



FACULTAD DE INGENIERÍA Y CIENCIAS APLICADAS

PREDICCIÓN DE COMPORTAMIENTO DE PM2.5 DEL ÁREA URBANA DE  
QUITO CON MÉTODOS DE INTELIGENCIA ARTIFICIAL.

AUTOR

Marco Gabriel Bastidas Puentestar

AÑO

2019



FACULTAD DE INGENIERÍA Y CIENCIAS APLICADAS

PREDICCIÓN DE COMPORTAMIENTO DE PM2.5 DEL ÁREA URBANA DE QUITO  
CON MÉTODOS DE INTELIGENCIA ARTIFICIAL.

Trabajo de Titulación presentado en conformidad con los requisitos  
establecidos para optar por el título de Ingeniero Ambiental en Prevención y  
Remediación.

Profesora Guía  
PhD. Rasa Zalakeviciute

Autor  
Marco Gabriel Bastidas Puentestar

Año  
2019

## **DECLARACIÓN DEL PROFESOR GUÍA**

Declaro haber dirigido el trabajo, 'Predicción de comportamiento de PM2.5 del área urbana de Quito con métodos de inteligencia artificial.', a través de reuniones periódicas con el estudiante Marco Gabriel Bastidas Puentestar, en el semestre 201920, orientando sus conocimientos y competencias para un eficiente desarrollo del tema escogido y dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

---

Rasa Zalakeviciute  
Doctor of Engineering Science  
C.C.:1757076797

## **DECLARACIÓN DEL PROFESOR CORRECTOR**

Declaro haber revisado este trabajo, 'Predicción de comportamiento de PM2.5 del área urbana de Quito con métodos de inteligencia artificial.', del estudiante Marco Gabriel Bastidas Puentestar, en el semestre 201920, dando cumplimiento a todas las disposiciones vigentes que regulan los Trabajos de Titulación.

---

Santiago Daniel Piedra Burgos  
Master of Science in Hydro Science and Engineering  
C.C.: 1715384150

## **DECLARACIÓN DE AUTORÍA DEL ESTUDIANTE**

Declaro que este trabajo es original, de mi autoría, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales que protegen los derechos de autor vigentes.

---

Marco Gabriel Bastidas Puentestar

C.C.: 1726931064

## **AGRADECIMIENTOS**

Doy gracias a Dios, por su amistad, ayuda y bendiciones, al igual que su apoyo incondicional en mi vida y en esta carrera; a toda mi 'enorme' y bella familia por su amor, enseñanzas, ayuda y paciencia en todo momento; a mis docentes Pao, Ale, Rasa, Indi, Pavi, Santy, Dany, Sebas, Doc. Gualoto, y mis demás profes, por ser mis educadores y geniales amigos; y a mis amigos Nicole, Pepe, Mateo, Angie, Lizz, Verito, el grupo de los 'D', Mishu y Johan, con quienes sufrí, reí, me divertí, conté a lo largo de la carrera y aún son mi apoyo cuando lo necesito.

Mi sincera gratitud a todos ustedes.

## **Dedicatoria**

A mi buen Dios, a mi familia, a mis abuelitos, a mis amigos, a mis profes... por TODO, y...

A toda persona que desee aprender cómo funciona una pequeña parte del mundo que nuestro Dios creó (y que necesita urgente de nuestra ayuda para su recuperación); a quien quiera aprender un poquito de ciencia en base a evidencia, ayudar a construir un mundo mejor antes de que nuestro mundo se vuelva inhabitable, o el que quiera disfrutar con la famosísima tesis que no permite dormir en paz

Esta tesis es para ustedes :D.

## RESUMEN

Debido a la cantidad y gravedad de los diferentes efectos que puede producir el PM<sub>2.5</sub> el presente trabajo genera modelos de predicción para dicho contaminante en el área urbana de Quito, a partir de las variables disponibles en los registros históricos de la Secretaría del Ambiente, mediante métodos de inteligencia artificial (usando el programa WEKA) y con el método de interpolación de distancia inversa ponderada (en el programa QGIS). Tras la depuración y normalización de los datos se estableció modelos, mediante regresión lineal múltiple, con una correlación del 0.52 de aproximación a los registros; y árboles de aprendizaje, mediante el clasificador J48, para determinar si la concentración superaba a la mediana de los registros, obteniendo una clasificación de 64% de instancias correctas. Se registró como contaminantes claves para la regresión lineal el CO, NO<sub>2</sub>, Humedad Relativa y Temperatura Ambiental; mientras que los árboles fueron claves el CO, NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>, PM<sub>10</sub>, Humedad Relativa, Velocidad del Viento, Radiación Solar y la Hora. Los mapas interpolados a partir de los algoritmos indican que la mayor concentración del contaminante ocurre de 06:00 a 11:59 (Mañana), luego de 17:00 a 20:59 (Noche), en tercer lugar, de 21:00 a 05:59 (Madrugada) y finalmente de 12:00 a 16:59 (Tarde). Desafortunadamente, el método no ofrece la resolución adecuada para discriminar las variaciones de concentraciones altas y bajas, por lo que una alternativa para predecir con mayor exactitud la concentración del contaminante en cuestión es mediante la densidad vegetal presente, capaz de reducir el PM<sub>2.5</sub> mediante evapotranspiración; y la intensidad del tráfico y flujo vehicular según el tipo de calle, debido a que actualmente es la principal fuente de contaminación en el Área Urbana de Quito.



## ABSTRACT

Due to the quantity and severity of the different effects that PM<sub>2.5</sub> can generate, this work generates predictive models for this pollutant in the urban area of Quito, based on the variables available in the historical records of '*Secretaría del Ambiente de Quito*', by means of artificial intelligence methods (using WEKA program) and with the weighted inverse distance interpolation method (in QGIS program). After the quality check and data normalization, multiple linear regression model was used, resulting in a correlation of 0.52 of approximation to records; and a decision tree algorithm, using J48 classifier, to determine if the concentration exceeds the median of records, which resulted in correct classification of 64% of instances. The CO, NO<sub>2</sub>, Relative Humidity and Ambient Temperature are classified as key pollutants for the linear regression; while the decision tree identifies the importance of CO, NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub> and PM<sub>10</sub>, Relative Humidity, Wind Speed and Solar Radiation and Hour. The maps interpolated from the algorithms indicate the highest concentration occurs from 06:00 to 11:59 (Morning), then from 05:00 to 20:59 (Night), and 21:00 to 05:59 (Early morning), and last from 12:00 to 16:59 (Afternoon). Unfortunately, the method does not offer the adequate resolution to discriminate the variations of high and low concentrations, so an alternative to more accurately predict the concentration of the pollutant in question is through plant density, capable of reducing PM<sub>2.5</sub> by evapotranspiration, and the traffic intensity and vehicular flow according to the street type, because traffic is a major source of pollution in the Urban Area of Quito.

# ÍNDICE

1.	Capítulo I: Introducción .....	1
1.1.	Antecedentes .....	1
1.2.	Objetivos .....	6
1.2.1.	Objetivo General.....	6
1.2.2.	Objetivos Específicos .....	6
2.	Capítulo II: Marco Teórico .....	6
2.1.	Contaminación ambiental .....	6
2.2.	Contaminantes atmosféricos .....	8
2.2.1.	Óxidos de nitrógeno (NO <sub>2</sub> , NO, NO <sub>x</sub> ).....	9
2.2.2.	Monóxido de Carbono (CO).....	10
2.2.3.	Dióxido de azufre (SO <sub>2</sub> ) .....	10
2.2.4.	Ozono (O <sub>3</sub> ) .....	11
2.2.5.	Material particulado (PM <sub>10</sub> , PM <sub>2.5</sub> , PM <sub>1</sub> ) .....	12
2.3.	Material particulado fino (PM <sub>2.5</sub> ) .....	14
2.4.	Combustibles .....	15
2.5.	Medición de contaminación .....	16
2.6.	Interpolación de la contaminación .....	18
2.7.	Predicción de contaminación.....	21
2.8.	Aprendizaje computacional .....	22
3.	Capítulo III: Alcance .....	23
3.1.	Alcance.....	23
3.2.	Situación actual de Quito.....	24

4.	Capítulo IV: Justificación .....	24
5.	Capítulo V: Metodología.....	25
5.1.	Revisión previa de equipos .....	25
5.2.	Aplicación de Aprendizaje computacional.....	25
5.2.1.	Depuración de base de datos .....	25
5.2.2.	Selección de modelos de aprendizaje computacional.....	27
5.3.	Generación del mapa de concentración de $PM_{2.5}$ .....	27
5.4.	Validación del mapa de concentración de contaminantes de $PM_{2.5}$ .....	28
6.	Capítulo VI: Resultados y análisis.....	28
6.1.	Revisión previa de equipos .....	28
6.2.	Regresión Lineal Múltiple .....	31
6.3.	Árboles de Aprendizaje.....	34
6.4.	Mapas de Interpolados de $PM_{2.5}$ .....	41
6.5.	Validación de mapas en campo .....	43
7.	Conclusiones y Recomendaciones .....	47
7.1.	Conclusiones.....	47
7.2.	Recomendaciones .....	47
	Referencias.....	49
	ANEXOS .....	61

## 1. Capítulo I: Introducción

### 1.1. Antecedentes

“El amor está en el aire, pero el aire está altamente contaminado” (Amit Abraham). Más allá de una sátira es una realidad para quienes viven en la ciudad, junto a fábricas con chimeneas activas, o cerca de una concurrida carretera. No hay problemas por el amor, sino por los problemas de calidad de aire que se presentan en sitios urbanos.

El estudio de la contaminación del aire dio inicio con eventos críticos como fue en Mause Valley en Bélgica, en 1930, donde se generó problemas respiratorios en cerca de 1.400 personas y el fallecimiento de 60 personas, debido a una combinación de condiciones atmosféricas poco favorables, y una mezcla de gas de óxidos de azufre, hollín y monóxido de carbono de fábricas (Mata & Quevedo, 2005); luego ocurrió la niebla tóxica de Donora (Estados Unidos) en 1948, debido a un fenómeno meteorológico y las concentraciones de azufre, flúor, y otros contaminantes, que resultó en la muerte de 20 personas y problemas de salud en 7000 personas aproximadamente; finalmente está la Gran Niebla de Londres del 4 al 8 de diciembre de 1952 (Reino Unido), donde una mezcla entre niebla y humo tomó la vida de 12.000 londinenses y enfermó alrededor de 100.000 personas más, esto debido a la quema de combustibles con altas cantidades de azufre por la helada del invierno (Miller & Miller, 1989). Y un posterior incidente en 1962, en las industrias alemanas, donde 156 personas murieron (O’Hara, 2007). Tras el problema en Gran Bretaña, Sir Hug Beaver, presidente del Comité de Contaminación del Aire del Gobierno del Reino Británico, tomó medidas para estudiar y recomendar el control de las emisiones industriales y las calefacciones de hogar con objeto de eliminar el humo negro, medidas que se tornaron en ley de Gran Bretaña. (O’Hara, 2007; Rusznak, Devalia, & Davies, 1994).

Estudios de contaminantes como SO<sub>2</sub>, NO<sub>2</sub>, CO, Ozono, ciertos metales pesados y material particulado fino determinaron que estos elementos son perjudiciales para la salud de las personas por estudios como en España el

Estudio Multicéntrico Español sobre la relación entre la Contaminación del Aire y la Mortalidad (EMECAM), el Estudio Nacional de Morbilidad, mortalidad y contaminación del Aire de Estados Unidos (NMMAPS en inglés) y el Sistema de Información Europeo: contaminación del Aire y Salud (APHEIS) entre otros (Martí Valls, 2017).

Como una referencia sobre los problemas por contaminantes atmosféricos, de acuerdo con los datos registrados por la ONU, 7 millones de personas murieron en 2012 debido a exposición a aire contaminado, de las cuales 2,6 millones se deben a contaminación atmosférica en exteriores. Las enfermedades que contraen debido a la calidad de aire en exteriores son “Enfermedad coronaria (40%), Accidentes cerebrovasculares (40%), Enfermedades pulmonares obstructivas crónicas (11%), Cáncer de Pulmón (6%) e Infección aguda de las vías respiratorias (3%)” (Romero, 2014).

Existen eventos que afectan la calidad del aire de las personas, como son incendios forestales, que aumentan significativamente a los problemas cardiovasculares y cerebrovasculares, especialmente en adultos mayores a 65 años (Wettstein et al., 2018), o las temporadas estacionales donde inflorescencias liberan polen en gran cantidad, las cuales producen rinitis alérgica, asma y dermatitis atópica, todas estas asociadas con PM<sub>10</sub>, además de NO<sub>2</sub>, SO<sub>2</sub>, O<sub>3</sub> y CO (Kim, Park, Park, & Yoo, 2016). El ozono presente a nivel terrestre, técnicamente ozono troposférico, es un contaminante atmosférico producido por reacciones fotosintéticas de óxidos nitrosos y compuestos orgánicos volátiles con influencia de la radiación solar, y tiene la capacidad de generar alteraciones inmunológicas, funcionales, morfológicas y bioquímicas tanto en personas como en animales (Lippmann, 1989). Los metales en el aire como el hierro, sodio, sílice y magnesio generalmente son de origen natural, mientras que los metales derivados de la combustión varían según la fuente, así debido a la combustión del aceite se puede emitir Vanadio, Níquel, Zinc y Hierro; mientras que del carbón puede existir Zinc y Selenio (Amdur, Doull, & Klaassen, 1993).

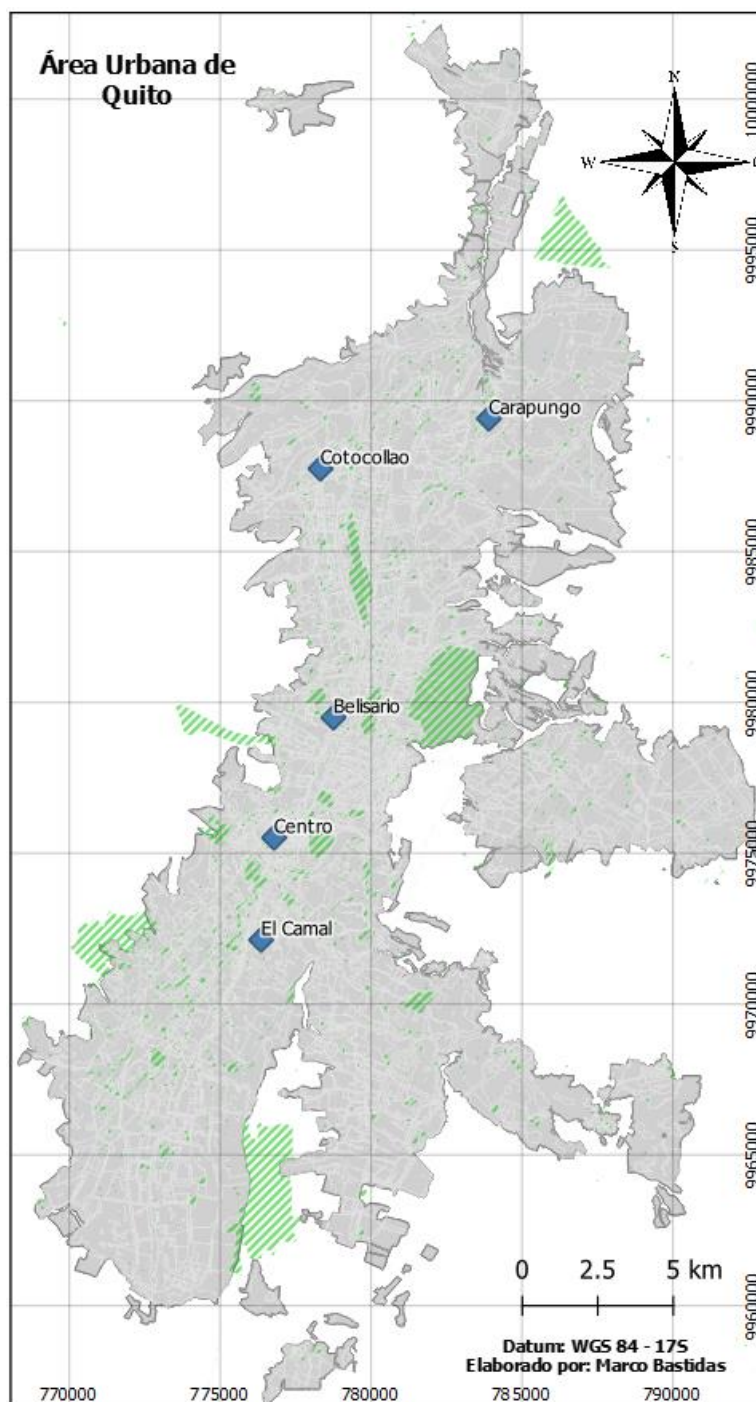
Por la quema incompleta de materia orgánica se generan aldehídos (formaldehídos y acroleínas) que también son liberados al ambiente, causando problemas de asma, alergias y problemas en el tracto respiratorio bajo de los niños (Krzyzanowski, Quackenboss, & Lebowitz, 1990; McGwin, Lienert, & Kennedy, 2010; Roda et al., 2011). Es necesario considerar que los metales, así como el material particulado (generalmente como óxidos, sulfatos o cloruros), atraviesan las vías respiratorias y en conjunto generan una determinada toxicidad en el cuerpo humano (Costa & Dreher, 1997).

Las dos principales fuentes de los contaminantes mencionados en Sudamérica, para países en vías de desarrollo, son el parque automotriz y el consumo doméstico de combustibles (Karagulian et al., 2015). En Quito, según la Secretaría del Ambiente del Distrito Metropolitano de Quito se puede describir que otras fuentes de contaminantes importantes, además de las anteriores, se consideran: la erosión eólica, actividades de construcción, molienda de rocas, procesos de combustión, incendios forestales, fundición y purificación de metales, generadores de energía, termoeléctricas, erupciones volcánicas, fertilizantes, baterías a la intemperie, plásticos, incineración de basura, humo de cigarrillo y quema de materiales orgánicos (Secretaría del Ambiente/ DMQ, 2015).

Toda actividad que realizamos es una potencial fuente generadora de contaminación para el medio ambiente, la diferencia radica en la concentración para que se vuelva perjudicial para el ser humano o el ambiente (Amdur et al., 1993), y dicha concentración varía según la facilidad de dispersión del contaminante en la atmósfera. Por esto, el determinar la calidad del aire en una ciudad requiere el tomar en cuenta factores físicos (topografía, meteorología), el nivel de desarrollo del sitio (transporte, industria, sistema de cuidado a la salud, canasta básica, etc.) y características de la población (distribución por edades y línea base del estado de salud), lo cual permite tener una idea de las afectaciones que puede tener la calidad de aire (Siabato, Sánchez, & Cely, 2002).

Sin embargo, con la tecnologías y recursos disponibles, actualmente no es viable determinar constantemente la concentración de cada elemento en el ambiente en áreas extensas, por lo cual es muy útil el uso de herramientas de modelamiento matemático de contaminantes, esto es posible mediante el ingreso de variables y datos clave, su procesamiento mediante algoritmos y ecuaciones, y el análisis final de resultados (De Visscher, 2013). De hecho, el modelamiento de contaminantes previo a la construcción de un proyecto con afectación a la atmósfera (Ej.: una fábrica, una cantera, etc.) permite determinar con mayor certeza los impactos que tendrá el proyecto en el entorno, y las medidas a tomar evitar dichos problemas (Puigcerver Zanón & Carrascal Triola, 2008). Para ello, se requiere de un inventario detallado y actualizado continuamente respecto a fuentes contaminantes, a fin de obtener una alta precisión en cuanto a los resultados; y en caso de existir terrenos o condiciones complejas los modelos que se aplican no predicen con exactitud la información en todos los casos, pero se puede obtener una aproximación cercana a la realidad dependiendo de las mejores características y condiciones que se ajusten al entorno (Bluett et al., 2004).

Una de las principales iniciativas propuestas por la Organización Panamericana de la Salud, al mismo nivel de importancia que se da al Cambio Climático, es desarrollar planes de acción que permitan una mejora en la calidad del aire para grandes centros urbanos, que surge como la Iniciativa de “Aire Limpio para las Ciudades de América Latina” (M. E. Korc, 2000). Los datos recuperados para Quito demuestran que existe normativas, políticas y controles de la calidad de aire de la ciudad de Quito, que son los solicitados por el programa. La Secretaría del Ambiente de Quito genera información de las condiciones en tiempo real cada hora, la cual es procesada para generar el índice Quiteño de Calidad de Aire en función de los contaminantes  $PM_{2.5}$ ,  $PM_{10}$ ,  $NO_2$ ,  $SO_2$ ,  $O_3$  y  $CO$ , cuyos datos para el cálculo son extraídos directamente de las estaciones de monitoreo automático en tiempo real (Secretaría del Ambiente/ DMQ, 2015).



*Figura 1.* Estaciones de monitoreo para PM<sub>2.5</sub> de la Red Metropolitana de Monitoreo Atmosférico de Quito.

En ciertas zonas del área urbana de Quito el nivel de PM<sub>2.5</sub> sobrepasa los criterios establecidos por la OMS respecto al nivel aceptable para la salud ( $25\mu\text{g}/\text{m}^3$  en 24 horas y  $10\mu\text{g}/\text{m}^3$  promedio anual), pero se desconoce cuáles son los sitios exactos donde esto ocurre. Una propuesta a esto es la generación de mapas de dispersión y concentración de contaminantes, los cuales se han



llevado a cabo anteriormente. De hecho, en Cuenca a partir de datos de monitoreo y con los modelos DISPER5.2 y WRF/Chem se ha generado simulaciones entre los datos de geografía (con ArcGIS), meteorología y química de los contaminantes (Moscoso Vanegas, Astudillo Alemán, & Vázquez Freire, 2015) y en Quito se generó un mapa de contaminación atmosférica de SO<sub>2</sub>, NO<sub>x</sub>, PM<sub>2.5</sub> y Ozono (sin interacciones entre gases) con CALMET-CALPUFF (Soraya, 2013). Por lo tanto:

¿Es posible generar un mapa de PM<sub>2.5</sub> adaptado a la realidad de Quito?

## 1.2. Objetivos

### 1.2.1. Objetivo General

Predecir el comportamiento de PM<sub>2.5</sub> del área urbana de Quito con métodos de inteligencia artificial.

### 1.2.2. Objetivos Específicos

- Identificar los factores meteorológicos y químicos influyentes en la concentración de PM<sub>2.5</sub> de Quito (con datos de química y meteorología) con métodos de aprendizaje computacional.
- Generar un mapa de concentración de PM<sub>2.5</sub> de Quito mediante la interpolación mediante distancia inversa ponderada.
- Validar el mapa generado de PM<sub>2.5</sub> en una zona de Quito mediante observaciones.

## 2. Capítulo II: Marco Teórico

### 2.1. Contaminación ambiental

“Los seres humanos constituyen el centro de las preocupaciones relacionadas con el desarrollo sostenible. Tienen derecho a una vida saludable y productiva

en armonía con la naturaleza.” (Conferencia de las Naciones Unidas sobre el Medio Ambiente y el Desarrollo, 1992). Es el principio, con el cual inicia la declaración de Río de Janeiro, reafirmando la declaración de los derechos humanos y con prospectiva hacia un desarrollo a futuro que salvaguarde la integridad del ambiente. A pesar de ser la segunda cumbre respecto al medio ambiente, es en ella donde se plantea los problemas y se busca establecer e integrar soluciones entre las diferentes naciones. En cumbres posteriores (Johannesburgo en 2002 y Río+20 en 2012) aumenta la presión social y la evidencia científica de una urgente necesidad de implantar cambios en el estilo de vida mundial para salvaguardar y recuperar la salud del planeta.

El evento histórico que marca la partida del desarrollo humano, la explotación a gran escala y la generación de desechos y contaminantes al medio ambiente, es la revolución industrial alrededor del año 1760. Es un punto clave para el comienzo del desarrollo de la tecnología que poseemos actualmente, pero a la vez, es el punto en el que la humanidad, debido a la contaminación generada, cruza el límite de la naturaleza para regenerar los recursos y equilibrar las condiciones naturales. Por supuesto, no es correcto el generalizar que toda industria es contaminante, pues existen industrias que no emiten contaminación, y existe contaminación que es de origen natural (Martí, 2003).

La contaminación se define como “alteración nociva de las condiciones normales por agentes químicos o físicos” (Real Academia Española, 2019). Se considera la contaminación del aire, agua y suelo cuando este ha perdido sus características originales y/o le es imposible cumplir las funciones que debería, volviéndose insostenible para la vida. En materia atmosférica, con el desarrollo como objetivo en la mira de las grandes ciudades, la actividad industrial aumenta, así también el uso de combustibles fósiles (especialmente carbono y derivados del petróleo), y así también el parque automotriz. Para el último siglo se ha relacionado fuertemente eventos climáticos y atmosféricos con la actividad humana (Seoánez Calvo, 2008):

- Contaminación de vegetación y fauna en las cercanías de vías altamente transitadas.
- Lluvias ácidas sobre lagos y bosques en el Centro y Norte de Europa alrededor de los años 60).
- Depósitos ácidos en las cercanías de instalación de combustión en España alrededor de los años 70.
- Hielo con plomo de origen atmosférico en Groenlandia (1960)
- Muerte de 80 000 aves y malformaciones congénitas por emisiones accidentales de dioxina en una fábrica de triclorofenol en Seveso, Italia (1976).
- Daños en zonas agrícolas, lluvias negras, reducción de calidad notable en reservas de agua y valles cercanos a causa del incendio de 600 pozos de petróleo en la guerra del Golfo por parte del ejército iraquí (1991).

La contaminación atmosférica o polución del aire es un hecho, y en primer lugar es necesario entender lo que la compone.

## 2.2. Contaminantes atmosféricos

Los contaminantes pueden clasificarse según a partir de su origen (primarios si se emiten directamente de la fuente o secundarios si se producen a partir de los primarios), por su toxicidad, o por su fuente (natural o antrópica). Desde la primera acta en favor del aire limpio se han considerado seis contaminantes criterio (Parlamento del Reino Unido, 1956), y hasta la actualidad se consideran estos por parte de la Organización Mundial de la Salud, la Agencia de Protección del Ambiente de Estados Unidos [EPA], y demás organismos internacionales; pero fuera de estos existe una variedad de compuestos contaminantes dañinos para la naturaleza y/o el ser humano. La concentración y tiempo de exposición son factores clave al determinar los efectos, siendo las unidades usadas  $\text{mg}/\text{m}^3$ ,  $\mu\text{g}/\text{m}^3$ , ppm o ppb (partes por millón o billón en volumen) para los contaminantes. A continuación, se describen las propiedades y problemas generados por los contaminantes criterio del presente estudio, así como un resumen en Figura 2.

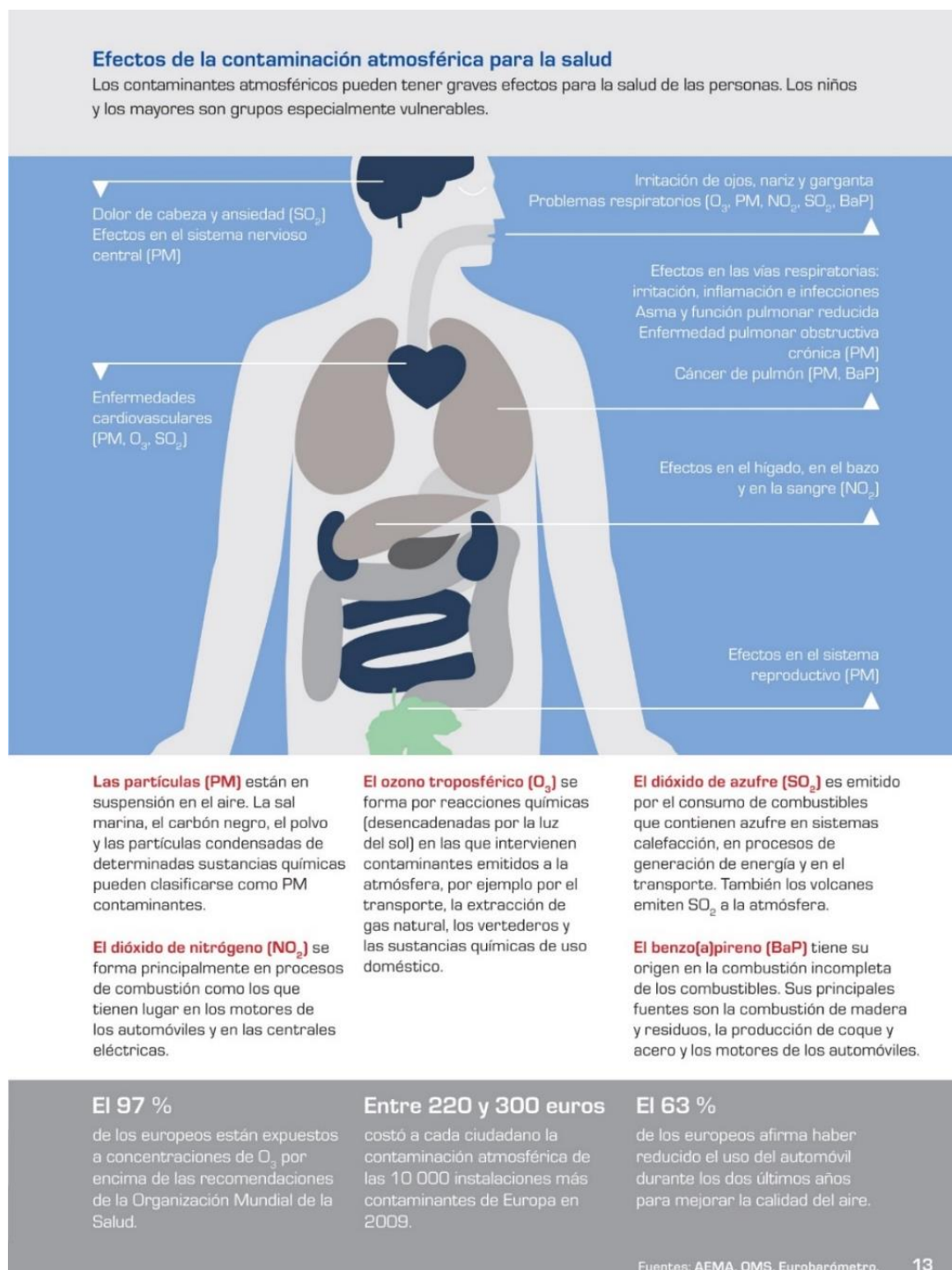


Figura 2. Resumen de fuentes, efectos y estadísticas de contaminantes atmosféricos. Tomado de Agencia Europea de Medio Ambiente, 2014.

#### 2.2.1. Óxidos de nitrógeno ( $\text{NO}_2$ , $\text{NO}$ , $\text{NO}_x$ )

El dióxido de nitrógeno es un gas irritante que puede producir una variada cantidad de problemas pulmonares cuando se inhala en altas concentraciones,

y puede provenir de motores a combustión (según la cantidad de oxígeno), cocinas de keroseno, calentadores de casa o incluso de combustibles de biomasa, y aumenta o disminuye respecto al ratio de aire ingresado para la combustión (Amdur et al., 1993). Se ha determinado que puede afectar los bronquiolos terminales (incluso con mayor potencia que el ozono) y, a altas concentraciones, los ductos alveolares y alveolos son afectados presentando una sensibilidad al cambio oxidante, resultando en un deterioro de las funciones respiratorias, cambios en la respuesta del sistema inmune y en una reducción de anticuerpos especialmente en neonatos a partir de 0,6 ppm (Chen, Gokhale, Shofer, & Kuschner, 2007).

El monóxido de nitrógeno es un compuesto altamente reactivo que es parte del proceso intermediario del ozono (debido a que es un radical libre) y acarrea consecuencias más graves que la exposición al dióxido de nitrógeno (Pacher, Beckman, & Liaudet, 2007).

#### 2.2.2. Monóxido de Carbono (CO)

Es un gas no irritante incoloro e inodoro que al entrar en contacto con la sangre produce carboxihemoglobina, un asfixiante químico, que impide la absorción de oxígeno en la sangre (evitando la oxigenación del cuerpo), lo cual disminuye la concentración de oxígeno en la sangre y provoca efectos negativos en el sistema cardiovascular nervioso, así como cambios en el comportamiento, y en el peor de los casos muerte debido a la falta de oxígeno en el cerebro (Amdur et al., 1993). Es generado en la combustión cuando el aire utilizado no es suficiente para la reacción, por lo que sus fuentes son estufas domésticas, motores de combustión, plantas energéticas e incendios (Mihelcic & Zimmerman, 2012).

#### 2.2.3. Dióxido de azufre (SO<sub>2</sub>)

Es un gas irritante soluble en agua. Puede estimular constricciones bronquiales y secreción de mucosas, se absorbe con mayor facilidad al respirar por la boca y genera mayor irritación especialmente en personas asmáticas. Ocurre porque

se disuelve en las paredes mucosas como bisulfito o sulfito y se redistribuye por el cuerpo. Se cree que puede interactuar con los receptores sensoriales en las vías respiratorias generando broncoconstricción, y resultando en asfixia al combinarse con otros problemas respiratorios en los peores casos. Se han observado estos efectos con concentración de 1 a 3 ppm durante 10 minutos de exposición, y de 0,25 a 1 ppm durante una hora causa los mismos efectos en personas con asma de nivel medio. (Amdur et al., 1993).

#### 2.2.4. Ozono (O<sub>3</sub>)

El ozono solo se considera como un contaminante secundario cuando se encuentra a nivel troposférico (llega hasta 17 km verticales de la atmósfera a partir del nivel del mar en la longitud del ecuador terrestre) y es formado gracias a la acción de la luz ultravioleta a partir de reacciones complejas de óxidos nitrosos con compuestos volátiles orgánicos como alcanos o cadenas aromáticas (Figura 3) así como en muy raras ocasiones por una intrusión por parte de la capa de ozono de la estratósfera en la tropósfera (Mamane, 2008).

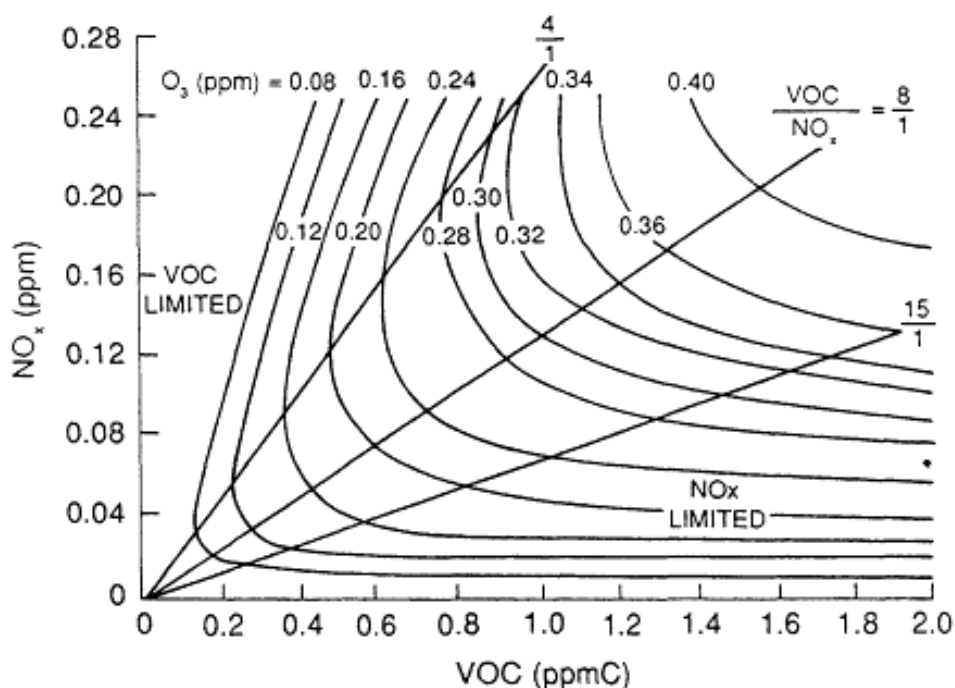


Figura 3. Isoleptas de ozono típicas utilizadas en EKMA de EPA. Tomado de Swackhamer, 2003.

Es considerado como el principal contaminante del smog fotoquímico y posee una elevada bio-reactividad, en comparación a los demás contaminantes, debido a que genera radicales libres en una mayor concentración que los que puede generar los óxidos nitrosos debido a su interacción y formación en presencia de compuestos orgánicos (Lippmann, 1989).

Se conoce que los efectos son morfológicos, funcionales, inmunológicos y alteraciones biológicas, pero que atacan especialmente a las vías respiratorias (Environmental Protection Agency, 1999). A corto plazo genera daños y anomalías en la función pulmonar y cardiovascular (Amdur et al., 1993), así como a largo plazo daños graves en las vías respiratorias, capacidad pulmonar reducida, síntomas de aumento de frecuencia respiratoria, asma de inicio reciente en niños y agravamiento de síntomas en asmáticos e incluso mortalidad respiratoria (Nuvolone, Petri, & Voller, 2018). Se determinó que no causa ningún efecto a la salud a partir de 20 ppb, y se recomienda que el nivel seguro es de 10 ppb (Bell, Peng, & Dominici, 2006).

#### 2.2.5. Material particulado (PM<sub>10</sub>, PM<sub>2.5</sub>, PM<sub>1</sub>)

Aunque generalmente se cree que es solo 'hollín', no está compuesto únicamente por el carbón residual procedente de combustión incompleta, está conformado también por otros materiales carbonáceos, sulfatos de ácido, silicatos, sal marina, polvo, polen, metales pesados, etc. Todo esto varía en cuanto al sitio, las condiciones meteorológicas, atmosféricas, ambientales, climáticas, históricas y al desarrollo de un sitio. Así mismo pueden ocurrir transformaciones secundarias incluso con los gases que los circundan. El material particulado puede estar en estado sólido, líquido o una mezcla de estos conformado por compuestos orgánicos e inorgánicos.

Este contaminante se investigó y controló con mayor fuerza, aún más que el que el ozono, en los últimos 20 años debido a los datos consistentes del aumento de mortalidad y morbilidad, así como un mayor impacto adverso en relación con el análisis costo-beneficio. Se considera que no existe un "límite" saludable que se

pueda establecer para el Material particulado (Amdur et al., 1993). Se ha considerado múltiples formas de evaluar su toxicidad en relación con su peso y la toxicidad específicamente química entre otras, pero no se ha llegado a un consenso general aún.

Se ha clasificado el material particulado en función de los diámetros aerodinámicos: iguales o menores a 10 micrómetros (material particulado grueso o PM<sub>10</sub>), iguales o menores a 2.5 micrómetros (material particulado fino o PM<sub>2.5</sub>), iguales o menores a 1 micrómetro (material particulado ultrafino o PM<sub>1</sub>).

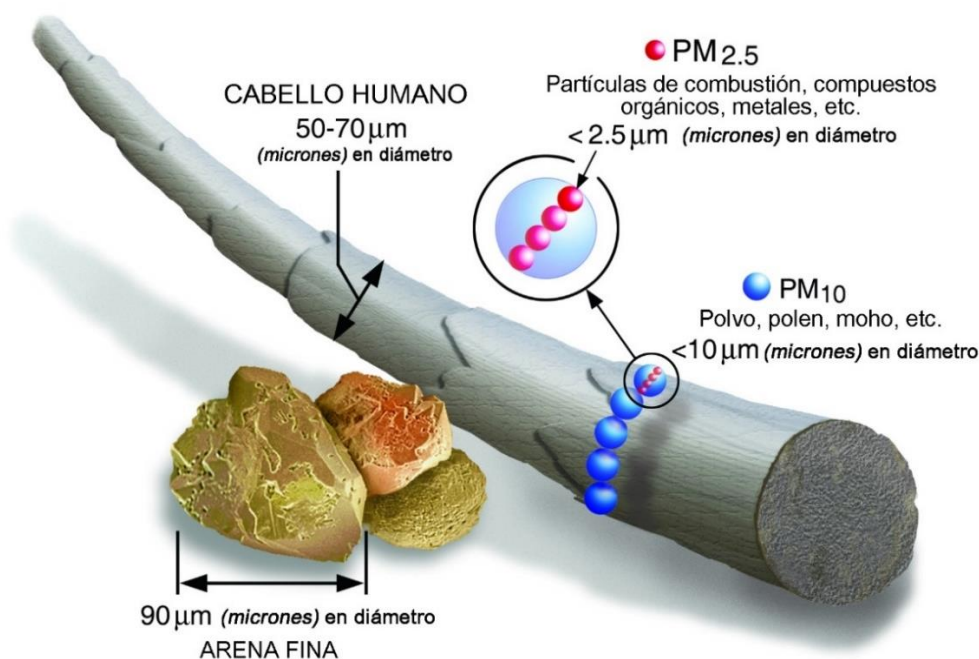


Figura 4. Comparación de tamaño de partículas del material particulado. Adaptado de US EPA, s/f.

Aparte de la formación natural, el material particulado grueso se forma por disrupción mecánica, evaporación de aerosoles, suspensión de polvo entre otras, y el material particulado fino a causa de pirólisis a altas temperaturas, así como la transformación que involucra luz del sol y vapores orgánicos (Amdur et al., 1993).

Según el Programa de Desarrollo de las Naciones Unidas se determinó que a nivel mundial 7 millones de personas mueren anualmente a causa de material particulado fino en las grandes ciudades, es una cifra alarmante por lo que se



incluye como indicador del tercer objetivo de desarrollo sostenible: Salud y Bienestar (PNUD, 2019). Incluso, se estima que en el 2016 existieron aproximadamente 4.206 muertes en la capital de Ecuador a causa del material particulado (BREATHHELIFE, 2016; OMS, 2017).

### 2.3. Material particulado fino (PM<sub>2.5</sub>)

El material particulado con diámetro aerodinámico igual o menor a 2.5 micras de metro (PM<sub>2.5</sub>) se genera por reacciones químicas, condensación, coagulación y evaporación de múltiples compuestos, así como de fuentes naturales como son incendios, erupciones volcánicas y tormentas de arena. Su composición varía según su fuente, pero puede contener cenizas, carbón, nitratos, cloratos, sulfatos, metales pesados y carbón elemental. Estudios sobre las principales especies encontradas de PM<sub>2.5</sub> (en Chongju, Korea del Sur y en Nanjing, China) sugiere que son sulfatos (11 a 23%), carbón elemental y orgánico (17 a 28%), nitratos y amonios sus principales componentes (Lee & Kang, 2001; H. Yang et al., 2005). También

Una forma de evaluar indirectamente la presencia y concentración de PM<sub>2.5</sub> es mediante otros contaminantes que se emitan en conjunto con el contaminante estudiado, por lo tanto: Si su origen se origina de la combustión química puede poseer una relación directa con las emisiones de contaminantes como SO<sub>x</sub>, NO<sub>x</sub>, CO, O<sub>3</sub> (Chaparro, Cuervo, Gómez, & Toro, 2001); si su origen es mecánico (por acción del viento, maquinaria o herramientas de trabajo) se verá focalizado alrededor de sitios específicos de generación (Arrieta Fuentes, 2016), y su concentración es afectada en función de las condiciones meteorológicas del momento, por lo tanto las condiciones de dispersión como la lluvia, la humedad relativa (influenciada a su vez por presión y temperatura), el viento, la radiación solar y la capa borde límite del planeta (influye directamente por inversión térmica en las mañanas) son factores que se deben considerar al realizar el modelamiento (Mamane, 2008; Mills, 2014; Wallace & Kanaroglou, 2009). Pueden darse interacciones entre el PM<sub>2.5</sub> y los demás gases contaminantes en presencia de una alta humedad relativa, la cual puede acidificar al material

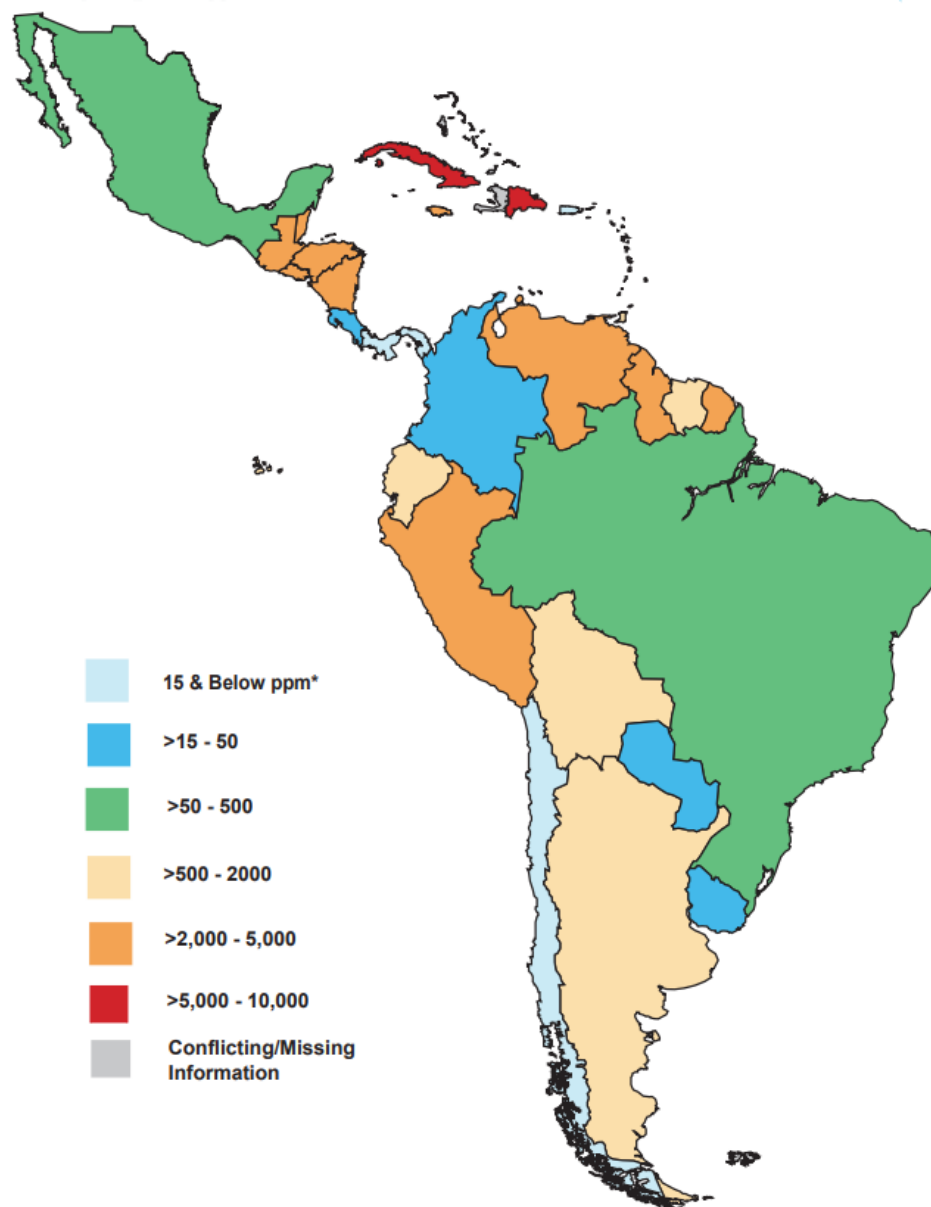
particulado con iones de nitratos, sulfatos, y la proporción entre carbono elemental y carbón orgánico (Zhang et al., 2016).

Por su tamaño y peso son capaces de viajar entre 100 y 1000 km desde el sitio en el que se formaron. Generan y/o agravan problemas cardiopulmonares, resultando en complicaciones que reducen el tiempo de vida considerablemente y son causas principales de mortalidad en ciudades altamente contaminadas (Lippmann, 2012).

#### 2.4. Combustibles

Considerando que en Latinoamérica la principal fuente de contaminación son los vehículos (M. E. Korc, 2000), es importante tomar en cuenta características clave del combustible en relación con la contaminación atmosférica.

El nivel de azufre que compone a un combustible es un elemento clave para determinar la producción de dióxido de azufre (Figura 5). Por otra parte, la combustión incompleta, el exceso de aire y la efectividad de los motores de transporte en base a diésel, gasolina u otros combustibles determina notablemente la formación de material particulado, compuestos orgánicos volátiles, monóxido de carbono y óxidos nitrosos (Mamane, 2008). Pero aun así que el uso de combustibles líquidos reduce la muerte prematura (en 51%), el cáncer de pulmón (en 65%), y accidentes cerebrovasculares (en 28%) en comparación con combustibles sólidos como el carbón (CCAC & UNEP, 2017).



*Figura 5.* Niveles de azufre del combustible diésel en ppm por país en Latino América. Tomado de PNUD & Asociación para combustibles limpios y vehículos, 2017.

## 2.5. Medición de contaminación

Dado que existe una gran variedad de contaminantes generados por el ser humano y su afección al sistema respiratorio y cardiovascular, se ha generado la suficiente conciencia para crear normativas y tomar medidas para el cuidado del aire. Los primeros estudios en Latinoamérica se realizaron en 1950 por parte de universidades y dependencias de ciertos ministerios de salud, seguidos por

el establecimiento de una red de estaciones de monitoreo de contaminación del aire para los estados miembro de la Organización Panamericana de Salud (M. Korc & Sáenz, 1999).

El monitoreo es definido como el muestreo, análisis y procesamiento de información sobre las sustancias concentradas en el aire en un sitio y momento determinados (Martínez & Romieu, 2003). Esta acción es clave para establecer estudios epidemiológicos relacionados con las sustancias nocivas, determinar los estándares de calidad, las condiciones reales y actualizadas del sitio, así como información clave para los programas y estrategias a implementar.

Los pasos de la estrategia sugerida por Martínez & Romieu (2003) para establecer una red de monitoreo atmosférico adecuado son:

- Establecer objetivos del monitoreo
- Definir parámetros ambientales a medir
- Determinación de número y sitios de muestreo
- Determinación de tiempos de muestreo
- Elección de técnicas de análisis y equipo

Las técnicas utilizadas para el análisis de cada contaminante varían según la planificación y los recursos asignados, pueden ser clasificadas en 5 categorías: muestreadores pasivos, reactivos, automáticos, sensores remotos y bioindicadores (Inche, 2004).

Los muestreadores pasivos son dispositivos que colectan mediante adsorción y/o absorción un contaminante, al estar expuestos durante un intervalo de muestreo, para ser analizados cuantitativamente en el laboratorio tras su recuperación. El dato que representan es un valor promedio del intervalo de exposición.

Los muestreadores reactivos son equipos que requieren el paso del aire a través de un medio para su recolección y/o análisis físico químico. Dichos equipos

tienen una alta sensibilidad a las variaciones de concentración, a partir de las cuales también pueden generarse promedios de intervalos.

Los monitores automáticos, al igual que el anterior método, se basan en medios fisicoquímicos para realizar el análisis de los gases, pero poseen una mayor precisión respecto a los intervalos analizados, y dado que son automatizados entregan los resultados en tiempo real.

Existen sensores remotos, que son similares a los monitores automáticos, pero a diferencia de estos generan mediciones integradas de múltiples componentes en trayectorias específicas de la atmósfera, permitiendo así la generación de mapas tridimensionales de los contaminantes y de su dispersión.

Finalmente, el término Bioindicadores es usado para referirse al uso de seres vivos como medio de monitoreo de calidad y concentración de los contaminantes atmosféricos, los cuales reaccionan según su morfología y fisiología a las variaciones de condición del entorno. Pueden considerarse plantas, animales y microorganismos de la zona, según los estudios previos de toxicidad e influencia de contaminantes en ellos.

## 2.6. Interpolación de la contaminación

De acuerdo con las necesidades y capacidades que disponen las autoridades zonales es posible implementar una o varias técnicas de monitoreo según la cantidad de sitios requeridos. Pero, el obtener datos de un área geográfica extensa de la ciudad en tiempo real implica costos elevados y el mantenimiento de una serie de instrumentos de monitoreo de forma constante, por lo que una alternativa viable es calcular un valor aproximado de la concentración de los contaminantes a partir de un grupo de mediciones estratégicas en la zona, que es a lo que se llama Interpolación. Sin embargo, el tipo de interpolación utilizada puede ser determinante sobre los resultados, puesto que tienen sus propias fortalezas y debilidades (Ikechukwu, Ebinne, Idorenyin, & Raphael, 2017). Los Sistemas de Información Geográficos (SIG) actualmente han incorporado estos

métodos vinculando los cálculos a la información geográfica, entre los cuales los métodos más utilizados son: distancia inversa ponderada, interpolación basada en una red de triángulos irregulares, Kriging y Spline.

La interpolación basada en una red de triángulos irregulares (*TNI* en inglés), genera un sistema de triángulos a partir de la distribución de triángulos y posteriormente se calcula mediante el método de orden por distancia o por el método de triangulación de Delanuey, el cual permite el cálculo de la interpolación cuando círculos tangentes a 3 puntos no se encuentren con otro, volviéndolos tan equiláteros como sea posible (X. Yang, 2010). Una variante de este método es una red basada en rectángulos irregulares, pero funciona similarmente (Mitas & Mitasova, 2005).

El método Kriging está basado en funciones aleatorias con respecto a la superficie o el volumen asumido tras una realización de una función aleatoria a partir de una covarianza espacial preestablecida. Para ello asume un componente estructural de la tendencia general de la superficie, desviaciones aleatorias relacionadas a la tendencia y un ruido aleatorio (se estima aleatoriamente cada uno de estos), sin embargo este modelo pierde eficacia en comparación con otros métodos cuando no puede hallar puntos cercanos para construir su variograma (X. Yang, 2010).

*Spline* son un grupo de interpoladores llamados Funciones en Bases Radiales, los cuales utilizan funciones matemáticas para conectar los valores y generar superficies de elevación continuas. Son óptimos en superficies lisas amplias con una gran cantidad de datos, y a diferencia del método de distancia inversa interpolada, no obliga a las estimaciones a mantener los datos muestreados (Ikechukwu et al., 2017).

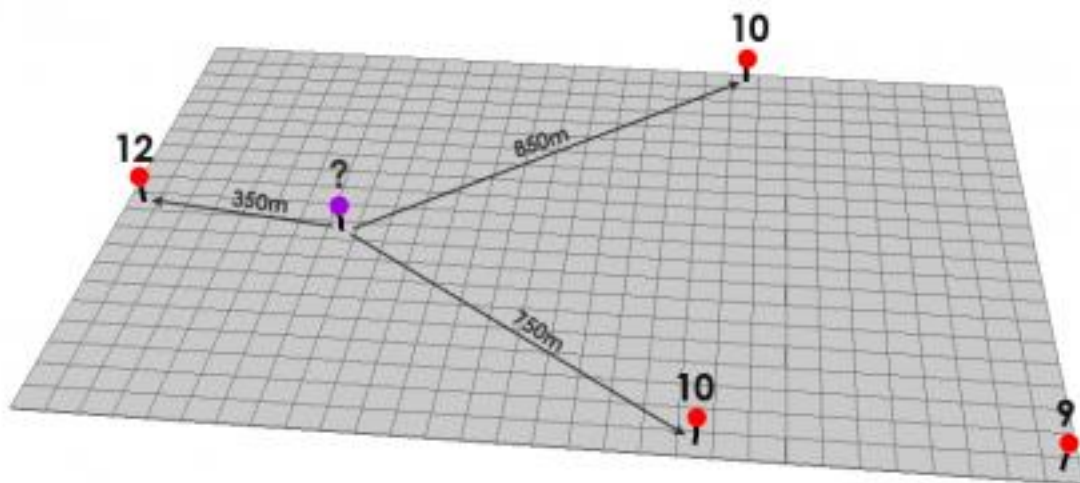
Distancia inversa ponderada (*IDW* en inglés), es un método local que aplica la Primera ley de Tolber de Geoestadística, e implica que el peso del valor de un punto es inversamente proporcional a la distancia geométrica del valor estimado elevado a una potencia específica. Su Fórmula de estimación es:

$$F(s) = \sum_{i=1}^n w_i z(s_i) = \frac{\sum_{i=1}^m w_i z(s_i)}{\sum_{i=1}^m \frac{1}{|s - s_i|^p}}$$

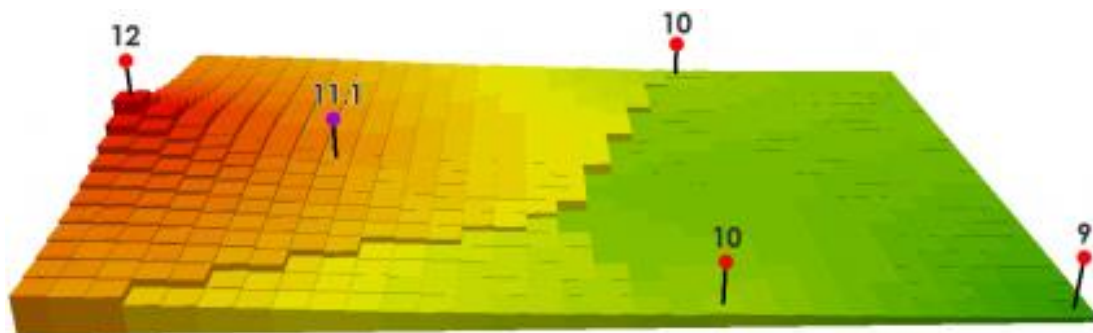
*Ecuación 1.* Ecuación para calcular Distancia Inversa Ponderada.

Donde  $s$  es una ubicación no muestreada y  $p$  es un parámetro (generalmente 2).

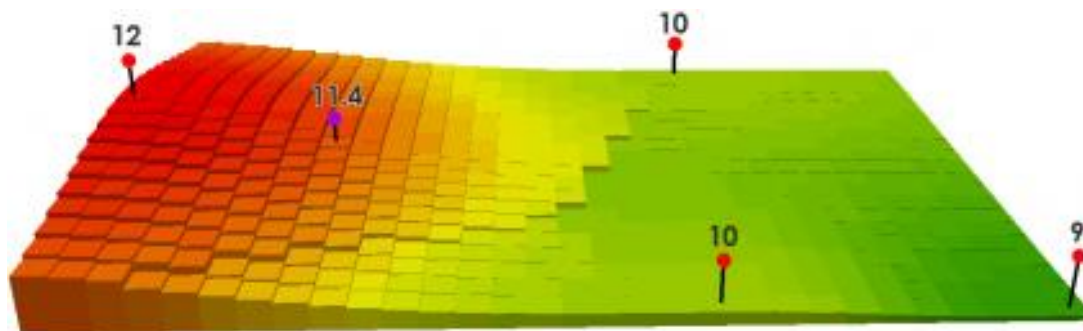
Como se aprecia en Figura 6 se requieren puntos de medición georreferenciados, y al algoritmo de IDW se calcula el valor de las celdas de las cuales no se dispone información, se aprecian los resultados suavizados al usar la potencia 1 (Figura 7) y adquieren una mayor influencia cada punto al usar la potencia de 2 (Figura 8).



*Figura 6.* Valores para iniciar interpolación. Tomado de GISGeography, 2018.



*Figura 7.* Interpolación mediante Distancia Inversa Interpolada a la potencia 1. Tomado de GISGeography, 2018.



*Figura 8.* Interpolación mediante Distancia Inversa Interpolada a la potencia 2. Tomado de GISGeogephy, 2018.

Es un método de sencillo uso y está disponible en la mayoría de los SIG. Aunque en general no produce la forma local implicada por los datos, produce extremos locales de los puntos entregados (Mitas & Mitasova, 2005) por lo cual es un modelo adecuado cuando existe poca información.

## 2.7. Predicción de contaminación

Los modelos para la predicción de contaminantes atmosféricos pueden ser clasificados según el área de aplicación (global, regional, nacional, local), la escala de tiempo (horario, diario, mensual, temporal, ocasional), según el tipo de ecuación de cálculo de transporte (euleriano o lagrangiano) y las interacciones entre los contaminantes (nulas, deposición húmeda o química entre contaminantes). Puede establecerse también si se encuentran orientados a procesos (también llamados deterministas) donde se relaciona directamente los efectos de la emisión a su fuente, causa y origen. Ejemplos de modelos deterministas son los modelos meteorológicos, de penacho de contaminación, de difusión, de deposición y químicos. La segunda clase de orientación está basada en modelos estadísticos que mediante una base de datos pueden determinar concentraciones a partir de condiciones específicas. Ejemplos de estos son los modelos de análisis de series temporales, de regresión, de distribución uniforme o multivariantes (Gallego Picó, González Fernández, & Sánchez Gimeno, 2000).



Los modelos usados con mayor frecuencia para estudios locales suelen ser Semi-Empíricos y Gaussianos, los cuales requieren una reducida cantidad de condiciones para generar resultados, pero aplican una gran cantidad de condiciones teóricas (Sbarato, Sbarato, & Ortega, 2007).

## 2.8. Aprendizaje computacional

El campo de estudio interesado en transformar datos en acciones inteligentes mediante el uso de algoritmos informáticos a lo que se le conoce como aprendizaje computacional, aprendizaje automático o '*Machine Learning*' en inglés (Lantz, 2015). Son clave para completar procesos como tratar gran cantidad de datos, hacer predicciones precisas, manejar una variedad de problemas de aprendizaje y alcanzar un aprendizaje derivado de las condiciones establecidas para el aprendizaje (Tripathi, 2017).

Los algoritmos que utiliza para el análisis de datos son divididos en 3 tipos: Supervisados, no supervisados y de refuerzo.

Los algoritmos supervisados, requieren la variable de respuesta delimitada (generalmente un vector) y mapean dentro de las demás variables entregadas una predicción de la variable establecida (Burger, 2018; Yu-Wei, 2015). Estos a su vez se subdividen en regresión, clasificación y mixtos.

Los de regresión se utilizan para determinar los cambios y evolución de los datos, y finalmente generar una predicción de los valores a futuro. Ejemplos de estos son los algoritmos: Árboles de regresión, Regresión Lineal y Árboles Modelos entre otros. Los de clasificación se los utiliza para categorizar los datos en esquemas de forma que tenga un sentido de sistematización. Ejemplos de estos son los algoritmos: Vecino más cercano, Árboles de decisión e "Ingenuo Bayes" entre otros. Por último, los algoritmos mixtos combinan los métodos de regresión y clasificación, como son: Redes Neuronales y Máquinas de vectores de soporte.

Los no supervisados, no requieren de ninguna directriz respecto a las variables (Burger, 2018). Y debido a que no posee una variable o dato objetivo, no se utiliza para realizar predicciones, pero si para la minería de datos mediante aprendizaje computacional. Ejemplos de estos modelos son algoritmos como: Reglas de asociación (para determinar relaciones y patrones) y agrupamiento por k-medias.

Finalmente, y no menos importante, son los algoritmos de refuerzo donde los algoritmos trabajan en base a recompensas (maximizar o minimizar una variable numérica) respecto a las interacciones y retroalimentación recibida según las acciones que toma. Es un algoritmo de descubrimiento mediante prueba y error en lugar de establecer de antemano la serie de acciones que debe tomar para procesar la información (Witten, Frank, Hall, & Pal, 2016).

Tras la construcción de los modelos según los datos implementados es de vital importancia el generar una evaluación del rendimiento del modelo, para lo cual no es adecuado utilizar los mismos datos, sino una porción separada de antemano del conjunto de datos o una nueva base de datos (Witten et al., 2016; Yu-Wei, 2015). Es clave este punto para evaluar un buen acople a los datos, a fin de generalizar el modelo para datos nuevos y evitar un modelo muy específico para cada caso analizado, llamado sobreajuste, o un modelo que no se permita discernir la realidad, que es el sub ajuste (Burger, 2018; Lesmeister, 2017), es preferible el uso de algoritmos que mantengan un equilibrio a fin de evitar sub ajustes y sobreajustes de los modelos (Kleine Deters, Zalakeviciute, Gonzalez, & Rybarczyk, 2017).

### **3. Capítulo III: Alcance**

#### **3.1. Alcance**

Este estudio busca examinar las diferentes relaciones que presentan los contaminantes atmosféricos ( $\text{SO}_2$ ,  $\text{CO}$ ,  $\text{NO}_x$ , y  $\text{O}_3$ ), y la influencia de las variables climáticas, con relación al  $\text{PM}_{2.5}$  en el área urbana de Quito a partir del registro

de datos de las estaciones automáticas de monitoreo de la Secretaría del Ambiente de Quito, por un período superior o igual a 10 años.

### 3.2. Situación actual de Quito

Las estaciones de monitoreo funcionan para brindar datos de la calidad de aire en Quito, más la información indicada no puede ser interpolada para toda la ciudad porque varían las condiciones entre puntos (Schneider et al., 2017), y en Quito existe mayor complejidad debido a las variaciones entre microclimas, meteorología, topografía, fuentes de contaminación atmosférica, contaminantes y vegetación. Es de igual de importante el considerar la radiación elevada, a causa de la ubicación geográfica y altitud del sitio, lo que favorece los procesos para la generación de ozono troposférico y requiere ser considerado para la generación de sus precursores (Derwent et al., 2018).

## 4. Capítulo IV: Justificación

Cuando existen múltiples fuentes que liberen contaminantes atmosféricos en sitios con condiciones de baja dispersión de gases, se aumenta la concentración de estos con un alto riesgo para la salud (Barría, Calvo, & Pino, 2016). En Quito, fuera del parque automotriz, existen variadas fuentes que aportan con diferentes contaminantes, como son las industrias, centros de comercio y residencias, por lo que es necesario un monitoreo constante (Baca, 2011; Secretaría del Ambiente/ DMQ, 2015). La red de monitoreo de contaminantes atmosféricos es un sistema de vigilancia clave para la calidad del aire en la ciudad y sus mediciones son puntuales para cada sitio de cada estación de monitoreo. Sin embargo, la certeza de los datos decae a medida que se aleja del sitio de medición. Las redes de monitoreo en Quito no cubren la ciudad completamente, y el implementar nuevas estaciones de monitoreo repercutiría en fuertes costos de implementación y mantenimiento extras para la ciudad.

Por otra parte, el poseer un mapa base o un mapa dinámico de la calidad de aire en la ciudad sería un gran beneficio, pues mostraría un valor aproximado de la

calidad de aire en los diferentes sectores de la ciudad, además de identificar puntos críticos y zonas con menor contaminación para que se realicen diferentes actividades; así como ser una herramienta en el futuro para planes de mitigación y control. Este puede generarse a partir de modelos de dispersión de contaminantes adaptados a la realidad de la urbe.

El generar un modelo de  $PM_{2.5}$  adaptado para la ciudad de Quito considerando variables como interacciones meteorológicas, química de los contaminantes y tráfico de la zona a partir de los históricos disponibles por la Secretaría del Ambiente de Quito es una excelente opción viable para la generación del mapa, el cual puede ser alcanzado mediante métodos de aprendizaje computacional.

## **5. Capítulo V: Metodología**

### **5.1. Revisión previa de equipos**

Se realizó una comparación del equipo Microdust Pro y la estación automática de monitoreo atmosférico 'Belisario' de la Secretaría del Ambiente el 14 de febrero de 2019, a fin de verificar que las mediciones del equipo Microdust se encuentran dentro del rango de la estación automática.

Además, se analizó el histórico de datos de  $PM_{2.5}$  utilizado por las estaciones con el propósito de establecer los horarios a trabajar en función de las horas pico con mayor concentración de  $PM_{2.5}$ . Se analizó estadísticamente el histórico de las estaciones para determinar tendencias notables de crecimiento o decrecimiento de  $PM_{2.5}$  a fin de ajustar las bases de datos previo a su análisis.

### **5.2. Aplicación de Aprendizaje computacional**

#### **5.2.1. Depuración de base de datos**

Se analizó un histórico de 10 años de registro de datos de contaminantes atmosféricos y meteorológicos de la Secretaría del Ambiente de Quito para las

estaciones Belisario (78°29'24" O, 0°10'48" S), Carapungo (78°26'50" O, 0°5'54" S), Cotocollao (78°29'50" O, 0°6'28" S) y El Camal (78°30'36" O, 0°15'00" S). Se trabajó los datos por estación respecto a todas las variables disponibles en el desarrollo del aprendizaje computacional (ver Tabla 1).

Se eliminó de todas y cada una de las variables los datos con valores negativos, los cuales no existen en las escalas trabajadas para los parámetros de dirección del viento [grados], humedad relativa [porcentaje], lluvia [mm/h], temperatura ambiental [grados centígrados], presión [bares], índice ultravioleta, radiación solar [ $\text{w}/\text{m}^2$ ], rapidez del viento [m/s], CO [ $\text{mg}/\text{m}^3$ ], NO<sub>2</sub> [ $\mu\text{g}/\text{m}^3$ ], O<sub>3</sub> [ $\mu\text{g}/\text{m}^3$ ], SO<sub>2</sub> [ $\mu\text{g}/\text{m}^3$ ], distancia a capa borde del planeta [km], PM<sub>10</sub> [ $\mu\text{g}/\text{m}^3$ ] o PM<sub>2.5</sub> [ $\mu\text{g}/\text{m}^3$ ].

Tabla 1. *Contaminantes registrados por estación en el período 2005 a 2018.*

<b>ESTACIÓN</b>	<b>Belisario</b>	<b>Carapungo</b>	<b>Cotocollao</b>	<b>El Camal</b>
<b>DIR_VEC</b>	X	X	X	X
<b>HUM_REL</b>	X	X	X	X
<b>LLUVIA</b>	X	X	X	X
<b>TEMP_AMB</b>	X	X	X	X
<b>PRES_BAR</b>	X	X	X	X
<b>IUV</b>			X	
<b>RAD_SOLAR</b>	X	X	X	X
<b>RAP_VEC</b>	X	X	X	X
<b>CO_mg</b>	X	X	X	X
<b>NO2_ug</b>	X	X	X	X
<b>OZONO_ug</b>	X	X	X	X
<b>SO2_ug</b>	X	X	X	X
<b>PBL_km</b>	X			
<b>PM10_ug</b>		X		
<b>PM2.5_ug</b>	X	X	X	X

Nota: PBL fue generado por WRF en colaboración con USFQ para junio-julio de 2017.

Posterior a ello, se realizó una segunda depuración eliminando datos atípicos que se encontraran por encima o debajo de tres desviaciones estándar a fin de mantener el 99,73% de la distribución normal (Galindo, 2015). El proceso se

realizó por estación y en los intervalos de horarios definidos (06:00 a 11:59, 12:00 a 16:59, 17:00 a 20:59 y 21:00 a 05:59 como Mañana, Tarde, Noche y Madrugada respectivamente). Finalmente, se verificó una distribución normalizada de cada variable con el estadístico de Kolmogorov-Smirnov en el programa RStudio (RStudio Team, 2016), que en caso de no cumplirse se normalizó toda la tabla en valores de 0 a 1 en función de su mínimo y máximo (sin los datos atípicos retirados en el paso previo).

### 5.2.2. Selección de modelos de aprendizaje computacional

Mediante Weka 3.8.3 (Frank, Hall, & Witten, 2016) se aplicó un modelo para aprendizaje supervisado de regresión lineal múltiple, donde se determinó las variables clave para la predicción de  $PM_{2.5}$  con el clasificador 'Linear regression', y al mismo tiempo se generó el modelo del árbol de decisiones para determinar concentración superior (Hi) o inferior (Lo) a la mediana con el clasificador 'J48' tomando como parámetro una clasificación mínima de 2000 instancias para cada nodo terminal.

Se analizó únicamente un 70% de los datos (tomados aleatoriamente) por estación y horario determinados para el entrenamiento del modelo. Con los patrones establecidos se realizó una validación comparando el modelo planteado y con el 30% de los datos restantes se evaluó su precisión en cuanto al coeficiente de correlación entre datos modelados y reales.

### 5.3. Generación del mapa de concentración de $PM_{2.5}$

A partir de la mediana de las condiciones en la mañana (06:00 a 11:59), tarde (12:00 a 16:59), noche (17:00 a 20:59) y madrugada (21:00 a 05:59) de cada estación y los algoritmos resultantes se generó un mapa de concentración de  $PM_{2.5}$  para cada horario, mediante la interpolación "Distancia Inversa Interpolada" (Ma et al., 2016) de los datos en las estaciones automáticas en el software de Q-GIS (Team, 2013).

#### 5.4. Validación del mapa de concentración de contaminantes de PM<sub>2.5</sub>

Con el uso del equipo Microdust Pro (Casella) - medidor de PM<sub>2.5</sub> - con filtro para material particulado de 2.5 µm, un GPS, y el equipo Aeroqual S500 -medidor de PM<sub>10</sub> y PM<sub>2.5</sub> - se realizó mediciones continuas transitando en las vías cercanas a la estación Belisario las fechas 28 de mayo, 30 de mayo y 1 de junio de 2019 desde la 7:30 hasta las 10:30, horario donde no existe dilución de contaminantes en la atmósfera. A la vez, mediante un teléfono celular con una aplicación desarrollada por el equipo de investigación de la UDLA – Sistemas Interactivos e inteligentes - SI2, se tomó datos de velocidad y congestión vehicular de las vías para aumentar la precisión del mapa. Esto a su vez se comparó con una interpolación de las medianas de concentración de PM<sub>2.5</sub> por cada estación estudiada.

### **6. Capítulo VI: Resultados y análisis**

#### 6.1. Revisión previa de equipos

Se realizó la validación desde las 08:12 hasta las 16:24 del jueves 14 de febrero, con mediciones de PM<sub>2.5</sub> cada 10 segundos con el equipo Microdust Pro, y cada hora para la estación de monitoreo de acuerdo con la configuración del equipo de la estación de monitoreo “Belisario” de la Secretaría del Ambiente de Quito.

Una comparación entre los valores arrojados por ambos equipos (Figura 9) indica que las mediciones del equipo de la Secretaría del Ambiente están dentro de la desviación estándar del equipo Microdust, y, por lo tanto, el equipo Microdust está dentro del rango válido para la toma de datos para el presente proyecto.

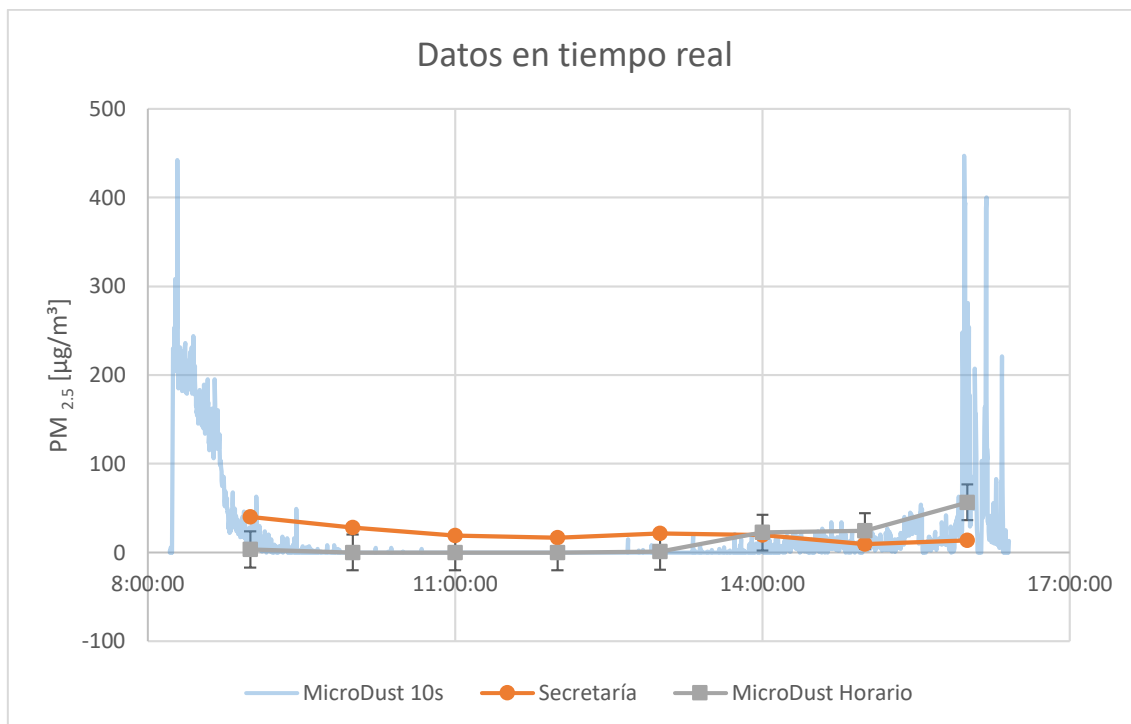


Figura 9. Mediciones de equipo Microdust Pro y estación Belisario de Secretaría del Ambiente.



Figura 10. Ubicación de equipo para validación de Datos.

Se efectuó un promedio mensual de la variación horaria de  $PM_{2.5}$  de abril de 2018 (Figura 11), que indica que es óptimo realizar las mediciones de validación entre 08:00 y 11:00 de la mañana.



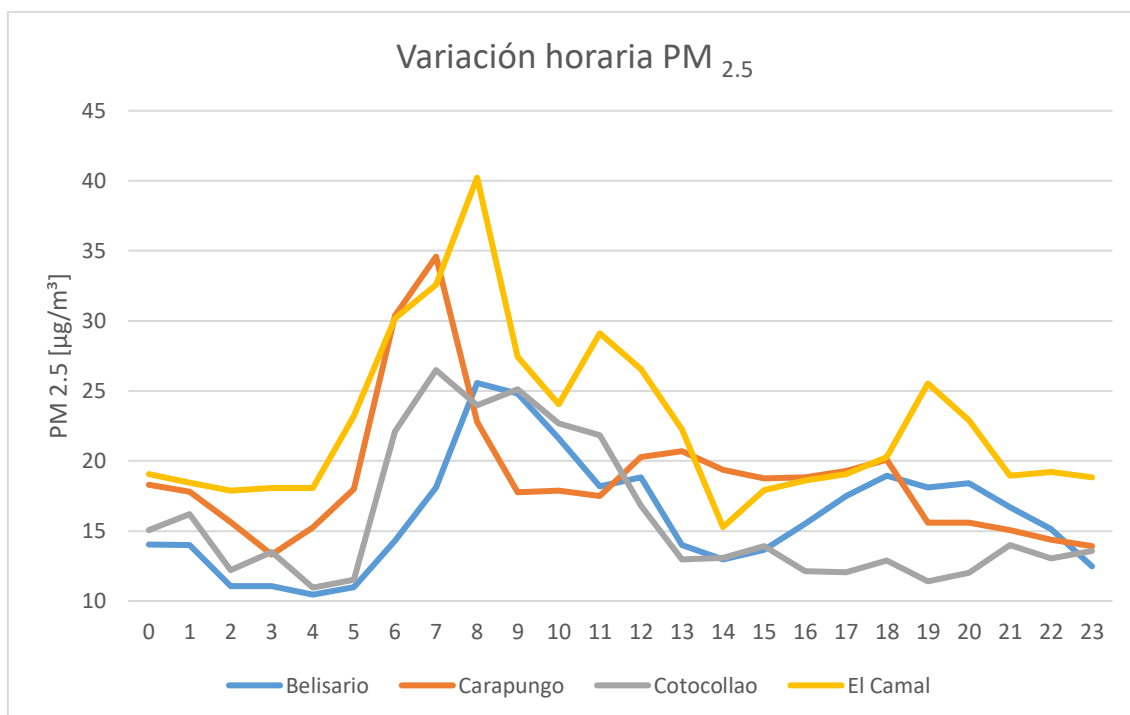


Figura 11. Variación horaria de PM 2.5 en µg/m<sup>3</sup> por estación de abril 2018.

A la vez se determinó las “horas pico” donde aumenta la concentración de PM<sub>2.5</sub>, las cuales son de 06:00 a 11:00 en la mañana y de 17:00 a 20:00 para el pico de la noche. Por lo tanto, se determinó los horarios:

- Mañana: 06:00 a 11:59
- Tarde: 12:00 a 16:59
- Noche: 17:00 a 20:59
- Madrugada 21:00 a 05:59

No se aprecia una tendencia notable respecto al crecimiento anual en el histórico de PM<sub>2.5</sub> de acuerdo con el registro histórico de la Secretaría del Ambiente (Figura 12).

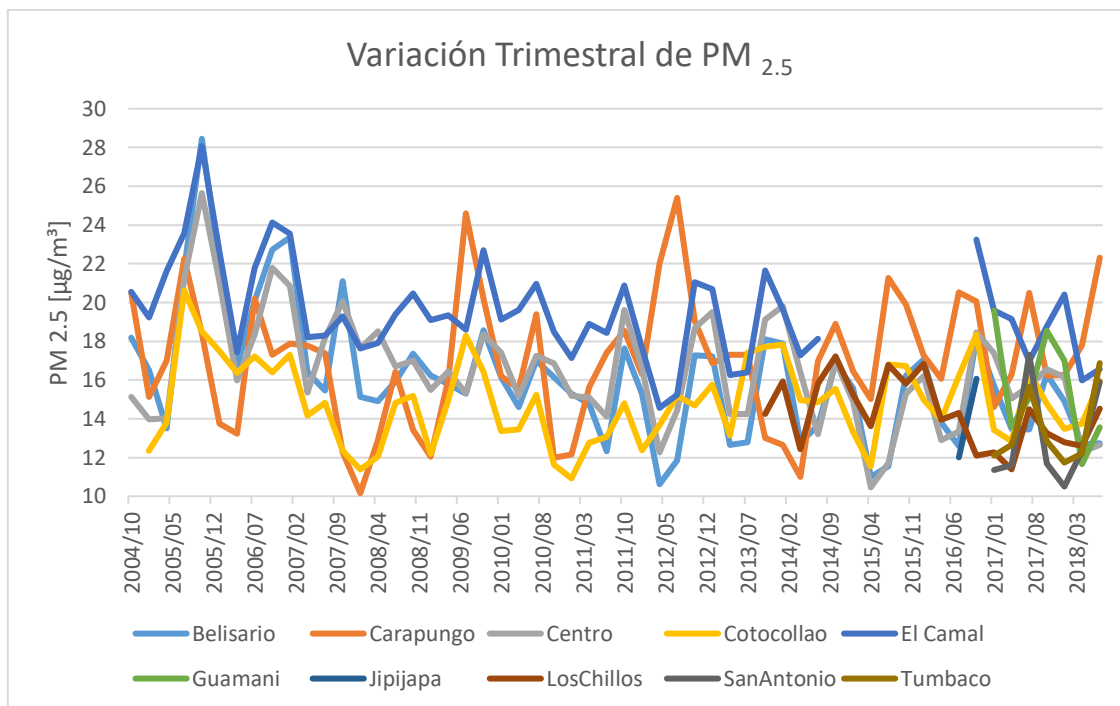


Figura 12. Histórico de variación trimestral de PM<sub>2.5</sub> por estación.

Tras determinar las horas pico se dividió la base de datos en los 4 segmentos del día establecidos y se realizó el aprendizaje computacional para cada segmento y estación.

## 6.2. Regresión Lineal Múltiple

Para el horario de 06:00 a 11:00 se determinó como factores clave en al menos 3 estaciones: la temperatura ambiental, el dióxido de nitrógeno, el monóxido de carbono y el ozono (Tabla 2).

Tabla 2. Regresión lineal múltiple normalizada para 06:00 a 11:00.

Belisario	Carapungo	Cotocollao	El Camal
$PM_{2.5\_ug} =$ $0.19464 \cdot HUM\_REL +$ $0.37682 \cdot TEMP\_AMB$ $+0.73853 \cdot NO_2\_ug +$ $-0.22518 \cdot OZONO\_ug$ $-0.23129$	$PM_{2.5\_ug} =$ $-0.08115 \cdot$ $TEMP\_AMB +0.31209$ $\cdot CO\_mg +0.39679 \cdot$ $NO_2\_ug +0.16027 \cdot$ $PM_{10\_ug} +0.0464$	$PM_{2.5\_ug} =$ $0.223 \cdot HUM\_REL +$ $0.23253 \cdot TEMP\_AMB$ $+0.20851 \cdot RAD\_$ $SOLAR +0.22878 \cdot$ $CO\_mg +0.50946 \cdot$ $NO_2\_ug -0.14439 \cdot$ $OZONO\_ug -0.21401$	$PM_{2.5\_ug} =$ $0.16786 \cdot RAD\_SOLAR$ $+0.29392 \cdot CO\_mg$ $+0.33237 \cdot NO_2\_ug$ $-0.25658 \cdot OZONO\_ug$ $+0.29653 \cdot SO_2\_ug$ $+0.06188$
<b>Correlación =</b> <b>0.6419</b>	<b>Correlación=</b> <b>0.5898</b>	<b>Correlación:</b> <b>0.6318</b>	<b>Correlación:</b> <b>0.6896</b>

Para el horario de 12:00 a 16:00 se determinó como factores clave en al menos 3 estaciones: la humedad relativa, el dióxido de nitrógeno y el ozono (Tabla 3).

Tabla 3. Regresión lineal múltiple normalizada para 12:00 a 16:00.

<b>Belisario</b>	<b>Carapungo</b>	<b>Cotocollao</b>	<b>El Camal</b>
$PM2.5_{ug} =$ $0.41681 \cdot HUM\_REL +$ $0.17571 \cdot TEMP\_AMB$ $-0.18472 \cdot$ $RAD\_SOLAR$ $+0.36944 \cdot NO2_{ug} +$ $0.17634 \cdot OZONO_{ug}$ $+ -0.14044$ <b>Correlación =</b> <b>0.5124</b>	$PM2.5_{ug} =$ $0.10872 \cdot HUM\_REL +$ $0.32784 \cdot CO_{mg} +$ $0.43349 \cdot NO2_{ug} +$ $0.07821 \cdot OZONO_{ug}$ $+ 0.17182 \cdot PM10_{ug}$ $-0.10569$ <b>Correlación=</b> <b>0.5898</b>	$PM2.5_{ug} =$ $0.31859 \cdot HUM\_REL +$ $0.13638 \cdot TEMP\_AMB$ $-0.08544 \cdot RAP\_VEC +$ $0.21837 \cdot NO2_{ug} +$ $0.24601 \cdot OZONO_{ug} -$ $0.04076$ <b>Correlación:</b> <b>0.329</b>	$PM2.5_{ug} =$ $0.27887 \cdot HUM\_REL$ $-0.04514 \cdot CO_{mg} +$ $0.42809 \cdot NO2_{ug}$ $+0.14992 \cdot$ $OZONO_{ug} -0.03737$ <b>Correlación:</b> <b>0.4767</b>

Para el horario de 17:00 a 20:00 se determinó como factores clave en al menos 3 estaciones: la humedad relativa, la temperatura ambiental, el monóxido de carbono, el dióxido de nitrógeno y el dióxido de azufre (Tabla 4).

Tabla 4. Regresión lineal múltiple normalizada para 17:00 a 20:00.

<b>Belisario</b>	<b>Carapungo</b>	<b>Cotocollao</b>	<b>El Camal</b>
$PM2.5_{ug} =$ $0.32263 \cdot HUM\_REL +$ $0.19155 \cdot TEMP\_AMB$ $+$ $0.18159 \cdot RAP\_VEC +$ $0.1617 \cdot CO_{mg} +$ $0.29844 \cdot NO2_{ug} +$ $0.11478 \cdot SO2_{ug}$ $-0.24041$ <b>Correlación:</b> <b>0.4865</b>	$PM2.5_{ug} =$ $0.05775 \cdot CO_{mg} +$ $0.28368 \cdot NO2_{ug} +$ $0.14531 \cdot OZONO_{ug}$ $+$ $0.35234 \cdot PM10_{ug} +$ $0.04034$ <b>Correlación=</b> <b>0.3204</b>	$PM2.5_{ug} =$ $0.21376 \cdot HUM\_REL +$ $0.20382 \cdot TEMP\_AMB$ $+ 0.34044 \cdot NO2_{ug} +$ $0.16095 \cdot OZONO_{ug}$ $+ 0.16919 \cdot SO2_{ug} +$ $-0.11034$ <b>Correlación:</b> <b>0.3616</b>	$PM2.5_{ug} =$ $0.14001 \cdot HUM\_REL +$ $0.09611 \cdot TEMP\_AMB$ $+ 0.14246 \cdot CO_{mg} +$ $0.325 \cdot NO2_{ug} +$ $0.1377 \cdot SO2_{ug} +$ $0.00217$ <b>Correlación:</b> <b>0.4671</b>

Para el horario de 21:00 a 05:00 se determinó como factores clave en al menos 3 estaciones: la humedad relativa, la temperatura ambiental, el monóxido de carbono y el dióxido de nitrógeno (Tabla 5).

Tabla 5. *Regresión lineal múltiple normalizada para 21:00 a 05:00.*

<b>Belisario</b>	<b>Carapungo</b>	<b>Cotocollao</b>	<b>El Camal</b>
$PM_{2.5\_ug} =$ $0.18929 \cdot HUM\_REL$ $+ 0.37723 \cdot$ $TEMP\_AMB +$ $0.73453 \cdot NO_2\_ug -$ $0.22564 \cdot OZONO\_ug$ $+ -0.22598$	$PM_{2.5\_ug} =$ $-0.14336 \cdot TEMP\_AMB +$ $0.15923 \cdot CO\_mg +$ $0.32883 \cdot NO_2\_ug +$ $0.31675 \cdot PM_{10\_ug} +$ $0.10696$	$PM_{2.5\_ug} =$ $0.17375 \cdot HUM\_REL$ $-0.0816 \cdot RAP\_VEC +$ $0.08749 \cdot CO\_mg +$ $0.34908 \cdot NO_2\_ug +$ $0.10321 \cdot SO_2\_ug +$ $0.00237$	$PM_{2.5\_ug} =$ $0.07564 \cdot HUM\_REL$ $+ 0.15777 \cdot CO\_mg$ $+ 0.23741 \cdot NO_2\_ug$ $-0.14029 \cdot$ $OZONO\_ug +$ $0.26714 \cdot SO_2\_ug +$ $0.07283$
<b>Correlación:</b> <b>0.642</b>	<b>Correlación=0.5201</b>	<b>Correlación:</b> <b>0.493</b>	<b>Correlación:</b> <b>0.5776</b>

La correlación promedio es 0.52 para los todos modelos generados por regresión lineal múltiple. De los modelos generados, los modelos de la mañana, tienen un coeficiente de correlación promedio de 0.638 que son los que se aproximan más a la realidad, principalmente debido a que no existe dilución de los gases y emisiones de PM<sub>2.5</sub> en la atmósfera (Rybarczyk & Zalakeviciute, 2018).

En conjunto los parámetros con mayor influencia de los estudiados son la Humedad Relativa, la Temperatura Ambiental, el NO<sub>2</sub> y el CO. De estos el único en cuya correlación es negativo es la Temperatura Ambiental en Carapungo.

Si se analiza la relación entre la Humedad Relativa y la Temperatura Ambiental se puede ver que son inversamente proporcionales (ver Anexo 2), y los estudios indican que cuando se reduce la temperatura y aumenta la Humedad Relativa del sitio se produce efectos de higroscopía (Ryu, Kim, Byeon, Go, & Lee, 2019), debido a que la atmósfera reduce su capacidad de contener una mayor cantidad de agua, por lo que los equipos registran una mayor concentración de PM<sub>2.5</sub>.

Por otra parte, el dióxido de nitrógeno y monóxido de carbono son dos contaminantes formados por la combustión incompleta o con exceso de oxígeno de los motores de combustibles fósiles, por lo cual es posible relacionar la concentración de los contaminantes el parque automotor.

El  $O_3$  se utiliza en la predicción de más de 11 de los modelos generados, por lo cual también es posible asociarlo directamente a la concentración de  $PM_{2.5}$ , sin embargo, es notable la presencia de este en las horas de madrugada, donde indica una reducción de la concentración. En teoría, debido a la falta de radiación solar no debe existir registros de alguna concentración de  $O_3$ , sin embargo, puede ser generada por otras fuentes o generarse por interacción de la misma atmósfera.

### 6.3. Árboles de Aprendizaje

Al determinar los resultados de árboles de decisión se estableció los siguientes árboles para las horas de la Mañana entre 06:00 a 11:59 (Figura 13), de la Tarde entre 12:00 a 16:59 (Figura 14), de la Noche entre 17:00 a 20:59 (Figura 15) y de la Madrugada entre 21:00 a 05:00 (Figura 16). En general el algoritmo determinó 4 instancias para cada uno de los árboles de decisión, a excepción de las estaciones Cotocollao y el Camal entre 21:00 y 05:59.

Para el horario de 06:00 a 11:59 (ver figura 13) el CO es el primer y principal clasificador de las estaciones, a excepción de Cotocollao, donde es el  $NO_2$ . Existe mayor cantidad de nodos terminales de concentración inferior a la mediana que superior a esta en todos los árboles, a excepción de Carapungo donde están equiparados. Los árboles se aproximan a un 70% de clasificación correcta de las instancias del histórico de 10 años.

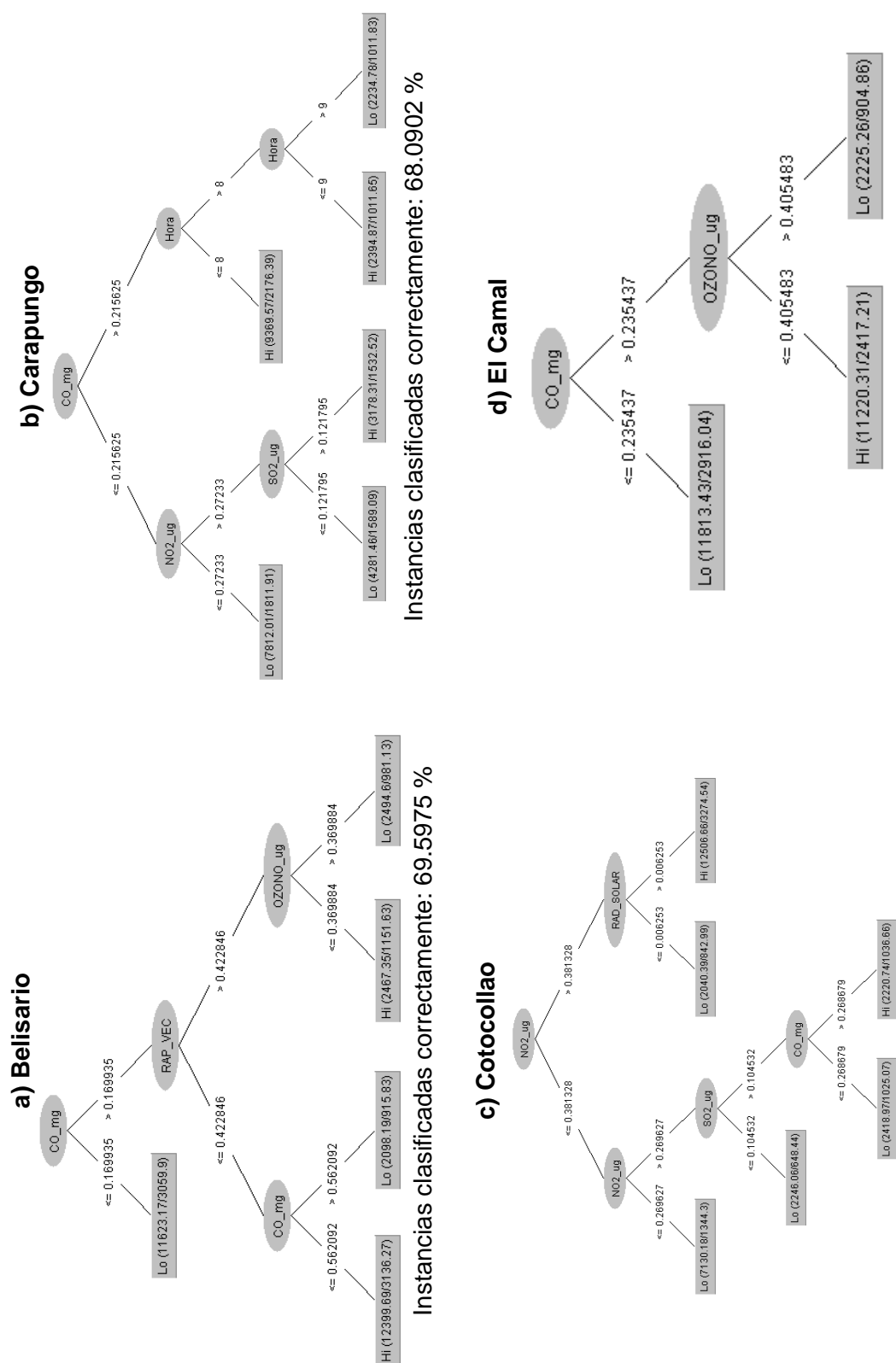
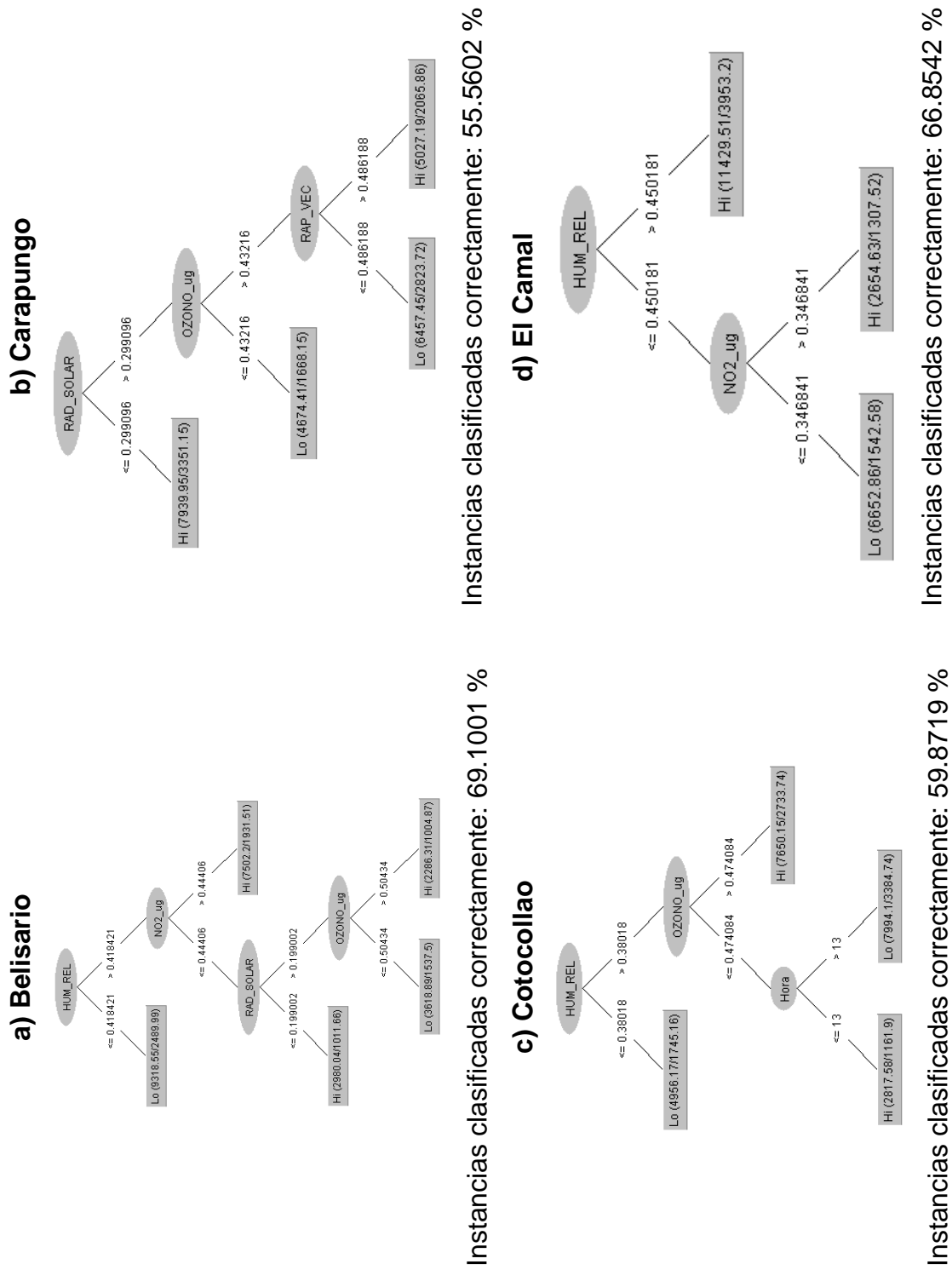


Figura 13. Árboles de decisión entre 06:00 a 11:59 para a) Belisario, b) Carapungo, c) Cotocollao, d) El Camal.

Figura 14. Árboles de decisión entre 12:00 a 16:59 para a) Belisario, b) Carapungo, c) Cotocollao, d) El Camal.



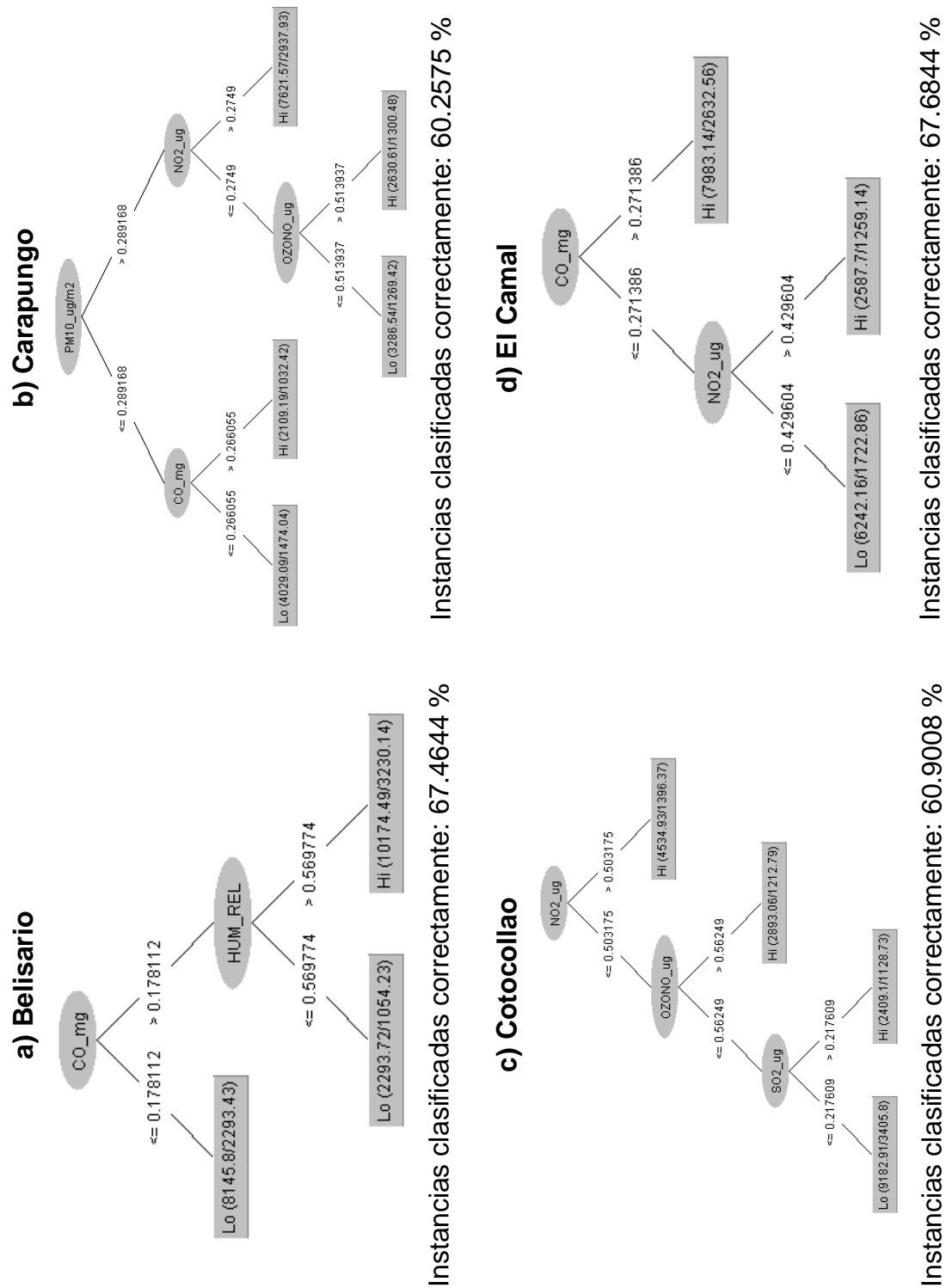


Figura 15. Árboles de decisión entre 17:00 a 20:59 para a) Belisario, b) Carapungo, c) Cotocollao, d) El Camal.



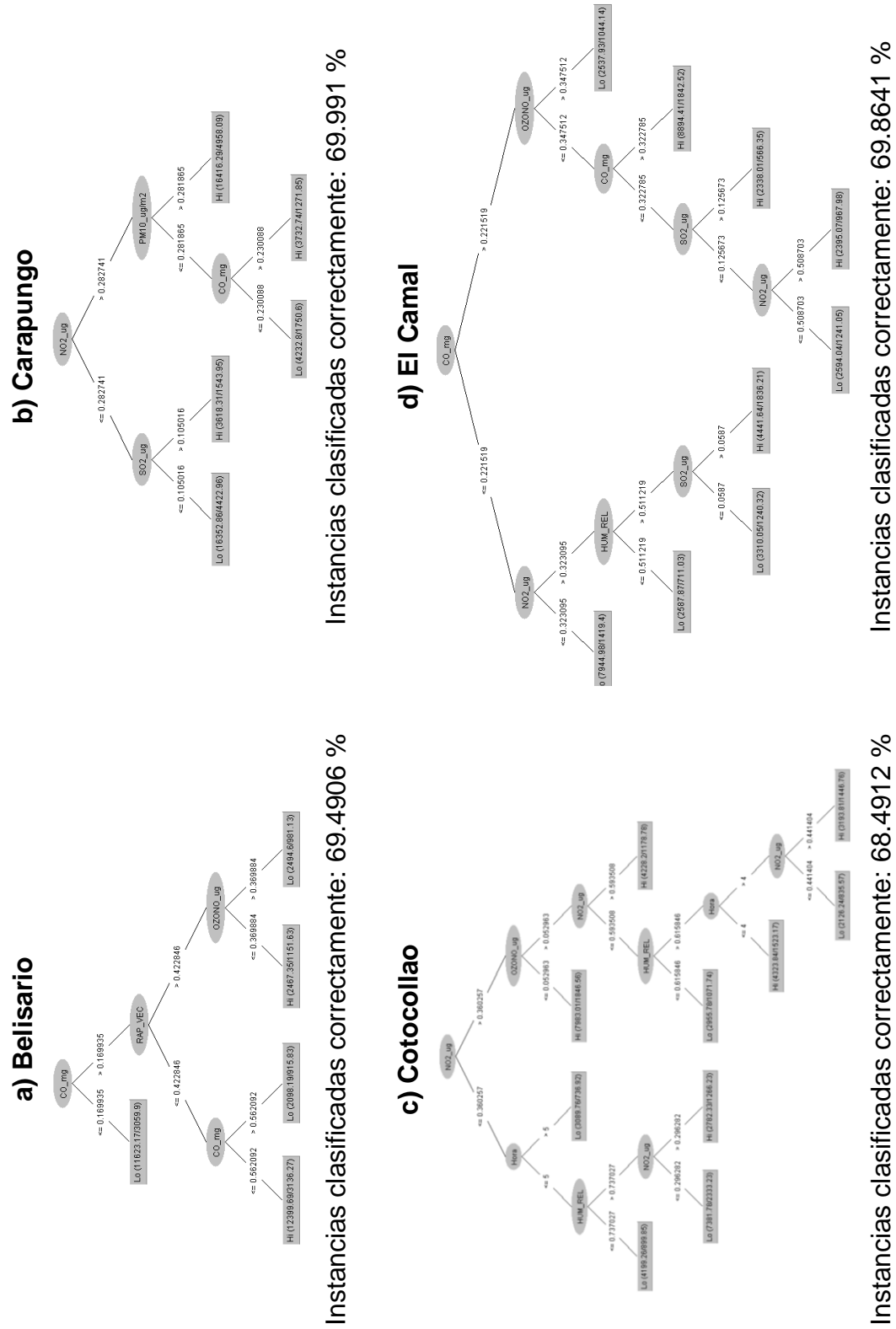


Figura 16. Árboles de decisión entre 21:00 a 05:59 para a) Belisario, b) Carapungo, c) Cotocollao, d) El Camal.

Belisario posee como clasificadores, tras el CO, la velocidad del viento y para sus ramales nuevamente el CO y el O<sub>3</sub>. Carapungo por su parte, después de determinar la concentración de CO, se divide nuevamente respecto a la concentración de NO<sub>2</sub> y SO<sub>2</sub>, y en su segundo ramal indica que la hora entre 6:00 hasta las 9:00 la concentración se mantendrá alta. Cotocollao, tras decidir el NO<sub>2</sub> en su primer ramal, dedujo la concentración alta en función del NO<sub>2</sub>, SO<sub>2</sub> y el CO (pero como último nodo de decisión del árbol), mientras que el segundo ramal del primer nodo indica que es importante la Radiación Solar, por lo cual puede verse implicado indirectamente la formación de O<sub>3</sub>. En El Camal tras la concentración de CO el O<sub>3</sub> es un elemento clave que define la concentración de forma directa.

Para el segundo horario de 12:00 a 16:59 (ver figura 14) la Humedad Relativa es el primer y principal clasificador de las estaciones, a excepción de Carapungo, donde es la Radiación Solar. Las estaciones de Carapungo y Cotocollao presentan un número de nodos igual para concentraciones bajas y altas respecto a la mediana, mientras que el Camal y Belisario indican una mayor cantidad de nodos terminales de concentración superior. Los árboles se aproximan a un 62% de clasificación correcta de las instancias del histórico de 10 años.

Belisario determina la concentración baja en el primer nodo en función de la Humedad relativa, y en el segundo ramal donde influye el NO<sub>2</sub>, la Radiación Solar y el O<sub>3</sub> en orden. Carapungo presenta concentraciones altas cuando existe una baja Radiación Solar, o si fuera el caso contrario, en función del O<sub>3</sub> y la Rapidez del Viento. Cotocollao presenta una concentración baja cuando la humedad relativa es baja, o en la segunda rama una concentración baja en función del ozono una hora superior a las 13:00. El Camal por su parte solo presenta concentraciones bajas en función de la humedad relativa y el NO<sub>2</sub>.

Para el tercer horario de 17:00 a 20:59 (ver figura 15) el CO es el primer nodo en Belisario y El Camal, mientras que en Carapungo y Cotocollao son el PM<sub>10</sub> y el NO<sub>2</sub> respectivamente. Las estaciones de Carapungo, Cotocollao y El Camal presentan un número de nodos mayor para concentraciones superiores respecto

a la mediana, mientras que Belisario presenta uno de tres nodos terminales con concentración superior. Los árboles se aproximan a un 64% de clasificación correcta de las instancias del histórico de 10 años.

Belisario presenta concentraciones altas en función del CO y luego según la Humedad Relativa. Carapungo presenta condiciones inferiores a la mediana en función del PM<sub>10</sub> y el CO, o en función del PM<sub>10</sub>, el NO<sub>2</sub> y el O<sub>3</sub>. Cotocollao clasifica la concentración baja en función del NO<sub>2</sub>, el O<sub>3</sub> y el SO<sub>2</sub>. El Camal por su parte presenta concentraciones bajas en función del CO y luego del NO<sub>2</sub>.

Para el cuarto y último horario de 21:00 a 05:59 (ver figura 16) el CO es el primer nodo en Belisario y El Camal, y el NO<sub>2</sub> para Carapungo y Cotocollao. En las estaciones de Belisario y El Camal presentan por un nodo una mayor cantidad de nodos con concentraciones superiores respecto a la media, Cotocollao tiene la misma cantidad de nodos terminales superiores e inferiores a la media, y Carapungo por un nodo posee mayor cantidad de nodos terminales inferiores a la media. Los árboles se aproximan a un 69% de clasificación correcta de las instancias del histórico de 10 años.

Belisario obtiene los valores superiores a la mediana en función de la concentración de CO, la velocidad del viento y de este nodo en función del CO nuevamente o del O<sub>3</sub>. Carapungo inicia con el nodo de NO<sub>2</sub>, y determina los nodos terminales con concentración inferior a la media en función del SO<sub>2</sub> o del PM<sub>10</sub> y seguido del CO. Cotocollao presenta un árbol más complejo que inicia también con el NO<sub>2</sub>, del cual su primer nodo identifica solo una concentración alta cuando la hora es inferior a las 05:00, según la humedad relativa y el NO<sub>2</sub>; en cuanto a la segunda rama del nodo inicial del NO<sub>2</sub> se puede definir las concentraciones inferiores a la media en función del O<sub>3</sub>, nuevamente el NO<sub>2</sub> y directamente en el nodo de la Humedad Relativa o de este mismo nodo cuando la hora es mayor a las 04:00 y a la concentración del NO<sub>2</sub>. El Camal presenta también un árbol complejo, su primer nodo es el CO, su primer ramal define como concentración alta en función del NO<sub>2</sub>, la Humedad Relativa y el SO<sub>2</sub>, y en su

segundo ramal define como concentración baja en el segundo nodo determinado por el  $O_3$ , o por la segunda rama del  $O_3$  en función del CO, el  $SO_2$  y el  $NO_2$ .

En general se obtuvo un 64% instancias correctamente clasificadas para los modelos propuestos. También puede apreciarse la presencia del CO y el  $NO_2$  en todos los árboles, lo cual concuerda con los resultados de las regresiones lineales múltiples, mientras que solo en las horas donde existe poco tráfico (12:00 a 16:59 y de 21:00 a 05:59) se visualizó la presencia de la Humedad Relativa como nodo clasificador, y la temperatura ambiental no es un factor decisivo para la clasificación de los árboles.

Es notable que las condiciones meteorológicas, exceptuando la Humedad Relativa, la Velocidad del Viento y la Radiación Solar, se vuelven factores poco relevantes para determinar si la concentración del  $PM_{2.5}$  supera la mediana, pero toma en cuenta con mayor frecuencia contaminantes como el CO,  $NO_2$ ,  $O_3$ ,  $SO_2$  y  $PM_{10}$ .

#### 6.4. Mapas de Interpolados de $PM_{2.5}$ .

En base a los algoritmos previos y las medianas de la ciudad se generó los mapas de contaminación de la ciudad de Quito en Figura 17, 18, 19 y 20. Puede apreciarse que la mayor concentración ocurre de 06:00 a 11:59 (Figura 20), luego de 17:00 a 20:59 (Figura 18), en tercer lugar, de 21:00 a 05:59 (Figura 17) y en último lugar de 12:00 a 16:59 (Figura 19).

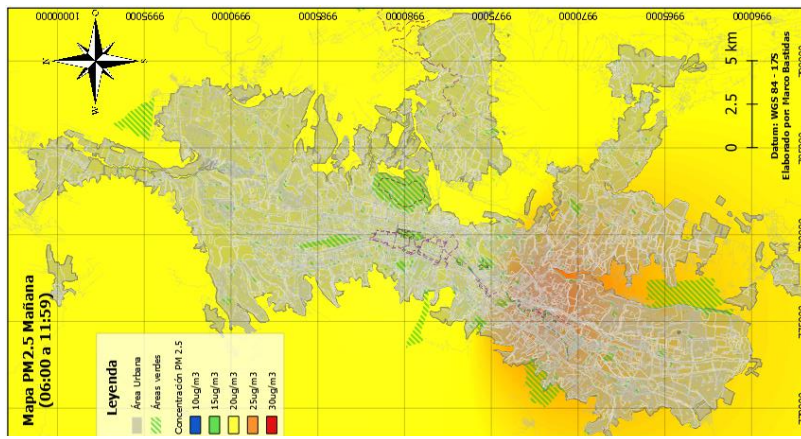
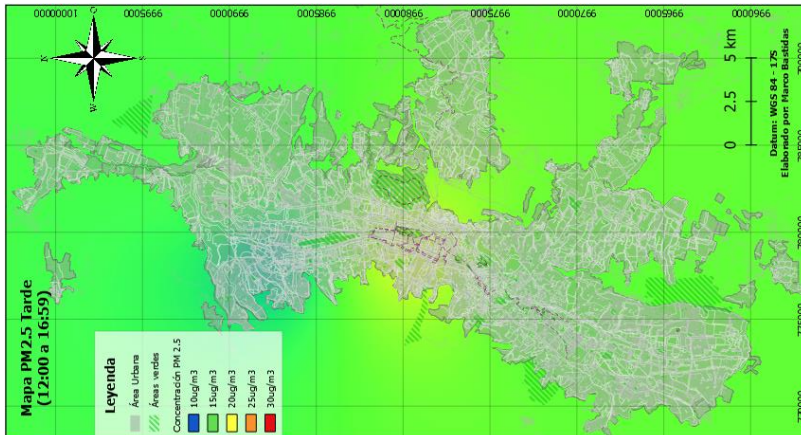
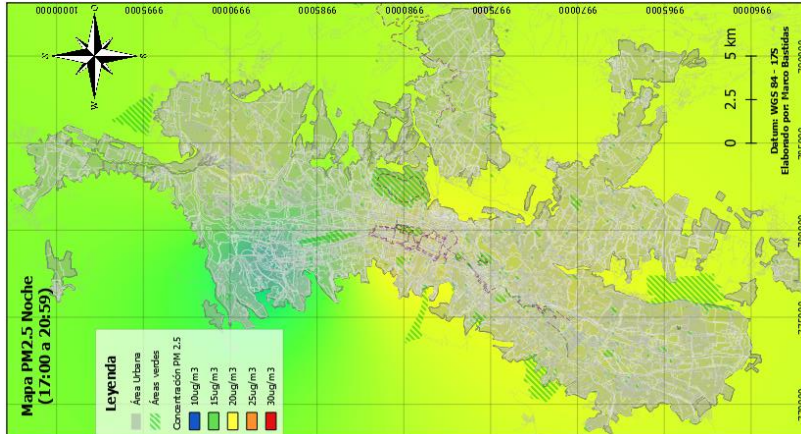
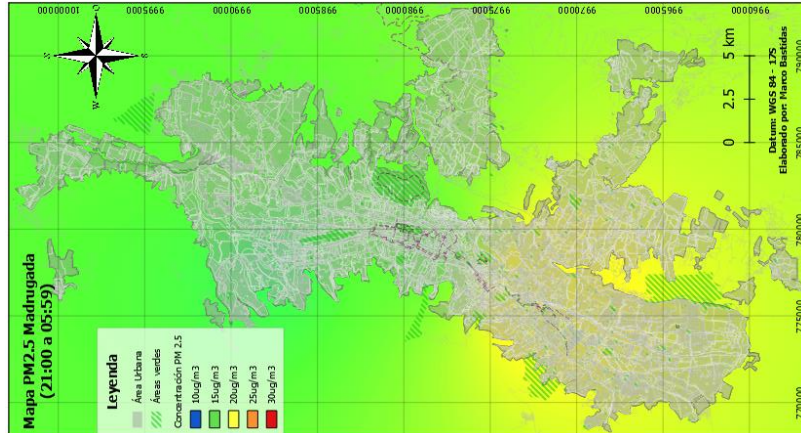


Figura 20. Mapa interpolado de PM2.5 de Quito para la Mañana (06:00 a 11:59).  
 Figura 19. Mapa interpolado de PM2.5 de Quito para la Tarde (12:00 a 16:59).  
 Figura 18. Mapa interpolado de PM2.5 de Quito para la Noche (17:00 a 20:59).  
 Figura 17. Mapa interpolado de PM2.5 de Quito para la Madrugada (21:00 a 05:59).

Las áreas donde se aprecia una mayor concentración de contaminantes son el área central de la ciudad de Quito en base a la interpolación. Lo cual puede ser lógico debido a la intensidad de flujo vehicular registrada, en especial por el área del centro de la ciudad (Figura 21).

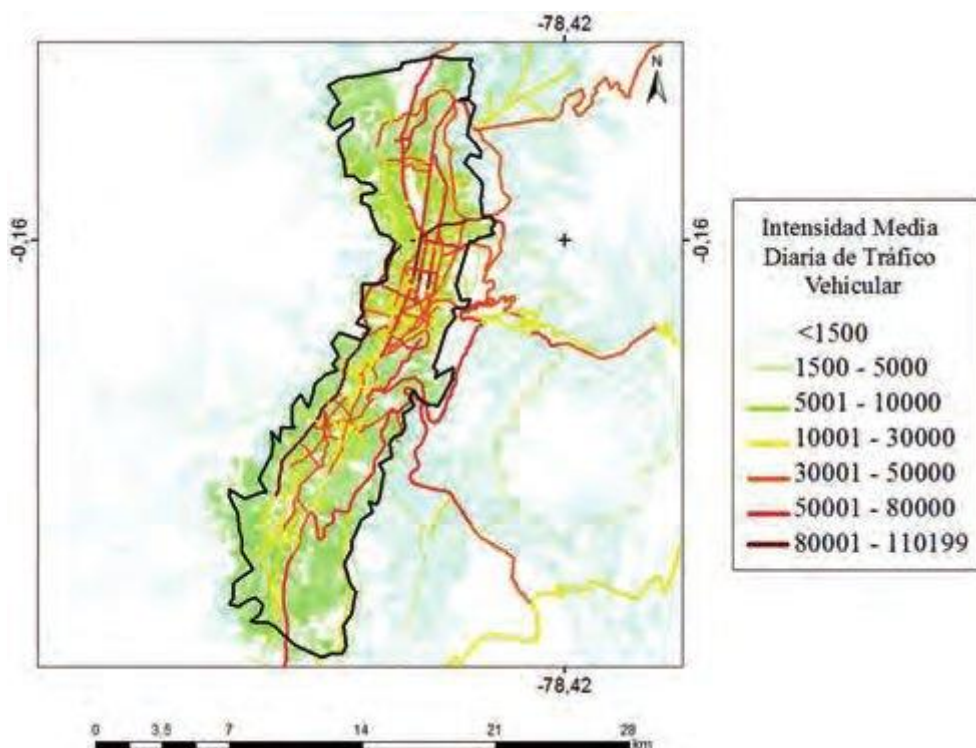


Figura 21. Mapa de la intensidad media diaria de tráfico vehicular del Distrito Metropolitano de Quito. Tomado de Vega & Parra, 2016.

### 6.5. Validación de mapas en campo

Las rutas de muestreo para validación del mapa se realizaron alrededor de la estación Belisario, debido a que es la que posee mayor cantidad de tráfico entre las horas pico de estudio y las mayores concentraciones de  $PM_{2.5}$ . El pico de contaminación se determinó entre las 07:30 y 10:30, horario en el que se realizó la validación mediante 3 rutas (Figura 22).

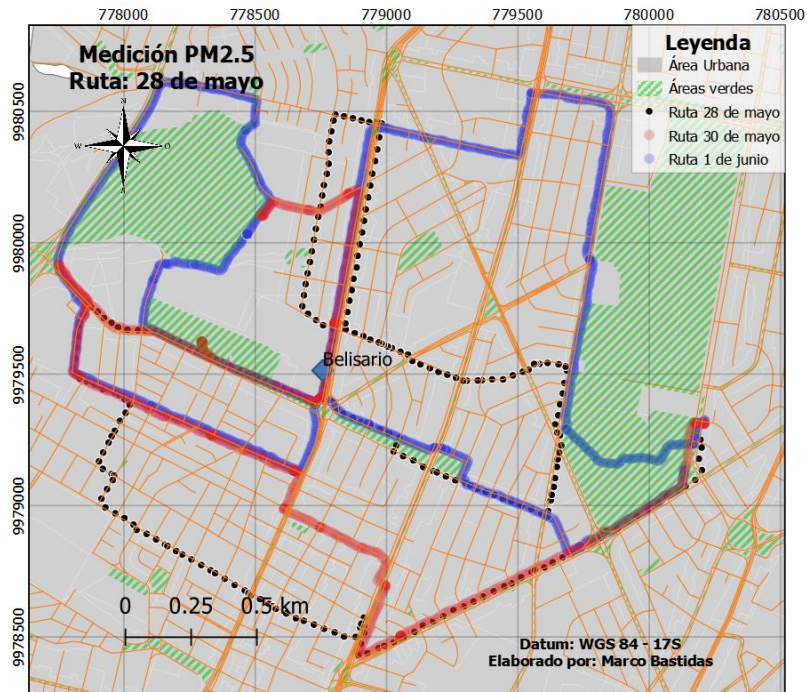


Figura 22. Mapa de rutas de muestreo para validación de Mapa de PM<sub>2.5</sub> entre 7:30 a 10:30.

Puede apreciarse una diferencia notable entre la interpolación modelada y las mediciones en todas en cada validación (Figuras 23, 24 y 25).

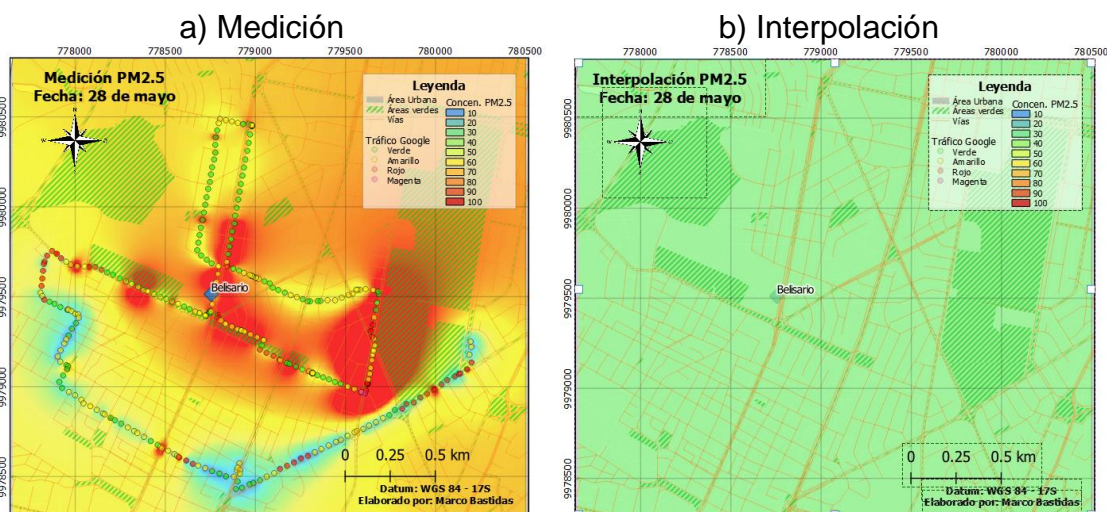


Figura 23. Comparación de concentración de PM<sub>2.5</sub> en muestreo (a) e interpolación (b) del 28 de mayo de 2019.

La primera validación durante el 28 de mayo registró que la concentración modelada de PM<sub>2.5</sub> se encontraba entre 30 y 40  $\mu\text{g}/\text{m}^3$ , mientras que la validación

presenta un rango entre 10 y 100  $\mu\text{g}/\text{m}^3$ , el error cuadrático medio entre la medición y la interpolación es de 36.47 (Figura 23).

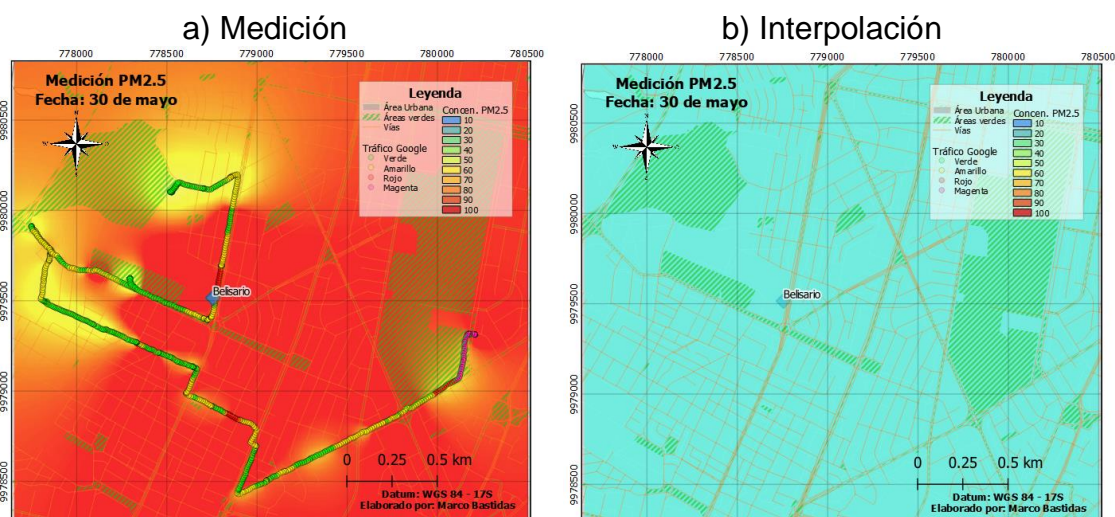


Figura 24. Comparación de concentración de PM<sub>2.5</sub> en muestreo (a) e interpolación (b) del 30 de mayo de 2019.

La segunda validación (30 de mayo), indica que la concentración modelada de PM<sub>2.5</sub> varió entre 20 y 30  $\mu\text{g}/\text{m}^3$ , pero la validación indica un rango superior a 50  $\mu\text{g}/\text{m}^3$ , el error cuadrático medio es 71,24, el cual es el más alto (Figura 24). Durante la validación existió una elevada humedad relativa y temperatura baja, lo cual debido a los efectos de higroscopía generó lecturas erróneas, pero a pesar de ello se puede apreciar una variación de concentración al separarse de avenidas principales.

La tercera y última validación durante el 3 de junio, presenta que la concentración modelada de PM<sub>2.5</sub> se encontraba cercano a 20  $\mu\text{g}/\text{m}^3$ , mientras que la validación presenta un rango entre 10 y 100  $\mu\text{g}/\text{m}^3$ , el error cuadrático medio es de 21,8 (Figura 25).



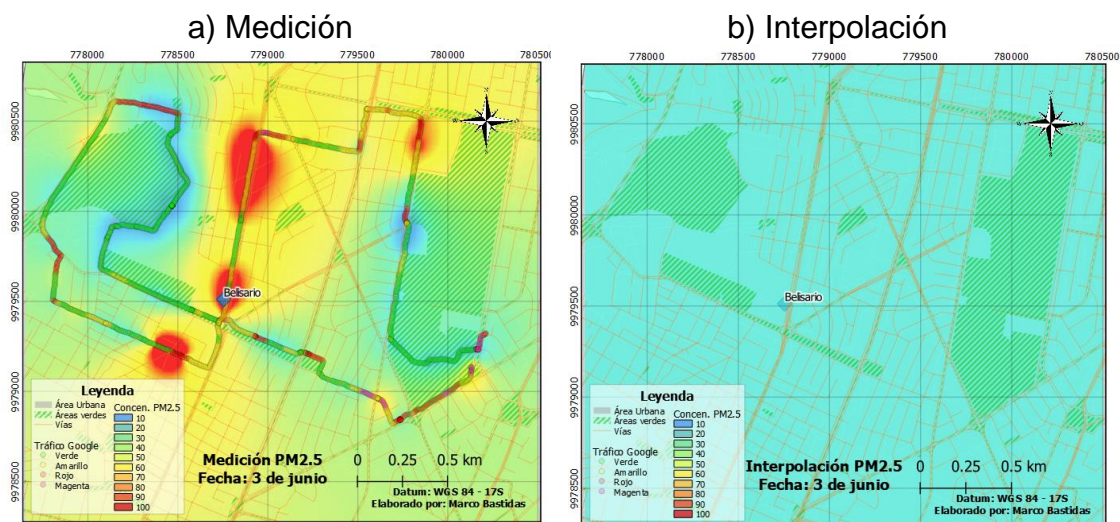


Figura 25. Comparación de concentración de PM<sub>2.5</sub> en muestreo (a) e interpolación (b) de 3 de junio de 2019.

Puede apreciarse respecto al tráfico, especialmente en Figura 23 y 25, que cuando las mayores concentraciones se presentan en las calles principales, o intersecciones con estas, cuando el tráfico está elevado (representado en la aplicación de *Google Traffic* por colores amarillo, rojo y magenta), confirmando que en la ciudad de Quito la principal fuente actual de PM<sub>2.5</sub> es el parque automotriz (CCAC & UNEP, 2017); mientras que en la cercanía o dentro de las áreas verdes (especialmente con una densidad vegetal alta) la concentración del PM<sub>2.5</sub> se ve reducida aún por debajo de lo estimado en la interpolación, lo cual se debe a la evapotranspiración de las plantas (Ryu et al., 2019), y aunque no se ha determinado claramente el mecanismo de funcionamiento para la remoción del PM<sub>2.5</sub>, es claro que las plantas poseen la capacidad para mitigar dicha concentración (Cohen, Potchter, & Schnell, 2014).

## 7. Conclusiones y Recomendaciones

### 7.1. Conclusiones

En síntesis, la concentración de  $PM_{2.5}$  puede predecirse numéricamente con regresión lineal principalmente a partir de las concentraciones de CO y  $NO_2$ , y las variables meteorológicas de humedad relativa y temperatura ambiental con una correlación de 0.52 a las mediciones de las estaciones. Mientras que los árboles de aprendizaje indican un 64% de instancias correctamente clasificadas mediante el uso de contaminantes como el CO,  $NO_2$ ,  $O_3$ ,  $SO_2$  y  $PM_{10}$  y variables meteorológicas como Humedad Relativa, Velocidad del Viento y Radiación Solar. Los mapas interpolados a partir de los algoritmos indican que la mayor concentración ocurre de 06:00 a 11:59 (Mañana), luego de 17:00 a 20:59 (Noche), en tercer lugar de 21:00 a 05:59 (Madrugada) y en último lugar de 12:00 a 16:59 (Tarde), pero mediante interpolación no poseen la suficiente resolución para discriminar concentraciones altas y bajas de  $PM_{2.5}$ , por lo que una alternativa para predecir con mayor exactitud la concentración del contaminante en cuestión es mediante la densidad vegetal presente, capaz de reducir el  $PM_{2.5}$  mediante evapotranspiración, y la intensidad del tráfico y flujo vehicular según el tipo de calle, debido a que es la principal fuente de contaminación en el DMQ.

### 7.2. Recomendaciones

Por lo descrito en el documento, se recomienda el implementar programas para reducir el tráfico vehicular, así también verificar frecuentemente la combustión vehicular a fin de reducir concentraciones elevadas de  $PM_{2.5}$ . Se recomienda el reforestar y aumentar las áreas verdes para reducir las concentraciones de  $PM_{2.5}$  mediante evapotranspiración. Dos alternativas para mejorar el monitoreo son implementar sensores de bajo costo de  $PM_{2.5}$  a altura de un ser humano promedio a fin de obtener un mapa más fiable con un mayor detalle; y en segundo lugar generar un modelo local basado en la densidad vegetal o forestal, el tráfico y el flujo vehiculares (el cual puede brindar mayor cantidad de detalles

que uno generado con las estaciones de medición actuales), a fin de ser una fuente de información antes de generar estrategias de prevención.

## Referencias

- Agencia Europea de Medio Ambiente. (2014). *Efectos de la contaminación atmosférica para la salud*. Recuperado de <https://www.eea.europa.eu/es/senales/senales-2013/infografia/efectos-de-la-contaminacion-atmosferica-2/view#tab-basado-en-datos>
- Amdur, M. O., Doull, J., & Klaassen, C. D. (1993). *Casarett and Doull's Toxicology: The Basic Science of Poisons*. (4.<sup>a</sup> ed.). En *Journal of Occupational and Environmental Medicine* (Vol. 35). Recuperado de <https://www.acgih.org/forms/store/ProductFormPublic/casarett-and-doull-s-toxicology-the-basic-science-of-poisons-8th-ed>
- Arrieta Fuentes, A. J. (2016). Dispersión de Material Particulado (PM10), con interrelación de factores Meteorológicos y Topográficos. *Ingeniería Investigación y Desarrollo*, 16(2), 43–54. <https://doi.org/10.19053/1900771x.v16.n2.2016.5445>
- Baca, J. (2011). *Secretaría de Ambiente Inventario de Emisiones de Gases del Efecto de Invernadero en el Distrito Metropolitano de Quito*. Recuperado el 25 de junio de 2018 de <http://revistas.usfq.edu.ec/index.php/avances/article/view/270/271>
- Barría, R. M., Calvo, M., & Pino, P. (2016). Contaminación intradomiciliaria por material particulado fino (MP2,5) en hogares de recién nacidos. *Revista Chilena de Pediatría*, 87(5), 343–350. <https://doi.org/10.1016/j.rchipe.2016.04.007>
- Bell, M. L., Peng, R. D., & Dominici, F. (2006). *The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations*. *Environmental health perspectives*, 114(4), 532–536. <https://doi.org/10.1289/ehp.8816>

- Bluett, J., Gimson, N., Fisher, G., Heydenrych, C., Freeman, T., & Godfrey, J. (2004). *Good practice guide for atmospheric dispersion modelling*. 522, 152. <https://doi.org/0-478-18941-9>
- BREATHELIFE. (2016). *City Data Page - BreatheLife 2030*. Recuperado 23 de abril de 2019, de <http://breathelife2030.org/city-data-page/?city=1465>
- Burger, S. (2018). *Introduction to machine learning with R: rigorous mathematical analysis* (1ª. ed.). Recuperado el 7 de mayo de 2019 de [https://books.google.com.ec/books?id=buhPDwAAQBAJ&printsec=frontcover&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](https://books.google.com.ec/books?id=buhPDwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false)
- CCAC & UNEP (2017). *Integrated Assessment of Short-Lived Climate Pollutants in LAC - Summary for Decision Makers*. En *Integrated Assessment of Short-Lived Climate Pollutants in Latin America and the Caribbean*. <https://doi.org/10.18356/d34a8b75-en>
- Chaparro, R., Cuervo, M., Gómez, J., & Toro, M. (2001). Emisiones al ambiente en Colombia. En *El Medido Ambiente en Colombia* (pp. 530–543). Recuperado el 25 de julio de 2019 de <http://documentacion.ideam.gov.co/openbiblio/bvirtual/000001/cap13.pdf>
- Chen, T. M., Gokhale, J., Shofer, S., & Kuschner, W. G. (2007). *Outdoor air pollution: Nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects*. *American Journal of the Medical Sciences*, 333(4), 249–256. <https://doi.org/10.1097/MAJ.0b013e31803b900f>
- Cohen, P., Potchter, O., & Schnell, I. (2014). *The impact of an urban park on air pollution and noise levels in the Mediterranean city of Tel-Aviv, Israel*. *Environmental Pollution*, 195, 73–83. <https://doi.org/10.1016/j.envpol.2014.08.015>

Conferencia de las Naciones Unidas sobre el Medio Ambiente y el Desarrollo. (1992). *Declaración de Río sobre el Medio Ambiente*. Recuperado el 16 de abril de 2019 de [http://www.unesco.org/education/pdf/RIO\\_S.PDF](http://www.unesco.org/education/pdf/RIO_S.PDF)

Costa, D. L., & Dreher, K. L. (1997). *Bioavailable transition metals in particulate matter mediate cardiopulmonary injury in healthy and compromised animal models*. *Environmental health perspectives*, 105(5), 1053–1060. <https://doi.org/10.2307/3433509>

De Visscher, A. (2013). *Air Dispersion Modelling: Foundations and Applications*. John Wiley & Sons.

Derwent, R. G., Parrish, D. D., Galbally, I. E., Stevenson, D. S., Doherty, R. M., Naik, V., & Young, P. J. (2018). *Uncertainties in models of tropospheric ozone based on Monte Carlo analysis: Tropospheric ozone burdens, atmospheric lifetimes and surface distributions*. *Atmospheric Environment*, 180(February), 93–102. <https://doi.org/10.1016/j.atmosenv.2018.02.047>

*Environmental Protection Agency [EPA]*. (1999). *Smog-Who Does It Hurt?* Recuperado el 14 de mayo de 2019 de <http://www.epa.gov/oar/>

Frank, E., Hall, M. A., & Witten, I. H. (2016). *The WEKA Workbench*. Recuperado el 17 de diciembre de 2018 de <https://www.cs.waikato.ac.nz/ml/weka/citing.html>

Galindo, E. (2015). *Estadística Métodos y aplicaciones* (3ª. ed.). Quito: ProCiencia Editores.

Gallego Picó, A., González Fernández, I., & Sánchez Gimeno, B. (2000). *Contaminación atmosférica*. Editorial UNED. Recuperado el 17 de junio de 2019 de:

[https://books.google.com.ec/books?id=7xYvtMyRP7EC&hl=es&source=gbs\\_navlinks\\_s](https://books.google.com.ec/books?id=7xYvtMyRP7EC&hl=es&source=gbs_navlinks_s)

GISGeography. (2018). *Inverse Distance Weighting (IDW) Interpolation - GIS Geography*. Recuperado 12 de mayo de 2019 de <https://gisgeography.com/inverse-distance-weighting-idw-interpolation/>

Ikechukwu, M. N., Ebinne, E., Idorenyin, U., & Raphael, N. I. (2017). *Accuracy Assessment and Comparative Analysis of IDW, Spline and Kriging in Spatial Interpolation of Landform (Topography): An Experimental Study*. *Journal of Geographic Information System*, 09(03), 354–371. <https://doi.org/10.4236/jgis.2017.93022>

Inche, J. L. (2004). Monitoreo Atmosférico. En *National Geographic* (pp. 69–82). Recuperado el 3 de mayo de 2019 de [http://sisbib.unmsm.edu.pe/bibvirtualdata/libros/geologia/gestion\\_calidad/Cap09.pdf](http://sisbib.unmsm.edu.pe/bibvirtualdata/libros/geologia/gestion_calidad/Cap09.pdf)

Karagulian, F., Belis, C. A., Dora, C. F. C., Prüss-Ustün, A. M., Bonjour, S., Adair-Rohani, H., & Amann, M. (2015). *Contributions to citie's ambient particulate matter (PM): A systematic review of local source contributions at global level*. *Atmospheric Environment*, Vol. 120, pp. 475–483. <https://doi.org/10.1016/j.atmosenv.2015.08.087>

Kim, H., Park, Y., Park, K., & Yoo, B. (2016). *Association between Pollen Risk Indexes, Air Pollutants, and Allergic Diseases in Korea*. *Osong Public Health and Research Perspectives*, 7(3), 172–179. <https://doi.org/10.1016/j.phrp.2016.04.003>

Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., & Rybarczyk, Y. (2017). *Modeling PM2.5Urban Pollution Using Machine Learning and Selected Meteorological Parameters*. *Journal of Electrical and*

*Computer Engineering*, 2017, 1–14.  
<https://doi.org/10.1155/2017/5106045>

Korc, M. E. (2000). *Situación de los Programas de Gestión de Calidad del Aire Urbano en América Latina y el Caribe*. Recuperado el 12 de junio de 2018 de [https://www.researchgate.net/publication/267839428\\_Situacion\\_de\\_los\\_Programas\\_de\\_Gestion\\_de\\_Calidad\\_del\\_Aire\\_Urbano\\_en\\_America\\_Latina\\_y\\_el\\_Caribe](https://www.researchgate.net/publication/267839428_Situacion_de_los_Programas_de_Gestion_de_Calidad_del_Aire_Urbano_en_America_Latina_y_el_Caribe)

Korc, M., & Sáenz, R. (1999). *Monitoreo de la Calidad del Aire en el distrito de Comas*. p. 20. Recuperado el 3 de mayo de 2019 de [https://www.researchgate.net/profile/Marcelo\\_Korc/publication/238739661\\_MONITOREO\\_DE\\_LA\\_CALIDAD\\_DEL\\_AIRE\\_EN\\_AMERICA\\_LATINA/links/02e7e53a85b76bf4bc000000.pdf](https://www.researchgate.net/profile/Marcelo_Korc/publication/238739661_MONITOREO_DE_LA_CALIDAD_DEL_AIRE_EN_AMERICA_LATINA/links/02e7e53a85b76bf4bc000000.pdf)

Krzyzanowski, M., Quackenboss, J. J., & Lebowitz, M. D. (1990). *Chronic respiratory effects of indoor formaldehyde exposure*. *Environmental Research*, 52(2), 117–125. [https://doi.org/10.1016/S0013-9351\(05\)80247-6](https://doi.org/10.1016/S0013-9351(05)80247-6)

Lantz, B. (2015). *Machine learning with R: discover how to build machine learning algorithms, prepare data, and dig deep into data prediction techniques with R*. Recuperado el 7 de mayo de 2019 de <http://www.worldcat.org/title/machine-learning-with-r-discover-how-to-build-machine-learning-algorithms-prepare-data-and-dig-deep-into-data-prediction-techniques-with-r/oclc/970590218>

Lee, H. S., & Kang, B.-W. (2001). *Chemical characteristics of principal PM<sub>2.5</sub> species in Chongju, South Korea*. *Atmospheric Environment*, 35(4), 739–746. [https://doi.org/10.1016/S1352-2310\(00\)00267-3](https://doi.org/10.1016/S1352-2310(00)00267-3)

Lesmeister, C. (2017). *Mastering Machine Learning with R Second Edition*



*Advanced prediction, algorithms, and learning methods with R 3.x.*  
 Recuperado el 7 de mayo de 2019 de  
[https://books.google.com.ec/books?id=50EwDwAAQBAJ&source=gb\\_s\\_similarbooks](https://books.google.com.ec/books?id=50EwDwAAQBAJ&source=gb_s_similarbooks)

Lippmann, M. (1989). *Health effects of ozone. A critical review.* *JAPCA*, 39(5), 672–695.  
 Recuperado de  
<http://www.ncbi.nlm.nih.gov/pubmed/2659744>

Lippmann, M. (2012). *Environmental toxicants: human exposures and their health effects.* En *Occupational and Environmental Medicine* (Vol. 57). New Jersey: John Wiley & Sons.

Ma, Z., Liu, Y., Zhao, Q., Liu, M., Zhou, Y., & Bi, J. (2016). *Satellite-derived high resolution PM2.5 concentrations in Yangtze River Delta Region of China using improved linear mixed effects model.* *Atmospheric Environment*, 133, 156–164.  
<https://doi.org/10.1016/j.atmosenv.2016.03.040>

Mamane, Y. (2008). *Fundamentals of air Pollution - Meteorology.* 303–304.  
 Recuperado el 28 de mayo de 2018 de  
<http://biblioteca.udla.edu.ec/client/default/search/results?qu=Vallero%2C+Daniel>

Martí, M. A. C. (2003). La ecotoxicología, una ciencia de hoy. *Medicina Balear*, 18. Recuperado el 26 de abril de 2019 de  
[http://ibdigital.uib.es/greenstone/collect/medicinaBalear/import/2003\\_v18\\_n3/Medicina\\_Balear\\_2003v18n3\\_p101.pdf](http://ibdigital.uib.es/greenstone/collect/medicinaBalear/import/2003_v18_n3/Medicina_Balear_2003v18n3_p101.pdf)

Martí Valls, J. (2017). Efectos de la calidad del aire sobre la salud. *FMC Formacion Medica Continuada en Atencion Primaria*, 24(9), 511–514.  
<https://doi.org/10.1016/j.fmc.2017.03.004>

- Martínez, A., & Romieu, I. (2003). Estrategia De Monitoreo. *Introducción al monitoreo atmosférico.*, 7–37. Recuperado el 3 de mayo de 2019 de <http://www.bvsde.ops-oms.org/bvsci/e/fulltext/intromon/cap2.pdf>
- Mata, A., & Quevedo, F. (2005). *Diccionario didáctico de ecología*. Editorial de la Universidad de Costa Rica.
- McGwin, G., Lienert, J., & Kennedy, J. I. (2010, marzo). *Formaldehyde exposure and asthma in children: A systematic review. Environmental Health Perspectives*, Vol. 118, pp. 313–317. <https://doi.org/10.1289/ehp.0901143>
- Mihelcic, J., & Zimmerman, J. B. (2012). *Ingeniería Ambiental: fundamentos, sustentabilidad, diseño* (1ª. ed.). México: Alfaomega Grupo Editor.
- Miller, E. W. (Eugene W., & Miller, R. M. (1989). *Environmental hazards. Air pollution: a reference handbook. ABC-CLIO*.
- Mills, G. (2014). *Urban climatology: History, status and prospects. Urban Climate*, 10, 479–489. <https://doi.org/10.1016/J.UCLIM.2014.06.004>
- Mitas, L., & Mitasova, H. (2005). *Spatial Interpolation*. En D. J. Longley, P.A. Goodchild, M.F. Maguire & D. W. Rhind (Eds.), *Geographic Information Systems: Principles, Techniques, Management and Applications* (2ª. ed., pp. 129–136). <https://doi.org/10.4018/978-1-59140-995-3.ch017>
- Moscoso Vanegas, D. L., Astudillo Alemán, A. L., & Vázquez Freire, V. E. (2015). Modelamiento de la calidad del aire en la ciudad de Cuenca-Ecuador. *ITECKNE*, 12(2), 188–197. <https://doi.org/10.15332/iteckne.v12i2.1245>
- Nuvolone, D., Petri, D., & Voller, F. (2018). *The effects of ozone on human health*.

*Environmental Science and Pollution Research*, 25(9), 8074–8088.  
<https://doi.org/10.1007/s11356-017-9239-3>

O'Hara, G. (2007). *An Economic History of Twentieth-Century Europe*. En *The English Historical Review*.  
<https://doi.org/10.1017/CBO9780511800627>

Organización Mundial de la Salud [OMS]. (2017). *Respira la vida. Aire limpio, futuro saludable*. WHO. Recuperado el 16 de enero de 2019 de <https://www.who.int/phe/breathe-life/es/>

Pacher, P., Beckman, J. S., & Liaudet, L. (2007). *Nitric Oxide and Peroxynitrite in Health and Disease*. *Physiological Reviews*, 87(1), 315–424.  
<https://doi.org/10.1152/physrev.00029.2006>

Parlamento del Reino Unido. (1956). *Clean Air Act*. *Nature*, Vol. 179, p. 615.  
<https://doi.org/10.1038/179615e0>

Programa de las Naciones Unidas para el Desarrollo [PNUD] (2019). *Objetivo 3: Salud y bienestar*. Recuperado 29 de abril de 2019, de <https://www.undp.org/content/undp/es/home/sustainable-development-goals/goal-3-good-health-and-well-being.html>

Programa de las Naciones Unidas para el Desarrollo [PNUD] & Asociación para combustibles limpios y vehículos. (2017). *Niveles de azufre del combustible diesel* (p. 1). Recuperado de [https://wedocs.unep.org/bitstream/handle/20.500.11822/17532/MapL\\_ACSulph\\_March2017.pdf?sequence=1&isAllowed=y](https://wedocs.unep.org/bitstream/handle/20.500.11822/17532/MapL_ACSulph_March2017.pdf?sequence=1&isAllowed=y)

Puigcerver Zanón, M., & Carrascal Triola, M. D. (2008). *El Medio atmosférico: meteorología y contaminación*. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.

- Real Academia Española [RAE]. (2019). *Diccionario de la Lengua Española*. Recuperado el 27 de mayo de 2019 de <https://dle.rae.es/?id=AU1m1dd>
- Roda, C., Kousignian, I., Guihenneuc-Jouyaux, C., Dassonville, C., Nicolis, I., Just, J., & Momas, I. (2011). *Formaldehyde exposure and lower respiratory infections in infants: Findings from the paris cohort study*. *Environmental Health Perspectives*, 119(11), 1653–1658. <https://doi.org/10.1289/ehp.1003222>
- Romero, S. (2014). La contaminación provoca 7 millones de muertes al año. *MUY INTERESANTE*, 42, 1–3. Recuperado el 22 de mayo de 2018 de <http://www.who.int/mediacentre/news/releases/2014/air-pollution/es/>
- RStudio Team. (2016). *RStudio server: Integrated Development for R*. Recuperado el 27 de noviembre de 2018 de <http://www.rstudio.com/>
- Rusznak, C., Devalia, J. L., & Davies, R. J. (1994). *The impact of pollution on allergic disease*. *Allergy*, 49(s18), 21–27. <https://doi.org/10.1111/j.1398-9995.1994.tb04234.x>
- Rybarczyk, Y., & Zalakeviciute, R. (2018). *Regression Models to Predict Air Pollution from Affordable Data Collections*. En *Machine Learning - Advanced Techniques and Emerging Applications*. <https://doi.org/10.5772/intechopen.71848>
- Ryu, J., Kim, J. J., Byeon, H., Go, T., & Lee, S. J. (2019). *Removal of fine particulate matter (PM 2.5 ) via atmospheric humidity caused by evapotranspiration*. *Environmental Pollution*, 245, 253–259. <https://doi.org/10.1016/j.envpol.2018.11.004>
- Sbarato, D., Sbarato, V. M., & Ortega, J. E. (2007). Predicción y evaluación de

impactos ambientales sobre la atmosfera/*Prediction and assessment of environmental impacts about atmosphere*. Recuperado el 7 de mayo de 2019 de [https://books.google.com.ec/books?id=xNnHxgwU\\_toC&printsec=frontcover&hl=es&source=gbs\\_vpt\\_reviews#v=onepage&q&f=false](https://books.google.com.ec/books?id=xNnHxgwU_toC&printsec=frontcover&hl=es&source=gbs_vpt_reviews#v=onepage&q&f=false)

Schneider, P., Castell, N., Vogt, M., Dauge, F. R., Lahoz, W. A., & Bartonova, A. (2017). *Mapping urban air quality in near real-time using observations from low-cost sensors and model information*. *Environment International*. <https://doi.org/10.1016/j.envint.2017.05.005>

Secretaría del Ambiente DMQ. (2015). *Informe Anual 2014 Calidad del Aire del Distrito Metropolitano de Quito* (Vol. 1). <https://doi.org/10.1017/CBO9781107415324.004>

Seoánez Calvo, M. (2008). *Tratado de la contaminación atmosférica: problemas, tratamiento y gestión*. p. 1111. Recuperado el 26 de abril de 2019 de [https://books.google.com.ec/books/about/Tratado\\_de\\_la\\_contaminación\\_atmosférica.html?id=hqf\\_wAEACAAJ&source=kp\\_cover&redir\\_esc=y](https://books.google.com.ec/books/about/Tratado_de_la_contaminación_atmosférica.html?id=hqf_wAEACAAJ&source=kp_cover&redir_esc=y)

Siabato, W. L., Sánchez, A. H., & Cely, J. W. (2002). Modelo de calidad del aire para Bogotá, Bogotá D.C. *Ingeniería*, 7(1), 65–71. Recuperado el 22 de mayo de 2018 de <https://dialnet.unirioja.es/descarga/articulo/4797179.pdf>

Soraya, E. (2013). *Modelación de la dispersión de las emisiones de SO<sub>2</sub> y NO<sub>x</sub> de las termoeléctricas en la ciudad de Quito en el software CALPUFF* (Quito, 2013.). Recuperado el 3 de julio de 2018 de <http://bibdigital.epn.edu.ec/handle/15000/8069>

Swackhamer, D. L. (2003). *Rethinking the Ozone Problem in Urban and Regional Air Pollution*. *Journal of Aerosol Science*, 24(7), 977–978.

[https://doi.org/10.1016/0021-8502\(93\)90076-1](https://doi.org/10.1016/0021-8502(93)90076-1)

Team, Q. D. C. (2013). *QGIS Geographic Information System. Open Source Geospatial Foundation Project*. Vienna, Austria: <http://qgis.osgeo.org>. Recuperado el 17 de diciembre de 2018 de <https://www.qgis.org>

Tripathi, A. (2017). *Practical Machine Learning Cookbook - Resolving and offering solutions to your machine learning problems with R*. Recuperado el 7 de mayo de 2019 de [https://books.google.com.ec/books?id=HUIwDwAAQBAJ&source=gb\\_s\\_similarbooks](https://books.google.com.ec/books?id=HUIwDwAAQBAJ&source=gb_s_similarbooks)

US EPA, O. (s. f.). *Particulate Matter (PM) Basics*. Recuperado de <https://www.epa.gov/pm-pollution/particulate-matter-pm-basics#PM>

Vega, D., & Parra, R. (2016). Caracterización de la intensidad media diaria y de los perfiles horarios del tráfico vehicular del Distrito Metropolitano de Quito. *Avances en Ciencias e Ingeniería*, 6(2). <https://doi.org/10.18272/aci.v6i2.186>

Wallace, J., & Kanaroglou, P. (2009). *The effect of temperature inversions on ground-level nitrogen dioxide (NO<sub>2</sub>) and fine particulate matter (PM<sub>2.5</sub>) using temperature profiles from the Atmospheric Infrared Sounder (AIRS)*. *Science of The Total Environment*, 407(18), 5085–5095. <https://doi.org/10.1016/J.SCITOTENV.2009.05.050>

Wettstein, Z. S., Hoshiko, S., Fahimi, J., Harrison, R. J., Cascio, W. E., & Rappold, A. G. (2018). *Cardiovascular and cerebrovascular emergency department visits associated with wildfire smoke exposure in California in 2015*. *Journal of the American Heart Association*, 7(8), 213–224. <https://doi.org/10.1161/JAHA.117.007492>

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical*

*Machine Learning Tools and Techniques*. En *Data Mining: Practical Machine Learning Tools and Techniques*.  
<https://doi.org/10.1016/c2009-0-19715-5>

Yang, H., Yu, J. Z., Ho, S. S. H., Xu, J., Wu, W.-S., Wan, C. H., ... Wang, L. (2005). *The chemical composition of inorganic and carbonaceous materials in PM<sub>2.5</sub> in Nanjing, China*. *Atmospheric Environment*, 39(20), 3735–3749.  
<https://doi.org/10.1016/J.ATMOSENV.2005.03.010>

Yang, X. (2010). *Spatial Interpolation*. En *Handbook of Research on Geoinformatics* (pp. 129–136). <https://doi.org/10.4018/978-1-59140-995-3.ch017>

Yu-Wei, C. (2015). *Machine learning with R cookbook: explore over 110 recipes to analyze data and build predictive models with the simple and easy-to-use R code*. Recuperado el 7 de mayo de 2019 de [https://books.google.com.ec/books?id=iPirBwAAQBAJ&source=gbs\\_similarbooks](https://books.google.com.ec/books?id=iPirBwAAQBAJ&source=gbs_similarbooks)

Zhang, K., & Batterman, S. (2013). *Air pollution and health risks due to vehicle traffic*. *Science of the Total Environment*, 450–451, 307–316.  
<https://doi.org/10.1016/j.scitotenv.2013.01.074>

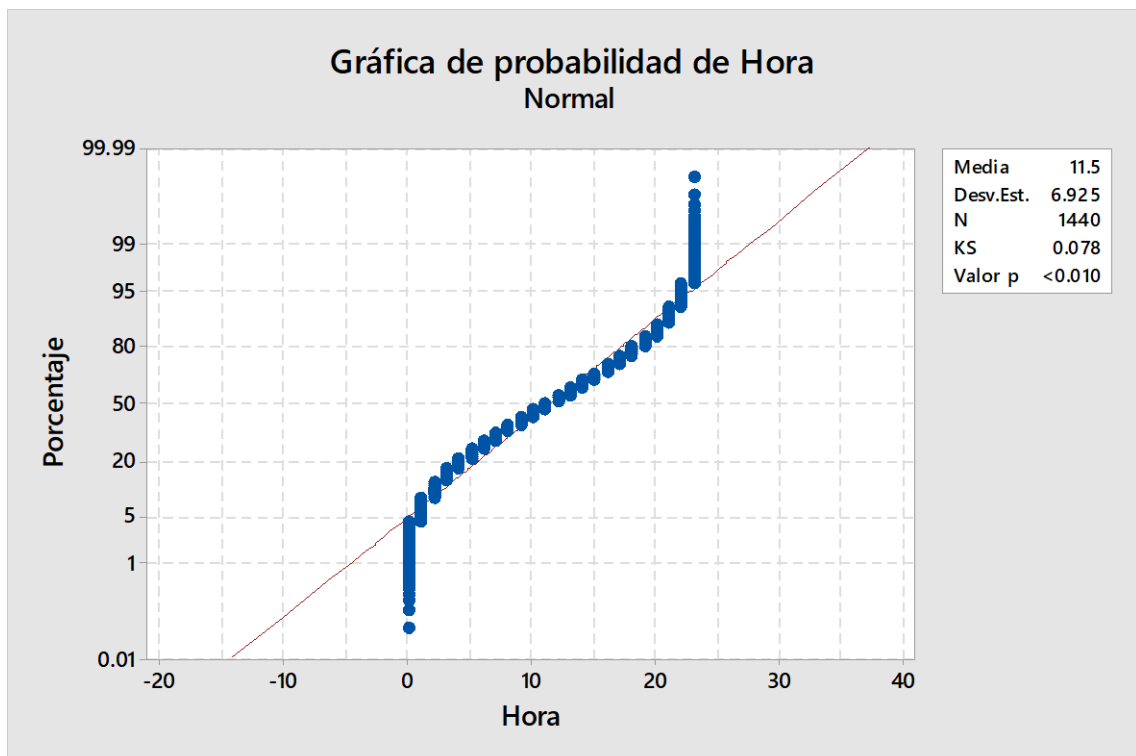
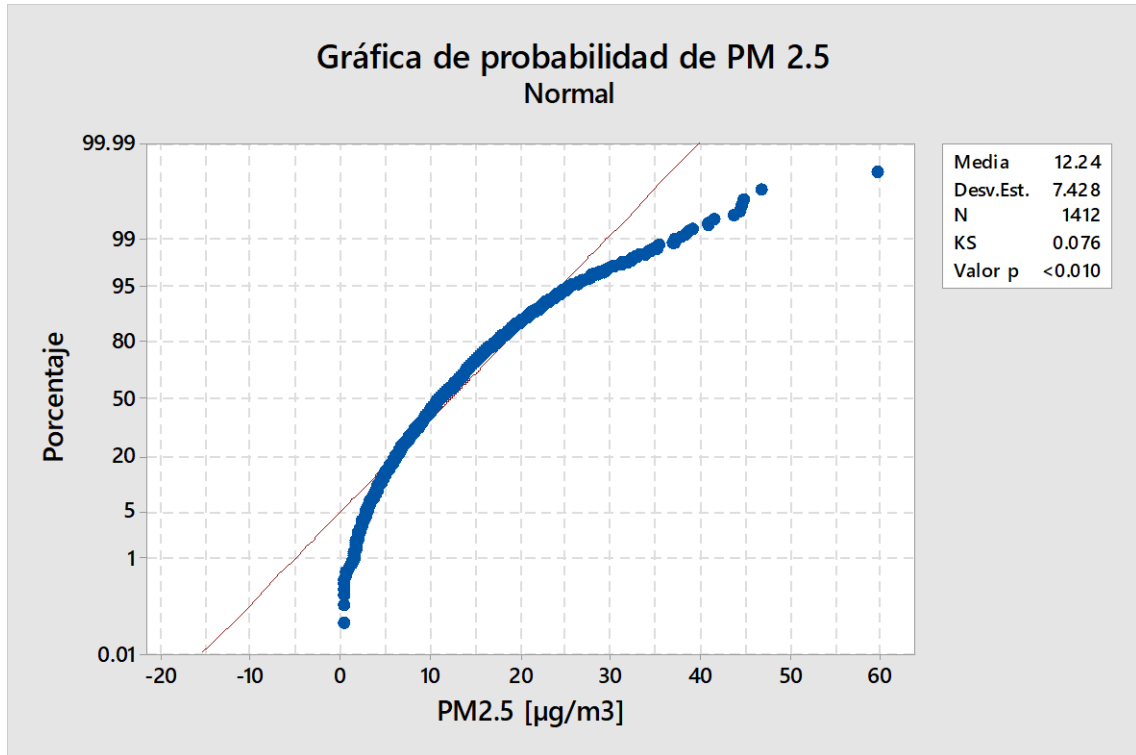
Zhang, Y., Huang, W., Cai, T., Fang, D., Wang, Y., Song, J., Hu M. & Zhang, Y. (2016). *Concentrations and chemical compositions of fine particles (PM<sub>2.5</sub>) during haze and non-haze days in Beijing*. *Atmospheric Research*, 174–175, 62–69.  
<https://doi.org/10.1016/J.ATMOSRES.2016.02.003>

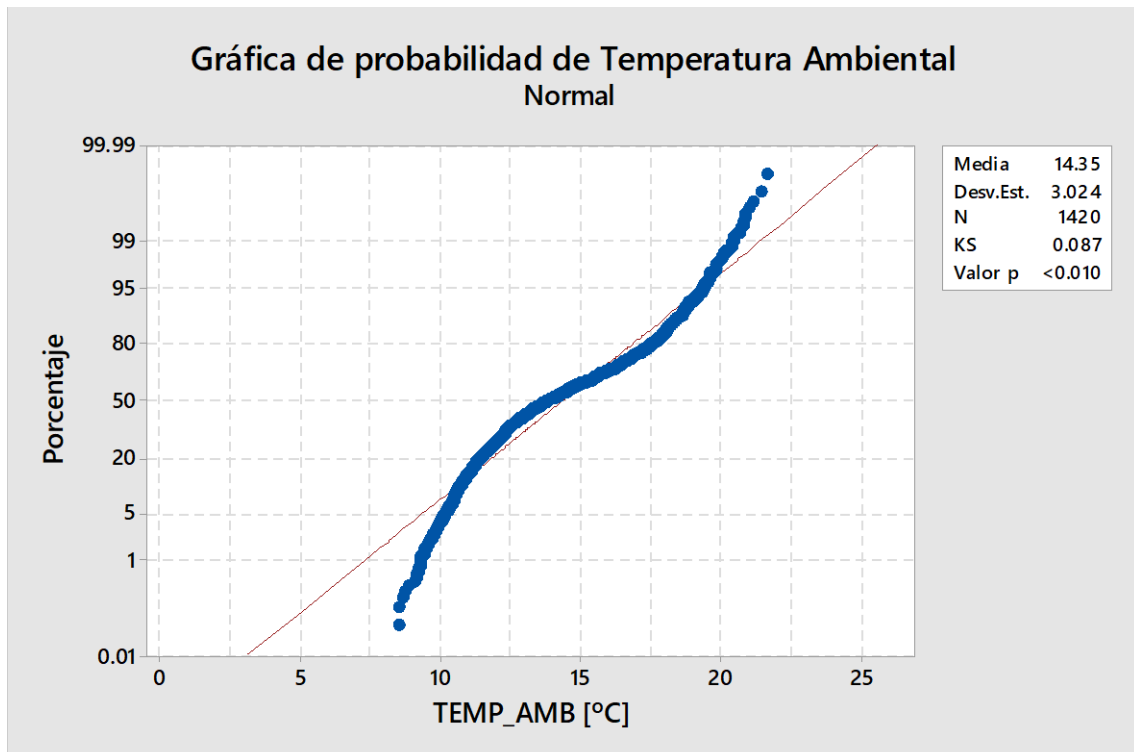
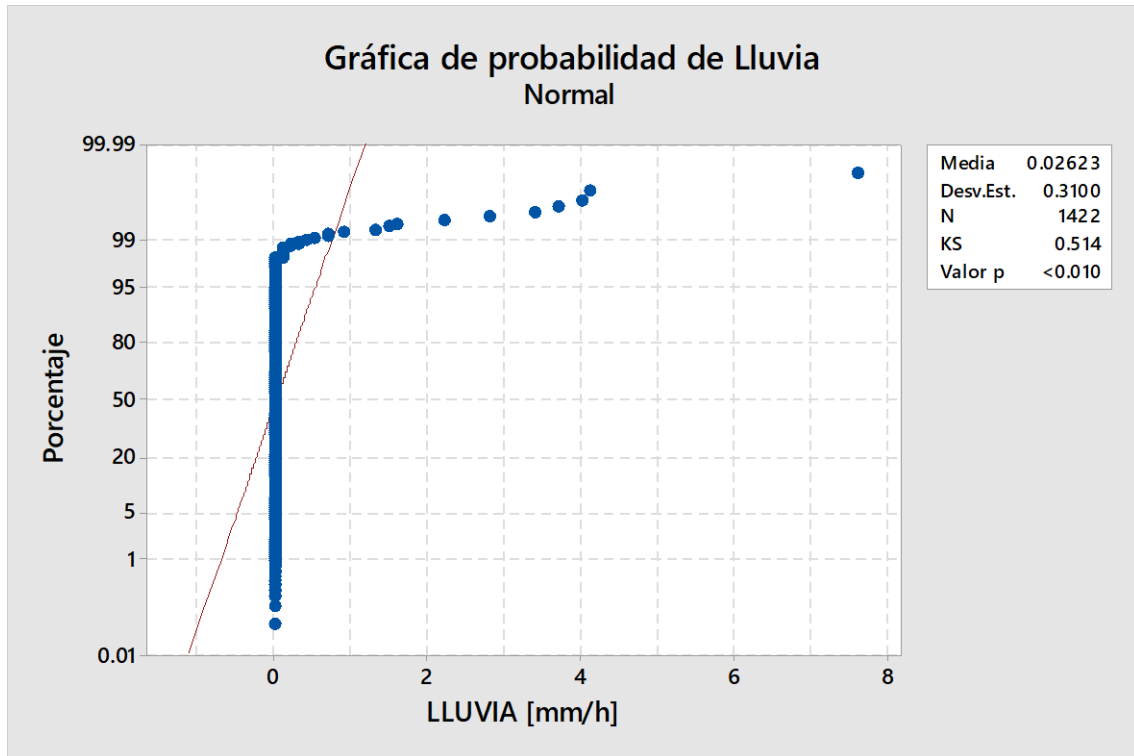
## **ANEXOS**



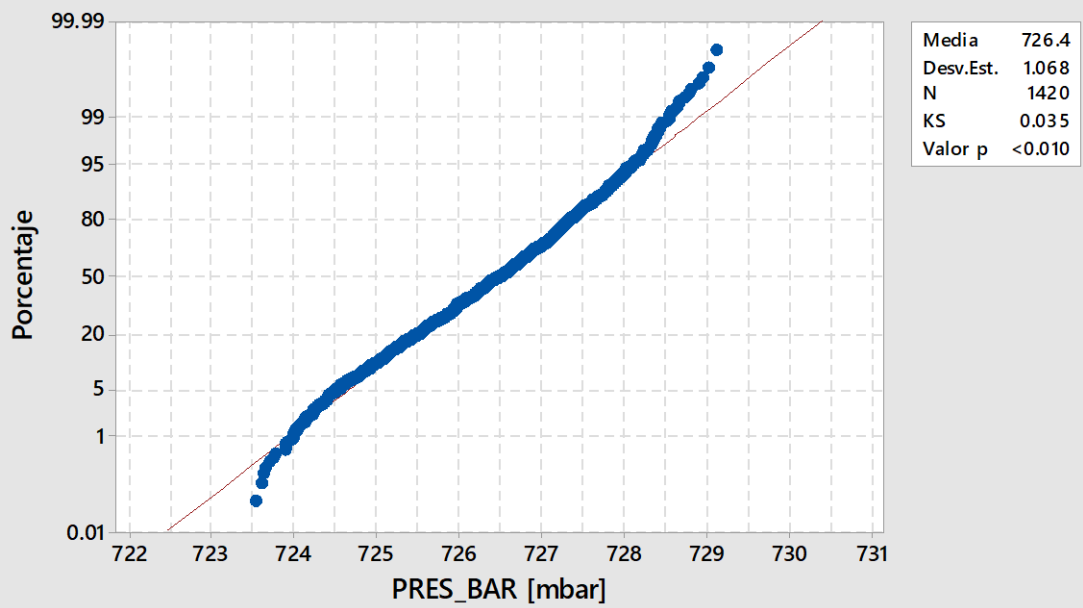
### Anexo 1: Pruebas de normalidad por componente

Las pruebas de normalidad se realizaron para el mes de abril en la Estación de Belisario.

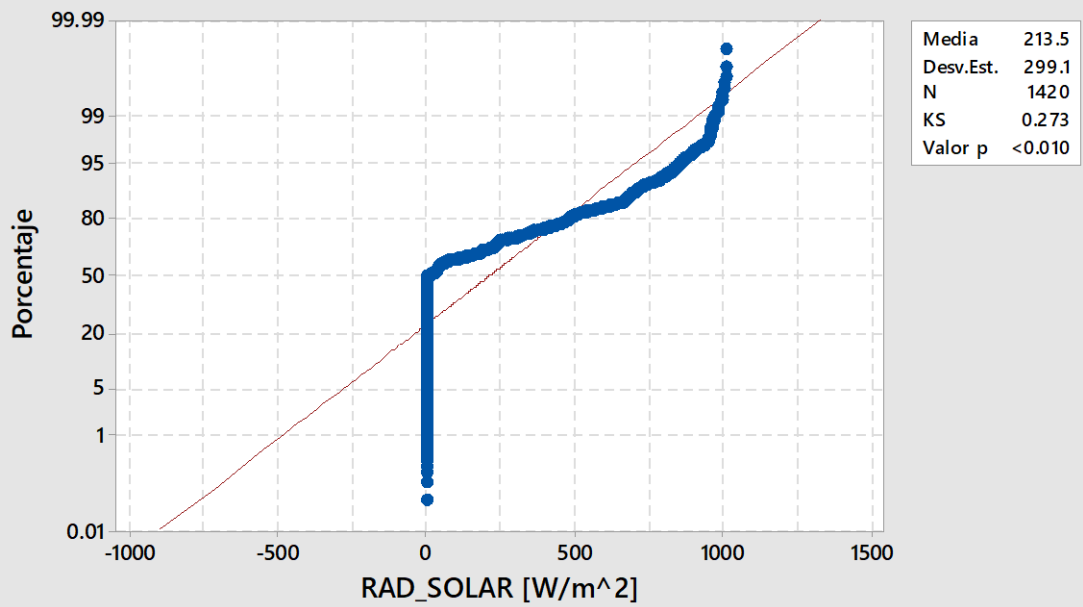




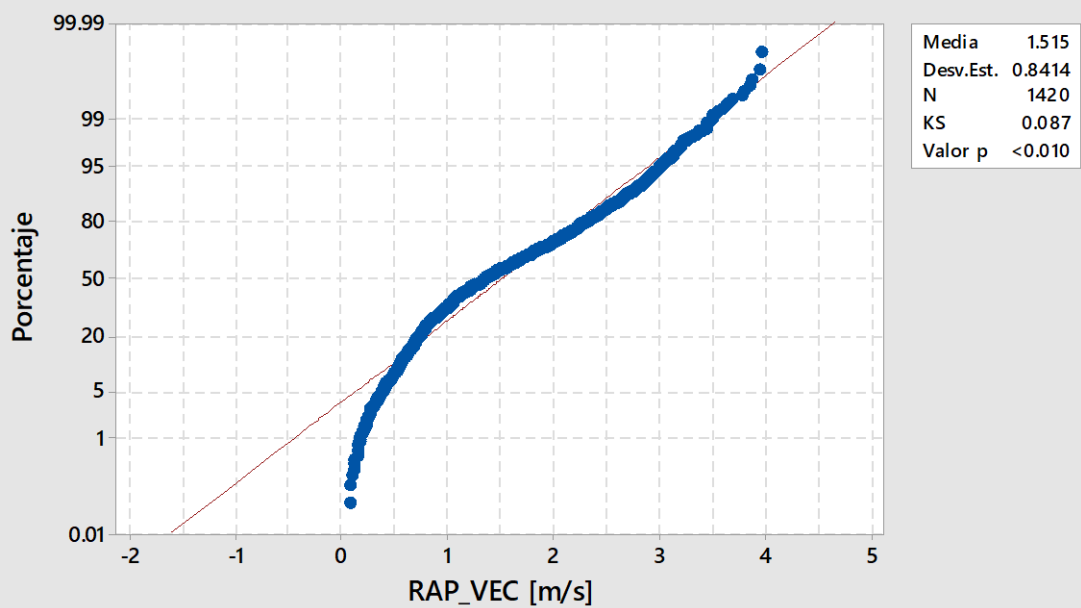
Gráfica de probabilidad de Presión Barométrica Normal



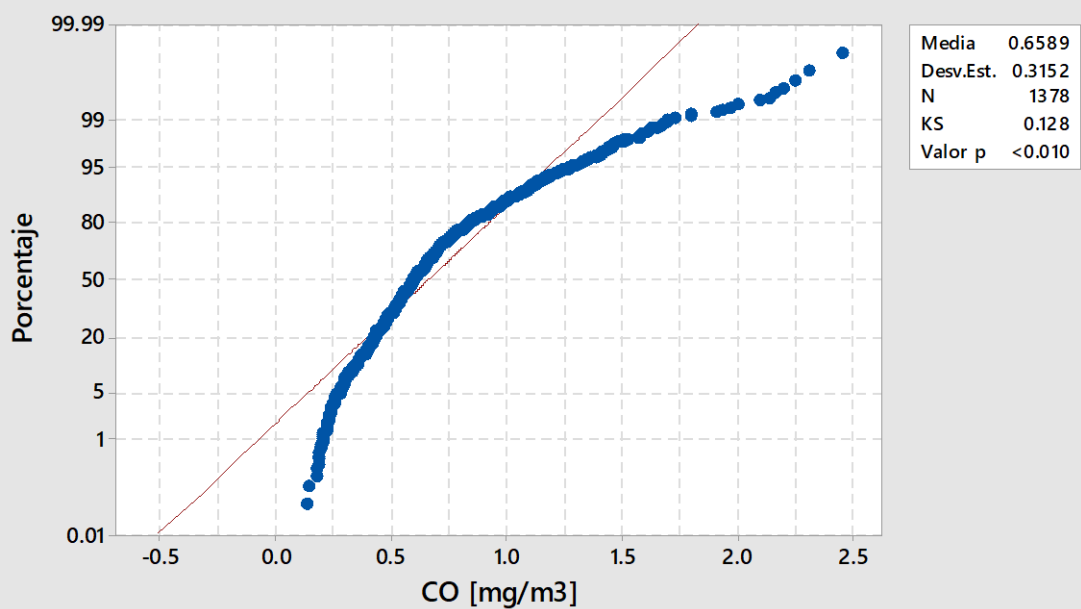
Gráfica de probabilidad de Radiación Solar Normal

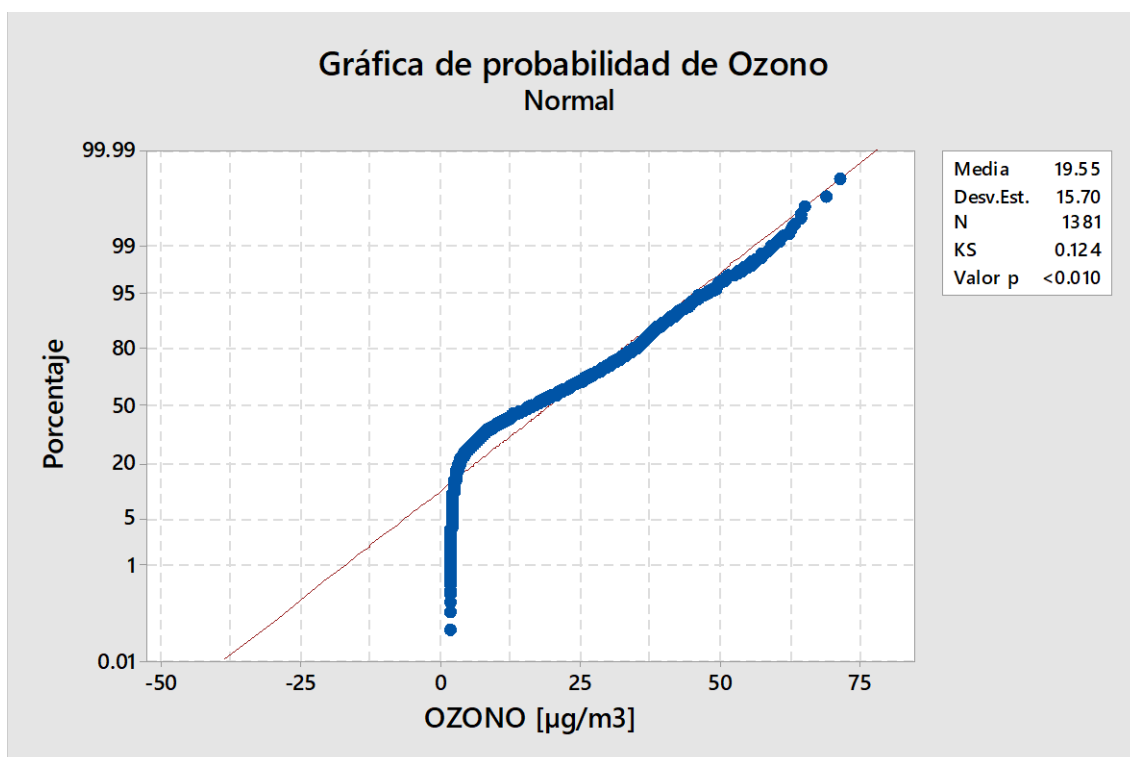
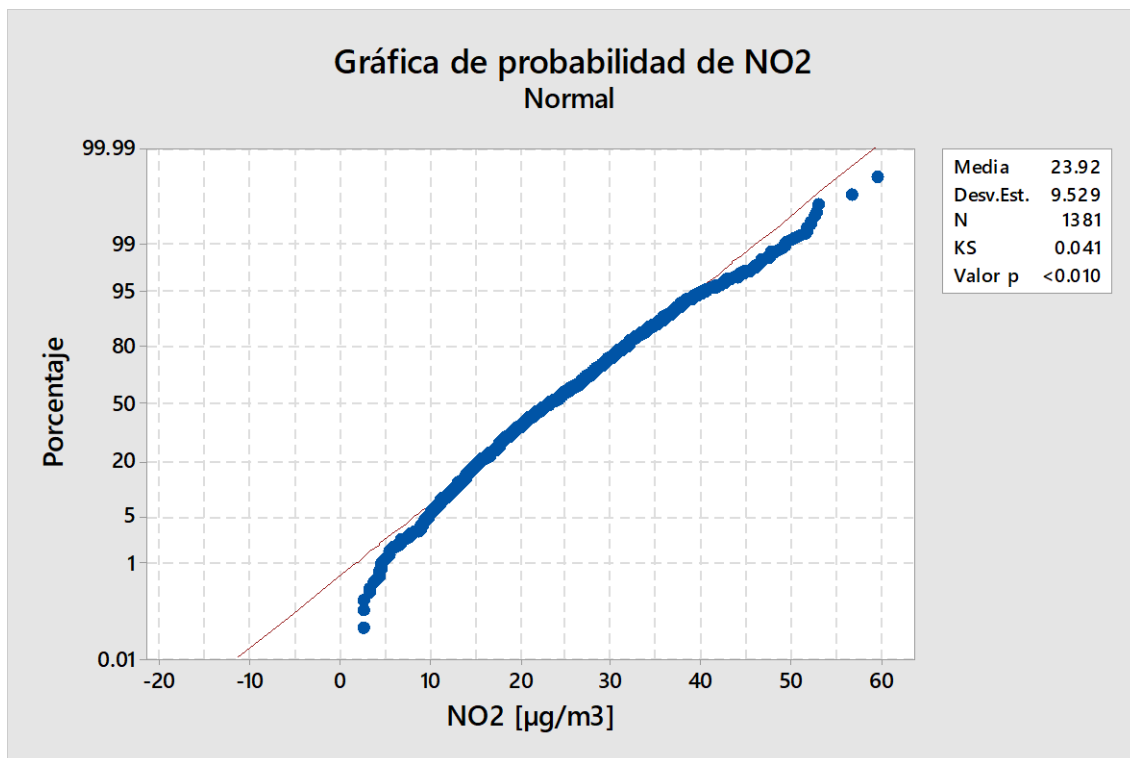


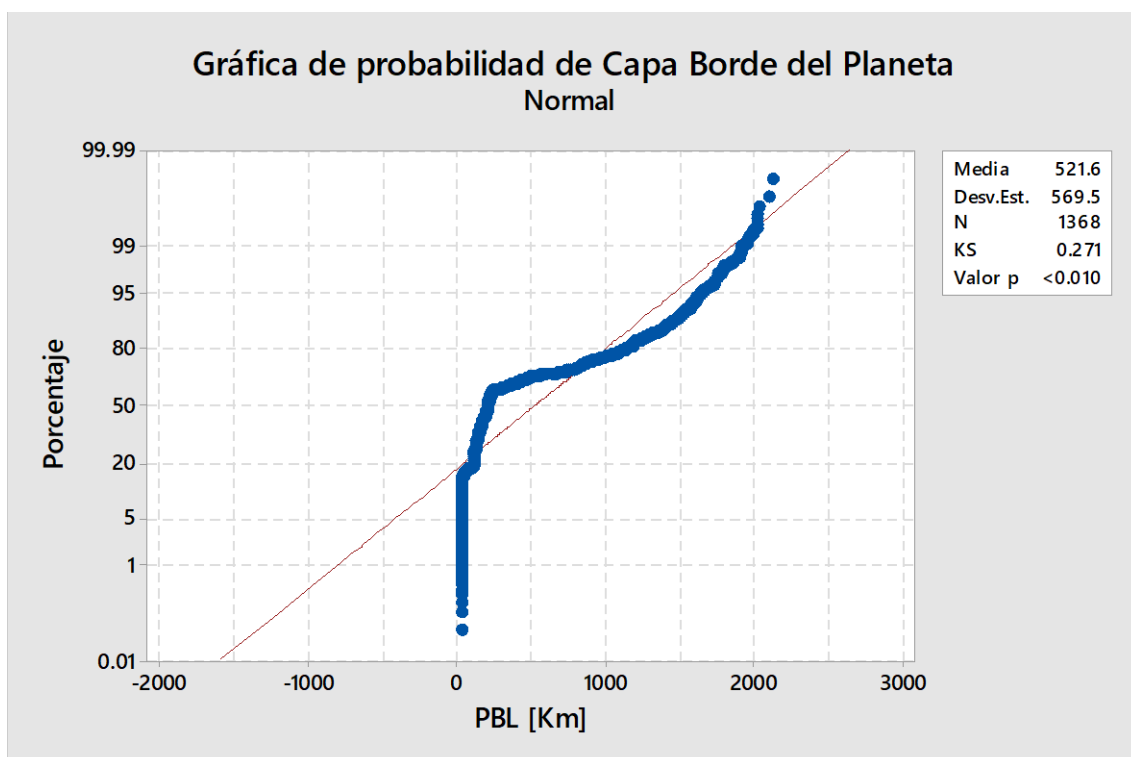
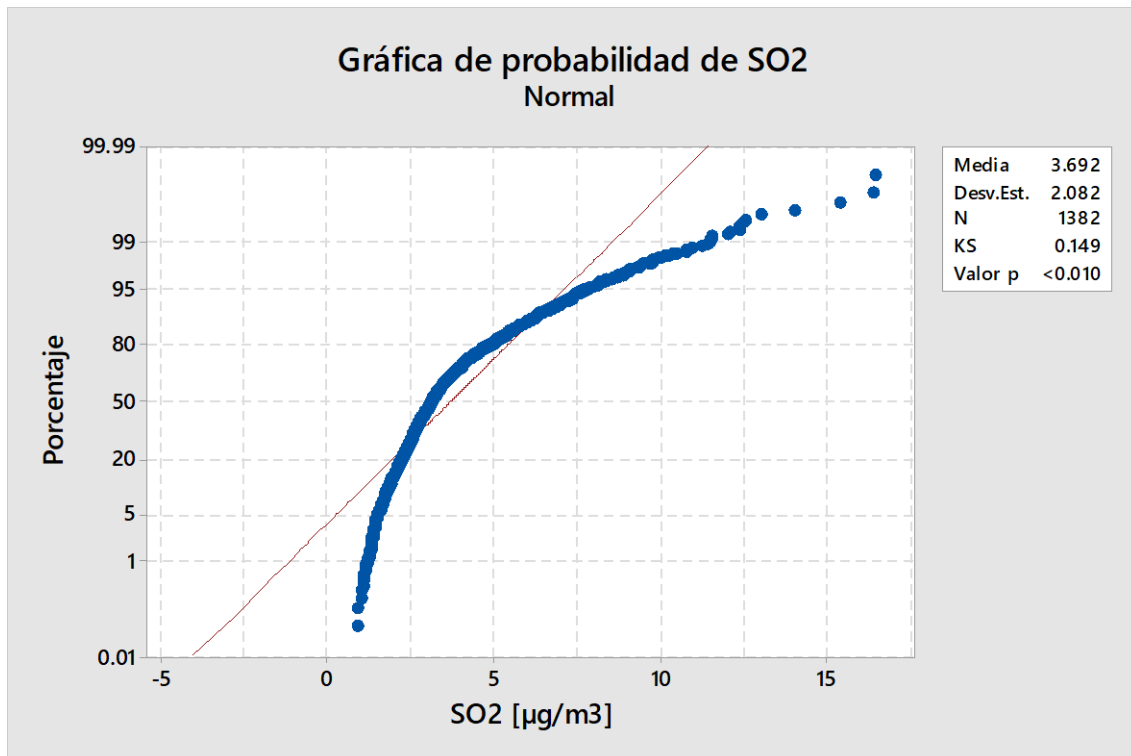
Gráfica de probabilidad de Rapidez del Viento  
Normal



Gráfica de probabilidad de CO  
Normal









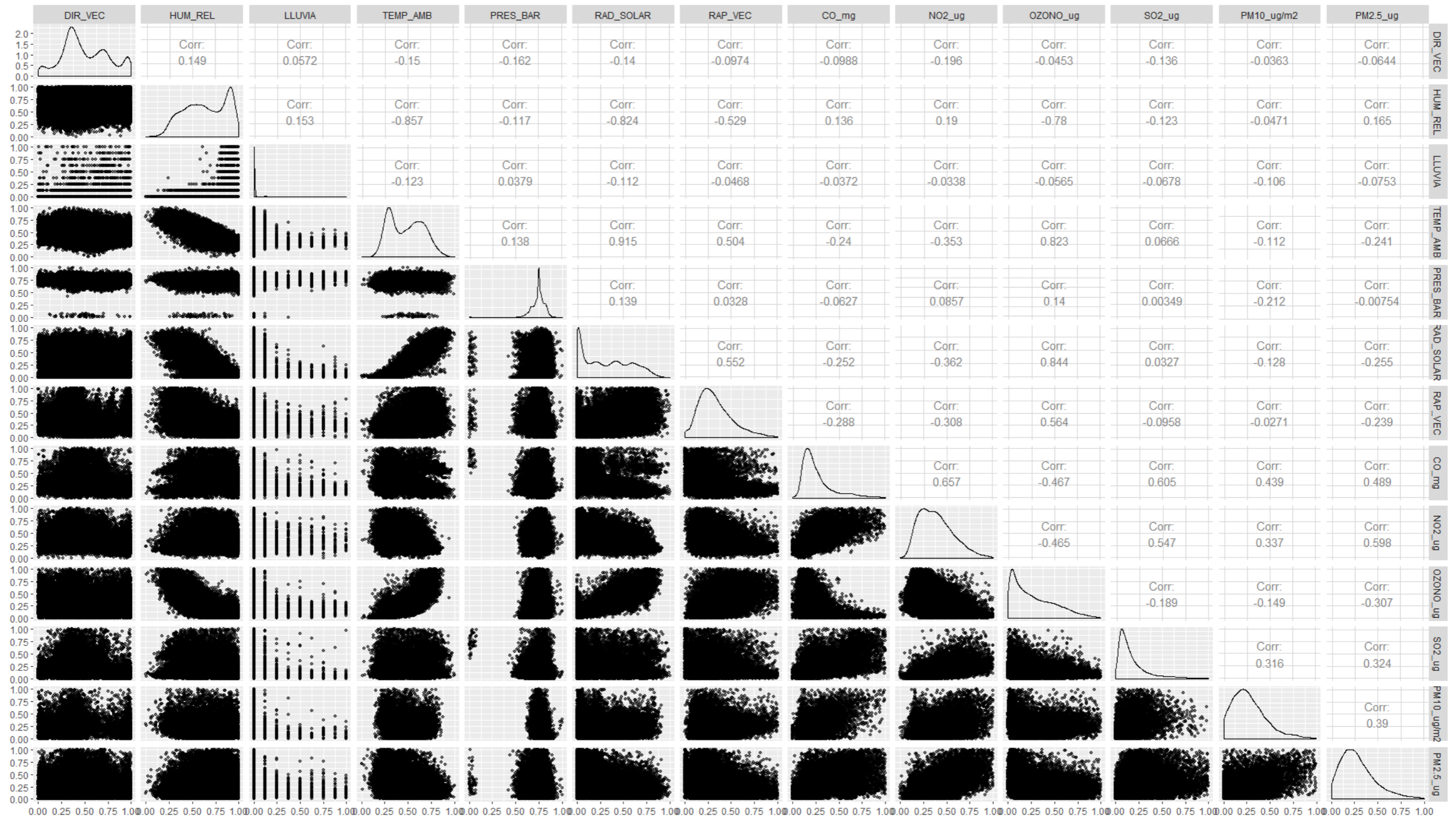




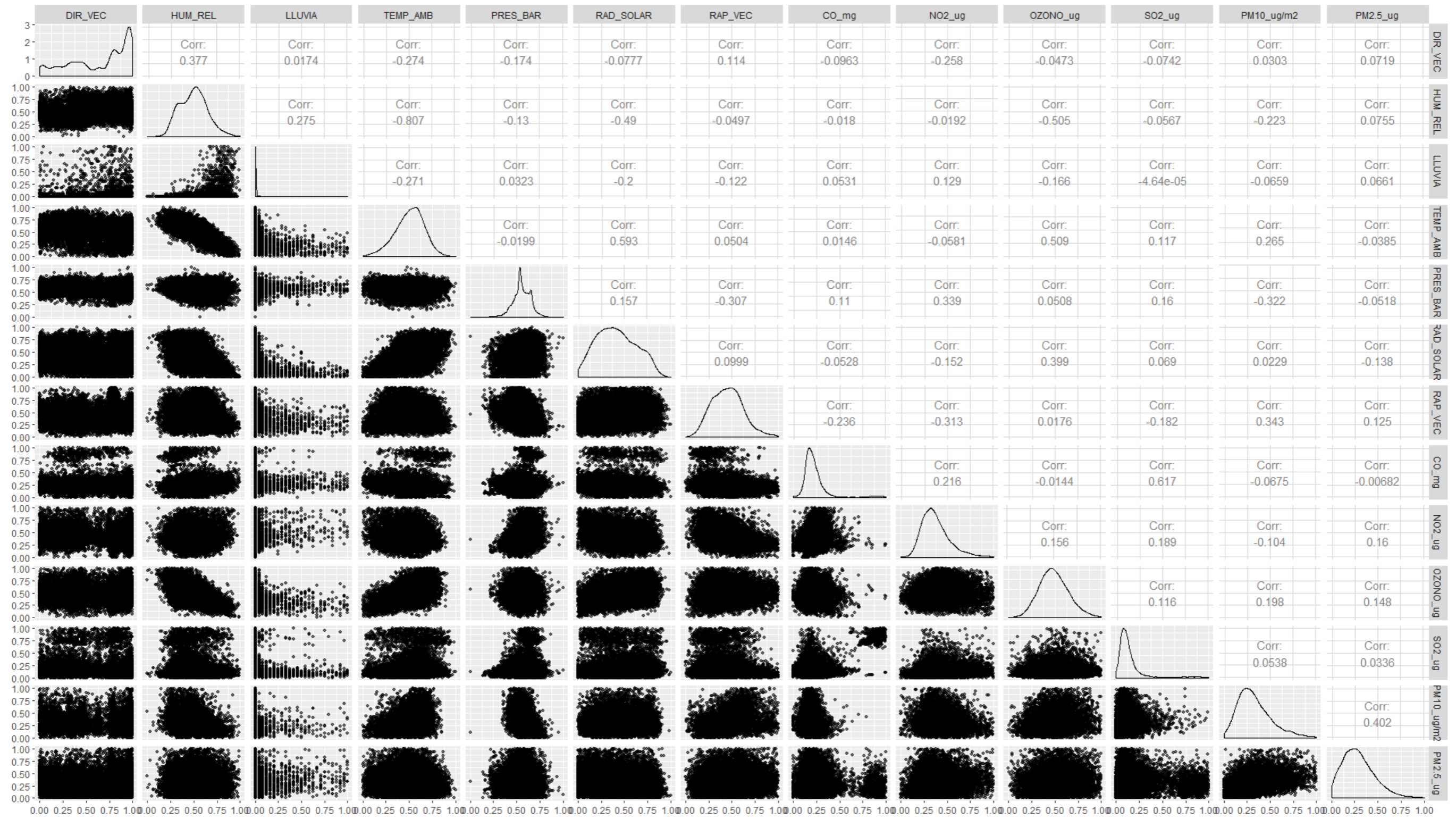




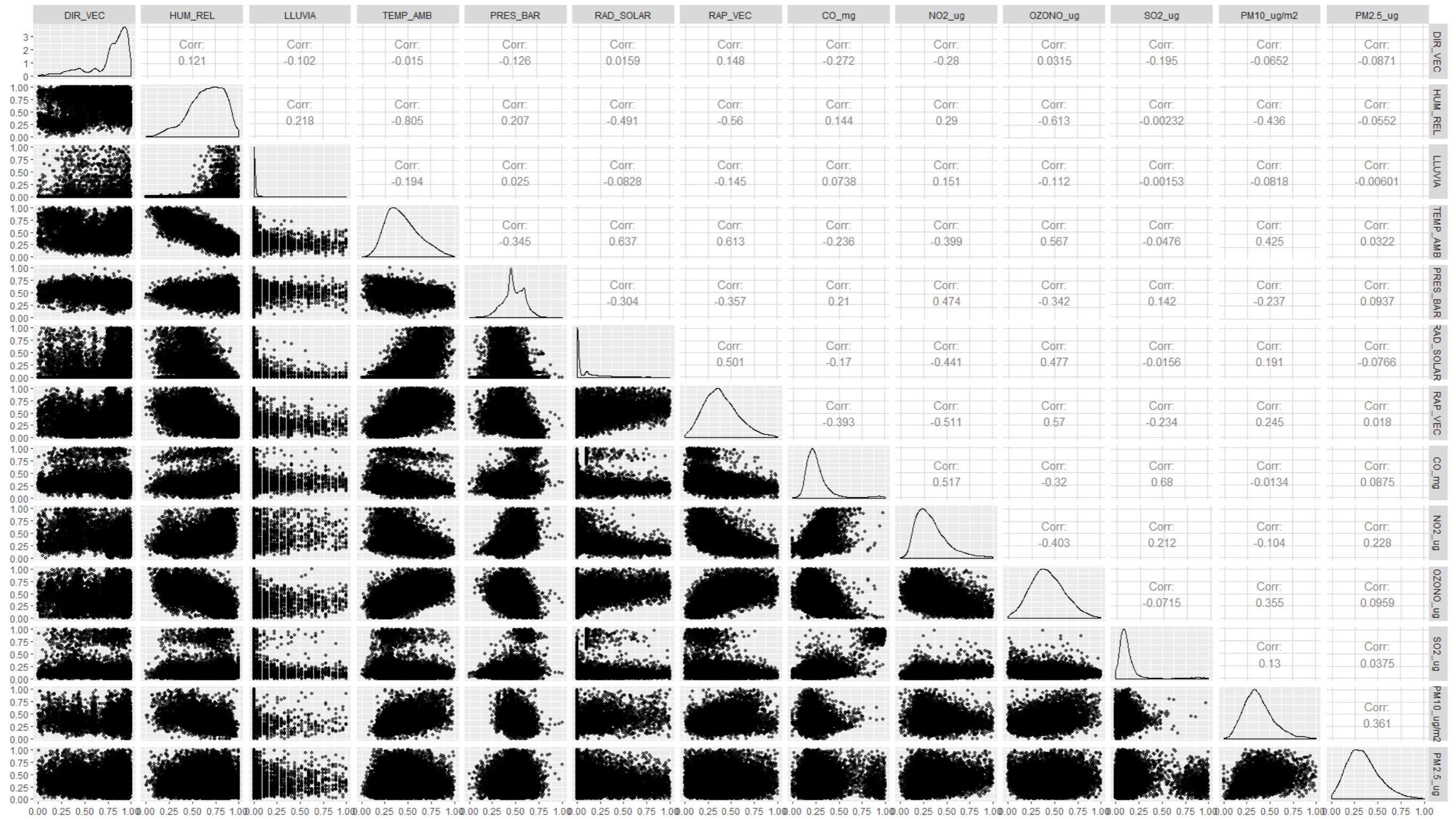
Estación Carapungo 06:00 a 11:00



Estación Carapungo 12:00 a 16:00



Estación Carapungo 17:00 a 20:00









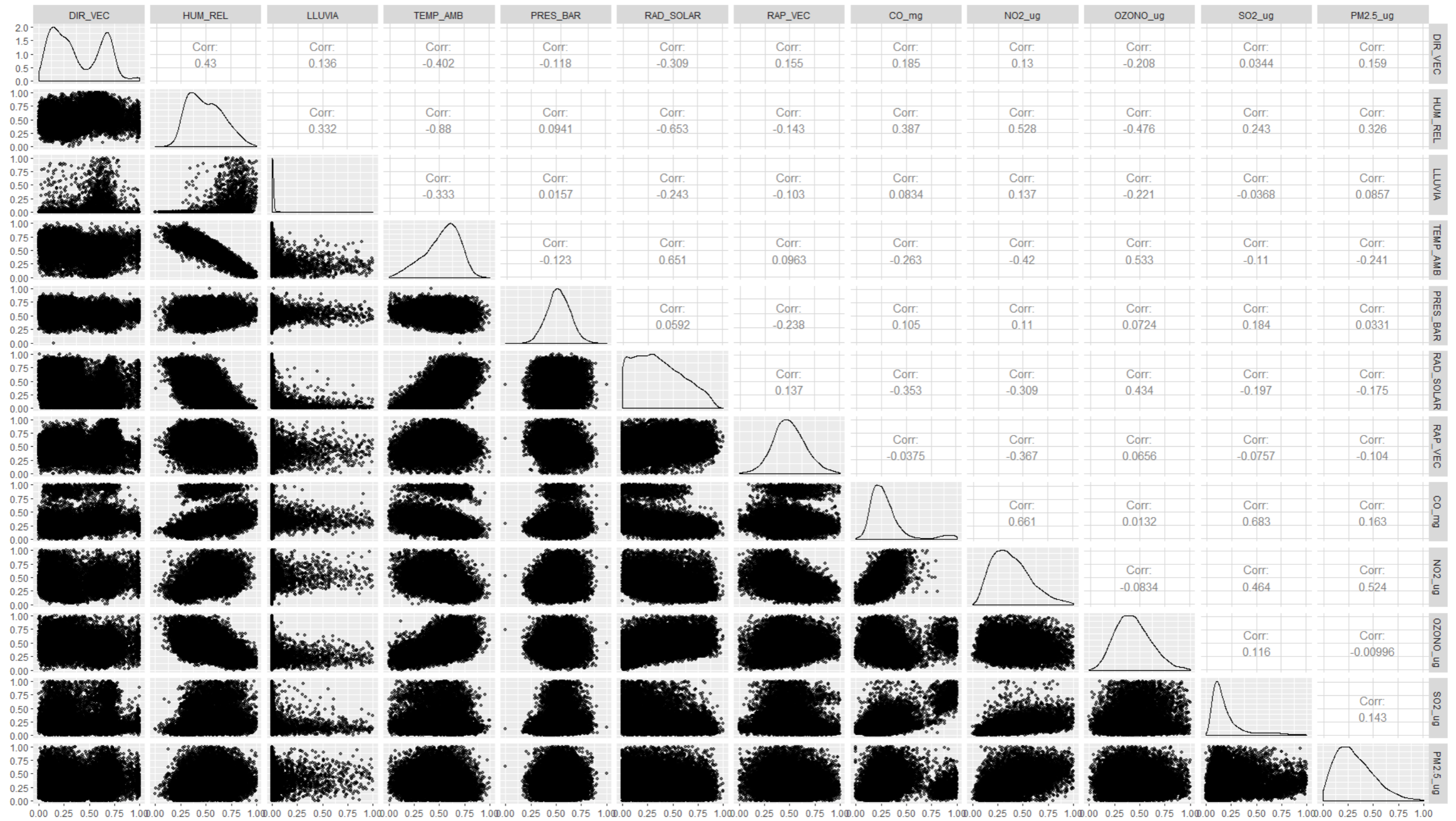








Estación El Camal 12:00 a 16:00







## Anexo 3: Resultados Regresión Lineal Múltiple sin Normalización

Estación	06:00 a 11:59	12:00 a 16:59	17:00 a 20:59	21:00 a 05:59
Belisario	$PM2.5_{ug} =$ $0.12715 * HUM\_REL +$ $1.2624 * TEMP\_AMB +$ $0.61007 * NO2_{ug} +$ $-0.16 * OZONO_{ug} +$ $-21.87235$ <b>Correlación = 0.6419</b>	$PM2.5_{ug} =$ $0.2203 * HUM\_REL +$ $0.70298 * TEMP\_AMB +$ $-0.00695 * RAD\_SOLAR +$ $0.338 * NO2_{ug} +$ $0.08936 * OZONO_{ug} +$ $-16.59284$ <b>Correlación = 0.5124</b>	$PM2.5_{ug} =$ $0.21502 * HUM\_REL +$ $0.87077 * TEMP\_AMB +$ $2.21878 * RAP\_VEC +$ $1.99693 * CO\_mg +$ $0.24519 * NO2_{ug} +$ $0.16665 * SO2_{ug} +$ $-24.06781$ <b>Correlación: 0.4865</b>	$PM2.5_{ug} =$ $0.12697 * HUM\_REL +$ $1.26379 * TEMP\_AMB +$ $0.61001 * NO2_{ug} +$ $-0.16032 * OZONO_{ug} +$ $-21.86464$ <b>Correlación: 0.642</b>
Carapungo	$PM2.5_{ug} =$ $-0.32983 * TEMP\_AMB +$ $7.6307 * CO\_mg +$ $0.48477 * NO2_{ug} +$ $0.0793 * PM10_{ug/m2} +$ $5.58951$ <b>Correlación= 0.5898</b>	$PM2.5_{ug} =$ $0.11853 * HUM\_REL +$ $-0.39321 * CO\_mg +$ $0.36317 * NO2_{ug} +$ $0.11272 * OZONO_{ug} +$ $0.11706 * PM10_{ug/m2} +$ $-6.42189$ <b>Correlación= 0.5898</b>	$PM2.5_{ug} =$ $1.41293 * CO\_mg +$ $0.33492 * NO2_{ug} +$ $0.11678 * OZONO_{ug} +$ $0.12909 * PM10_{ug/m2} +$ $2.1519$ <b>Correlación=</b> <b>0.3204</b>	$PM2.5_{ug} =$ $-1.15718 * TEMP\_AMB +$ $4.123 * CO\_mg +$ $0.35445 * NO2_{ug} +$ $0.16875 * PM10_{ug/m2} +$ $15.79593$ <b>Correlación=0.5201</b>
Cotocollao	$PM2.5_{ug} =$ $0.15895 * HUM\_REL +$ $0.95025 * TEMP\_AMB +$ $0.01037 * RAD\_SOLAR +$ $5.58482 * CO\_mg +$ $0.4995 * NO2_{ug} +$ $-0.10492 * OZONO_{ug} +$ $-21.49496$ <b>Correlación: 0.6318</b>	$PM2.5_{ug} =$ $0.13422 * HUM\_REL +$ $0.44374 * TEMP\_AMB +$ $-0.50192 * RAP\_VEC +$ $0.29037 * NO2_{ug} +$ $0.10468 * OZONO_{ug} +$ $-7.89302$ <b>Correlación: 0.329</b>	$PM2.5_{ug} =$ $0.13499 * HUM\_REL +$ $0.40672 * TEMP\_AMB +$ $0.31267 * NO2_{ug} +$ $0.10184 * OZONO_{ug} +$ $0.09779 * SO2_{ug} +$ $-9.22987$ <b>Correlación: 0.3616</b>	$PM2.5_{ug} =$ $0.12496 * HUM\_REL +$ $-1.26253 * RAP\_VEC +$ $2.25355 * CO\_mg +$ $0.32411 * NO2_{ug} +$ $0.28852 * SO2_{ug} +$ $-3.8306$ <b>Correlación: 0.493</b>
El Camal	$PM2.5_{ug} =$ $0.01324 * RAD\_SOLAR +$ $6.77515 * CO\_mg +$ $0.41791 * NO2_{ug} +$ $-0.26937 * OZONO_{ug} +$	$PM2.5_{ug} =$ $0.16291 * HUM\_REL +$ $-0.91184 * CO\_mg +$ $0.42514 * NO2_{ug} +$ $0.08259 * OZONO_{ug} +$	$PM2.5_{ug} =$ $0.10312 * HUM\_REL +$ $0.4508 * TEMP\_AMB +$ $2.30626 * CO\_mg +$ $0.26857 * NO2_{ug} +$	$PM2.5_{ug} =$ $0.10327 * HUM\_REL +$ $3.79753 * CO\_mg +$ $0.26412 * NO2_{ug} +$ $-0.20821 * OZONO_{ug} +$

	$0.32129 * SO2_{ug} +$ $4.62276$ <b>Correlación: 0.6896</b>	$-5.4585$ <b>Correlación: 0.4767</b>	$0.256 * SO2_{ug} +$ $-7.26638$ <b>Correlación: 0.4671</b>	$0.26682 * SO2_{ug} +$ $0.91033$ <b>Correlación: 0.5776</b>
--	---	---	--	---



