



UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Cuantificación de la reducción de la concentración de la columna de NO_2 troposférico en la Ciudad Autónoma de Buenos Aires durante la pandemia de COVID-19

Tesis presentada para optar al título de Magister en Explotación de Datos y
Descubrimiento de Conocimiento

Tesista: Lic. Matias Poullain

Director: Dr. Rodrigo Castro

Codirectora: Dra. Andrea Pineda Rojas

Buenos Aires, 30 de agosto de 2024

RESUMEN

En 2020, el brote del nuevo virus SARS-CoV-2 obligó a diversos gobiernos alrededor del mundo a tomar medidas para prevenir y mitigar su transmisión. Una de las medidas más importantes fue la restricción al movimiento. Esta medida tuvo efectos positivos adicionales, como la reducción de la contaminación atmosférica por dióxido de nitrógeno (NO_2). América Latina es una de las regiones con menos estudios realizados sobre esta temática, y la Ciudad Autónoma de Buenos Aires (CABA) no es la excepción.

En este trabajo nos proponemos cuantificar la reducción de la columna de NO_2 durante la restricción a la circulación en CABA entre el 20/03/2020 y el 06/11/2020. Utilizando datos históricos de conteos de vehículos en peajes, información meteorológica e imágenes satelitales de la columna de NO_2 troposférico, se evaluaron algoritmos de machine learning y luego de seleccionar el mejor se simuló la columna de NO_2 en un escenario contrafactual sin restricciones, para finalmente comparar con los datos observados con las restricciones aplicadas.

En este trabajo se logró desarrollar satisfactoriamente modelos de machine learning de buen desempeño relativo tanto para la estimación de la circulación vehicular como para la concentración de la columna de NO_2 . Se estimó que la concentración de la columna de NO_2 disminuyó un 39 % en el primer mes y un 9 % durante todo el período.

Palabras claves: Contaminación atmosférica, NO_2 , restricciones al movimiento, Machine Learning, Prophet, imágenes satelitales, series temporales

ABSTRACT

In 2020, the outbreak of the new SARS-CoV-2 virus forced various governments around the world to take measures to prevent and mitigate its transmission. One of the most important measures was the restriction of movement. This measure had additional positive effects, such as the reduction of atmospheric pollution by nitrogen dioxide (NO_2). Latin America is one of the regions with the fewest studies conducted on this topic, and the Autonomous City of Buenos Aires (CABA) is no exception.

In this work, we aim to quantify the reduction of the NO_2 column during the movement restriction in CABA between 03/20/2020 and 11/06/2020. Using historical vehicle count data at tolls, meteorological information, and satellite images of the tropospheric NO_2 column, machine learning algorithms were evaluated. After selecting the best algorithm, the NO_2 column was simulated in a counterfactual scenario without restrictions, to finally compare with the observed data with restrictions. In this work, machine learning models with good relative performance were successfully developed for both the estimation of vehicle circulation and the concentration of the NO_2 column. It was estimated that the concentration of the NO_2 column decreased by 39 % in the first month and by 9 % throughout the entire period.

Keywords: Air pollution, NO_2 , movement restrictions, Machine Learning, Prophet, satellite imagery, time-series

Índice general

1..	Introducción	1
1.1.	Descripción del problema y motivación	1
1.2.	Objetivos	2
2..	Materiales y Métodos	5
2.1.	Datos	5
2.1.1.	Columna de NO_2 troposférico	5
2.1.2.	Variables meteorológicas	6
2.1.3.	Conteo Vehicular y Días Feriados	7
2.2.	Análisis exploratorio	8
2.3.	Cuantificación de la reducción de la columna de NO_2	9
2.3.1.	Estimación del conteo vehicular durante el período de restricción al movimiento si no lo hubiese habido	9
2.3.2.	Estimación de la columna de NO_2 durante el período de restricción al movimiento si no lo hubiese habido	10
2.3.3.	Comparaciones de las columnas de NO_2	12
3..	Resultados	13
3.0.1.	Análisis exploratorio	13
3.0.2.	Cuantificación de la reducción de la columna de NO_2	24
4..	Discusión y Conclusiones	29

1. INTRODUCCIÓN

1.1. Descripción del problema y motivación

La pandemia del virus SARS-CoV-2 obligó a los diferentes gobiernos a establecer aislamientos sociales obligatorios y restricciones al movimiento dentro y entre ciudades. Estas estrategias de control son efectivas contra enfermedades infecciosas, especialmente contra aquellas de sintomatología tardía o escasa [1], como lo es el COVID-19. En la República Argentina se estableció el 20 de marzo del 2020 el llamado *Aislamiento Social, Preventivo y Obligatorio (ASPO)* para combatir la enfermedad, a fin de mitigar su propagación y su impacto en el sistema sanitario [2]. En la Ciudad Autónoma de Buenos Aires (CABA), la restricción de circulación finalizó el 6 de noviembre del 2020, con una duración de poco más de 7 meses. Sin embargo, las restricciones a la circulación tuvieron diversos efectos secundarios, entre ellos la disminución de las emisiones de gases contaminantes y de efecto invernadero, dada la reducción de la actividad industrial y de la circulación vehicular [3, 4].

Uno de los gases contaminantes más relacionados a la circulación vehicular es el dióxido de nitrógeno (NO_2) [5, 6]. Este es un contaminante atmosférico común en áreas urbanas donde la principal fuente de emisión es el transporte vehicular [7]. Los impactos de la calidad del aire tanto en la salud humana como en el medio ambiente son importantes áreas de investigación [8, 9, 10]. Distintos estudios de corte epidemiológico evidenciaron asociaciones causales entre exposición ambiental al NO_2 y efectos adversos en la salud, como aumento de la mortalidad [11, 12, 13], aumento de casos de cáncer de pulmón [14] y asociaciones a diferentes enfermedades, inclusive a partir de exposiciones a corto plazo [15] trayendo entre otras consecuencias un aumento en la admisión hospitalaria. Los valores actuales máximos recomendados por la Organización Mundial de la Salud (OMS) son $10 \mu g/m^3$ (media anual), $25 \mu g/m^3$ (media diaria) y $200 \mu g/m^3$ (media horaria) [16]. Para CABA son de 40 PPB (media anual) y 200 PPB (media horaria) [17]. Contrariamente a otros gases, como el dióxido de carbono (CO_2 , el principal gas de efecto invernadero), el tiempo de vida (o tiempo de permanencia en la atmósfera) del NO_2 es muy corto, entre uno y dos días [18, 19], alcanzando su máximo en invierno [20]. Esto evita que las moléculas de este gas sean transportadas lejos de su fuente de emisión y hace que variaciones en la cantidad de gas emitido tengan un efecto casi inmediato sobre las concentraciones observadas. Esto quiere decir que, de observar zonas de alta concentración de NO_2 , su emisión ocurrió muy recientemente y que su fuente de emisión se encuentra muy cercana.

Para predecir cómo cambian las concentraciones de contaminantes existen modelos de calidad de aire de distintas escalas [21, 22, 23]. Entre ellas se encuentran las que se basan en modelos químicos y procesos que interrelacionan las fuentes, los sumideros y otros procesos transformativos clave. Algunos modelos suelen necesitar de una considerable potencia computacional, difícilmente accesible. Por otro lado, existen modelos que se basan en representaciones estadísticas o de aprendizaje automático que se ajustan a observaciones empíricas [24] y que pueden implementarse con requisitos computacionales más modestos. La capacidad de pronosticar los niveles de contaminación es importante para una amplia gama de partes interesadas; entre ellas, las autoridades regionales, que podrían querer aplicar estrategias adaptativas de gestión del tráfico que minimicen la exposición de los

grupos vulnerables.

Los métodos de predicción por algoritmos de machine learning se han utilizado a menudo para mitigar algunos de los retos asociados a la construcción de modelos de enfoque químico de alta resolución para su uso a escala local. Si bien tales algoritmos no buscan reemplazarlos en la modelización de las contribuciones de las fuentes y el destino del contaminante, sí ofrecen un método alternativo para utilizar las mediciones locales históricas, comprender los efectos de otras variables y predecir los cambios de concentración del gas [25].

Existen pocos intentos de realizar modelos de machine learning de la concentración de NO_2 en función de variables ambientales y antropogénicas previos al año 2020 [26, 27]. Sin embargo, las grandes fluctuaciones de la concentración de NO_2 y del tráfico vehicular en las ciudades producidas por las restricciones al movimiento impuestas en el 2020, presentaron una oportunidad sin paralelo para abordar esta problemática, como se observa en el gran número de nuevos trabajos publicados a partir de ese año [28, 29, 30, 31, 32, 33, 34, 35, 36, 37]. Como los modelos de inteligencia artificial permiten hacer predicciones aún sin tener datos de emisión, estos han sido ampliamente utilizados para estimar el impacto de la abrupta disminución del flujo vehicular durante el aislamiento en la concentración de NO_2 en distintas ciudades del mundo.

En general, estos trabajos utilizan la información relacionada a la concentración de NO_2 obtenida de dos principales fuentes de medición: mediciones en superficie y mediciones a distancia. Las primeras provienen de estaciones de referencia distribuidas en la zona de interés. Tienen como ventaja una gran resolución temporal y que las mediciones son directas. Si bien la cantidad de este tipo de instrumentos instalados está en aumento a fin de mejorar la salud pública, existen grandes desigualdades entre su distribución en los países en vía de desarrollo en comparación a los desarrollados. Por otro lado, las mediciones a distancia se refieren a las imágenes satelitales. De ellas, se destaca el sensor TROPospheric Monitoring Instrument (TROPOMI) integrado en el satélite Sentinel-5P del programa Copernicus perteneciente a la Agencia Espacial Europea (ESA) [38], cuyas grandes ventajas son su gran cobertura y resolución espacial, su cantidad de instrumentos medidores y la disponibilidad de la información.

Por otro lado, los trabajos que describen la construcción de modelos predictivos para la concentración de NO_2 suelen usar la estrategia de *Business As Usual* (BAU), donde se realiza una predicción en todo el período de estudio con datos históricos y manteniendo la tendencia observada hasta un momento previo. Los algoritmos más utilizados encontrados en la bibliografía fueron Random Forest, Redes Neuronales, Regresiones Lineales Multivariadas y/o Boosting. Si bien, en la mayoría de los casos, se integran componentes temporales en los modelos, son pocos los que utilizan modelos de series temporales [39, 25]. Ejemplos de estos últimos utilizan un algoritmo relativamente novedoso llamado *Prophet* [40]. Sin embargo, no se encontraron trabajos publicados donde modelo fuese utilizado para hacer predicciones CABA.

1.2. Objetivos

El objetivo general de este trabajo es cuantificar la reducción de la concentración de la columna de NO_2 sobre la Ciudad Autónoma de Buenos Aires durante las restricciones al movimiento del 2020 (20/03 al 06/11) impuestas para combatir la propagación del virus SARS-CoV-2. De él, se desprenden cinco objetivos específicos:

-
1. Realizar un análisis exploratorio de la columna de NO_2 troposférico medida sobre CABA, el conteo de pases de vehículos en peajes, como indicador de la circulación vehicular, y la meteorología, previo y durante las restricciones al movimiento.
 2. Desarrollar un modelo que permita pronosticar cómo hubiera sido el flujo vehicular de no haber habido restricciones al movimiento.
 3. Desarrollar un modelo que permita pronosticar la concentración de la columna de NO_2 de no haber habido restricciones al movimiento a partir de los datos históricos de NO_2 y meteorológicos y del pronóstico del conteo vehicular durante las restricciones al movimiento de no haberla habido.
 4. Obtener una estimación de la reducción de la columna de NO_2 comparando el pronóstico en un escenario BAU con su valor observado.

2. MATERIALES Y MÉTODOS

2.1. Datos

2.1.1. Columna de NO_2 troposférico

La información de la columna de NO_2 troposférico (OFFL/L3_ NO_2) fue medida por el satélite Sentinel-5P del programa Copernicus perteneciente a la ESA. El conjunto de datos L3 se descargó desde la API Javascript de Google Earth Engine (GEE), con una resolución espacial de $0.01^\circ \times 0.01^\circ$ (aproximadamente $1 \text{ Km} \times 1 \text{ Km}$) como resultado del sobremuestreo de los datos L2 OFFL NO_2 de TROPOMI por GEE utilizando la herramienta harpconvert. Su extensión espacial fue la suficiente para englobar a CABA (Vértices del polígono en la proyección WGS84: $-34.52^\circ; -58.33^\circ$ / $-34.73^\circ; -58.33^\circ$ / $-34.73^\circ; -58.56^\circ$ / $-34.52^\circ; -58.56^\circ$), conteniendo 528 píxeles. El conjunto de datos GEE OFFL NO_2 (en unidades de $\mu\text{mol}/\text{m}^2$) fue previamente filtrado para eliminar los píxeles con valores de calidad inferiores al 75 %. El conjunto de datos está disponible desde el 08/07/2018 hasta la actualidad. Sin embargo, el 29 de noviembre de 2020, la versión disponible de los datos fue actualizada a 1.4.0V presentando diferencias con la versión anterior en los criterios de recogida y preprocesamiento [41]. Por esta razón, los datos posteriores al 29 de noviembre de 2020 se excluyeron del análisis. Como resultado, se descargaron todos los rásters disponibles entre las fechas 08/07/2018 hasta el 26/11/2020 con una resolución espacial de aproximadamente un día (teniendo una observación en la mayoría de los días pero habiendo días sin observaciones, mientras que otros presentaron dos).

Para el análisis de las series temporales, se utilizó la media de la columna de NO_2 troposférico en una región específica de interés, denominada *máscara*. La *máscara* fue definida siguiendo los pasos siguientes:

1. Para cada mes, se calculó un nuevo ráster de columnas promedio de NO_2 troposférico.
2. De esos 29 rásters obtenidos, se identificaron los píxeles cuyo valor superaron el percentil 70 de su correspondiente ráster.
3. Se seleccionaron todos los píxeles identificados en el punto anterior, comunes para los 29 rásters del promedio de columna de NO_2 mensual.

Estos pasos fueron representados en la Figura 2.1

La necesidad de la construcción y utilización de la *máscara* queda en evidencia al comparar las observaciones mensuales medias de los meses de abril 2019 y 2020 (primer mes completo en restricciones al movimiento), donde la distribución espacial de su cociente presentó una heterogeneidad destacable (Figura 2.2).

La *máscara* obtenida para el análisis está presentada en la Figura 2.3, conteniendo 259 píxeles. De esta forma, se pudo obtener una *región de mayor columna de NO_2 troposférico*, es decir, aquella zona donde más se concentra el gas, que consideramos la potencialmente más afectada, excluyendo del análisis zonas con valor bajo de concentración en la columna.

Para cada día se calculó el valor promedio de los píxeles contenidos dentro de la *máscara* obteniendo así un valor de columna de NO_2 promedio diario con una resolución temporal

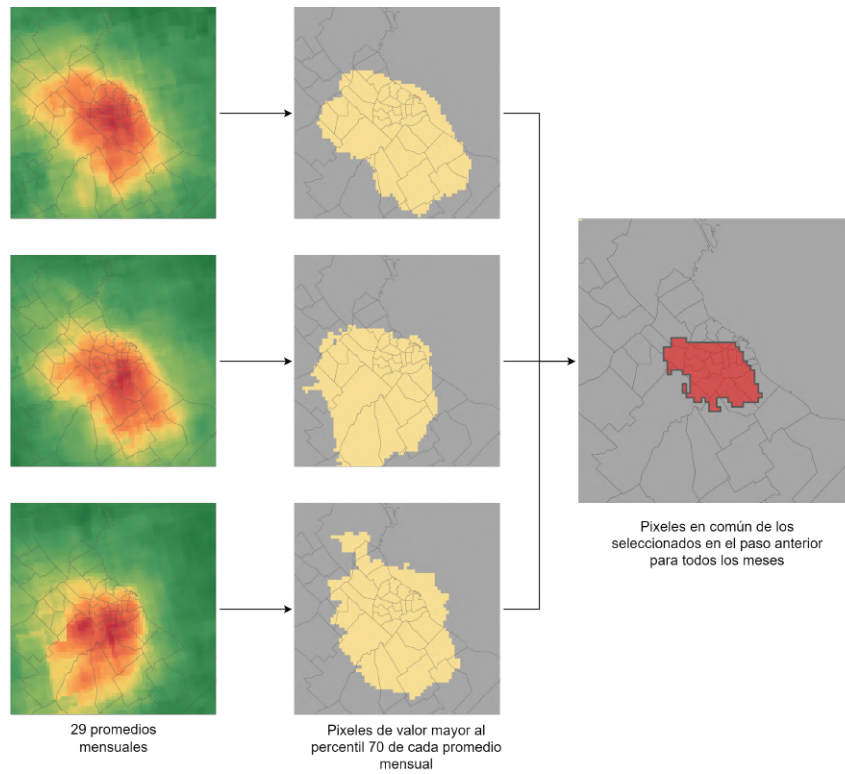


Fig. 2.1: Diagrama de los pasos realizados para la construcción de la máscara

casi diaria desde el 08/07/2018 hasta el 26/11/2020. Por último, se realizaron distintas transformaciones de esta variable (logaritmo, raíz cuadrada, etc.) para realizar distintos tipos de ensayos que serán descritos más adelante.

2.1.2. Variables meteorológicas

La información meteorológica fue compartida por el Servicio Meteorológico Nacional (SMN). El pedido se realizó en su web de contacto (<https://www.argentina.gob.ar/smn/institucional/contacto>). Esta información fue medida en la estación “Observatorio” ubicada en el barrio de Villa Ortúzar de CABA. La información obtenida junto con las variables nuevas creadas a partir de ella se presentan en la tabla 2.1.

Tab. 2.1: Variables meteorológicas utilizadas.

Nombre	Abreviación	Resolución temporal	Origen
Temperatura Horaria (°C)	T_h	Horaria	SMN
Tiempo presente	TP	Horaria	SMN
Temperatura Media Diaria (°C)	T_d	Diaria	Calculada
Ocurrencia de Precipitaciones horaria	OP_h	Horaria	Calculada
Ocurrencia de Precipitaciones diaria	OP_d	Diaria	Calculada
Dirección Cardinal del Viento		Horaria	SMN
Intensidad del Viento horario (Km/h)	IV_h	Horaria	SMN
Intensidad del Viento diaria (Km/h)	IV_d	Diaria	Calculada

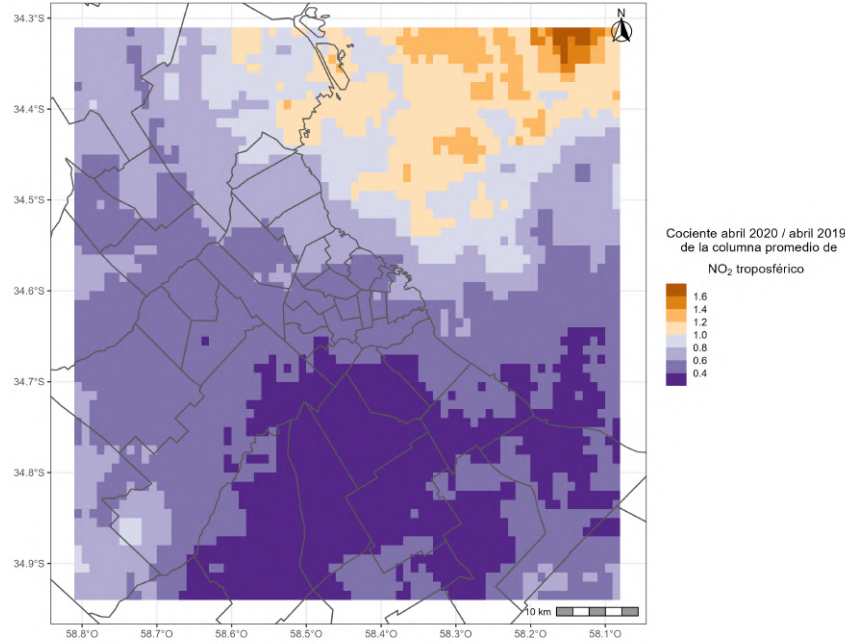


Fig. 2.2: Mapa del cociente abril 2020 sobre abril 2019 de las columnas mensuales promedio de NO_2 .

En todos los casos, se pudo obtener información desde el 08/07/2018. La T_h es la temperatura medida a principio de cada hora, mientras que el TP es la descripción cualitativa del tiempo siguiendo el código Synop [42]. A partir de la T_h se pudo calcular la T_d como el promedio diario de la primera. Además, a partir del TP , se generaron las variables OP_h y OP_d . Estas últimas indican si en esa hora o día respectivamente precipitó. Se consideró que precipitó cuando TP tomó uno de los valores siguientes: 20-27, 29, 31, 50-69 ó 80-98.

Por último, la información relacionada al viento compartida por el SMN se encontraba descripta de forma vectorial por su intensidad y dirección resultante medida por hora. La dirección estaba definida por 16 rangos cardinales discretos que fueron cuantificados en ángulos tomando como su dirección puntal el centro de cada uno de los 16 rangos cardinales. A fin de obtener la IV_d , se calculó la resultante vectorial diaria del viento y se guardó únicamente su intensidad.

Además, de las variables T_d y IV_d se construyeron las mismas variables a varios lags, denominadas T_{d-n} y IV_{d-n} donde n corresponde a la cantidad de días de lag aplicados.

2.1.3. Conteo Vehicular y Días Feriados

La información del conteo vehicular se encuentra disponible en la plataforma de datos abiertos de CABA [43]. Contiene información sobre la cantidad de pasos horarios de vehículos según su categoría (motocicletas, automóviles con y sin tráiler o vehículos pesados de dos a seis ejes), tipo de cobro (Efectivo, Mercado Pago, Rec. Deuda, Tag, Exento, Tarjeta Magnética, Violación o CPP) y sentido (Provincia o Centro). Se descargaron los datos de 2019 y 2020 y se calcularon los pasos totales en CABA entre el 01/01/2019 y el 31/12/2020 por hora y por día, obteniendo de esta forma una misma variable en dos resoluciones temporales distintas: V_h y V_d (conteo vehicular horario y diario respectivamente).

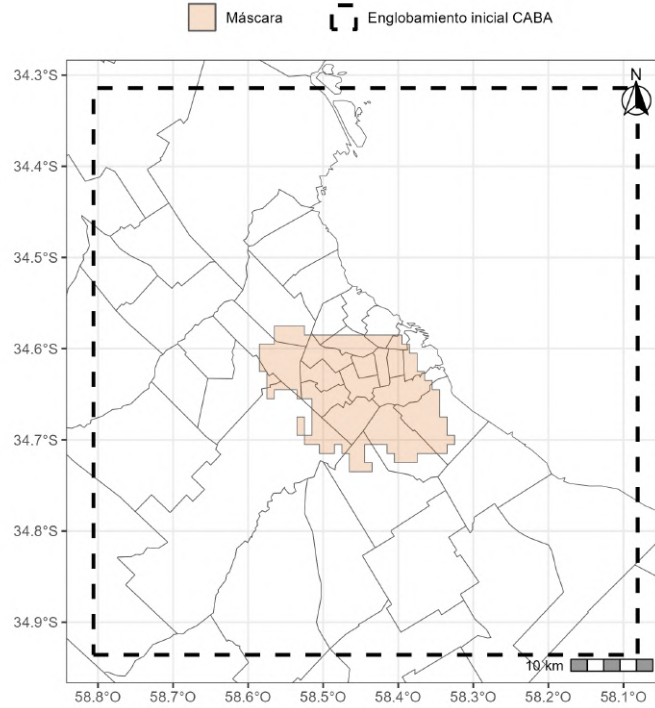


Fig. 2.3: Mapa de la extensión de la máscara.

El listado de días feriados y no laborables (denominados en conjunto como *feriados*) de los años 2019 y 2020 fue obtenido del Ministerio de Educación de la Nación [44] y del Ministerio de Interior de la Nación [45].

2.2. Análisis exploratorio

En primera instancia, se realizó un análisis descriptivo de los datos, se mapeó la columna de NO_2 troposférico promedio mensual para los meses de febrero y abril del 2019 y 2020 sobre la región de estudio. Se eligieron estas fechas ya que febrero y abril 2020 representan los meses enteros inmediatos previo y posterior al inicio de las restricciones, a su vez, los del año 2019 permiten realizar una comparación manteniendo las variables climáticas aproximadamente similares. Luego se compararon distintos tipos de medidas resumen para los períodos antes y durante las restricciones al movimiento. En particular, siendo que la columna de NO_2 troposférico y el conteo vehicular presentaron comportamientos cíclicos observables a simple vista, se realizó un análisis de series temporales aplicando transformaciones de Fourier, autocorrelogramas y gráficos agrupados por hora del día o día de la semana (según corresponda) a fin de entender cuales son las frecuencias más importantes que componen ambas series temporales.

En segunda instancia, se calcularon correlaciones de Spearman entre todas las variables cuantitativas de misma resolución temporal con distintas transformaciones y a distintos lags. Este análisis permitió observar de forma sencilla las asociaciones existentes entre las variables y cómo estas relaciones se modifican al aplicar las distintas transformaciones.

2.3. Cuantificación de la reducción de la columna de NO_2

De forma general, para realizar la cuantificación de la reducción de la columna del NO_2 troposférico se generó un pronóstico del conteo vehicular esperado durante el período posterior al 20/03/2020 de no haberse instaurado las restricciones al movimiento. Luego, se calculó la columna esperada de NO_2 sin restricciones al movimiento posterior al 20/03/2020. Por último, se analizaron tales predicciones y se realizaron comparaciones entre ella y los valores observados, cuantificando la reducción de la columna de NO_2 que consideraremos *debida a las restricciones*.

En las siguientes secciones se utilizará un algoritmo relativamente novedoso llamado Prophet [40], construido para determinar una combinación robusta de hiperparámetros con el fin de generar predicciones precisas para datos con tendencias y estructuras periódicas. Prophet es utilizado en diversas aplicaciones de Meta (ex Facebook) para generar previsiones precisas para la planificación y el establecimiento de objetivos. Se ha comprobado que supera otras estrategias [46] en ciertos escenarios. Los valores atípicos, los datos faltantes y los cambios bruscos en las series temporales son manejados eficientemente. Prophet está basado en un *modelo generalizado aditivo* en el que se ajustan tendencias no lineales con periodicidad anual, semanal y diaria.

A diferencia de la validación cruzada de los métodos estándar de regresión o clasificación, el proceso integrado y recomendado por los autores [40] incluye el ajuste a los datos históricos durante un período determinado, proporcionando previsiones continuas con N días de antelación y repitiendo el proceso durante una ventana móvil de períodos de medición. También hay parámetros que se pueden ajustar manualmente y que pueden mejorar el rendimiento. Entre ellos se encuentra la definición manual de los puntos de cambio (*change points*, o puntos en el tiempo en los que la tendencia general cambia significativamente).

2.3.1. Estimación del conteo vehicular durante el período de restricción al movimiento si no lo hubiese habido

En esta primera parte, se construyó el llamado Modelo 1: un modelo Prophet cuya fórmula es la siguiente:

$$\begin{aligned} \text{Log}_{10}(V_h)(t_h, T_h, OP_h) = & \beta_0 + S_a(t_h) + S_s(t_h) + S_{ds}(t_h) + S_{df}(t_h) + F(t_h) + \\ & \beta_1 \times T_h + \beta_2 \times OP_h \end{aligned} \quad (2.1)$$

donde $\text{Log}_{10}(V_h)$ es el conteo vehicular horario con una transformación logarítmica en base 10, t_h es la hora de la observación (distinguiendo por día y año), β_0 es la tendencia general constante, S_a es la periodicidad anual, S_s es la periodicidad semanal, S_{ds} es la periodicidad diaria de los días de semana, S_{df} es la periodicidad diaria de los días del fin de semana, F es el efecto de los feriados y β_1 y β_2 son los coeficientes de regresión de las variables T_h y OP_h respectivamente (variables descritas en la sección 2.1.2).

Primero, se realizó la hiperparametrización del modelo. Esta consistió en una optimización bayesiana de 1000 iteraciones a minimizar el error cuadrático medio (RMSE por sus siglas en inglés). Dentro de cada iteración, se realizó una validación cruzada a futuro [40] forzando que los conjuntos de entrenamiento abarquen al menos un año y usando tres folds cuyos conjuntos de test sean de igual tiempo, traducándose en conjuntos de validación de 632 horas (26.33 días) desde el 01/01/2019 hasta el 20/03/2020 (período previo a las

restricciones), como puede observarse en la Figura 2.4 A). No se realizaron repeticiones internas para reducir el efecto de las semillas ya que se observó por experimentos previos que los resultados de las validaciones fueron muy poco variantes. Los hiperparámetros optimizados fueron *changepoint.prior.scale*, *holidays.prior.scale* y *seasonality.prior.scale*. Se seleccionó la combinación de tales hiperparámetros cuyo RMSE fue el menor.

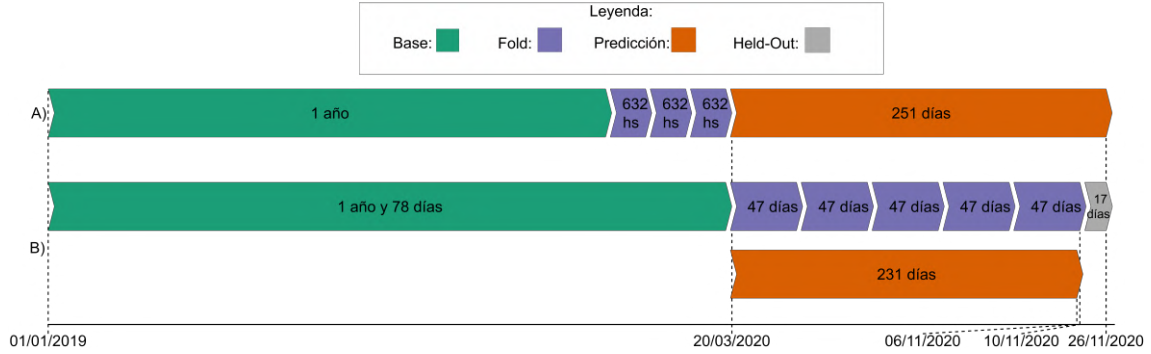


Fig. 2.4: Diagramas de los períodos de tiempo utilizados para la validación cruzada, las predicciones y como held-out para A) el Modelo 1 (conteo vehicular) y B) el Modelo 2 (columna de NO_2 troposférico)

Segundo, se entrenó el Modelo 1 con todos los datos desde el 01/01/2019 hasta el 20/03/2020 y se confirmó la bondad de ajuste observando distintos gráficos resumen.

Por último, se realizó una estimación del conteo vehicular a futuro, en el período de 20/03/2020 hasta el 26/11/2020, correspondiente al conteo vehicular esperado de no haberse instaurado las restricciones al movimiento. Siendo que el modelo fue entrenado con $\log_{10}(V_h)$ y no con V_h , los valores de V_h calculados a partir de las estimaciones en escala logarítmica del modelo fueron forzosamente positivos.

2.3.2. Estimación de la columna de NO_2 durante el período de restricción al movimiento si no lo hubiese habido

En esta segunda parte, se construyó el llamado Modelo 2.

Para definir este Modelo, se evaluaron distintos algoritmos de machine learning, utilizando en todos casi las mismas variables explicativas. Estas variables explicativas fueron agrupadas de la siguiente forma:

- Temporales:
 - Día de la semana: valores del 1 al 7.
 - Día del mes: valores del 1 al 31
 - Día del año: valores del 1 al 366
 - Días feriados: dummy de todos los dias feriados (18 variables)
 - Fecha: solo para el algoritmo Prophet, que obtiene la información de los días de la semana, mes y año de esta variable, por lo que estas últimas no son incorporadas.
- Meteorológicas:

- $\text{Log}_{10}(T_{d-2})$: T_d con un lag de 2 días y con una transformación logarítmica en base 10.
- $\text{Log}_{10}(IV_d)$: IV_d con una transformación logarítmica en base 10.
- OP_d : ocurrencia de Precipitaciones en el día (dicotómica).
- $\text{Log}_{10}(V_d)$: conteo vehicular diario con una transformación logarítmica en base 10.

Las transformaciones y lags utilizados en las variables del Modelo 2 fueron elegidos por ser los de mayor correlación de Spearman observada en la última parte del análisis descriptivo (sección 2.2).

Los algoritmos utilizados se presentan en la Tabla 2.2 junto a sus hiperparámetros elegidos para ser optimizados. En todos los casos, se realizó una optimización bayesiana de los hiperparámetros de 1000 iteraciones a minimizar el RMSE. La validación cruzada empleada para cada iteración fue similar a la del Modelo 1 pero utilizando la información disponible en el período 01/01/2019 a 08/11/2020. Siendo este superior al del Modelo 1, se pudo implementar una validación cruzada de 5 folds en vez de 3, siempre forzando a que los conjuntos de entrenamiento de cada fold contengan al menos un año de información y que los períodos de validación sean de igual largo, como puede observarse en la Figura 2.4 B).

Tab. 2.2: Algoritmos de Machine Learning evaluados y sus respectivos hiperparámetros que fueron optimizados

Nombre	Abreviación	Hiperparámetros Optimizados e intervalos/opciones de búsqueda
K-Nearest Neighbors	KNN	n_neighbors [1; 50] weights [uniform; distance] p [1.0; 2.0]
Prophet	Prophet	changepoint_prior_scale [0.001; 5.0] seasonality_prior_scale [0.01; 10.0] holidays_prior_scale [0.01; 10.0]
Ridge	Ridge	alpha [1e-10; 500.0]
Support Vector Machine	SVM	C [1e-3; 100.0] epsilon [0.001; 0.1] kernel [linear; poly; rbf; sigmoid] gamma [scale; auto] si kernel $\neq$ 'linear' degree [2; 4] si kernel = 'poly'
Extreme Gradient Boosting	XGB	subsample [0.2; 1.0] colsample_bytree [0.2; 1.0] max_depth [3; 101] min_child_weight [0; 10] eta [1e-4; 1.0]

La información del período del 09/11/2020 al 26/11/2020 fue utilizada como held-out (Figura 2.4 B)). Una vez obtenidos los hiperparámetros optimizados y entrenado el modelo final para cada algoritmo, se realizaron predicciones sobre el período held-out y se calcularon métricas de rendimiento de la predicción. Estas métricas fueron RMSE, MAPE (Error del Porcentaje Medio Absoluto por sus siglas en inglés) y R^2 . Estas fueron

las métricas elegidas para la comparación de la rendimiento de los distintos algoritmos ya que muestran una perspectiva completa de distintos aspectos del rendimiento. Estas métricas, en su conjunto, proporcionan una indicación de la precisión general del modelo en términos de unidades de la variable de respuesta y en términos porcentuales además de evaluar la bondad de ajuste en términos de la variabilidad explicada. El algoritmo elegido para continuar con el trabajo fue aquel que minimizó el RMSE y MAPE y maximizó el R^2 . Ese algoritmo fue KNN (ver sección 3.0.2 para la comparación del rendimiento de los algoritmos).

El siguiente paso fue realizar la estimación de la columna de NO_2 en el período de 01/01/2019 hasta el 26/11/2020 utilizando el algoritmo elegido a partir del conteo vehicular diario esperado de no haberse instaurado las restricciones al movimiento (predicción del Modelo 1) como variable V_d . De esta forma, se obtuvo la columna de NO_2 troposférico esperada Sin Restricciones (NO_2SR).

2.3.3. Comparaciones de las columnas de NO_2

En esta última sección se cuantificó la reducción de la columna de NO_2 calculando un **cociente** Q que definimos como el *promedio de los valores diarios de NO_2 observado sobre los de NO_2SR* .

Se calculó un Q por cada mes que contuvo días de restricciones (marzo a noviembre 2020 inclusive) y un Q general para todo el período de restricciones.

Por último, se realizó un cuadro comparativo de las reducciones promedio de la columna de NO_2 y el conteo vehicular horario.

3. RESULTADOS

3.0.1. Análisis exploratorio

Conteo vehicular

El análisis descriptivo del conteo vehicular mostró que esta variable presenta un rango muy amplio de valores observados y de una variabilidad muy grande diaria y semanal y con un cambio en la tendencia general muy marcado a partir del inicio de las restricciones al movimiento. En promedio, se contaron 14470 vehículos/hora en un rango entre 792 y 29335 vehículos/hora durante el período sin restricciones y un promedio de 8673 vehículos/hora en un rango de 184 a 28013 durante las restricciones.

La figura 3.1 visualiza la variación del conteo vehicular semanal total a lo largo de todo el período de estudio. Antes del inicio a las restricciones al movimiento, la tendencia pareciera ser constante, con conteos de vehículos menores en las últimas y primeras semanas del año especialmente. A partir del 20 de marzo de 2020, el conteo vehicular semanal total se redujo a valores medios no observados anteriormente en el período de estudio. A partir de esa fecha, se distinguió una tendencia general creciente en los valores medios y en la variabilidad, alcanzando valores similares a los del año anterior recién en las últimas semanas de diciembre.

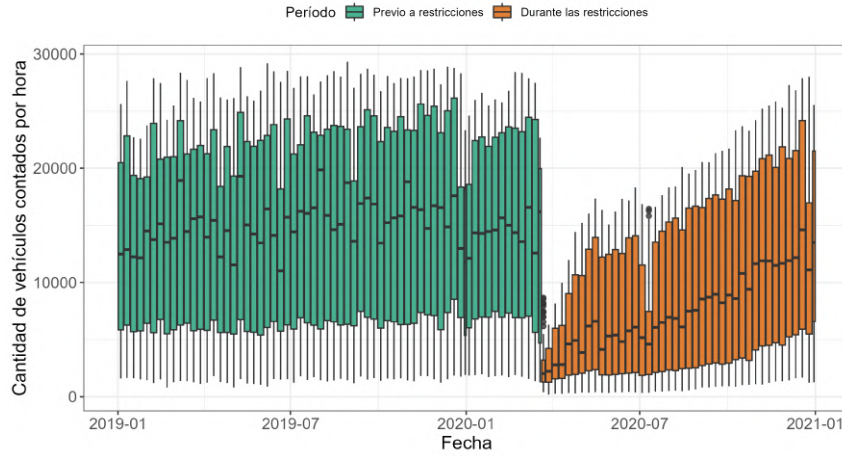


Fig. 3.1: Boxplot del conteo vehicular horario agrupado por semana y pintado según su período en relación a las restricciones al movimiento.

La figura 3.2 profundiza en las observaciones anteriores. Previo a marzo 2020, la distribución del conteo vehicular horario presentaba a los valores más bajos (cerca de 300 vehículos/hora) y a los más altos (cerca de 2500 vehículos/hora) como los más frecuentes en todos los meses, sin observarse diferencias marcadas entre meses. Sin embargo, desde marzo 2020 en adelante, los valores de vehículos contados más frecuentes son cercanos a 100 vehículos/hora exhibiendo una distribución muy asimétrica, con valores altos muy poco frecuentes. A medida que pasan los meses, los valores altos empiezan a ser más frecuentes, la distribución a ser más simétrica y parecida a las observadas en los meses previos a las restricciones al movimiento.

El comportamiento semanal y diario del conteo vehicular presentó un comportamiento cíclico a lo largo de todo el período de estudio. A partir de la Figura 3.3 A y B se pueden observar los valores medios de conteos vehiculares diarios y horarios agrupados por los días de la semana o por la hora del día respectivamente. La variación semanal se caracteriza por valores máximos de conteo de vehículos los miércoles y viernes, mientras que estos disminuyen notablemente los fines de semana. En cuanto a la variaciones diarias, aquellas de los días de semana pareciera ser diferente a las de los fines de semana y feriados, siendo estas dos similares a grandes rasgos. En los días de semana, el conteo vehicular es muy reducido a horas de la madrugada y muy alto desde las 7 hasta las 19 Hs. En los fines de semana y feriados, la variación es similar pero con una amplitud reducida: los vehículos contados en la madrugada son los más bajos del día pero no tan reducidos como los de un día de semana, en cambio los valores máximos se observan entre las 10 y las 17 Hs, siendo estos menores a los de los días de semana. Durante el período de restricciones se observó una disminución del conteo de vehículos, sin embargo, la variación relativa se mantuvo similar a la observada previa a la cuarentena durante los días de semana, fin de semana y feriados.

El autocorrelograma de la serie de tiempo de conteo de vehículos por hora (figura 3.4) muestra máximos locales significativos en todos los lags múltiplos de 24 Hs, especialmente altos en los múltiplos de 168 Hs ($24 \text{ Hs} \times 7$, es decir a lags de semanas). También, se pudieron observar valles en los lags múltiplos de 12 Hs que no son múltiplos de 24 Hs. Por último, se puede ver que las autocorrelaciones de lags cercanos son similares. Tales observaciones se ven reflejadas en el espectro de la amplitud del dominio de la frecuencia, obtenido por la transformada de Fourier, exhibida en la Figura 3.5. En ella destacan, de mayor a menor importancia, los picos de amplitud a las 24 Hs, 17544 Hs (aprox. 2 años), 12 Hs, 168.7 Hs (aprox. 1 semana), 8 Hs y 6 Hs.

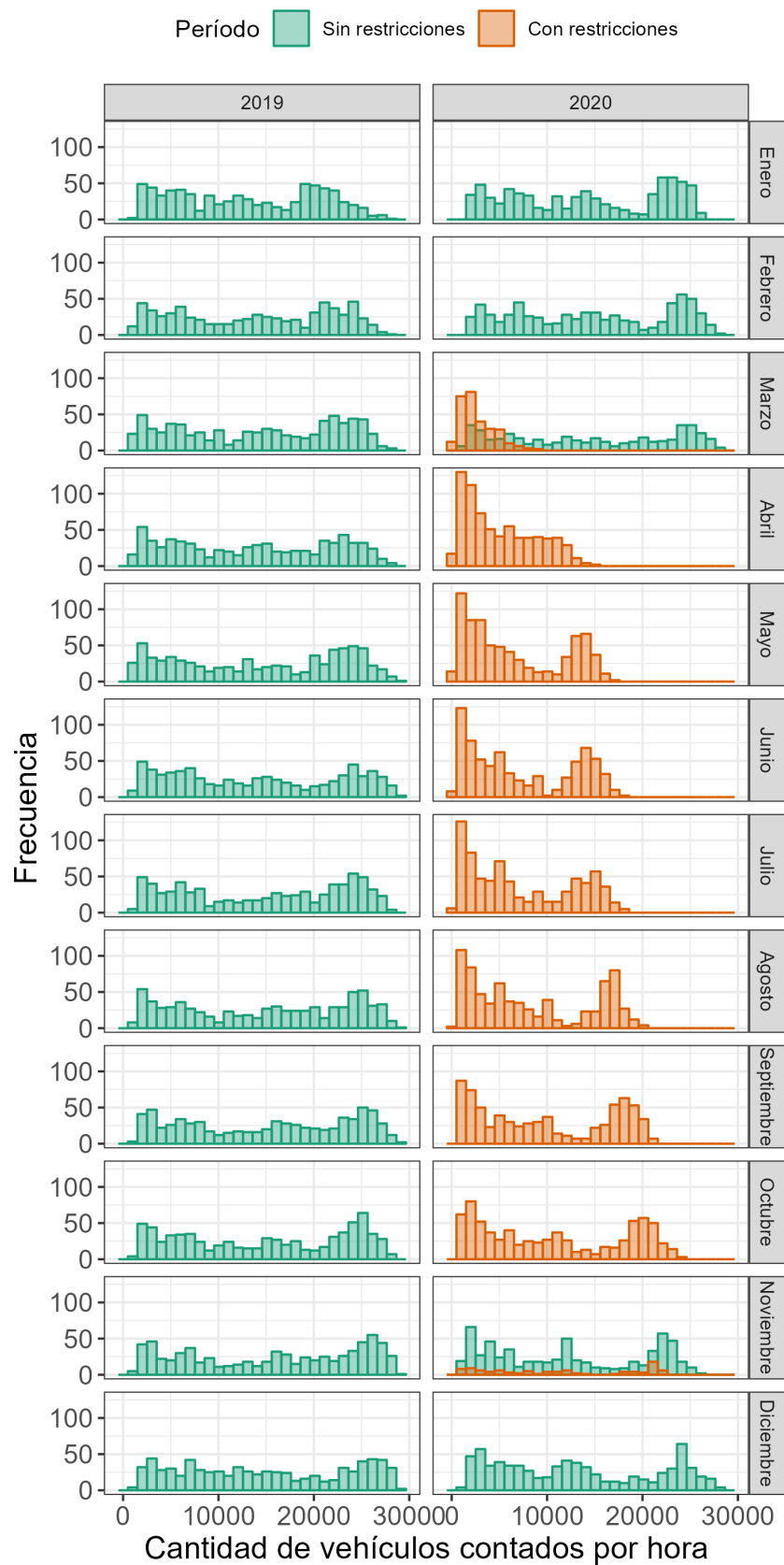


Fig. 3.2: Histogramas del conteo vehicular diario por mes y año y pintado según su período en relación a las restricciones al movimiento.

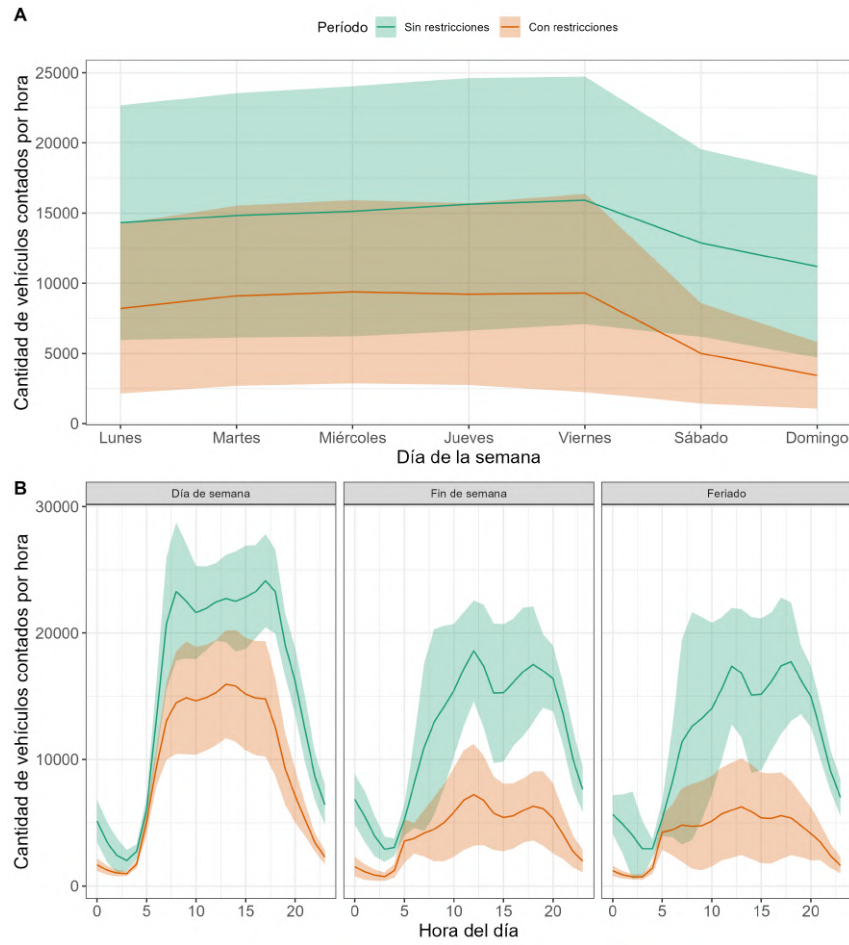


Fig. 3.3: Variaciones semanales y diarias del conteo vehicular en los períodos sin (verde) y con (naranja) restricciones al movimiento. La línea más oscura representa el promedio y el sombreado que lo rodea al desvío estándar. A: Variación semanal; promedios del conteo vehicular horario por día de la semana. B: Variación diaria: promedios del conteo vehicular horario por hora del día y agrupado por característica del día (de semana, de fin de semana, feriado).

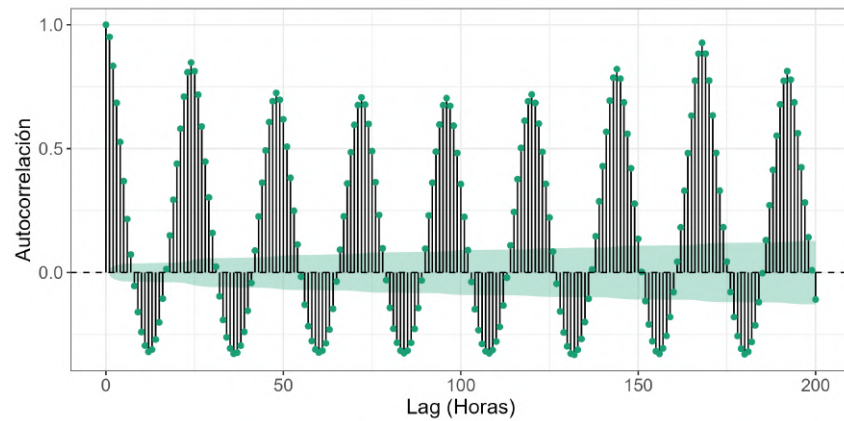


Fig. 3.4: Autocorrelograma de la serie de tiempo del conteo vehicular horario. Los puntos que se encuentran fuera del área sombreada representan correlaciones significativas.

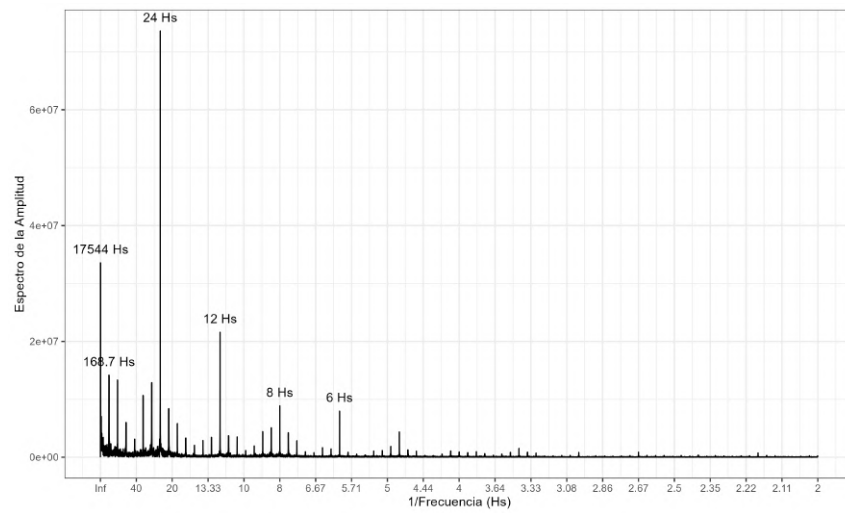


Fig. 3.5: Espectro de la amplitud del dominio de la frecuencia de la serie temporal del conteo vehicular horario

NO_2 troposférico

En las imágenes satelitales diarias descargadas (previo a la delimitación de la *máscara*), ya se pudo observar un fuerte descenso de la columna de NO_2 troposférico sobre CABA. En la figura 3.6 se mapearon los promedios mensuales para los meses de febrero y abril 2019 y 2020. En ella destaca la fuerte disminución de la columna en abril 2020 comparada a febrero 2020 y a abril 2019. También es importante notar que de febrero a abril 2019, la columna de NO_2 pareciera aumentar en general, mientras que en el 2020 ocurre lo contrario. Además, la zona de valores de columna mayor pareciera encontrarse al sur de la ciudad, mostrando un gran contraste con los valores observados sobre el Río de la Plata, que pareciera mantenerse relativamente constante en estos meses.

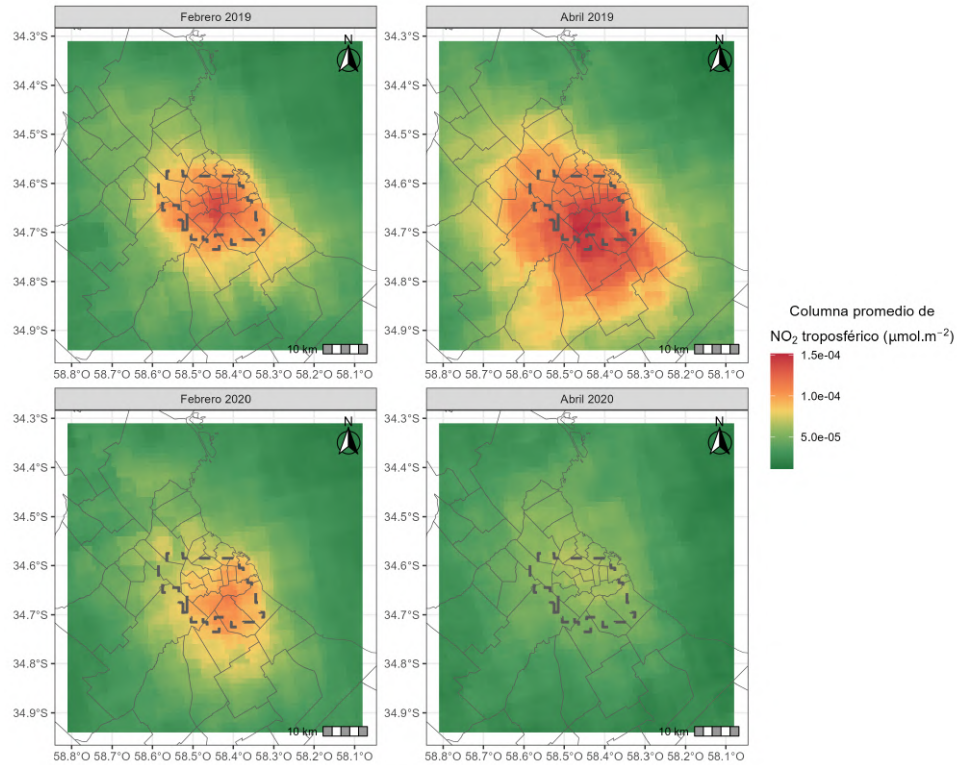


Fig. 3.6: Mapas de la columna de NO_2 troposférico promedio mensual para los febreros y abril 2019 y 2020

La figura 3.7 destaca el comportamiento semanal de la serie de tiempo de la columna de NO_2 troposférico. Se pueden observar valores más altos en los días de semana, especialmente en los jueves y viernes, y valores más bajos los sábados y domingos.

En las Figuras 3.8 y 3.9, se presentan dos autocorrelogramas de esta serie de tiempo. En el primero, los lags son diarios, y pueden observarse picos de correlación significativas cercanos a lags múltiplos de 7 días. En el segundo, se autocorrelacionó las medias semanales de la serie de tiempo. En este se observan correlaciones significativas únicamente a lags menores a 9 semanas, sin embargo, se destacan máximos locales cercanos a lags múltiplos de 50 semanas (aproximadamente un año) y valles cercanos a las semanas 25-30 y 75-80 (aproximadamente un medio año y un año y medio respectivamente). En cuanto al análisis del dominio de la frecuencia (Figura 3.10), se aprecian resultados similares, con

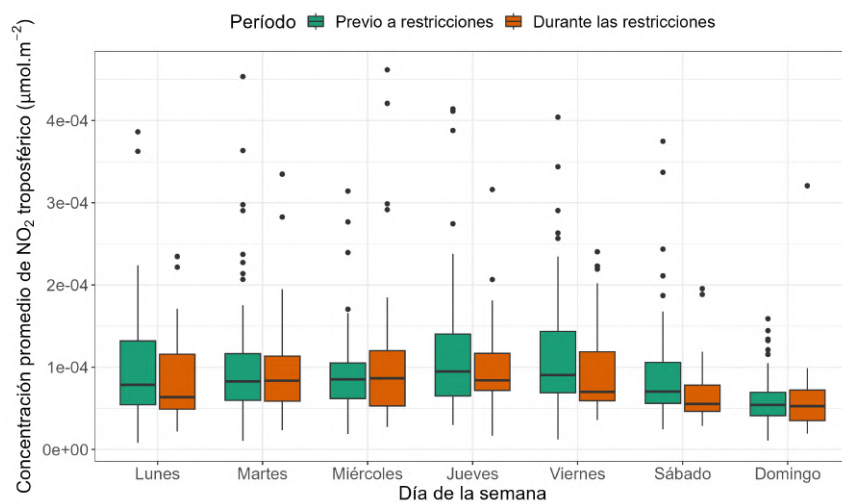


Fig. 3.7: Boxplot de la columna media diaria observada de NO_2 troposférico según el día de la semana y su período en relación a las restricciones.

picos cercanos a los 7 y 28 días aunque la señal es apreciablemente más ruidosa que aquella del conteo vehicular.

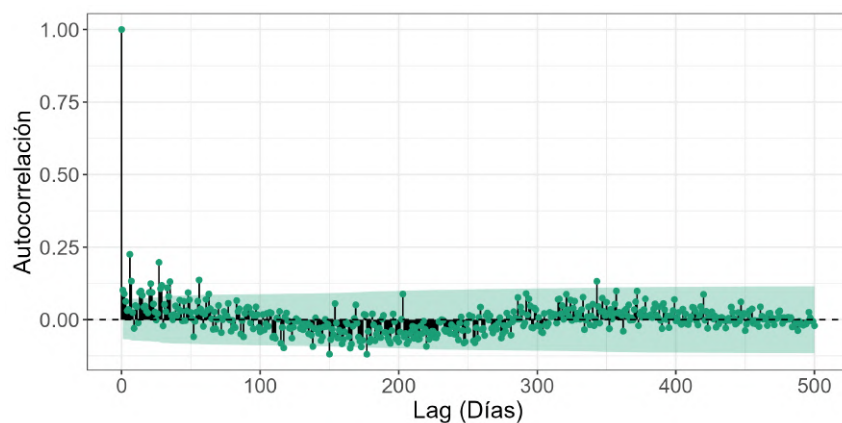


Fig. 3.8: Autocorrelograma de la serie de tiempo de la columna diaria de NO_2 troposférico. Los puntos que se encuentran fuera del área sombreada representan correlaciones significativas.

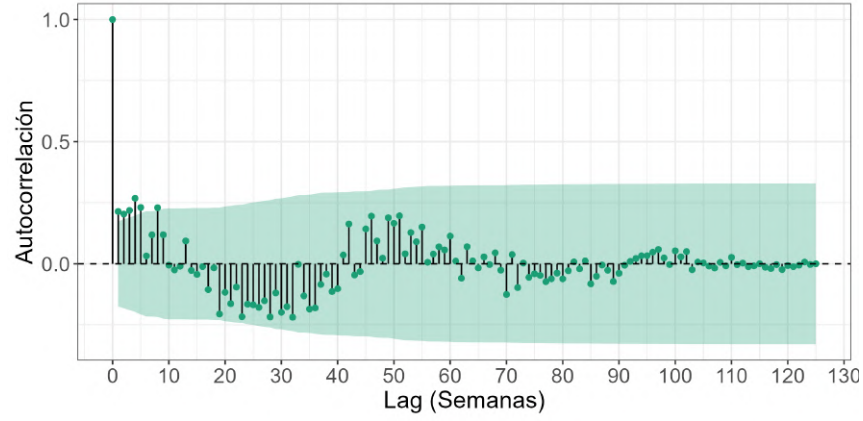


Fig. 3.9: Autocorrelograma de la serie de tiempo del promedio semanal de la columna diaria de NO_2 troposférico. Los puntos que se encuentran fuera del área sombreada representan correlaciones significativas

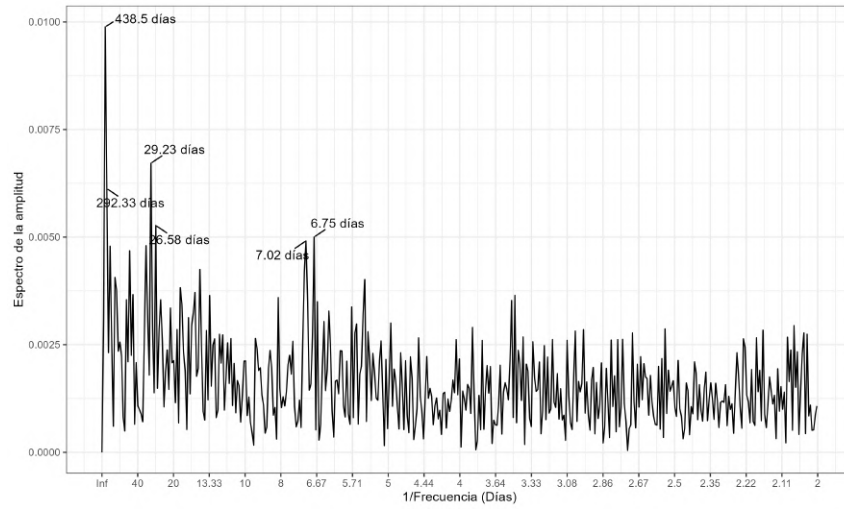


Fig. 3.10: Espectro de la amplitud del dominio de la frecuencia de la serie temporal de la columna diaria de NO_2

Variables meteorológicas

En cuanto a las variables meteorológicas, T_h presentó un comportamiento esperable (Figura 3.11), con máximos diarios cercanos a las 15 Hs, mínimos diarios cercanos a las 6 Hs máximos anuales a fin del verano y mínimos anuales en el invierno. Se observaron precipitaciones en 29.4 % de los días entre el 2019 y 2020, con un promedio de 4.78 días seguidos sin lluvia y 1.99 días seguidos con lluvia.

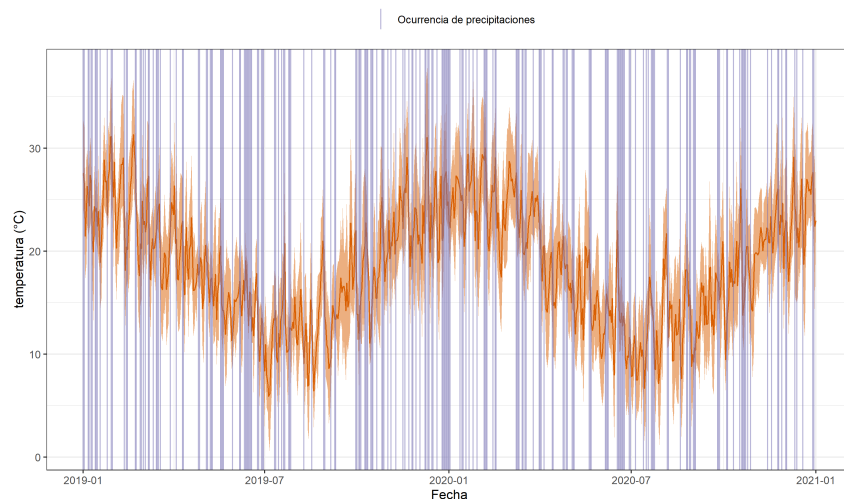


Fig. 3.11: Variación temporal de la temperatura y ocurrencia de precipitaciones diarias. La línea más oscura representa la temperatura diaria promedio, el sombreado que lo rodea rango de temperatura máxima y mínima del día y las líneas verticales azules a la ocurrencia de precipitaciones en el día.

Cruzamiento de variables

Se realizó un cruzamiento entre tres variables: columna de NO_2 troposférico, T_d y V_d (Figura 3.12). En ella se observa un gradiente de la columna de NO_2 , que parte de valores bajos a temperaturas altas y a baja cantidad de conteos vehiculares y aumenta cuando la temperatura baja y la cantidad de vehículos contados es mayor.

En la figura 3.13 se muestran distintos valores de correlación de Spearman obtenidos al correlacionar tres conversiones (identidad, logaritmo en base 10 y raíz cuadrada) de la variable columna de NO_2 troposférico frente a V_d (conteo vehicular diario), IV_d (intensidad del viento), T_d (temperatura media diaria) y las temperaturas mínimas y máximas diarias con las mismas conversiones y a distintos lags (0 a 30 días). Se pudo observar que, entre el NO_2 y V_d , el valor absoluto más alto de la correlación se obtuvo al correlacionar $Log_{10}(NO_2)$ con $Log_{10}(V_d)$ a un lag de 0 días ($\rho = 0,29$). Para el caso de IV_d , la correlación más importante se observó a 0 días de lag y entre $\sqrt{NO_2}$ con $Log_{10}(IV_d)$ ($\rho = -0,53$). Por último, todas las variables relacionadas a las temperaturas mostraron un pico de correlación con $\sqrt{NO_2}$ todas bajo una transformación logarítmica a un lag de dos días. Si bien, para la mayoría de los casos las correlaciones más importantes fueron observadas con $\sqrt{NO_2}$, aquellas calculadas para $Log_{10}(NO_2)$ a mismos lags son muy similares por lo que esta última transformación fue utilizada en los posteriores modelos.

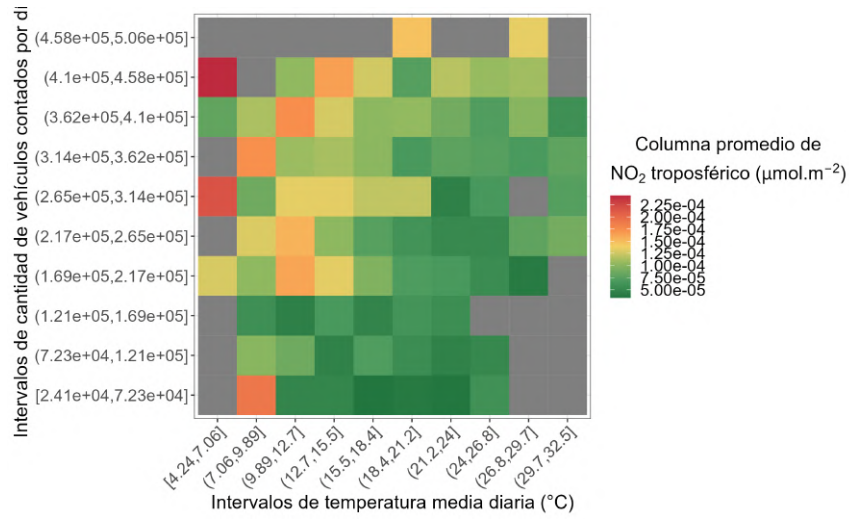


Fig. 3.12: Mapa de calor de la columna promedio de NO_2 troposférico en función de intervalos equiespaciados de V_d y de T_d . Los valores en gris son combinaciones no observadas de las tres variables

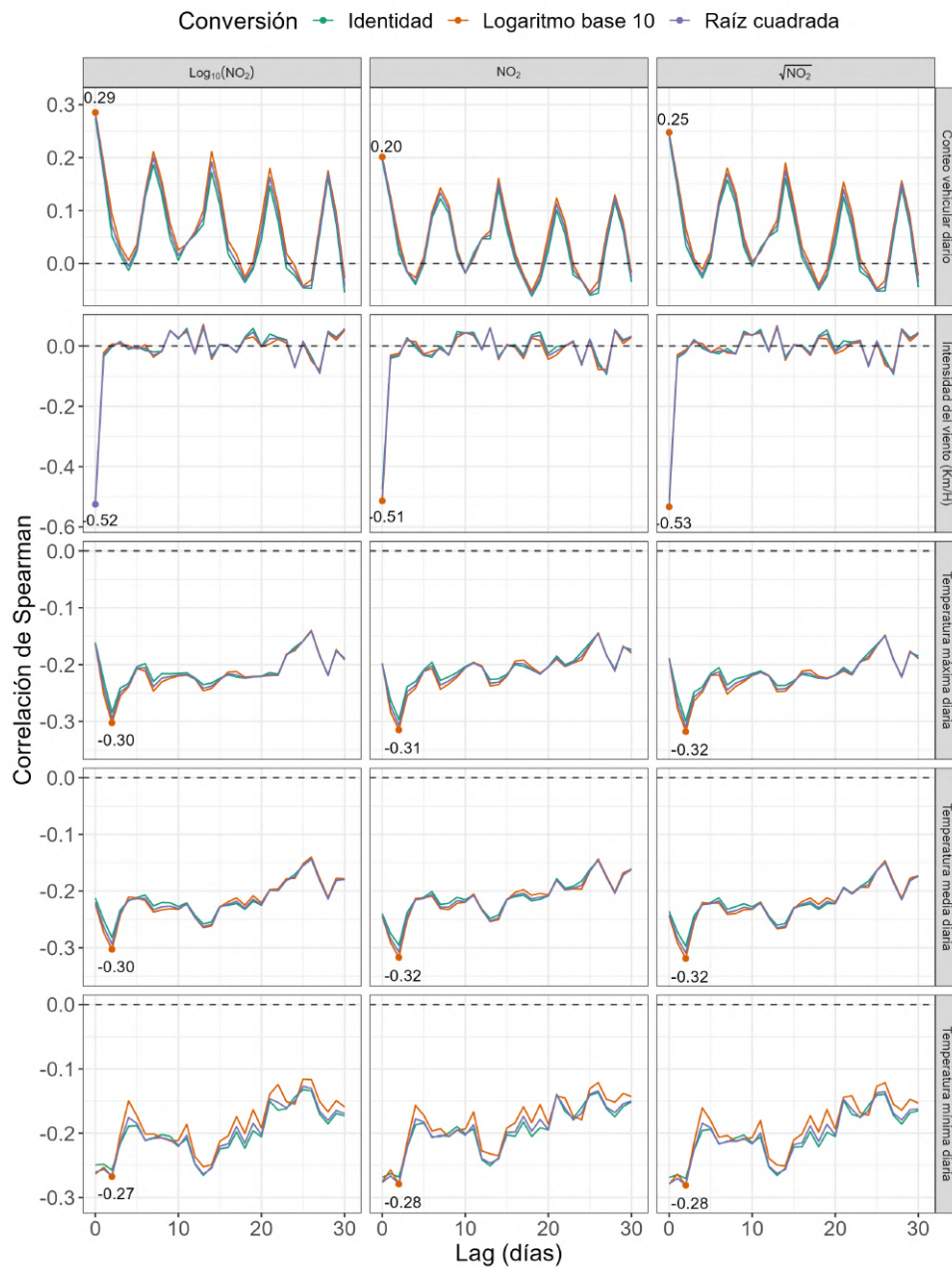


Fig. 3.13: Correlaciones de Spearman calculadas entre la columna de NO_2 troposférico (con distintas conversiones) contra distintas posibles variables explicativas (con distintas conversiones) a distintos lags y sus valores absolutos máximos.

3.0.2. Cuantificación de la reducción de la columna de NO_2

Estimación del conteo vehicular durante el período de restricción al movimiento si no lo hubiese habido

La validación cruzada a 30 días en el futuro del Modelo 1 con los hiperparámetros óptimos presentó un RMSE promedio de 0.227. El resultado de las predicciones del Modelo 1, junto con el conteo vehicular observado para la semana en la cual se instauraron las restricciones al movimiento se presentan en la Figura 3.14, donde se puede observar que las predicciones durante el período previo a las restricciones al movimiento son cercanas a los valores observados, mientras que, las predicciones son mucho mayores durante los primeros meses de las restricciones. En la Figura 3.15 se destacan predicciones inferiores a los valores observados a valores de conteo de vehículos bajos y el caso contrario para valores de conteos de vehículos altos. En cuanto a la importancia de las variables adicionales, sólo OP_h fue detectada como importante (intervalo de credibilidad = $[-0.021767; -0.010471]$), mientras que T_h no lo fue (intervalo de credibilidad = $[-0.000102; 0.000849]$).

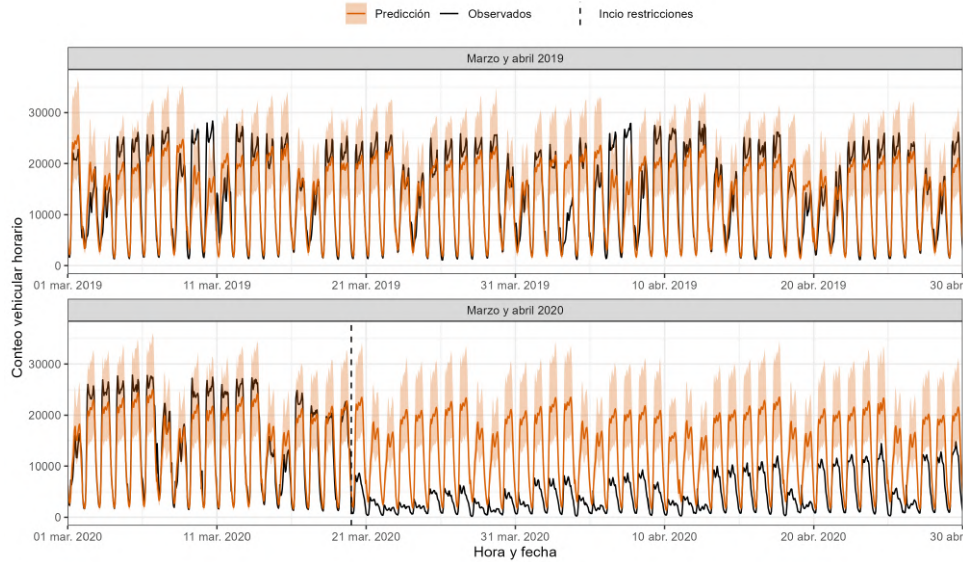


Fig. 3.14: Conteos vehiculares horarios observados y predichos por el Modelo 1 con su intervalo de incerteza durante los meses de marzo y abril de 2019 (arriba) y 2020 (abajo).

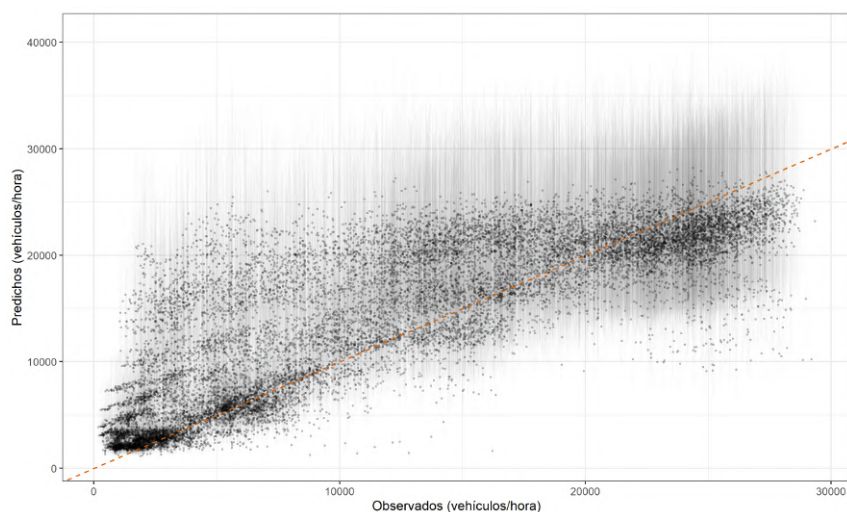


Fig. 3.15: Conteos vehiculares horarios observados vs. predichos por el Modelo 1 y sus intervalos de incerteza.

Estimación de la columna de NO_2 durante el período de restricciones si no lo hubiese habido.

El algoritmo que presentó el mejor rendimiento según todas las métricas utilizadas fue KNN. En la Figura 3.16 se muestran los resultados de dichas comparaciones donde se puede observar que el algoritmo KNN presentó los valores más bajos de RMSE ($2,84 \times 10^{-5}$) y MAPE (0,21) y el valor más alto de R^2 (0,41). Los resultados de las predicciones sobre el conjunto de Held-Out están plasmadas en la Figura 3.17.

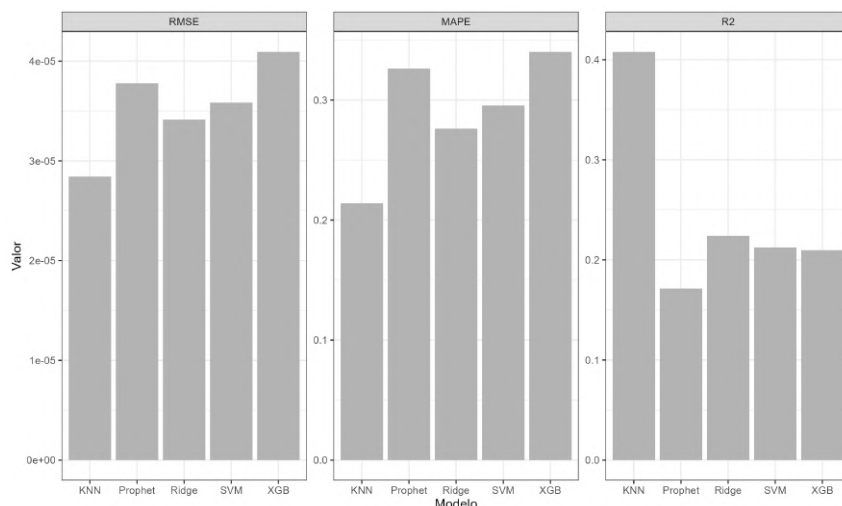


Fig. 3.16: Rendimiento comparado por tres métricas de los distintos algoritmos optimizados para la predicción de la columna de NO_2 troposférico.

La estimación de NO_2SR utilizando el algoritmo KNN con hiperparámetros optimizados se visualizó en la Figura 3.18. Los valores estimados para NO_2SR parecieran ser similares a los observados antes del inicio de las restricciones y estos se bifurcan luego

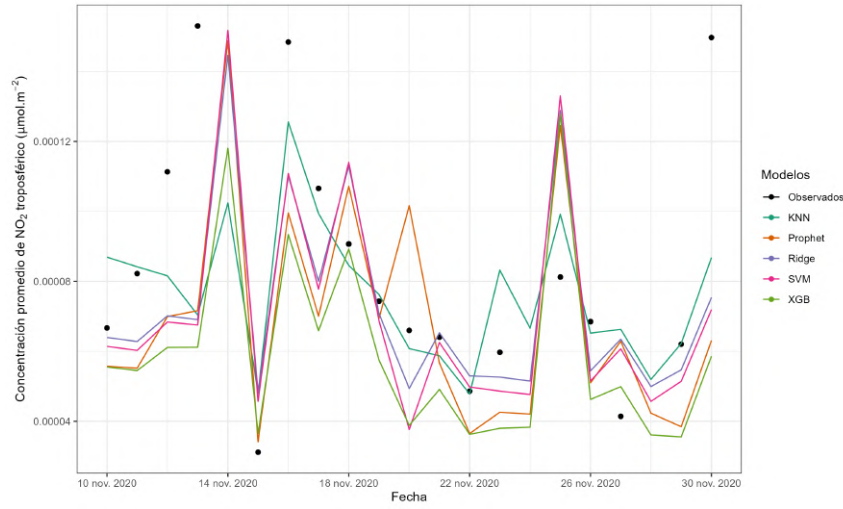


Fig. 3.17: Predicciones del NO_2 troposférico sobre el conjunto de Held-Out por los distintos algoritmos evaluados.

del 20 de marzo. Sin embargo, la variabilidad de la columna de NO_2 observado es mucho mayor a la estimada por NO_2SR como se muestra en la Figura 3.19: previo al inicio de las restricciones, a valores bajos y medios de la columna de NO_2 , los predichos son similares a los observados. Mientras que los valores observados altos son subestimados por el modelo.

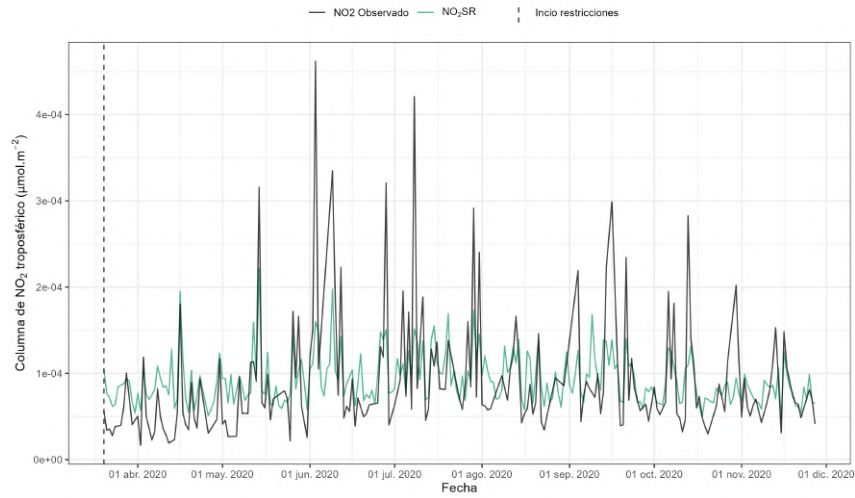


Fig. 3.18: Columna de NO_2 observada y predicha por el Modelo 2 durante el período de restricciones al movimiento.

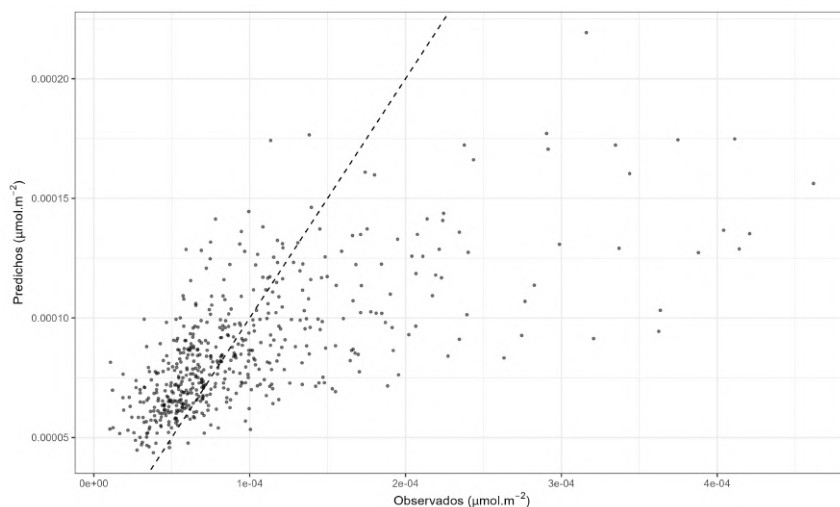


Fig. 3.19: Observados vs. predichos por el Modelo 2 previo al inicio del período de restricciones.

Comparaciones de las columnas de NO_2

En cuanto a la cuantificación de la reducción de la columna de NO_2 , se estimó que en el período de restricciones, esta se redujo un 9 %, aunque esta reducción no fue homogénea, siendo mayor en los primeros meses; siendo abril el primero completo en restricciones, este presentó una reducción cercana al 39 %, como puede observarse en la Figura 3.20. Además, al comparar las estimaciones de la reducción promedio de la columna de NO_2 y del conteo vehicular horario (Tabla 3.1), se observó que la del conteo vehicular fue más importante que la del NO_2 en todos los meses con restricciones.

Tab. 3.1: Comparación de las reducciones promedio estimadas de la columna de NO_2 en distintos períodos

Período	Reducción estimada NO_2 (%)	Reducción estimada conteo vehicular (%)
mar. 2020	2.68	26.27
abr. 2020	38.66	61.37
may. 2020	16.01	50.27
jun. 2020	-7.21	45.68
jul. 2020	-8.10	49.79
ago. 2020	21.80	40.43
sep. 2020	-0.66	33.69
oct. 2020	-3.12	25.84
nov. 2020	4.62	18.87
Total Período	9.26	44.85

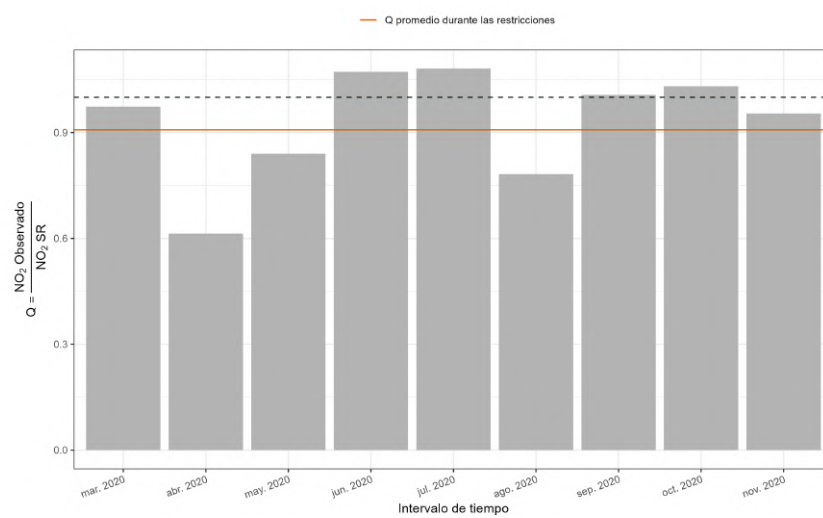


Fig. 3.20: Cociente Q calculado para distintos períodos de tiempo.

4. DISCUSIÓN Y CONCLUSIONES

En este trabajo se estimó la disminución del flujo vehicular y de la columna de NO_2 desde el inicio de las restricciones a la circulación impuesta el 20 de marzo de 2020 hasta su finalización el 6 de noviembre del mismo año, establecidas a fin de mitigar la transmisión del virus SARS-CoV-2.

El conteo vehicular en las casillas de peajes mostró una dinámica temporal diaria y semanal muy similar a la observada en otras grandes ciudades del mundo [47, 48, 49, 50, 51]. Estas dinámicas temporales presentaron características periódicas diarias y semanales marcadas que podrían ser explicadas por la actividad laboral de la ciudad. Esta relación se encuentra plasmada en la reducción de la circulación vehicular los fines de semana y feriados en comparación a los días hábiles con picos de circulación previos y posteriores al inicio y al final, respectivamente, del horario laboral típico (de 9 a 18 Hs). Esta inferencia está en concordancia con el Ministerio de Empleo, Trabajo y Seguridad Social de la Nación que estimó que en el 2018, en Argentina, el 30.7 % de las personas trabaja los fines de semana solo el 8.9 % de tiene un trabajo nocturno [52] y es consistente con lo observado en estaciones de referencia [53]. Si bien la instauración de las restricciones tuvo un impacto fuerte e inmediato en la intensidad de la cantidad de vehículos en circulación, las variaciones periódicas diarias y semanales observadas se mantuvieron similares a las ocurridas antes del 20 de marzo 2020, sin restricciones. Esta observación ya fue descripta en Chennai, India, donde además demostró que la circulación vehicular estuvo asociada a la severidad y duración de las restricciones impuestas [54].

El Modelo 1 pudo representar las variaciones del conteo vehicular a partir de las distintas periodicidades temporales y de variables meteorológicas, estimando una reducción del 45 % en todo el período. La asociación entre las variables meteorológicas y la circulación vehicular fue descripta anteriormente aunque con distintas conclusiones. En cuanto a las precipitaciones, en concordancia con Goodwin et al. 2002 [55] y Essien et al. 2018 [56], el Modelo 1 detectó una asociación negativa entre el conteo vehicular y la ocurrencia de precipitaciones. Además, en línea con las conclusiones de Essien et al. 2018 [56], no se observó un efecto importante de la T_h sobre el conteo vehicular. Si bien estos trabajos fueron realizados con casi 20 años de diferencia en distintas ciudades (varias de EE.UU y Manchester, Reino Unido respectivamente, con distintos regímenes meteorológicos, índices de desarrollo humano y diferencias culturales) la consistencia entre los resultados de tales estudios y el presente trabajo refuerza la robustez de las conclusiones desarrolladas.

En cuanto al Modelo 2, se lograron realizar predicciones para la columna de NO_2 troposférico sin la ocurrencia de las restricciones al movimiento. La concentración de NO_2 está determinada tanto por factores naturales como antropogénicos [57, 23, 22, 7] en los cuales, en línea con lo observado en este trabajo, la columna de NO_2 presentó una asociación positiva con la circulación vehicular y una negativa con la temperatura y la intensidad del viento.

En cuanto a la reducción de la concentración de NO_2 durante las restricciones al movimiento, se estimó una reducción promedio del 39 % durante el primer mes completo con restricciones (abril) y del 9 % durante todo el período. Represa et al. 2021 [58] estimaron una reducción del 30 % en CABA entre el 20/03/2020 y el 04/05/2020 utilizando el mismo satélite pero con otra metodología de estimación. Para ese mismo período, el presente

trabajo estima una reducción del 22 %. Por otro lado, Liu et al. 2021 [59] estimaron una caída inicial del 39 % en la ciudad durante un período de tiempo similar al de Represa et al. 2021. Estas dos publicaciones describen una reducción de la columna de NO_2 superior a la estimada por el presente trabajo. Esto podría deberse a las diferencias metodológicas utilizadas para la estimación. En ambas publicaciones, la reducción fue calculada comparando la columna de NO_2 del 2019 y 2020 para un mismo intervalo de fechas (es decir que lo observado en 2019 fue considerado como una estimación directa de lo que hubiese ocurrido en condiciones normales en 2020). En este trabajo se tomaron en consideración la influencia de otros factores para estimar la concentración de la columna de NO_2 en el período en condiciones normales.

Sólo se encontró un trabajo [57] en la bibliografía que consideró tales condiciones a fin de realizar una predicción de la concentración de NO_2 en CABA durante las restricciones al movimiento. Briganti 2021 estimó, por medio de un modelo de machine learning, que la reducción de la concentración de NO_2 en el mes de abril 2020 fue de 20-26 %. Si bien se implementó una metodología similar en el mismo período de tiempo y en la misma zona, la principal diferencia con el presente trabajo radica en el tipo los datos de NO_2 utilizados: Briganti 2021 utilizó mediciones directas de la concentración de NO_2 en superficie medidas en tres estaciones dentro de CABA. Dada la gran variabilidad espacio-temporal de la concentración de NO_2 en la atmósfera urbana, es de esperar que hayan diferencias entre ambas estimaciones. Las mediciones en superficie permiten medir directamente la calidad del aire a nivel del suelo, donde la contaminación atmosférica tiene efectos sobre la salud. Por su parte, las imágenes satelitales proporcionan una estimación la concentración de NO_2 en la columna de aire [60].

Este trabajo presenta algunas limitaciones. En primer lugar, el Modelo 2 fue entrenado sólo con dos años de información retrospectiva, en los cuales uno de ellos presentó condiciones muy particulares en cuanto a la circulación vehicular. Estas condiciones no permiten estimar fielmente una tendencia anual. Para hacerlo, es necesario utilizar mayor cantidad de años, sin embargo, esto no fue posible dado que el satélite TROPOMI fue lanzado a fines del 2018, por lo que la información utilizada en este trabajo fue toda la información previa existente de este satélite. Por otro lado, existen otros satélites más longevos que también miden la columna de NO_2 troposférico, como el Ozone Monitoring Instrument (OMI/AURA) [61] con información disponible desde el 2004. No obstante, el tamaño del píxel de estas imágenes es de $0.25^\circ \times 0.25^\circ$ que, siendo 624 veces mayor a la de TROPOMI ($0.01^\circ \times 0.01^\circ$), es demasiado grande para utilizarlo en zonas relativamente chicas, como lo es la extensión de CABA, que ocuparía parcialmente sólo 5 píxeles de OMI.

En segundo lugar, en la bibliografía ya mencionada, se describen otros factores naturales que pueden afectar la concentración del NO_2 como la radiación solar incidente. Estos factores no fueron incluidos en el presente trabajo ya que no se pudieron obtener los valores históricos necesarios.

En tercer lugar, en este trabajo se utilizó el conteo vehicular como indicador de la circulación vehicular en la ciudad y, a su vez, como indicador de intensidad de emisión del NO_2 . Sin embargo, la circulación vehicular no es la única fuente de emisión del gas. Otra importante fuente de emisión en CABA es la actividad industrial. Esta no fue incluida en el análisis ya que no se encontró información disponible en el período de estudio.

Por último, las mediciones de la intensidad del viento pueden estar sesgadas ya que el anemómetro situado en la estación meteorológica 'Observatorio' se encuentra cerca de árboles y esto puede afectar las mediciones en determinadas condiciones.

No obstante, las observaciones y estimaciones obtenidas en este trabajo son consistentes casi en su totalidad con las disponibles en la bibliografía relacionada.

En conclusión, las restricciones al movimiento en CABA tuvieron un objetivo principal que fue disminuir la transmisión del virus SARS-CoV-2 entre los habitantes. Estas medidas generaron diversos efectos secundarios, entre ellos la disminución de contaminantes atmosféricos. En este trabajo se cuantificó la reducción de la concentración de la columna de NO_2 sobre CABA durante las restricciones al movimiento del 2020. Se estimó que durante el primer mes de restricciones, la reducción fue del 39 % mientras que durante todo el período, la reducción fue del 9 %.

BIBLIOGRAFÍA

- [1] Corey M Peak et al. "Comparing nonpharmaceutical interventions for containing emerging epidemics". En: *Proceedings of the National Academy of Sciences* 114.15 (2017), págs. 4023-4028.
- [2] Presidencia de la Nación. "Aislamiento social preventivo y obligatorio". En: *Boletín Oficial de la República Argentina* (2020), págs. 1-7. URL: <https://www.boletinoficial.gob.ar/detalleAviso/primera/227042/20200320>.
- [3] Maria Inés de Casas et al. "Impacto del aislamiento social preventivo y obligatorio (ASPO) por la epidemia de COVID-19 sobre la calidad del aire: CABA, Marzo-Abril-Mayo 2020". En: *Impacto del aislamiento social preventivo y obligatorio (ASPO) por la epidemia de COVID-19 sobre la calidad del aire: CABA, Marzo-Abril-Mayo 2020*. 2020, a-f.
- [4] Lidia Ana Otero et al. "Mediciones de Pm10 y No2 durante los primeros meses de la Cuarentena de Covid-19 en el 2020 en Ciudad De Buenos Aires, Argentina". En: (2021).
- [5] Paula S Castesana et al. "Temporal and spatial variability of nitrous oxide emissions from agriculture in Argentina". En: *Carbon Management* 11.3 (2020), págs. 251-263.
- [6] S Enrique Puliafito et al. "High-resolution atmospheric emission inventory of the argentine energy sector. Comparison with edgar global emission database". En: *Helvion* 3.12 (2017), e00489.
- [7] Laura Venegas, Nicolás Mazzeo y Andrea Pineda Rojas. "Evaluation of an Emission Inventory and Air Pollution in the Metropolitan Area of Buenos Aires". En: jun. de 2011. ISBN: 978-953-307-307-1. DOI: 10.5772/18767.
- [8] Ramamohana Reddy Appannagari. "Environmental pollution causes and consequences: a study". En: *North Asian International Research Journal of Social Science and Humanities* 3.8 (2017), págs. 151-161.
- [9] Regina Ruckerl et al. "Air pollution and markers of inflammation and coagulation in patients with coronary heart disease". En: *American journal of respiratory and critical care medicine* 173.4 (2006), págs. 432-441.
- [10] Jonathan O Anderson, Josef G Thundiyil y Andrew Stolbach. "Clearing the air: a review of the effects of particulate matter air pollution on human health". En: *Journal of medical toxicology* 8.2 (2012), págs. 166-175.
- [11] Richard W Atkinson et al. "Long-term concentrations of nitrogen dioxide and mortality: a meta-analysis of cohort studies". En: *Epidemiology (Cambridge, Mass.)* 29.4 (2018), pág. 460.
- [12] Dan L Crouse et al. "Ambient PM2.5, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC)". En: *Environmental health perspectives* 123.11 (2015), págs. 1180-1186.
- [13] Gerard Hoek et al. "Long-term air pollution exposure and cardio-respiratory mortality: a review". En: *Environmental health* 12.1 (2013), págs. 1-16.

- [14] Ghassan B Hamra et al. "Lung cancer and exposure to nitrogen dioxide and traffic: a systematic review and meta-analysis". En: *Environmental health perspectives* 123.11 (2015), págs. 1107-1112.
- [15] Monica Chiusolo et al. "Short-term effects of nitrogen dioxide on mortality and susceptibility factors in 10 Italian cities: the EpiAir study". En: *Environmental health perspectives* 119.9 (2011), págs. 1233-1238.
- [16] World Health Organization et al. *WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization, 2021.
- [17] Agencia de Protección Ambiental (APRA). "Estándares de calidad del Aire - Métodos de medición de calidad del aire - Implementación de carácter progresivo". En: *Resolución 68 2021 Agencia de Protección Ambiental, Ley 2628* (2021). URL: <https://boletinoficial.buenosaires.gob.ar/normativaba/norma/550171>.
- [18] Daniel Vallero. *Fundamentals of air pollution*. Academic press, 2014.
- [19] Machenzie L Davis y Susan J Coautor Masten. *Ingeniería y ciencias ambientales*. 628 D2951i Ej. 1. McGraw-Hill, 2005.
- [20] Viral Shah et al. "Effect of changing NO_x lifetime on the seasonality and long-term trends of satellite-observed tropospheric NO₂ columns over China". En: *Atmospheric Chemistry and Physics* 20.3 (2020), págs. 1483-1495.
- [21] Carlos Silveira, Joana Ferreira y Ana Isabel Miranda. "The challenges of air quality modelling when crossing multiple spatial scales". En: *Air Quality, Atmosphere & Health* 12.9 (2019), págs. 1003-1017.
- [22] Andrea L. Pineda Rojas y Laura E. Venegas. "Upgrade of the DAUMOD atmospheric dispersion model to estimate urban background NO₂ concentrations". En: *Atmospheric Research* 120-121 (2013), págs. 147-154. ISSN: 0169-8095. DOI: <https://doi.org/10.1016/j.atmosres.2012.08.010>. URL: <https://www.sciencedirect.com/science/article/pii/S0169809512002773>.
- [23] Laura E. Venegas, Nicolás A. Mazzeo y Mariana C. Dezzutti. "A simple model for calculating air pollution within street canyons". En: *Atmospheric Environment* 87 (2014), págs. 77-86. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2014.01.005>. URL: <https://www.sciencedirect.com/science/article/pii/S1352231014000144>.
- [24] Yves Rybarczyk y Rasa Zalakeviciute. "Machine learning approaches for outdoor air quality modelling: A systematic review". En: *Applied Sciences* 8.12 (2018), pág. 2570.
- [25] David Topping et al. "Evaluating the use of Facebook's Prophet model v0. 6 in forecasting concentrations of NO₂ at single sites across the UK and in response to the COVID-19 lockdown in Manchester, England". En: *Geoscientific Model Development Discussions* (2020), págs. 1-22.
- [26] M.W. Gardner y S.R. Dorling. "Neural network modelling and prediction of hourly NO_x and NO₂ concentrations in urban air in London". En: *Atmospheric Environment* 33.5 (1999), págs. 709-719. ISSN: 1352-2310. DOI: [https://doi.org/10.1016/S1352-2310\(98\)00230-1](https://doi.org/10.1016/S1352-2310(98)00230-1). URL: <https://www.sciencedirect.com/science/article/pii/S1352231098002301>.

-
- [27] Vera van Zoest et al. "Spatio-temporal regression kriging for modelling urban NO₂ concentrations". En: *International journal of geographical information science* 34.5 (2019), págs. 851-865.
- [28] Jiani Yang et al. "From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model". En: *Proceedings of the National Academy of Sciences* 118.26 (2021), e2102705118. DOI: 10.1073/pnas.2102705118. eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.2102705118>. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.2102705118>.
- [29] Giancarlo Ciarelli et al. "Modeling the effect of reduced traffic due to COVID-19 measures on air quality using a chemical transport model: impacts on the Po Valley and the Swiss Plateau regions". En: *Environmental Science: Atmospheres* 1.5 (2021), págs. 228-240.
- [30] Riccardo Rossi, Riccardo Ceccato y Massimiliano Gastaldi. "Effect of road traffic on air pollution. Experimental evidence from COVID-19 lockdown". En: *Sustainability* 12.21 (2020), pág. 8984.
- [31] C. A. Keller et al. "Global impact of COVID-19 restrictions on the surface concentrations of nitrogen dioxide and ozone". En: *Atmospheric Chemistry and Physics* 21.5 (2021), págs. 3555-3592. DOI: 10.5194/acp-21-3555-2021. URL: <https://acp.copernicus.org/articles/21/3555/2021/>.
- [32] Jiani Yang et al. "From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model". En: *Proceedings of the National Academy of Sciences* 118.26 (2021), e2102705118.
- [33] Mario Lovrić et al. "Understanding the true effects of the COVID-19 lockdown on air pollution by means of machine learning". En: *Environmental Pollution* 274 (2021), pág. 115900.
- [34] Stuart K Grange et al. "COVID-19 lockdowns highlight a risk of increasing ozone pollution in European urban areas". En: *Atmospheric Chemistry and Physics* 21.5 (2021), págs. 4169-4185.
- [35] Marlon Brancher. "Increased ozone pollution alongside reduced nitrogen dioxide concentrations during Vienna's first COVID-19 lockdown: Significance for air quality management". En: *Environmental Pollution* 284 (2021), pág. 117153.
- [36] Hervé Petetin, Dene Bowdalo y Albert Soret. "Assessment of the impact of the Covid-19 lockdown on air pollution over Spain using machine learning". En: (2020).
- [37] Dimitris Akritidis et al. "Implications of COVID-19 Restriction Measures in Urban Air Quality of Thessaloniki, Greece: A Machine Learning Approach". En: *Atmosphere* 12.11 (2021), pág. 1500.
- [38] J.H van Geffen et al. "TROPOMI ATBD of the total and tropospheric NO₂ data products Jos van Geffen". En: 2.2.0 (2021), págs. 1-86. URL: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms>.
- [39] Thompson Stephan et al. "Machine learning analysis on the impacts of COVID-19 on India's renewable energy transitions and air quality". En: *Environmental Science and Pollution Research* (2022), págs. 1-23.

-
- [40] Sean J Taylor y Benjamin Letham. “Forecasting at scale”. En: *The American Statistician* 72.1 (2018), págs. 37-45.
- [41] TROPOMI. *Level 2 Products*. URL: <http://www.tropomi.eu/data-products/level-2-products>.
- [42] World Meteorological Organization. *Entry: FM 12 - SYNOP*. URL: http://codes.wmo.int/wmdr/DataFormat/_FM-12-synop.
- [43] Autopistas Urbanas (AUSA). Secretaria de Transporte y Obras Publicas. Jefatura de Gabinete de Ministros. *Flujo Vehicular por Unidades de Peaje AUSA*. URL: <https://data.buenosaires.gob.ar/dataset/flujo-vehicular-por-unidades-peaje-ausa>.
- [44] Ministerio de Educación de la Nación Argentina. *Feriatos 2019: Conocé el calendario completo*. 2019. URL: <https://www.argentina.gob.ar/noticias/feriatos-2019-conoce-el-calendario-completo>.
- [45] Ministerio del Interior de la Nación Argentina. “Feriatos nacionales 2020”. En: (2020). URL: <https://www.argentina.gob.ar/interior/feriatos-nacionales-2020>.
- [46] Christophorus Beneditto Aditya Satrio et al. “Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET”. En: *Procedia Computer Science* 179 (2021), págs. 524-532.
- [47] Li Chen et al. “Estimating vehicle and pedestrian activity from town and city traffic cameras”. En: *Sensors* 21.13 (2021), pág. 4564.
- [48] Tom Thomas, Wendy Weijermars y Eric van Berkum. “Variations in urban traffic volumes”. En: *European Journal of Transport and Infrastructure Research* 8.3 (2008).
- [49] Billy M Williams y Lester A Hoel. “Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results”. En: *Journal of transportation engineering* 129.6 (2003), págs. 664-672.
- [50] Wei Huang et al. “Real-time prediction of seasonal heteroscedasticity in vehicular traffic flow series”. En: *IEEE Transactions on Intelligent Transportation Systems* 19.10 (2017), págs. 3170-3180.
- [51] Marco Rapelli, Claudio Casetti y Giandomenico Gagliardi. “Vehicular traffic simulation in the city of turin from raw data”. En: *IEEE Transactions on Mobile Computing* (2021).
- [52] Ministerio de Producción y Trabajo de la Nación Argentina. *Encuesta Nacional a trabajadores sobre Condiciones de Empleo, Trabajo, Salud y Seguridad*. Inf. téc. 2018. URL: https://www.trabajo.gob.ar/downloads/estadisticas/ecetss/ecetss_informe.pdf.
- [53] Andrea L. Pineda Rojas et al. “High PM10 concentrations in the city of Buenos Aires and their relationship with meteorological conditions”. En: *Atmospheric Environment* 241 (2020), pág. 117773. ISSN: 1352-2310. DOI: <https://doi.org/10.1016/j.atmosenv.2020.117773>. URL: <https://www.sciencedirect.com/science/article/pii/S1352231020305057>.
- [54] Satya S Patra, Bhargava Rama Chilukuri y Lelitha Vanajakshi. “Analysis of road traffic pattern changes due to activity restrictions during COVID-19 pandemic in Chennai”. En: *Transportation letters* 13.5-6 (2021), págs. 473-481.

-
- [55] Lynette C Goodwin. “Weather impacts on arterial traffic flow”. En: *Mitretek systems inc* 24 (2002), págs. 4-8.
- [56] Aniekan Essien et al. “The impact of rainfall and temperature on peak and off-peak urban traffic”. En: *International Conference on Database and Expert Systems Applications*. Springer. 2018, págs. 399-407.
- [57] María Lucía Briganti. “Impacto del confinamiento debido a la COVID-19 en la concentración de contaminantes medidos en la ciudad de Buenos Aires”. Tesis de licenciatura en Ciencias de la Atmósfera. Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y de los Océanos, 2021.
- [58] Natacha S Represa et al. “Atmospheric Pollutants Assessment during the COVID-19 Lockdown Using Remote Sensing and Ground-based Measurements in Buenos Aires, Argentina”. En: *Aerosol and Air Quality Research* 21.3 (2021), pág. 200486.
- [59] Song Liu et al. “Nitrogen dioxide decline and rebound observed by GOME-2 and TROPOMI during COVID-19 pandemic”. En: *Air Quality, Atmosphere & Health* 14.11 (2021), págs. 1737-1755.
- [60] Seohyun Choi Susan M. O'Neill Mohammad Z. Al-Hamdan Aaron Van Donkelaar Randall V. Martin Xiaomeng Jin Arlene M. Fiore Daven K. Henze Forrest Lacey Patrick L. Kinney Frank Freedman Narasimhan K. Larkin Yufei Zou James T. Kelly Minghui Diao Tracey Holloway y Ambarish Vaidyanathan. “Methods, availability, and applications of PM2.5 exposure estimates derived from ground measurements, satellite, and atmospheric models”. En: *Journal of the Air & Waste Management Association* 69.12 (2019). PMID: 31526242, págs. 1391-1414. DOI: 10.1080/10962247.2019.1668498. eprint: <https://doi.org/10.1080/10962247.2019.1668498>. URL: <https://doi.org/10.1080/10962247.2019.1668498>.
- [61] NA Krotkov et al. *the OMI core team: OMI/Aura NO2 Cloud-Screened Total and Tropospheric Column L3 Global Gridded 0.25 degree× 0.25 degree V3, NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC)*. 2019.