcular los siguientes parámetros de la capa límite planetaria.

1. Velocidad de fricción, μ
2. Longitud de Monin-Obukhov, L
3. Escala de velocidad convectiva, w^*
4. Escala de temperatura, θ^*
5. Altura de la capa de mezcla convectiva (zic) y mecánica (zim)
6. Flujo de calor superficial, H

Tipo		Símbolo	Descripción
Terreno		ED	Elevación digital
		US	Uso de suelos
Meteorología	Parámetros meteorológicos	v	Velocidad del viento
		T	Temperatura ambiente
		Dv	Dirección del viento
		EA	Estabilidad Atmosférica
		Hrur	Altura de capa de mezcla rural
		Hurb	Altura de capa de mezcla urbana
		f v	Velocidad de fricción

Figura E.2: Tabla parámetros de entrada sistema AERMOD [69].

Anexo F

PID

Los sensores de foto ionización es un tipo de detector de gas que mide los compuestos orgánicos volátiles y otros gases. Son los detectores de gas más eficientes y asequibles, su funcionamiento se base en la ionización de un gas por medio de una luz ultravioleta y el par electrón-Catión formado es atraído hacia el ánodo y cátodo respectivamente. En la figura 11 se puede observar un detalle de este proceso de ionización. La diferencia de potencial generada es registrada por la electrónica del dispositivo, lo que permite identificar la concentración del gas en ppb o ppm.

Figura Esquema de foto ionización de COV en sensor PID [89]

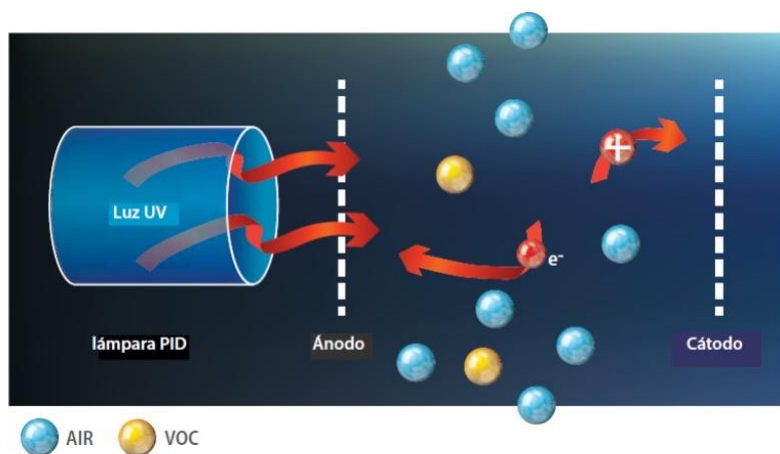


Figura F.1: Esquema de funcionamiento sensor de fotoionización.

Los PID pueden medir compuestos orgánicos volátiles (COV) a concentraciones muy bajas, los PID Alphasense son compactos y de bajo consumo, con componentes electrónicos mejorados recientemente y una mayor vida útil de la lámpara. El sensor Alphasense PID se ofrece en dos modelos, A12 y AH2. Son prácticamente insensibles a cambios de humedad, lo que proporciona un rendimiento incomparable en una variedad de aplicaciones [89].

- PID-A12 tiene un rango dinámico lineal de 50 ppb a 6.000 ppm (isobutileno).
- PID-AH2 tiene un rango dinámico lineal de 1 ppb a 50 ppm (isobutileno).

Los sensores PID-A1 y PID-AH miden compuestos orgánicos volátiles (COV) en el aire mediante detección de fotoionización (PID), que se muestra esquemáticamente en la figura 12. El gas de prueba (1) se presenta al filtro de membrana en la parte superior de la celda de fotoionización y se difunde libremente dentro y fuera de la cámara subyacente formada por el filtro, alojamiento paredes y una ventana con lámpara UV. La lámpara emite fotones (mostrados por flechas) de luz ultravioleta de alta energía, transmitido a través de la ventana. La fotoionización ocurre en la cámara cuando un fotón es adsorbido por la molécula, generando dos iones cargados eléctricamente, uno con carga positiva y otro con carga negativa. Un campo eléctrico, generado entre los electrodos del cátodo y el ánodo, atrae iones y la corriente resultante, que es proporcional a la concentración de COV, se mide y se utiliza para determinar la concentración de gas. El PID-A incluye un tercer electrodo de “Valla”, para asegurar que el amplificado corriente no incluya contribuciones significativas debidas a otras fuentes de corriente, como la condensación de agua en las paredes de la cámara.[61]

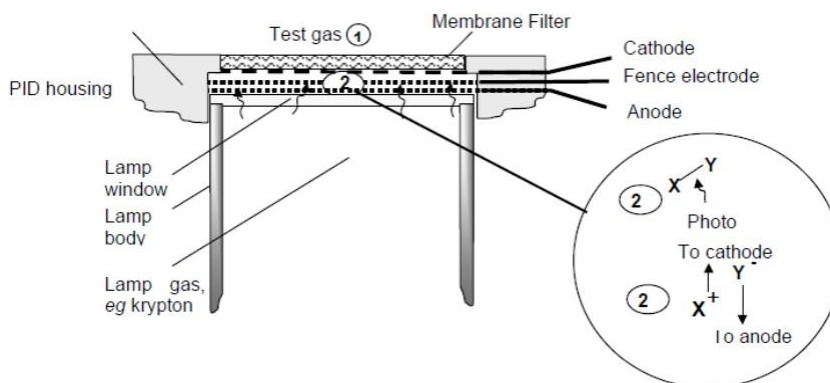


Figura F.2: Esquema funcionamiento interno sensor PID *alphasense*

La mayoría de los COV pueden detectarse mediante PID. Las excepciones notables son los hidrocarburos de bajo peso molecular. Cada El COV tiene una energía umbral característica de la luz (energía de fotones) que, cuando se dirige al COV, provoca que se fragmente en iones. Esto se denomina potencial de ionización o IP. Los COV se ionizan (y por tanto se detectan) si luz de energía fotónica mayor que el IP interactúa con la muestra de gas. La energía máxima de fotones generada en un detector depende de la lámpara PID utilizada: xenón = 9,6 eV, deuterio = 10,2 eV, criptón = 10,6 eV y Argón = 11,7 eV. Por lo tanto, el uso de una lámpara de argón conduce a la detección de la gama más amplia de compuestos volátiles, mientras que el uso de una lámpara de xenón puede aumentar la selectividad.

La elección de la lámpara depende de los gases objetivo, los requisitos de selectividad y las consideraciones sobre la vida útil de la lámpara. Por lo general, se utiliza la lámpara de criptón debido a su alta sensibilidad. Lámpara de xenón (9,6 eV) mide muchos compuestos aromáticos e insaturados de COV que contienen al menos 6 átomos de carbono. A veces es una ventaja utilizar la lámpara de xenón para asegurar una detección más selectiva de compuestos como BTEX. La lámpara de criptón (10,6 eV) detecta algunos COV de 2 carbonos, pero en general de 3 a más carbonos.

Anexo G

Resultados

G.O.1. Test 1

Verano

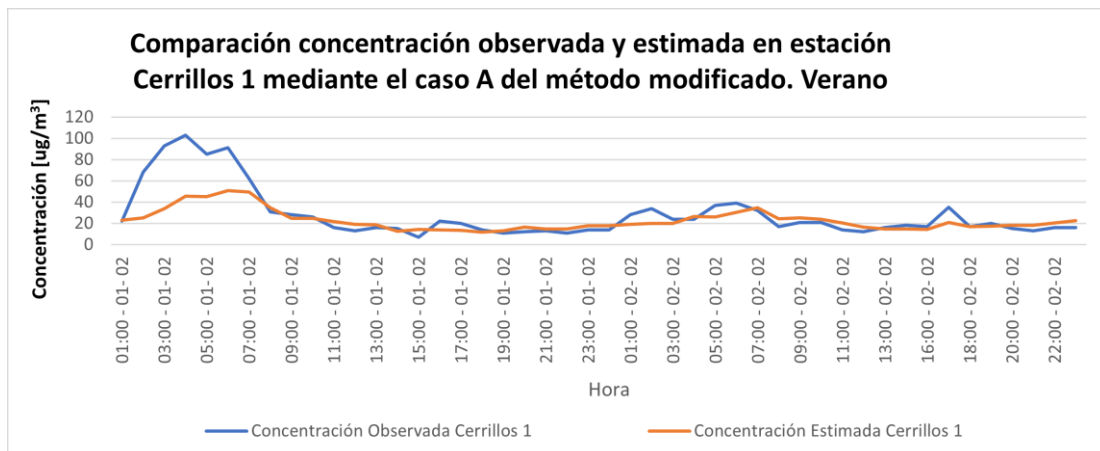


Figura G.1: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa A. Temporada Verano.

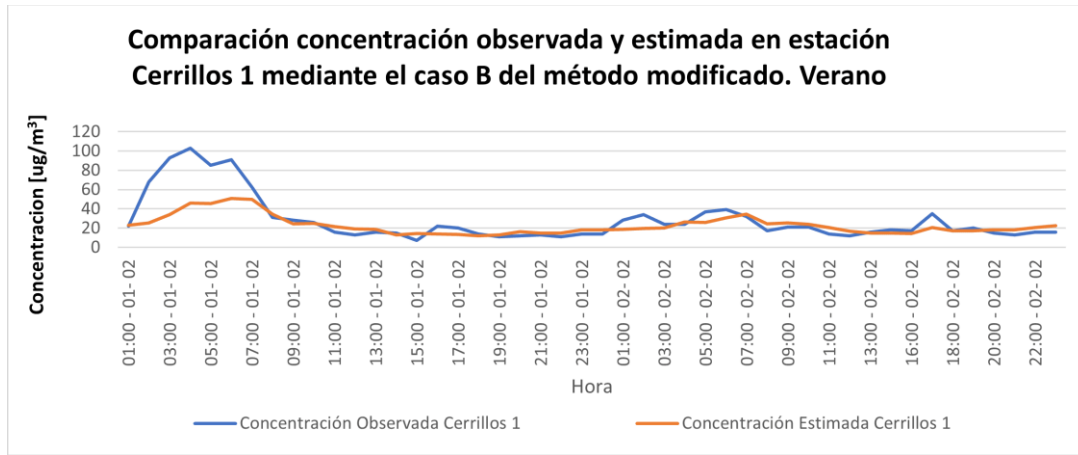


Figura G.2: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa B. Temporada Verano.

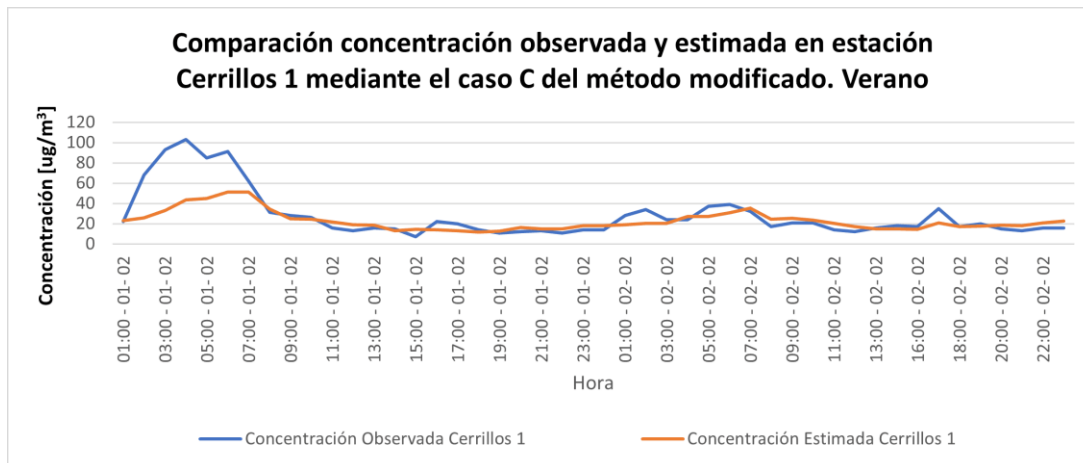


Figura G.3: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. Interpolación iterativa C. Temporada Verano.

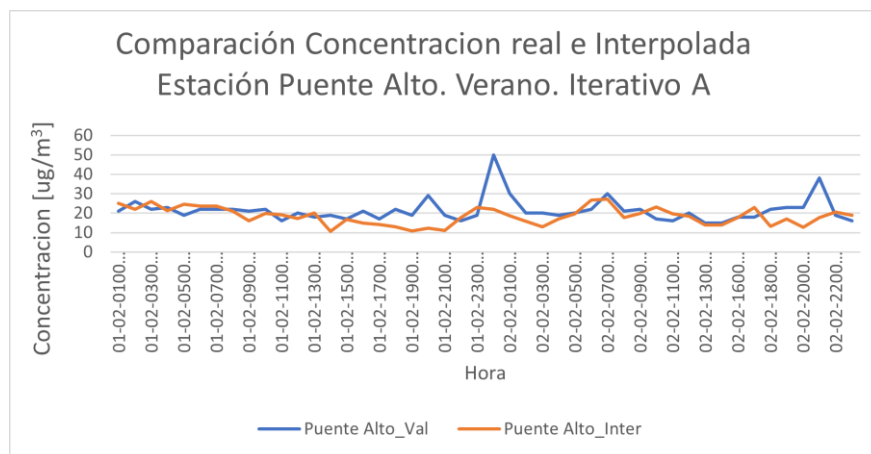


Figura G.4: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IA. Temporada Verano.

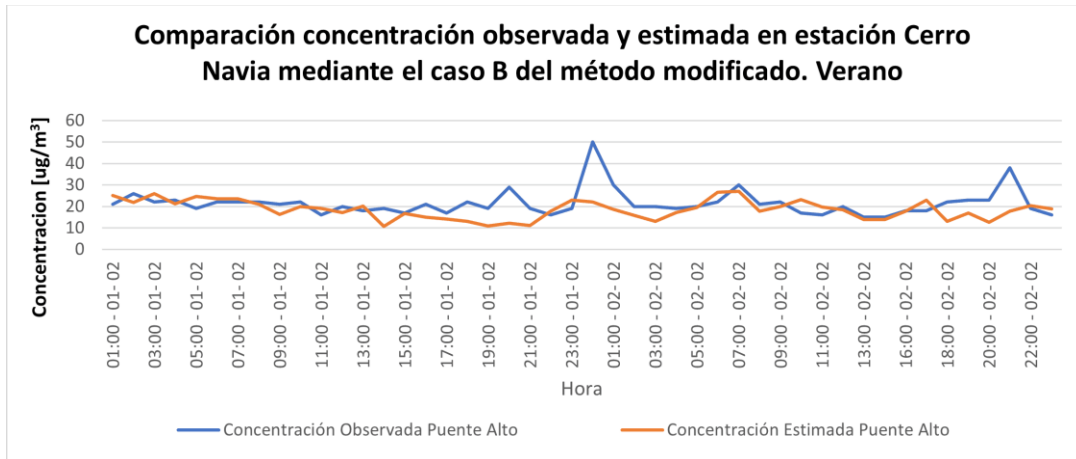


Figura G.5: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IB. Temporada Verano.

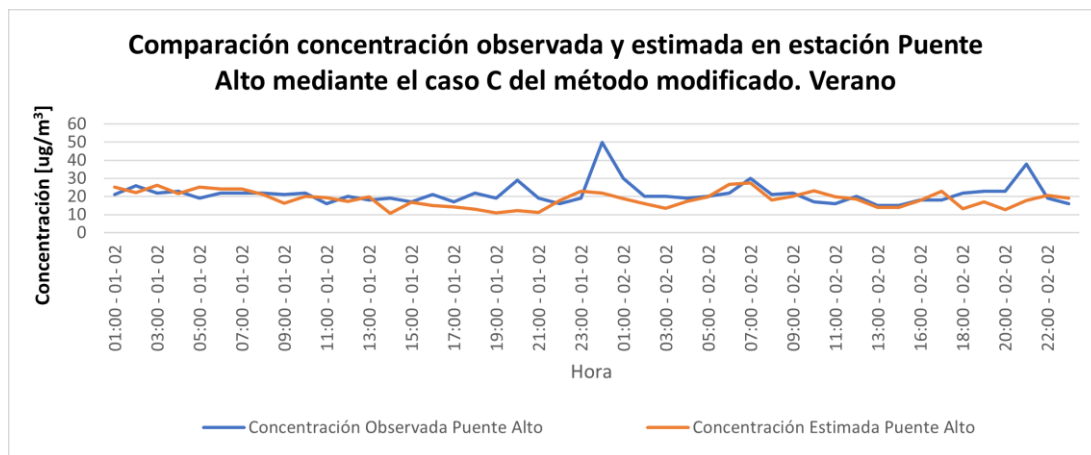


Figura G.6: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto. IC. Temporada Verano.

Invierno

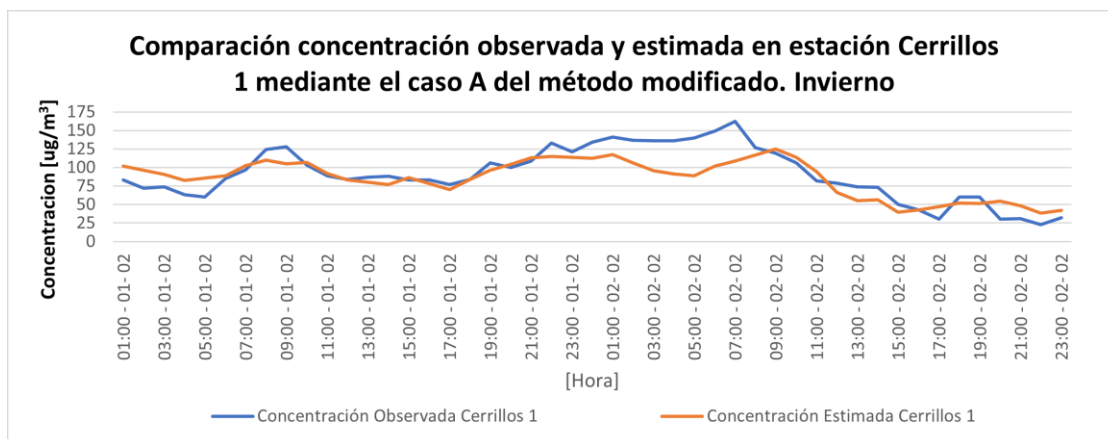


Figura G.7: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. IA. Temporada Invierno.

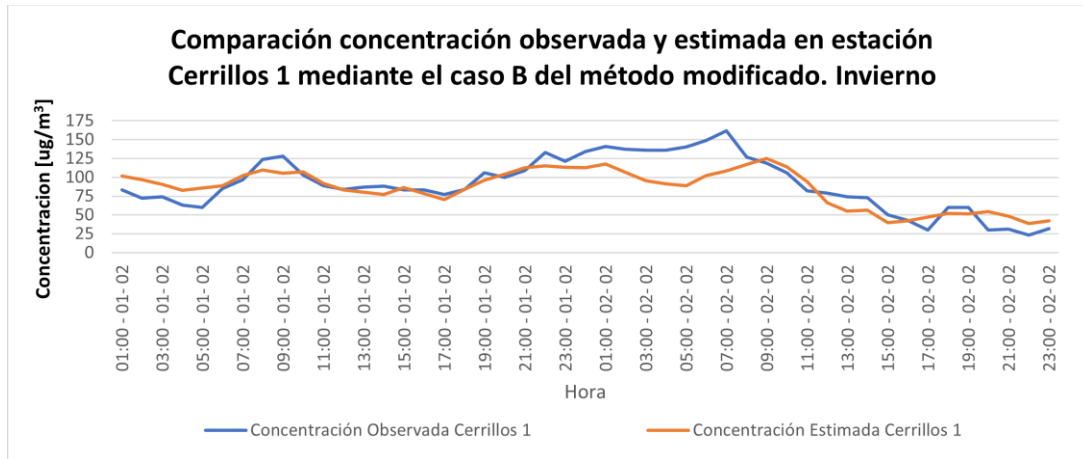


Figura G.8: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1. IB.Temporada Invierno.

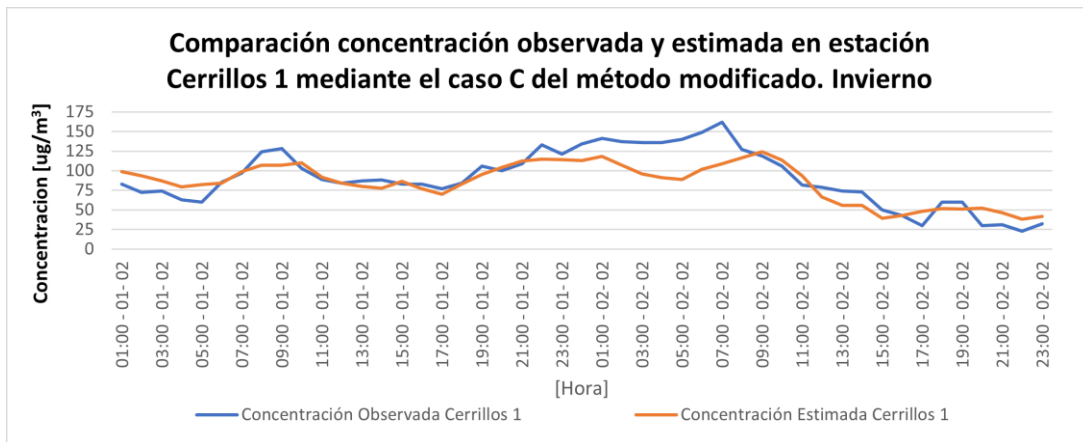


Figura G.9: Comparación Concentraciones reales e interpoladas en ubicación estación Cerrillos 1 Orden

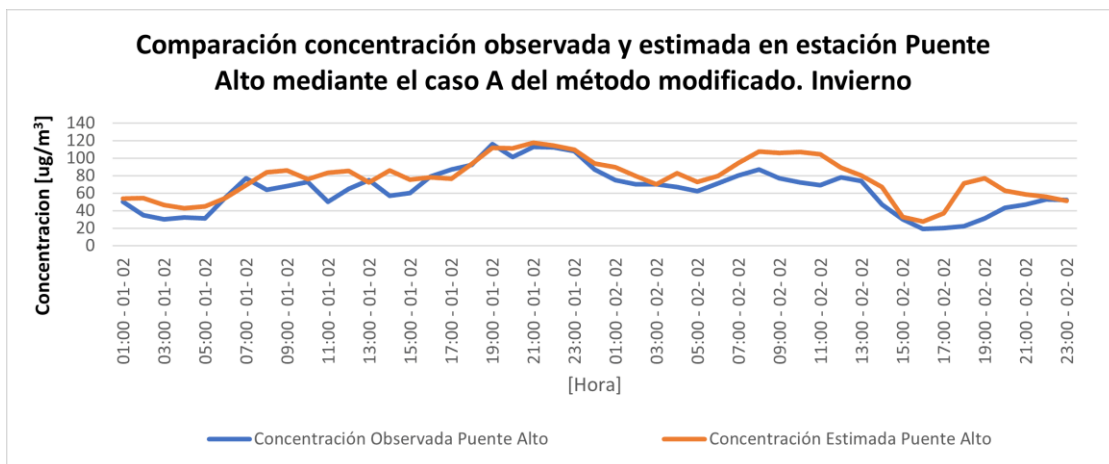


Figura G.10: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto A1

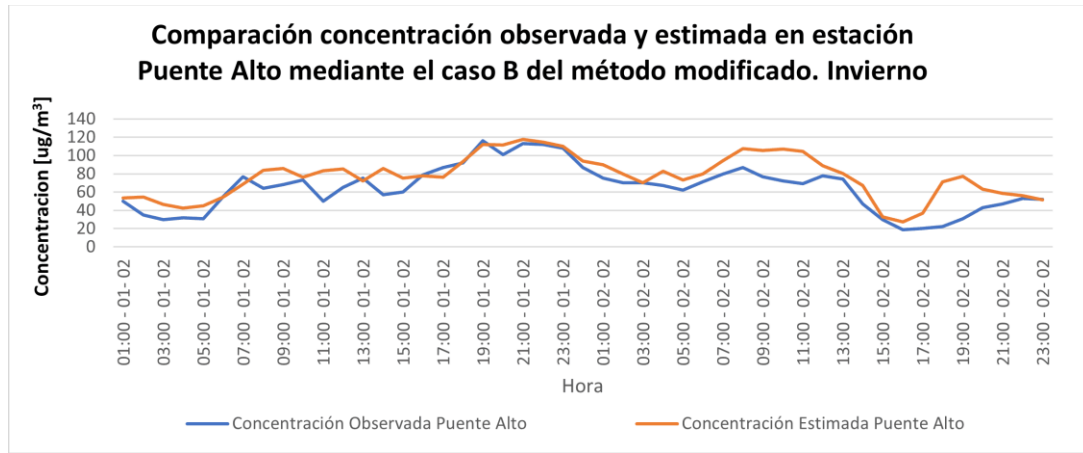


Figura G.11: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto A1

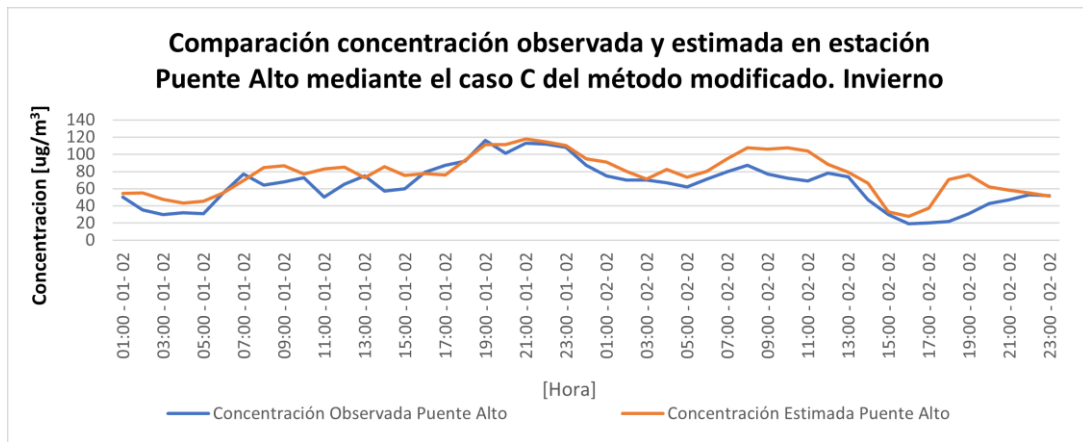


Figura G.12: Comparación Concentraciones reales e interpoladas en ubicación estación Puente Alto Orden

G.O.2. Test 2

Verano

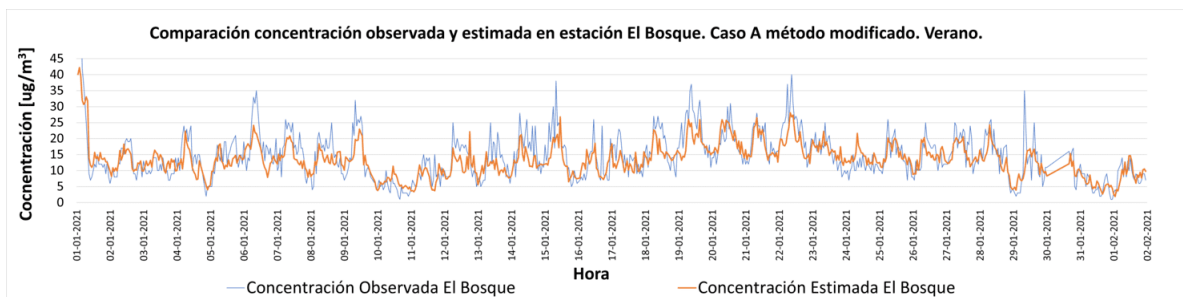


Figura G.13: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IA. Temporada Verano

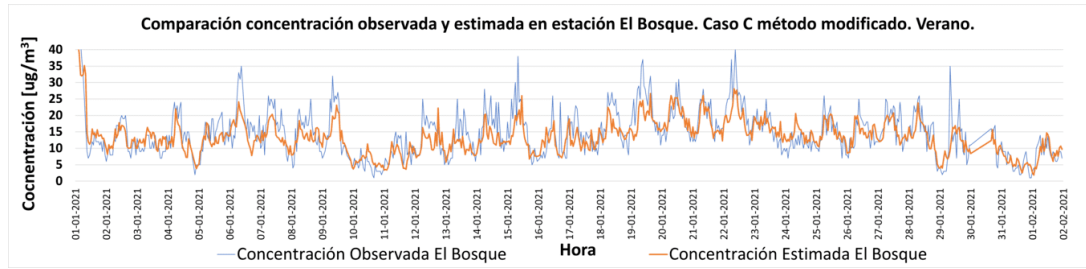


Figura G.14: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IC. Temporada Verano.

Invierno

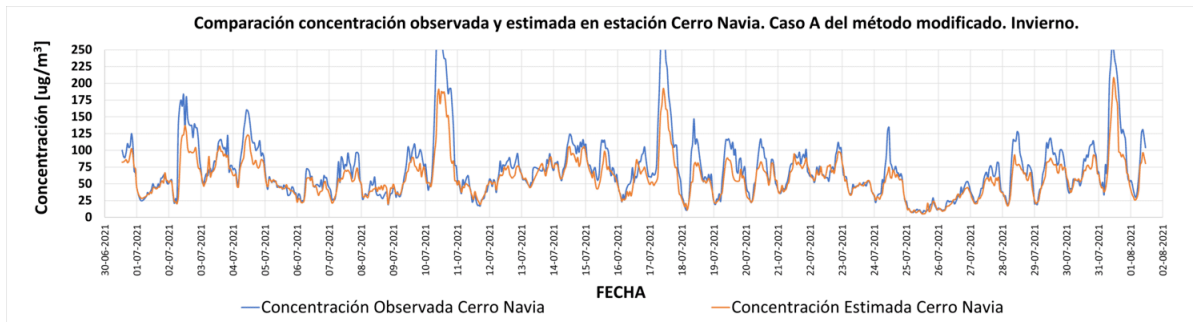


Figura G.15: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IA. Temporada Invierno.

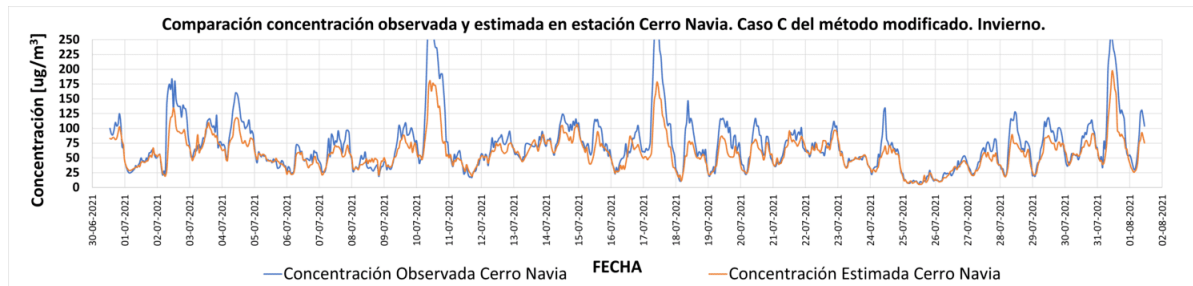


Figura G.16: Comparación Concentraciones reales e interpoladas en ubicación estación Cerro Navia para un mes de datos. IC. Temporada Invierno.

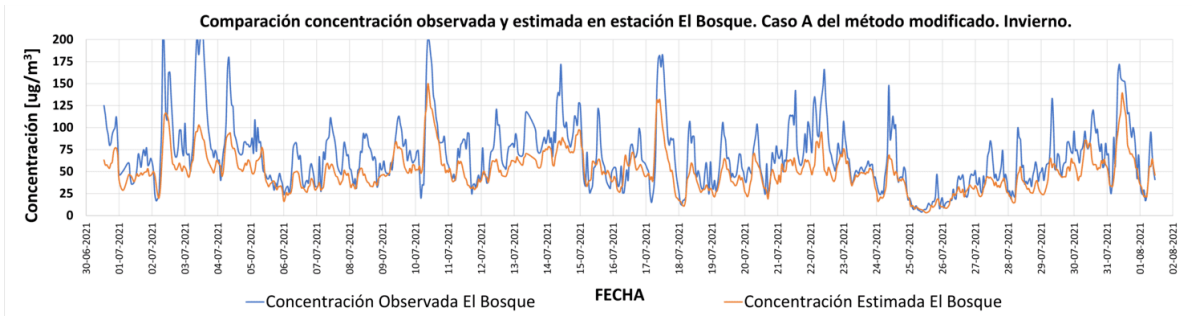


Figura G.17: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IA. Temporada Invierno.

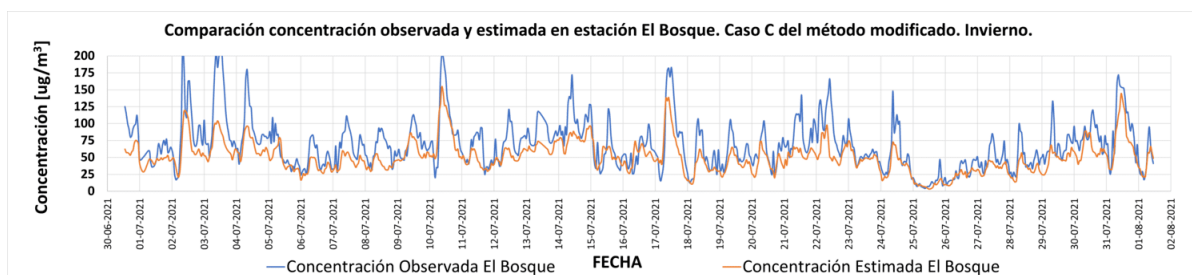


Figura G.18: Comparación Concentraciones reales e interpoladas en ubicación estación El Bosque para un mes de datos. IC. Temporada Invierno.