



UNIVERSIDAD NACIONAL DE COLOMBIA

# Una Aplicación en la Evaluación del Impacto del Confinamiento Estricto por la Covid-19 en la Calidad del Aire en la Ciudad de Medellín Basado en Modelos Bayesianos Dinámicos Multivariados y Espaciales

**Carlos Andrés Pérez Aguirre**

Universidad Nacional de Colombia  
Facultad de Ciencias, Escuela de Estadística  
Medellín, Colombia  
2022

# Una Aplicación en la Evaluación del Impacto del Confinamiento Estricto por la Covid-19 en la Calidad del Aire en la Ciudad de Medellín Basado en Modelos Bayesianos Dinámicos Multivariados y Espaciales

**Carlos Andrés Pérez Aguirre**

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:  
**Maestro de las ciencias estadística**

Directores:

Ph.D., Johnatan Cardona Jiménez

Docente, Facultad de Ingenierías, Institución Universitaria Pascual Bravo.

Ph.D., Isabel Cristina Ramírez Guevara

Profesora Asociada, Universidad Nacional de Colombia Sede Medellín.

Línea de Investigación:

Estadística Bayesiana.

Grupo de Investigación:

Grupo de Investigación en Estadística UNAL, sede Medellín

Universidad Nacional de Colombia  
Facultad de Ciencias, Escuela de Estadística  
Medellín, Colombia

2022

# Agradecimientos

Agradecemos al Sistema de Alerta Temprana de Medellín y el Valle de Aburrá (SIATA) por proveer la información utilizada en este estudio, y a los desarrolladores de R y STAN por el trabajo dedicado a construir software libremente disponible para el uso académico.



# Resumen

## **Una Aplicación en la Evaluación del Impacto del Confinamiento Estricto por la Covid-19 en la Calidad del Aire en la Ciudad de Medellín Basado en Modelos Bayesianos Dinámicos Multivariados y Espaciales.**

En la ciudad de Medellín y demás municipios del Valle de Aburrá los episodios de altos niveles de contaminación en el aire han sido recurrentes en los últimos años. Múltiples estudios realizados por universidades de la región y entes gubernamentales <sup>1</sup> han presentado evidencia del impacto negativo en la salud de los habitantes de la región asociado a varios tipos de contaminantes presentes en el aire. Las medidas de aislamiento provocadas por la pandemia SARS-CoV-2 (Covid-19) durante inicios del 2020 provocaron una reducción drástica del flujo vehicular. Así, en este proyecto proponemos e implementamos mediante un paquete en R un método estadístico basado en modelos dinámicos espacio-temporales para evaluar el impacto de la reducción del flujo vehicular en la presencia de contaminantes (PM 10, PM 2.5, NO, NO<sub>2</sub>, NO<sub>x</sub>). En esta tesis encontramos que se presenta una reducción en la concentración en el aire de los contaminantes previamente mencionados.

**Palabras clave:** Series de tiempo, Análisis de intervención, Inferencia Bayesiana, Modelos dinámicos espacio-temporales, Calidad del aire..

---

<sup>1</sup><http://www.cgm.gov.co/cgm/Paginaweb/IP/Informe20Ambiental202019/Informe20Ambiental20Vigencia202019.pdf>

# Abstract

## **An Application in the Evaluation of the Impact of Strict Confinement by Covid-19 on Air Quality in the City of Medellín Based on Multivariate and Spatial Dynamic Bayesian Models.**

In the city of Medellín and other municipalities in the Aburrá Valley, episodes of high levels of air pollution have been recurrent in recent years. Multiple studies carried out by universities in the region and government entities <sup>1</sup>, have presented evidence of the negative impact on the health of the inhabitants of the region associated with various types of pollutants present in the air. The isolation measures caused by the SARS-CoV-2 (Covid-19) pandemic during early 2020 caused a drastic reduction in vehicular flow. Thus, in this project we propose and implement in R package a statistical method based on space-time dynamic models to assess the impact of reducing vehicular flow in the presence of pollutants (PM 10, PM 2.5, NO<sub>2</sub>, NO<sub>x</sub>). In this thesis we find that there is a reduction in the concentration in the air of the previously mentioned pollutants.

**Keywords:** Time series, Intervention analysis, Bayesian inference, Space-time dynamic models, Air quality.

# Contenido

<b>Agradecimientos</b>	<b>iii</b>
<b>Resumen</b>	<b>v</b>
<b>Contenido</b>	<b>vii</b>
<b>Lista de figuras</b>	<b>ix</b>
<b>Lista de tablas</b>	<b>xiii</b>
<b>1 Introducción</b>	<b>2</b>
1.0.1 Pregunta de investigación . . . . .	3
1.0.2 Antecedentes: calidad del aire en el valle de Aburrá . . . . .	4
1.1 Estado del arte . . . . .	6
1.1.1 Evaluación de impacto . . . . .	8
1.1.2 Evaluación de impacto en series de tiempo . . . . .	8
1.1.3 Diferencias en diferencias . . . . .	9
1.1.4 Evaluación punto a punto . . . . .	10
1.1.5 Contrafactual Artificial - ArCo . . . . .	11
1.2 Impacto agregado . . . . .	12
<b>2 Modelo Lineal Dinámico Matriz-Variado</b>	<b>13</b>
2.0.1 Modelo conjugado . . . . .	14
2.0.2 Aproximación posterior y predicción mediante MCMC . . . . .	18
2.1 Modelo Lineal Dinámico Matriz-Variado espacial . . . . .	19
2.2 Modelo Lineal Dinámico Matriz-Variado espacial modificado . . . . .	20
<b>3 Simulación</b>	<b>22</b>
3.1 Algoritmos de simulación para el modelo (2-1) . . . . .	22
3.2 Impacto en la misma dirección de la tendencia . . . . .	25
3.3 Caso en el cual no se presenta impacto . . . . .	26
3.4 Impacto en dirección contraria a la tendencia . . . . .	27
3.5 Análisis de simulación del impacto agregado (1.2) empleando el modelo (2-1)	27
3.6 Resultado para el modelo espacial definido en (2-19). . . . .	32

---

<b>4</b>	<b>Aplicación: Evaluación del impacto relacionado con la hipótesis de investigación.</b>	<b>35</b>
4.1	Resultados para PM 10 . . . . .	35
4.2	Resultados para PM 2.5 . . . . .	41
4.3	Resultados para NO <sub>x</sub> . . . . .	44
4.4	Discusión . . . . .	46
<b>5</b>	<b>Paquete DynamicIpackt</b>	<b>48</b>
<b>6</b>	<b>Conclusiones y recomendaciones</b>	<b>52</b>
6.1	Conclusiones . . . . .	52
6.2	Recomendaciones . . . . .	52
6.3	Limitaciones y trabajos futuros . . . . .	53
<b>7</b>	<b>Anexos</b>	<b>55</b>
7.1	Resultados adicionales . . . . .	55
7.1.1	PM 10 . . . . .	55
7.1.2	NO <sub>x</sub> . . . . .	59
7.1.3	NO . . . . .	61
7.1.4	NO <sub>2</sub> . . . . .	62
7.1.5	PM 2.5 . . . . .	63
7.2	Resultados adicionales para el modelo espacial . . . . .	70
7.2.1	PM 2.5 . . . . .	70
7.2.2	NO . . . . .	72
7.2.3	NO <sub>2</sub> . . . . .	73
7.2.4	NO <sub>x</sub> . . . . .	73
7.2.5	PM 10 . . . . .	74
7.3	Estaciones de monitoreo de la calidad de aire . . . . .	75
7.3.1	PM 10 . . . . .	75
7.3.2	PM 2.5 . . . . .	76
7.3.3	NO . . . . .	78
7.3.4	NO <sub>2</sub> . . . . .	79
7.3.5	NO <sub>x</sub> . . . . .	80
	<b>Bibliografía</b>	<b>81</b>



# Lista de Figuras

<b>1-1</b>	Factores que incrementan la contaminación en el Valle de Aburrá. Fuente: Área Metropolitana del Valle de Aburrá . . . . .	2
<b>1-2</b>	Estaciones de monitoreo de la calidad del aire. . . . .	4
<b>1-3</b>	Series de tiempo para material particulado PM 2.5 y PM 10 medida en una estación de monitoreo (Centro) de la ciudad de Medellín. Fuente: SIATA. . .	5
<b>1-4</b>	Ejemplo del impacto en series de tiempo utilizando datos simulados. . . . .	9
<b>3-1</b>	Resultados de simulación cuando se presenta un impacto en una estación simulada. En el panel (1) se presenta la serie real simulada y su contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5) respectivamente. . . . .	25
<b>3-2</b>	Ejemplo resultado de simulación cuando no se presenta un impacto en una estación simulada. En el panel (1) se presenta la serie real simulada y su contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5), presentados en capítulo 1.	26
<b>3-3</b>	Ejemplo resultado de simulación cuando se presenta un impacto con dirección contraria a la tendencia en una estación simulada. En el panel (1) se puede ver que después del $T_0 = 73$ la serie real sufre un cambio de nivel en dirección contraria a la tendencia, lo cual hace que se separe del contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5), presentados en capítulo 1. . . . .	27
<b>3-4</b>	Resultado de la simulación en función del número de elementos, para los casos en los cuales existe un impacto. . . . .	29
<b>3-5</b>	Resultado de la simulación en función del número de elementos, para los casos en los cuales no existe un impacto. . . . .	29
<b>3-6</b>	Resultado de la simulación en función del porcentaje de datos usados antes de la intervención, para los casos en los cuales existe un impacto. . . . .	30
<b>3-7</b>	Resultado de la simulación en función del porcentaje de datos usados antes de la intervención, para los casos en los cuales no existe un impacto. . . . .	31
<b>3-8</b>	Resultados de simulación cuando se presenta un impacto en la dirección de la tendencia para el modelo espacial definido en (2-19). . . . .	32
<b>3-9</b>	Resultados de simulación cuando no se presenta un impacto para el modelo espacial definido en (2-19). . . . .	33

<b>3-10</b>	Resultados de simulación cuando se presenta un impacto en la dirección contraria de la tendencia para el modelo espacial definido en (2-19). . . . .	33
<b>4-1</b>	Grupos de estaciones para el contaminante PM 10 obtenidos utilizando al algoritmo de clusterización presentado por Papenberg y Klau (2020). . . . .	36
<b>4-2</b>	Resultados agregados PM 10 para el grupo de estaciones 1, utilizando la ecuación (1-6). . . . .	37
<b>4-3</b>	Resultados PM 10 para la estación Tráfico Centro. . . . .	37
<b>4-4</b>	Resultados agregados PM 10 para el grupo de estaciones 2, utilizando la ecuación (1-6). . . . .	38
<b>4-5</b>	Resultados PM 10 para la estación de Policía Los Gómez. . . . .	39
<b>4-6</b>	Resultados agregados PM 10 para el grupo de estaciones 1, utilizando la ecuación (1-6) para el modelo (2-19). . . . .	40
<b>4-7</b>	Resultados agregados PM 10 para el grupo de estaciones 2, utilizando la ecuación (1-6) para el modelo (2-19). . . . .	40
<b>4-8</b>	Grupos de estaciones para el contaminante PM 2.5 divididas utilizando al algoritmo de clúster desarrollado por Papenberg y Klau (2020). . . . .	41
<b>4-9</b>	Resultados agregados PM 2.5 grupo de estaciones 1, utilizando la ecuación (1-6). . . . .	42
<b>4-10</b>	Resultados PM 2.5 para la estación La Estrella - Hospital. . . . .	42
<b>4-11</b>	Resultados agregados PM 2.5 para el grupo de estaciones 1, utilizando la ecuación (1-6) para el modelo (2-19) . . . . .	43
<b>4-12</b>	Estaciones para los contaminantes NO, NO <sub>2</sub> y NO <sub>x</sub> . . . . .	44
<b>4-13</b>	Resultados agregados NO <sub>x</sub> , utilizando la ecuación (1-6). . . . .	45
<b>4-14</b>	Resultados agregados NO <sub>x</sub> , utilizando la ecuación (1-6) para el modelo (2-19). . . . .	46
<b>5-1</b>	Instalación del paquete <b>dynamicimpact</b> . . . . .	48
<b>5-2</b>	Ejemplo de uso del paquete DynamicImpact 1. . . . .	49
<b>5-3</b>	Ejemplo de uso del paquete DynamicImpact 2. . . . .	50
<b>5-4</b>	Ejemplo de uso del paquete DynamicImpact 3. . . . .	50
<b>5-5</b>	Ejemplo de uso del paquete DynamicImpact 2. . . . .	51
<b>7-1</b>	Resultados PM 10 para la estación Universidad San Buenaventura. . . . .	55
<b>7-2</b>	Resultados PM 10 para la estación Institucion Universitaria ITM Robledo. . . . .	55
<b>7-3</b>	Resultados PM 10 para la estación Institución Educativa Colombia. . . . .	56
<b>7-4</b>	Resultados PM 10 para la estación Exito San Antonio. . . . .	56
<b>7-5</b>	Resultados PM 10 para la estación Politecnico Colombiano Jaime Isaza Cadavid. . . . .	58
<b>7-6</b>	Resultados PM 10 para la estación I.E. Conejo Municipal de Itagüí. . . . .	58
<b>7-7</b>	Resultados PM 10 para la estación estación Tráfico Sur. . . . .	58
<b>7-8</b>	Resultados NO <sub>x</sub> para la estación Tráfico Sur. . . . .	59
<b>7-9</b>	Resultados NO <sub>x</sub> para la estación ITM Robledo. . . . .	59
<b>7-10</b>	Resultados NO <sub>x</sub> Casa de Justicia Itagüí. . . . .	60

---

<b>7-11</b> Resultados $\text{NO}_x$ Politecnico Colombiano Jaime Isaza Cadavid. . . . .	60
<b>7-12</b> Resultados agregados NO, utlizando la ecuación (1-6) . . . . .	61
<b>7-13</b> Resultados agregados $\text{NO}_2$ , utlizando la ecuación (1-6) . . . . .	62
<b>7-14</b> Resultados agregados PM 2.5 grupo de estaciones 2, utlizando la ecuación (1-6)	63
<b>7-15</b> Resultados agregados PM 2.5 grupo de estaciones 3, utlizando la ecuación (1-6)	64
<b>7-16</b> Resultados agregados PM 2.5 grupo de estaciones 4, utlizando la ecuación (1-6)	64
<b>7-17</b> Resultados PM 2.5 para la estación Corporacion Universitaria Lasallista. . .	65
<b>7-18</b> Resultados PM 2.5 para la estación Corporacion Universitaria Lasallista. . .	65
<b>7-19</b> Resultados PM 2.5 para la estación Tráfico Sur. . . . .	65
<b>7-20</b> Resultados PM 2.5 para la estación I.E. Rafael J. Mejía. . . . .	66
<b>7-21</b> Resultados PM 2.5 para la estación E.S.E. Santa Gertrudis. . . . .	66
<b>7-22</b> Resultados PM 2.5 para la estación Casa de Justicia Itagüí. . . . .	67
<b>7-23</b> Resultados PM 2.5 para la estación Concejo Municipal de Itagüí. . . . .	67
<b>7-24</b> Resultados PM 2.5 para la estación I.E. Pedro Octavio Amado. . . . .	67
<b>7-25</b> Resultados PM 2.5 para la estación I.E Pedro Justo Berrio. . . . .	68
<b>7-26</b> Estaciones PM 10 parte 1 . . . . .	75
<b>7-27</b> Estaciones PM 10 parte 2 . . . . .	76
<b>7-28</b> Estaciones PM 2.5 parte 1 . . . . .	76
<b>7-29</b> Estaciones PM 2.5 parte 2 . . . . .	77
<b>7-30</b> Estaciones PM 2.5 parte 3 . . . . .	77
<b>7-31</b> Estaciones PM 2.5 parte 4 . . . . .	78
<b>7-32</b> Estaciones NO . . . . .	78
<b>7-33</b> Estaciones $\text{NO}_2$ . . . . .	79
<b>7-34</b> Estaciones $\text{NO}_x$ . . . . .	80



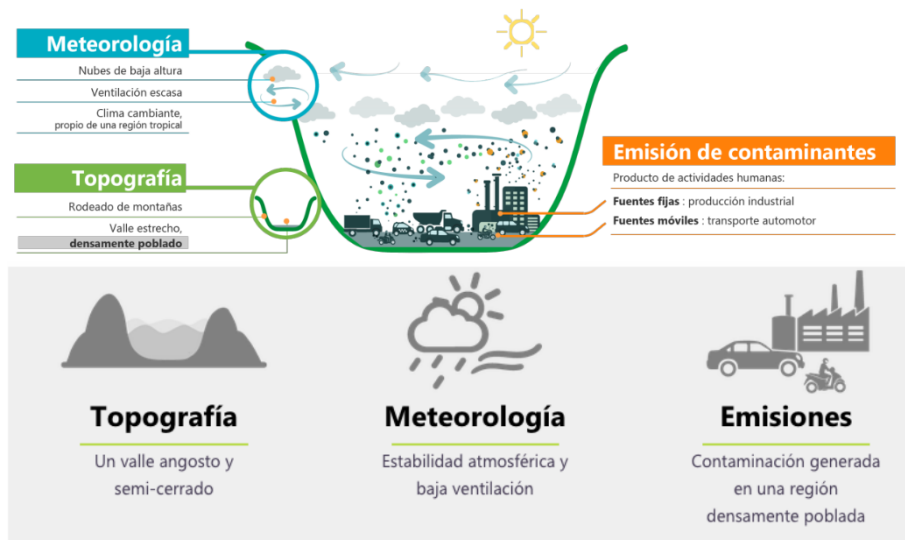
# Lista de Tablas

1-1	Estaciones y contaminantes utilizados en el análisis. . . . .	6
3-1	Resultados en función del número de elementos de la matriz $Y_t$ . . . . .	30
3-2	Resultados en función del porcentaje de datos usados antes de la intervención	31
4-1	Resultados PM 10 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 1”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15. . . . .	38
4-2	Resultados PM 10 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 2”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15. . . . .	39
4-3	Resultados PM 2.5 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 1”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15. . . . .	43
4-4	Resultados NO <sub>x</sub> agregados e individuales con tiempo de corte 2020-03-31. Los resultados agregados aparecen con nombre de estación “Global”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-16 y 2020-03-31. . . . .	45
5-1	Funciones disponible para el usuario en el paquete <b>dynamicimpact</b> . . . . .	48
7-1	Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global” grupo 1.	57
7-2	Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 2.	57
7-3	Resultados NO <sub>x</sub> agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”. . . . .	61
7-4	Resultados NO agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”. . . . .	62
7-5	Resultados NO <sub>2</sub> agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”. . . . .	63

<b>7-6</b>	Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 1. . .	68
<b>7-7</b>	Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 2. . .	69
<b>7-8</b>	Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 3. . .	69
<b>7-9</b>	Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 4. . .	70
<b>7-10</b>	Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 1. . . . .	70
<b>7-11</b>	Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 2. . . . .	71
<b>7-12</b>	Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 3. . . . .	71
<b>7-13</b>	Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 4. . . . .	72
<b>7-14</b>	Resultados NO agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”. . . . .	72
<b>7-15</b>	Resultados NO <sub>2</sub> agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”. . . . .	73
<b>7-16</b>	Resultados NO <sub>x</sub> agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”. . . . .	73
<b>7-17</b>	Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, parte 1. . . . .	74
<b>7-18</b>	Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, parte 2. . . . .	74

# 1 Introducción

El deterioro de la calidad de aire es un problema común en las grandes y medianas ciudades. Este fenómeno puede ser explicado por diferentes factores relacionados con las actividades humanas (industria, movilidad, agricultura, etc), factores geográficos y procesos naturales. Por ejemplo, en la ciudad de Medellín, a través del sistema de estaciones de monitoreo del Sistema de Alerta Temprana (SIATA, 2019)<sup>1</sup>, en los últimos años se han tornado comunes los episodios de altos niveles de degradación en la calidad del aire. Este problema se ha agudizado al punto de convertirse en un problema de salud pública. Los altos niveles de material particulado de tamaños PM 10, PM 2.5 y menores comprometen la salud de los habitantes de la ciudad de Medellín, con mayor incidencia en menores de 5 años y mayores de 65 años (Área Metropolitana Valle de Aburrá, 2019; Contraloría General de Medellín, 2019).



**Figura 1-1:** Factores que incrementan la contaminación en el Valle de Aburrá. Fuente: Área Metropolitana del Valle de Aburrá

En la Figura 1, se presentan algunos de los factores más relevantes relacionados con los episodios de niveles altos de contaminación del aire en el Valle de Aburrá. Se destaca que, existen factores no controlables como las condiciones topográficas y fenómenos meteorológicos, y factores controlables como las fuentes fijas y móviles. Por factores no controlables se entiende

<sup>1</sup>[https://siata.gov.co/siata\\_nuevo/](https://siata.gov.co/siata_nuevo/)

como aquellas variables que no pueden ser controladas por la voluntad humana. Por ejemplo, no es posible definir políticas para controlar los incendios forestales que ocurren fuera de la jurisdicción del Área Metropolitana (Hernández, 2019), los vientos que traen material particulado desde el desierto del Sahara (Méndez et al., 2018), la lluvia o el viento. Pero sí es posible definir políticas públicas para mitigar las emisiones generadas por las actividades industriales y/o el parque automotor. Este tipo de decisiones o políticas encaminadas a controlar ciertas fuentes contaminantes son complejas e impopulares por el impacto que ellas pueden tener en las actividades económicas y los hábitos de movilidad adoptados por una fracción importante de la población. Por eso es, de gran importancia realizar una evaluación científica del impacto relacionado con las fuentes controlables en la calidad del aire para crear conciencia sobre el efecto que este problema tiene sobre nuestra calidad de vida y ganar el respaldo de la población en la construcción de políticas encaminadas a crear hábitos más amigables con el medio ambiente y mejoramiento de los sistemas de movilidad. Vale resaltar que esta última no es una tarea trivial dada la complejidad del problema por los múltiples factores involucrados en el fenómeno de la contaminación del aire. Sin embargo, durante los meses de marzo y abril del 2020 se dieron unas condiciones atípicas que permitirían realizar una evaluación aislada del efecto del parque automotor y de algunas fuentes fijas relacionadas con actividades industriales en la calidad del aire. Con la llegada de la Covid-19 al país, el desconocimiento sobre sus mecanismos de transmisión y los antecedentes de los episodios vividos en países como España e Italia, el gobierno nacional decretó una medida de confinamiento estricto que llevó a una pausa total de gran parte de las actividades económicas y sociales del país. Como consecuencia de esta medida, el parque automotor y algunas actividades industriales (las catalogadas como no esenciales) cesaron, lo que derivó en una disminución o incluso en algunos casos, eliminación completa de las emisiones de durante poco más de un mes. Así, en este proyecto proponemos una metodología estadística para evaluar el impacto de fuentes fijas y móviles en los niveles de contaminación que se presenta en uno de los periodos críticos del año.

### **1.0.1. Pregunta de investigación**

¿Durante los periodos de cuarentena estricta (marzo - abril del 2020) (Ministerio de Salud y Protección Social, 2020), (Alcaldía de Medellín, 2020) en los cuales se restringió la movilidad de vehículos y las actividades industriales, se produjo una reducción en la concentración de contaminantes PM 10, PM 2.5, NO, NO<sub>x</sub> y NO<sub>2</sub> en el Valle de Aburrá?

La Figura 1-2 presenta la distribución espacial de las estaciones en el Valle de Aburrá, podemos observar que la mayoría de estas están concentradas en el área metropolitana.



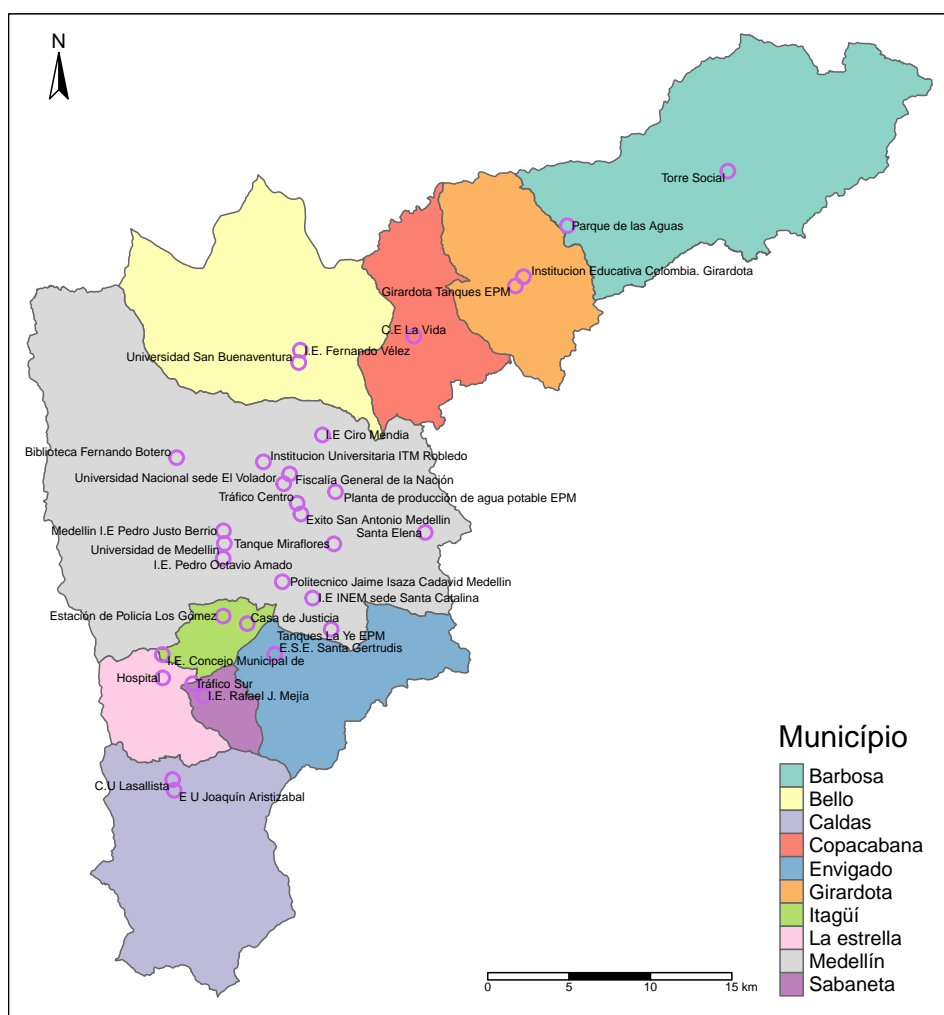


Figura 1-2: Estaciones de monitoreo de la calidad del aire.

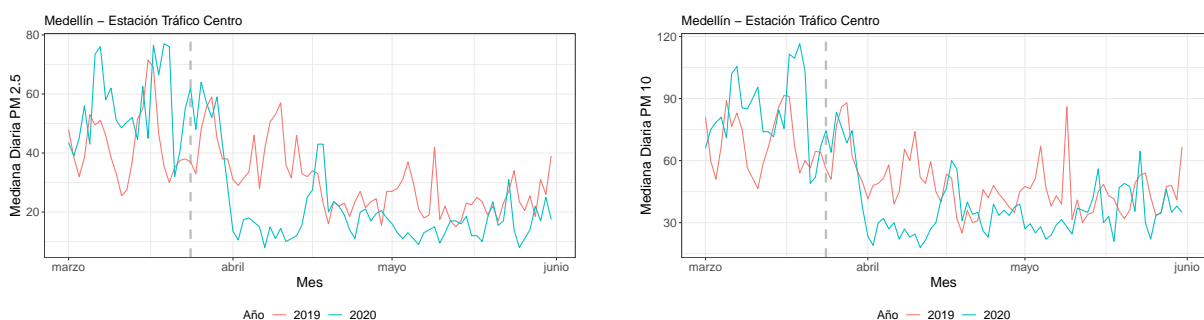
### 1.0.2. Antecedentes: calidad del aire en el valle de Aburrá

La ubicación geográfica de la ciudad Medellín y en el área metropolitana del Valle de Aburrá generan condiciones de poca ventilación (Área Metropolitana del Valle de Aburrá, 2018), las cuales facilitan la concentración de contaminantes que causan efectos adversos a la salud (Kampa y Castanas, 2008). Por ejemplo, Bedoya y Martínez (2009) reportan concentraciones de material particulado respirable PM 10 de  $70\mu\text{g}/\text{m}^3$  más de tres veces el límite de  $20\mu\text{g}/\text{m}^3$  recomendado por la organización mundial de la salud. Adicionalmente, a los límites definidos por la organización mundial de la salud existen distintos indicadores de calidad del aire construidos con base a uno o múltiples contaminantes, algunos de estos como el AQHI (Air Quality Health Index) están definidos teniendo en cuenta perspectivas desde el punto de vista de salud pública (Gorai y Goyal, 2015).

El efecto del parque automotor sobre la presencia de contaminantes en la calidad del aire de

la ciudad de Medellín ha sido discutido múltiples veces por autores como Posada et al. (2017) y Cardenas (2017) entre otros. Las cuarentenas aplicadas al inicio de la emergencia sanitaria en Colombia restringieron la movilidad de personas y vehículos, lo cual provee un conjunto de condiciones tan cercano como es posible al escenario de remover el parque automotor. Lo anterior permite estimar, mediante modelos estadísticos, el efecto de remover el parque automotor utilizando como base los periodos en los cuales se ejercieron cuarentenas estrictas. En este trabajo queremos estimar estadísticamente el impacto (si este existe) definido como la reducción en la presencia de contaminantes en el aire, causada por la remoción del parque automotor y la suspensión de las actividades industriales, durante el periodo de cuarentena estricta en Medellín y el Valle de Aburrá. Para nuestro caso, las restricciones de datos por el Covid-19 nos permiten construir un experimento incompleto en el cual las cuarentenas actúan como un tratamiento y el grupo de control no está presente (Bernal et al., 2016). Para remplazar el grupo de control se utiliza un modelo contrafactual (Winship y Morgan, 1999) el cual puede ser construido mediante la información previa al tratamiento.

Para la estimación de la concentración de contaminantes se construye el modelo contrafactual, el cual representa el escenario donde no se presenta el evento. En este caso, el modelo contrafactual representa un escenario hipotético en el cual nunca se imponen restricciones, lo que implica que no hay remoción del parque automotor y no hay reducción de la capacidad industrial. Este modelo es ajustado utilizando únicamente información previa a la cuarentena. Nuestra estimación del impacto proviene de utilizar la serie de tiempo obtenida mediante la distribución predictiva (West y Harrison, 1997a) del modelo contrafactual y compararla con los valores observados durante el periodo de cuarentena. Brodersen et al. (2015) proponen una metodología basada en modelos espacio-estado para la estimación de impacto aplicado en campañas de publicidad, el cual será adaptado a este caso de estudio. Es posible utilizar otras metodologías basadas en algoritmos de descubrimiento de relaciones causales basados en rezagos, ver Maathuis y Nandy (2015) como base para la construcción de un modelo contrafactual, aunque estos análisis no son utilizados en este trabajo. Una vez se posee el modelo contrafactual es posible hacer la estimación del impacto utilizando la diferencia de los valores observados y los predichos como se presenta en la sección 1.1.1.



**Figura 1-3:** Series de tiempo para material particulado PM 2.5 y PM 10 medida en una estación de monitoreo (Centro) de la ciudad de Medellín. Fuente: SIATA.

En la Figura 1-3 se puede apreciar una disminución del material particulado PM 2.5 y PM 10 durante el mes de abril, con respecto a ese mismo periodo en el año anterior, en la estación tráfico centro de Medellín. La disminución observada es consecuencia de las restricciones al parque automotor y la actividad industrial generadas por el COVID-19, este escenario ideal para evaluar el efecto de interés sería impensable en condiciones normales (sin pandemia). En la tabla 1.0.2 se presenta la lista de estaciones utilizadas en este trabajo y los contaminantes monitoreados en cada una de estas.

**Tabla 1-1:** Estaciones y contaminantes utilizados en el análisis.

Estación	PM 2.5	Ozono	PM 10	NO	NO <sub>2</sub>	NO <sub>x</sub>
Barbosa Torre Social	X					
Bello I.E. Fernando Vélez	X					
Caldas Corporación Universitaria Lasallista	X	X				
Caldas E U Joaquín Aristizabal	X					
Copacabana Ciudadela Educativa La Vida	X					
Envigado E.S.E. Santa Gertrudis	X					
Estación Tráfico Centro	X		X			
Estación Tráfico Sur	X		X	X	X	X
Exito San Antonio			X			
Institución Educativa Colombia. Girardota			X			
Institución Universitaria ITM Robledo			X	X	X	X
Itagüí Casa de Justicia Itagüí	X			X	X	X
Itagüí Estación de Policía Los Gómez			X			
Itagüí I.E. Concejo Municipal de Itagüí	X	X	X			
La Estrella Hospital	X					
Altavista I.E. Pedro Octavio Amado	X					
Aranjuez I.E. Ciro Mendía	X					
Belén I.E. Pedro Justo Berrio	X					
El Poblado I.E. INEM sede Santa Catalina	X					
El Poblado Tanques La Ye EPM	X	X				
San Cristobal Parque Biblioteca Fernando Botero	X					
Villahermosa Planta de producción de agua potable EPM	X					
Parque de las Aguas		X				
Politécnico Colombiano Jaime Isaza Cadavid			X	X	X	X
Sabaneta I.E. Rafael J. Mejía	X					
Tanque Miraflores		X				
Universidad San Buenaventura		X	X			

## 1.1. Estado del arte

Durante el 2020 se realizaron múltiples estudios para medir los impactos de las cuarentenas estrictas sobre la calidad del aire. En el estudio realizado en Estados Unidos y China por Shakoor et al. (2020) se analizan los contaminantes CO, NO<sub>2</sub>, PM 2.5, PM 10, SO<sub>2</sub> en 10

estaciones (5 para Estados Unidos y 5 para China), para cada estación se calcula la reducción promedio en la concentración de contaminantes entre el periodo anterior a las restricciones y el periodo con restricciones. En el estudio realizado por Ropkins y Tate (2020) en el Reino Unido se estudian los contaminantes  $\text{NO}_2$ ,  $\text{NO}_x$ ,  $\text{NO}$ ,  $\text{PM}_{2.5}$ ,  $\text{PM}_{10}$  y  $\text{O}_3$  en 153 estaciones. Para determinar si existe una reducción en la contaminación, se ajustan modelos en diferentes periodos de tiempos y se analizan los cambios en la tendencia de estos. El estudio realizado por Briz-Redón et al. (2021) en España utiliza 11 estaciones cada una de ellas en una ciudad diferente, los contaminantes analizados son  $\text{NO}_2$ ,  $\text{CO}$ ,  $\text{O}_3$  y  $\text{SO}_2$ . Para determinar si existe una reducción en la contaminación se ajusta un modelo de regresión, en el cual una de sus variables explicativas representa el periodo de confinamiento. En Colombia Mendez-Espinosa et al. (2020) realizan un estudio en 4 estaciones (2 para Medellín y 2 Bogotá) para los contaminantes  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$  y  $\text{NO}_2$ . En este análisis se compara la concentración promedio antes y después de las restricciones por la cuarentena, en este estudio no se tiene un modelo que explique la relación espacial o temporal entre las estaciones.

En este trabajo utilizamos una metodología basada en modelos dinámicos bayesianos, los cuales permiten la evolución a través del tiempo de los parámetros que los conforman (Quintana, 1985), para evaluar el impacto de las restricciones generadas por la cuarentena sobre la concentración de contaminantes en el aire. El enfoque bayesiano utilizado en este trabajo permite de manera natural realizar estimaciones sobre el rango en el cual se encuentra el cambio en la concentración de contaminantes. Los modelos dinámicos multivariados permiten realizar un análisis conjunto de un grupo de contaminantes o estaciones, por lo tanto, es posible incluir en el modelo las covarianzas entre grupos de estaciones, lo cual es ignorado en los análisis anteriormente citados, en los cuales únicamente se estudia la diferencia en la concentración entre dos periodos de tiempo.

Box y Tiao (1975) comentan que un problema frecuente en el análisis aplicado de series tiempo es determinar el efecto que tiene una intervención sobre las observaciones futuras. La estimación del impacto de la intervención es equivalente a estimar el efecto de un tratamiento en un entorno en el que no se posee un grupo de control (Linden y Adams, 2011; Bernal et al., 2016). Para remplazar el grupo de control se construye la serie contrafactual (Winship y Morgan, 1999), la cual es obtenida al ajustar un modelo para algunos periodos antes de la intervención y predecir con este para tiempos después de la intervención. De esta forma, la serie contrafactual puede ser interpretada como un control que no ha recibido el tratamiento bajo estudio. Así, la evaluación del impacto es obtenida al comparar los valores predichos con los observados para el periodo de tiempo posterior (y durante) a la intervención. Con el fin de obtener una estimación robusta del efecto producido por la intervención, es recomendable incluir variables explicativas que no se vean afectadas por el evento de interés en el modelo (Morgan y Winship, 2014). Siguiendo esta metodología para evaluación de impacto con series temporales, definimos dos tipos de modelos para datos con dependencia espacial y temporal definidos en las secciones 2.1 y 2 respectivamente. Con estos modelos y la identificación de las

variables de control adecuadas se obtendrán las series contrafactuales. Estas series servirán para realizar una estimación del impacto de la cuarentena estricta (en los meses de marzo y abril del 2020) en la disminución observada de diferentes contaminantes presentes en el aire de la ciudad de Medellín y algunos municipios de Área Metropolitana.

### 1.1.1. Evaluación de impacto

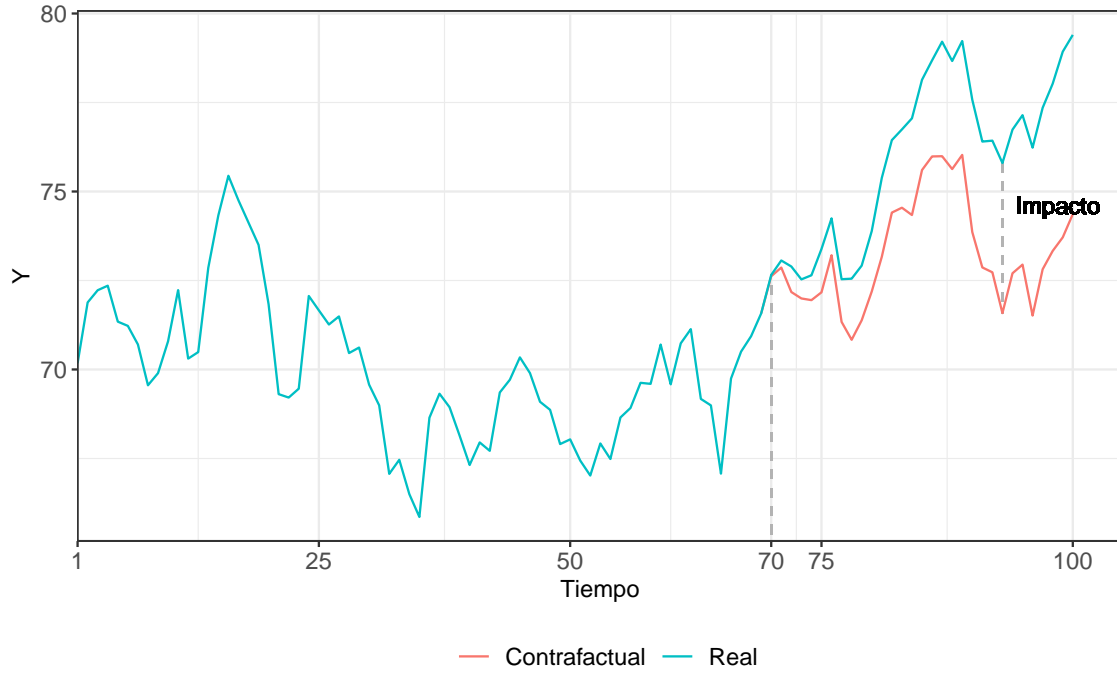
El objetivo de evaluación de impacto es caracterizar y valorar los cambios generados por un evento o intervención. En este trabajo nos centraremos en la evaluación de impacto cuantitativa, la cual permite medir de manera objetiva los sujetos de estudio. En la evaluación de impacto se requiere un contrafactual el cual representa lo que hubiese sucedido cuando no se presenta el evento o intervención de interés. El contrafactual puede ser obtenido mediante métodos experimentales y no experimentales (Nguyen y Bloom, 2006).

**Métodos experimentales:** En esta metodología los participantes del estudio son asignados de manera aleatoria a un grupo de tratamiento (en el cual se presenta el evento de interés) o a un grupo de control (en el cual no se presenta el evento de interés), este último es el contrafactual

**Métodos no experimentales:** En esta metodología no posee un diseño experimental, el contrafactual es obtenido de manera estadística utilizando modelos predictivos construidos con información de los sujetos de estudio antes del evento de interés, o utilizando criterios de selección para encontrar sujetos de estudio que no son afectados por el evento de interés (Lance et al., 2014).

### 1.1.2. Evaluación de impacto en series de tiempo

En múltiples casos de interés, el sujeto de estudio es un sistema del cual se tienen una o más mediciones en el tiempo. Cuando sucede una intervención o evento, el objetivo es cuantificar el impacto que este tiene sobre el sistema.



**Figura 1-4:** Ejemplo del impacto en series de tiempo utilizando datos simulados.

En La Figura 1-4 se presenta un ejemplo del efecto de una intervención que comienza en el tiempo 70 sobre la serie en color azul claro. En color rojo se presenta el contrafactual para los tiempos después de la intervención. El impacto para un tiempo determinado se define como la diferencia entre estos. El modelo contrafactual es la predicción construida con información previa a la intervención para tiempos posteriores a la intervención. Esta predicción representa lo que hubiese sucedido bajo un escenario sin intervención.

### 1.1.3. Diferencias en diferencias

Una manera de identificar si existe un impacto es comparar el sistema de estudio (grupo de tratamiento) con otro sistema que no fue afectado por el evento o intervención (grupo de intervención) (Belasen y Polachek, 2008). El estimador de diferencias en diferencias se define como la diferencia de la resta entre el promedio del valor observador en el sistema afectado ( $\bar{Y}^T$ ) y promedio del valor observador en el sistema no afectado ( $\bar{Y}^C$ ) (Abadie, 2010). Tanto  $\bar{Y}^T$  como  $\bar{Y}^C$  deben tener la misma longitud,

$$\hat{\beta} = (\bar{Y}_{T_0+1, \dots, T}^T - \bar{Y}_{T_0+1, \dots, T}^C) - (\bar{Y}_{1, \dots, T_0}^T - \bar{Y}_{1, \dots, T_0}^C), \quad (1-1)$$

donde:

$1, \dots, T_0$  : es el periodo antes de la intervención o evento.

$T_0 + 1, \dots, T$  : es el periodo después de la intervención o evento.

La distribución de  $\hat{\beta}$  puede ser obtenida a partir de las distribuciones de  $\bar{Y}^C$  y  $\bar{Y}^T$  y con ella es posible construir un intervalo de confianza para un nivel de significancia  $\alpha$  definido por los investigadores. Este método ofrece una solución sencilla al problema de la evaluación de impacto, pero asume que la tendencia a través del tiempo del sistema afectado y la tendencia del sistema no afectado son paralelas (Paskov, 2019), lo cual hace que este método no sea aplicable en múltiples escenarios, incluyendo este trabajo, debido a que la variable de entrada (valores observados de los contaminantes en 2019) y la variable a predecir (valores observados de los contaminantes en 2020) no son paralelas.

#### 1.1.4. Evaluación punto a punto

Para realizar la evaluación de impacto de interés se deben identificar variables predictoras que no se vean afectadas por el tratamiento bajo estudio (Morgan y Winship, 2014). Como ya se ha mencionado anteriormente, el tratamiento de interés en este estudio es el aislamiento preventivo decretado entre los meses de marzo y abril del 2020. Identificar dichas variables predictoras será una de las tareas más complejas y retadoras a realizar en este trabajo. Solo para citar posibles candidatas a variables predictoras tenemos factores atmosféricos (la temperatura, la cantidad de radiación solar, el viento, la humedad, etc.) e información histórica de los contaminantes bajo estudio observados durante el mismo periodo, pero en años anteriores. Por ahora supongamos que ya contamos con un buen predictor para la construcción de la serie temporal contrafactual. Sea  $\tilde{\mathbf{Y}}_{t,i}$  un vector de  $q$  estaciones o una matriz de  $r$  contaminantes por  $q$  estaciones, dependiendo del caso, del  $i$ -ésimo valor simulado de la serie contrafactual a partir de la distribución predictiva para cada tiempo  $t = T_0 + 1, \dots, T$ , donde  $T_0$  es el periodo inicial de la intervención y  $T$  el tamaño de la serie. Así, para obtener muestras de la distribución aproximada predictiva posterior del efecto de interés (impacto del tratamiento) se define la variable:

$$\delta_{t,i} = \mathbf{Y}_t - \tilde{\mathbf{Y}}_{t,i} \quad (1-2)$$

donde  $\mathbf{Y}_t$  representaría el valor observado del contaminante en el aire bajo el estado de cuarentena o aislamiento preventivo por la Covid-19, y  $\tilde{\mathbf{Y}}_{t,i}$  representa la misma variable, pero bajo un escenario sin intervención o sin cuarentena. Adicional a esta estimación de impacto puntual, también puede ser de interés el impacto acumulado del efecto de la intervención o tratamiento

$$\text{impacto acumulado}_{t,i} = \sum_{t=T_0}^T \delta_{t,i} \quad (1-3)$$

Donde:

$i$  : Índice de la simulación.

$T_0$  : Inicio de las restricciones.

$T$  : Fin del periodo de estudio.

$\mathbf{Y}_t$  : Es un vector de  $q$  estaciones o una matriz de  $r$  contaminantes por  $q$  estaciones con los valores observados, para el tiempo  $t$ .

$\tilde{\mathbf{Y}}_{t,i}$  : E un vector de  $q$  estaciones o una matriz de  $r$  contaminantes por  $q$  estaciones, para el tiempo  $t$  en la simulación  $i$ .

### 1.1.5. Contrafactual Artificial - ArCo

El Contrafactual Artificial (ArCo) permite obtener una estimación del impacto promedio y provee las herramientas para realizar pruebas de hipótesis para verificar la significancia del impacto (Carvalho et al., 2018a). con el fin de calcular ArCo definimos,  $Y_t$  en función de la variable  $Y'_t$ , la cual representa la serie si no hubiese sucedido la intervención, de la siguiente manera.

$$\mathbf{Y}_t = \begin{cases} Y'_t & t = 1, \dots, T_0 \\ Y'_t + \delta_t & t = T_0, \dots, T \end{cases}$$

De lo anterior podemos concluir que para tiempos mayores o iguales a  $T_0$  se tiene la siguiente igualdad:

$$\delta_t = Y_t - Y'_t$$

Definimos la variable  $\Delta_T$  como:

$$\Delta_T = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T \delta_t$$

La hipótesis nula en la cual se asume que no existe un impacto en la serie  $Y_t$  después del periodo de intervención  $T_0$  se define como:

$$H_0 : \Delta_T = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T \delta_t = 0 \quad (1-4)$$

El estimador para el contrafactual ArCo se define a continuación:

$$\begin{aligned} \hat{\Delta}_T &= \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T \hat{\delta}_t \\ \hat{\delta}_t &= \frac{1}{Q} \sum_{q=1}^Q \hat{\delta}_{q,t} \\ \hat{\delta}_{q,t} &= \frac{\hat{y}_{q,t} - y_{q,t}}{y_{q,t}} \end{aligned} \quad (1-5)$$

Las propiedades estadísticas de este estimador se describen en Carvalho et al. (2018b).



- $T$  : índice de tiempo,  $T > T_0$ ,
- $T_0$  : Tiempo inicial de la intervención,
- $y_{q,t}$  : Valor observado en la estación  $q$  en el tiempo  $t$ ,
- $\hat{y}_{q,t}$  : Valor estimado para la estación  $q$  en el tiempo  $t$ ,

## 1.2. Impacto agregado

Con el objetivo de tener un resultado global para un contaminante específico, en esta tesis proponemos una medida de impacto agregada. Esta medida resume el impacto medido en  $q$  estaciones a una medida global  $\phi_r$  como se define a continuación:

$$\begin{aligned}\phi_{t,r,q} &= Y_{t,r,q} - \tilde{Y}_{t,r,q} \\ \phi_{t,r} &= \frac{1}{Q} \sum_{q=1}^Q \phi_{t,r,q} \\ \phi_r &= \sum_{t=T_0}^T \phi_{t,r}\end{aligned}\tag{1-6}$$

Donde:

$Y_{t,r,q}$  : es el valor observado en el tiempo  $t$  para la estación  $q$  en el contaminante  $r$ .

$\tilde{Y}_{t,r,q}$  : es valor de la serie contrafactual en el tiempo  $t$  para la estación  $q$  en el contaminante  $r$ .

El objetivo de la ecuación (1-6) es obtener una estimación global de la disminución en la concentración de un contaminante a partir de las estaciones estudiadas.

En el siguiente capítulo se presentarán los modelos estadísticos que se emplearán para aplicar los métodos de evaluación de impacto descritos en este capítulo. En el capítulo 3 estas herramientas de evaluación de impacto serán caracterizadas empleando métodos de simulación. En el capítulo 4 dichas medidas serán abordadas nuevamente para responder la pregunta de investigación planteada en 1.0.1. En el capítulo 5 se presenta el paquete de R (R Core Team, 2022) **DynamicImpact**, el cual es una implementación de los métodos desarrollados en esta tesis. Finalmente, en el capítulo 6 se presentan las conclusiones generales y algunas líneas de trabajo futuro.

## 2 Modelo Lineal Dinámico Matriz-Variado

Sea  $\mathbf{Y}_t$  una matriz aleatoria de dimensión  $r \times q$ , donde  $r$  representa el número de contaminantes monitoreados y  $q$  representa el número de estaciones de monitoreo donde el contaminante ha sido medido. Siguiendo el enfoque definido en Quintana (1985), Quintana (1987) y West y Harrison (1997b) la respuesta  $\mathbf{Y}_t$  es modelada bajo una estructura dinámica o de espacio-estado de la siguiente forma:

$$\begin{aligned} \text{Ecuación de Observación: } \mathbf{Y}_t &= \mathbf{F}_t \boldsymbol{\Theta}_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \mathbf{V}_t, \boldsymbol{\Sigma}) \\ \text{Ecuación de Evolución: } \boldsymbol{\Theta}_t &= \mathbf{G}_t \boldsymbol{\Theta}_{t-1} + \boldsymbol{\Omega}_t, & \boldsymbol{\Omega}_t &\sim N(0, \mathbf{W}_t, \boldsymbol{\Sigma}) \end{aligned} \quad (2-1)$$

donde

- $\mathbf{F}_t$  : es una matriz de dimensión  $r \times p$  de covariables independientes,
- $\boldsymbol{\Theta}_t$  : es una matriz desconocida de parámetros de regresión de dimensión  $p \times q$ ,
- $\varepsilon_t$  : es la matriz de errores de la ecuación de observación de dimensión  $r \times q$ ,
- $\mathbf{V}_t$  : es la matriz de varianzas de dimensión  $r \times r$  asociada con  $\varepsilon_t$ ,
- $\mathbf{G}_t$  : es la matriz de evolución de orden  $p \times p$ ,
- $\boldsymbol{\Omega}_t$  : es la matriz de ruido de evolución de orden  $p \times q$ ,
- $\mathbf{W}_t$  : es la matriz de varianza de orden  $p \times p$  asociada con  $\boldsymbol{\Omega}_t$ ,
- $\boldsymbol{\Sigma}$  : es la matriz de escala del sistema de orden  $q \times q$ .
- $T$ : periodo total estudio, 122 días en nuestro caso.
- $T_0$ : Tiempo desde el cual se acumula el impacto.

$$t = 1 \dots T.$$

Las matrices  $\mathbf{F}_t$ ,  $\mathbf{V}_t$ ,  $\mathbf{G}_t$  y  $\mathbf{W}_t$  se asumen conocidas antes de observar  $\mathbf{Y}_t$ , para nuestro caso específico tomamos  $\mathbf{G}_t = \mathbb{I}_p$ , y hacemos  $\mathbf{V}_t = \mathbf{V}_{ML}$ , donde  $\mathbf{V}_{ML}$  es el estimador de máxima

verosimilitud para  $\mathbf{V}_t$ . Las filas de la matriz  $\mathbf{F}_t$  representa los contaminantes en el año 2019 y  $\mathbf{Y}_t$  la serie de tiempo de contaminantes para el año 2020.  $\mathbf{W}_t$  es estimado mediante factores descuento como se presenta posteriormente en la ecuación (2-7).

### 2.0.1. Modelo conjugado

Se supone que la matriz  $\Sigma$  es desconocida y su incertidumbre es descrita por una distribución a priori Wishart invertida, en el tiempo  $t - 1$  (después de observar  $\mathbf{Y}_{t-1}$ ), dada por la expresión:

$$\Sigma \sim W^{-1}(\mathbf{S}_{t-1}, d_{t-1}). \quad (2-2)$$

Donde  $W^{-1}$  representa la distribución inversa de Wishart,  $d_{t-1}$  representa los grados de libertad y  $\mathbf{S}_{t-1}$  representa la matriz de escala.  $d_{t-1}$  y  $\mathbf{S}_{t-1}$  son los hiperparámetros en el tiempo  $t - 1$ , para el tiempo  $t = 0$  se fijan valores iniciales que en nuestro caso definirán a priori no informativas. Para los tiempos  $t > 0$  estos hiperparámetros se actualizan secuencialmente como se presenta en (2-11).

Regularmente, se utiliza como distribución a priori para  $\Theta_{t-1}$  la distribución normal matriz-variada:

$$\Theta_{t-1} \sim N(\mathbf{M}_{t-1}, \mathbf{C}_{t-1}, \Sigma). \quad (2-3)$$

Donde  $\mathbf{M}_{t-1}$ ,  $\mathbf{C}_{t-1}$  y  $\Sigma$  son hiperparámetros que representan la media, la matriz de covarianza por filas y la matriz de covarianza por columnas respectivamente. En el tiempo  $t = 0$  se toman valores fijos que en nuestro caso nuevamente representarán a priori no informativas y se actualizarán de manera secuencial siguiendo las ecuaciones (2-4), (2-5) y (2-6). Para el lector interesado en conocer más sobre distribuciones matriz-variadas ver Gupta y Nagar (2018).

Las distribuciones posteriores relacionadas con el modelo (2-1) podrían ser aproximadas mediante métodos MCMC (Markov chain Monte Carlo), sin embargo, como consecuencia de los supuestos dados por (2-1), (2-3) y asumiendo  $\Sigma$  conocida, es posible realizar el ajuste utilizando las ecuaciones (2-4), (2-5) y (2-6) (Quintana, 1987). Emplear este enfoque puede ser útil en situaciones en las cuales se cuenta con un poder computacional limitado. Como consecuencia de los supuestos definidos para el modelo (2-1) y la a priori definida en (2-3) se tienen los siguientes resultados:

$$\Theta_t \sim N(\mathbf{M}_t^*, \mathbf{C}_t^*, \Sigma) \quad \text{donde} \quad \mathbf{C}_t^* = \mathbf{W}_t + \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t^\top, \quad \mathbf{M}_t^* = \mathbf{G}_t \mathbf{M}_{t-1} \quad (2-4)$$

$$\mathbf{Y}_t \sim N(\hat{\mathbf{Y}}_t, \check{\mathbf{Y}}_t, \Sigma) \quad \text{donde} \quad \check{\mathbf{Y}}_t = \mathbf{V}_t + \mathbf{F}_t \mathbf{C}_t^* \mathbf{F}_t^\top, \quad \hat{\mathbf{Y}}_t = \mathbf{F}_t \mathbf{M}_t^* \quad (2-5)$$

$$\Theta_t | \mathbf{Y}_t \sim N(\mathbf{M}_t, \mathbf{C}_t, \Sigma) \quad \text{donde} \quad \mathbf{C}_t = \mathbf{C}_t^* - \mathbf{A}_t \check{\mathbf{Y}}_t \mathbf{A}_t^\top, \quad \mathbf{M}_t = \mathbf{M}_t^* + \mathbf{A}_t \hat{\mathbf{E}}, \quad (2-6)$$

$$\hat{\mathbf{E}} = \mathbf{Y} - \hat{\mathbf{Y}}, \quad \mathbf{A}_t = \mathbf{C}_t^* \mathbf{F}_t^\top \check{\mathbf{Y}}_t^{-1}.$$

Donde  $\hat{\mathbf{Y}}_t$  es la predicción de  $\mathbf{Y}_t$ , y  $\check{\mathbf{Y}}_t$  es la matriz de covarianza por filas para la predicción.

Dado que  $\mathbf{W}_t$  es una matriz de covarianza asociada a las filas de  $\Theta_t$  definirla de manera objetiva representa un reto en el ajuste del modelo (2-1). Adaptamos el enfoque de factores descuento propuesto por Ameen y Harrison (1985), siguiendo la implementación utilizada por West y Harrison (1985). Expresamos  $\mathbf{W}_t$  como se define en (2-8), donde  $\mathbf{B}$  es una matriz diagonal de factores de descuento  $1/\beta^{1/2}$  definida en (2-7) y  $\beta \in (0, 1)$ .

$$\mathbf{B} = (\beta \mathbb{I}_p)^{-\frac{1}{2}}. \quad (2-7)$$

Como consecuencia de (2-7) tenemos que:

$$\mathbf{W}_t = \mathbf{B} \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t^\top \mathbf{B} - \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t^\top. \quad (2-8)$$

De la ecuación (2-8) vemos que el parámetro  $\beta$  controla a la variabilidad del proceso evolutivo (Cardona-Jiménez y de B. Pereira, 2021), lo cual lo hace fundamental porque puede afectar la detección de impacto. El rango de valores usual para el hiperparámetro  $\beta$  es (0.8, 0.95). En nuestro caso  $\beta$  es definido a través de un proceso de selección de modelos. Se estiman 100 modelos para diferentes valores de  $\beta$  y se seleccionan el  $\beta$  del mejor modelo según el criterio del factor de Bayes. Este proceso no es requerido para los casos en los cuales se utiliza

MCMC, ya que la distribución posterior de  $\mathbf{W}_t$  es aproximada junto con las distribuciones de los demás parámetros del modelo.

Ahora, suponiendo  $\Sigma$  desconocida bajo la distribución en (2-2) se tiene que:

$$\begin{bmatrix} \mathbf{Y}_t \\ \Sigma \end{bmatrix} \sim NW^{-1}(\hat{\mathbf{Y}}_t, \check{\mathbf{Y}}_t, \mathbf{S}_{t-1}, d_{t-1}), \quad (2-9)$$

Donde  $NW^{-1}$  Representa la distribución normal inversa Wishart en su versión matricial.  $\hat{\mathbf{Y}}_t$  es el parámetro de localización definido en (2-5),  $d_{t-1}$  los grados libertad definidos en (2-11),  $\check{\mathbf{Y}}_t$  y  $\mathbf{S}_{t-1}$  son las matrices de covarianza por filas y columnas respectivamente definidas en (2-5) y (2-11). Consecuentemente tenemos que:

$$\Sigma | \mathbf{Y}_t \sim W^{-1}(\mathbf{S}_t, d_t), \quad (2-10)$$

Donde

$$\begin{aligned} \mathbf{S}_t &= \mathbf{S}_{t-1} + \hat{\mathbf{E}}_t^\top \check{\mathbf{Y}}^{-1} \hat{\mathbf{E}}_t, \\ d_t &= d_{t-1} + 1. \end{aligned} \quad (2-11)$$

De esta forma la distribución conjunta de  $\Theta_t$  y  $\Sigma$ , la distribución predictiva de  $\mathbf{Y}_t$ , la distribución posterior de  $\Theta_t$  y  $\Sigma$  en el tiempo  $t$  están dadas por:

$$\begin{bmatrix} \Theta_t \\ \Sigma \end{bmatrix} \sim NW^{-1}(\mathbf{M}_t^*, \mathbf{C}_t^*, \mathbf{S}_{t-1}, d_{t-1}). \quad (2-12)$$

Distribución Predictiva:

La ecuación para la predicción, equivalente a (2-5), tiene una distribución T la cual es la versión matricial de la distribución t-student con  $d_{t-1}$  grados de libertad, parámetro de localización  $\hat{\mathbf{Y}}_t$ , matriz de covarianza por filas  $\check{\mathbf{Y}}_t$  definidos en (2-5) y matriz de covarianza por columnas  $\mathbf{S}_{t-1}$  definida en (2-11).

$$\mathbf{Y}_t \sim T(\hat{\mathbf{Y}}_t, \check{\mathbf{Y}}_t, \mathbf{S}_{t-1}, d_{t-1}). \quad (2-13)$$

Distribución posterior conjunta:

$$\begin{bmatrix} \Theta_t \\ \Sigma \end{bmatrix} | \mathbf{Y}_t \sim NW^{-1}(\mathbf{M}_t, \mathbf{C}_t, \mathbf{S}_t, d_t). \quad (2-14)$$

El proceso de actualización para  $\mathbf{M}_t^*$ ,  $\mathbf{C}_t^*$ ,  $\mathbf{C}_t$ ,  $\mathbf{A}_t$ ,  $\hat{\mathbf{E}}_t$ ,  $\hat{\mathbf{Y}}_t$  y  $\check{\mathbf{Y}}_t$  es el mismo presentado en (2-4), (2-5) y (2-6).

Las relaciones (2-10), (2-12), (2-13) y (2-14) permiten actualizar las distribuciones de  $\Theta_t$ ,  $\Sigma_t$  y  $\mathbf{Y}_t$  de manera sistemática con información nueva. Es importante resaltar que  $\mathbf{M}_t$ ,  $\mathbf{C}_t$ , y  $\mathbf{A}_t$ , son independientes del valor actual de  $\Sigma$  (Quintana, 1985).

### Predicción

Bajo las condiciones previas es posible demostrar que la predicción  $K$  unidades de tiempo adelante para un modelo ajustado utilizado, las ecuaciones recurrentes (2-4), (2-5) y (2-6), está dada por (West y Harrison, 1997b):

$$\mathbf{Y}_{T_0+K} \sim T(\hat{\mathbf{Y}}_{T_0+K}, \check{\mathbf{Y}}_{T_0+K}, \mathbf{S}_{T_0+K-1}, d_{T_0+K-1}) \quad (2-15)$$

Donde:

$$\begin{aligned} \check{\mathbf{Y}}_{T_0+K} &= \mathbf{V} + \mathbf{F}_{T_0}^\top \mathbf{C}_{T_0+K-1}^* \mathbf{F}_{T_0} \\ \mathbf{M}_{T_0+K}^* &= \mathbf{G}_t \mathbf{M}_{T_0+K-1}^* \\ \hat{\mathbf{Y}}_{T_0+K} &= \mathbf{F}_{T_0+K} \mathbf{M}_{T_0+K}^* \\ \mathbf{A}_{T_0+K} &= \mathbf{C}_{T_0+K-1} \mathbf{F}'_{T_0+K} \check{\mathbf{Y}}_{T_0+K}^{-1} \\ \mathbf{C}_{T_0+K}^* &= \mathbf{W}_{T_0+k} + \mathbf{G}_{T_0+K} \mathbf{C}_{T_0+K-1}^* \mathbf{G}_{T_0+K}^\top \\ \mathbf{S}_{T_0+K} &= \check{\mathbf{Y}}_{T_0+K} \mathbf{S}_{T_0} \end{aligned}$$

Es necesario resaltar que para realizar la predicción en tiempo  $T_0 + K$ , es necesario realizar las predicciones desde el tiempo  $T_0 + 1$  hasta  $T_0 + K - 1$ . Los valores iniciales ( $K = 0$ ) son  $\mathbf{M}_{T_0+K}^* = \mathbf{M}_{T_0}$  y  $\mathbf{C}_{T_0+K}^* = \mathbf{C}_{T_0}$ .

### Estimación de la incertidumbre para el impacto

Con el fin de obtener intervalos de probabilidad posterior para la estimación del impacto, empleamos el algoritmo FSTS (Forward State Trajectories Sampler) desarrollado por Cardona-Jiménez y de B. Pereira (2021). A diferencia de la definición original en la cual únicamente se toman simulaciones de  $\Theta_t^i$ , tomamos adicionalmente muestras simuladas de  $\mathbf{Y}_t^i$ . Este algoritmo nos permite obtener muestras para  $\hat{\mathbf{Y}}$  y  $\Theta$  mantenido la dependencia temporal derivada de las ecuaciones de recurrencia definidas en (2-4), (2-5) y (2-6).

---

**Algoritmo 1** Forward state trajectories sampler - Modificado
 

---

```

1: procedure
2:   for  $i \leftarrow 1$  to  $M$  do
3:     Calcular  $\mathbf{W}_t = \mathbf{B}_t \mathbf{C}_{t-1} \mathbf{B}_t - \mathbf{C}_{t-1}$    for  $t = T_0, \dots, T$ 
4:     Simular  $\Theta_{t-1}^i = N(\mathbf{M}_{t-1}, \mathbf{C}_{t-1}, \mathbf{S}_{t-1})$    for  $t = T_0, \dots, T$ 
5:     Simular  $\Omega_t^i = N(\mathbf{0}, \mathbf{W}_t, \mathbf{S}_t)$    for  $t = T_0, \dots, T$ 
6:     Calcular  $\Theta_t^i = \Theta_{t-1}^i + \Omega_t^i$    for  $t = T_0, \dots, T$ 
7:     Simular  $\mathbf{Y}_t^i = N(\mathbf{F}_t^i \Theta_t^i, \mathbf{S}_t)$    for  $t = T_0, \dots, T$ 
8:   end for
9:    $\hat{\mathbf{Y}}_t = \{\mathbf{Y}_t^1, \dots, \mathbf{Y}_t^M\}$ 
10: end procedure

```

---

Con las muestras resultantes para  $\hat{\mathbf{Y}}_t$  podemos evaluar el impacto con la metodología presentada en la sección 1.1.1 las cuales dependen de sumas acumuladas de las predicciones después del inicio de la intervención dado por  $T_0$ .

## 2.0.2. Aproximación posterior y predicción mediante MCMC

De manera alternativa es posible ajustar el modelo definido en (2-1) utilizando Markov Chain Monte Carlo (MCMC), en cuyo caso es posible utilizar remover la restricción de compartir  $\Sigma$  en las ecuaciones de observación y evolución, además de estimar  $\mathbf{W}_t$  sin necesidad de emplear factores de descuento. El algoritmo 2 permite obtener la predicción  $K$  unidades de tiempo adelante:

---

**Algoritmo 2** Predicción: Monte carlo
 

---

```

1: procedure
2:   for  $j \leftarrow 1$  to  $K$  do
3:     Simular  $\Theta_{T_0+j}^i = N(\Theta_{T_0}^i, \mathbf{W}^i, \Sigma^i)$ 
4:     Calcular  $\mu_{T_0+j}^i = \mathbf{F}_{T_0+j} \Theta_{T_0+j}^i$ 
5:     Simular  $\mathbf{Y}_{T_0+j}^i = N(\mu_{T_0+j}^i, \mathbf{V}^i, \Sigma^i)$ 
6:   end for
7:    $\hat{\mathbf{Y}}^i = \{\mathbf{Y}_{T_0+1}^i, \dots, \mathbf{Y}_{T_0+K}^i\}$ 
8: end procedure

```

---

En el algoritmo 2  $i$  representa el índice de la simulación, con  $i = 1, \dots, M$ , donde  $M$  es el número de simulaciones. En la estimación por MCMC realizada en este trabajo estimamos  $\Sigma$ ,  $\mathbf{W}$  y  $\mathbf{V}$  como valores constantes para  $1 < t < T$ .

Es importante resaltar que  $\Theta_{T_0+j}^i \sim N(\Theta_{T_0}^i, \mathbf{W}^i, \Sigma^i)$  lo que implica que no se continúa el proceso evolutivo de  $\Theta$  después de la intervención en el tiempo  $T_0$  (Brodersen et al., 2015).

En este caso, la variabilidad y el intervalo de probabilidad posterior de la predicción después  $T_0$  pueden ser estimados utilizando las muestras de la distribución posterior y no se requiere utilizar el algoritmo 1.

## 2.1. Modelo Lineal Dinámico Matriz-Variado espacial

La estructura espacio-estado propuesto por (Stroud et al., 2001), esta dada por:

$$\begin{aligned} \text{Ecuación de Observación: } \mathbf{Y}_t &= \mathbf{S}(\mathbf{F}_{t,j}; \boldsymbol{\Theta}_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \mathbf{V}_t, \boldsymbol{\Sigma}) \\ \text{Ecuación de Evolución: } \boldsymbol{\Theta}_{t,j} &= \mathbf{G}_t \boldsymbol{\Theta}_{t-1,j} + \boldsymbol{\Omega}_t, \quad \boldsymbol{\Omega}_t \sim N(0, \mathbf{W}_t, \boldsymbol{\Sigma}) \end{aligned} \quad (2-16)$$

$$\mathbf{S}(t; \boldsymbol{\Theta}_t) = \sum_{j=1}^J \pi_j \mathbf{F}_{t,j} \boldsymbol{\Theta}_{t,j} \quad (2-17)$$

$\mathbf{S}(t; \boldsymbol{\Theta}_t)$  es una suma ponderada en  $\pi_j$  de la matriz de covariables  $\mathbf{F}_{t,j}$  para elemento  $j$  de la mezcla multiplicada por su respectiva matriz de parámetros  $\boldsymbol{\Theta}_{t,j}$ .

$$\pi_{j,q} \propto |\Sigma_{\text{mixture}}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (x_q - \mu_j) \Sigma_{\text{mixture}} (x_q - \mu_j)^\top \right\} \quad (2-18)$$

$$j = 1 \dots J, \quad q = 1 \dots Q, \quad r = 1 \dots R$$

Donde:

- $Q$  : representa el número de estaciones,
- $J$  : Representa el número de elementos en la mezcla,
- $\pi_j : [\pi_{j,q}, \dots, \pi_{j,q}]^\top$  : Representan los pesos de cada uno de los elementos de la mezcla.
- $\mathbf{x}_q$  : Ubicación de la estación  $q$  en coordenadas planares,
- $\mathbf{X} : [\mathbf{x}_1, \dots, \mathbf{x}_q]$ ,
- $\Sigma_{\text{mixture}}$  : es la matriz de varianzas de dimensión  $2 \times 2$  asociada con a las coordenadas planares,
- $\mu_j$  : es vector de dimensiones  $2 \times 1$  con las coordenadas del kernel  $j$ ,
- $\boldsymbol{\Theta}_{t,j}$  : es una matriz desconocida de parámetros de regresión de dimensión  $p \times q$ ,
- $\boldsymbol{\Theta}_t : [\boldsymbol{\Theta}_{t,j}, \dots, \boldsymbol{\Theta}_{t,J}]$ ,
- $\mathbf{F}_{t,j}$  : es una matriz de dimensión  $r \times p$  de covariables independientes para la componente  $j$  de la mezcla,



- $\varepsilon_t$  : es la matriz de errores de la ecuación de observación de dimensión  $r \times q$ ,
- $\mathbf{V}_t$  : es la matriz de varianzas de dimensión  $r \times r$  asociada con  $\varepsilon_t$ ,
- $\mathbf{G}_t$  : es la matriz de evolución de orden  $p \times p$ ,
- $\mathbf{\Omega}_t$  : es la matriz de ruido de evolución de orden  $p \times q$ ,
- $\mathbf{W}_t$  : es la matriz de varianza de orden  $p \times p$  asociada con  $\mathbf{\Omega}_t$ ,
- $\mathbf{\Sigma}$  : es la matriz de escala del sistema de orden  $q \times q$ ,
- $\mathbf{Y}_t$  : es una matriz de dimensiones dimensión  $r \times q$ .

Dada la inclusión de  $\Sigma_{\text{mixture}}$  y  $\mu_j$  en el modelo no es posible ajustar el modelo de manera conjugada utilizando las ecuaciones recurrentes (2-4), (2-5) y (2-6) por lo cual utilizamos MCMC para realizar la predicción.

---

**Algoritmo 3** Predicción Monte Carlo - modelo espacial
 

---

```

1: procedure
2:   for  $k \leftarrow 1$  to  $K$  do
3:     Simular  $\Theta_{t+k,j}^i = N(\Theta_{t,j}^i, \mathbf{W}_{t+k}^i, \mathbf{\Sigma}^i)$  for  $j = 1, \dots, J$ 
4:     Calcular  $\mu_{t+k}^i = \sum_{j=1}^J \pi_j \mathbf{F}(X)_{t+k} \Theta_{t+k,j}^i$ 
5:     Simular  $\mathbf{Y}_{t+k}^i = N(\mu_{t+k}^i, \mathbf{V}_{t+k}^i, \mathbf{\Sigma}^i)$ 
6:   end for
7:    $\hat{\mathbf{Y}}^i = \{\mathbf{Y}_{T+1}^i, \dots, \mathbf{Y}_{T+K}^i\}$ 
8: end procedure

```

---

Similarmente, a lo expuesto en el Modelo Lineal Dinámico Matriz-Variado, tenemos que  $\Theta_{t+k,j}^i \sim N(\Theta_{t,j}^i, \mathbf{W}_{t+k}^i, \mathbf{\Sigma}^i)$  lo que implica que este no continua el proceso evolutivo de  $\Theta$  después de la intervención  $T_0$ .

El modelo definido en (2-16) permite utilizar una matriz de covariables diferente para cada elemento de la mezcla, en nuestro caso solo contamos con una matriz de covariables, por lo cual en este trabajo utilizamos una modificación de este modelo presentada en (2-19).

## 2.2. Modelo Lineal Dinámico Matriz-Variado espacial modificado

Tomando el modelo (2-16) como base, en este trabajo agregamos el componente no espacial  $\mu_t^{\text{no espacial}}$  con su respectiva matriz de evolución  $\Theta_t^{\text{no espacial}}$ . Adicionalmente, a esto se incluye el parámetro,  $\lambda \in (0, 1)$  el cual define la importancia de la parte espacial en el modelo. De la

definición podemos ver que valores cercanos a cero en  $\lambda$  representan una baja importancia para la componente espacial del modelo.

$$\begin{aligned}
\text{Observación:} \quad \mathbf{Y}_t &= (1 - \lambda)\mu_t^{\text{no espacial}} + \lambda\mu_t^{\text{espacial}} \\
\text{Media espacial:} \quad \mu_t^{\text{espacial}} &= \mathbf{S}(\mathbf{F}_t; \Theta_{t,1}^{\text{espacial}}, \dots, \Theta_{t,J}^{\text{espacial}}) \\
\text{Media no espacial:} \quad \mu_t^{\text{no espacial}} &= \mathbf{F}_t \Theta_t^{\text{no espacial}} \\
\text{Evolución espacial:} \quad \Theta_{t,j}^{\text{espacial}} &= \Theta_{t-1,j}^{\text{espacial}} + \Omega_t^{\text{espacial}}, \\
\text{Evolución no espacial:} \quad \Theta_t^{\text{no espacial}} &= \Theta_{t-1}^{\text{no espacial}} + \Omega_t^{\text{no espacial}}
\end{aligned} \tag{2-19}$$

Donde:

$$\begin{aligned}
\varepsilon_t &\sim N(0, \mathbf{V}_t, \Sigma) \\
\Omega_t^{\text{no espacial}} &\sim N(0, \mathbf{W}_t, \Sigma) \\
\Omega_t^{\text{espacial}} &\sim N(0, \mathbf{W}_t, \Sigma)
\end{aligned}$$

$$\mathbf{S}(\mathbf{F}_t; \Theta_{t,1}^{\text{espacial}}, \dots, \Theta_{t,J}^{\text{espacial}}) = \sum_{j=1}^J \pi_j \mathbf{F}(X)_t \Theta_{t,j}^{\text{espacial}}$$

$$\pi_{j,q} \propto |\Sigma_{\text{mixture}}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (x_q - \mu_j) \Sigma_{\text{mixture}} (x_q - \mu_j)^\top \right\}$$

Donde:  $\mathbf{F}_t$  es una matriz de dimensión  $r \times p$  de covariables independientes, la misma utilizada en el modelo (2-1).

El modelo definido en la ecuación (2-19) es producto de esta tesis, es el modelo espacial usado en este trabajo. Los modelos (2-1) y (2-19) serán utilizados para construir series contrafactuales, con las cuales se evaluará el impacto en la calidad del aire debido a las restricciones de movilidad en el parque automotor en la ciudad de Medellín durante los meses de marzo y abril del año 2020.

## 3 Simulación

Con el fin de estudiar el desempeño del método propuesto en esta tesis, se realiza un estudio de simulación en el cual se evalúa la capacidad de los modelos presentados para estimar el impacto de un evento que comienza en el tiempo  $T_0$ . Se evalúan tres escenarios: Cuando se presenta un impacto en la misma dirección de la tendencia, en dirección contraria a la tendencia y cuando no se presenta impacto.

### 3.1. Algoritmos de simulación para el modelo (2-1)

---

**Algoritmo 4** Simulación para cada estación

---

```
1: procedure SIMULACION ESTACION( $T_0, \Phi$ )
2:   for  $i \leftarrow 1$  to  $M$  do
3:     for  $t \leftarrow 1$  to  $T$  do
4:       Simular  $\Theta_t^i$  Desde  $N(\Theta_{t-1}^i, \mathbf{W}, \Sigma)$ 
5:       if  $t \leq T_0$  then
6:         calcular  $\mu_t^i \leftarrow \mathbf{F}_t \Theta_t^i$ 
7:       else if  $t > T_0$  then
8:         Simular  $\Phi_t^i$  Desde  $N(\Phi_{t-1}^i, \mathbf{W}, \Sigma)$ 
9:         calcular  $\mu_t^i \leftarrow \mathbf{F}_t \Theta_t^i + \mathbf{F}_t \Phi_t^i$ 
10:      end if
11:      Simular  $\mathbf{Y}_y^i$  Desde  $N(\mu_t^i, \mathbf{V}, \Sigma)$ 
12:    end for
13:  end for
14:   $\mathbf{Y}^i = \{\mathbf{Y}_1^i, \dots, \mathbf{Y}_T^i\}$ 
15: end procedure
```

---

**Algoritmo 5** Simulación global

---

```

1: procedure
2:   for  $\Phi \in \Xi$  do
3:     for  $q \in Q$  do
4:       for  $r \in R$  do
5:         for  $t_0 \in \tau$  do
6:           calcular  $\text{SIMUL}_{q,r,t_0} \leftarrow \text{Simulación Estación}(T_0 = t_0, \Phi, q, r)$ 
7:         end for
8:       end for
9:     end for
10:   end for
11: end procedure

```

---

El proceso de simulación implementado se divide en dos partes. En la primera parte, el algoritmo 4 genera una serie de tiempo con media  $\mu_t = \mathbf{F}_t \Theta_t$  antes del inicio del evento artificial y con media  $\mu_t = \mathbf{F}_t \Theta_t + \mathbf{F}_t \Phi_t$  después del inicio del evento. En la segunda parte, el algoritmo 5 genera conjuntos de  $q$  estaciones, cada una de ellas con  $r$  contaminantes. Los parámetros utilizados en la simulación se definen a continuación:

- $\mu_t$ , media de la serie simulada en el tiempo  $t$ , dimensión  $r \times q$ , donde  $r$  es el número de estaciones y  $q$  el número de contaminantes simulados.
- $\Theta_t$  : Es la matriz de coeficientes de dimensión  $p \times q$ .
- $\Phi_t$  : es la matriz de coeficientes sumada al proceso después de tiempo  $T_0$ , su dimensión es  $p \times q$ . Estos parámetros representan el efecto del impacto.
- $M$  : Número total de simulaciones.
- $T$  : Tamaño de la serie simulada.
- $T_0$  : Tiempo de inicio del impacto ( $T_0 < T$ ).
- $Q$  : Conjunto de posibles valores para el número de estaciones.
- $R$  : Conjunto de posibles valores para el número de contaminantes.
- $\tau$  : Conjunto de posibles valores simulados hasta el tiempo de inicio del evento en  $T_0$ .

La definición para  $\mathbf{V}$ ,  $\mathbf{W}$ ,  $\Sigma$  y  $\mathbf{F}_t$ , es la misma utilizada en la ecuación (2-1). El algoritmo 4 realiza la simulación de una estación individual. Para este algoritmo requerimos definir los valores que tomarán  $\mathbf{F}_t$ ,  $\mathbf{W}$ ,  $\mathbf{V}$ , y  $\Sigma$ , ya que estos valores no serán generados en el proceso de simulación, únicamente se fijan dentro este.

- $\mathbf{F}_t$  : se construye utilizando valores observados de las series de contaminantes para 2019 y 2020, y es únicamente utilizada para construir las simulaciones de  $\mathbf{Y}_t$ .
- $\mathbf{W}$  : Es la identidad de dimensiones  $p \times p$ .
- $\mathbf{\Sigma}$  : Es la estimación por máxima verosimilitud de la matriz de covarianza por columnas de  $\mathbf{F}_t$ .
- $\mathbf{V}$  : Es la estimación por máxima verosimilitud de la matriz de covarianza por filas de  $\mathbf{F}_t$ .

El uso de estimación por máxima para la definición de las matrices de covarianza es consecuencia de las dificultades que se tiene para definir valores razonables para  $\mathbf{W}$ ,  $\mathbf{V}$ , y  $\mathbf{\Sigma}$ . Las definiciones de  $\mathbf{V}$ ,  $\mathbf{W}$ ,  $\mathbf{\Sigma}$  y  $\mathbf{F}_t$  se toman por conveniencia numérica, es posible utilizar otros valores. Con el objetivo de simular “los centros de monitoreo de una ciudad” se genera un grupo de estaciones, en las cuales los valores simulados están correlacionados. Los escenarios de simulación usados para evaluar el modelo son los siguientes:

- $Q = \{2, 3, 4\}$ : Conjunto con posibles valores para el número de estaciones.
- $R = \{2, 3\}$ : Conjunto con posibles valores para el número de contaminantes.
- $\Xi = \{\Phi_1, \Phi_2, \Phi_3\}$ , representan los casos en los cuales el impacto se da en la dirección de la tendencia, en dirección contraria a la tendencia o cuando no hay impacto.
- $\tau = \{[0.5 \times T], [0.7 \times T], [0.8 \times T], [0.9 \times T]\}$ , representa el tiempo transcurrido hasta el momento de inicio de la intervención. Proponemos estos valores con el objetivo de estudiar el efecto del tiempo de inicio de la intervención sobre la capacidad de detección de impacto.

El tamaño de los conjuntos de parámetros está sujeto a restricciones computacionales, en futuros trabajos esperamos expandir sus dimensiones, las limitaciones computacionales encontradas en este trabajo se presentan en la sección 6.3.

El algoritmo 5 permite simular un grupo de estaciones  $q$  con una cantidad de contaminantes  $r$ , con un tiempo de inicio de intervención  $t_0$ . La simulación es realizada utilizando el software estadístico R (R Core Team, 2022), en su versión 4.2 y Stan en su versión 2.29.0 (Stan Development Team, 2022). Las simulaciones son ejecutadas en dos equipos cada uno con 128GB RAM DDR4 y un procesador Intel Core i7-12700KF. La simulación realizada en este trabajo toma aproximadamente dos semanas. Una de las principales limitaciones es aumentar el número de estaciones y contaminantes, ya que al hacer esto se generan problemas numéricos para garantizar que las estimaciones posteriores de  $\mathbf{\Sigma}$ ,  $\mathbf{W}$  y  $\mathbf{V}$  sean semidefinidas positivas.

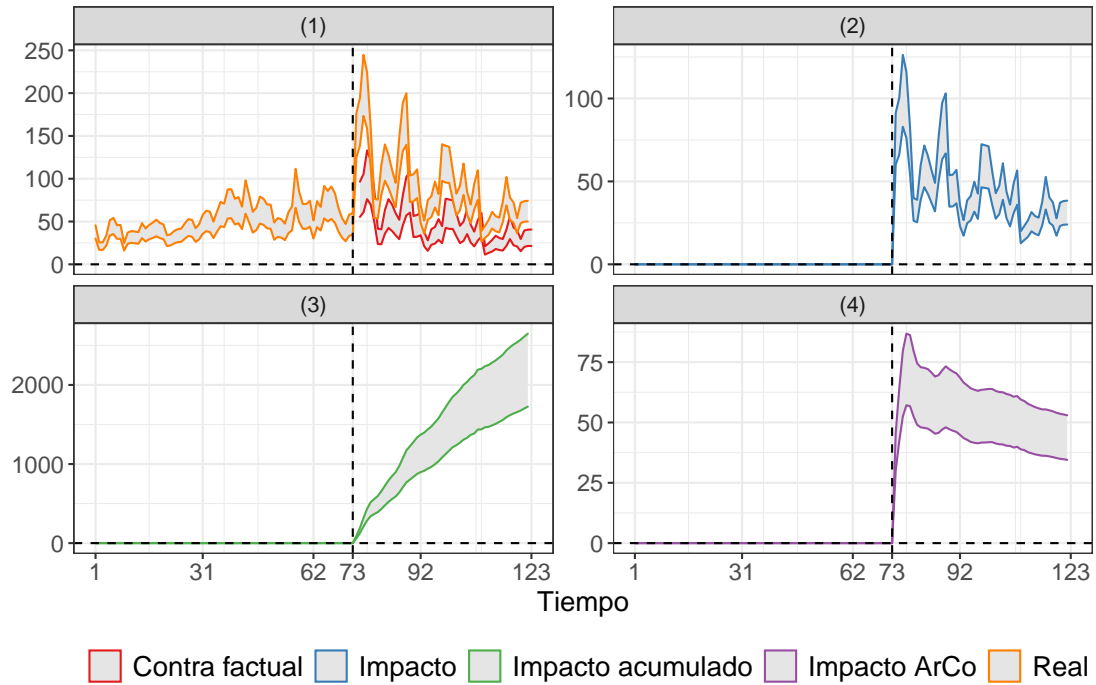
## 3.2. Impacto en la misma dirección de la tendencia

Este es el caso en el cual se tiene una serie con una tendencia creciente y el impacto observado también es creciente; o si se tiene una serie con una tendencia decreciente y el impacto observado también es decreciente.

Si la media de la serie antes del impacto está dada por  $\mu_t = \mathbf{F}_t \boldsymbol{\Theta}_t$  y después del tiempo  $T_0$  está dada por  $\mu_t = \mathbf{F}_t \boldsymbol{\Theta}_t + \mathbf{F}_t \boldsymbol{\Phi}_t$ , se tiene que:  $\text{sign}(\boldsymbol{\Theta}_t) = \text{sign}(\boldsymbol{\Phi}_t)$ , donde  $\text{sign}(\cdot)$  es la función signo aplicada a cada uno de los elementos de la matriz. Esta función se define como:

$$\text{sign}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Un ejemplo del caso para el cual se presenta un impacto simulado se puede observar en la Figura 3-1. La línea punteada vertical representa el tiempo en el cual comienza el evento. Se puede observar que todos los impactos son positivos y significativos (el valor cero cae fuera de los intervalos), lo cual es consistente con el escenario simulado.

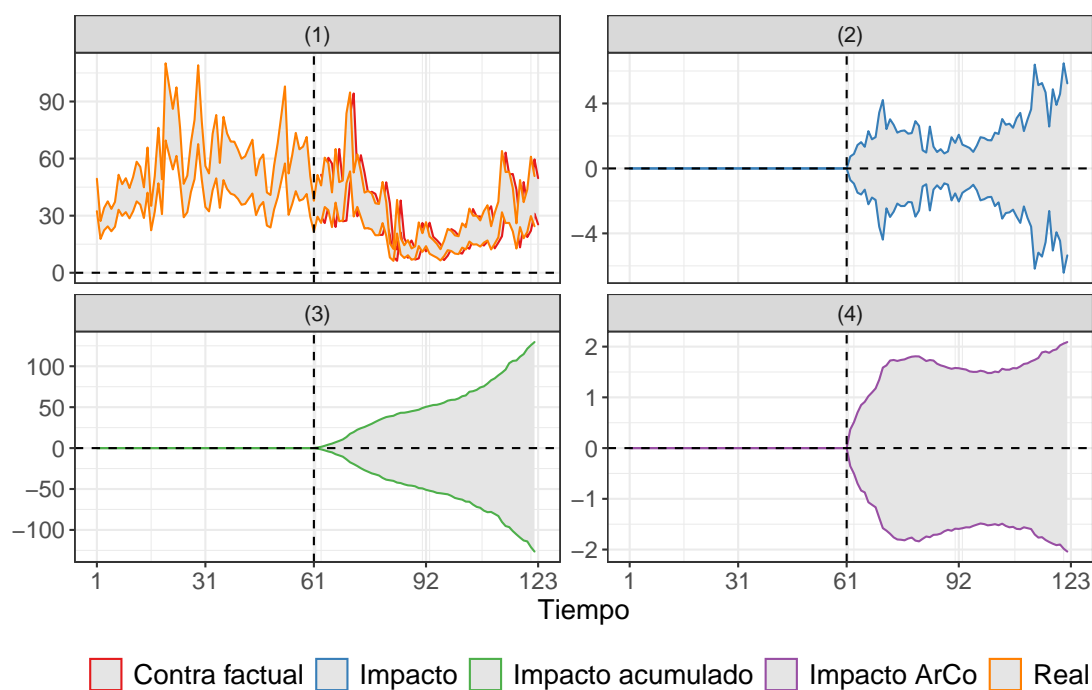


**Figura 3-1:** Resultados de simulación cuando se presenta un impacto en una estación simulada. En el panel (1) se presenta la serie real simulada y su contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5) respectivamente.

En el panel (1) se puede ver que después del  $T_0 = 73$  la serie real sufre un cambio de nivel, lo cual hace que se separe del contrafactual. Esta separación se puede en el panel (2) en el cual se tiene un impacto positivo a partir del inicio de la intervención, en los paneles (3) y (4) se observan valores positivos para el impacto acumulado y el ArCo. En ambos casos los límites superior e inferior de los intervalos de probabilidad posterior están por encima de cero, lo que implica que se detecta un impacto.

### 3.3. Caso en el cual no se presenta impacto

Similarmente, en la Figura 3-2, se presenta el escenario sin impacto para una estación simulada. Vemos que los intervalos de probabilidad posterior del 90 % de la serie contrafactual y la serie real se solapan, esto es consistente con el impacto, el impacto acumulado y el ArCo ya que sus intervalos contiene el cero.

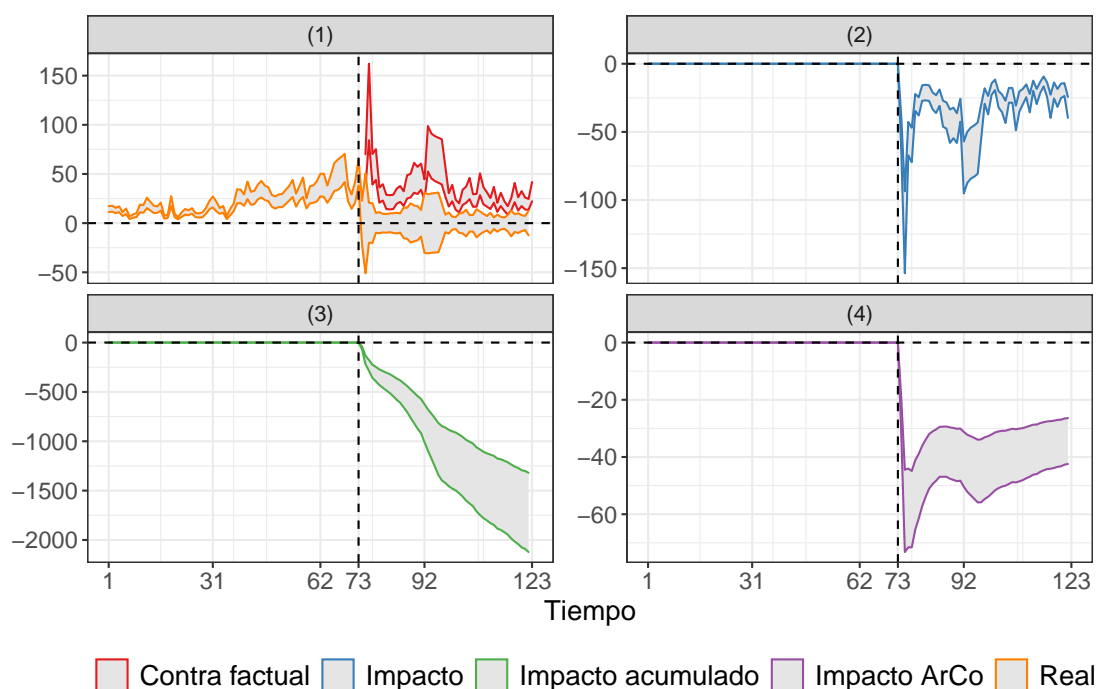


**Figura 3-2:** Ejemplo resultado de simulación cuando no se presenta un impacto en una estación simulada. En el panel (1) se presenta la serie real simulada y su contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5), presentados en capítulo 1.

En el panel (1) se puede ver que después del  $T_0 = 61$  los intervalos de probabilidad posterior de la serie real y contrafactual se solapan, lo que indica que no existe un impacto significativo.

### 3.4. Impacto en dirección contraria a la tendencia

En este caso se tiene una serie con una tendencia creciente y un impacto decreciente o viceversa. Este caso se puede expresar matemáticamente como:  $\text{sign}(\Theta_t) \neq \text{sign}(\Phi_t)$ . En la Figura 3-3 vemos que después de la línea punteada en  $t = 73$  todos los impactos son negativos, lo cual es consistente con el escenario simulado.



**Figura 3-3:** Ejemplo resultado de simulación cuando se presenta un impacto con dirección contraria a la tendencia en una estación simulada. En el panel (1) se puede ver que después del  $T_0 = 73$  la serie real sufre un cambio de nivel en dirección contraria a la tendencia, lo cual hace que se separe del contrafactual. Los paneles (2), (3) y (4) representan el impacto (1-2), el impacto acumulado (1-3) y el impacto ArCo (1-5), presentados en capítulo 1.

### 3.5. Análisis de simulación del impacto agregado (1.2) empleando el modelo (2-1)

Cada uno de los escenarios definidos en 3.1 es evaluado en una serie tiempo del mismo tamaño que el periodo de estudio (o periodo de restricciones de movilidad) sobre el cual se realiza el análisis (122 días). Se evalúan matrices con  $q \times r = 4, 6, 8, 10, y, 12$  elementos (número de contaminantes por estaciones), con el evento de interés comenzando en  $T_0$ .



Para evaluar si se detectó un impacto significativo se define el periodo de intervención (PI) como el conjunto de observaciones entre el inicio de la intervención  $T_0$  y el final de la serie  $T$ .

$$PI = \{T_0, T_0 + 1, \dots, T\}.$$

Para cada tiempo  $t$  en el periodo de intervención (PI) se construye un intervalo de probabilidad posterior del 90 % para el impacto acumulado definido en (1-3), usando la distribución posterior de la predicción. Cuando el límite inferior y superior del intervalo tienen el mismo signo se concluye que existe un impacto significativo, si los límites inferior y superior tienen signos diferentes se concluye que no hay evidencia de un impacto. Definimos la variable indicadora  $d_t$  como:

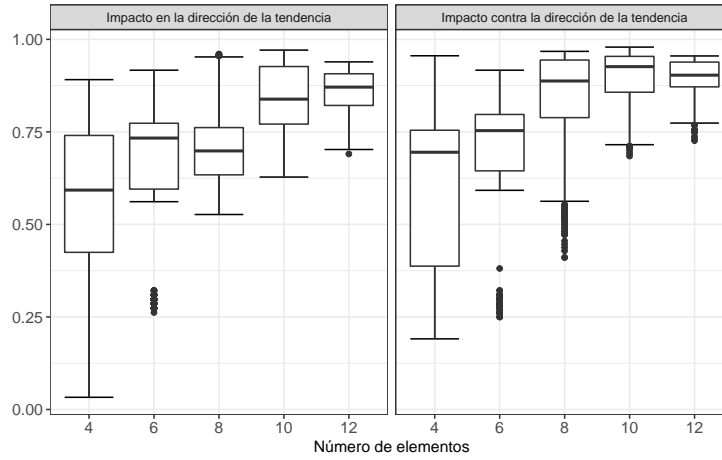
$$d_t = \begin{cases} 1 & \text{si se detecta un impacto,} \\ 0 & \text{en caso contrario} \end{cases}$$

Utilizando  $d_t$  podemos obtener la proporción de observaciones en las cuales se detectó un impacto como:

$$\text{proporción de impacto} = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T d_t \quad (3-1)$$

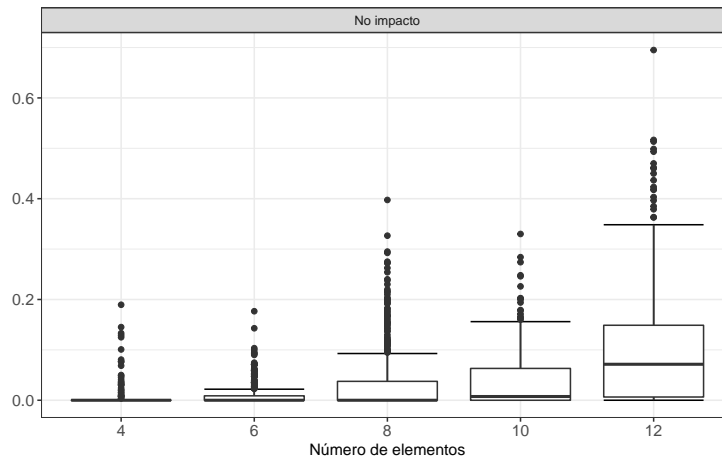
La ecuación (3-1) representa la proporción de tiempos en el periodo de intervención PI en los cuales se detecta un impacto.

Una vez tenemos la proporción de impacto (3-1), podemos observar su comportamiento en función de los parámetros de la simulación. Por ejemplo, en la Figura **3-4** podemos ver la distribución de la proporción del impacto para diferente número de elementos, donde un elemento se define como: número de estaciones por número contaminantes. Se observa entonces en este gráfico que la variabilidad en la proporción de observaciones en las cuales se detecta un impacto se reduce en función del número de elementos. Este resultado nos indica que la detección del impacto puede ser más precisa cuando se tiene en cuenta un mayor número de elementos, es decir, mayor número de estaciones y contaminantes modelados conjuntamente.



**Figura 3-4:** Resultado de la simulación en función del número de elementos, para los casos en los cuales existe un impacto.

En la Figura 3-5 tenemos el número de elementos contra la proporción de impacto definida en (3-1). Podemos observar que a medida que aumenta el número de elementos incrementa la variabilidad en la proporción de observaciones en las cuales se detecta impacto. Este resultado implica que posiblemente se tiene una mayor cantidad de falsos positivos cuando se tienen situaciones con alto número de contaminantes o estaciones. Aun así, en la Tabla 3-1 vemos que todos los casos en los cuales no se presenta impacto, el cero está incluido en el intervalo de probabilidad posterior, lo cual implica que la proporción de impacto 3-1 para este escenario es estadísticamente igual a cero.



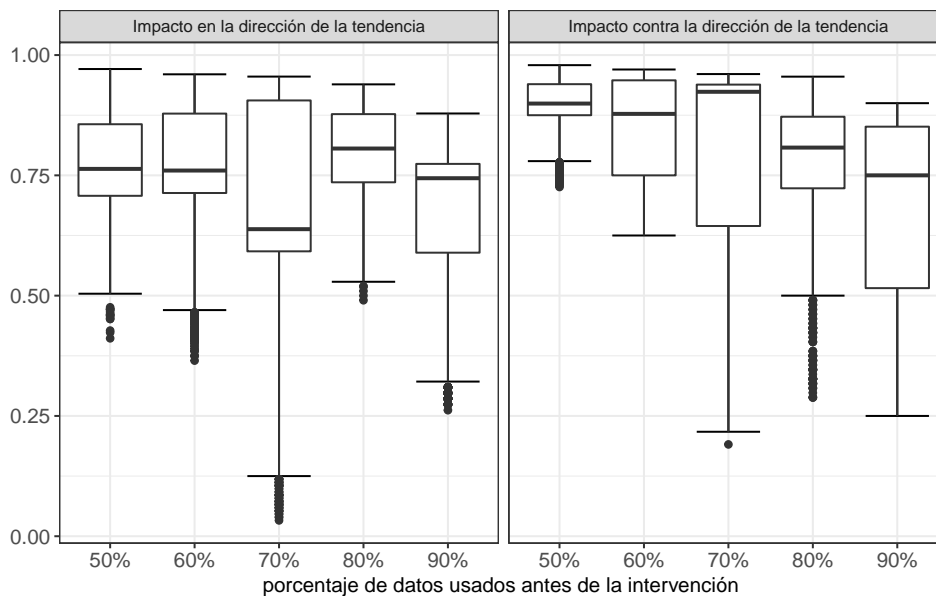
**Figura 3-5:** Resultado de la simulación en función del número de elementos, para los casos en los cuales no existe un impacto.

**Tabla 3-1:** Resultados en función del número de elementos de la matriz  $\mathbf{Y}_t$ .

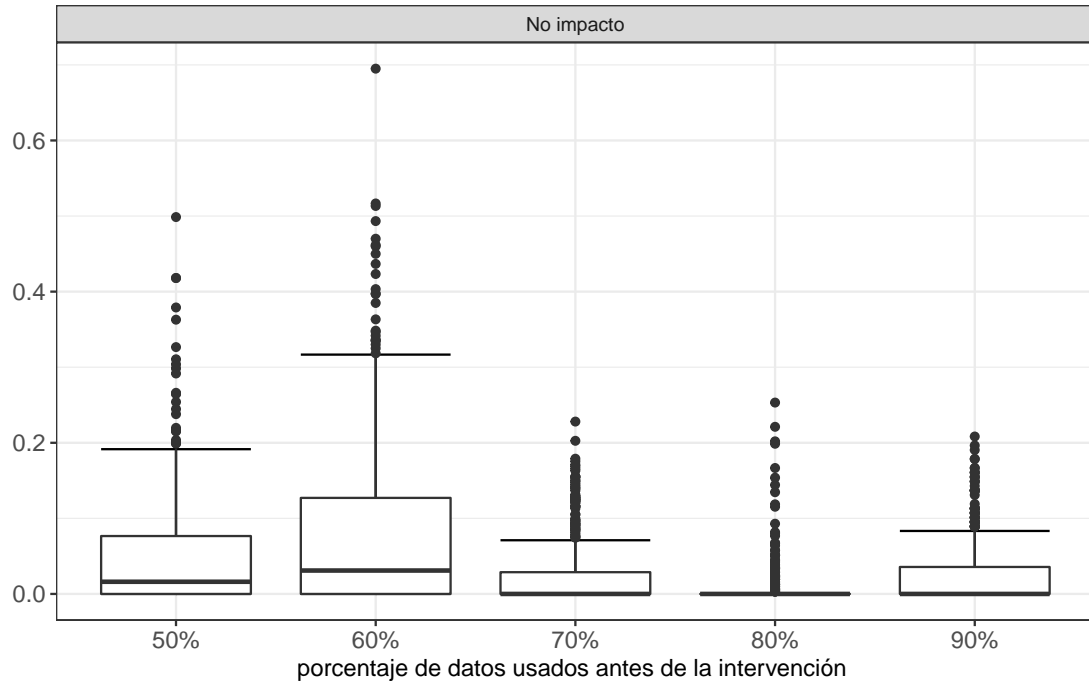
	Escenario	Número de elementos	Probabilidad de detección	q0.05	q0.95
1	Impacto en la dirección de la tendencia	4	0.501	0.072	0.768
2	Impacto en la dirección de la tendencia	6	0.702	0.286	0.872
3	Impacto en la dirección de la tendencia	8	0.752	0.58	0.945
4	Impacto en la dirección de la tendencia	10	0.85	0.708	0.947
5	Impacto en la dirección de la tendencia	12	0.864	0.744	0.926
6	Impacto contra la dirección de la tendencia	4	0.669	0.315	0.94
7	Impacto contra la dirección de la tendencia	6	0.728	0.274	0.872
8	Impacto contra la dirección de la tendencia	8	0.877	0.517	0.955
9	Impacto contra la dirección de la tendencia	10	0.913	0.723	0.97
10	Impacto contra la dirección de la tendencia	12	0.901	0.804	0.946
11	No impacto	4	0.005	0	0.012
12	No impacto	6	0.009	0	0.06
13	No impacto	8	0.047	0	0.178
14	No impacto	10	0.048	0	0.142
15	No impacto	12	0.116	0	0.325

En la Tabla **3-1** vemos que la menor detección para los escenarios en los cuales se presenta un impacto, se da cuando el impacto tiene la misma dirección de la tendencia. Este resultado nos indica que el proceso de detección para este escenario puede ser más retador comparado con el caso en el cual el impacto se da en dirección contraria a la tendencia.

Para los casos en los cuales se presenta un impacto en la misma dirección de la tendencia en la Figura **3-6**, no vemos una tendencia clara en la proporción del impacto (3-1) para los diferentes niveles de  $\tau$ . Por el contrario, en el escenario en el cual se presenta un impacto en dirección contraria de la tendencia, parece tener una detección más precisa cuando el porcentaje de datos disponibles para calcular el impacto es del 50%.

**Figura 3-6:** Resultado de la simulación en función del porcentaje de datos usados antes de la intervención, para los casos en los cuales existe un impacto.

Tampoco observamos una tendencia clara en la Figura 3-7 para los casos en los cuales no se presenta impacto en función del porcentaje de datos usados antes de la intervención.



**Figura 3-7:** Resultado de la simulación en función del porcentaje de datos usados antes de la intervención, para los casos en los cuales no existe un impacto.

Los resultados presentados en las Figuras 3-7, 3-6 y la Tabla 3-2, nos indica que la metodología es robusta respecto al porcentaje de datos disponibles para calcular el impacto.

**Tabla 3-2:** Resultados en función del porcentaje de datos usados antes de la intervención

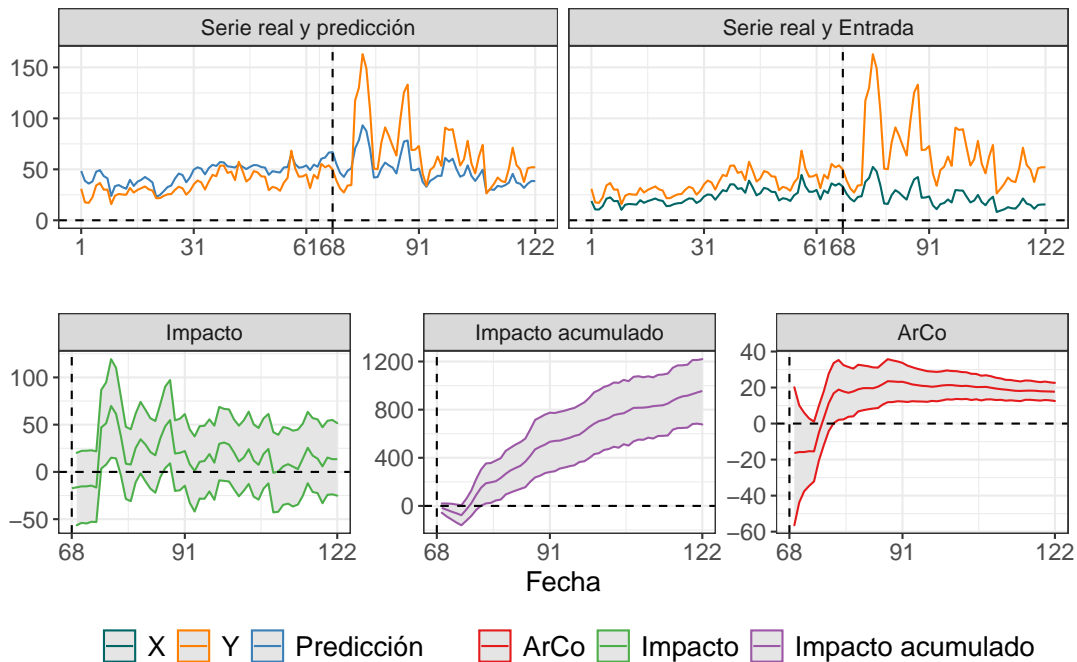
Escenario		Porcentaje de datos usados antes de la intervención	Probabilidad de detección	q0.05	q0.95
1	Impacto en la dirección de la tendencia	50 %	0.799	0.556	0.937
2	Impacto en la dirección de la tendencia	60 %	0.789	0.425	0.945
3	Impacto en la dirección de la tendencia	70 %	0.735	0.072	0.939
4	Impacto en la dirección de la tendencia	80 %	0.821	0.654	0.926
5	Impacto en la dirección de la tendencia	90 %	0.676	0.286	0.857
6	Impacto contra la dirección de la tendencia	50 %	0.895	0.761	0.969
7	Impacto contra la dirección de la tendencia	60 %	0.874	0.67	0.96
8	Impacto contra la dirección de la tendencia	70 %	0.837	0.329	0.951
9	Impacto contra la dirección de la tendencia	80 %	0.798	0.346	0.946
10	Impacto contra la dirección de la tendencia	90 %	0.694	0.274	0.886
11	No impacto	50 %	0.063	0	0.192
12	No impacto	60 %	0.107	0	0.307
13	No impacto	70 %	0.029	0	0.126
14	No impacto	80 %	0.01	0	0.043
15	No impacto	90 %	0.031	0	0.113

### 3.6. Resultado para el modelo espacial definido en (2-19).

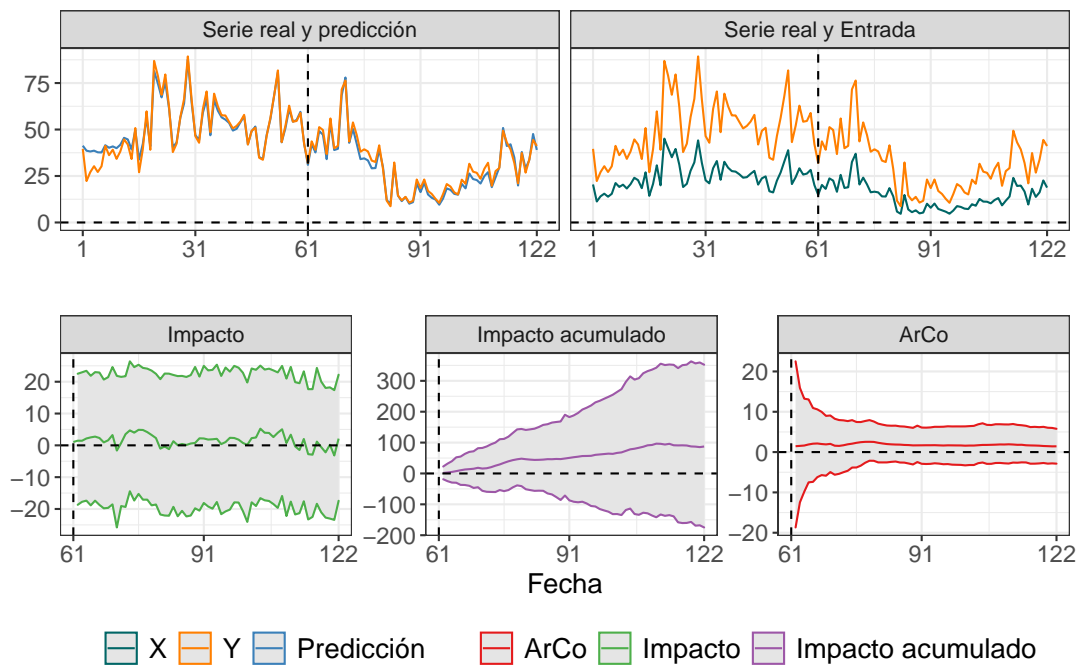
Por limitaciones computacionales y dado el límite tiempo que posee este trabajo, únicamente presentamos ejemplos particulares de resultados de simulación agregados utilizando la ecuación (1-6) para el modelo espacial. Las limitaciones al aproximar las distribuciones posteriores utilizando este modelo provienen de la cantidad adicional de parámetros que deben ser estimados, ya que se requiere aproximar las distribuciones posteriores de  $J$  matrices de parámetros (número de elementos en la mezcla)  $\Theta_{t,j}^{\text{espacial}}$  adicionales.

Una limitación adicional proviene de la estimación de la distribución posterior de  $\lambda$ , ya que la estimación de  $\mathbf{Y}$  depende de los productos:  $(1 - \lambda)\mu_t^{\text{no espacial}}$  y  $\lambda\mu_t^{\text{espacial}}$ , lo que es equivalente a  $(1 - \lambda)\mathbf{F}_t\Theta_t^{\text{no espacial}}$  y  $\lambda\sum_{j=1}^J\pi_j\mathbf{F}(X)_t\Theta_{t,j}^{\text{espacial}}$ , lo que implica que el parámetro  $\lambda$  multiplica tanto a  $\Theta_t^{\text{no espacial}}$  como a todos los  $\Theta_{t,j}^{\text{espacial}}$ . En trabajos futuros esperamos extender el proceso de simulación para incluir este modelo.

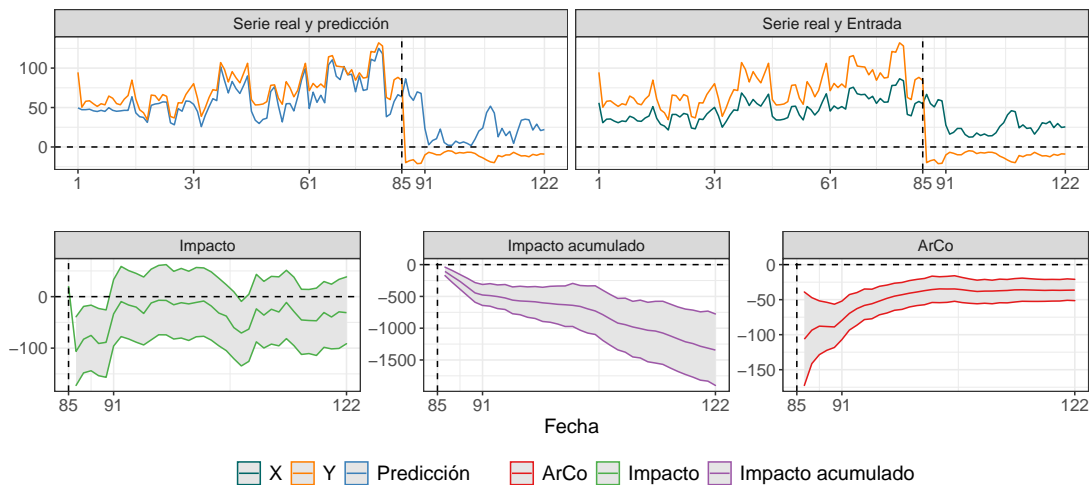
Las Figuras 3-8, 3-10 y 3-9 presentan ejemplos de detección de impacto en el modelo espacial definido en (2-19) para los casos en los cuales se presenta un impacto en dirección de la tendencia, en dirección contraria a la tendencia y cuando no se presenta un impacto respectivamente.



**Figura 3-8:** Resultados de simulación cuando se presenta un impacto en la dirección de la tendencia para el modelo espacial definido en (2-19).



**Figura 3-9:** Resultados de simulación cuando no se presenta un impacto para el modelo espacial definido en (2-19).



**Figura 3-10:** Resultados de simulación cuando se presenta un impacto en la dirección contraria de la tendencia para el modelo espacial definido en (2-19).

Los resultados de simulación presentados en este capítulo nos dan una idea de las ventajas y limitaciones que pueden tenerse al momento de realizar evaluaciones de impacto mediante modelos contrafactuales construidos empleando Modelos Dinámicos Bayesianos. Debemos

tener en cuenta que resultados solo sirven como guía y que dejamos para trabajos futuros la exploración de conjuntos de parámetros de simulación más amplios.

## 4 Aplicación: Evaluación del impacto relacionado con la hipótesis de investigación.

La pregunta de investigación para este trabajo definida en el capítulo 1 es la siguiente: “¿Durante los periodos de cuarentena estricta (marzo - abril del 2020) en los cuales se restringió la movilidad de vehículos y las actividades industriales, se produjo una reducción en la concentración de contaminantes PM 10, PM 2.5, NO, NO<sub>x</sub> y NO<sub>2</sub> en el Valle de Aburrá?”. En este capítulo esperamos responder a esta pregunta utilizando el paquete **dynamicimpact** para aproximar las distribuciones posteriores del modelo (2-1). Dicho paquete es un producto desarrollado en este trabajo y será introducido en el capítulo 5.

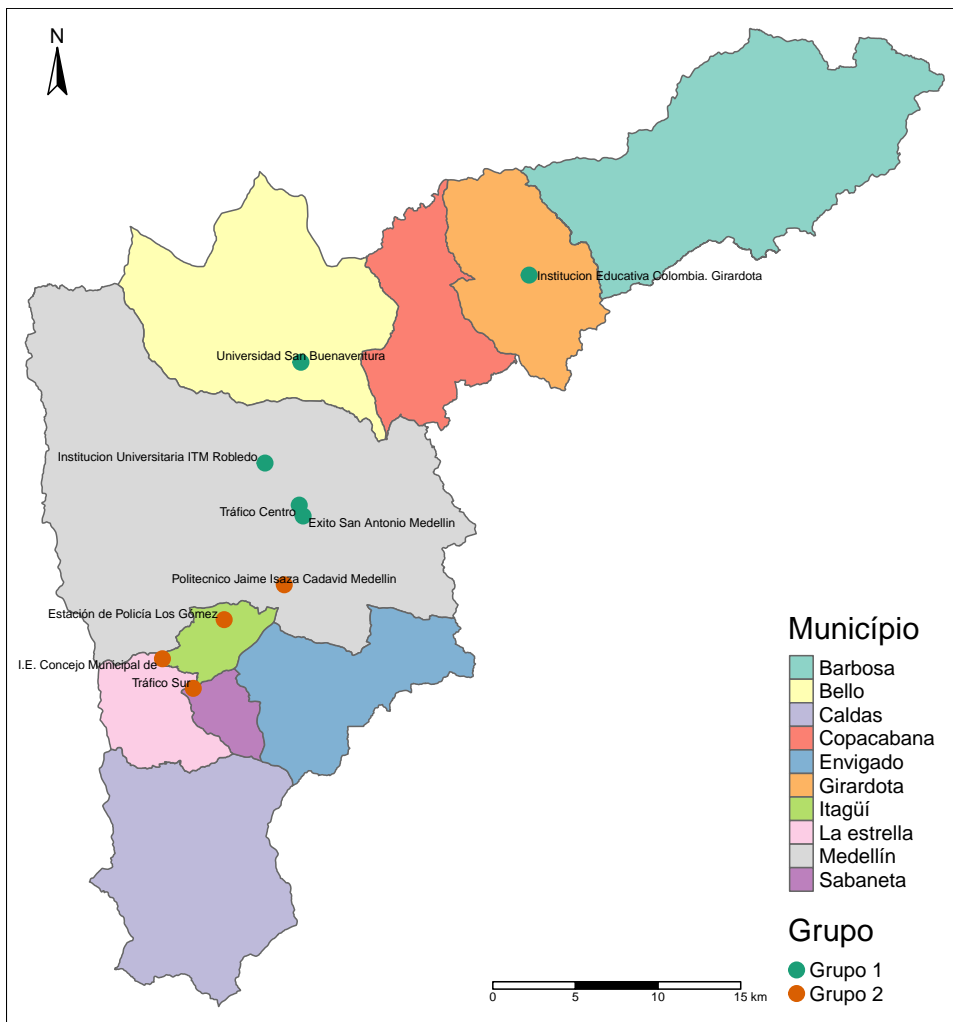
Dado que es posible que sea requerido esperar un periodo de alrededor de 10 días para ver una reducción en la concentración de contaminantes en el aire (Mendez-Espinosa et al., 2020), en este trabajo analizamos 3 periodos de tiempo para evaluar el impacto: [2020-03-16, 2020-03-31], [2020-03-31, 2020-04-15] y [2020-04-15, 2020-05-01]. Esto nos permite ubicar el periodo de tiempo en el cuál se presenta el impacto en caso de que este realmente exista.

### 4.1. Resultados para PM 10

El material particulado PM 10 está compuesto de partículas respirables con diámetro menor a  $10\mu m$ . Las fuentes usuales de estas partículas son el parque automotor y las actividades industriales (Tian et al., 2014). Múltiples estudios han encontrado efectos adversos en la salud de las personas expuestas de manera constante a este contaminante (Lu et al., 2015).

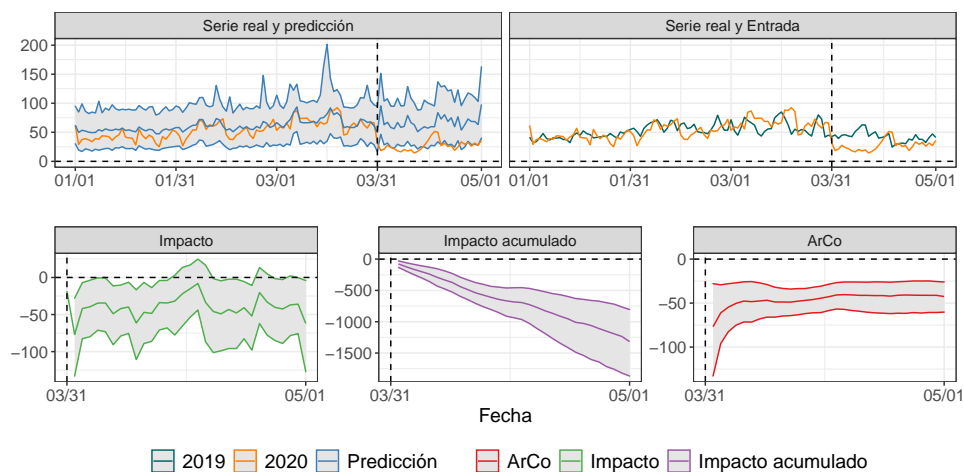
El proceso de estimación mediante MCMC puede generar problemas a medida que aumentan las dimensiones de las matrices de covarianza  $\Sigma$ ,  $\mathbf{W}$  y  $\mathbf{V}$ , debido a las dificultades numéricas que se generan para garantizar que estas matrices sean positivas semidefinidas. Para aliviar esta situación separamos las estaciones en dos grupos partiendo el conjunto de datos de manera balanceada, utilizando como variables de entrada la ubicación de las estaciones en coordenadas planares (IGAC, 2004). En la Figura 4-1 se presenta la ubicación geográfica de estos grupos de estaciones. El lector interesado en el proceso de clusterización empleado en este trabajo puede referirse a Papenberg y Klau (2020).





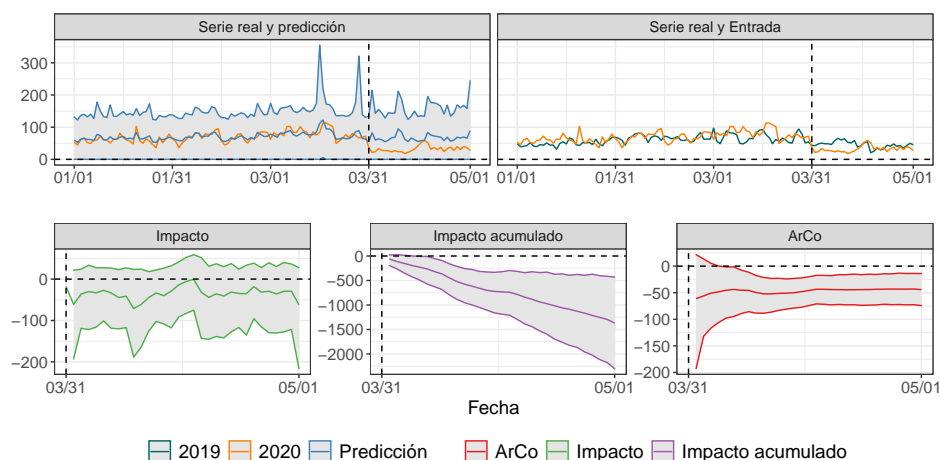
**Figura 4-1:** Grupos de estaciones para el contaminante PM 10 obtenidos utilizando el algoritmo de clusterización presentado por Papenberg y Klau (2020).

En la Figura 4-2 vemos los resultados obtenidos al aplicar el impacto agregado (1-6) relacionado con el contaminante PM 10 para el grupo de estaciones 1. Se toma como fecha de inicio para el cálculo del impacto el 31 de marzo de 2020 y fecha final el primero de mayo de 2020. En esta figura y en la Tabla 4-1 podemos ver que existe un impacto significativo para los resultados agregados en el periodo [2020-03-31, 2020-04-15], el cual está contenido en el periodo de tiempo de las restricciones estrictas impuestas por la pandemia del Covid-19. Este resultado representa el impacto global para el grupo de estaciones 1.



**Figura 4-2:** Resultados agregados PM 10 para el grupo de estaciones 1, utilizando la ecuación (1-6).

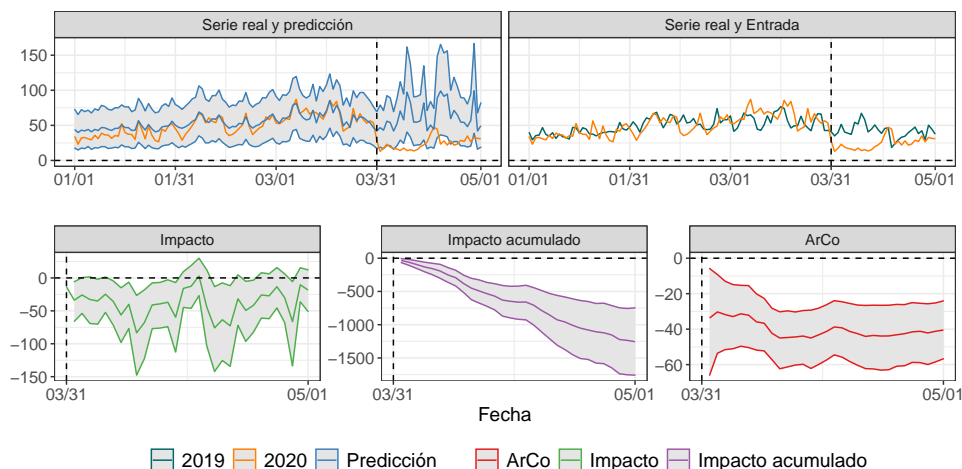
En la Figura 4-3 vemos los resultados para una de las estaciones pertenecientes al grupo 1 (estación Tráfico Centro). Observamos un comportamiento consistente con el observado en el impacto agregado (Figura 4-2), aunque los intervalos de probabilidad posterior para la predicción son más amplios. El impacto individual es significativo para esta estación, pues el cero no está contenido en el intervalo de probabilidad posterior (ver Tabla 4-1). Para el resto de estaciones pertenecientes al grupo 1 los resultados son bastante similares: impacto significativo y mayor variabilidad en la estimación (ver Figuras 7-1, 7-2, 7-3, 7-4). Sin embargo, al evaluar el impacto agregado (4-2) obtenido al modelar conjuntamente dicho grupo de estaciones, podemos ver que la estimación del impacto se torna más precisa. Este resultado es consistente con lo observado en el estudio de simulación presentado en el capítulo anterior, donde se observó que un mayor número de estaciones consideradas en el modelo conjunto puede mejorar la precisión con la que se calcula el impacto.



**Figura 4-3:** Resultados PM 10 para la estación Tráfico Centro.

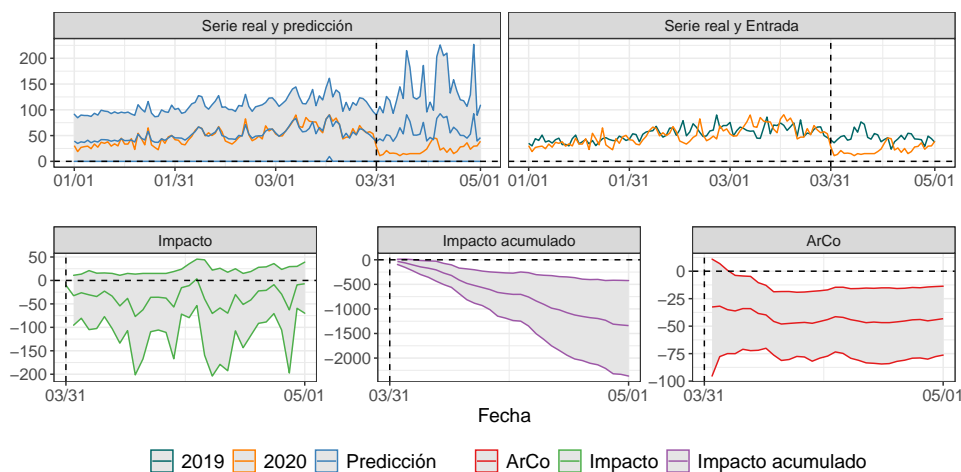
	Tiempo de corte	IC ArCo	IC Cumsum	Estación
13	2020-04-15	[-60.60, -30.61]	[-908.96, -459.09]	Global grupo 1
14	2020-04-15	[-77.44, -22.20]	[-1161.65, -332.95]	Estación Tráfico Centro
15	2020-04-15	[-60.18, -18.28]	[-902.67, -274.23]	Exito San Antonio - Medellin
16	2020-04-15	[-66.13, -18.69]	[-992.00, -280.40]	Institucion Educativa Colombia. Girardota
17	2020-04-15	[-96.51, -14.72]	[-1447.65, -220.79]	Institucion Universitaria ITM Robledo
18	2020-04-15	[-62.36, -17.24]	[-935.46, -258.61]	Universidad San Buenaventura

**Tabla 4-1:** Resultados PM 10 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 1”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15.



**Figura 4-4:** Resultados agregados PM 10 para el grupo de estaciones 2, utilizando la ecuación (1-6).

En la Figura 4-4 vemos los resultados obtenidos al aplicar el impacto agregado (1-6) relacionado con el contaminante PM 10 para el grupo de estaciones 2. Se toma nuevamente como fecha de inicio para el cálculo del impacto el 31 de marzo de 2020 y fecha final el primero de mayo de 2020. En esta Figura y en la Tabla 4-2 podemos ver que existe un impacto significativo para los resultados agregados en el periodo [2020-03-31, 2020-04-15], el cual está contenido en el periodo de tiempo de las restricciones estrictas impuestas por la pandemia del Covid-19. Este resultado representa el impacto global para el grupo de estaciones 2.



**Figura 4-5:** Resultados PM 10 para la estación de Policía Los Gómez.

En la Figura 4-5 vemos los resultados para una de las estaciones pertenecientes al grupo 2 (Estación de Policía Los Gómez). Observamos un comportamiento consistente con el observado en el impacto agregado (Figura 4-4). Para el resto de estaciones pertenecientes al grupo 2 los resultados son bastante similares: impacto significativo y mayor variabilidad en la estimación (ver Tablas 7-2 y 7-1). Nuevamente, vemos que se tiene impacto significativo individual en todas las estaciones, aunque el impacto agregado presenta una menor variabilidad (4-2). Este resultado también es coherente con lo visto en el capítulo 3 donde la capacidad de detección aumentaba en función del número de estaciones consideradas. Este poder de detección puede ser relevante en casos donde el impacto, relacionado con el evento o tratamiento bajo estudio, presente un tamaño de efecto muy pequeño. En el caso de la aplicación estudiada en este capítulo el efecto de la intervención presenta un efecto bastante pronunciado, lo cual lo hace detectable inclusive en estaciones individuales.

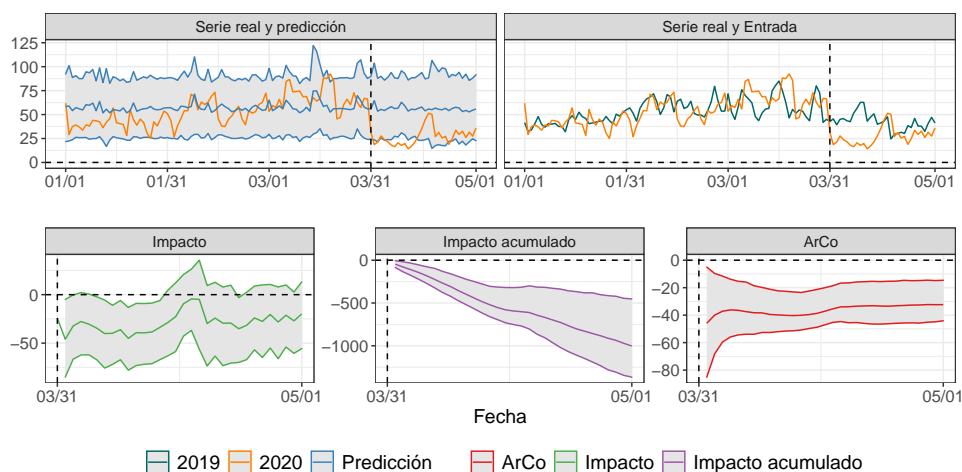
	Tiempo de corte	IC ArCo	IC Cumsum	Estación
11	2020-04-15	[-59.31, -27.72]	[-889.61, -415.82]	Global grupo 2
12	2020-04-15	[-57.92, -9.41]	[-868.81, -141.22]	Estación Tráfico Sur
13	2020-04-15	[-79.06, -17.78]	[-1185.97, -266.71]	Itagüí - Estación de Policía Los Gómez
14	2020-04-15	[-69.72, -12.16]	[-1045.74, -182.42]	Itagüí - I.E. Concejo Municipal de Itagüí
15	2020-04-15	[-83.08, -20.15]	[-1246.19, -302.22]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 4-2:** Resultados PM 10 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 2”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15.

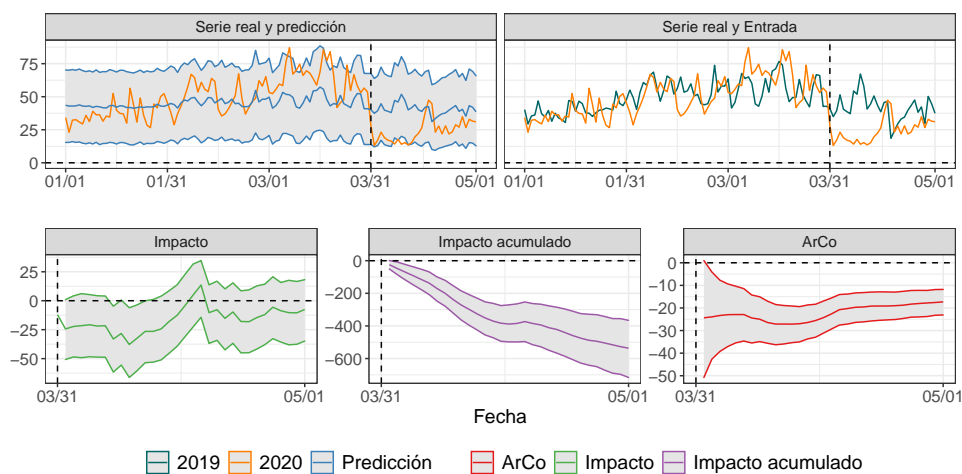
Para la aproximación mediante MCMC vemos que ambos conjuntos de estaciones detectan un impacto agregado en el mismo periodo de tiempo, lo cual quiere decir que ambos grupos de estaciones poseen comportamientos similares y que durante la cuarentena generada por el

Covid-19 hubo una disminución significativa en la concentración de PM 10 en el aire, al usar un intervalo de probabilidad posterior del 90%.

En las Figuras 4-6 y 4-7, se presentan los resultados para el modelo espacial (2-19) para los grupos de estaciones 1 y 2 respectivamente. Podemos ver que claramente se presenta un impacto significativo en el mismo periodo de tiempo en el cual se detecta un impacto para el modelo (2-1).



**Figura 4-6:** Resultados agregados PM 10 para el grupo de estaciones 1, utilizando la ecuación (1-6) para el modelo (2-19).

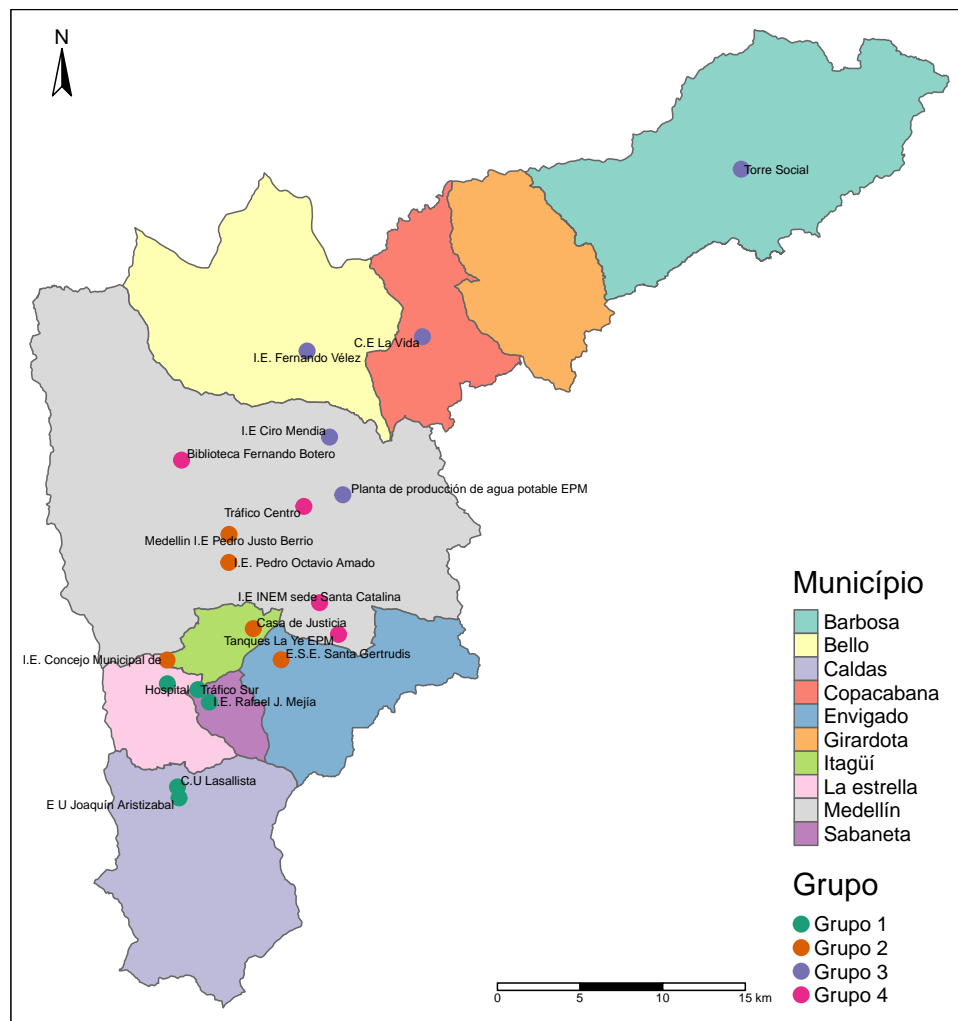


**Figura 4-7:** Resultados agregados PM 10 para el grupo de estaciones 2, utilizando la ecuación (1-6) para el modelo (2-19).

Las tablas de resultados para este modelo se presentan en los anexos 7-18 y 7-17.

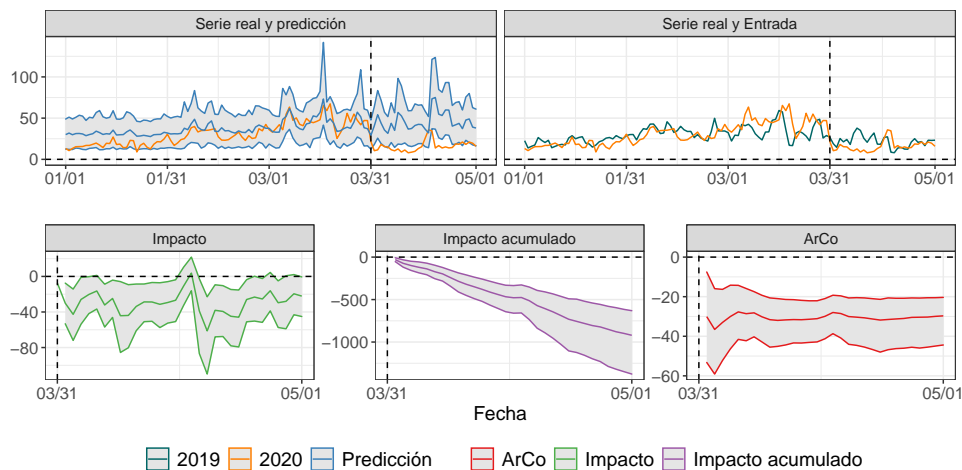
## 4.2. Resultados para PM 2.5

El material particulado PM 2.5 cuenta con un diámetro menor a  $2.5\mu m$  y sus fuentes de origen son las mismas de PM 10. A diferencia de PM 10 las partículas de PM 2.5 penetran las partes profundas de los pulmones, lo cual puede generar un mayor detrimento a la salud (Karimi et al., 2019). Para ajustar los resultados por MCMC realizamos el mismo proceso de clusterización realizado anteriormente para PM 10. El resultado del agrupamiento se presenta en la Figura 4-8.

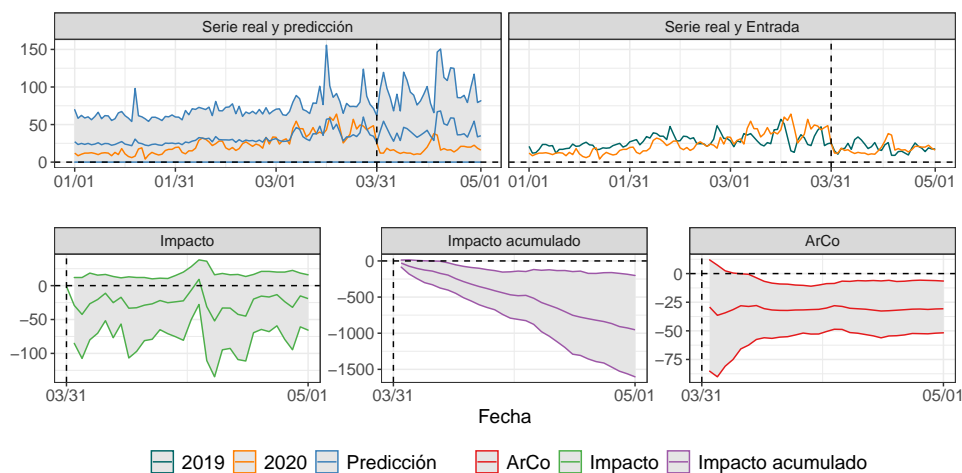


**Figura 4-8:** Grupos de estaciones para el contaminante PM 2.5 divididas utilizando al algoritmo de clúster desarrollado por Papenberg y Klau (2020).

En la Figura 4-9 se presentan los resultados agregados para el grupo 1 de estaciones. Observamos que los intervalos de probabilidad posterior para el impacto acumulado y el ArCo no contienen el valor cero, lo cual indica que existe un impacto significativo.



**Figura 4-9:** Resultados agregados PM 2.5 grupo de estaciones 1, utilizando la ecuación (1-6).



**Figura 4-10:** Resultados PM 2.5 para la estación La Estrella - Hospital.

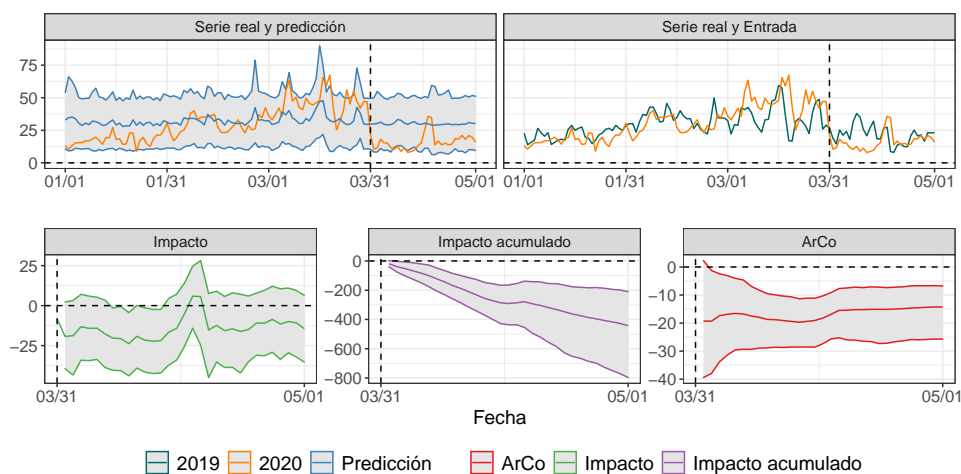
La Figura 4-10 presenta los resultados para la estación Estrella - Hospital, y en ella observamos que se presenta un impacto significativo. En la Tabla 4-3 vemos que se detecta un impacto a nivel global en el periodo de tiempo [2020-03-31, 2020-04-15], este es el mismo periodo en el cual se detecta un impacto para PM 10, lo cual es consistente, ya que ambos contaminantes comparten las mismas fuentes de origen. Vemos que el patrón de encontrar impactos a nivel global e individual observado en el contaminante PM 10 se presenta también en PM 2.5. Esto es consistente con los resultados de simulación presentados en el Capítulo 3.

	Tiempo de corte	IC ArCo	IC Cumsum	Estación
13	2020-04-15	[-42.86, -22.07]	[-642.97, -331.01]	Global grupo 1
14	2020-04-15	[-53.76, -12.79]	[-806.44, -191.84]	Caldas - Corporación Universitaria Lasallista
15	2020-04-15	[-52.04, -10.27]	[-780.64, -154.07]	Caldas - E U Joaquín Aristizabal
16	2020-04-15	[-53.04, -10.87]	[-795.60, -162.98]	Estación Tráfico Sur
17	2020-04-15	[-52.76, -10.02]	[-791.39, -150.29]	La Estrella - Hospital
18	2020-04-15	[-53.37, -13.07]	[-800.53, -196.07]	Sabaneta - I.E. Rafael J. Mejía

**Tabla 4-3:** Resultados PM 2.5 agregados e individuales con tiempo de corte 2020-04-15. Los resultados agregados aparecen con nombre de estación “Global grupo 1”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-31 y 2020-04-15.

En el Capítulo 7 se presenta las tablas de resultados 7-6, 7-7, 7-8 y 7-9 las cuales contienen los resultados para demás grupos de estaciones de PM 2.5. En general, para este contaminante siempre se detecta un impacto significativo en el periodo [2020-03-31, 2020-04-15].

En la Figura 4-11, se presentan los resultados para el modelo espacial (2-19). Las predicciones en este modelo están fuertemente influenciadas por la variable predictora (Entrada). Podemos ver que se presenta un impacto negativo muy marcado, en el mismo periodo de tiempo en el cual se detecta un impacto para PM 10 y PM 2.5 con el modelo (2-1).



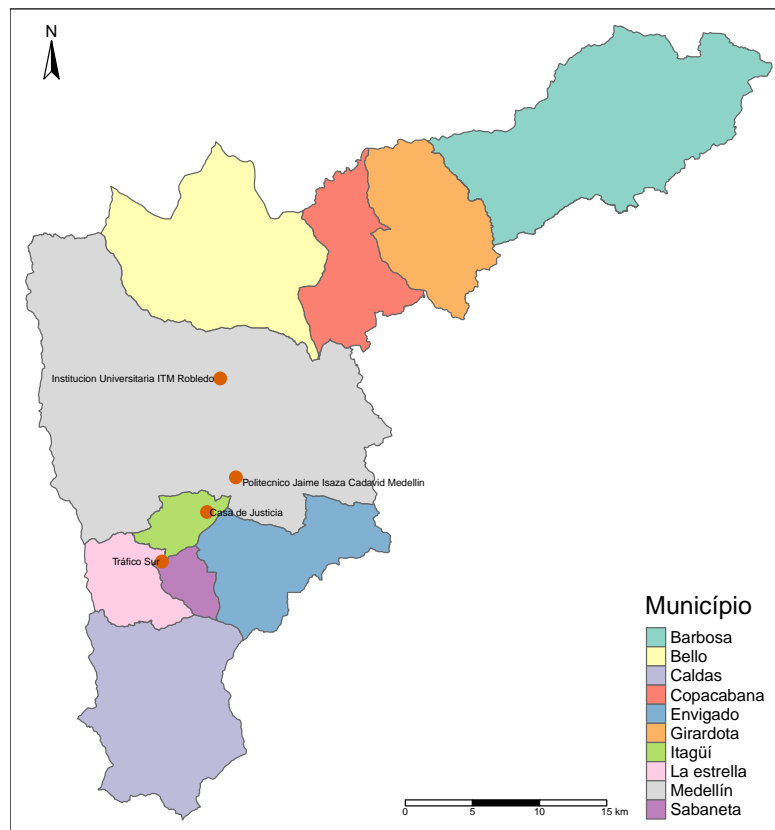
**Figura 4-11:** Resultados agregados PM 2.5 para el grupo de estaciones 1, utilizando la ecuación (1-6) para el modelo (2-19)

Los resultados para los demás grupos de estaciones se presentan el capítulo de anexos 7 en las Tablas 7-10, 7-11, 7-12 y 7-13.



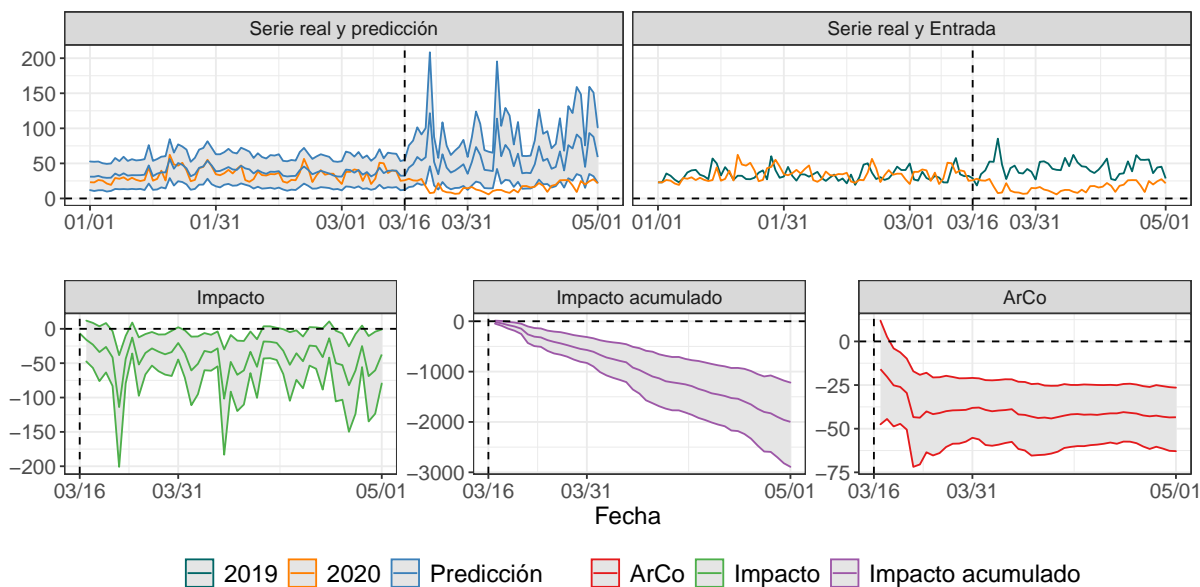
### 4.3. Resultados para NO<sub>x</sub>

La notación NO<sub>x</sub> incluye tanto el monóxido de nitrógeno(NO) como al dióxido de nitrógeno NO<sub>2</sub>. Una exposición anual promedio de 40µg/m<sup>3</sup> puede tener efectos largo plazo en la salud e incrementar la probabilidad de muerte por causas cardiopulmonares (Latz et al., 2009). La Figura 4-12 presenta la ubicación de las estaciones que monitorean el contaminante NO<sub>x</sub>. Estas son las mismas que hacen seguimiento a los contaminantes NO y NO<sub>2</sub>.



**Figura 4-12:** Estaciones para los contaminantes NO, NO<sub>2</sub> y NO<sub>x</sub>.

En la Figura 4-13 vemos los resultados agregados para las estaciones Tráfico Sur, Institución Universitaria ITM Robledo, Casa de Justicia Itagüi y Politécnico Colombiano Jaime Isaza Cadavid. Observamos que los intervalos de probabilidad posterior tanto para el impacto acumulado como para el ArCo están por debajo de cero, lo cual implica que se detecta un impacto. El impacto acumulado tiene una tendencia claramente negativa y al ArCo rápidamente se posiciona al rededor de un valor constante por debajo de cero.



**Figura 4-13:** Resultados agregados NO<sub>x</sub>, utilizando la ecuación (1-6).

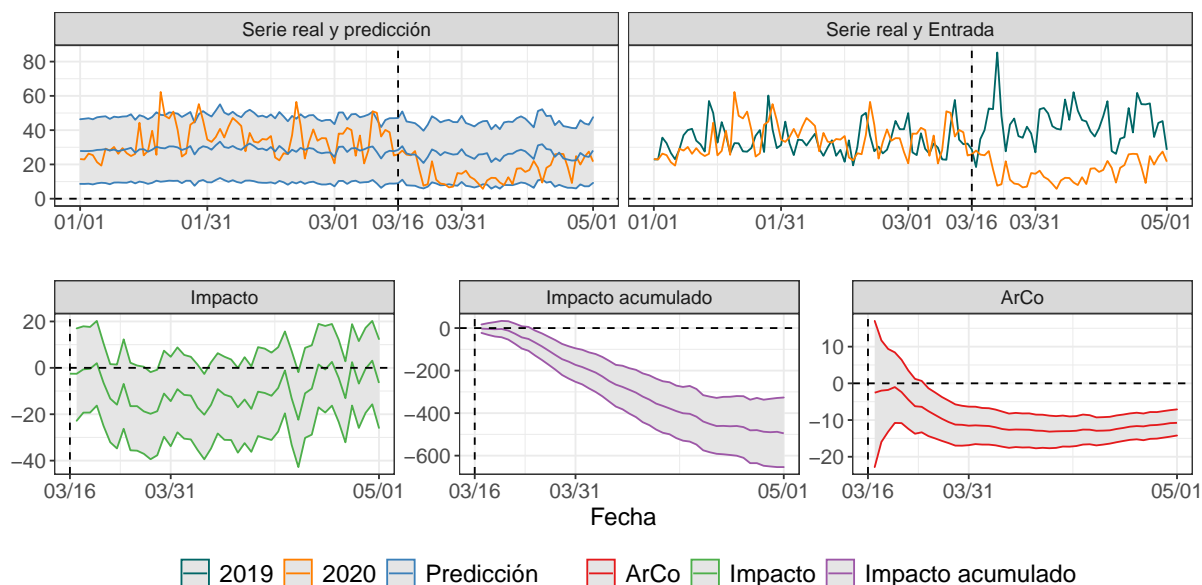
En la Tabla 4-4 vemos que se detecta un impacto para el contaminante NO<sub>x</sub>, tanto de manera global como de manera individual. Este resultado nos lleva concluir que la cuarentena generada por el Covid-19 genero una disminución en la concentración de NO<sub>x</sub> en el aire en el periodo de tiempo [2020-03-16, 2020-03-31]. Este mismo resultado se presenta en los periodos de tiempo [2020-03-31, 2020-04-15] y [2020-04-15, 2020-05-01], como se muestra en la tabla 7-3. Las Figuras 7-8, 7-9, 7-10 y 7-11 presentan los resultados individuales para las estaciones Estación Tráfico Sur, Institución Universitaria ITM Robledo, Casa de Justicia Itagüí, Politécnico Colombiano Jaime Isaza Cadavid. En todas ellas se presenta un impacto significativo y se mantiene la misma tendencia global presentada en la Figura 4-13.

	Tiempo de corte	IC ArCo	IC Cumsum	Estación
6	2020-03-31	[-55.21, -20.91]	[-828.15, -313.69]	Global
7	2020-03-31	[-77.67, -10.55]	[-1164.99, -158.19]	Estación Tráfico Sur
8	2020-03-31	[-43.37, -4.46]	[-650.58, -66.84]	Institucion Universitaria ITM Robledo
9	2020-03-31	[-53.45, -8.72]	[-801.70, -130.80]	Itagüí - Casa de Justicia Itagüí
10	2020-03-31	[-101.51, -10.53]	[-1522.67, -157.96]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 4-4:** Resultados NO<sub>x</sub> agregados e individuales con tiempo de corte 2020-03-31. Los resultados agregados aparecen con nombre de estación “Global”. Estos intervalos representan el impacto acumulado en el intervalo 2020-03-16 y 2020-03-31.

En el capítulo 7 se presenta las tablas de resultados 7-4, 7-5 y 7-3 para NO, NO<sub>2</sub> y NO<sub>x</sub> respectivamente. Para cada uno de estos contaminantes se detecta un impacto significativo tanto a nivel individual como global, lo cual es consistente con los resultados presentados para NO<sub>x</sub> ya que este cubre tanto al monóxido como dióxido de nitrógeno.

En la Figura 4-14 tenemos los resultados del modelo espacial (2-19), observamos un resultado similar al obtenido con PM 10 y PM 2.5, donde el resultado de la predicción está fuertemente marcada por la variable predictora (Entrada), pero con una pobre descripción de la variabilidad del proceso. A pesar del impacto ser significativo, este resultado no es confiable por las limitaciones en el ajuste ya mencionadas.



**Figura 4-14:** Resultados agregados  $\text{NO}_x$ , utilizando la ecuación (1-6) para el modelo (2-19).

Los resultados del modelo (2-19) para los contaminantes NO,  $\text{NO}_2$  y  $\text{NO}_x$  se presenta en las tablas 7-14, 7-15 y 7-16.

## 4.4. Discusión

En el capítulo 3 se realiza un estudio de simulación para el modelo (2-1). En este sentido vale resaltar la relevancia del método propuesto en esta tesis al ser contrastado con el método empleado por Mendez-Espinosa et al. (2020). En este último, los autores además de sólo emplear dos estaciones en sus análisis de evaluación impacto, realizan los análisis de forma independiente por estación. Por el contrario, en el método empleado en esta tesis se modela de manera conjunta la información para un número diferente de estaciones (para  $q = 4, 5, 6$ ) y contaminantes como se presentará en este capítulo.

Realizar un análisis univariado para este problema es equivalente a asumir que no existen correlaciones entre estaciones estudiadas. Debido a que el estudio se realizó dentro de una misma área metropolitana conformada por un conjunto de ciudades adyacentes y que estas comparten condiciones de ventilación por la estructura de valle en la cual están ubicadas,

sería extremadamente difícil justificar su independencia.

En su informe anual de calidad del aire el Área metropolitana del Valle de Aburrá (2020) muestra una reducción desde el mes de abril en los contaminantes de PM 10 y PM 2.5, esto no se observa en el caso  $\text{NO}_2$ , en el informe no se presenta en ciclo de contaminación anual para  $\text{NO}_x$  o  $\text{NO}$ . Es importante aclarar que los resultados presentados por el Área Metropolitana miden únicamente la concentración, no el impacto acumulado o el ArCo, tampoco realizan evaluación de impacto como se hace en este trabajo. En esto mismo informe se advierte también que en el mes de marzo hubo número de incendios ambientales significativamente superior a los años anteriores.

## 5 Paquete DynamicIpnact

El paquete **dynamicIpnact** disponible en <https://github.com/caaperezag/dynamicIpnactR>, es una implementación del método propuesto en esta tesis. La funcionalidad de este software es determinar si existe un impacto significado en un sistema después de que se presenta un evento de interés con fecha  $T_0$  de inicio conocida. El paquete **CausalImpact** construido por Brodersen et al. (2015) presenta una funcionalidad similar basada en modelos Bayesianos, pero al momento de escribir este documento dicho paquete aún no ofrece soporte para modelos vectoriales o matriciales como los empleados en esta tesis. Actualmente, el paquete **dynamicIpnact** no se encuentra disponible en CRAN, por lo cual debe ser instalado desde GitHub, y el proceso de instalación puede realizarse con una herramienta como **remotes**.

```
install.packages("remotes")
remotes::install_github("https://github.com/caaperezag/dynamicIpnactR")
```

**Figura 5-1:** Instalacion del paquete **dynamicIpnact**

El paquete **dynamicIpnact** posee funcionalidad para ajustar modelos mediante la aproximación posterior utilizando en MCMC y de manera conjugada utilizando las ecuaciones de recurrencia (2-4), (2-5) y (2-6). Las funciones presentes en el paquete y su descripción se presentan en la tabla 5-1. Opciones adicionales para cada una de las funciones presentadas, pueden ser exploradas mediante el comando **help**.

Función	Descripción
<b>StanModelVector</b>	Es la función utilizada para crear modelos en los cuales $\mathbf{Y}_t$ es un vector con $q$ elementos (estaciones de monitoreo en nuestro caso). La distribución posterior para este modelo es aproximada mediante MCMC a través del software stan.
<b>StanModelMatrix</b>	Es la función utilizada para crear modelos en los cuales $\mathbf{Y}_t$ es un matiz con dimensiones $r \times q$ (contaminantes $\times$ estaciones). Esta modelo utiliza stan para aproximar la distribución posterior.
<b>SpatialModel</b>	Implementa el modelo espacial presentado en la sección 2.2 en stan.
<b>ConjugateModel</b>	Implementa el ajuste de manera conjugada utilizando las ecuaciones de recurrencia (2-4), (2-5) y (2-6), no depende de stan.

**Tabla 5-1:** Funciones disponible para el usuario en el paquete **dynamicIpnact**.

La Figura 5-2 presenta un ejemplo de uso para el modelo en su versión vectorial. El parámetro `event_initial` define el inicio de la intervención, `predefined_cov_matrix_type` controla la definición de  $\Sigma$ : cuando toma el valor `mcmc` se aproxima su distribución posterior mediante MCMC, cuando `predefined_cov_matrix_type` toma los valores `identity` o `ml`,  $\Sigma$  se hace un término fijo que toma el valor de la identidad o su estimación mediante de máxima verosimilitud respectivamente. Cuando se utiliza máxima verosimilitud únicamente se realiza la estimación utilizando datos previos al inicio de la intervención.

```
library(tidyverse)
library(dynamicimpact)
m_model <- StanModelVector$new(
  X_data = X_vector,
  Y_data = Y_vector,
  variables_names = c("Est 1", "Est 2"),
  confidence_level=0.9,
  n_simul = 10000,
  n_chains = 4,
  n_cores = 4,
  thin = 1,
  predefined_cov_matrix_type = "mcmc",
  event_initial = 97
)

m_model$fit()
m_model$plot()
```

**Figura 5-2:** Ejemplo de uso del paquete DynamicImpact 1.

Para el caso del modelo vectorial, `X_vector` y `Y_vector` son matrices de dimensiones  $T \times q$ , donde  $T$  es el horizonte de tiempo del estudio y  $q$  es el número de estaciones, este modelo es usado para el caso en el cual  $\mathbf{Y}_t$  es un vector con  $q$  elementos.

El método `fit` realiza la aproximación por MCMC, y el método `plot` construye gráficos de impacto similares a los presentados en el capítulo 4. La figura 5-3 presenta un ejemplo de uso para el modelo matricial donde  $\mathbf{Y}_t$  es una matriz con  $r$  contaminates y  $q$  estaciones.

```
library(dynamicimpact)
m_model <- StanModelMatrix$new(
  X_data = X_matrix,
  Y_data = Y_matrix,
  variables_names = c("Est 1", "Est 2"),
  confidence_level=0.9,
```

```

n_simul = 1000,
n_chains = 4,
n_cores = 4,
thin = 1,
predefined_cov_matrix_type = "mcmc",
event_initial = 97
)

m_model$fit()
m_model$plot()

```

**Figura 5-3:** Ejemplo de uso del paquete DynamicImpact 2.

La Figura 5-4 se presenta el ejemplo de uso para el modelo conjugado, en este caso el modelo es ajustado utilizando las ecuaciones recurrentes (2-4), (2-4) y (2-4). Para este modelo el parámetro `n_simul` representa el número de iteraciones utilizadas en el algoritmo FSTS (1).

```

library(dynamicimpact)
m_model <- ConjugateModel$new(
  X_data = X_matrix,
  Y_data = Y_matrix,
  vector_name = "conjugate model",
  variables_names = c("Est 1", "Est 2"),
  confidence_level=0.9,
  n_simul = 10000,
  event_initial = 97
)

m_model$fit()
m_model$plot()

```

**Figura 5-4:** Ejemplo de uso del paquete DynamicImpact 3.

La Figura 5-5 presenta el ejemplo de uso para el modelo espacial, para este caso es necesario usar el parametro `coordinates` para especificar la ubicacion de cada unas las estaciones en coordenadas planares. El lector interesado en profundaizar sobre el sistema de coordenadas puede referirse a IGAC (2004).

```

library(dynamicimpact)
m_model <- SpatialModel$new(
  X_data = X_vector,
  Y_data = Y_vector,

```

```
vector_name = "example_1",
variables_names = c("Estación 1", "Estación 2"),
confidence_level=CONFIDENCE_LEVEL,
confidence_level=0.9,
n_simul = 1000,
n_chains = 4,
n_cores = 4,
thin = 1,
predefined_cov_matrix_type = "mcmc",
coordinates = coordinates_matrix
)

m_model$fit()
m_model$plot()
```

**Figura 5-5:** Ejemplo de uso del paquete DynamicImpact 2.

Aunque el paquete **dynamicimpact** fue construido modelar estaciones de calidad del aire, cada una con cantidad definida de contaminantes, el software puede ser usado en cualquier situación que pueda ser representada utilizando el modelo (2-1) presentado en el capítulo 2. Al momento de escribir este documento el paquete **dynamicimpact** únicamente está disponible a través de github. Esperamos subir este paquete al CRAN una vez se tenga la suficiente retroalimentación por parte de los usuarios.

El paquete se publica bajo una licencia MIT, la cual permite el uso comercial sin necesidad de compartir las modificaciones al código fuente (Mamedov, 2018).



# 6 Conclusiones y recomendaciones

## 6.1. Conclusiones

En este trabajo propusimos una metodología basada en modelos Dinámicos Bayesianos para detección de impacto en series de tiempo matriciales y vectoriales. Realizamos su implementación en el paquete de R **dynamicimpact**, presentado en el capítulo 5. Evaluamos las capacidades de la metodología propuesta para el modelo (2-1) en el capítulo 3 y presentamos ejemplos de detección para el modelo (2-19) utilizando el impacto acumulado definido en (1-6).

También propusimos el modelo espacial (2-19) el cual extiende el modelo (2-1) agregando el elemento  $\lambda\mu_t^{\text{espacial}}$  en la ecuación de observación de  $\mathbf{Y}_t$ . El modelo (2-19) nos permite estudiar un conjunto de  $q$  puntos fijos (estaciones) a través del tiempo, cada uno con  $r$  observaciones (contaminantes) en un tiempo  $t$ . Esto difiere del modelo (2-16) propuesto por Stroud et al. (2001) en el cual en un tiempo  $t$  solo se tienen una ubicación.

Finalmente, en el capítulo 4 utilizando el paquete **dynamicimpact** realizamos el análisis de impacto en la concentración de contaminantes en la calidad del aire. En general, observamos que se presenta una reducción en los niveles de contaminación en el aire. Para los contaminantes  $\text{NO}_2$ ,  $\text{NO}$ , y  $\text{NO}_x$  esta disminución comienza en el periodo [2020-03-16, 2020-03-31]. Para el caso de  $\text{PM}_{10}$  y  $\text{PM}_{2.5}$  se presenta un impacto significativo desde el periodo [2020-03-31, 2020-04-15]. El alcance de este trabajo no permite identificar si la reducción detectada para los contaminantes  $\text{NO}_2$ ,  $\text{NO}$ ,  $\text{NO}_x$ ,  $\text{PM}_{2.5}$  y  $\text{PM}_{10}$  es significativa desde el punto de vista médico.

## 6.2. Recomendaciones

Para los casos en los cuales  $\mathbf{Y}_t$  es vector de  $q$  elementos, recomendamos utilizar el modelo vectorial implementado en el paquete **dynamicimpact** en la función **StanModelVector** (5-2), ya que esta es más rápida para realizar estimaciones de las distribuciones posteriores.

El sistema de coordenadas utilizadas en la función **SpatialModel** (5-5) debe ser planar, si se utiliza otro tipo de coordenadas (como las geodésicas) las estimaciones de las distribuciones posteriores de para el modelo (2-19) no tendrían sentido. Usar coordenadas geodésicas(latitud

y longitud) sería equivalente a utilizar la distancia euclidiana, para estimar la distancia entre dos vectores cuyos elementos son ángulos.

Si el computador en el cual ejecuta los modelos tiene la suficiente cantidad núcleos, recomendamos hacer el parámetro `n_cores` igual al parámetro `n_chain`, en las funciones `StanModelVector`, `StanModelMatrix` y `SpatialModel`.

Si el equipo en el cual ejecuta los modelos tiene un procesador AMD(Advanced Micro Devices) recomendamos la instalación de `OpenBLAS` disponible en <https://www.openblas.net/>.

En caso de que el computador en el cual se ejecutan los modelos tenga un procesador Intel(integrated and electronics) recomendamos compilar R con soporte para MKL(Math Kernel Library). Las instrucciones para este proceso se encuentran en <https://www.intel.com/content/www/us/en/developer/articles/technical/using-onemkl-with-r.html>. Es importante tener en cuenta que utilizar MKL en un procesador AMD puede degradar el rendimiento.

### 6.3. Limitaciones y trabajos futuros

En el capítulo 3 se mencionan que existen limitaciones para garantizar que las matrices  $\Sigma$ ,  $\mathbf{W}$  y  $\mathbf{V}$  sean semidefinidas positivas. este problema puede solucionarse sumando un valor pequeño  $\delta$  a la diagonal de cada una de las matrices de covarianza, como se sugiere en Rousseeuw y Molenberghs (1993).

Otra limitación es el tiempo de cómputo que puede tomar realizar las simulaciones presentadas en el capítulo 3, o estimar la distribución de probabilidad posterior en el capítulo 4. Limitándose a `stan` existen algunas opciones no implementadas en el paquete `dynamicimpact` que pueden incrementar la velocidad de cómputo:

**Opción 1**, paralelizar al interior de la cadena: Desde su versión 2.23 `stan` implementa el algoritmo MapReduce (Dean y Ghemawat, 2004) el cual permite utilizar múltiples hilos al interior de cada cadena, lo cual, aunque un genera gasto mayor de memoria, puede incrementar la velocidad de ajuste del modelo.

**Opción 2**, utilizar el soporte de `stan` para GPU: Desde su versión 2.26.1 `stan` adquiere soporte para realizar cálculos utilizando GPU(Graphics processing unit) mediante `OpenCL` (Češnovar et al., 2019), aunque las GPU con la suficiente memoria y procesamiento para este modelo en específico son costosas, probablemente hubiese sido posible con ellas solucionar la totalidad de problemas de cómputo relacionados con el tiempo de ejecución.

**Opción 3**, utilizar Bayes variacional: **stan** implementa de manera automática inferencia variacional (Kucukelbir et al., 2015), este método para estimar la distribución posterior es mucho más rápido y requiere una menor cantidad de memoria, en futuros trabajos queremos evaluar la calidad de las sus estimaciones para la evaluación de impacto.

Tomando como base la metodología y el software construidos en este trabajo, se podrían explorar las siguientes posibilidades en futuros proyectos:

- Analizar si las reducciones en la concentración de PM 2.5 y PM 10 son relevantes desde el punto visto médico.
- Extender el trabajo realizado para evaluar los efectos de la calidad del aire en la salud y en los costos incurridos por las EPS.
- Utilizar el paquete en estudios no relacionados con la calidad de Aire.
- Expandir las capacidades del paquete **dynamicimpact** para estimar el impacto de múltiples intervenciones con diferentes tiempos de inicio.
- Estimación automática del tiempo de inicio de la intervención.
- Estimación automática del tiempo en el cual la intervención deja de tener un impacto.
- Expandir las combinaciones de parámetros estudiados presentadas en el capítulo 3.1.
- Agregar soporte para distribuciones diferentes a la normal en los modelos estimados mediante MCMC.
- Estimación de impactos que generen únicamente cambios en la variabilidad.
- Integrar el paquete **dynamicimpact** con el paquete **tidymodels**, para facilitar y estandarizar su uso.
- Implementar soporte para validación cruzada en el paquete **dynamicimpact**.
- Evaluar la estimación de impacto utilizando inferencia variacional.

# 7 Anexos

En esta sesión se presentan gráficos y resultados adicionales relacionados con los análisis y conclusiones presentados en este documento.

## 7.1. Resultados adicionales

### 7.1.1. PM 10

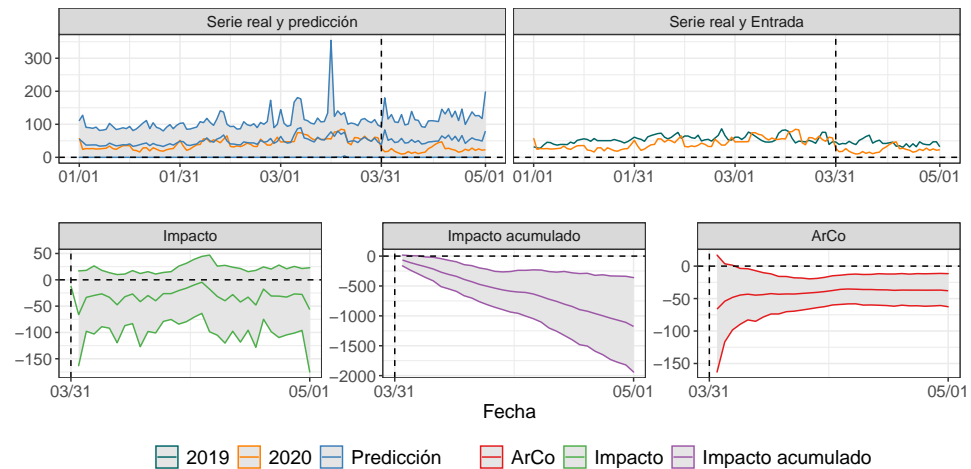
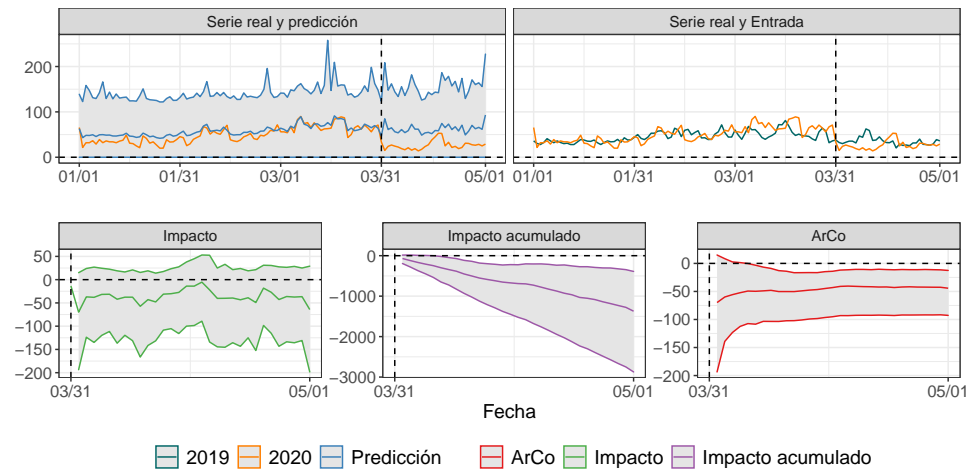
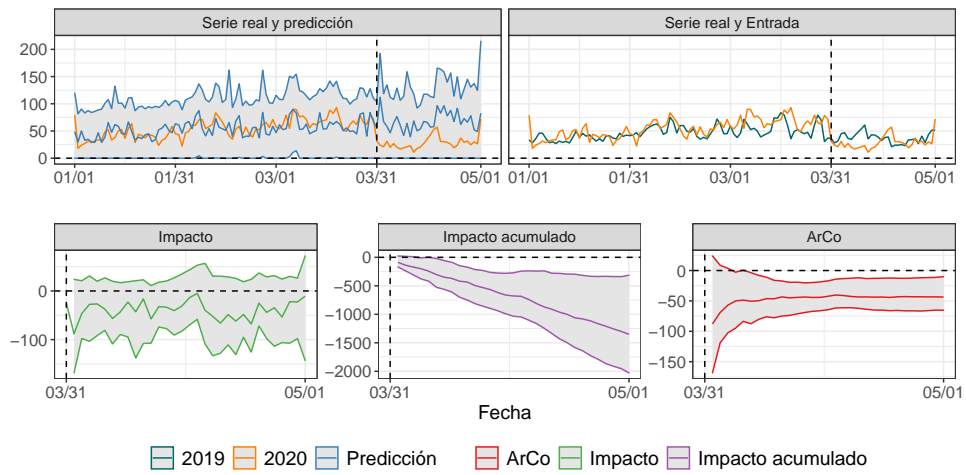


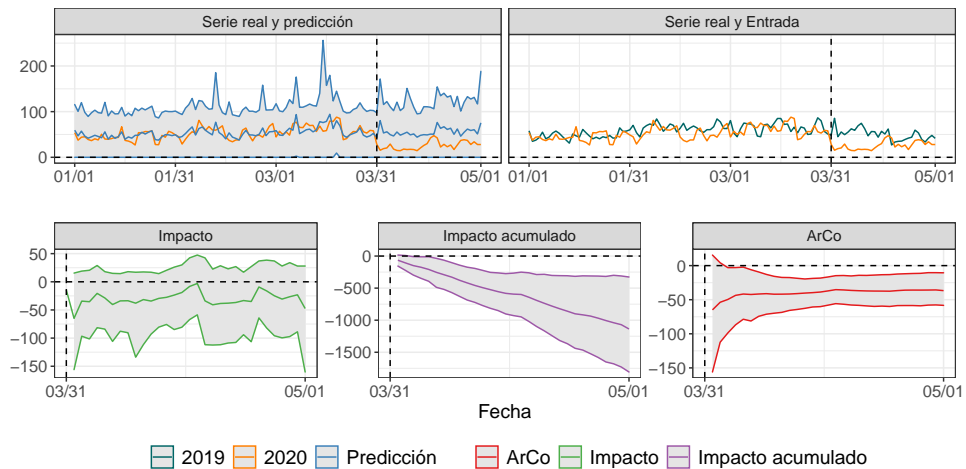
Figura 7-1: Resultados PM 10 para la estación Universidad San Buenaventura.



**Figura 7-2:** Resultados PM 10 para la estación Institucion Universitaria ITM Robledo.



**Figura 7-3:** Resultados PM 10 para la estación Institución Educativa Colombia.



**Figura 7-4:** Resultados PM 10 para la estación Exito San Antonio.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-28.18, 9.91]	[-422.73, 148.65]	Global
8	[2020-03-16, 2020-03-31]	[-59.94, 24.12]	[-899.09, 361.81]	Estación Tráfico Centro
9	[2020-03-16, 2020-03-31]	[-26.11, 19.55]	[-391.65, 293.26]	Exito San Antonio - Medellin
10	[2020-03-16, 2020-03-31]	[-22.19, 39.27]	[-332.78, 589.12]	Institucion Educativa Colombia. Girardota
11	[2020-03-16, 2020-03-31]	[-51.61, 27.90]	[-774.10, 418.56]	Institucion Universitaria ITM Robledo
12	[2020-03-16, 2020-03-31]	[-59.42, 22.71]	[-891.36, 340.66]	Universidad San Buenaventura
13	[2020-03-31, 2020-04-15]	[-60.60, -30.61]	[-908.96, -459.09]	Global grupo 1
14	[2020-03-31, 2020-04-15]	[-77.44, -22.20]	[-1161.65, -332.95]	Estación Tráfico Centro
15	[2020-03-31, 2020-04-15]	[-60.18, -18.28]	[-902.67, -274.23]	Exito San Antonio - Medellin
16	[2020-03-31, 2020-04-15]	[-66.13, -18.69]	[-992.00, -280.40]	Institucion Educativa Colombia. Girardota
17	[2020-03-31, 2020-04-15]	[-96.51, -14.72]	[-1447.65, -220.79]	Institucion Universitaria ITM Robledo
18	[2020-03-31, 2020-04-15]	[-62.36, -17.24]	[-935.46, -258.61]	Universidad San Buenaventura
19	[2020-04-15, 2020-05-01]	[-58.73, -20.45]	[-939.61, -327.16]	Global
20	[2020-04-15, 2020-05-01]	[-88.00, -1.35]	[-1408.06, -21.55]	Estación Tráfico Centro
21	[2020-04-15, 2020-05-01]	[-64.91, 2.55]	[-1038.52, 40.80]	Exito San Antonio - Medellin
22	[2020-04-15, 2020-05-01]	[-69.53, 9.00]	[-1112.48, 143.95]	Institucion Educativa Colombia. Girardota
23	[2020-04-15, 2020-05-01]	[-93.45, -3.51]	[-1495.22, -56.20]	Institucion Universitaria ITM Robledo
24	[2020-04-15, 2020-05-01]	[-69.83, -1.68]	[-1117.33, -26.91]	Universidad San Buenaventura

**Tabla 7-1:** Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global” grupo 1.

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-23.14, 6.65]	[-347.13, 99.78]	Global
7	[2020-03-16, 2020-03-31]	[-31.83, 13.42]	[-477.38, 201.26]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-39.17, 18.94]	[-587.61, 284.12]	Itagüí - Estación de Policía Los Gómez
9	[2020-03-16, 2020-03-31]	[-32.05, 22.95]	[-480.73, 344.19]	Itagüí - I.E. Concejo Municipal de Itagüí
10	[2020-03-16, 2020-03-31]	[-39.77, 21.24]	[-596.48, 318.54]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-59.31, -27.72]	[-889.61, -415.82]	Global
12	[2020-03-31, 2020-04-15]	[-57.92, -9.41]	[-868.81, -141.22]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-79.06, -17.78]	[-1185.97, -266.71]	Itagüí - Estación de Policía Los Gómez
14	[2020-03-31, 2020-04-15]	[-69.72, -12.16]	[-1045.74, -182.42]	Itagüí - I.E. Concejo Municipal de Itagüí
15	[2020-03-31, 2020-04-15]	[-83.08, -20.15]	[-1246.19, -302.22]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-61.60, -14.07]	[-985.64, -225.04]	Global
17	[2020-04-15, 2020-05-01]	[-61.52, 12.68]	[-984.27, 202.93]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-89.13, 0.38]	[-1426.01, 6.05]	Itagüí - Estación de Policía Los Gómez
19	[2020-04-15, 2020-05-01]	[-77.87, 4.15]	[-1245.95, 66.46]	Itagüí - I.E. Concejo Municipal de Itagüí
20	[2020-04-15, 2020-05-01]	[-96.17, -4.64]	[-1538.77, -74.17]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-2:** Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 2.

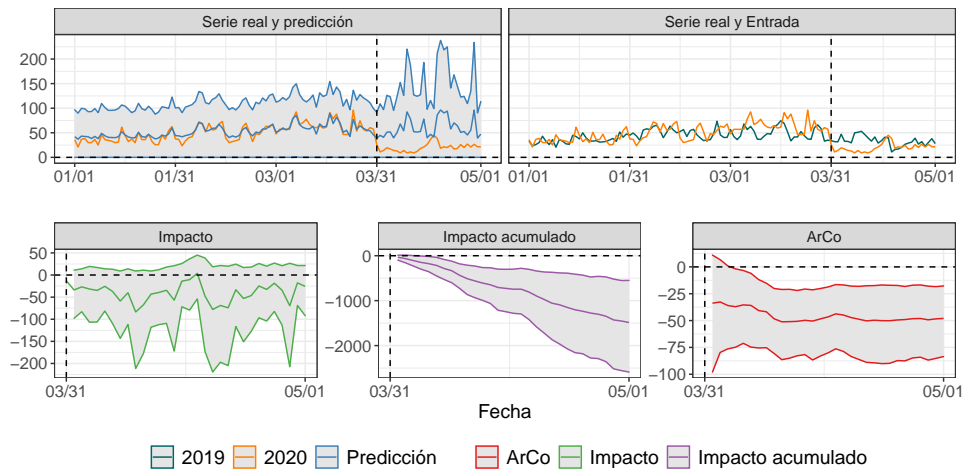


Figura 7-5: Resultados PM 10 para la estación Politecnico Colombiano Jaime Isaza Cadavid.

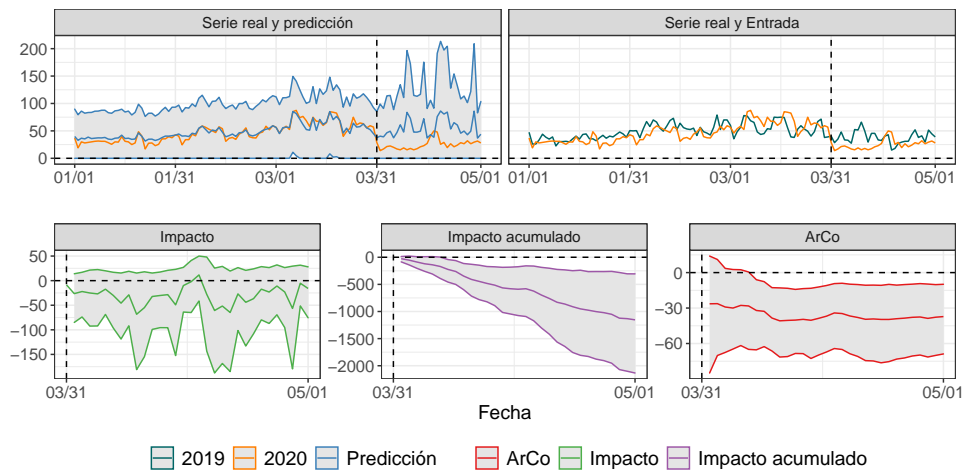
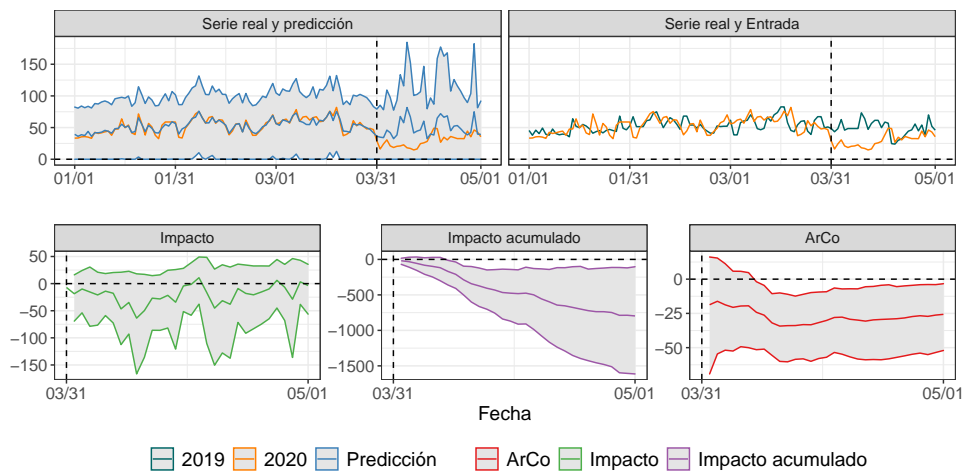
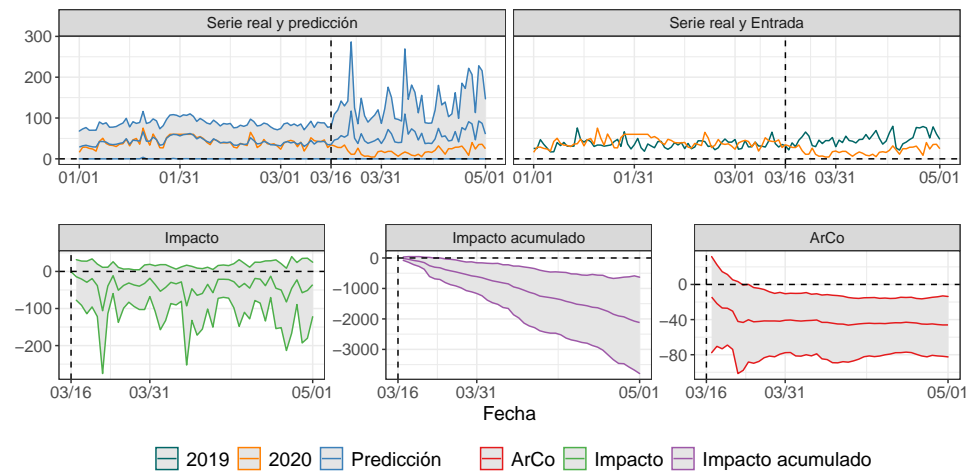


Figura 7-6: Resultados PM 10 para la estación I.E. Concejo Municipal de Itagüí.

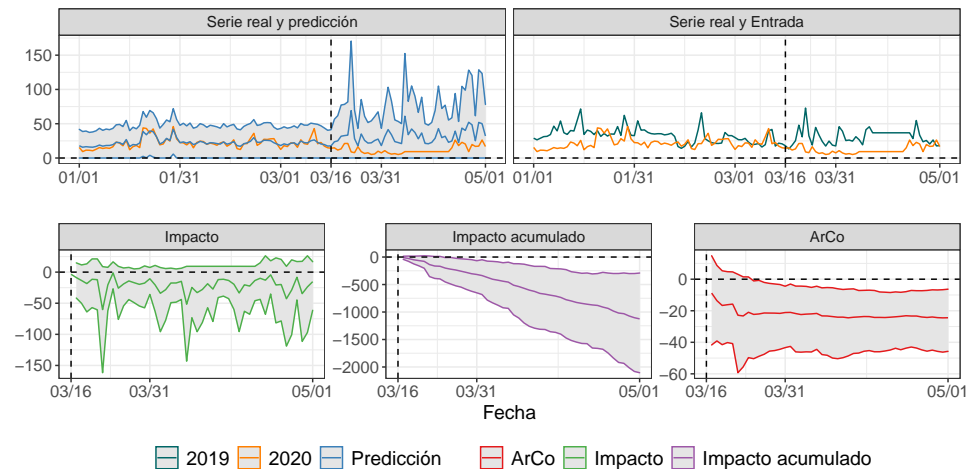


**Figura 7-7:** Resultados PM 10 para la estación estación Tráfico Sur.

### 7.1.2. NO<sub>x</sub>



**Figura 7-8:** Resultados NO<sub>x</sub> para la estación Tráfico Sur.



**Figura 7-9:** Resultados NO<sub>x</sub> para la estación ITM Robledo.



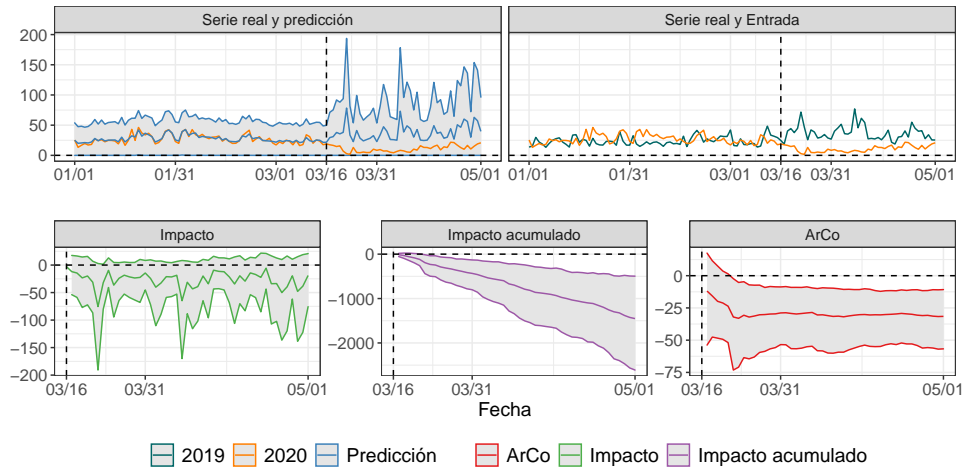


Figura 7-10: Resultados NO<sub>x</sub> Casa de Justicia Itagüí.

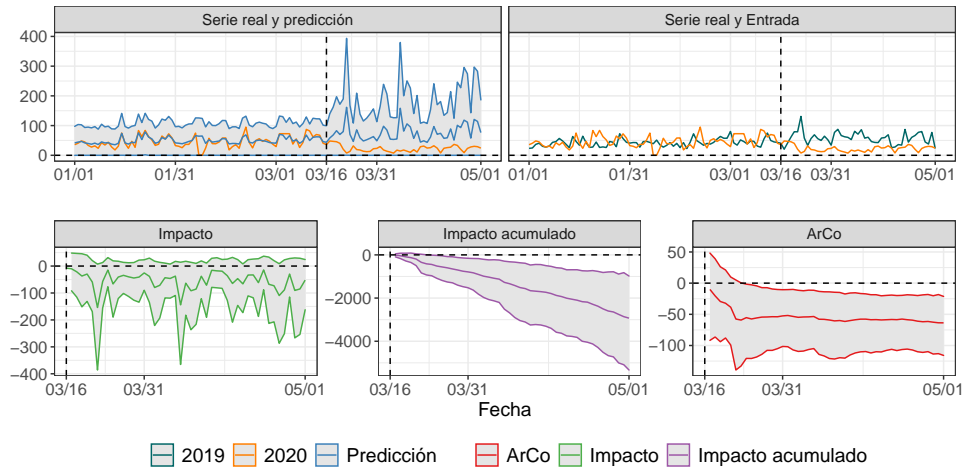
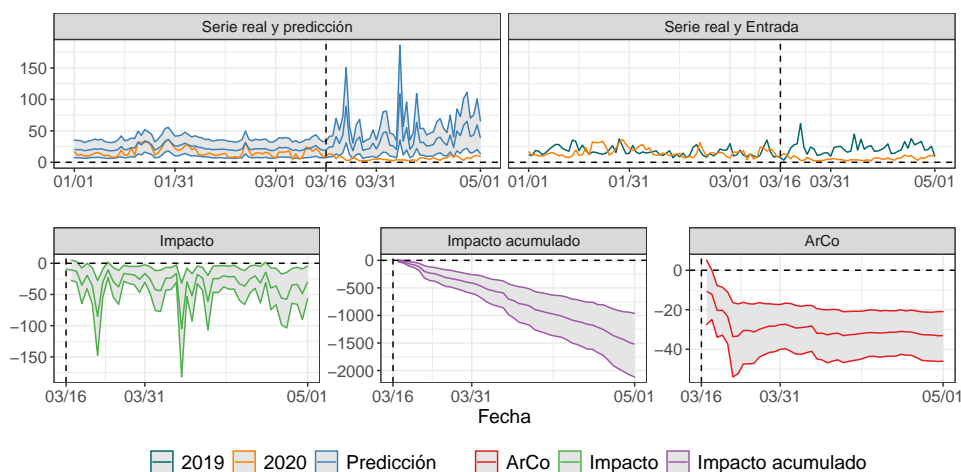


Figura 7-11: Resultados NO<sub>x</sub> Politecnico Colombiano Jaime Isaza Cadavid.

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-55.21, -20.91]	[-828.15, -313.69]	Global
7	[2020-03-16, 2020-03-31]	[-77.67, -10.55]	[-1164.99, -158.19]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-43.37, -4.46]	[-650.58, -66.84]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[-53.45, -8.72]	[-801.70, -130.80]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-101.51, -10.53]	[-1522.67, -157.96]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-61.25, -24.50]	[-918.78, -367.45]	Global
12	[2020-03-31, 2020-04-15]	[-85.29, -14.50]	[-1279.31, -217.44]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-46.46, -7.19]	[-696.93, -107.90]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[-57.17, -9.89]	[-857.49, -148.34]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[-109.81, -14.74]	[-1647.08, -221.05]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-59.61, -19.86]	[-953.73, -317.81]	Global
17	[2020-04-15, 2020-05-01]	[-83.33, -3.53]	[-1333.27, -56.41]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-46.20, -1.56]	[-739.24, -25.02]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[-56.28, -4.78]	[-900.54, -76.54]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[-114.92, -11.75]	[-1838.68, -187.96]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-3:** Resultados  $\text{NO}_x$  agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”.

### 7.1.3. NO

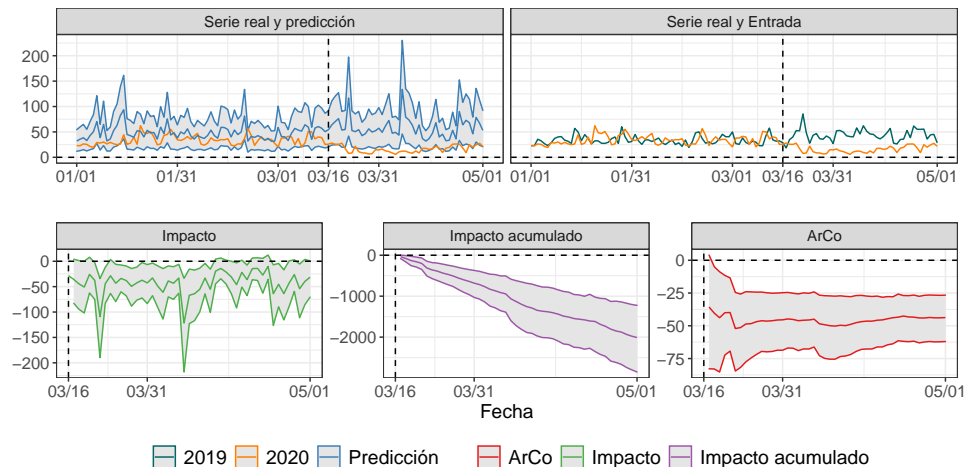


**Figura 7-12:** Resultados agregados  $\text{NO}$ , utilizando la ecuación (1-6)

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-39.99, -17.33]	[-599.78, -259.90]	Global
7	[2020-03-16, 2020-03-31]	[-61.00, -13.21]	[-915.00, -198.09]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-22.44, -1.67]	[-336.63, -24.99]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[-33.84, -5.15]	[-507.64, -77.26]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-73.37, -15.25]	[-1100.49, -228.71]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-45.43, -21.07]	[-681.49, -316.10]	Global
12	[2020-03-31, 2020-04-15]	[-69.35, -17.82]	[-1040.21, -267.32]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-27.66, -6.04]	[-414.87, -90.56]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[-37.01, -6.42]	[-555.22, -96.33]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[-81.93, -19.71]	[-1228.97, -295.70]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-43.63, -18.23]	[-698.12, -291.76]	Global
17	[2020-04-15, 2020-05-01]	[-66.72, -11.40]	[-1067.46, -182.44]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-27.08, -3.48]	[-433.21, -55.71]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[-37.74, -5.63]	[-603.83, -90.06]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[-80.86, -15.83]	[-1293.72, -253.27]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-4:** Resultados NO agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”.

#### 7.1.4. NO<sub>2</sub>

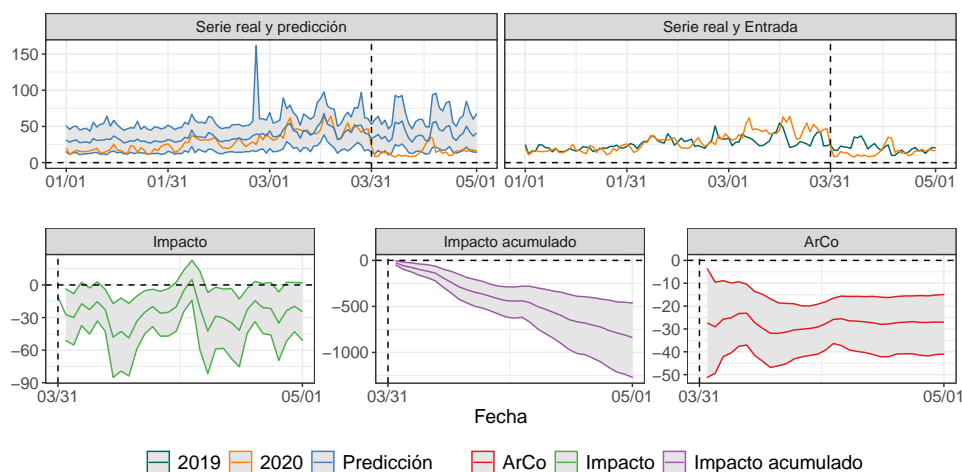


**Figura 7-13:** Resultados agregados NO<sub>2</sub>, utilizando la ecuación (1-6)

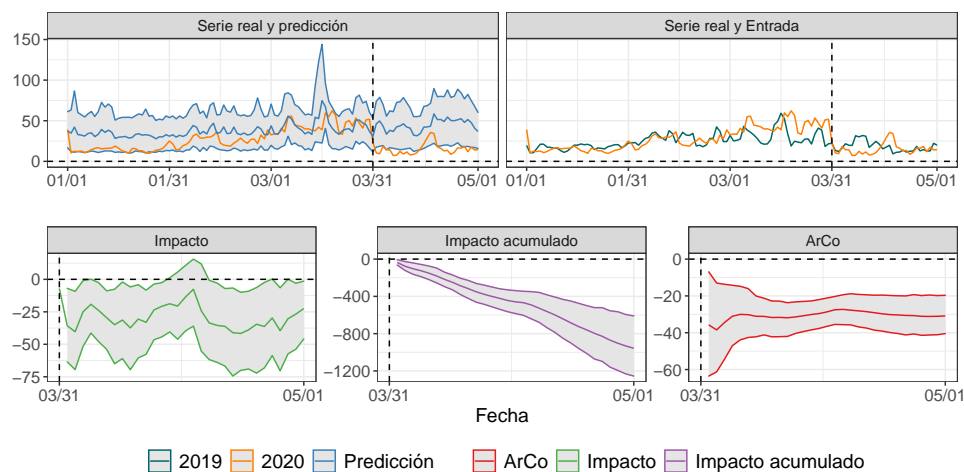
	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-68.63, -24.77]	[-1029.50, -371.52]	Global
7	[2020-03-16, 2020-03-31]	[-89.84, -9.58]	[-1347.55, -143.67]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-48.40, -3.97]	[-726.06, -59.58]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[-64.77, -9.79]	[-971.59, -146.91]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-135.33, -12.93]	[-2029.97, -193.98]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-65.87, -25.62]	[-988.11, -384.31]	Global
12	[2020-03-31, 2020-04-15]	[-92.61, -14.63]	[-1389.12, -219.49]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-50.70, -7.19]	[-760.52, -107.87]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[-60.99, -9.35]	[-914.90, -140.32]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[-122.29, -17.84]	[-1834.42, -267.59]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-65.84, -17.65]	[-1053.50, -282.34]	Global
17	[2020-04-15, 2020-05-01]	[-97.40, -2.29]	[-1558.38, -36.63]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-51.37, 0.69]	[-821.89, 11.10]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[-62.70, -2.92]	[-1003.19, -46.78]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[-130.74, -8.41]	[-2091.87, -134.61]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-5:** Resultados NO<sub>2</sub> agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”.

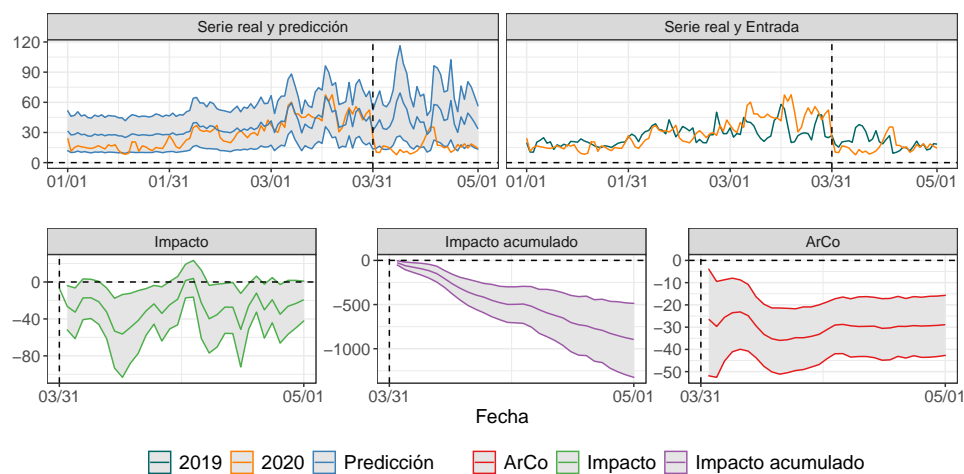
### 7.1.5. PM 2.5



**Figura 7-14:** Resultados agregados PM 2.5 grupo de estaciones 2, utilizando la ecuación (1-6)



**Figura 7-15:** Resultados agregados PM 2.5 grupo de estaciones 3, utilizando la ecuación (1-6)



**Figura 7-16:** Resultados agregados PM 2.5 grupo de estaciones 4, utilizando la ecuación (1-6)

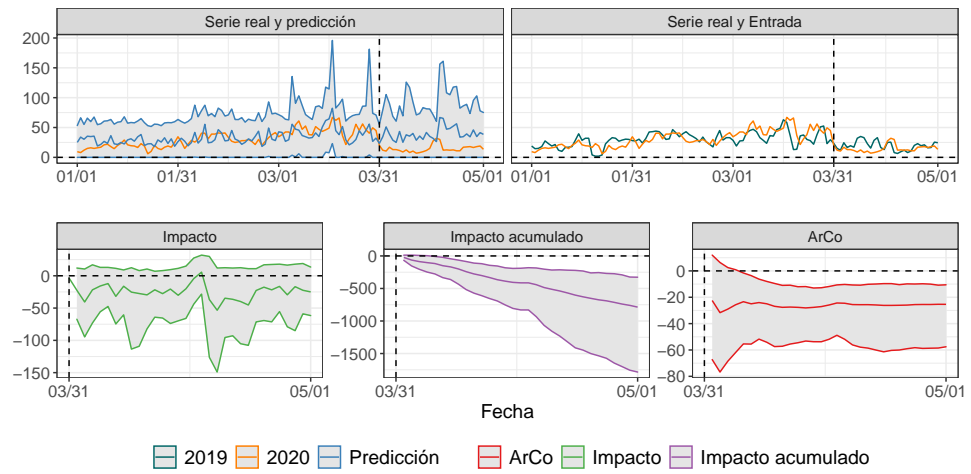


Figura 7-17: Resultados PM 2.5 para la estación Corporacion Universitaria Lasallista.

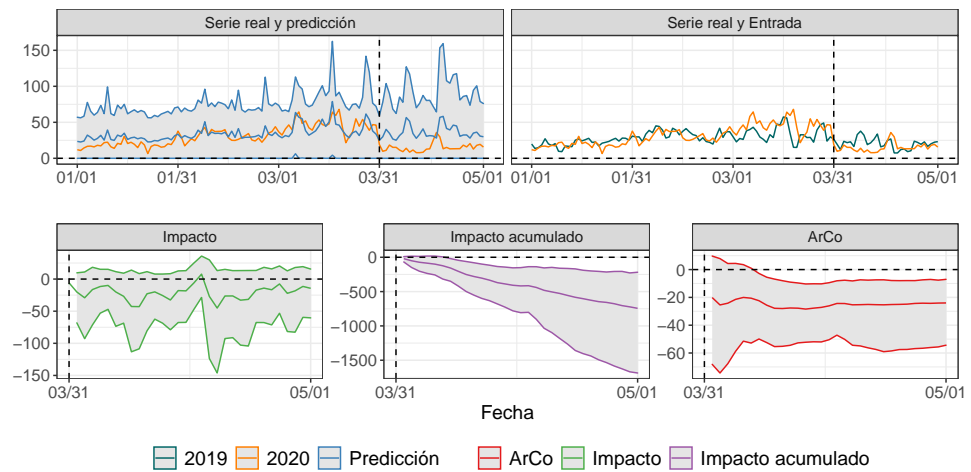
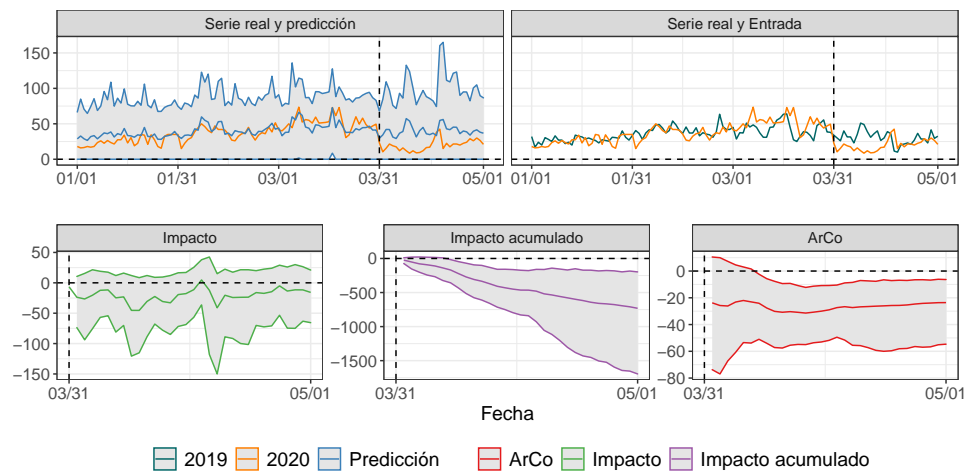
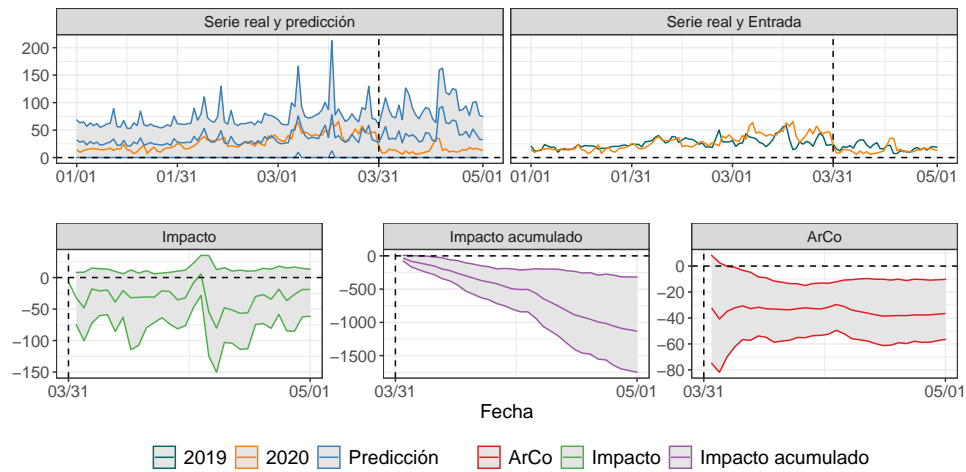


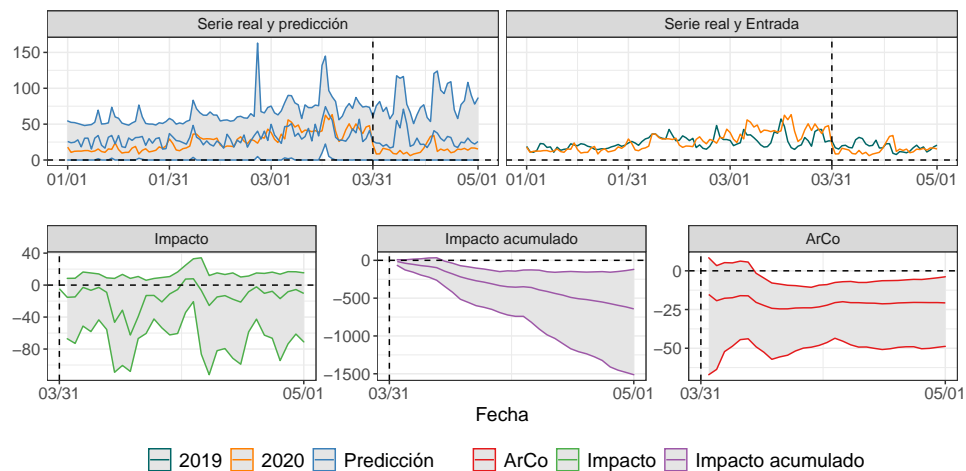
Figura 7-18: Resultados PM 2.5 para la estación Corporacion Universitaria Lasallista.



**Figura 7-19:** Resultados PM 2.5 para la estación Tráfico Sur.



**Figura 7-20:** Resultados PM 2.5 para la estación I.E. Rafael J. Mejía.



**Figura 7-21:** Resultados PM 2.5 para la estación E.S.E. Santa Gertrudis.

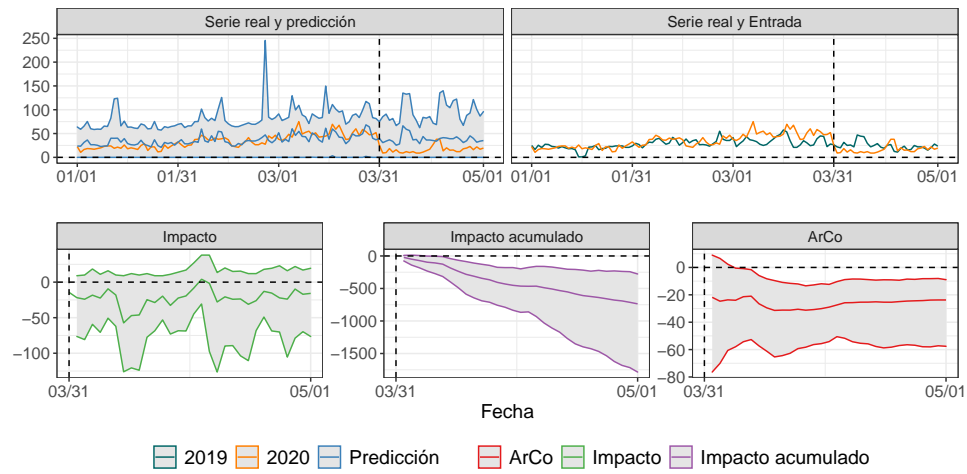


Figura 7-22: Resultados PM 2.5 para la estación Casa de Justicia Itagüí.

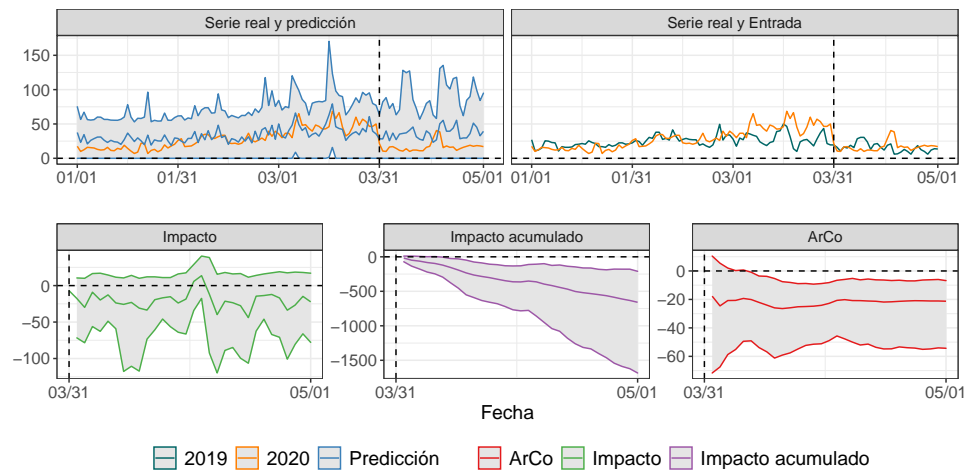
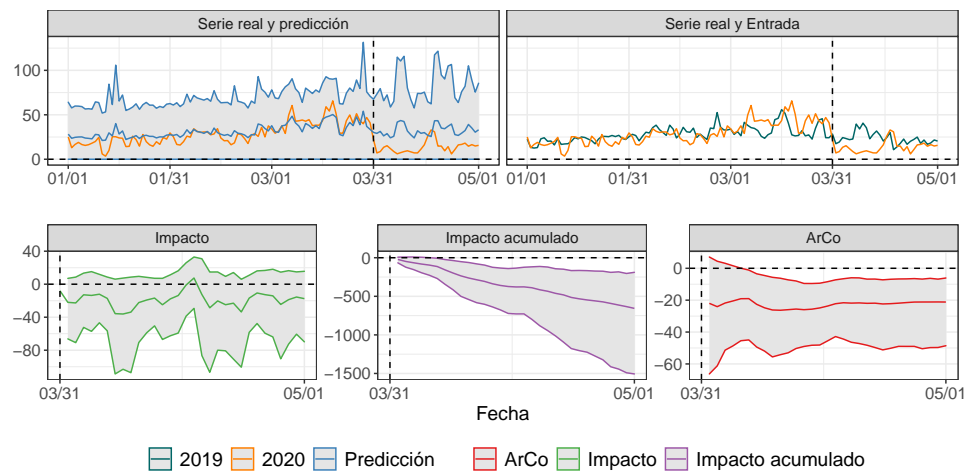
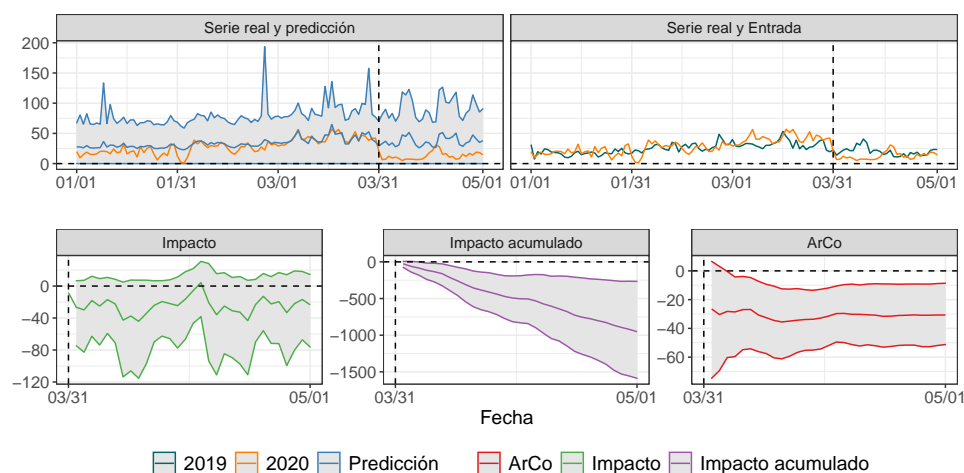


Figura 7-23: Resultados PM 2.5 para la estación Concejo Municipal de Itagüí.





**Figura 7-24:** Resultados PM 2.5 para la estación I.E. Pedro Octavio Amado.



**Figura 7-25:** Resultados PM 2.5 para la estación I.E Pedro Justo Berrio.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-1.45, 11.97]	[-21.70, 179.58]	Global
8	[2020-03-16, 2020-03-31]	[-8.52, 17.98]	[-127.82, 269.64]	Caldas - Corporacion Universitaria Lasallista
9	[2020-03-16, 2020-03-31]	[-9.14, 17.98]	[-137.12, 269.75]	Caldas - E U Joaquín Aristizabal
10	[2020-03-16, 2020-03-31]	[-8.82, 19.72]	[-132.32, 295.75]	Estación Tráfico Sur
11	[2020-03-16, 2020-03-31]	[-15.43, 20.49]	[-231.39, 307.42]	La Estrella - Hospital
12	[2020-03-16, 2020-03-31]	[-6.42, 24.80]	[-96.27, 371.95]	Sabaneta - I.E. Rafael J. Mejía
13	[2020-03-31, 2020-04-15]	[-42.86, -22.07]	[-642.97, -331.01]	Global
14	[2020-03-31, 2020-04-15]	[-53.76, -12.79]	[-806.44, -191.84]	Caldas - Corporacion Universitaria Lasallista
15	[2020-03-31, 2020-04-15]	[-52.04, -10.27]	[-780.64, -154.07]	Caldas - E U Joaquín Aristizabal
16	[2020-03-31, 2020-04-15]	[-53.04, -10.87]	[-795.60, -162.98]	Estación Tráfico Sur
17	[2020-03-31, 2020-04-15]	[-52.76, -10.02]	[-791.39, -150.29]	La Estrella - Hospital
18	[2020-03-31, 2020-04-15]	[-53.37, -13.07]	[-800.53, -196.07]	Sabaneta - I.E. Rafael J. Mejía
19	[2020-04-15, 2020-05-01]	[-38.21, -12.43]	[-611.35, -198.81]	Global
20	[2020-04-15, 2020-05-01]	[-45.96, -1.78]	[-735.39, -28.56]	Caldas - Corporacion Universitaria Lasallista
21	[2020-04-15, 2020-05-01]	[-45.53, -1.06]	[-728.48, -17.03]	Caldas - E U Joaquín Aristizabal
22	[2020-04-15, 2020-05-01]	[-82.76, 2.75]	[-1324.22, 44.07]	Estación Tráfico Sur
23	[2020-04-15, 2020-05-01]	[-40.89, 2.03]	[-654.20, 32.54]	La Estrella - Hospital
24	[2020-04-15, 2020-05-01]	[-45.68, 1.20]	[-730.85, 19.15]	Sabaneta - I.E. Rafael J. Mejía

**Tabla 7-6:** Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 1.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-9.90, 11.09]	[-148.51, 166.33]	Global
8	[2020-03-16, 2020-03-31]	[-18.99, 16.19]	[-284.84, 242.79]	Envigado - E.S.E. Santa Gertrudis
9	[2020-03-16, 2020-03-31]	[-19.52, 28.56]	[-292.81, 428.37]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-18.78, 19.06]	[-281.77, 285.91]	Itagüí - I.E. Concejo Municipal de Itagüí
11	[2020-03-16, 2020-03-31]	[-17.67, 21.13]	[-265.09, 316.93]	Medellín, Altavista - I.E. Pedro Octavio Amado
12	[2020-03-16, 2020-03-31]	[-19.85, 15.53]	[-297.70, 232.89]	Medellín, Belén - I.E. Pedro Justo Berrio
13	[2020-03-31, 2020-04-15]	[-41.02, -19.15]	[-615.30, -287.21]	Global
14	[2020-03-31, 2020-04-15]	[-48.17, -9.25]	[-722.56, -138.71]	Envigado - E.S.E. Santa Gertrudis
15	[2020-03-31, 2020-04-15]	[-55.65, -11.91]	[-834.82, -178.69]	Itagüí - Casa de Justicia Itagüí
16	[2020-03-31, 2020-04-15]	[-51.18, -8.84]	[-767.75, -132.66]	Itagüí - I.E. Concejo Municipal de Itagüí
17	[2020-03-31, 2020-04-15]	[-48.15, -9.36]	[-722.26, -140.43]	Medellín, Altavista - I.E. Pedro Octavio Amado
18	[2020-03-31, 2020-04-15]	[-54.36, -12.80]	[-815.33, -191.94]	Medellín, Belén - I.E. Pedro Justo Berrio
19	[2020-04-15, 2020-05-01]	[-35.93, -12.44]	[-574.95, -199.06]	Global
20	[2020-04-15, 2020-05-01]	[-71.19, -7.28]	[-1138.98, -116.49]	Envigado - E.S.E. Santa Gertrudis
21	[2020-04-15, 2020-05-01]	[-60.72, -5.04]	[-971.50, -80.58]	Itagüí - Casa de Justicia Itagüí
22	[2020-04-15, 2020-05-01]	[-37.28, 10.45]	[-596.48, 167.21]	Itagüí - I.E. Concejo Municipal de Itagüí
23	[2020-04-15, 2020-05-01]	[-35.23, -1.43]	[-563.63, -22.88]	Medellín, Altavista - I.E. Pedro Octavio Amado
24	[2020-04-15, 2020-05-01]	[-42.39, -3.20]	[-678.18, -51.12]	Medellín, Belén - I.E. Pedro Justo Berrio

**Tabla 7-7:** Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 2.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-7.33, 14.20]	[-110.01, 213.06]	Global
8	[2020-03-16, 2020-03-31]	[-8.39, 30.48]	[-125.79, 457.25]	Barbosa - Torre Social
9	[2020-03-16, 2020-03-31]	[-4.23, 19.92]	[-63.49, 298.75]	Bello - I.E. Fernando Vélez
10	[2020-03-16, 2020-03-31]	[-7.87, 20.41]	[-117.99, 306.10]	Copacabana - Ciudadela Educativa La Vida
11	[2020-03-16, 2020-03-31]	[-57.67, 27.28]	[-865.00, 409.14]	Medellín, Aranjuez - I.E. Ciro Mendía
12	[2020-03-16, 2020-03-31]	[-19.10, 21.89]	[-286.47, 328.32]	Medellín, Villahermosa - Planta de producción de agua potable EPM
13	[2020-03-31, 2020-04-15]	[-37.66, -21.98]	[-564.97, -329.74]	Global
14	[2020-03-31, 2020-04-15]	[-39.26, -10.11]	[-588.83, -151.58]	Barbosa - Torre Social
15	[2020-03-31, 2020-04-15]	[-36.99, -10.17]	[-554.89, -152.54]	Bello - I.E. Fernando Vélez
16	[2020-03-31, 2020-04-15]	[-47.12, -13.04]	[-706.76, -195.60]	Copacabana - Ciudadela Educativa La Vida
17	[2020-03-31, 2020-04-15]	[-74.19, -9.23]	[-1112.83, -138.40]	Medellín, Aranjuez - I.E. Ciro Mendía
18	[2020-03-31, 2020-04-15]	[-53.15, -14.14]	[-797.22, -212.10]	Medellín, Villahermosa - Planta de producción de agua potable EPM
19	[2020-04-15, 2020-05-01]	[-38.92, -15.47]	[-622.79, -247.52]	Global
20	[2020-04-15, 2020-05-01]	[-44.64, -3.59]	[-714.20, -57.39]	Barbosa - Torre Social
21	[2020-04-15, 2020-05-01]	[-45.46, -6.18]	[-727.44, -98.86]	Bello - I.E. Fernando Vélez
22	[2020-04-15, 2020-05-01]	[-42.26, -3.55]	[-676.14, -56.79]	Copacabana - Ciudadela Educativa La Vida
23	[2020-04-15, 2020-05-01]	[-68.71, -2.23]	[-1099.30, -35.72]	Medellín, Aranjuez - I.E. Ciro Mendía
24	[2020-04-15, 2020-05-01]	[-53.46, -3.91]	[-855.40, -62.48]	Medellín, Villahermosa - Planta de producción de agua potable EPM

**Tabla 7-8:** Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 3.

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[ 0.53, 16.68]	[ 8.01, 250.21]	Global
7	[2020-03-16, 2020-03-31]	[ -6.66, 25.86]	[ -99.96, 387.88]	Estación Tráfico Centro
8	[2020-03-16, 2020-03-31]	[ -5.19, 22.69]	[ -77.88, 340.38]	Medellín, El Poblado - I.E INEM sede Santa Catalina
9	[2020-03-16, 2020-03-31]	[ -7.30, 26.79]	[ -109.57, 401.92]	Medellín, El Poblado - Tanques La Ye EPM
10	[2020-03-16, 2020-03-31]	[ -4.28, 20.23]	[ -64.18, 303.42]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero
11	[2020-03-31, 2020-04-15]	[ -46.81, -19.96]	[ -702.12, -299.46]	Global
12	[2020-03-31, 2020-04-15]	[ -64.67, -10.39]	[ -970.11, -155.86]	Estación Tráfico Centro
13	[2020-03-31, 2020-04-15]	[ -54.40, -8.71]	[ -815.99, -130.66]	Medellín, El Poblado - I.E INEM sede Santa Catalina
14	[2020-03-31, 2020-04-15]	[ -63.78, -9.50]	[ -956.70, -142.48]	Medellín, El Poblado - Tanques La Ye EPM
15	[2020-03-31, 2020-04-15]	[ -47.38, -6.56]	[ -710.74, -98.45]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero
16	[2020-04-15, 2020-05-01]	[ -43.84, -8.50]	[ -701.44, -136.02]	Global
17	[2020-04-15, 2020-05-01]	[ -63.81, 3.99]	[ -1020.92, 63.86]	Estación Tráfico Centro
18	[2020-04-15, 2020-05-01]	[ -55.12, 0.62]	[ -881.86, 9.89]	Medellín, El Poblado - I.E INEM sede Santa Catalina
19	[2020-04-15, 2020-05-01]	[ -67.85, 0.35]	[ -1085.63, 5.57]	Medellín, El Poblado - Tanques La Ye EPM
20	[2020-04-15, 2020-05-01]	[ -48.53, 1.97]	[ -776.42, 31.51]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero

**Tabla 7-9:** Resultados PM 2.5 agregados e individuales para diferentes periodos de tiempo, los resultados agregados aparecen con nombre de estación “global”, parte 4

## 7.2. Resultados adicionales para el modelo espacial

### 7.2.1. PM 2.5

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[ 0.73, 19.71]	[ 11.00, 295.60]	Global
8	[2020-03-16, 2020-03-31]	[ 2.95, 22.35]	[ 44.20, 335.30]	Caldas - Corporacion Universitaria Lasallista
9	[2020-03-16, 2020-03-31]	[ 6.28, 22.13]	[ 94.18, 331.90]	Caldas - E U Joaquín Aristizabal
10	[2020-03-16, 2020-03-31]	[ 3.30, 24.81]	[ 49.43, 372.10]	Estación Tráfico Sur
11	[2020-03-16, 2020-03-31]	[ -25.66, 26.87]	[ -384.86, 403.10]	La Estrella - Hospital
12	[2020-03-16, 2020-03-31]	[ -8.08, 23.43]	[ -121.17, 351.51]	Sabaneta - I.E. Rafael J. Mejía
13	[2020-03-31, 2020-04-15]	[ -28.57, -11.09]	[ -428.55, -166.41]	Global
14	[2020-03-31, 2020-04-15]	[ -23.17, -9.99]	[ -347.55, -149.87]	Caldas - Corporacion Universitaria Lasallista
15	[2020-03-31, 2020-04-15]	[ -24.20, -8.76]	[ -363.04, -131.44]	Caldas - E U Joaquín Aristizabal
16	[2020-03-31, 2020-04-15]	[ -38.81, -11.54]	[ -582.10, -173.03]	Estación Tráfico Sur
17	[2020-03-31, 2020-04-15]	[ -42.11, -3.03]	[ -631.67, -45.42]	La Estrella - Hospital
18	[2020-03-31, 2020-04-15]	[ -40.97, -7.76]	[ -614.49, -116.34]	Sabaneta - I.E. Rafael J. Mejía
19	[2020-04-15, 2020-05-01]	[ -16.38, -0.90]	[ -262.13, -14.46]	Global
20	[2020-04-15, 2020-05-01]	[ -12.29, 3.02]	[ -196.61, 48.34]	Caldas - Corporacion Universitaria Lasallista
21	[2020-04-15, 2020-05-01]	[ -33.09, -0.04]	[ -529.37, -0.68]	Caldas - E U Joaquín Aristizabal
22	[2020-04-15, 2020-05-01]	[ -14.51, 7.71]	[ -232.15, 123.38]	Estación Tráfico Sur
23	[2020-04-15, 2020-05-01]	[ -33.43, 5.52]	[ -534.84, 88.37]	La Estrella - Hospital
24	[2020-04-15, 2020-05-01]	[ -33.99, 1.61]	[ -543.92, 25.82]	Sabaneta - I.E. Rafael J. Mejía

**Tabla 7-10:** Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 1.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-4.37, 20.28]	[-65.57, 304.22]	Global
8	[2020-03-16, 2020-03-31]	[11.24, 23.27]	[168.56, 348.99]	Envigado - E.S.E. Santa Gertrudis
9	[2020-03-16, 2020-03-31]	[-29.66, 29.56]	[-444.90, 443.44]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-22.86, 27.36]	[-342.88, 410.41]	Itagüí - I.E. Concejo Municipal de Itagüí
11	[2020-03-16, 2020-03-31]	[-2.89, 22.57]	[-43.36, 338.62]	Medellín, Altavista - I.E. Pedro Octavio Amado
12	[2020-03-16, 2020-03-31]	[-10.91, 21.26]	[-163.70, 318.84]	Medellin, Belén - I.E. Pedro Justo Berrio
13	[2020-03-31, 2020-04-15]	[-34.98, -13.24]	[-524.72, -198.57]	Global
14	[2020-03-31, 2020-04-15]	[-22.13, -7.43]	[-331.97, -111.52]	Envigado - E.S.E. Santa Gertrudis
15	[2020-03-31, 2020-04-15]	[-70.28, -9.11]	[-1054.17, -136.68]	Itagüí - Casa de Justicia Itagüí
16	[2020-03-31, 2020-04-15]	[-60.00, -6.97]	[-900.06, -104.49]	Itagüí - I.E. Concejo Municipal de Itagüí
17	[2020-03-31, 2020-04-15]	[-31.55, -10.48]	[-473.21, -157.25]	Medellín, Altavista - I.E. Pedro Octavio Amado
18	[2020-03-31, 2020-04-15]	[-44.75, -11.50]	[-671.28, -172.46]	Medellin, Belén - I.E. Pedro Justo Berrio
19	[2020-04-15, 2020-05-01]	[-28.13, -2.48]	[-450.10, -39.63]	Global
20	[2020-04-15, 2020-05-01]	[-16.23, 0.78]	[-259.75, 12.42]	Envigado - E.S.E. Santa Gertrudis
21	[2020-04-15, 2020-05-01]	[-62.20, 5.30]	[-995.27, 84.85]	Itagüí - Casa de Justicia Itagüí
22	[2020-04-15, 2020-05-01]	[-53.93, 2.10]	[-862.96, 33.52]	Itagüí - I.E. Concejo Municipal de Itagüí
23	[2020-04-15, 2020-05-01]	[-31.29, -1.43]	[-500.72, -22.83]	Medellín, Altavista - I.E. Pedro Octavio Amado
24	[2020-04-15, 2020-05-01]	[-36.62, -2.06]	[-585.95, -33.02]	Medellin, Belén - I.E. Pedro Justo Berrio

**Tabla 7-11:** Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 2.

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[2.55, 22.21]	[38.19, 333.14]	Global
8	[2020-03-16, 2020-03-31]	[8.16, 29.94]	[122.42, 449.09]	Barbosa - Torre Social
9	[2020-03-16, 2020-03-31]	[-4.87, 24.97]	[-73.01, 374.51]	Bello - I.E. Fernando Vélez
10	[2020-03-16, 2020-03-31]	[3.05, 24.41]	[45.68, 366.20]	Copacabana - Ciudadela Educativa La Vida
11	[2020-03-16, 2020-03-31]	[-3.33, 25.99]	[-50.01, 389.86]	Medellín, Aranjuez - I.E. Ciro Mendia
12	[2020-03-16, 2020-03-31]	[-24.66, 27.80]	[-369.97, 416.99]	Medellín, Villahermosa - Planta de producción de agua potable EPM
13	[2020-03-31, 2020-04-15]	[-29.13, -11.40]	[-436.92, -171.00]	Global
14	[2020-03-31, 2020-04-15]	[-31.92, -6.54]	[-478.76, -98.12]	Barbosa - Torre Social
15	[2020-03-31, 2020-04-15]	[-36.91, -7.66]	[-553.63, -114.93]	Bello - I.E. Fernando Vélez
16	[2020-03-31, 2020-04-15]	[-31.47, -9.28]	[-472.02, -139.26]	Copacabana - Ciudadela Educativa La Vida
17	[2020-03-31, 2020-04-15]	[-40.94, -10.70]	[-614.09, -160.44]	Medellín, Aranjuez - I.E. Ciro Mendia
18	[2020-03-31, 2020-04-15]	[-32.72, -9.38]	[-490.83, -140.77]	Medellín, Villahermosa - Planta de producción de agua potable EPM
19	[2020-04-15, 2020-05-01]	[-20.06, -1.84]	[-321.02, -29.48]	Global
20	[2020-04-15, 2020-05-01]	[-11.81, 0.59]	[-188.96, 9.47]	Barbosa - Torre Social
21	[2020-04-15, 2020-05-01]	[-32.28, 5.16]	[-516.54, 82.57]	Bello - I.E. Fernando Vélez
22	[2020-04-15, 2020-05-01]	[-24.65, -0.70]	[-394.37, -11.18]	Copacabana - Ciudadela Educativa La Vida
23	[2020-04-15, 2020-05-01]	[-26.73, 0.53]	[-427.74, 8.42]	Medellín, Aranjuez - I.E. Ciro Mendia
24	[2020-04-15, 2020-05-01]	[-25.71, -1.28]	[-411.31, -20.48]	Medellín, Villahermosa - Planta de producción de agua potable EPM

**Tabla 7-12:** Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 3.

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[ 11.70, 23.59]	[ 175.49, 353.85]	Global
7	[2020-03-16, 2020-03-31]	[ 8.61, 25.99]	[ 129.14, 389.85]	Estación Tráfico Centro
8	[2020-03-16, 2020-03-31]	[ 9.54, 24.16]	[ 143.12, 362.44]	Medellín, El Poblado - I.E INEM sede Santa Catalina
9	[2020-03-16, 2020-03-31]	[ 12.20, 29.37]	[ 182.93, 440.56]	Medellín, El Poblado - Tanques La Ye EPM
10	[2020-03-16, 2020-03-31]	[ 8.08, 21.36]	[ 121.25, 320.47]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero
11	[2020-03-31, 2020-04-15]	[ -20.31, -9.65]	[ -304.67, -144.76]	Global
12	[2020-03-31, 2020-04-15]	[ -27.07, -10.49]	[ -406.08, -157.33]	Estación Tráfico Centro
13	[2020-03-31, 2020-04-15]	[ -20.65, -7.11]	[ -309.69, -106.60]	Medellín, El Poblado - I.E INEM sede Santa Catalina
14	[2020-03-31, 2020-04-15]	[ -22.81, -7.01]	[ -342.18, -105.18]	Medellín, El Poblado - Tanques La Ye EPM
15	[2020-03-31, 2020-04-15]	[ -18.85, -6.68]	[ -282.79, -100.14]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero
16	[2020-04-15, 2020-05-01]	[ -10.55, 0.38]	[ -168.82, 6.07]	Global
17	[2020-04-15, 2020-05-01]	[ -14.45, 2.20]	[ -231.12, 35.23]	Estación Tráfico Centro
18	[2020-04-15, 2020-05-01]	[ -11.30, 2.13]	[ -180.77, 34.13]	Medellín, El Poblado - I.E INEM sede Santa Catalina
19	[2020-04-15, 2020-05-01]	[ -12.95, 2.58]	[ -207.23, 41.21]	Medellín, El Poblado - Tanques La Ye EPM
20	[2020-04-15, 2020-05-01]	[ -9.74, 2.43]	[ -155.81, 38.85]	Medellín, San Cristobal - Parque Biblioteca Fernando Botero

**Tabla 7-13:** Resultados PM 2.5 agregados individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, Grupo 4.

## 7.2.2. NO

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[ -29.80, -4.62]	[ -447.04, -69.37]	Global
7	[2020-03-16, 2020-03-31]	[ -16.86, -6.81]	[ -252.96, -102.17]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[ -12.68, 1.46]	[ -190.14, 21.93]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[ -9.21, -0.09]	[ -138.12, -1.42]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[ -85.50, -7.14]	[ -1282.57, -107.12]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[ -29.50, -6.34]	[ -442.46, -95.03]	Global
12	[2020-03-31, 2020-04-15]	[ -17.84, -6.21]	[ -267.66, -93.20]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[ -16.07, -2.29]	[ -240.99, -34.29]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[ -10.61, -1.18]	[ -159.21, -17.73]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[ -84.54, -7.54]	[ -1268.11, -113.15]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[ -28.77, -3.80]	[ -460.30, -60.86]	Global
17	[2020-04-15, 2020-05-01]	[ -13.96, -4.20]	[ -223.37, -67.13]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[ -14.86, 0.09]	[ -237.83, 1.42]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[ -9.03, -0.20]	[ -144.55, -3.19]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[ -82.84, -4.76]	[ -1325.51, -76.21]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-14:** Resultados NO agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”.

### 7.2.3. NO<sub>2</sub>

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-4.19, 0.29]	[-62.90, 4.39]	Global
7	[2020-03-16, 2020-03-31]	[-6.29, 0.16]	[-94.33, 2.47]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-6.36, -0.67]	[-95.41, -10.00]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[-6.63, 0.32]	[-99.51, 4.84]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-2.22, 5.97]	[-33.23, 89.56]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-6.59, -1.77]	[-98.84, -26.54]	Global
12	[2020-03-31, 2020-04-15]	[-6.82, 0.27]	[-102.23, 4.07]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-6.03, -0.11]	[-90.45, -1.59]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[-10.22, -3.19]	[-153.36, -47.80]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[-7.86, 0.93]	[-117.90, 13.99]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-3.97, 0.25]	[-63.60, 4.03]	Global
17	[2020-04-15, 2020-05-01]	[-3.47, 2.79]	[-55.56, 44.58]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-2.87, 2.61]	[-45.99, 41.69]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[-7.24, -0.51]	[-115.89, -8.20]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[-7.38, 0.78]	[-118.06, 12.54]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-15:** Resultados NO<sub>2</sub> agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”.

### 7.2.4. NO<sub>x</sub>

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[-16.85, -6.35]	[-252.76, -95.27]	Global
7	[2020-03-16, 2020-03-31]	[-22.43, -5.80]	[-336.48, -86.95]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[-11.72, -2.43]	[-175.87, -36.43]	Institucion Universitaria ITM Robledo
9	[2020-03-16, 2020-03-31]	[-16.45, -5.59]	[-246.78, -83.84]	Itagüí - Casa de Justicia Itagüí
10	[2020-03-16, 2020-03-31]	[-24.59, -3.80]	[-368.92, -57.01]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-19.87, -8.56]	[-298.02, -128.38]	Global
12	[2020-03-31, 2020-04-15]	[-25.78, -8.38]	[-386.75, -125.72]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-14.11, -4.41]	[-211.62, -66.14]	Institucion Universitaria ITM Robledo
14	[2020-03-31, 2020-04-15]	[-16.22, -4.76]	[-243.25, -71.45]	Itagüí - Casa de Justicia Itagüí
15	[2020-03-31, 2020-04-15]	[-30.19, -8.67]	[-452.90, -130.02]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-11.82, -1.66]	[-189.04, -26.49]	Global
17	[2020-04-15, 2020-05-01]	[-14.17, 1.73]	[-226.67, 27.61]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-6.76, 2.36]	[-108.22, 37.82]	Institucion Universitaria ITM Robledo
19	[2020-04-15, 2020-05-01]	[-9.94, 0.58]	[-159.02, 9.27]	Itagüí - Casa de Justicia Itagüí
20	[2020-04-15, 2020-05-01]	[-23.71, -3.26]	[-379.29, -52.16]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-16:** Resultados NO<sub>x</sub> agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”.

## 7.2.5. PM 10

	Intervalo	IC ArCo	IC Cumsum	Estación
7	[2020-03-16, 2020-03-31]	[-3.88, 20.32]	[-58.27, 304.81]	Global
8	[2020-03-16, 2020-03-31]	[7.51, 25.27]	[112.68, 379.09]	Estación Tráfico Centro
9	[2020-03-16, 2020-03-31]	[3.18, 23.47]	[47.74, 352.00]	Exito San Antonio - Medellin
10	[2020-03-16, 2020-03-31]	[-36.15, 26.47]	[-542.25, 397.06]	Institucion Educativa Colombia. Girardota
11	[2020-03-16, 2020-03-31]	[-43.63, 33.37]	[-654.47, 500.56]	Institucion Universitaria ITM Robledo
12	[2020-03-16, 2020-03-31]	[-10.10, 27.44]	[-151.48, 411.53]	Universidad San Buenaventura
13	[2020-03-31, 2020-04-15]	[-48.78, -21.26]	[-731.77, -318.90]	Global
14	[2020-03-31, 2020-04-15]	[-41.75, -20.53]	[-626.18, -307.90]	Estación Tráfico Centro
15	[2020-03-31, 2020-04-15]	[-43.45, -18.74]	[-651.82, -281.14]	Exito San Antonio - Medellin
16	[2020-03-31, 2020-04-15]	[-86.77, -14.59]	[-1301.50, -218.86]	Institucion Educativa Colombia. Girardota
17	[2020-03-31, 2020-04-15]	[-87.68, -10.07]	[-1315.27, -151.12]	Institucion Universitaria ITM Robledo
18	[2020-03-31, 2020-04-15]	[-50.87, -13.42]	[-763.10, -201.27]	Universidad San Buenaventura
19	[2020-04-15, 2020-05-01]	[-34.43, -5.11]	[-550.87, -81.83]	Global
20	[2020-04-15, 2020-05-01]	[-56.98, 8.60]	[-911.75, 137.68]	Estación Tráfico Centro
21	[2020-04-15, 2020-05-01]	[-16.33, 7.63]	[-261.27, 122.16]	Exito San Antonio - Medellin
22	[2020-04-15, 2020-05-01]	[-83.68, 2.68]	[-1338.92, 42.87]	Institucion Educativa Colombia. Girardota
23	[2020-04-15, 2020-05-01]	[-69.39, 17.15]	[-1110.32, 274.40]	Institucion Universitaria ITM Robledo
24	[2020-04-15, 2020-05-01]	[-39.11, -0.89]	[-625.72, -14.21]	Universidad San Buenaventura

**Tabla 7-17:** Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, parte 1.

	Intervalo	IC ArCo	IC Cumsum	Estación
6	[2020-03-16, 2020-03-31]	[5.27, 20.41]	[79.06, 306.20]	Global
7	[2020-03-16, 2020-03-31]	[-3.76, 13.65]	[-56.47, 204.71]	Estación Tráfico Sur
8	[2020-03-16, 2020-03-31]	[2.63, 23.97]	[39.49, 359.53]	Itagüí - Estación de Policía Los Gómez
9	[2020-03-16, 2020-03-31]	[7.55, 27.12]	[113.31, 406.86]	Itagüí - I.E. Concejo Municipal de Itagüí
10	[2020-03-16, 2020-03-31]	[5.51, 27.14]	[82.59, 407.13]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
11	[2020-03-31, 2020-04-15]	[-32.95, -18.35]	[-494.20, -275.25]	Global
12	[2020-03-31, 2020-04-15]	[-33.56, -15.97]	[-503.36, -239.50]	Estación Tráfico Sur
13	[2020-03-31, 2020-04-15]	[-37.20, -16.80]	[-557.97, -251.94]	Itagüí - Estación de Policía Los Gómez
14	[2020-03-31, 2020-04-15]	[-31.17, -12.42]	[-467.51, -186.27]	Itagüí - I.E. Concejo Municipal de Itagüí
15	[2020-03-31, 2020-04-15]	[-38.70, -17.71]	[-580.51, -265.68]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin
16	[2020-04-15, 2020-05-01]	[-17.61, -1.91]	[-281.69, -30.63]	Global
17	[2020-04-15, 2020-05-01]	[-15.51, 2.61]	[-248.08, 41.78]	Estación Tráfico Sur
18	[2020-04-15, 2020-05-01]	[-21.36, -0.09]	[-341.80, -1.40]	Itagüí - Estación de Policía Los Gómez
19	[2020-04-15, 2020-05-01]	[-17.04, 2.46]	[-272.68, 39.38]	Itagüí - I.E. Concejo Municipal de Itagüí
20	[2020-04-15, 2020-05-01]	[-25.75, -3.49]	[-411.96, -55.90]	Politecnico Colombiano Jaime Isaza Cadavid - Medellin

**Tabla 7-18:** Resultados PM 10 agregados e individuales para diferentes periodos de tiempo, construido usando el modelo (2-19), los resultados agregados aparecen con nombre de estación “global”, parte 2.

## 7.3. Estaciones de monitoreo de la calidad de aire

### 7.3.1. PM 10

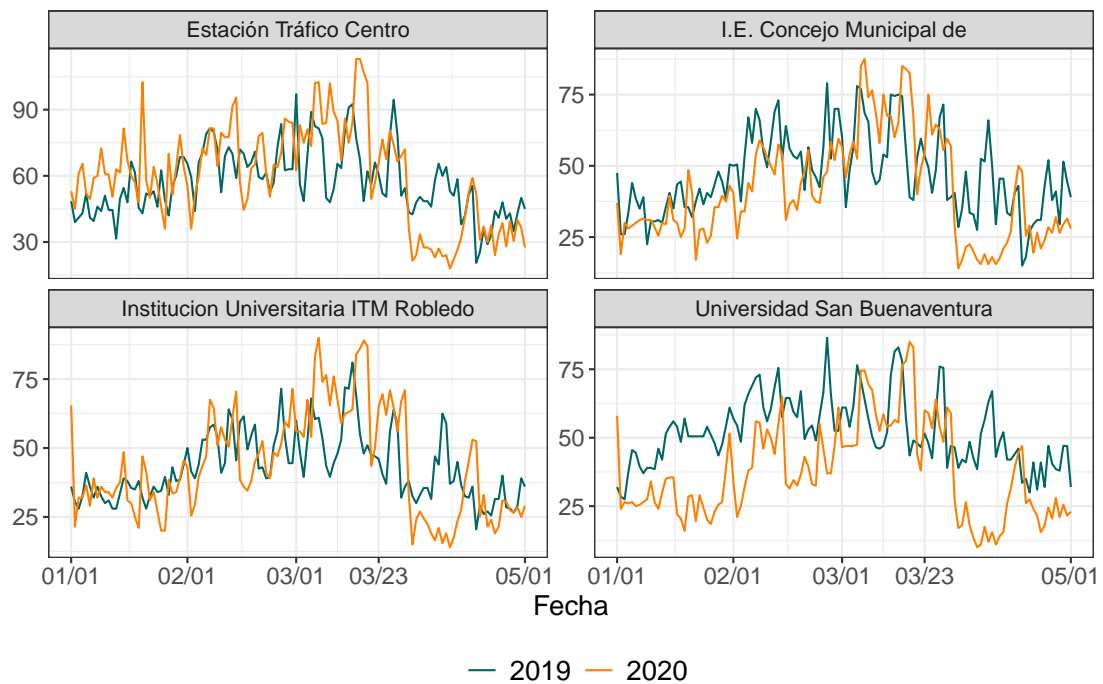


Figura 7-26: Estaciones PM 10 parte 1



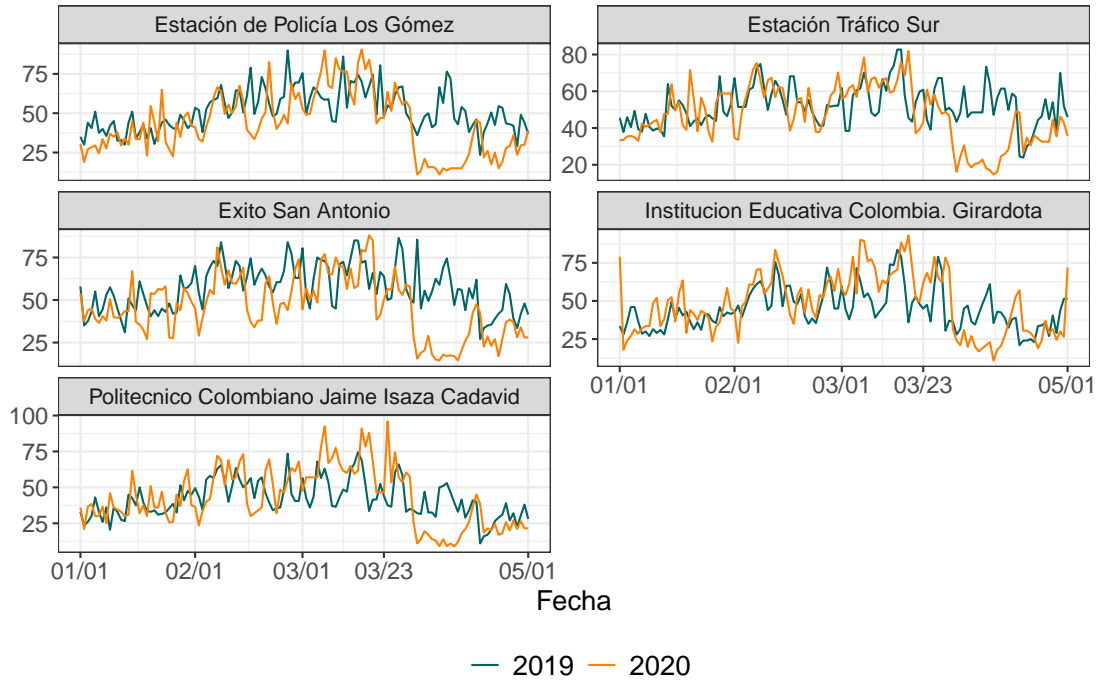


Figura 7-27: Estaciones PM 10 parte 2

### 7.3.2. PM 2.5

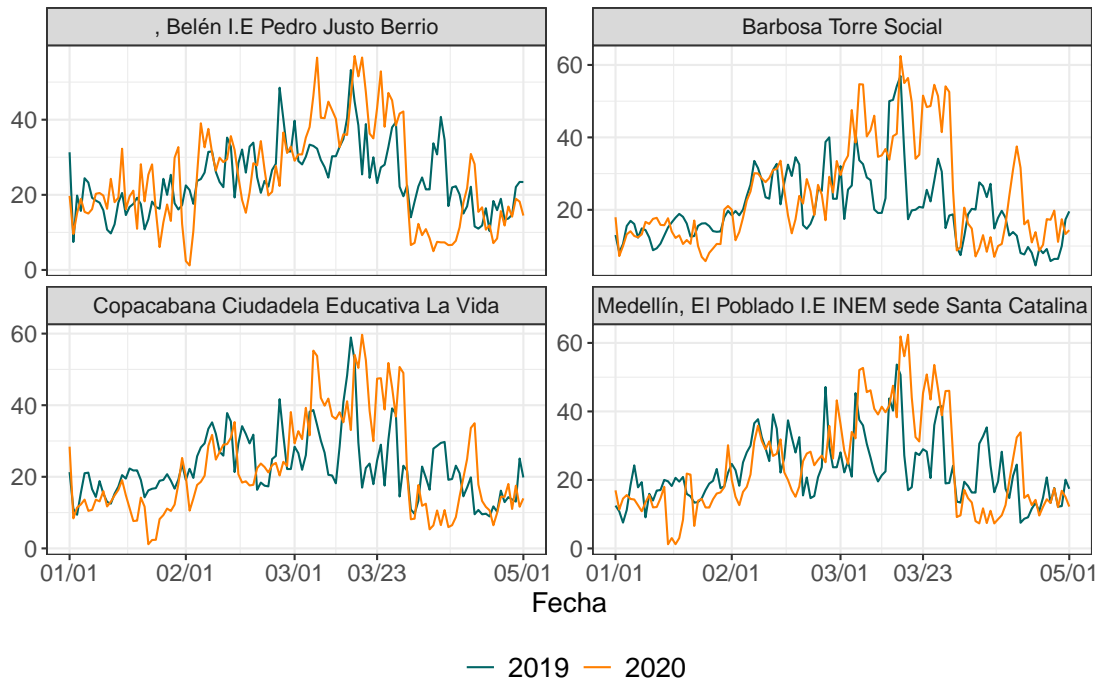
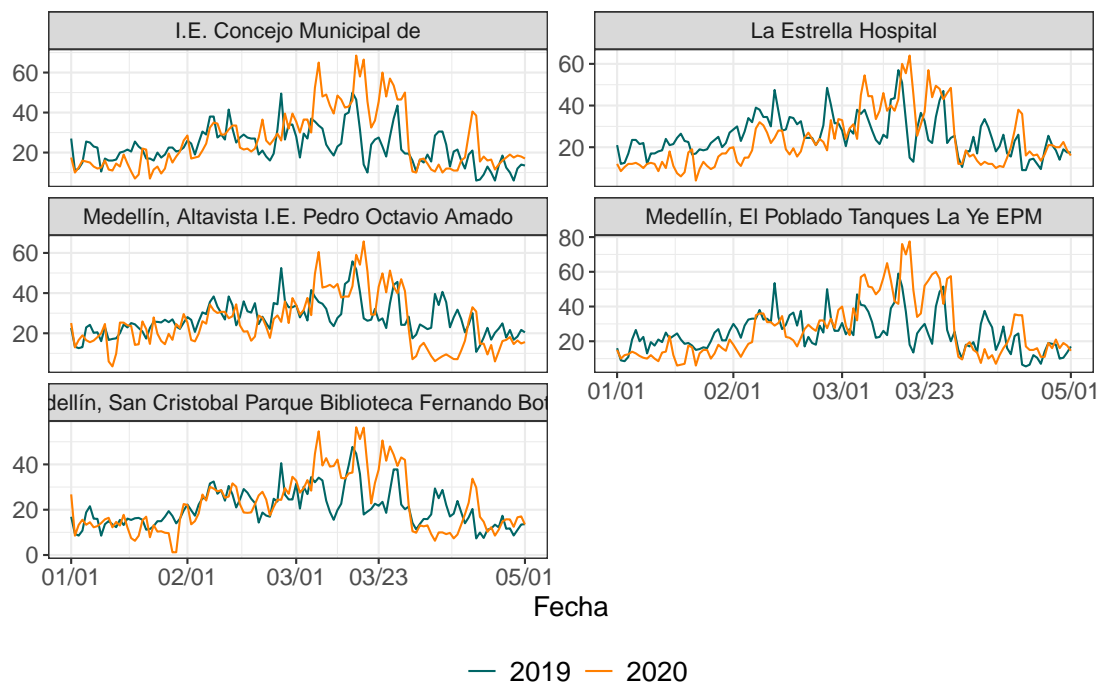
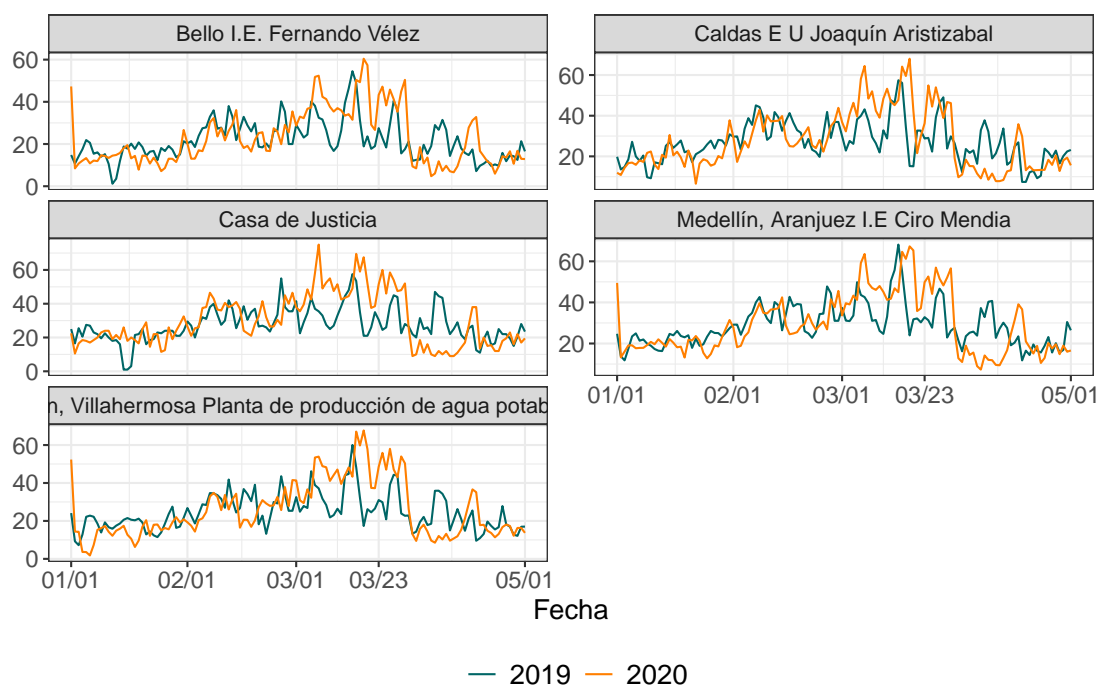


Figura 7-28: Estaciones PM 2.5 parte 1

**Figura 7-29:** Estaciones PM 2.5 parte 2**Figura 7-30:** Estaciones PM 2.5 parte 3

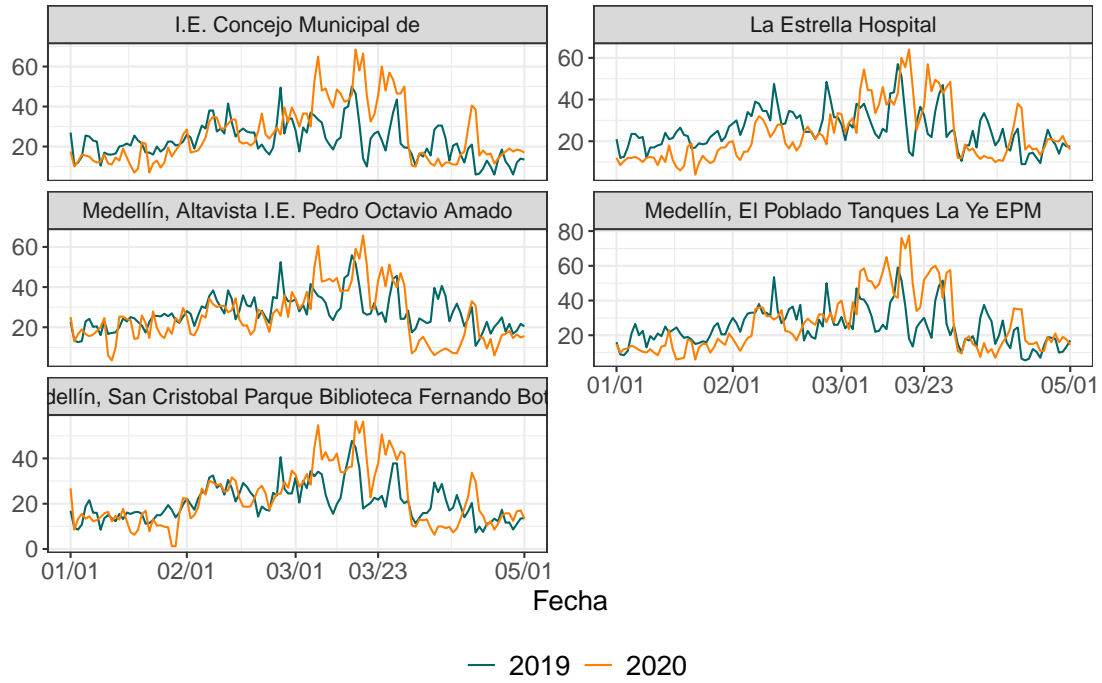


Figura 7-31: Estaciones PM 2.5 parte 4

7.3.3. NO

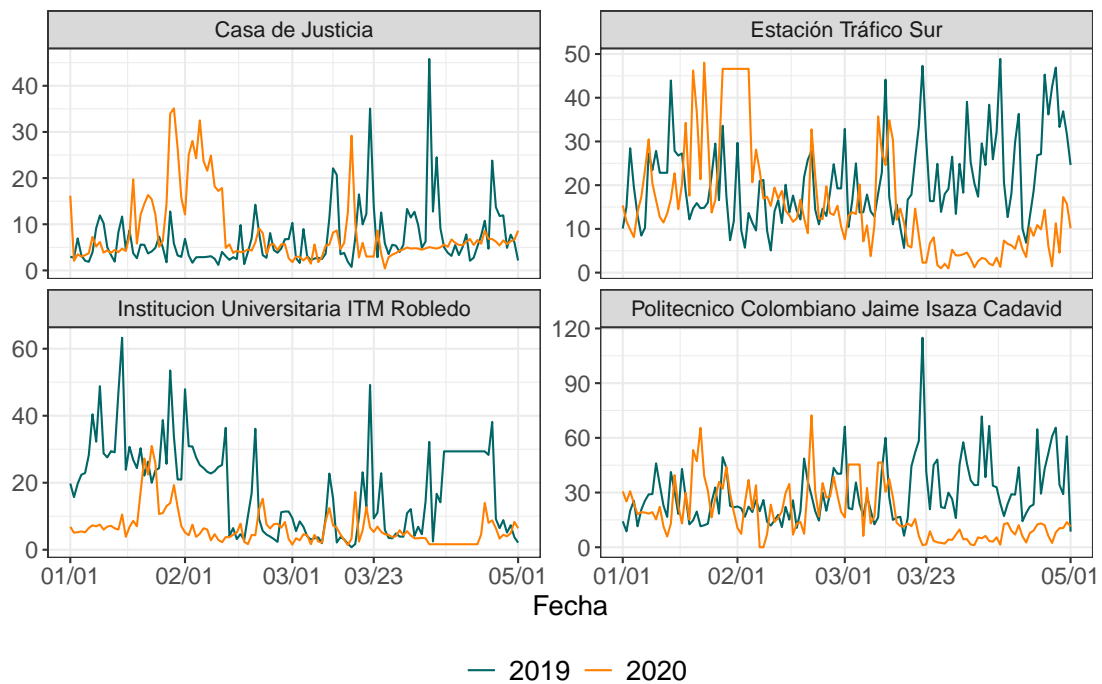


Figura 7-32: Estaciones NO

7.3.4. NO<sub>2</sub>

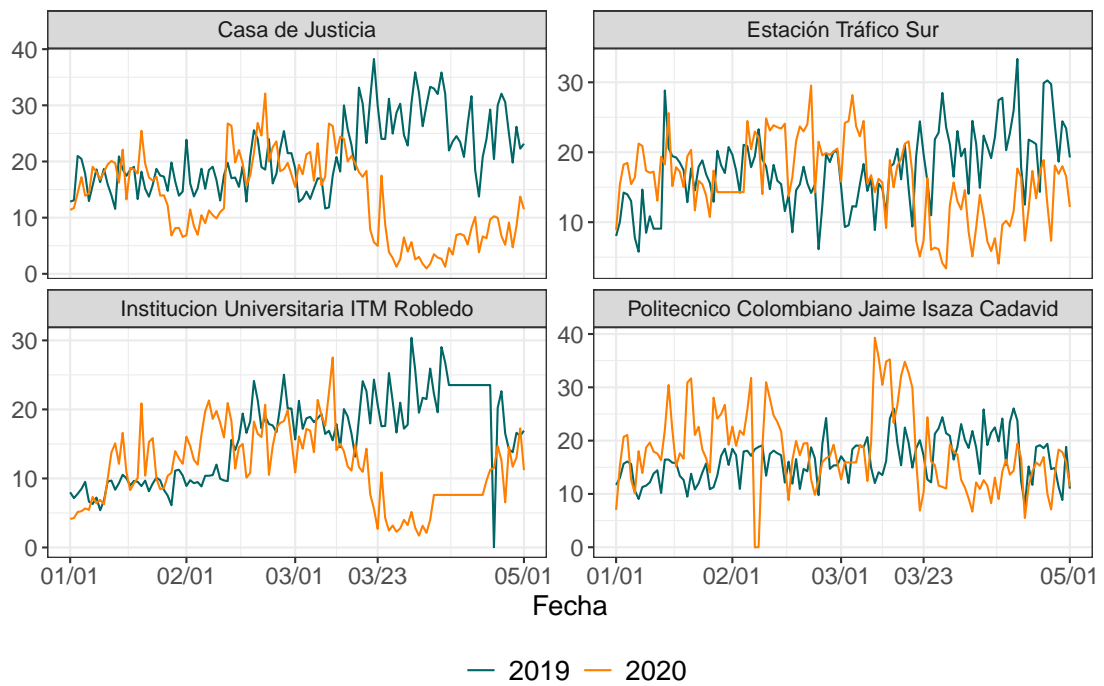
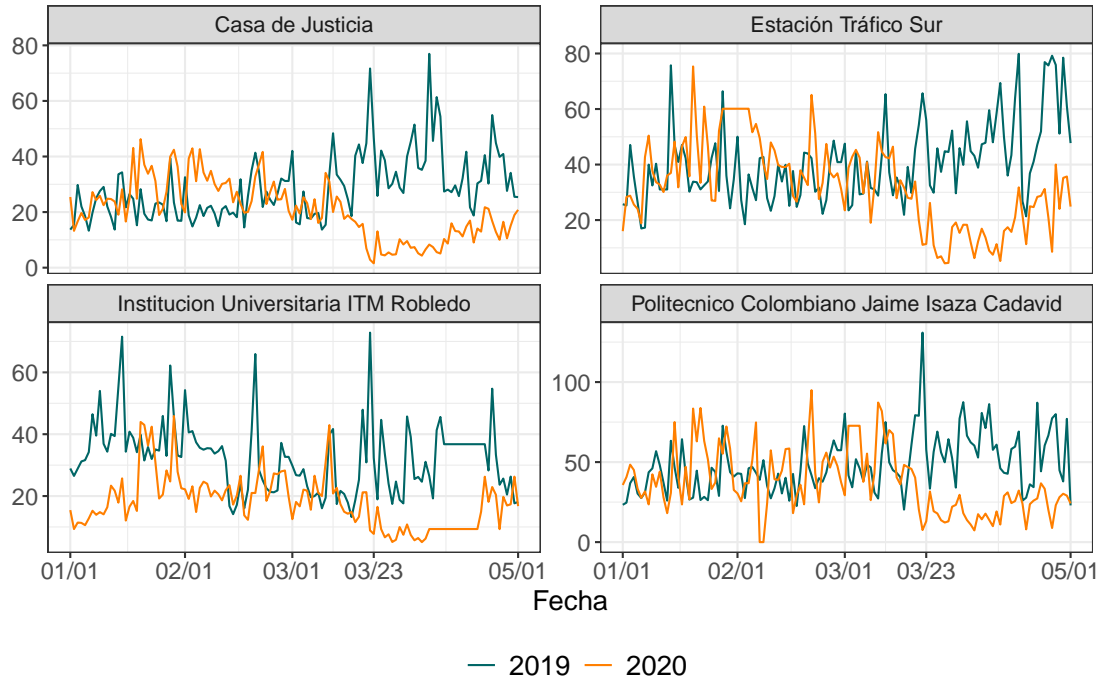


Figura 7-33: Estaciones NO<sub>2</sub>

### 7.3.5. NO<sub>x</sub>



**Figura 7-34:** Estaciones NO<sub>x</sub>

# Bibliografía

- Abadie, A. (2010). *Difference-in-Difference Estimators*, pages 36–39. Palgrave Macmillan UK, London.
- Alcaldía de Medellín (2020). Gaceta Oficial N°4677. 2020(Marzo 20).
- Ameen, J. & Harrison, P. (1985). Normal discount bayesian models. *Bayesian statistics*, 2(2):271–198.
- Bedoya, J. & Martínez, E. (2009). Calidad del aire en el valle de aburrá antioquia - colombia. page 9.
- Belasen, A. & Polachek, S. (2008). How hurricanes affect wages and employment in local labor markets. *American Economic Review*, 98.
- Bernal, J. L., Cummins, S., & Gasparrini, A. (2016). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, 46(1):348–355.
- Box, G. E. P. & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70(349):70–79.
- Briz-Redón, Á., Belenguer-Sapiña, C., & Serrano-Aroca, Á. (2021). Changes in air pollution during COVID-19 lockdown in Spain: A multi-city study. *Journal of Environmental Sciences (China)*, 101:16–26.
- Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1):247–274.
- Cardenas, J. E. P. (2017). La calidad del aire en colombia: un problema de salud pública, un problema de todos. page 2.
- Cardona-Jiménez, J. & de B. Pereira, C. A. (2021). Assessing dynamic effects on a bayesian matrix-variate dynamic linear model: An application to task-based fmri data analysis. *Computational Statistics and Data Analysis*, 163:107297.

- Carvalho, C., Masini, R., & Medeiros, M. C. (2018a). Arco: An artificial counterfactual approach for high-dimensional panel time-series data. *Journal of econometrics*, 207(2):352–380.
- Carvalho, C., Masini, R., & Medeiros, M. C. (2018b). Arco: An artificial counterfactual approach for high-dimensional panel time-series data. *Journal of econometrics*, 207(2):352–380.
- Contraloría General de Medellín (2019). Estado anual de los recursos naturales y del ambiente del municipio de medellín.
- Dean, J. & Ghemawat, S. (2004). Mapreduce: Simplified data processing on large clusters. In *OSDI'04: Sixth Symposium on Operating System Design and Implementation*, pages 137–150, San Francisco, CA.
- Gorai, A. K. & Goyal, P. (2015). A Review on Air Quality Indexing System. 9-2:13.
- Gupta, A. K. & Nagar, D. K. (2018). *Matrix variate distributions*. Chapman and Hall/CRC.
- Hernández, M. P. (2019). Incendio en el nordeste afecta calidad del aire en el aburrá.
- IGAC (2004). Aspectos Prácticos De La Adopción Del Marco Geocéntrico Nacional De Referencia Magna-Sirgas Como Datum Oficial De Colombia. page 107.
- Kampa, M. & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, 151(2):362–367.
- Karimi, A., Shirmardi, M., Hadei, M., Birgani, Y. T., Neisi, A., Takdastan, A., & Goudarzi, G. (2019). Concentrations and health effects of short- and long-term exposure to PM<sub>2.5</sub>, NO<sub>2</sub>, and O<sub>3</sub> in ambient air of Ahvaz city, Iran (2014–2017). *Ecotoxicology and Environmental Safety*, 180(May):542–548.
- Kucukelbir, A., Ranganath, R., Gelman, A., & Blei, D. M. (2015). Automatic variational inference in stan.
- Lance, P., Guilkey, D., Hattori, A., & Angeles, G. (2014). How do we know if a program made a difference. *A guide to statistical methods for program impact evaluation*. Chapel Hill (NC): MEASURE Evaluation.
- Latza, U., Gerdes, S., & Baur, X. (2009). Effects of nitrogen dioxide on human health: Systematic review of experimental and epidemiological studies conducted between 2002 and 2006. *International Journal of Hygiene and Environmental Health*, 212(3):271–287.
- Linden, A. & Adams, J. L. (2011). Applying a propensity score-based weighting model to interrupted time series data: improving causal inference in programme evaluation. *Journal of Evaluation in Clinical Practice*, 17(6):1231–1238.

- Lu, F., Xu, D., Cheng, Y., Dong, S., Guo, C., Jiang, X., & Zheng, X. (2015). Systematic review and meta-analysis of the adverse health effects of ambient PM<sub>2.5</sub> and PM<sub>10</sub> pollution in the Chinese population. *Environmental Research*, 136:196–204.
- Maathuis, M. H. & Nandy, P. (2015). A review of some recent advances in causal inference. *arXiv:1506.07669 [stat]*. arXiv: 1506.07669.
- Mamedov, R. (2018). The mit no attribution (mit-0) license. <https://github.com/aws/mit-0>. Last retrieved 2022-06-10.
- Méndez, J. F., Pinto-Herrera, L. C., & Belalcázar-Cerón, L. C. (2018). Estudio de una intrusión de polvo sahariano en la atmósfera de colombia. *Revista Ingenierías Universidad de Medellín*, 17(32):17–34.
- Mendez-Espinosa, J. F., Rojas, N. Y., Vargas, J., Pachón, J. E., Belalcazar, L. C., & Ramírez, O. (2020). Air quality variations in Northern South America during the COVID-19 lockdown. *Science of The Total Environment*, 749:141621.
- Ministerio de Salud y Protección Social (2020). Resolución No. 0385 de 2020.
- Morgan, S. L. & Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Analytical Methods for Social Research. Cambridge University Press, 2 edition.
- Nguyen, T. & Bloom, E. (2006). Impact evaluation-methodological and operational issues. *Economic Analysis and Operation Support Division*.
- Papenberg, M. & Klau, G. (2020). Using anticlustering to partition data sets into equivalent parts. *Psychological Methods*, 26.
- Paskov, P. (2019). Difference-in-Differences.
- Posada, E., Gómez, M., & Almanza, J. (2017). Análisis comparativo y modelación de las situaciones de calidad del aire en una muestra de ciudades del mundo. Comparación con el caso de Medellín. *Revista Politécnica*, 13(25):9–29.
- Quintana, J. M. (1985). A dynamic linear matrix–variate regression model. Master’s thesis.
- Quintana, J. M. (1987). *Multivariate Bayesian forecasting models*. PhD thesis, University of Warwick.
- R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ropkins, K. & Tate, J. E. (2020). Early observations on the impact of the covid-19 lockdown on air quality trends across the uk. *Science of The Total Environment*, 754:142374.



- Rousseeuw, P. J. & Molenberghs, G. (1993). Transformation of non positive semidefinite correlation matrices. *Communications in Statistics - Theory and Methods*, 22(4):965–984.
- Shakoor, A., Chen, X., Farooq, T. H., Shahzad, U., Ashraf, F., Rehman, A., Sahar, N. e., & Yan, W. (2020). Fluctuations in environmental pollutants and air quality during the lockdown in the USA and China: two sides of COVID-19 pandemic. *Air Quality, Atmosphere & Health*.
- Stan Development Team (2022). Stan. Version 2.29.0.
- Stroud, J. R., Müller, P., & Sansó, B. (2001). Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(4):673–689.
- Tian, G., Qiao, Z., & Xu, X. (2014). Characteristics of particulate matter (PM10) and its relationship with meteorological factors during 2001-2012 in Beijing. *Environmental Pollution*, 192(September 2018):266–274.
- West, M. & Harrison, J. (1997a). *Bayesian forecasting and dynamic models*. Springer series in statistics. Springer, New York, 2nd ed edition.
- West, M. & Harrison, J. (1997b). *Bayesian Forecasting and Dynamic Models (2Nd Ed.)*. Springer-Verlag New York, Inc., New York, NY, USA.
- West, M. & Harrison, P. J. (1985). *Bayesian forecasting and dynamic models*.
- Winship, C. & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25(1):659–706.
- Área Metropolitana del Valle de Aburrá (2018). Condiciones especiales del valle de aburrá.
- Área metropolitana del Valle de Aburrá (2020). Informe Anual de Calidad del Aire AMVA 2020. page 94.
- Área Metropolitana Valle de Aburra (2019). Calidad del aire, cuando inspiras respiras un mejor aire.
- Češnovar, R., Bronder, S., Sluga, D., Demšar, J., Ciglarič, T., Talts, S., & Štrumbelj, E. (2019). Gpu-based parallel computation support for stan.