



Análisis sobre la contaminación en la ciudad de Madrid

[Memoria Trabajo Fin de Grado]

Autor: Álvaro Gallego

Tutor responsable: Juan A. Fdez del Pozo

Agradecimientos

Con especial dedicación a mi familia, que siempre estuvo ahí en los momentos malos, y en memoria de mi difunto padre.

Resumen del trabajo realizado:

El objeto de este trabajo es la creación de una aplicación, la cual a través de manipular datos referentes a la contaminación atmosférica ofrecidos por los diferentes portales de datos abiertos, obtenga un análisis descriptivo de la situación actual y unas predicciones fiables a corto plazo para valorar si las medidas que se están llevando a cabo son eficaces y suficientes como para garantizar la salud y el bienestar de los ciudadanos de la ciudad de Madrid. Todo ello siguiendo la metodología CRISP-DM utilizada en proyectos de data science.

Por ello se ha dedicado gran parte del tiempo en el estudio del arte y la comprensión de los datos, el preprocessado de los datos, su análisis, el análisis dependencias y la tendencia, además de varios ajustes realizados en el entorno de trabajo para las diversas etapas en las que consistirá el proyecto. Para ello, lo primero fue documentarme sobre las principales variables de la contaminación atmosférica, las técnicas de análisis de series temporales y las distintas librerías que ofrece R para su manejo. A continuación, explorar toda la información disponible en los sitios web del Ayuntamiento de Madrid y de la AEMET (Agencia Estatal de Meteorología), obteniendo la información al respecto de la calidad del aire en Madrid y del tiempo.

Mediante el uso de tablas y gráficos ggplot, se muestran distintos agregados de datos según los años, los meses, las semanas y las horas, usando la media y la mediana como estimadores. También se muestra la tendencia, la cual luego se mostrará en estos gráficos mediante el método de cuadrados mínimos.

Se ha realizado el análisis de variables dependientes en distintos instantes del tiempo, la obtención de un modelo Vector Auto Regresivo para la obtención de predicciones durante el 2017, y el despliegue de la misma aplicación para obtener gráficos, estadísticos y predicciones de manera dinámica.

Finalmente se han tomado en consideración las medidas propuestas por el Plan de Acción A propuesto del Ayto. de Madrid para garantizar los objetivos medio y largo plazo.

Además se ha hecho uso de tecnologías como Docker, Anaconda y Github para facilitar el desarrollo y seguimiento de este proyecto.

Summary

The goal of this project is to create an application, which, through manipulating data related to air pollution offered by the different open data websites, produces a descriptive analysis of the current situation and reliable predictions in the short term to evaluate if the measures that are being executed are effective and sufficient to guarantee the health and wellness of the citizens of Madrid. All this following the CRISP-DM methodology used in data science projects.

Most of the time in this project has been spent on art study of environmental matters and data comprehension, data preprocessing, data analysis, as well as several adjustments on the work environment as long as the project progressed.

Needless to say, the first thing I did was to dig into the most important pollutants in our days and all the protocols and decrees that currently trying to reduce them. Later on, I explored some of the different analysis techniques widely used in time series and the different libraries that R provides. Then I started to look for more open data that could be useful for this study mainly weather and traffic, since I already knew that Madrid's city council had published this kind of information to public access.

By the use of tables and ggplot graphics, different aggregated are shown yearly, monthly, weekly, daily and hourly, using average and median as estimators. Trend lines have been also plotted on them using least squares.

As well as an analysis of dependent variables at different points of time, obtaining a Vector Auto Regressive model to obtain predictions during 2017, and the deployment of the same application to obtain graphs, statistics and predictions dynamically.

Finally, consideration has been given to the measures proposed by the government of Madrid in his Plan de Acción A to ensure medium and long-term objectives.

Last but not least, different dependent analysis had been done in different time instants.

All this work has been done by using technologies such as Docker to virtualize the environment, Anaconda for notebooks and as IDE, and Github to develop and share the whole project.

Explicación y justificación de las modificaciones al Plan de Trabajo.

No se ha modificado la lista de objetivos, ni la de tareas. Aunque se han pospuesto debido a falta de tiempo alguno de los plazos estimados en el anterior diagrama de Gantt.

La lista de objetivos y la lista de tareas cumplen con la metodología CRISP-DM que se ha escogido para la realización de este proyecto, como se explica brevemente en la introducción de la memoria.

Revisión de la lista de objetivos del trabajo.

- Investigar el estado del arte respecto al problema de la contaminación en Madrid.
- Análisis de los datos y selección de características.
- Modelado.
- Evaluación y obtención de resultados.
- Despliegue (puesta en producción)
- Proponer recomendaciones.

Revisión de la lista de tareas.

- Introducción
- Descripción de los datos (Fase I)
- Análisis de datos: comprensión de las variables de forma aislada (Fases II y III)
- Análisis de datos: análisis de dependencias (Fase III)
- Modelado y evaluación (Fases IV y V)
- Puesta en producción (Fase VI)
- Sistema de recomendaciones basándonos en fuentes externas.

Contenido

Índice

1.	<u>Introducción.....</u>	8
2.	<u>Entorno de trabajo.....</u>	9
3.	<u>Algoritmos empleados.....</u>	10
1.	<u>Algoritmos de selección de características.....</u>	10
2.	<u>Algoritmos de predicción.....</u>	11
4.	<u>Análisis descriptivo de variables.....</u>	11
1.	<u>Estaciones de medida.....</u>	12
2.	<u>Variables Monitorizadas.....</u>	14
3.	<u>Valores límite.....</u>	15
i.	<u>Umbrales de información y alerta para contaminantes.....</u>	16
5.	<u>Análisis de las series de contaminación registrados.....</u>	17
1.	<u>Variable Dióxido de Azufre (SO₂).....</u>	18
2.	<u>Variable Monóxido de Carbono (CO).....</u>	20
3.	<u>Variable Monóxido de Nitrógeno (NO).....</u>	23
4.	<u>Variable Dióxido de Nitrógeno (NO₂).....</u>	25
5.	<u>Variable Partículas en Suspensión de diámetro inferior a 2,5 mg (PM_{2,5}).....</u>	28
6.	<u>Variable Partículas en Suspensión de diámetro inferior a 10 mg (PM₁₀).....</u>	31
7.	<u>Variable Ozono (O₃).....</u>	33
8.	<u>Variable Tolueno (C₆H₅-CH₃).....</u>	37
9.	<u>Variable Benceno (C₆H₆).....</u>	40
10.	<u>Variable Etilbenceno (C₈H₁₀).....</u>	42
11.	<u>Variable Hexeno (C₆H₁₂).....</u>	45
12.	<u>Variable Hexano (C₆H₁₄).....</u>	48
13.	<u>Variables atmosféricas.....</u>	51
6.	<u>Análisis de dependencias.....</u>	52
1.	<u>Dependencias de variables en el mismo instante de tiempo.....</u>	52
2.	<u>Dependencias de variables en instantes de tiempo anteriores.....</u>	53
3.	<u>Dependencias de variables en instantes de tiempo posteriores</u>	55
7.	<u>Predicciones de variables.....</u>	56
1.	<u>Ánálsis PCA.....</u>	56
2.	<u>Modelado VAR.....</u>	59
8.	<u>Despliegue de la aplicación.....</u>	67
9.	<u>Sistema de Recomendaciones.....</u>	71
10.	<u>Conclusiones finales y líneas futuras.....</u>	80
11.	<u>Referencias y bibliografía.....</u>	81
12.	<u>Índice de tablas.....</u>	84
13.	<u>Índice de figuras.....</u>	85

1. Introducción

La contaminación medioambiental es uno de los principales problemas del siglo XXI. Con el paso de los años, hay más entidades a nivel local y global, que se comprometen y pocas veces logran cumplir, el respetar unos límites establecidos por expertos climatológicos.

Sin ir más lejos, en la ciudad de Madrid cada vez es más frecuente que se tomen medidas restrictivas respecto a la circulación de los vehículos debido a los altos niveles de contaminación. Por ello, el propósito de este trabajo, es detectar causas, poder predecir niveles de alerta y hacer recomendaciones sobre las concentraciones de estos gases.

Madrid es una ciudad que cuenta con una amplia red de sensores que registran información en tiempo real o histórico, de uso libre, sobre varios tipos de factores climatológicos, atmosféricos, como de la actividad diaria de los ciudadanos. El estudio de este proyecto se centrará en la creación de un modelo capaz de analizar y hacer predicciones no sólo en los puntos de medida como puedan ser los sensores, sino también en ubicaciones inferidas, así como la propuesta de recomendaciones en base a una situación de alerta.

Actualmente existen dos metodologías principales para hallar de manera normalizada, el conocimiento en los datos (Knowledge Discovery in Databases o KDD), en determinados procesos de minería de datos: CRISP-DM y SEMMA. Ambas especifican las tareas a realizar en cada fase descrita por KDD e indican lo que se precisa obtener en cada fase. Sin embargo, CRISP-DM es más completo, ya que tiene en cuenta la aplicación al entorno de negocio de los resultados.

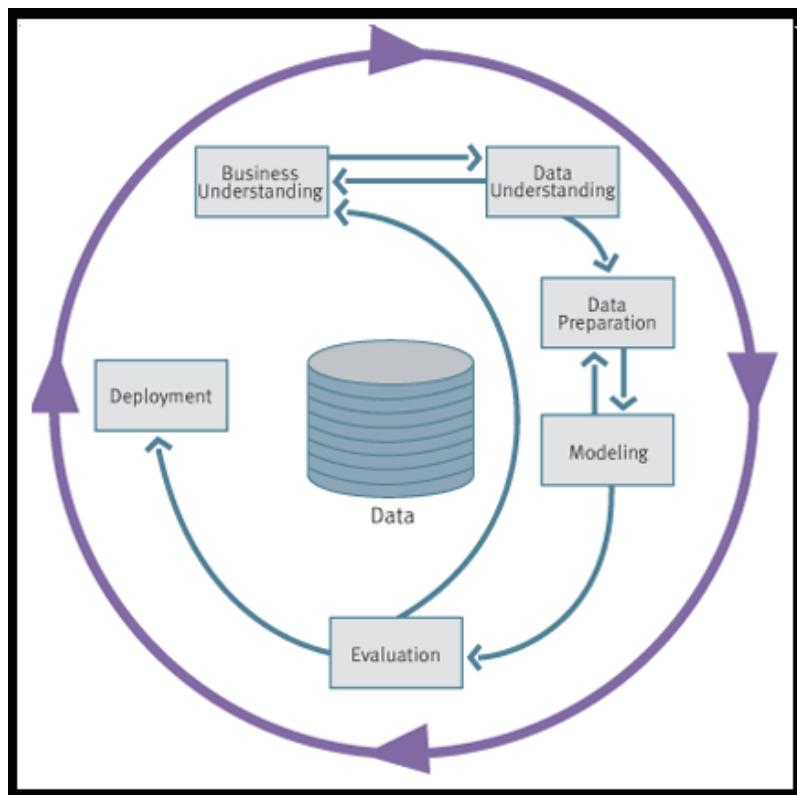


Figura 1. Metodología CRISP-DM

Se hará uso de CRISP-DM (Cross Industry Standard Process for Data Mining) para el desarrollo de este proyecto, que comprende las fases referentes a la Comprensión del negocio (Fase I), la Compresión de los datos (Fase II), la Preparación de los datos (Fase III), el Modelado (Fase IV), la Evaluación (Fase V) y el Despliegue (Fase VI), las cuales se implementan a lo largo de este proyecto¹. La fase del Despliegue consistirá en una aplicación interactiva la cual se encarga de recibir un fichero .csv o Excel, el cual contendrá información sobre los gases de manera mensual, y mostrar predicciones mediante distintos algoritmos. Los algoritmos de predicción que se han implementado para la aplicación han sido técnicas de suavizamiento exponencial, de la cual nos centraremos en suavizamientos exponenciales, redes neuronales y ARIMA⁶³. Además del modelo multivariante Vector Auto Regressive^{19,20} que es ampliamente utilizado en el mundo de la economía y permite obtener predicciones de todas las variables endógenas o internas al modelo.

Un suavizamiento exponencial simple y doble, únicamente suavizan los datos de la serie temporal. Un suavizamiento exponencial triple, conocido como Holt-Winters, suaviza datos y tendencia, y también puede suavizar la estacionalidad, si se indica para un modelo aditivo o multiplicativo.

- El Modelo multiplicativo cuando la magnitud del patrón estacional en los datos depende de la magnitud de los datos. En otras palabras, la magnitud del patrón estacional aumenta a medida que los valores de los datos se incrementan y disminuye a medida que los valores de los datos decrecen. Multiplicative: $x_t = \text{Trend} * \text{Seasonal} * \text{Random}$
- El modelo aditivo cuando la magnitud del patrón estacional en los datos no depende de la magnitud de los datos. En otras palabras, la magnitud del patrón de estación no cambia cuando la serie sube o baja. Additive: $x_t = \text{Trend} + \text{Seasonal} + \text{Random}$

Un modelo ARIMA (es una generalización para modelos sin estacionalidad, Additivo y Multiplicativo). El modelo ARIMA puede generalizarse aún más para considerar el efecto de la estacionalidad. En ese caso, se habla de un modelo SARIMA (seasonal autoregressive integrated moving average). La estimación de un modelo ARIMA solo funciona cuando todas las variables predictoras son deterministas y los errores son estacionarios. Por tanto, se requiere de la diferencia estocástica de las variables hasta que todas parezcan estacionarias, y entonces ajustar el modelo con errores ARIMA. La función en R auto.arima(), se encarga de seleccionar la selección del orden y la diferenciación, pero solo comprueba que los errores son estacionarios.

Las redes neuronales artificiales permiten el modelado de relaciones complejas no lineares sobre variables de entrada y salida. En el caso de un modelo de redes neuronales autoregresivas (NNAR), los valores de momentos temporales de series temporales son utilizados como entrada y salida de los valores de las series temporales. Una de las diferencias principales entre NNAR y el método HW es que las NNAR no imponen restricción sobre sus parámetros para garantizar la estacionalidad.

El uso del modelado VAR se explica en el apartado de Algoritmos de predicción.

2. Entorno de trabajo

Haciendo uso del portal de datos abiertos del Ayuntamiento de Madrid² y de la AEMET³, se han utilizado las herramientas disponibles en R^{4,5,6} para el análisis de los datos.

R es una implementación de software libre del lenguaje S pero con soporte de alcance estático⁴. R es un entorno y lenguaje de programación con un enfoque al análisis estadístico. Proporciona un amplio abanico de herramientas estadísticas (modelos lineales y no lineales, tests estadísticos, análisis de series temporales, algoritmos de clasificación y agrupamiento, etc.) y gráficas.

R es parte del sistema GNU y se distribuye bajo la licencia GNU GPL. Está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux.

Como entorno de trabajo se utiliza Anaconda⁷, que es una plataforma abierta de ciencia de datos impulsada por Python. La versión de código abierto de Anaconda es una distribución de alto rendimiento de Python y R e incluye más de 100 de los paquetes más populares de Python, R y Scala para la ciencia de los datos.

Para desplegar la aplicación se hace uso de la comunidad de shinyapps.io⁸, que es una plataforma como servicio (PaaS) donde alojaremos nuestra aplicación web. La cual utiliza el framework Shiny⁹ para aplicaciones web en R.

Todo el material de este proyecto se publicado en el sistema de control de versiones de github¹⁰, de manera pública, pudiéndose descargar o consultar¹¹. Además se incluye el anexo "Contaminacion_Madrid.pdf" con todo el proceso del proyecto.

Finalmente, se ha creado un container mediante el uso de docker¹², para la virtualización del entorno de trabajo desde el propio navegador¹³, a partir de una distribución Debian.

3. Algoritmos empleados

Este proyecto se ha basado en la creación de un modelado Vectorial Auto Regresivo, comúnmente conocido como modelo VAR, para la predicción de los principales gases contaminantes en el 2017. No obstante, además de un pre análisis de los datos de las variables, también se han utilizado algoritmos de selección de características que nos permiten observar la relevancia de una variable sobre el resto, así como posibles agrupaciones de estas variables.

3.1 Algoritmos de selección de características

Se ha utilizado el algoritmo PCA⁵⁰, para la reducción de dimensiones.

El análisis de reducción de dimensiones (PCA) es utilizado para la reducción de dimensiones, selección de características, y la representación de aprendizaje. Esto es útil cuando se tienen demasiadas variables redundantes. El objetivo es encontrar el valor intrínseco de la dimensionalidad de los datos.

El algoritmo k-medias⁴ (k-means, en inglés) por su parte, es un algoritmo de clustering que devuelve el agrupamiento natural de puntos de datos, en base a distancias euclídeas tomando como referencia centroides o puntos de referencia tomados en subconjuntos de datos. Es un caso especial de Gaussian mixture models. Sin embargo, nuestro objetivo es aportar el mayor número de subgrupos a nuestro modelo VAR, por lo que no se ha considerado de utilidad el uso de clustering.

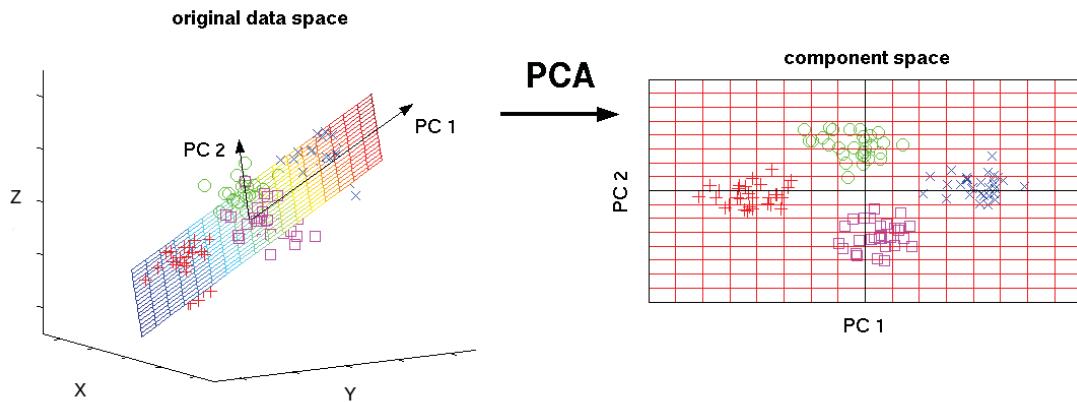


Figura 2. Ejemplo representativo de la aplicación PCA y k-means

A modo de ejemplo se puede observar cómo trabajan estos dos algoritmos, en la figura anterior.

3.2 Algoritmos de predicción

En cuanto a los algoritmos empleados para la obtención de predicciones se ha considerado en primer lugar la importancia en el tiempo de los datos, ya que estamos trabajando con series temporales. Esto implica considerar cada variable en el tiempo, en base a retardos en el tiempo que normalmente pueden ser significativos en las 24 horas previas a un momento dado, y la obtención de sus coeficientes que mejor se ajusten a la curva descrita por dicha variable.

Todo ello supone un modelo que retroalimente la predicción de estas variables, por lo cual el uso de técnicas como ARIMA, redes neuronales o suavizamiento exponencial sólo son útiles si disponemos previamente los valores del resto de variables que mejor definen dicha curva en el periodo a predecir.

Por lo que se ha optado por un modelo Vectorial Auto Regresivo (VAR), ampliamente utilizado en el mundo de la econometría. Para el cual habrá que comprobar previamente que estamos trabajando con datos estacionarios, donde se repiten patrones. En caso contrario, tendremos que realizar una diferenciación en las variables que tengan una raíz unitaria o un camino aleatorio, tantas veces como sean necesarias hasta la obtención de la estacionalidad. Esto se puede comprobar fácilmente realizando tests como pueden ser el test aumentado de Dicker-Fuller o el test KPSS.

El modelo VAR además requiere que se suministren variables linealmente independientes, donde una variable o más variables no impliquen el suceso de otra. Esto se puede comprobar mediante el test de Granger que analiza la causalidad entre variables. También es importante considerar que cuantas más variables linearmente independientes se aporten a dicho modelo, se obtendrá un mayor número de coeficientes, obteniendo así una mayor precisión en las predicciones.

4. Análisis descriptivo de variables

A partir de la información ofrecida por la AEMET y el Ayto. de Madrid:

4.1 Estaciones de medida

En este estudio se ha trabajado con dos tipos distintos de estaciones: las estaciones que nos ofrecen datos atmosféricos en la ciudad de Madrid proporcionados por la AEMET y las estaciones de medición de gases atmosféricos que nos ofrece la web del Ayuntamiento de Madrid.

La AEMET nos proporciona cuatro estaciones en la ciudad de Madrid, localizadas en Ciudad Universitaria, Retiro, Cuatro Vientos y Barajas. Sin embargo, debido a la inconsistencia de información en dos de las cuatro estaciones de medición ofrecidas por la AEMET, únicamente se ha trabajado con las estaciones de Cuatro Vientos y Barajas, ya que algunas de las variables de medición a considerar no estaban presentes durante los años 2011 y 2016 en algunas de las otras dos estaciones descartadas.

	Ciudad Universitaria	Cuatro Vientos
Latitud	402706N	402232N
Longitud	034327W	034710W
Altitud	664	690
Ind-Sinop	08220	08223

Tabla 1. Identificación de las estaciones meteorológicas

En cuanto a las estaciones de medición ofrecidas por la web de Ayuntamiento de Madrid, los códigos siguen el siguiente patrón, según el documento oficial "Interprete_ficheros_calidad_del_aire":

28 (código de la Comunidad de Madrid) que se repetirán en todas las estaciones
079 (código del municipio de Madrid) que se repetirán en todas las estaciones
004 (código de la estación, en este caso la no 4) diferentes según la estación

Las estaciones que miden la calidad del aire durante los años 2011 hasta 2016 son:

id	Ubicación
28079004	Pza. De España
28079008	Escuelas Aguirre
28079011	Av. Ramón y Cajal
28079016	Arturo Soria
28079017	Villaverde Alto
28079018	C/ Farolillo
28079024	Casa de Campo
28079027	Barajas
28079035	Pza. Del Carmen
28079036	Moratalaz
28079038	Cuatro Caminos
28079039	Barrio del Pilar
28079040	Vallecas
28079047	Méndez Álvaro
28079048	Paseo de la Castellana
28079049	Retiro
28079050	Pza. Castilla
28079054	Ensanche Vallecas
28079055	Urb. Embajada (Barajas)

4.1 Estaciones de medida

28079056	Pza. Fdez. Ladreda
28079057	Sanchinarro
28079058	El Pardo
28079059	Parque Juan Carlos I
28079060	Tres Olivos

Tabla 2. Identificación de las estaciones de medición de gases

En la [web⁷](#) puede ver algunos puntos de las estaciones, así como la obtención de un [boletín diario⁸](#) de los niveles de O3, NO2 y PM10 en estas estaciones



ANEXO:

El índice horario utilizado para información a la población es:

CALIDAD DEL AIRE	HORARIO		
	PM10 ($\mu\text{g}/\text{m}^3$)	NO _x ($\mu\text{g}/\text{m}^3$)	O ₃ ($\mu\text{g}/\text{m}^3$)
Bueno	0 - 50	0 - 100	0 - 90
Admisible	51 - 90	101 - 200	90 - 180
Deficiente	91 - 150	201 - 300	180 - 240
Malo	> 150	> 301	> 240

Distribución por zonas de la ciudad de Madrid:

ZONA	ESTACIONES
1 (interior M-10)	7 de tráfico (Escuelas Aguirre, Castellana, Plaza de Castilla, Ramón y Cajal, Cuatro Caminos, Plaza de España y Barrio del Pilar) + 3 de fondo (Plaza del Carmen, Méndez Álvarez y Retiro)
2 (sureste)	1 de tráfico (Moratalaz) + 2 de fondo (Vallecas y Ensanche de Vallecas)
3 (noreste)	5 de fondo (Arturo Soria, Sanchinarro, Urbanización Embajada, Barajas Pueblo y Tres Olivos) + 1 suburbana (Juan Carlos I)
4 (noroeste)	2 suburbanas (El Pardo y Casa de Campo)
5 (suroeste)	1 de tráfico (Fernández Ladreda) + 2 de fondo (Farolillo y Villaverde)



Figura 3. Boletín diario. Portal de Calidad del Aire

Consulta de datos

Portal de Calidad del Aire / Consulta de datos / Mapa de la Red de Vigilancia

- Mapa de la Red de Vigilancia
- » Boletín diario
- » Informes
- » Representación gráfica
- » Predicción
- » Contaminación acústica
- » Datos históricos



Figura 4. Consulta de datos. Portal de Calidad del Aire

4.2 Variables monitorizadas

Los datos utilizados en este trabajo son los datos históricos ofrecidos por la web del Ayuntamiento de Madrid y por la AEMET.

Los siguientes apartados presentan una descripción de las variables en estudio.

4.2 Variables monitorizadas



ANEXO II
Magnitudes, unidades y técnicas de medida

Magnitud	Abreviatura o fórmula	Unidad medida	Técnica de medida
01 Dióxido de Azufre	SO ₂	µg/m ³	38 Fluorescencia ultravioleta
06 Monóxido de Carbono	CO	mg/m ³	48 Absorción infrarroja
07 Monóxido de Nitrógeno	NO	µg/m ³	08 Quimiluminiscencia
08 Dióxido de Nitrógeno	NO ₂	µg/m ³	08 Id.
09 Partículas < 2.5 µm	PM2.5	µg/m ³	47 Microbalanza
10 Partículas < 10 µm	PM10	µg/m ³	47 Id.
12 Óxidos de Nitrógeno	NOx	µg/m ³	08 Quimiluminiscencia
14 Ozono	O ₃	µg/m ³	06 Absorción ultravioleta
20 Tolueno	TOL	µg/m ³	59 Cromatografía de gases
30 Benceno	BEN	µg/m ³	59 Id.
35 Etilbenceno	EBE	µg/m ³	59 Id.
37 Metaxileno	MXY	µg/m ³	59 Id.
38 Paraxileno	PXY	µg/m ³	59 Id.
39 Orthoxileno	OXY	µg/m ³	59 Id.
42 Hidrocarburos totales (hexano)	TCH	mg/m ³	02 Ionización de llama
44 Hidrocarburos no metálicos (hexano)	NMHC	mg/m ³	02 Id.

Figura 5.1. Listado de las variables contaminantes monitorizadas

Debido a que no todas las estaciones miden los mismos gases, y que en algún momento han dejado de medir alguno de estos para pasar a medir otros o simplemente han dejado de medirlos. Se ha optado por analizar la información de manera aleatoria entre las estaciones, para todos los gases. Los gases son medidas de manera horaria, durante 2011 hasta 2016, y se analizarán de forma horaria, diaria, mensual y anual.

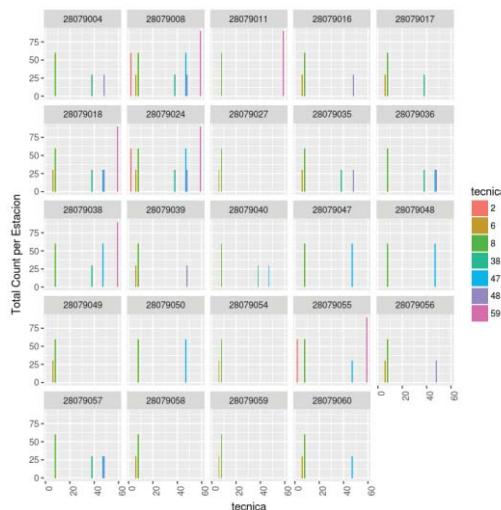


Figura 5.2. Representación gráfica de las mediciones en las estaciones no meteorológicas

4.3 Valores límite

Las variables atmosféricas son dadas de forma mensual y anual, de las cuales se han considerado:

id	Descripción
Direccion	Dirección del viento (decenas de grado) obtenida a través del primer elemento de w_racha
w_med	'Velocidad media mensual elaborada a partir de las observaciones de 7,13 y 18 UTC.(Km/h)'
tm_mes	Temperatura media mensual/anual (grados 15periodo)
hr	'Humedad relativa media mensual/anual (%)'
q_med	'Presión media mensual/anual al nivel de la estación (hPa)'
Inso	'Media mensual/anual de la insolación diaria (horas)'
p_mes	'Precipitación total mensual/anual (mm)'

Tabla 3. Listado de las variables atmosféricas consideradas

4.3 Valores límite

La directiva 2008/50/CE¹⁴ del Parlamento Europeo y del Consejo, de 21 de mayo de 2008, relativa a la calidad del aire ambiente y a una atmósfera más limpia en Europa, establece un conjunto de medidas orientadas principalmente a:

- reducir la contaminación a niveles que limiten al mínimo los efectos perjudiciales para la salud humana y el medio ambiente
- mejorar la información proporcionada a los ciudadanos sobre los riesgos a los que se exponen.

En el **Anexo XI – Valores límite para la protección de la salud humana**, de dicha directiva, se establecen los valores límite mostrados en la tabla siguiente.

Periodo medio	Valor límite
PM10	
1 día	50 µg/m ³ , que no podrá superarse más de 35 veces por año civil
Año civil	40 µg/m ³
Dióxido de azufre (SO₂)	
1 hora	350 µg/m ³ , que no podrá superarse más de 24 veces por año civil
1 día	125 µg/m ³ , que no podrá superarse más de 3 veces por año civil
Dióxido de Nitrógeno (NO₂)	
1 hora	200 µg/m ³ , que no podrá superarse más de 18 veces por año civil
Año civil	40 µg/m ³
Monóxido de carbono (CO)	
Máxima diaria cada 8 horas	10 mg/m ³
Benceno	
Año civil	5 µg/m ³

Tabla 4. Valores límites para la protección de la salud humana

En el **Anexo XIII – Niveles críticos para la protección de la vegetación**, se establecen los valores mostrados en la tabla siguiente.

Periodo medio	Valor límite
Dióxido de azufre	
Año civil e invierno (1 de octubre a 31 de marzo)	20 µg/m ³
Óxidos de nitrógeno	
Año civil	30 µg/m ³

Tabla 5.1 Valores límites para la protección de la vegetación

Por su parte el Ayto. de Madrid en su informe de la Calidad del Aire del 2016²⁶, aplica la vigente legislación del Real Decreto 102/2011, del 28 de enero²⁵, relativo a la mejora de la calidad del aire.

Por su parte, este decreto establece:

	Valor límite horario	Valor límite octohorario	Valor límite diario	Valor límite anual
Dióxido de Azufre	350µg/m ³ *no podrá superarse más de 24 veces por año		125µg/m ³ *que no podrá superarse en más de 3 ocasiones por año	
PM10			50µg/m ³ *no podrán superarse en más de 35 ocasiones al año	40µg/m ³
PM2.5				25 µg/m ³
Dióxido de Nitrógeno	200µg/m ³ *que no podrán superarse en más de 18 ocasiones al año			40µg/m ³
CO		10mg/m ³		
Benceno				5µg/m ³
Ozono				

Tabla 5.2 Valores límites para la protección de la salud humana R.D. 102/2011

4.3.1 Umbrales de información y alerta para contaminantes

Con el objeto de mejorar la información proporcionada a los ciudadanos, dicha directiva define dos tipos de umbrales, ya sea registrados o previstos, por encima de los cuales la población debería ser informada:

- Umbral de alerta: “nivel a partir del cual una exposición de breve duración supone un riesgo para la salud humana que afecta al conjunto de la población y que requiere la adopción de medidas inmediatas por parte de los Estados miembros”.
- Umbral de información: “nivel a partir del cual una exposición de breve duración supone un riesgo para la salud de los sectores especialmente vulnerables de la población y que requiere el suministro de información inmediata y apropiada”.

En su **Anexo XII – Umbrales de información y de alerta**, la directiva establece los umbrales de la tabla siguiente.

Valor límite	
Dióxido de azufre	
Umbral de alerta	500 µg/m ³ durante 3 horas consecutivas
Dióxido de nitrógeno	
Umbral de alerta	400 µg/m ³ durante 3 horas consecutivas
Ozono	
Umbral de información	180 µg/m ³ (promedio por hora)
Umbral de alerta	240 µg/m ³ (promedio por hora)

Tabla 6.1 Umbrales de información y de alerta

Por parte del R.D. 102/2011, el único umbral de información indicado hace referencia al nivel de Ozono, establecido como valor promedio por hora en 180µg/m³.

Umbral de alerta	
Dióxido de Azufre	500µg/m ³ *durante 3 horas consecutivas en un área > 100km ²
PM10	
PM2.5	
Dióxido de Nitrógeno	400µg/m ³ *durante 3 horas consecutivas en un área > 100km ²
CO	
Benceno	
Ozono	240µg/m ³ (valor promedio por hora)

Tabla 6.2 Umbrales de alerta R.D. 102/2011

Como se puede apreciar ambas legislaciones coinciden en los valores límites y umbrales.

5. Análisis de las series de contaminación registrados

En este apartado se presentan los principales estadísticos descriptivos de las variables de contaminación y meteorológicas, para lo cual se han analizado los estadísticos de la media y la mediana con agregados horarios, diarios, mensuales, y anuales.

Las series de datos se corresponden con los datos registrados en las estaciones de medida desde enero de 2011 hasta diciembre de 2016, ambos meses inclusive.

En el análisis clásico de series temporales se considera que una serie temporal queda formada por cuatro componentes:

5.1 Variable Dióxido de azufre (SO₂)

- Tendencia (T): Movimiento regular de la serie, a largo plazo.
- Variaciones Estacionales: Oscilaciones a corto plazo del período regular, de duración menor o igual a un año.
- Variaciones Cíclicas: Movimientos a medio plazo (superior a un año) en torno a la tendencia cuyo período y amplitud pueden presentar cierta regularidad.
- Variaciones regulares o Accidentales (A): Son ciertas fluctuaciones producidas por factores eventuales, esporádicos e imprevisibles, que no muestran una periodicidad predecible.

La tendencia se muestra mediante un suavizamiento lineal (LM) siguiendo una distribución t de Student.

5.1 Variable Dióxido de azufre (SO₂)

El dióxido de azufre es el principal causante de la lluvia ácida, se libera en muchos procesos de combustión, de combustibles fósiles, principalmente carbón y derivados del petróleo y afecta sobre todo a las mucosidades y a los pulmones, provocando ataques de tos.

Dióxido de Azufre (SO ₂)	Valor límite
Máxima horaria	350 $\mu\text{g}/\text{m}^3$
Máxima diaria	125 $\mu\text{g}/\text{m}^3$
Umbraal de Alerta	500 $\mu\text{g}/\text{m}^3$

Tabla 7. Valores límite SO₂

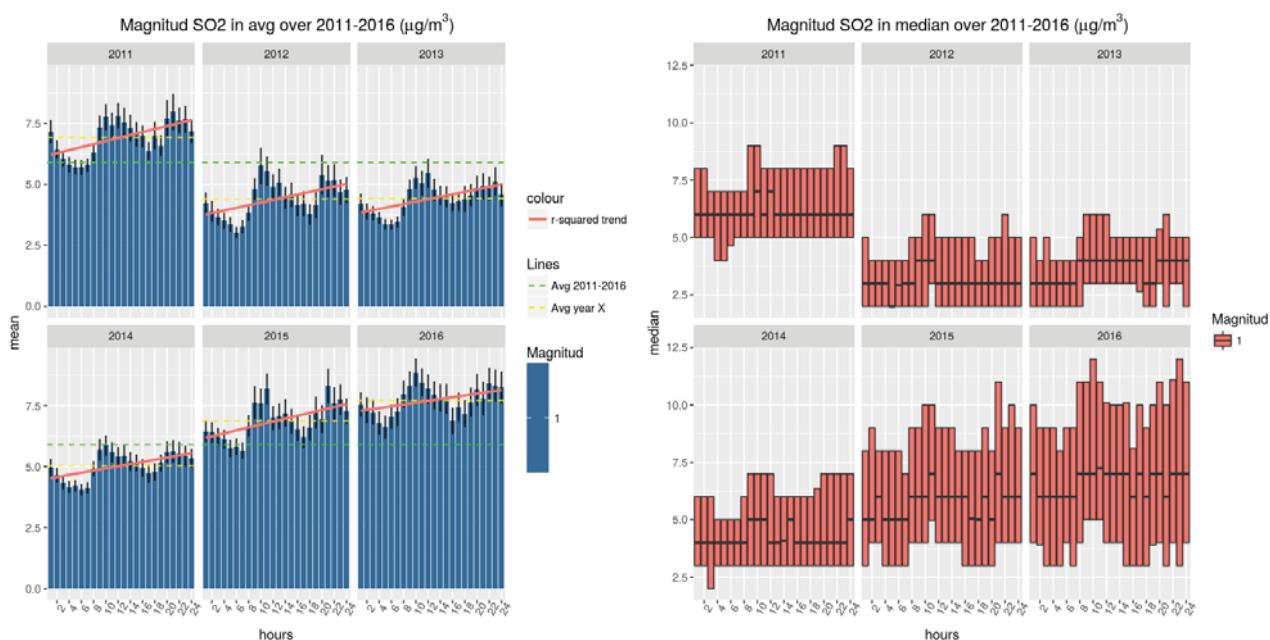


Figura 6.1. Agregado horario SO₂ utilizando la media y mediana como estadísticos

5.1 Variable Dióxido de azufre (SO₂)

A la vista de ambos estadísticos de los agregados horarios, se puede apreciar una estacionalidad en los entre las 9h y 10h de la mañana, y las 21h y 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado una hora antes.

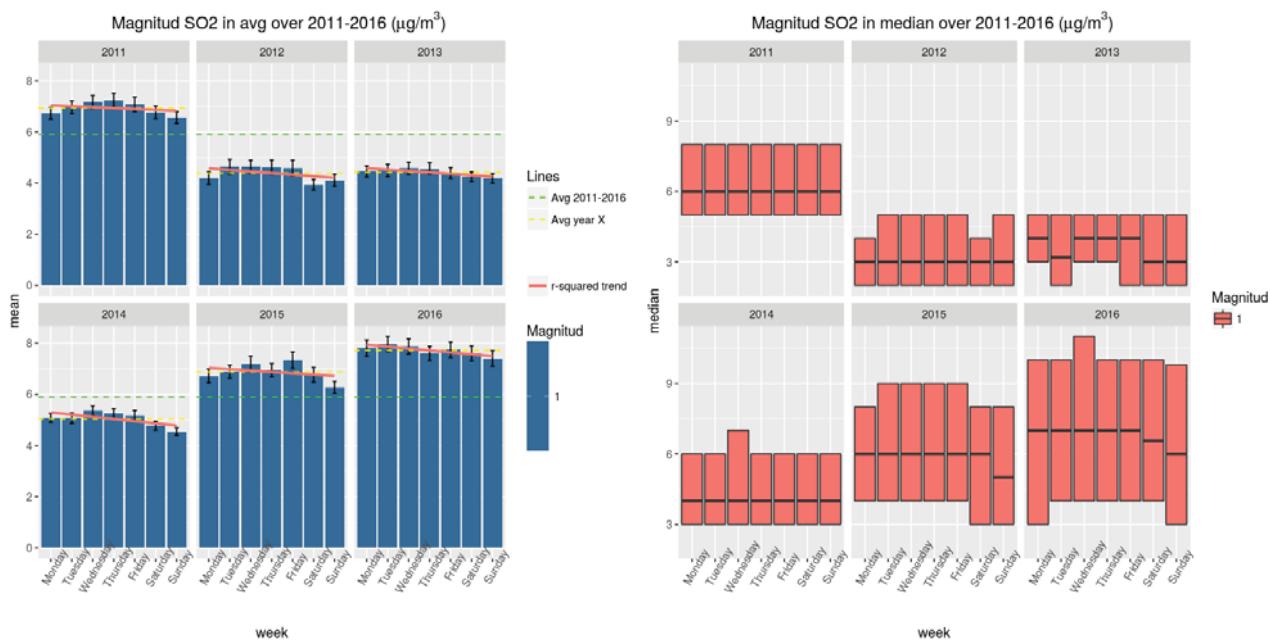


Figura 6.2. Agregado mensual SO₂ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos de los agregados semanales, se puede apreciar una ligera estacionalidad en los días laborales frente a los sábados y domingos, donde las emisiones son algo más bajas.

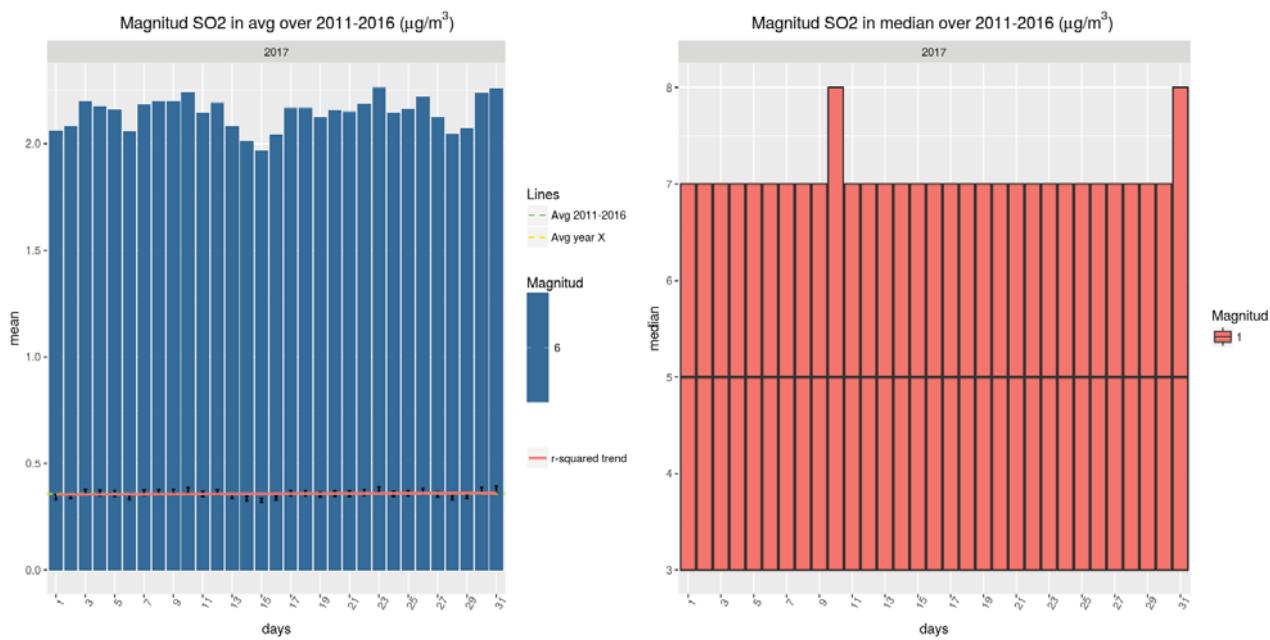


Figura 6.3. Agregado diario SO₂ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos de los agregados diarios, existen mínimos a mediados de mes y máximos el último día del mes, nótese que hay meses que terminan el 30 y meses que terminan el 31.

5.2 Variable Monóxido de carbono (CO)

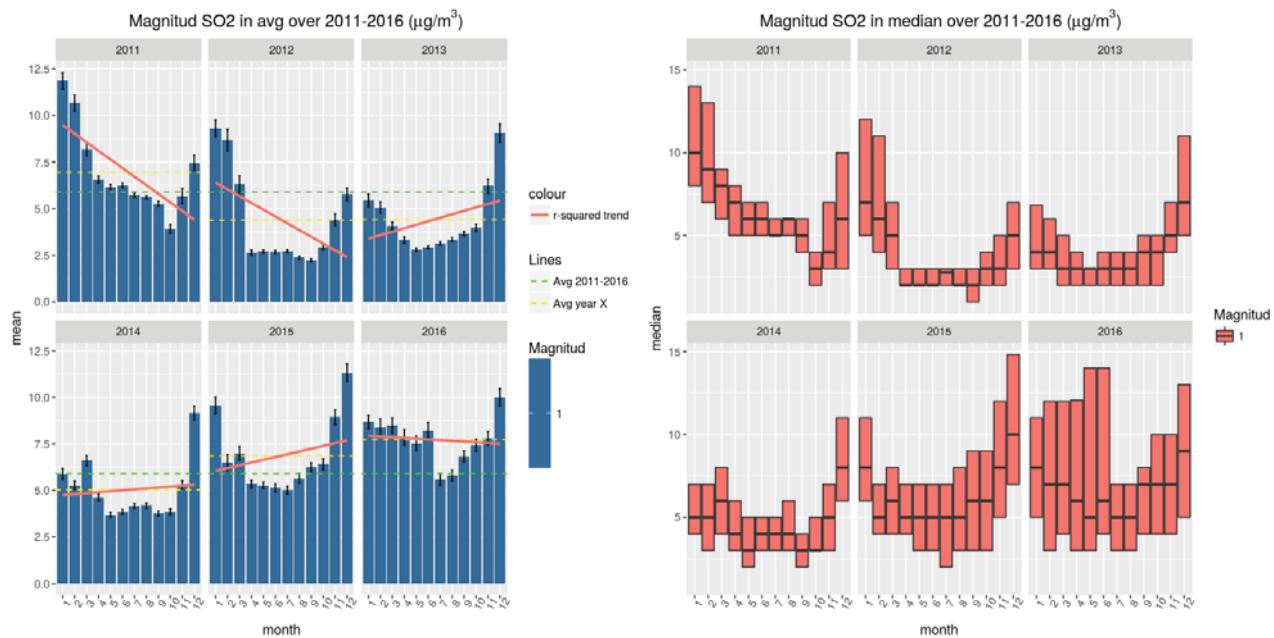


Figura 6.4. Agregado mensual SO₂ utilizando la media y mediana como estadísticos

En ambos agregados mensuales, se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional. Por último mencionar como las emisiones se redujeron en todos los agregados entre 2012 y 2014 por debajo de la media de los 6 años.

En 2016, los valores de SO₂ se sitúan muy por debajo del nivel crítico para la protección de la vegetación establecidos en 20µg/m³, como promedio anual, indicando una buena calidad del Aire respecto a este contaminante en la ciudad de Madrid.

5.2 Variable Monóxido de carbono (CO)

El monóxido de carbono es un subproducto no deseado resultado de numerosos procesos térmicos. Se genera durante todos los procesos de combustión del carbono y sus compuestos cuando no hay saturación de oxígeno. Las fuentes naturales de monóxido de carbono son responsables de más del 90% de la emisión global; el 10% restante se distribuye entre los gases de escape de vehículos (55%), la industria (11%) y otros emisores.

Las intoxicaciones agudas se manifiestan en dolores de cabeza, náuseas, debilidad muscular, pérdida del conocimiento y dificultad respiratoria y, finalmente, muerte, según sea la concentración y tiempo de exposición.

Monóxido de carbono (CO)	Valor límite
Máxima diaria cada 8 horas	10 mg/m ³

Tabla 8. Valores límite CO

5.2 Variable Monóxido de carbono (CO)

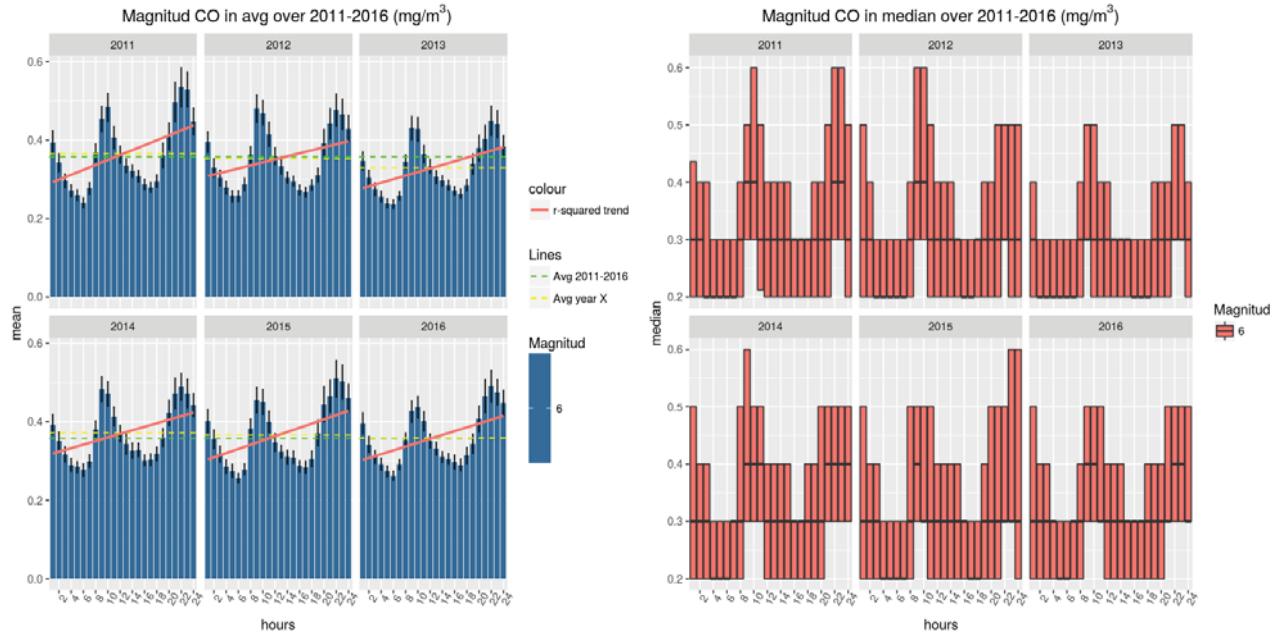


Figura 7.1. Agregado horario CO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 7.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y especialmente los máximos entre las 22h y 23h.

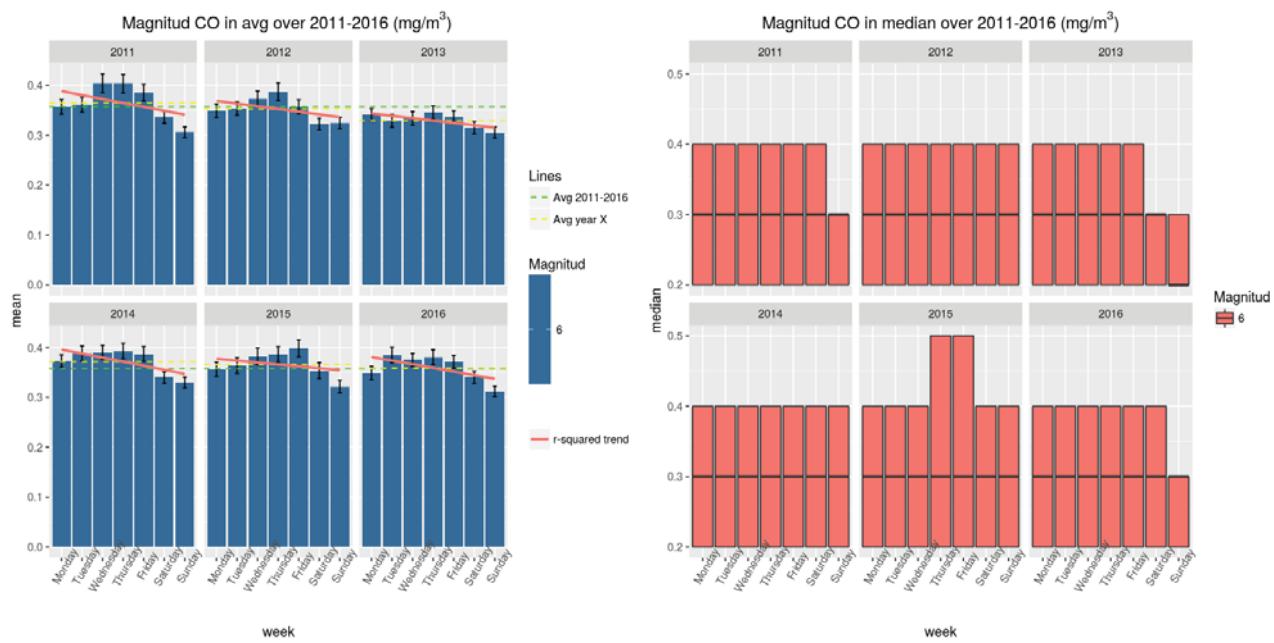


Figura 7.2. Agregado semanal CO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 7.2], se puede apreciar una estacionalidad durante de Martes a Viernes frente a los sábados y domingos, donde las emisiones son más bajas.

5.2 Variable Monóxido de carbono (CO)

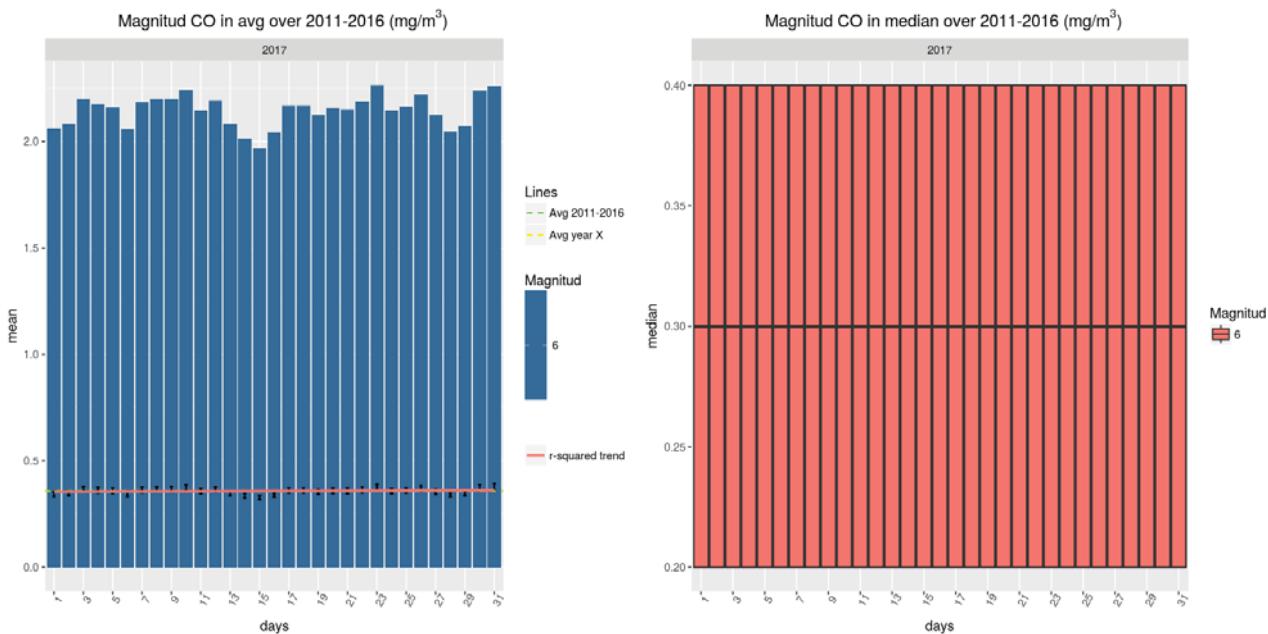


Figura 7.3. Agregado diario CO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 7.3], al igual que sucede con los niveles de SO_2 existen mínimos a mediados de mes y máximos el último día del mes.

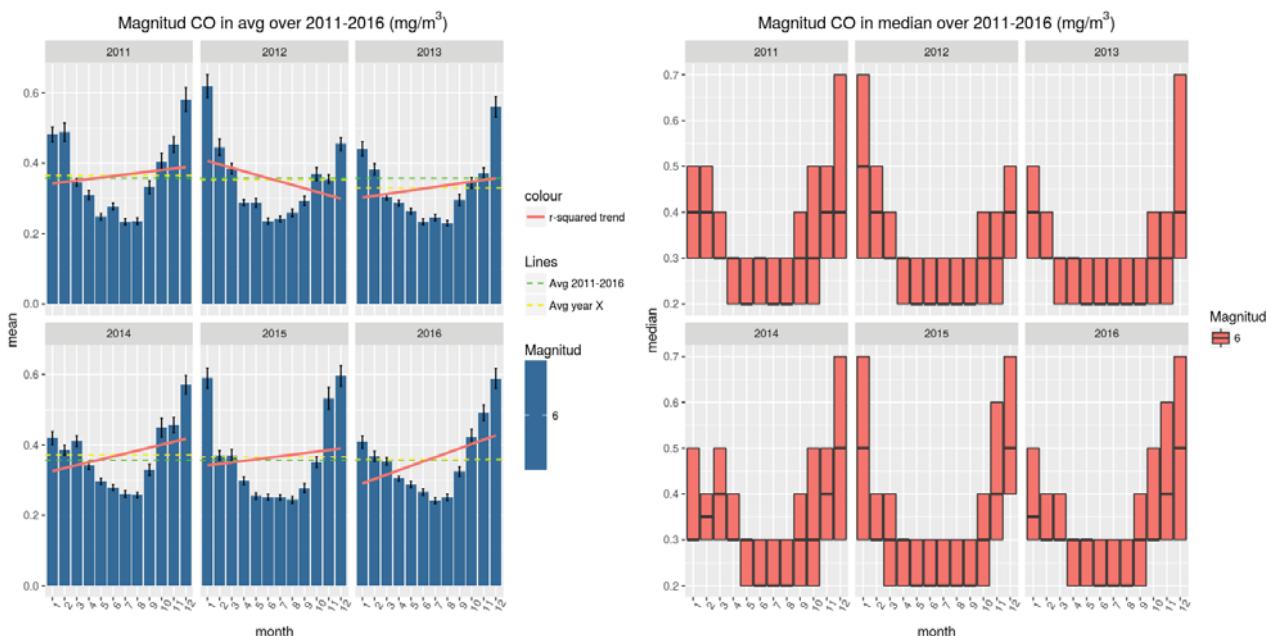


Figura 7.4. Agregado mensual CO utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 7.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional.

5.3 Variable Monóxido de nitrógeno (NO)

En contraste con los valores límites podemos decir que en 2016, no se alcanza el valor límite octohorario para los niveles de CO indicados por el R.D. 102/2011 establecidos en 10mg/m³.

5.3 Variable Monóxido de nitrógeno (NO)

Las altas temperaturas alcanzadas en las combustiones provocan la combinación directa del oxígeno y el nitrógeno del aire para formar óxido nítrico (NO), y éste luego se oxida parcialmente transformándose en NO₂. Por tanto, las instalaciones fijas de combustión, los vehículos de gasolina, y los motores diésel emiten óxidos de nitrógeno con proporciones variables de NO₂ y NO. Posteriormente, el NO inyectado en la atmósfera urbana por las fuentes emisoras se oxida, principalmente debido a reacciones fotoquímicas, y se transforma a NO₂.

La principal fuente emisora de óxidos de nitrógeno a la atmósfera urbana son los vehículos (especialmente los motores diésel) y en menor medida instalaciones de combustión como las calefacciones. Al estar tan íntimamente relacionado con el NO₂, este compuesto no tiene valores umbrales a partir de los cuales es nocivo para el ser humano. Sin embargo, se ha establecido que la media anual de los óxidos de nitrógeno tiene que estar por debajo de los 30µg/m³.

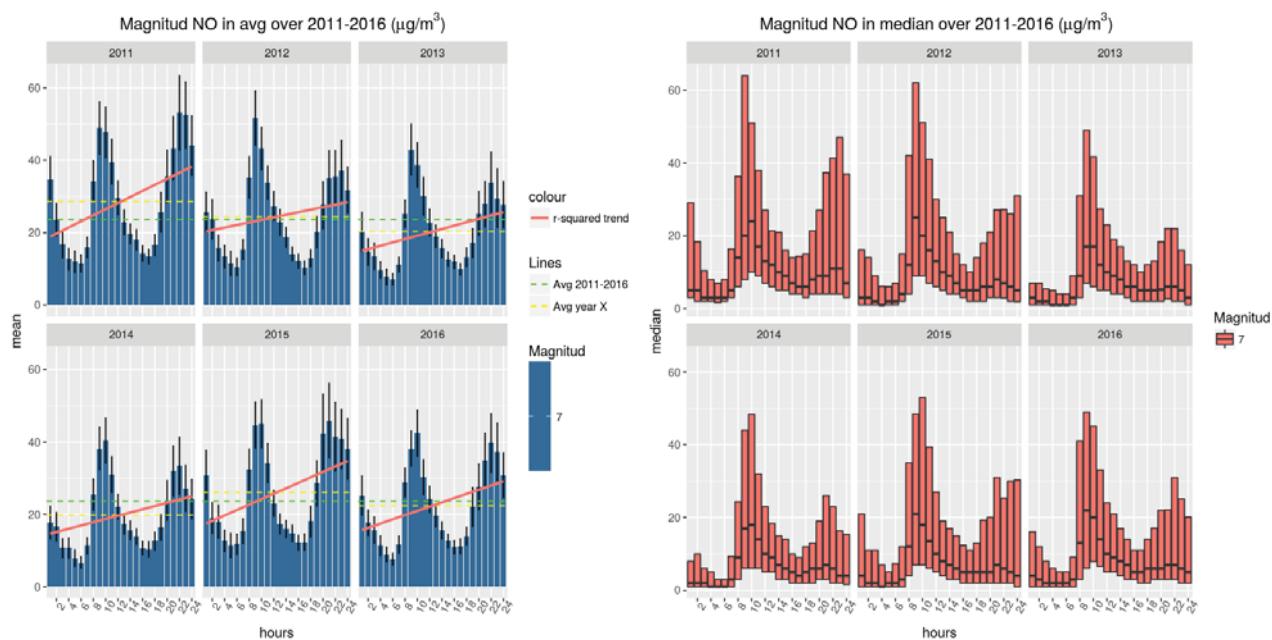


Figura 8.1. Agregado horario NO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 8.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y las 21h y 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado una hora antes.

5.3 Variable Monóxido de nitrógeno (NO)

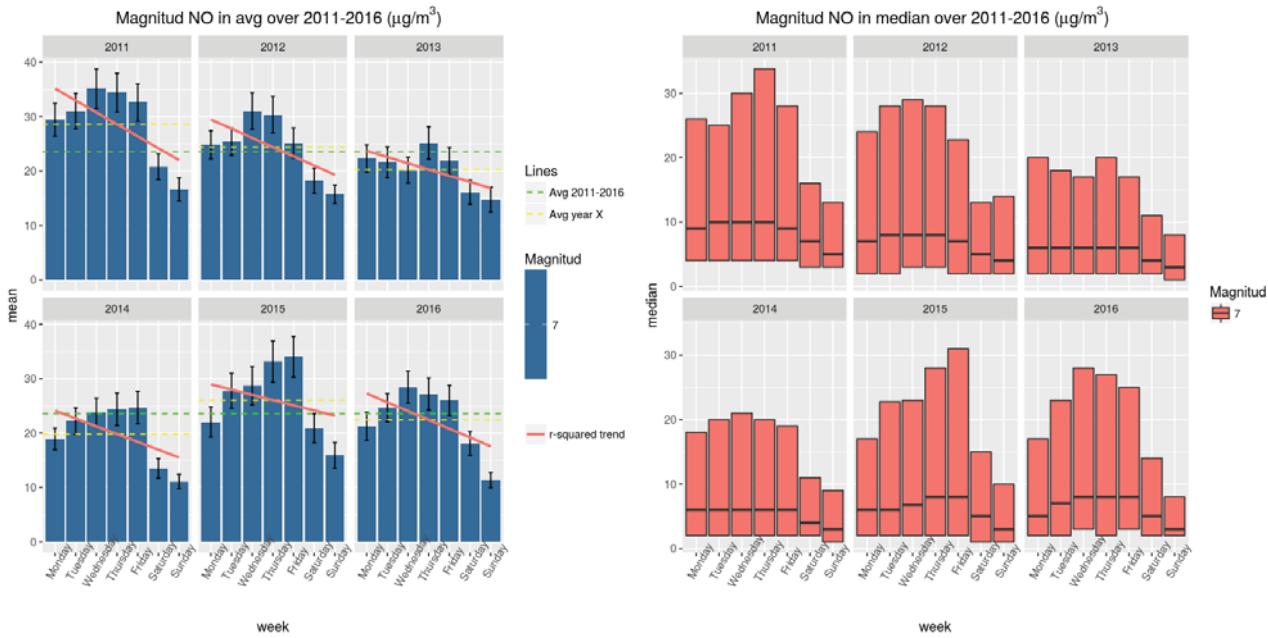


Figura 8.2. Agregado semanal NO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 8.2], se puede apreciar una ligera estacionalidad de Martes a Viernes frente a los sábados y domingos, donde las emisiones son algo más bajas.

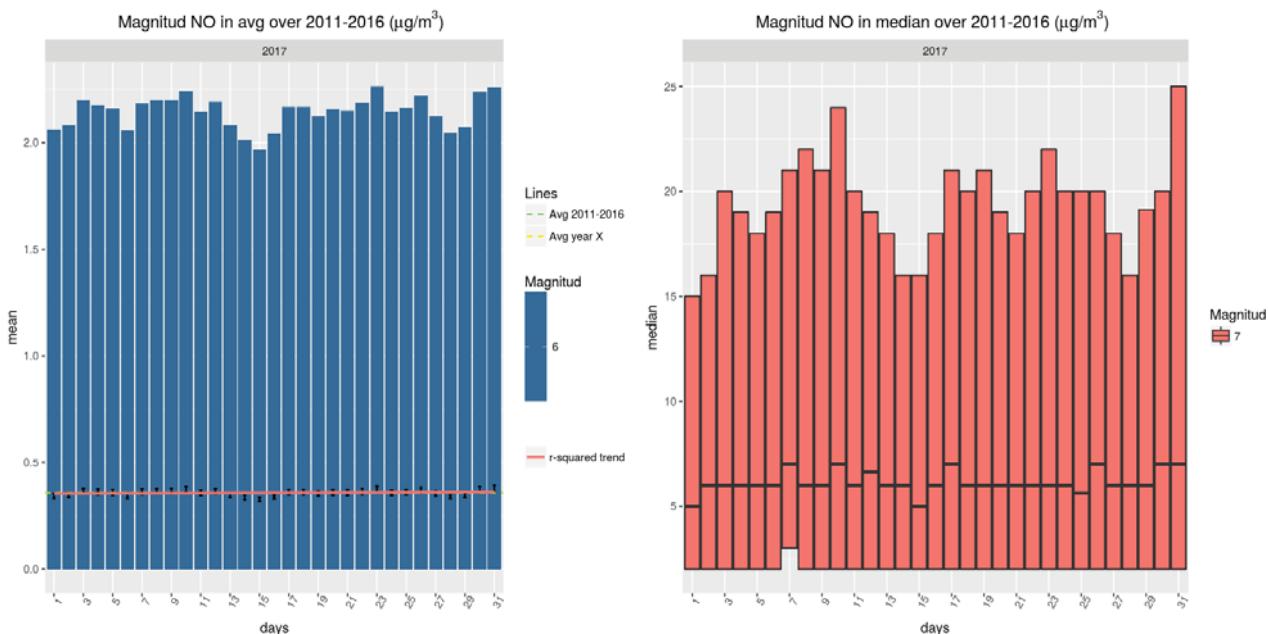


Figura 8.3. Agregado diario NO utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 8.3], como en el caso de los niveles de SO₂ hay mínimos a mediados de mes y máximos el último día del mes.

5.4 Variable Dióxido de nitrógeno (NO₂)

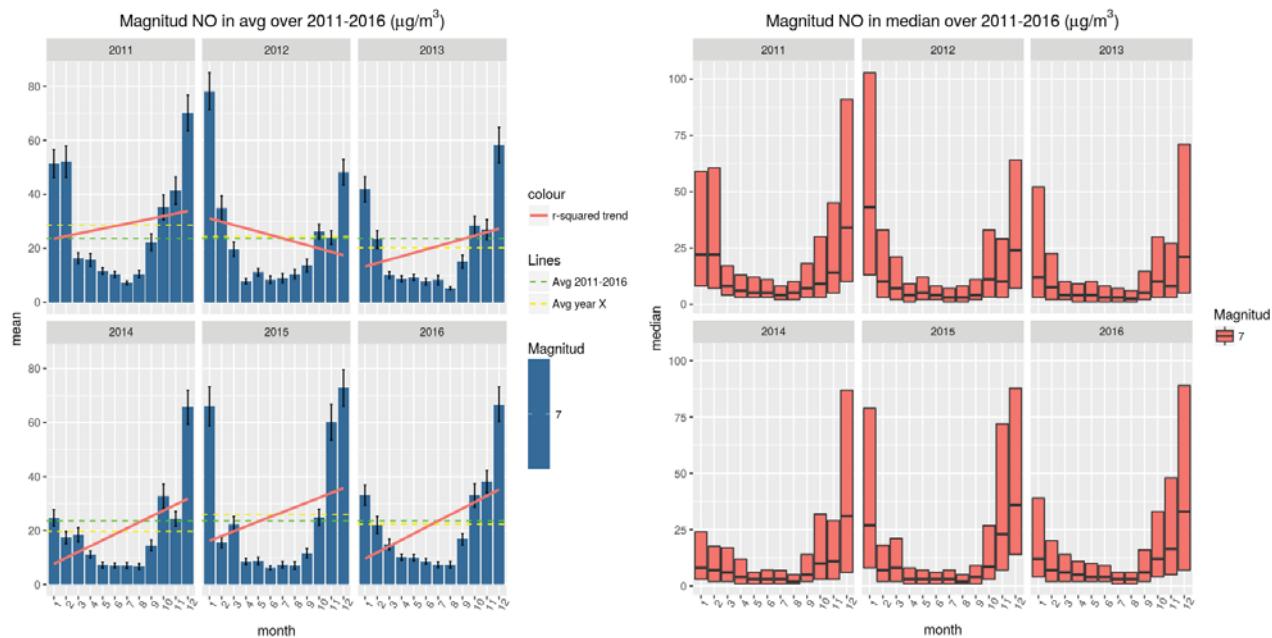


Figura 8.4. Agregado mensual NO utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 8.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional.

En contraste con los valores límites, podemos decir que la media anual en 2016 está por debajo del valor límite establecido en $30\mu\text{g}/\text{m}^3$. Aunque cabe mencionar que en el mes de diciembre se deberían tomar medidas ya que se alcanzan valores muy altos.

5.4 Variable Dióxido de nitrógeno (NO₂)

Las altas temperaturas alcanzadas en las combustiones provocan la combinación directa del oxígeno y el nitrógeno del aire para formar óxido nítrico (NO), y éste luego se oxida parcialmente transformándose en NO₂. Por tanto, las instalaciones fijas de combustión, los vehículos de gasolina, y los motores diésel emiten óxidos de nitrógeno con proporciones variables de NO₂ y NO. Posteriormente, el NO injectado en la atmósfera urbana por las fuentes emisoras se oxida, principalmente debido a reacciones fotoquímicas, y se transforma a NO₂.

La principal fuente emisora de óxidos de nitrógeno a la atmósfera urbana son los vehículos (especialmente los motores diésel) y en menor medida instalaciones de combustión como las calefacciones.

Los valores límites indicados por el R.D. 102/2011 son de $40\mu\text{g}/\text{m}^3$ como promedio anual, y de $200\mu\text{g}/\text{m}^3$ como valor límite horario, el cual sólo podrá superarse 18 ocasiones al año.

5.4 Variable Dióxido de nitrógeno (NO_2)

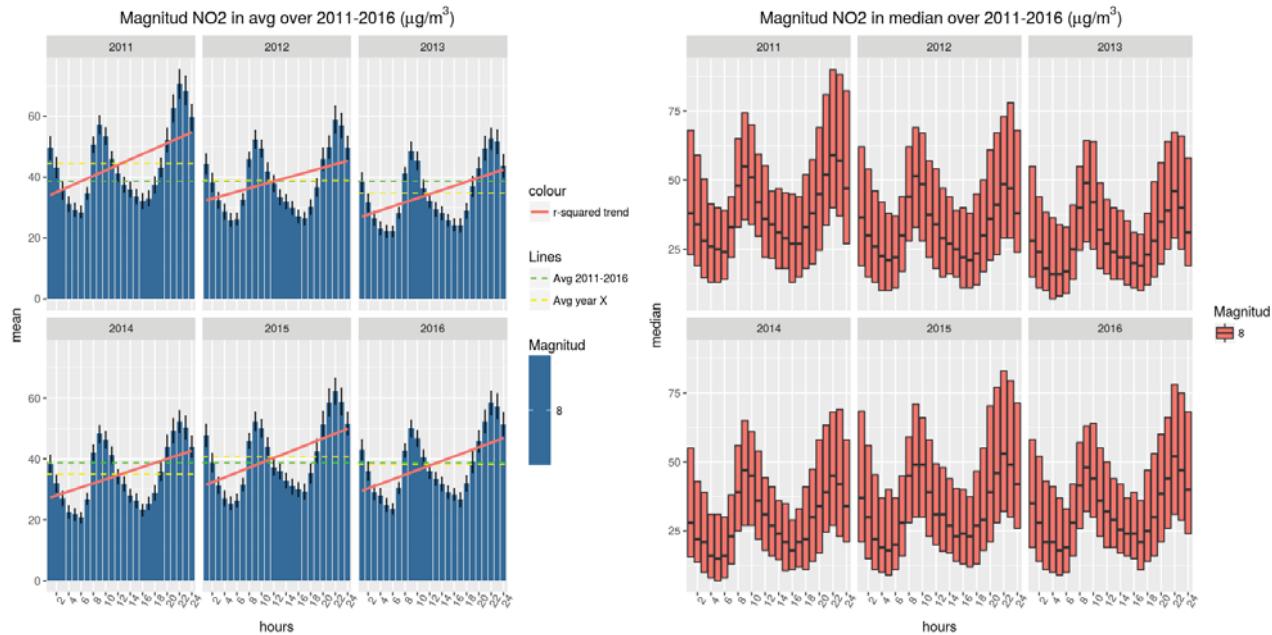


Figura 9.1. Agregado horario NO_2 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 9.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y las 22h y 23h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generadas dos o tres horas antes.

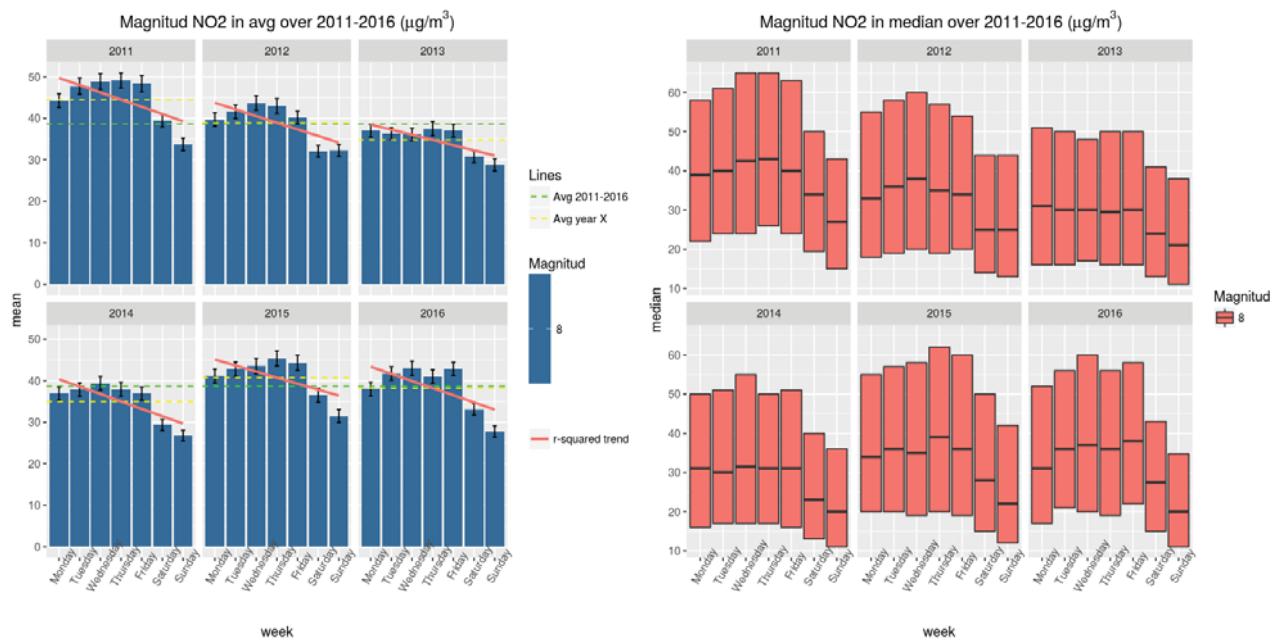


Figura 9.2. Agregado semanal NO_2 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 9.2], se puede apreciar una ligera estacionalidad de Martes a Viernes frente a los sábados y domingos, donde las emisiones son algo más bajas.

5.4 Variable Dióxido de nitrógeno (NO_2)

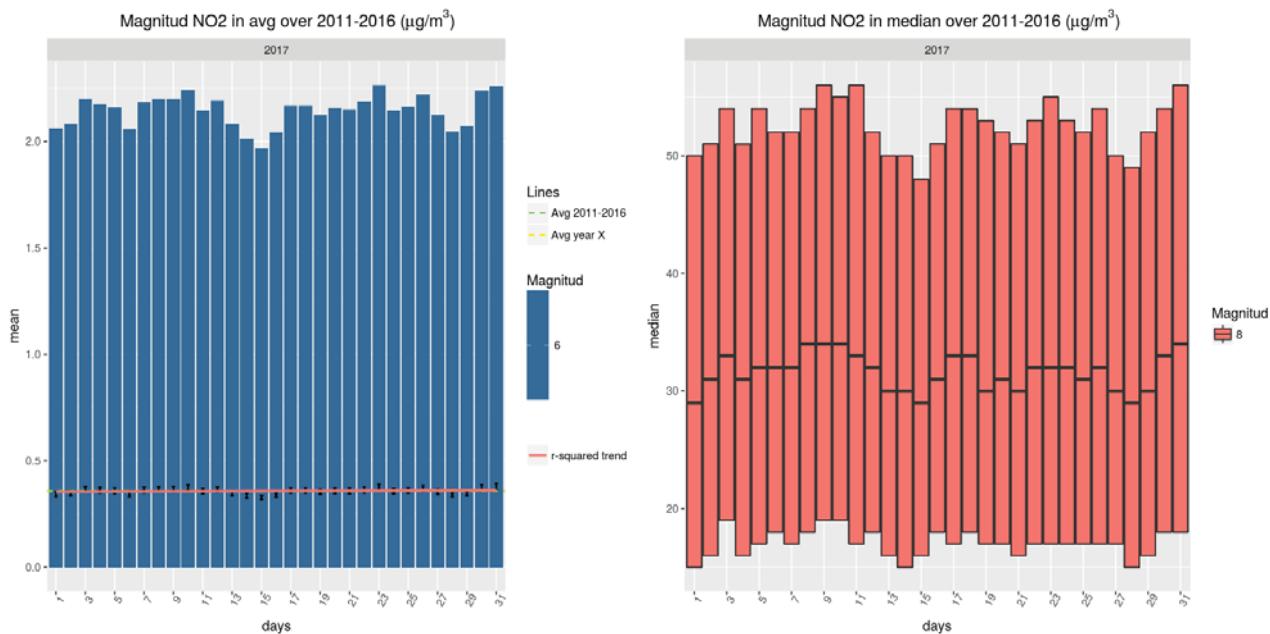


Figura 9.3. Agregado diario NO_2 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 9.3], al igual que sucede con los niveles de SO_2 existen mínimos a mediados de mes y máximos el último día del mes.

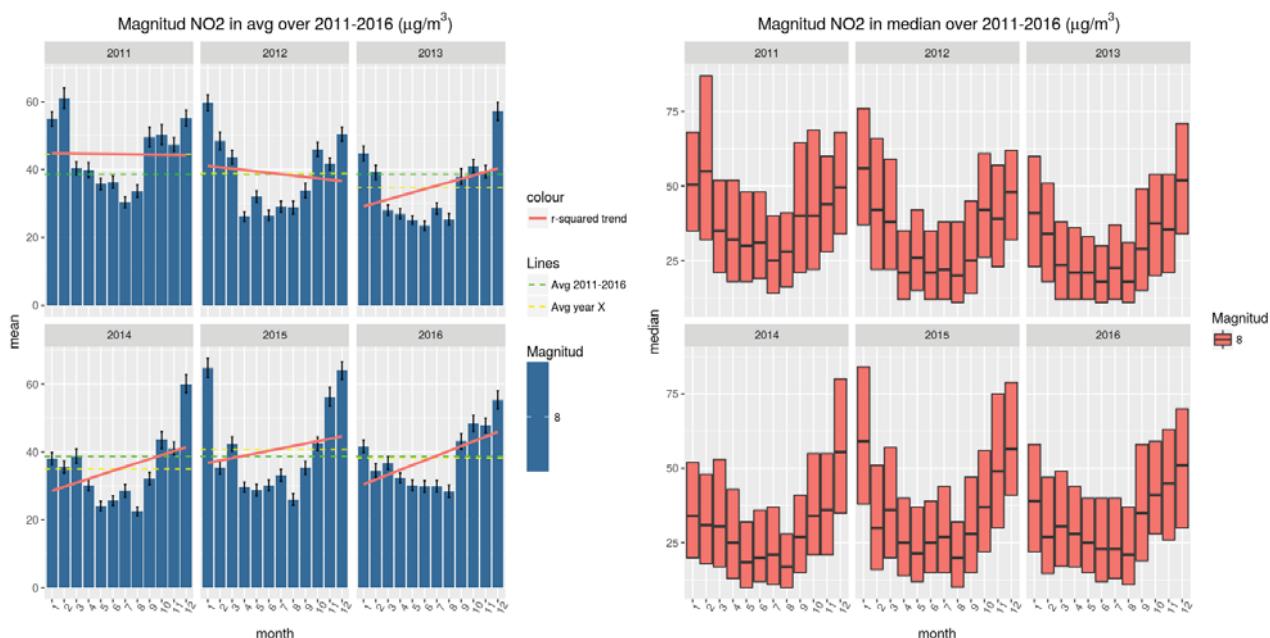


Figura 9.4. Agregado mensual NO_2 utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 9.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas. Sin embargo, en los dos primeros años es más acentuada en los dos primeros meses del año.

En 2016, los valores de NO_2 se sitúan en promedio anual ligeramente por debajo de $40 \mu\text{g}/\text{m}^3$, valor límite indicado por el R.D. 102/2011. Este valor se supera en el último cuatrimestre del año tanto en promedio como en mediana. El valor límite horario indicado por el R.D. 102/2011, es de $200 \mu\text{g}/\text{m}^3$, el cual sólo se podrá superar 18 ocasiones al año. Este valor se encuentra muy lejos del rango intercuartílico, al igual que el umbral de alerta establecido en $400 \mu\text{g}/\text{m}^3$. Cabe mencionar que sería recomendable tomar alguna medida

5.5 Variable Partículas en Suspensión de diámetro inferior a 2,5 mg (PM2.5)

para el último cuatrimestre del año, donde todos los meses superan el promedio anual y ponen en riesgo el objetivo del promedio anual.

5.5 Variable Partículas en Suspensión de diámetro inferior a 2,5 mg (PM2.5)

El material en partículas respirable presente en la atmósfera de nuestras ciudades en forma sólida o líquida (polvo, cenizas, hollín, partículas metálicas, cemento y polen, entre otras) se puede dividir, según su tamaño, en dos grupos principales. A las de diámetro aerodinámico igual o inferior a los 10 μm o 10 micrómetros (1 μm corresponde a la milésima parte de un milímetro) se las denomina PM10 y a la fracción respirable más pequeña, PM2,5. Estas últimas están constituidas por aquellas partículas de diámetro aerodinámico inferior o igual a los 2,5 micrómetros, es decir, son 100 veces más delgadas que un cabello humano.

Además, el tamaño no es la única diferencia. Cada tipo de partículas está compuesto de diferente material y puede provenir de diferentes fuentes. En el caso de las PM2,5, su origen está principalmente en fuentes de carácter antropogénico como las emisiones de los vehículos diésel, mientras que las partículas de mayor tamaño pueden tener en su composición un importante componente de tipo natural, como partículas de polvo procedente de las intrusiones de viento del norte de África (polvo sahariano), frecuente en nuestras latitudes.

Los últimos trabajos científicos sugieren que este tipo de contaminación, y particularmente las partículas procedentes del tráfico urbano, está asociado con incrementos en la morbi-mortalidad de la población expuesta y al creciente desarrollo del asma y alergias entre la población infantil.

Según los citados Valores Guía de protección para la salud de la OMS para valores medios anuales, los niveles más bajos a partir de los cuales se han detectado asociación entre efectos cardiopulmonares y mortalidad por cáncer de pulmón debido a la exposición prolongada a PM2,5.

PM2.5	Umbrales más bajos de riesgo
Media anual	10 $\mu\text{g}/\text{m}^3$
Media horaria (24h)	25 $\mu\text{g}/\text{m}^3$

Tabla 9. Valores límite PM2.5

5.5 Variable Partículas en Suspensión de diámetro inferior a 2,5 mg (PM2.5)

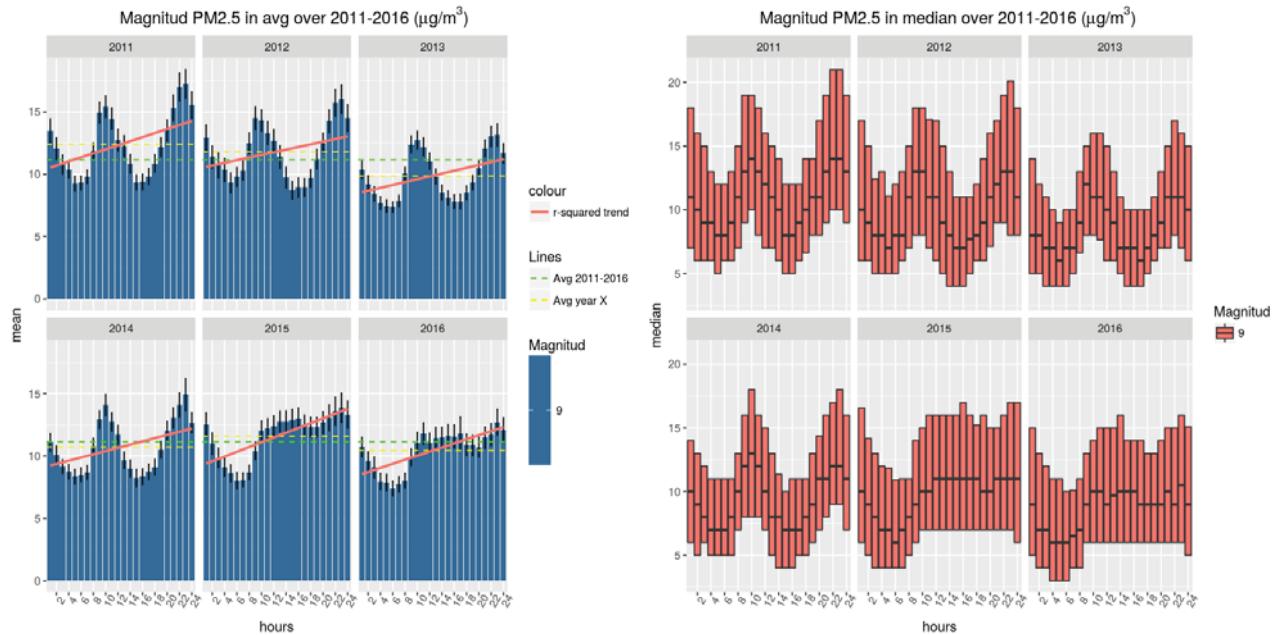


Figura 10.1. Agregado horario PM2.5 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 10.1], se puede apreciar una estacionalidad en a las 10h de la mañana, y las 23h. Sin embargo, a partir de 2015 las emisiones son bastante más altas a partir de las 10h.

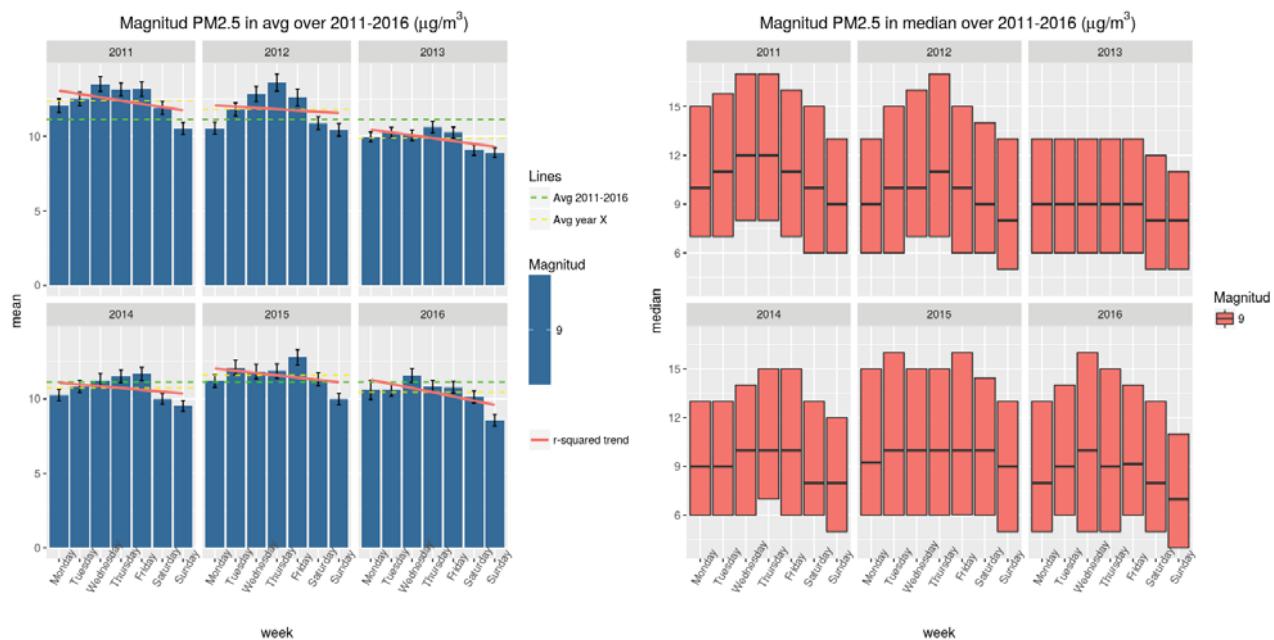


Figura 10.2. Agregado semanal PM2.5 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 10.2], se puede apreciar una ligera estacionalidad en los días laborales frente a los sábados y domingos, donde las emisiones son algo más bajas. Cabe destacar que en 2013, 2014 y 2016 las emisiones son algo más bajas.

5.5 Variable Partículas en Suspensión de diámetro inferior a 2,5 mg (PM2.5)

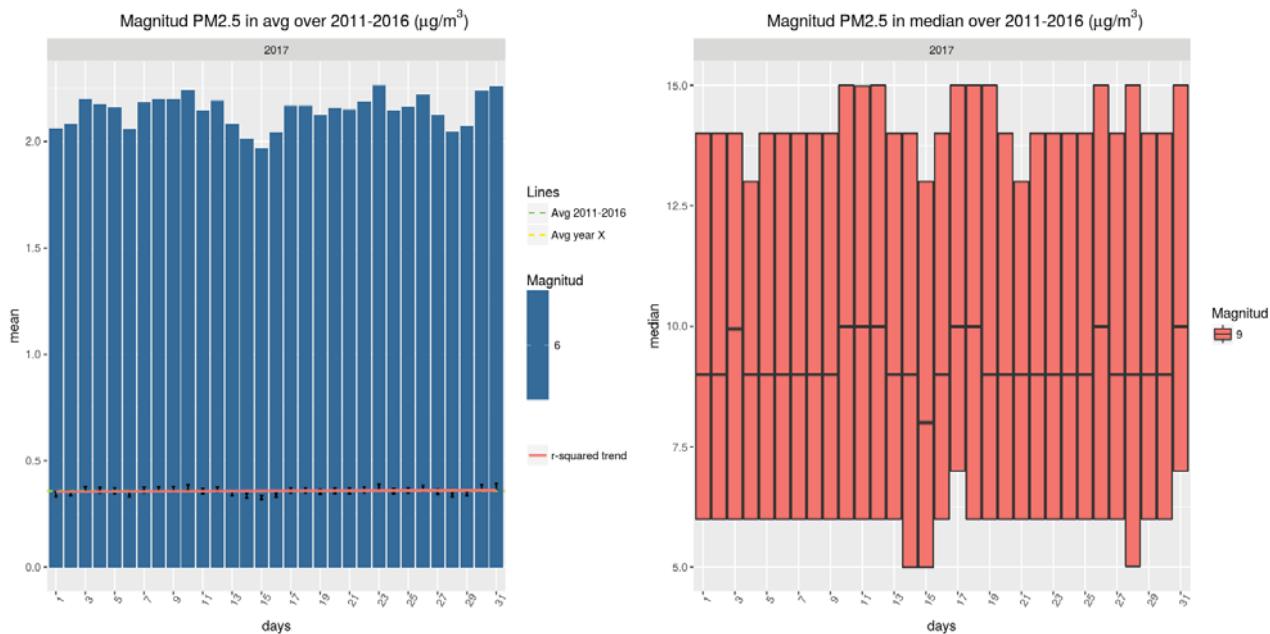


Figura 10.3. Agregado diario PM2.5 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 10.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

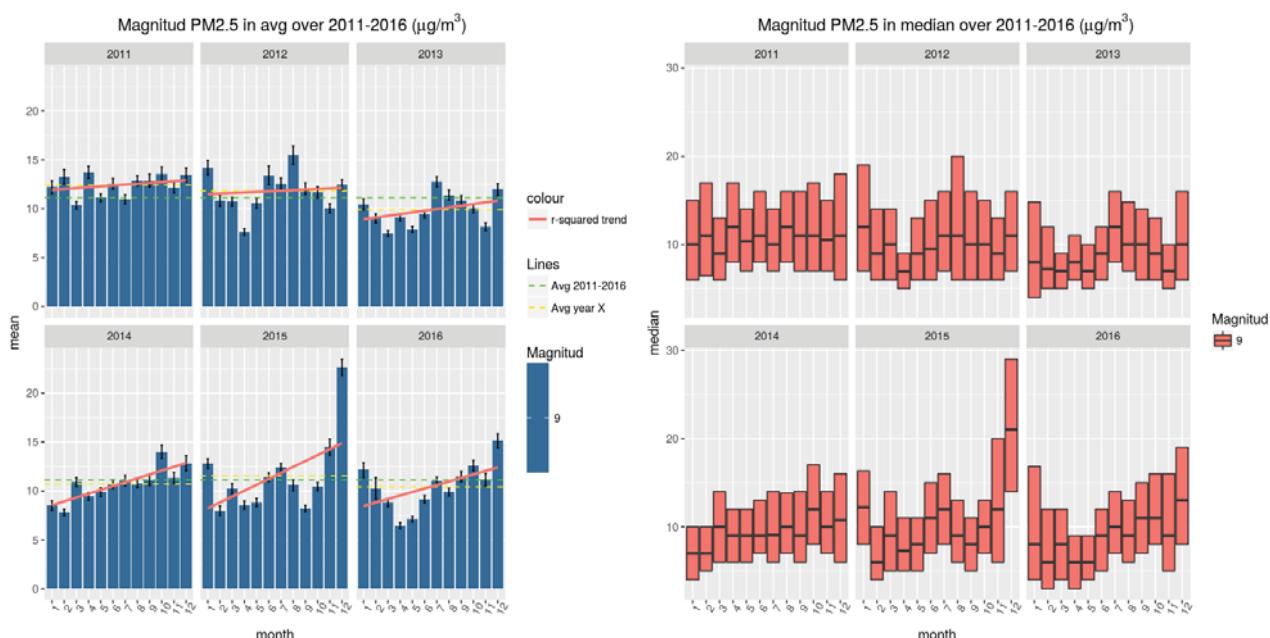


Figura 10.4. Agregado mensual PM2.5 utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 10.4], no se aprecia estacionalidad. No obstante, debido a los máximos en los meses de verano e invierno, se observa una tendencia creciente conforme se finaliza el año, posiblemente debido al periodo vacacional.

En 2016, los valores de PM2.5 están por debajo del valor límite indicado por el R.D. 102/2011, establecido en 25µg/m³ en promedio anual. El valor guía establecido por la OMS para este

5.6 Variable Partículas en Suspensión de diámetro superior a 10 mg (PM10)

contaminante es de $10\mu\text{g}/\text{m}^3$, el cual se supera ligeramente como promedio anual, especial atención en diciembre que alcanza hasta más de $15\mu\text{g}/\text{m}^3$.

5.6 Variable Partículas en Suspensión de diámetro superior a 10 mg (PM10)

Según la normativa europea y su transposición a la legislación española en el R.D. 1073/2002¹⁵, el cumplimiento del valor límite anual de PM10 –no superar más de 35 días al año el valor límite diario. Los valores fijados en las Directrices son:

PM10	Umbrales más bajos de riesgo
Media anual	$20\mu\text{g}/\text{m}^3$
Media diaria	$50\mu\text{g}/\text{m}^3$

Tabla 10. Valores límite PM10

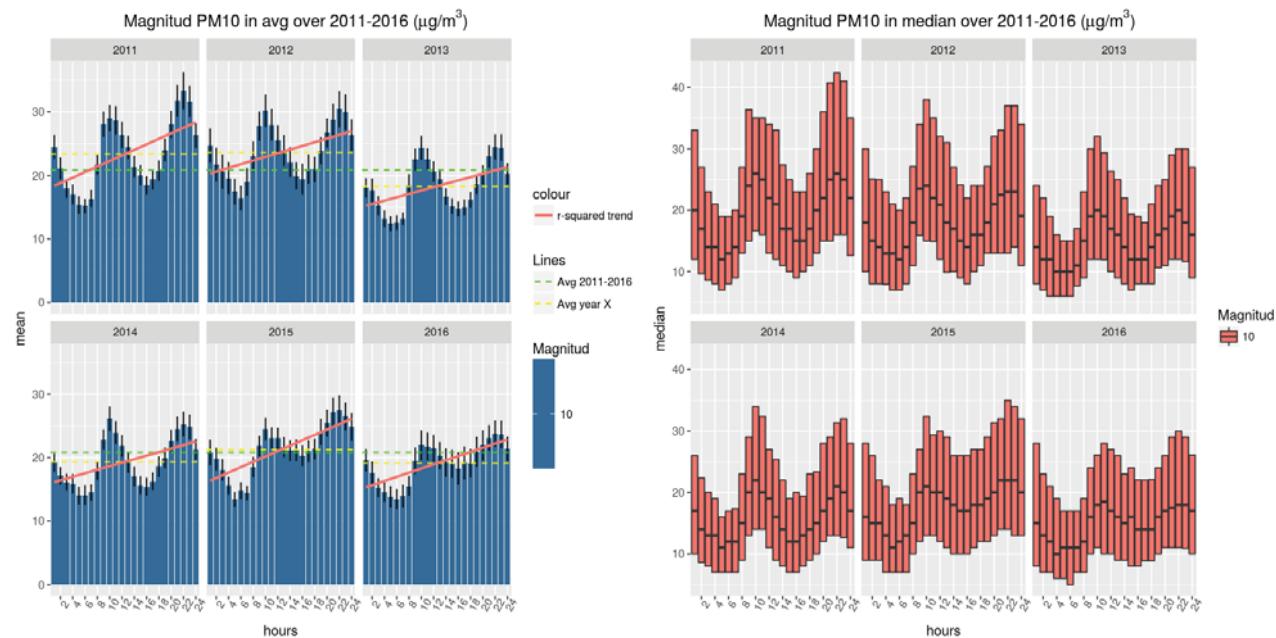


Figura 11.1. Agregado horario PM10 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 11.1], se puede apreciar una estacionalidad en entre las 10h de la mañana, y las 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado en horas previas.

5.6 Variable Partículas en Suspensión de diámetro superior a 10 mg (PM10)

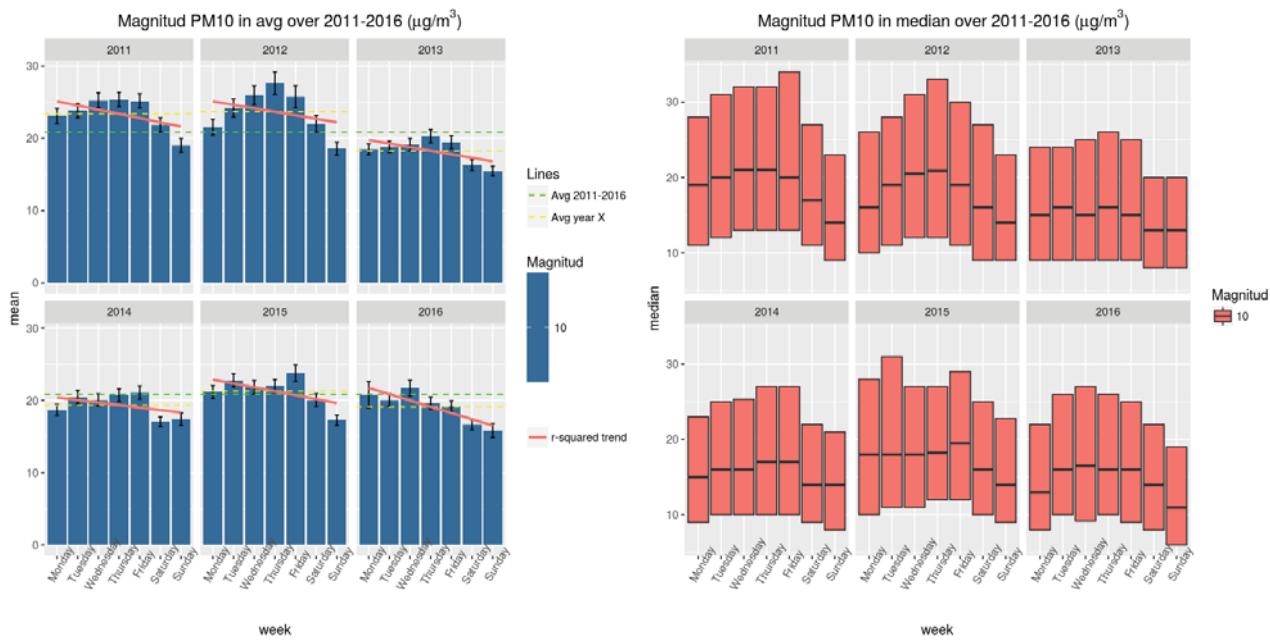


Figura 11.2. Agregado semanal PM10 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 11.2], se puede apreciar una ligera estacionalidad en los días laborales frente a los sábados y domingos, donde las emisiones son algo más bajas.

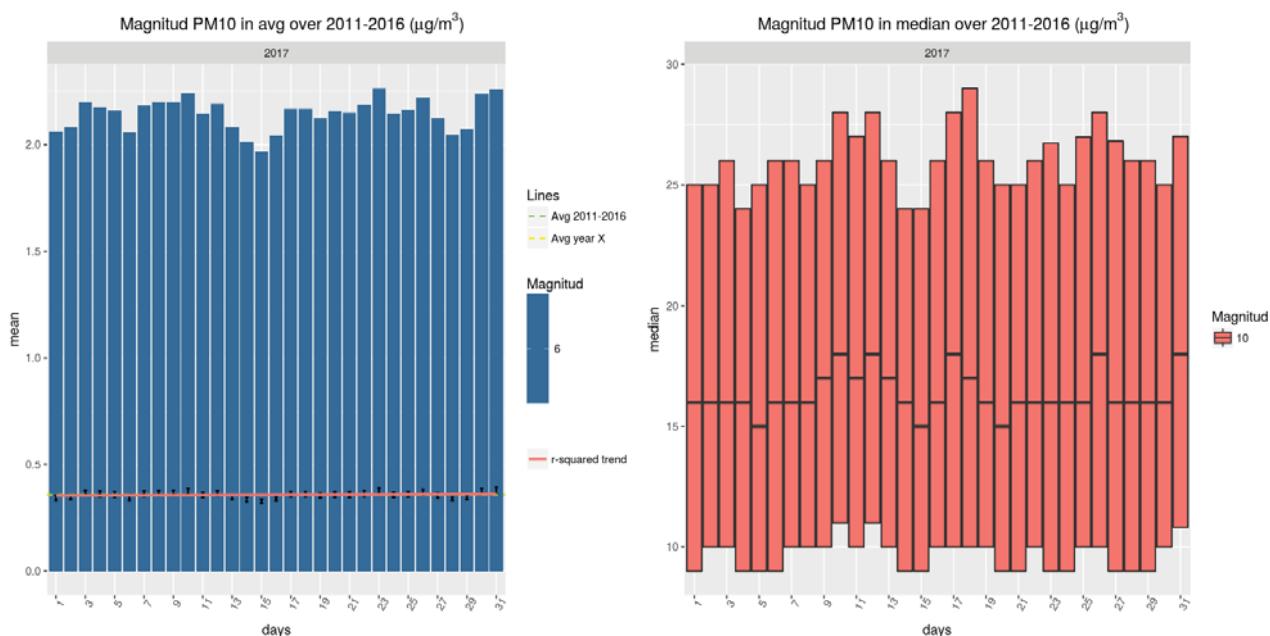


Figura 11.3. Agregado diario PM10 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 11.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

5.7 Variable Ozono (O3)

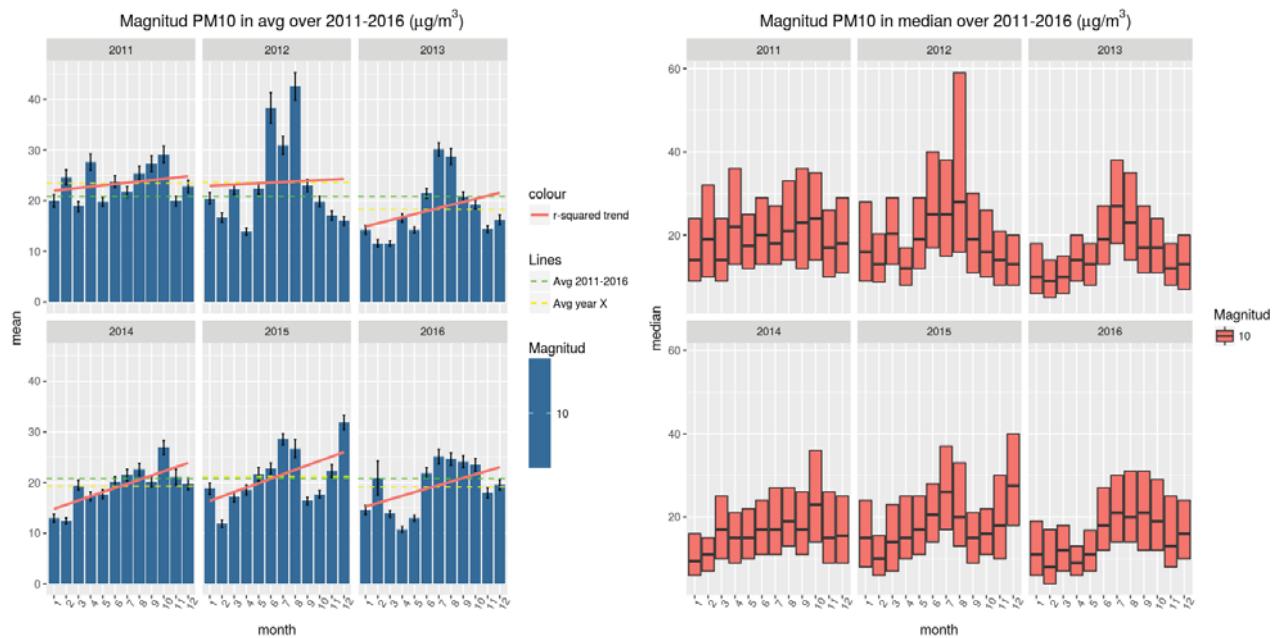


Figura 11.4. Agregado mensual PM10 utilizando la media y mediana como estadísticos

En contraste, aunque no se aprecia gran estacionalidad en los datos, en estos agregados mensuales [F. 11.4] los valores máximos, salvo en 2015, se encuentran en los meses de otoño y verano, con mínimos entre los cuatro primeros meses del año.

En 2016, los valores de PM10 se sitúan por debajo del valor límite indicado por el R.D. 102/2011 establecido en $40\mu\text{g}/\text{m}^3$ en promedio anual. El valor límite para el promedio diario es de $50\text{mg}/\text{m}^3$, el cual sólo se ha alcanzado el último día de mes aunque es en promedio de los 6 años. Por tanto, podemos decir que no es del todo relevante. El valor guía establecido por la OMS para este contaminante es de $20\mu\text{g}/\text{m}^3$, el cual se supera en promedio ligeramente los días entre semana, en especial los miércoles alcanzando casi $22\mu\text{g}/\text{m}^3$, sin embargo se debe a horas puntas pues la mediana nunca supera tal valor durante los días de la semana. Observando el estadístico de las horas se puede observar como la mediana llega a alcanzar dicho límite alrededor de las 9h y las 22h.

5.7 Variable Ozono (O3)

El ozono troposférico —que no debe confundirse con la capa de ozono en la atmósfera superior— es uno de los principales componentes de la niebla tóxica. Este se forma en la atmósfera mediante una reacción entre la luz solar (fotoquímica) con la emisión de contaminantes atmosféricos primarios, tales como el dióxido de azufre (SO_2), óxidos de nitrógeno (NO_x), compuestos orgánicos volátiles (COV) y amoníaco (NH_3), procedentes de las emisiones de vehículos, los disolventes o la industria. Los niveles de ozono más elevados se registran durante los períodos de tiempo soleado.

Los techos nacionales de emisión de estos contaminantes se encuentran regulados por la Directiva 2001/81/CE²⁴ del Parlamento Europeo y del Consejo, de 23 de octubre de 2001.

Dicha directiva establece los siguientes techos de emisión, expresados en kilo-toneladas, para España.

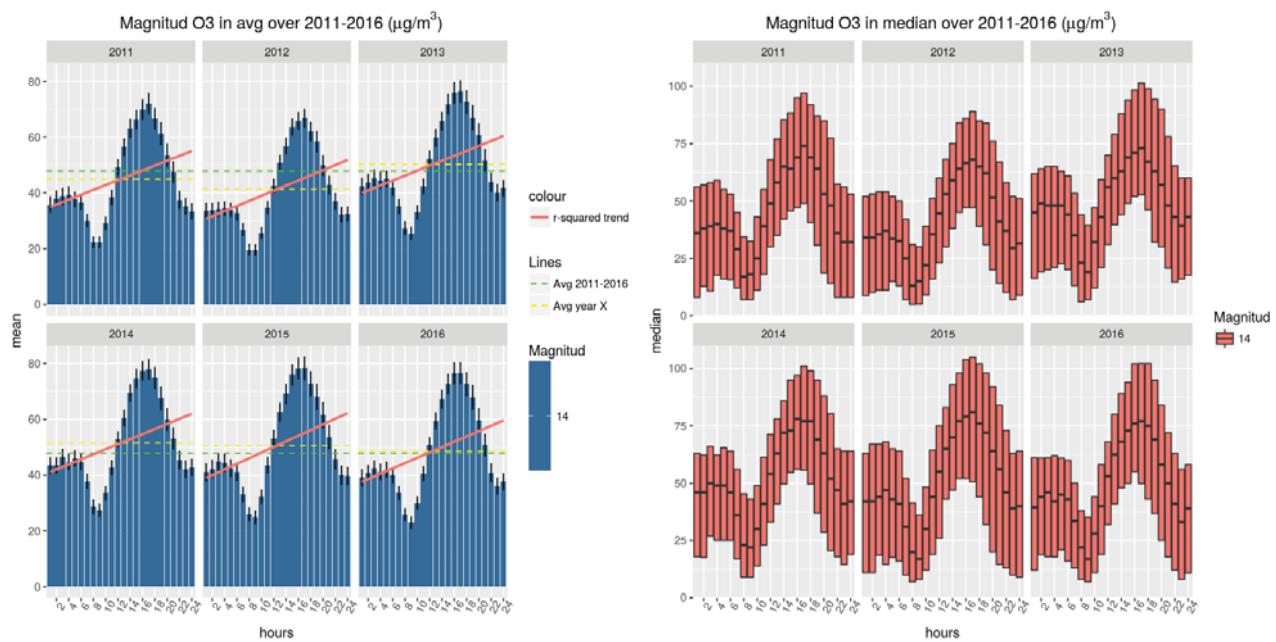
Dióxido de azufre (SO ₂)	Óxidos de nitrógeno (NO _x)	Compuestos orgánicos volátiles (COV)	Amoníaco (NH ₃)
746	847	662	353

Tabla 11.1. Valores límite de contaminantes que producen O_3

Un aumento en la concentración del ozono troposférico, localizado en la parte inferior de la atmósfera terrestre (donde se encuentra el aire que respiramos), puede afectar a las vías respiratorias, provocando tos, dolor de cabeza o náuseas.

En las Directrices de la OMS sobre la Calidad del Aire se estima como umbral de nivel de concentración $100\mu\text{g}/\text{m}^3$ de media móvil octohoraria, y según el R.D. el umbral de información y de alerta como promedio horario, y valor objetivo para 2014-2016 como media móvil octohoraria:

Umbral de información	Umbral de alerta	Valor objetivo
$180\mu\text{g}/\text{m}^3$	$240\mu\text{g}/\text{m}^3$	$125\mu\text{g}/\text{m}^3$

Tabla 11.2. Umbrales de O_3 **Figura 12.1.** Agregado horario O_3 utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 12.1], se puede apreciar una gran estacionalidad con mínimos entre las 8h y 9h de la mañana, y máximos en torno a las 16h conforme transcurren las horas de la tarde, donde las emisiones son más altas.

5.7 Variable Ozono (O₃)

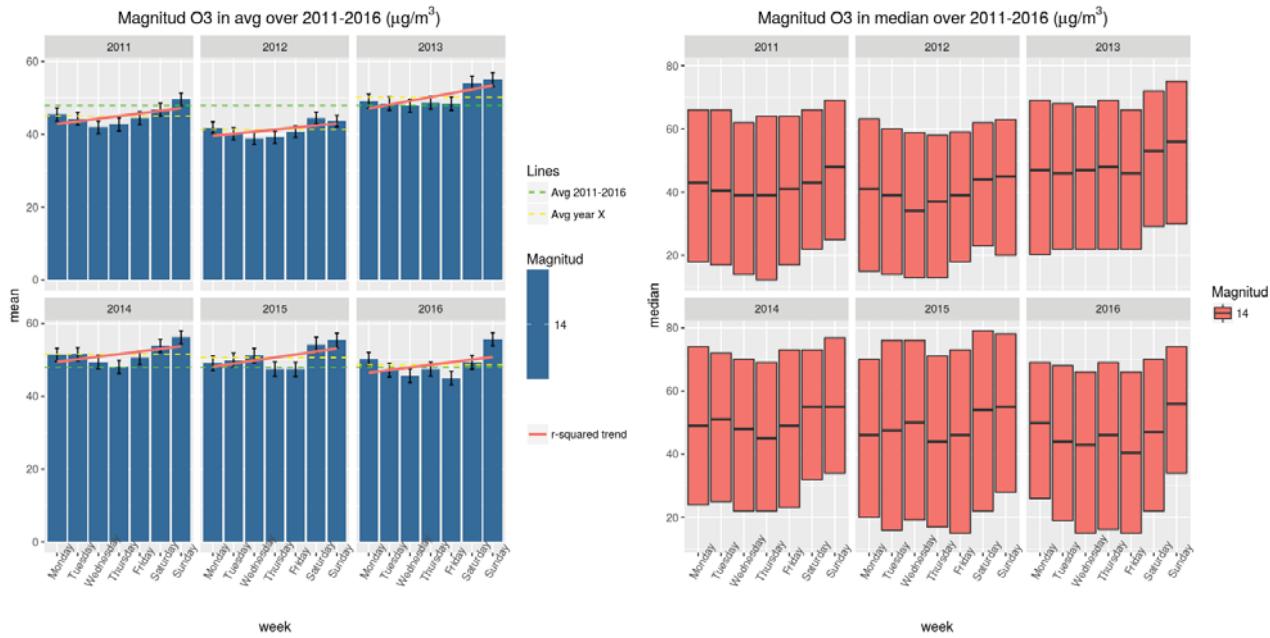


Figura 12.2. Agregado semanal O₃ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 12.2], se puede apreciar una estacionalidad con máximos en Sábados y Domingos, respecto a las emisiones durante la semana.

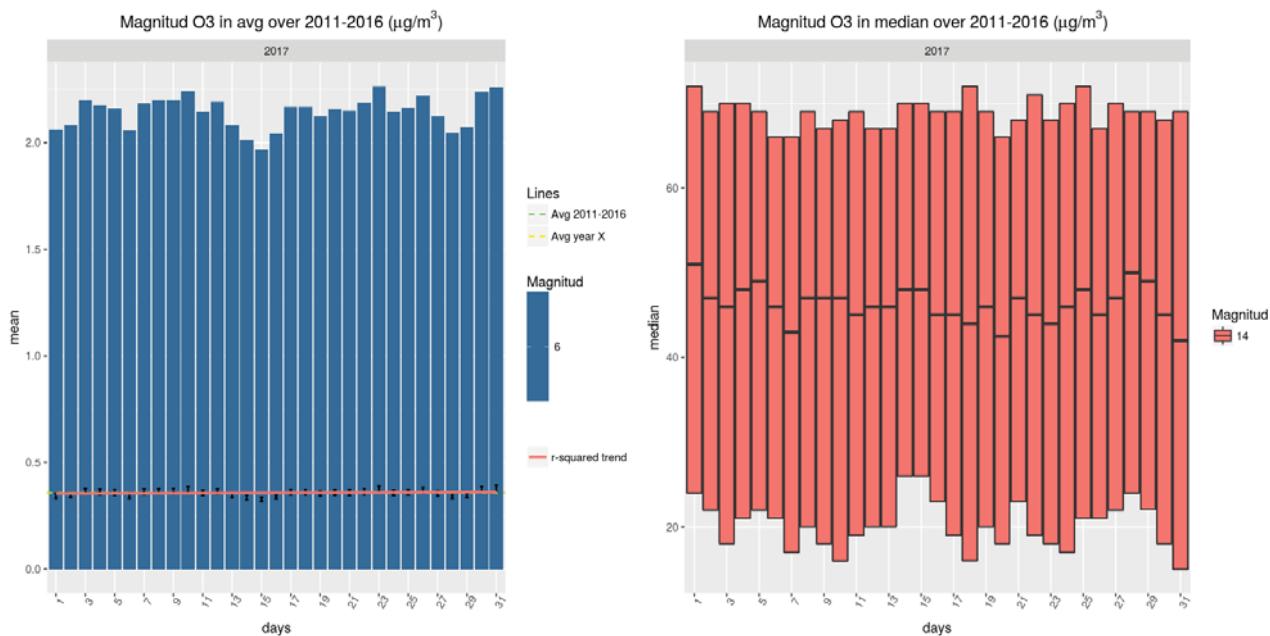


Figura 12.3. Agregado diario O₃ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 12.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

5.7 Variable Ozono (O₃)

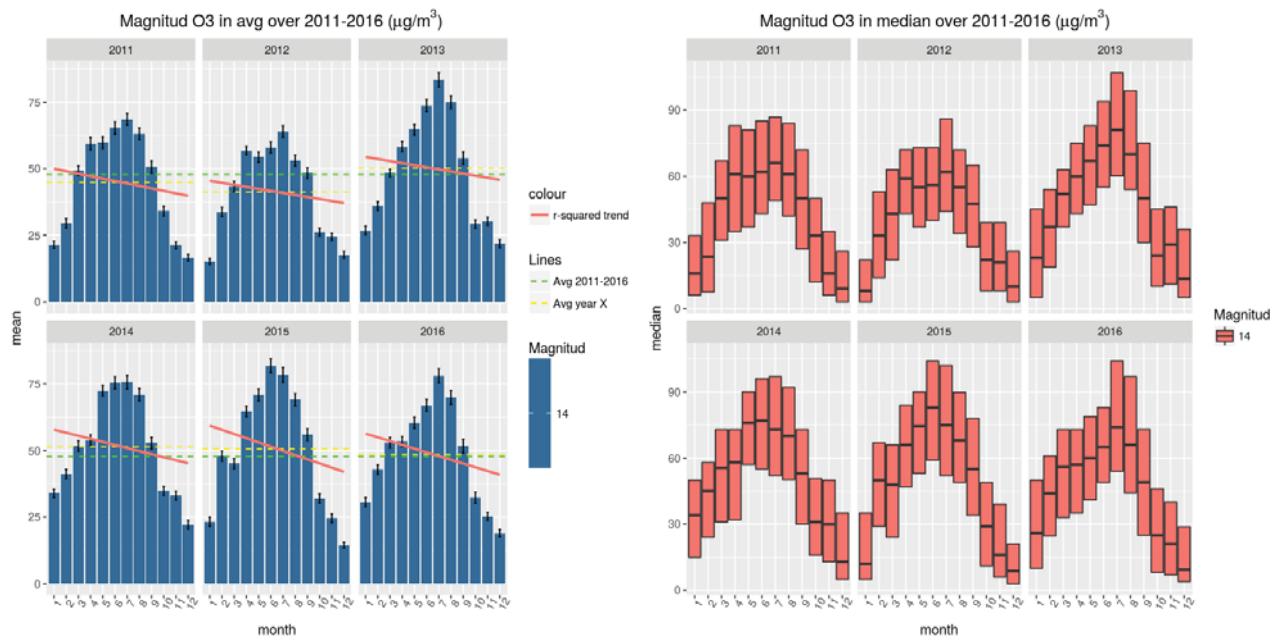


Figura 12.4. Agregado mensual O₃ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 12.4], se puede apreciar una gran estacionalidad en torno a los bajos valores registrados en los meses del periodo invernal, y máximos meses de verano.

En 2016, dichos valores en promedio son menos de la mitad del umbral de información, y el tercer cuartil apenas llega a los $100\mu\text{g}/\text{m}^3$ alrededor de las 16h. El objetivo establecido para los años 2014-2016, es de no superar los $120\mu\text{g}/\text{m}^3$ en media octohoraria, y que no se supere más de 25 días por año.

5.8 Variable Tolueno (C₆H₅-CH₃)

El tolueno, también conocido como metilbenceno, es la principal materia prima a partir de la cual se obtienen otras sustancias y productos derivados del benceno como por ejemplo, el ácido benzoico, el fenol, la sacarina, medicamentos, colorantes, perfumes, TNT y detergentes. Como fuentes de emisión de origen natural destacan el petróleo crudo y en el árbol tolú, aunque también es producido por el humo de los cigarrillos.

La exposición a grandes concentraciones puede provocar afecciones en el sistema nervioso de la persona, y a bajas y moderadas dosis se producen síntomas de cansancio, confusión, debilidad, pérdida de memoria, náuseas, pérdida del apetito y de vista. No obstante, la sintomatología generalmente desaparece cuando no existe exposición a la sustancia.

El tolueno junto con el benceno y el xileno forman parte de los Compuestos Orgánicos Volátiles que actúan como precursores del ozono.

El límite oficial de concentración de tolueno antes de que empiece a ser nocivo para el ser humano es de 500 $\mu\text{g}/\text{m}^3$.

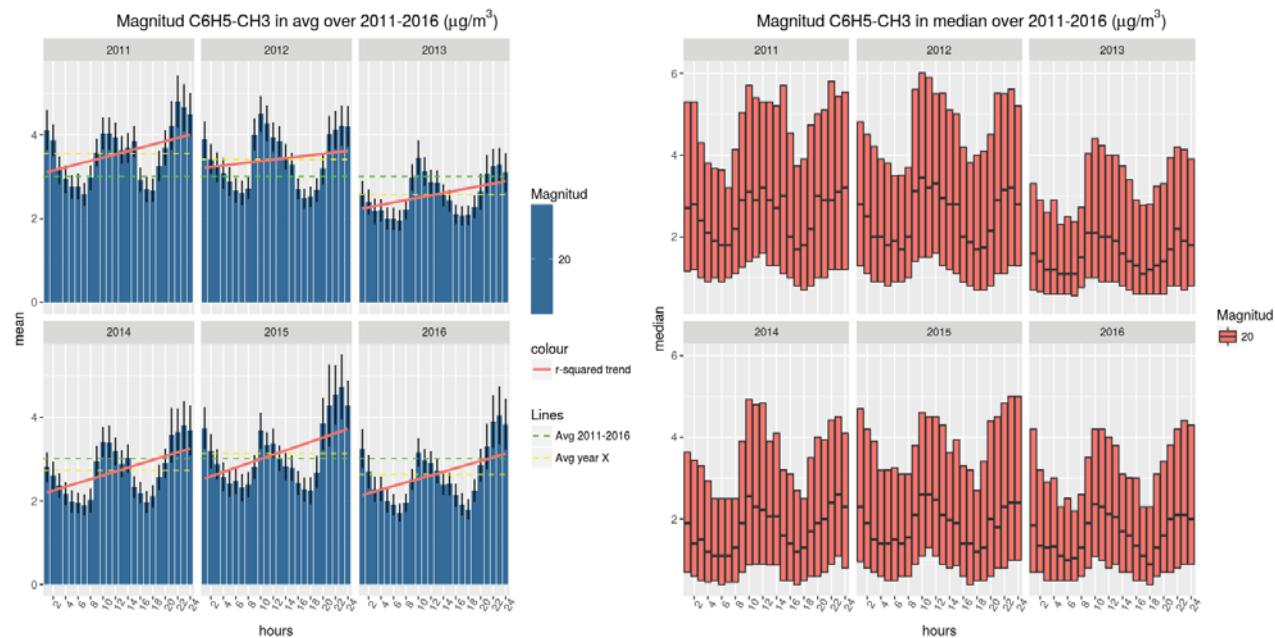


Figura 13.1. Agregado horario C₆H₅-CH₃ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 13.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y las 21h y 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado una hora antes.

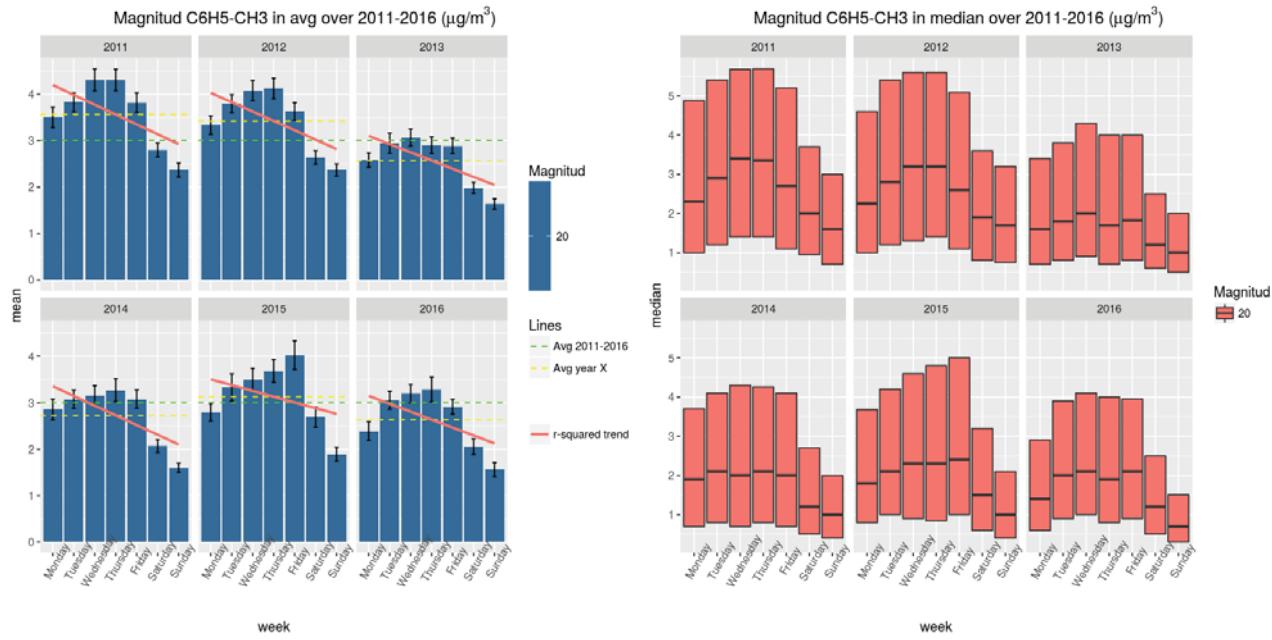


Figura 13.2. Agregado semanal $C_6H_5-CH_3$ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 13.2], se puede apreciar una ligera estacionalidad en los días laborales frente a los sábados y domingos, donde las emisiones son algo más bajas,

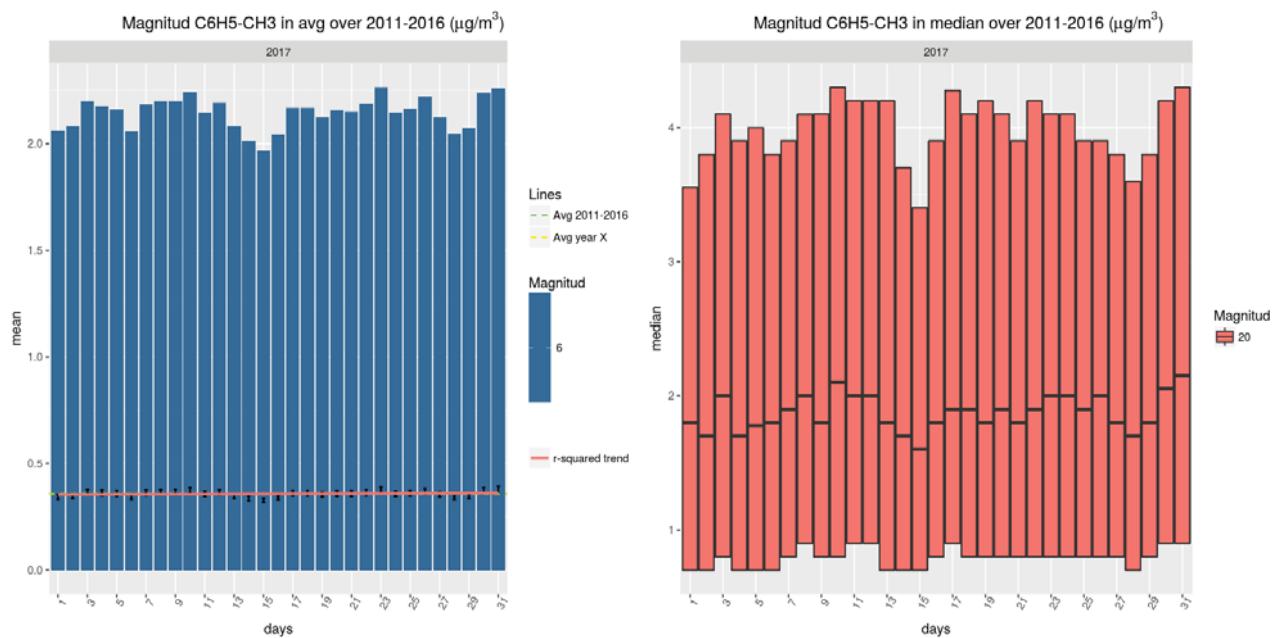


Figura 13.3. Agregado diario $C_6H_5-CH_3$ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 13.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

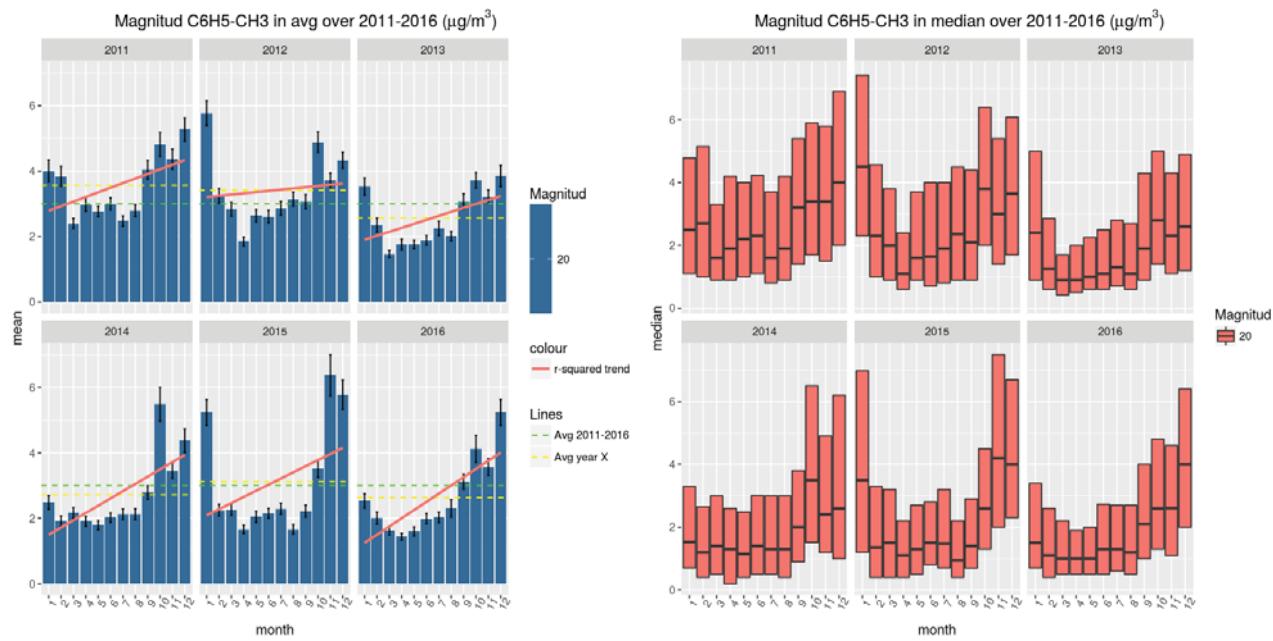


Figura 13.4. Agregado mensual C₆H₅-CH₃ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 13.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional.

Se puede concluir que no los niveles de Tolueno recogidos no representan ningún riesgo para la salud.

5.9 Variable Benceno (C6H6)

El benceno, conocido también como benzol, es un líquido incoloro de olor dulce que se evapora rápidamente. Es perjudicial para la salud debido a su carácter carcinógeno. Las principales fuentes de benceno en el ambiente son los procesos industriales y el tráfico de vehículos en las ciudades.

Los niveles de benceno en el aire pueden aumentar por emisiones generadas por la combustión de carbón y petróleo, operaciones que involucran residuos o almacenaje de benceno, el tubo de escape de automóviles y evaporación de gasolina en estaciones de servicio. El humo de tabaco es otra fuente de benceno en el aire, especialmente en el interior de viviendas.

El benceno junto con el tolueno y el xileno, forma parte de los llamados Compuestos Orgánicos Volátiles (COV). Estos compuestos orgánicos proceden de fuentes antropogénicas o de carácter artificial (tales como procesos industriales, combustión, tratamiento de residuos, transporte y maquinaria móvil) y de fuentes biogénicas o de carácter natural (principalmente vegetación y volcanes), con excepción del metano, capaces de producir oxidantes fotoquímicos por reacción con los óxidos de nitrógeno bajo el efecto de la luz solar, actúan como precursores del ozono.

El valor límite anual para la protección de la salud humana establecido en el R.D. 102/2011 es de $5\mu\text{g}/\text{m}^3$.

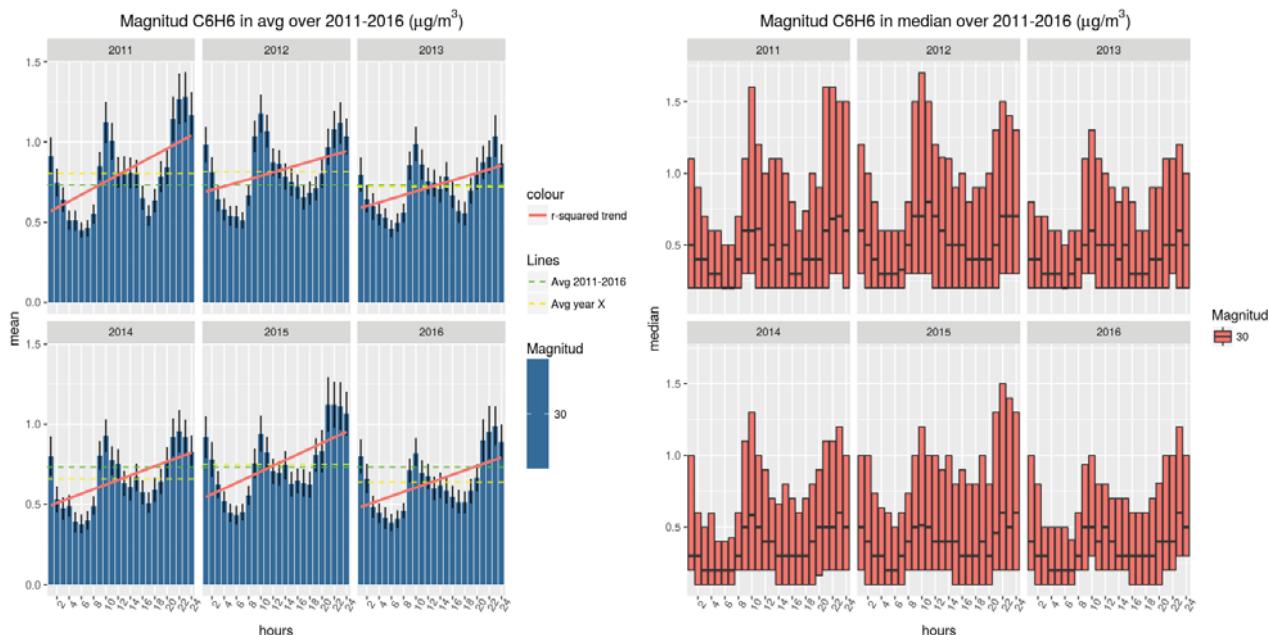


Figura 14.1. Agregado horario C₆H₆ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 14.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y a partir de las 22h.

5.9 Variable Benceno (C6H6)

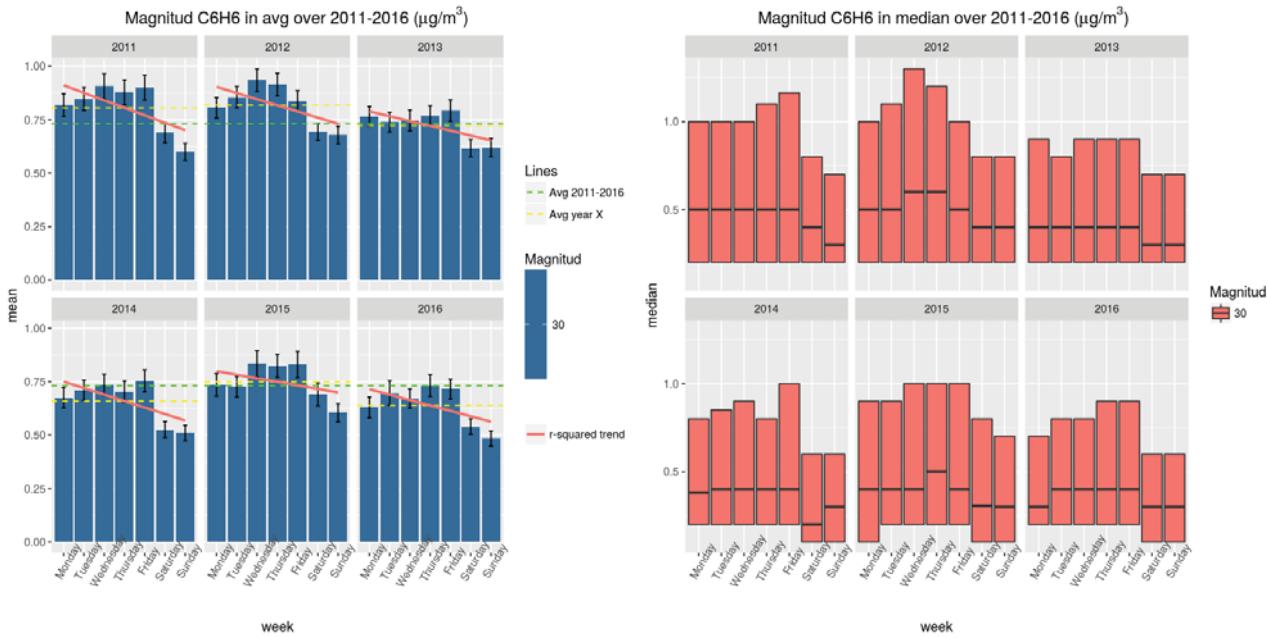


Figura 14.2. Agregado semanal C₆H₆ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 14.2], se puede apreciar una ligera estacionalidad de miércoles a viernes, registrando máximos.

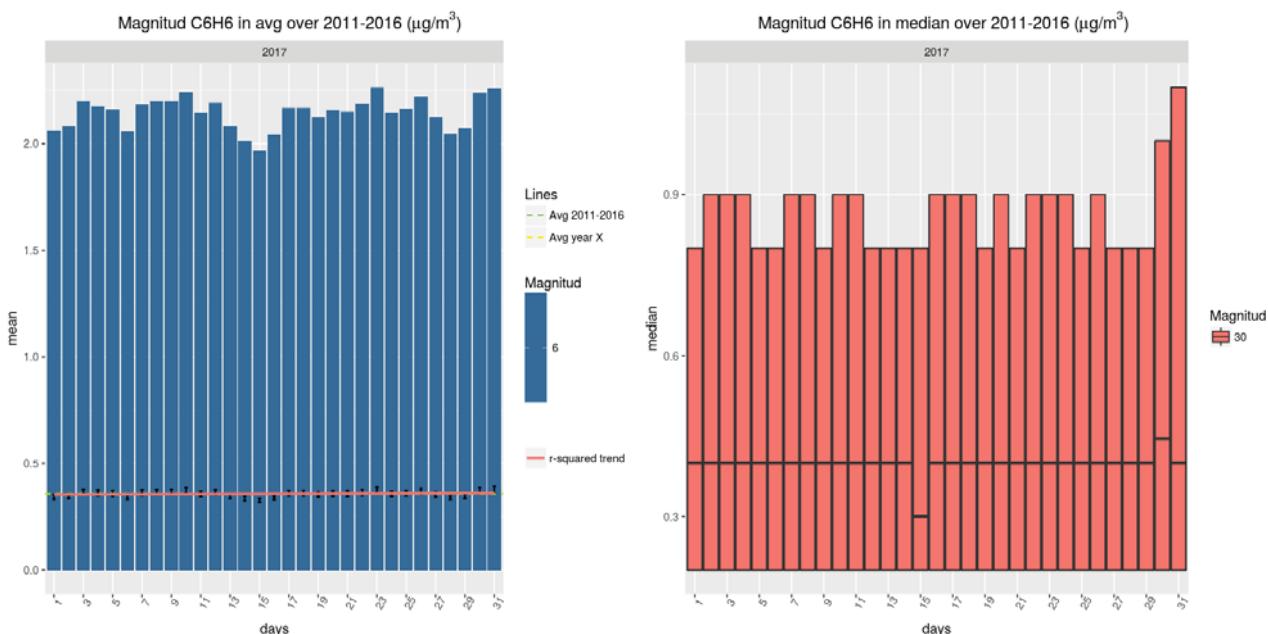


Figura 14.3. Agregado diario C₆H₆ utilizando la media y mediana como estadísticos

A la vista de los agregados diarios [F. 14.3], con mínimos a mediados de mes y máximos el último día del mes. Los valores dentro del rango intercuartílico R2-R3 están mucho más próximos a la mediana a excepción del día 8, 10 y en especial el 30 y 31.

5.10 Variable Etilbenceno (C8H10)

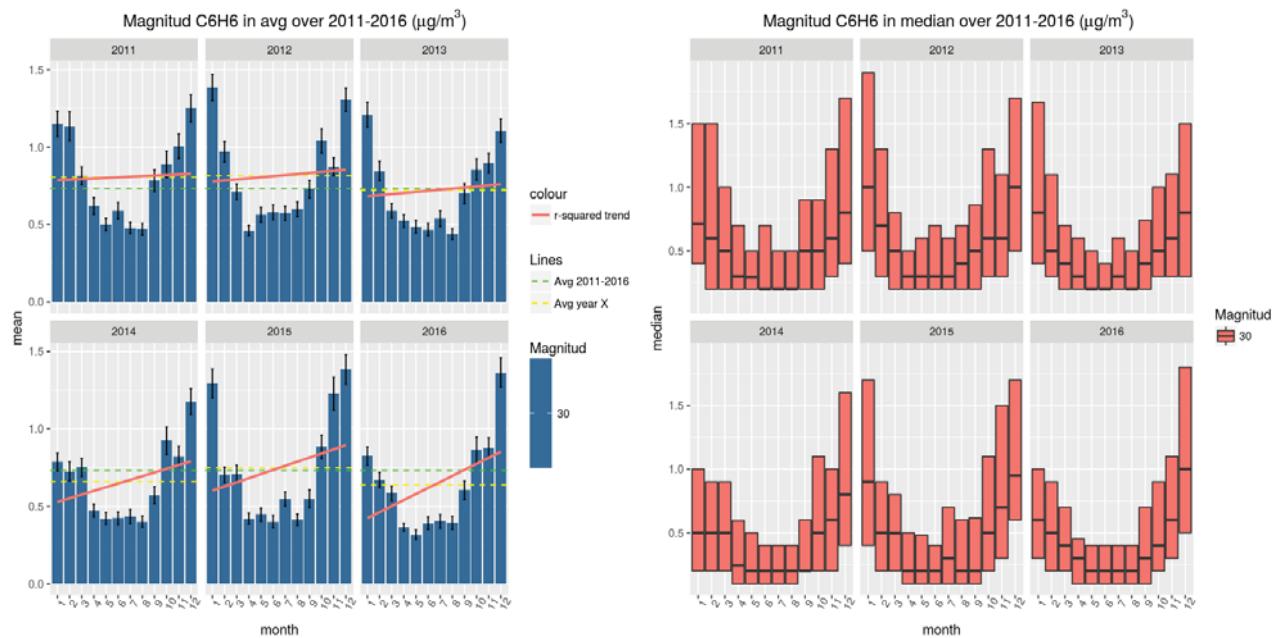


Figura 14.4. Agregado mensual C₆H₆ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 14.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional. Destacar también como la mediana en diciembre está mucho más próxima al cuarto cuartil R4, que a R2 en 2012, 2015 y 2016, por tanto asimétrica hacia la izquierda o negativa en ese mes.

En 2016 como se puede apreciar, los valores recogidos se encuentran bastante lejos del valor límite anual establecido para el Benceno por el R.D. 102/2011 de 5 $\mu\text{g}/\text{m}^3$.

5.10 Variable Etilbenceno (C8H10)

El etilbenceno es un líquido incoloro que se encuentra en numerosos productos entre los que se incluyen la gasolina y pinturas. Respirar niveles muy altos de etilbenceno produce mareo e irritación de los ojos y la garganta. La inhalación de niveles más bajos ha producido defectos de la audición y daño de los riñones en animales.

Su inhalación en ambientes con una concentración superior a 2000 $\mu\text{g}/\text{m}^3$ de este compuesto provoca irritación de ojos y garganta, mareo y fatiga.

La Administración de Salud y Seguridad Ocupacional (OSHA)¹⁶ ha establecido un límite promedio de exposición al etilbenceno en el aire del trabajo de 100mg/ m^3 (100ppm) en media móvil octohoraria, durante una jornada diaria de 8 horas, 40 horas semanales.

5.10 Variable Etilbenceno (C8H10)

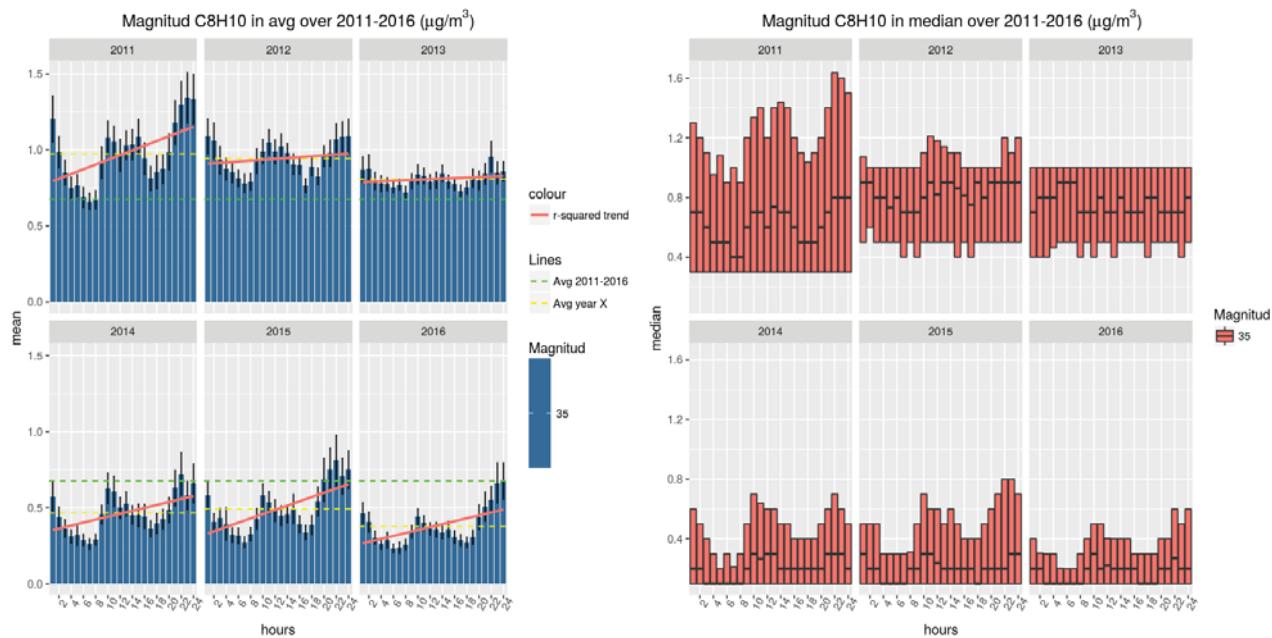


Figura 15.1. Agregado horario C₈H₁₀ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 15.1], se puede apreciar una estacionalidad en entre las 9h y 10h de la mañana, y las 21h y 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado una hora antes. Destacar como se han reducido en gran cantidad las emisiones en los tres últimos años.

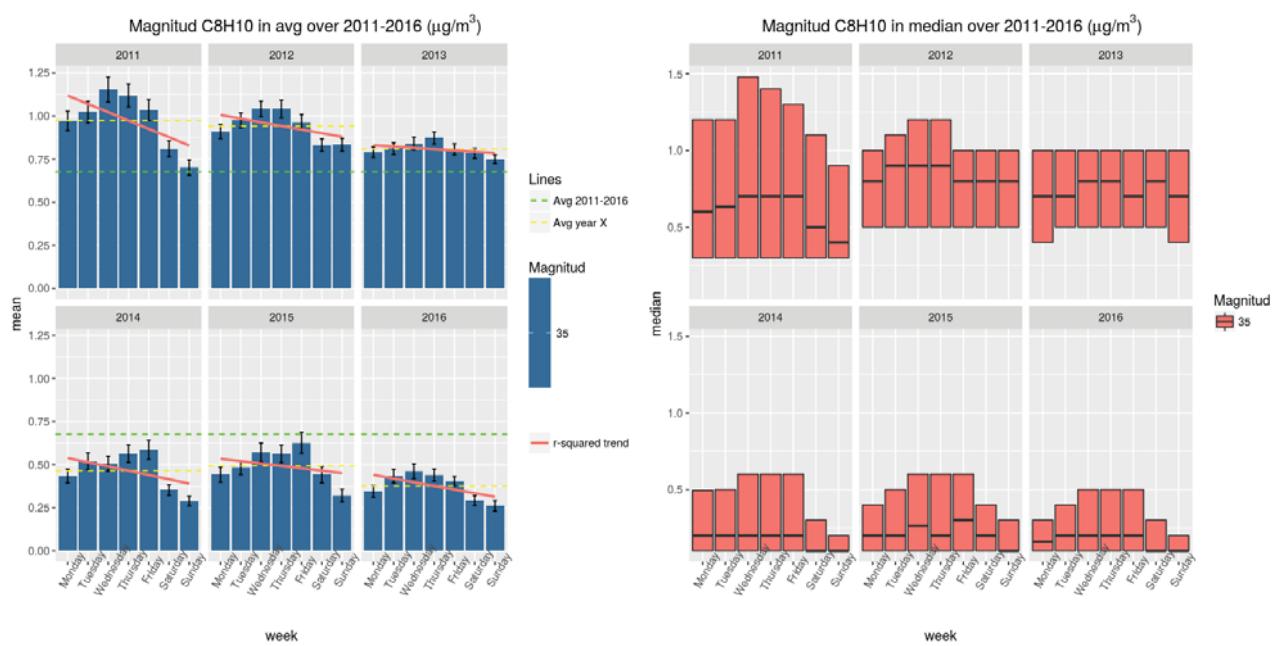


Figura 15.2. Agregado semanal C₈H₁₀ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 15.2], se puede apreciar una ligera estacionalidad en los días laborales frente a los sábados y domingos, donde las emisiones son algo más bajas.

5.10 Variable Etilbenceno (C8H10)

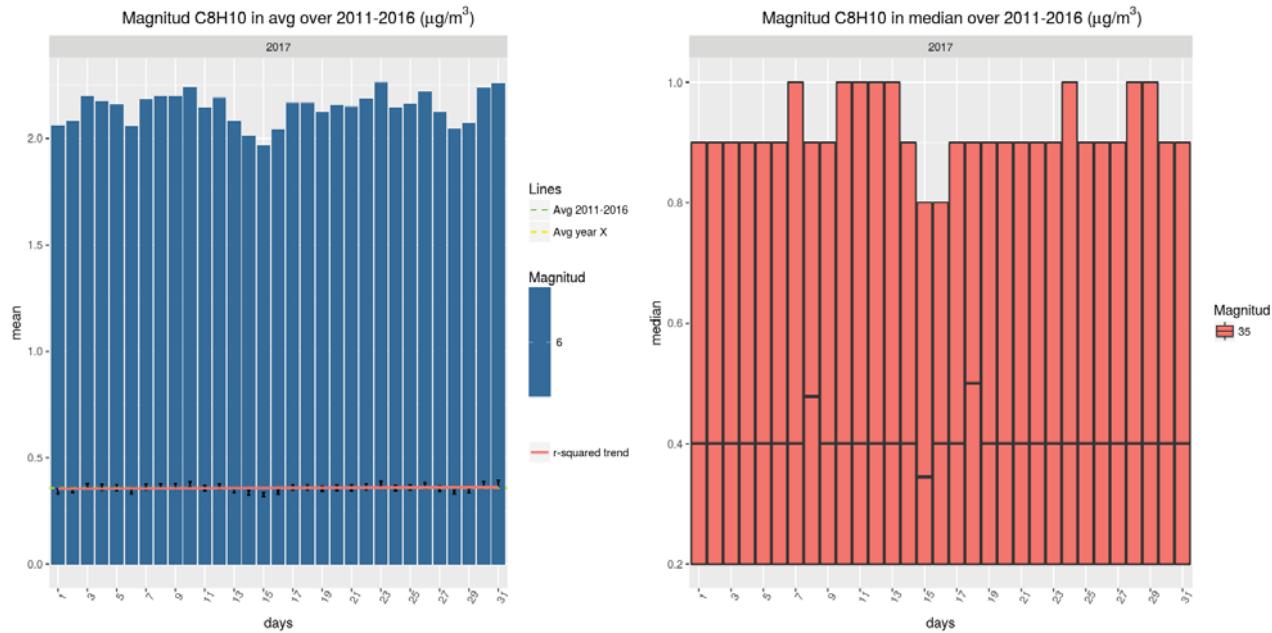


Figura 15.3. Agregado diario C₈H₁₀ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 15.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

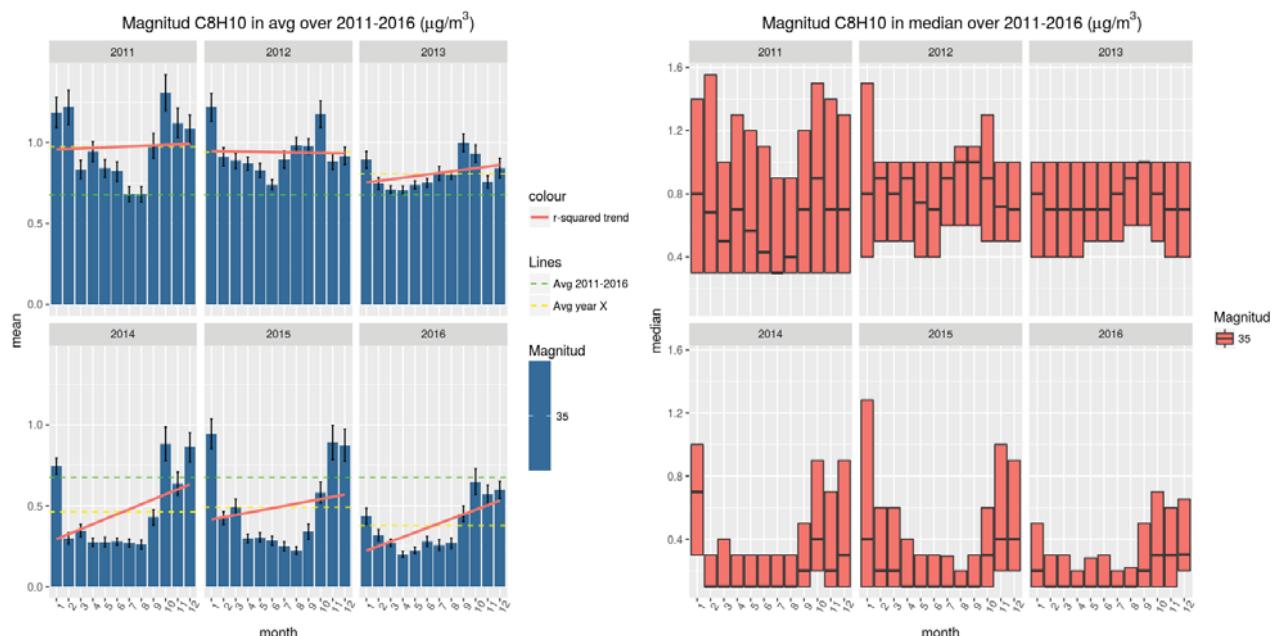


Figura 15.4. Agregado mensual C₈H₁₀ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 15.4], se puede apreciar una estacionalidad en los meses del periodo invernal, donde las emisiones son más altas, probablemente debido a las bajas temperaturas y el tráfico generado especialmente en el periodo vacacional.

Las emisiones de etilbenceno se han reducido drásticamente en los tres últimos años.

Se puede concluir que los valores de etilbenceno recogidos no suponen riesgo alguno para la salud, al estar muy por debajo de la recomendación indicada por la OSHA.

5.11 Variable Hidrocarburos totales (hexeno) (C6H12)

El principal uso del hexeno como un comonómero es en la producción de polietileno, uno de los plásticos más comunes debido a su bajo precio y simplicidad en su fabricación.

Su exposición en ambientes con una concentración superior a $50\text{mg}/\text{m}^3$ (50ppm)¹⁷ en media móvil octohoraria puede causar tos, vértigo, pérdida de conocimiento, somnolencia, dolor de garganta, vómitos, irritación de ojos, piel seca.

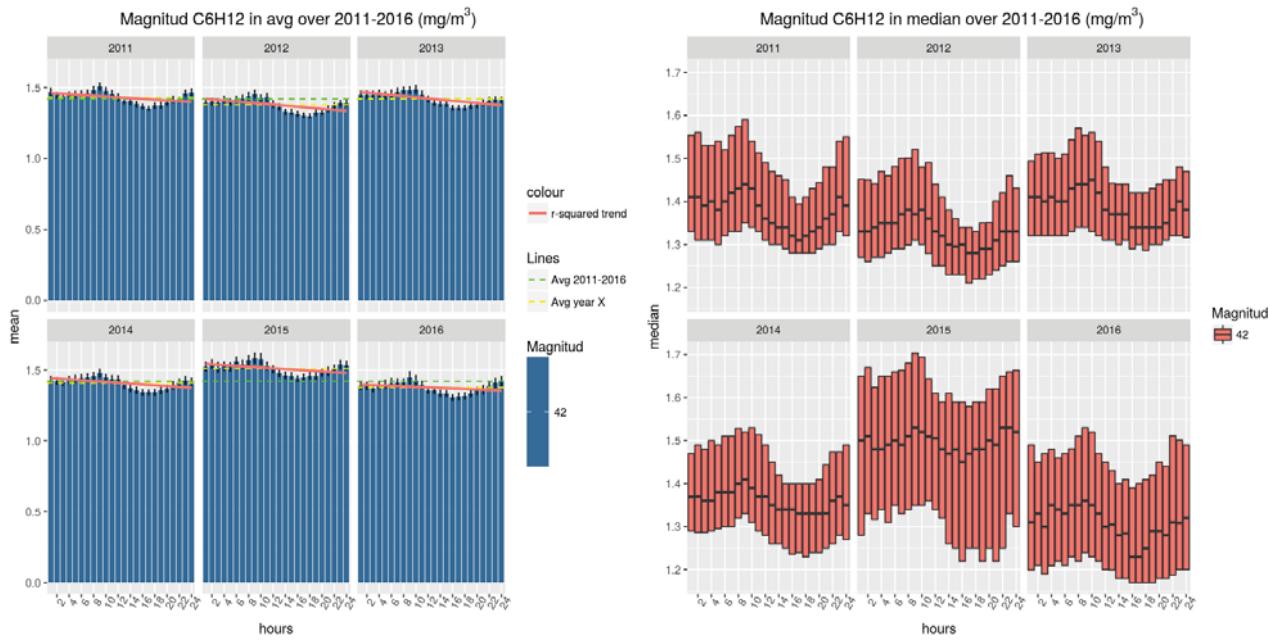


Figura 16.1. Agregado horario C₆H₁₂ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 16.1], se puede apreciar una ligera estacionalidad creciente desde las 6h a las 10h de la mañana y que disminuye hasta las 17h, y vuelve a aumentar hasta 22h, donde las emisiones son más altas, probablemente debido a la jornada laboral y las retenciones de tráfico generado una hora antes.

5.11 Variable Hidrocarburos totales (hexeno) (C6H12)

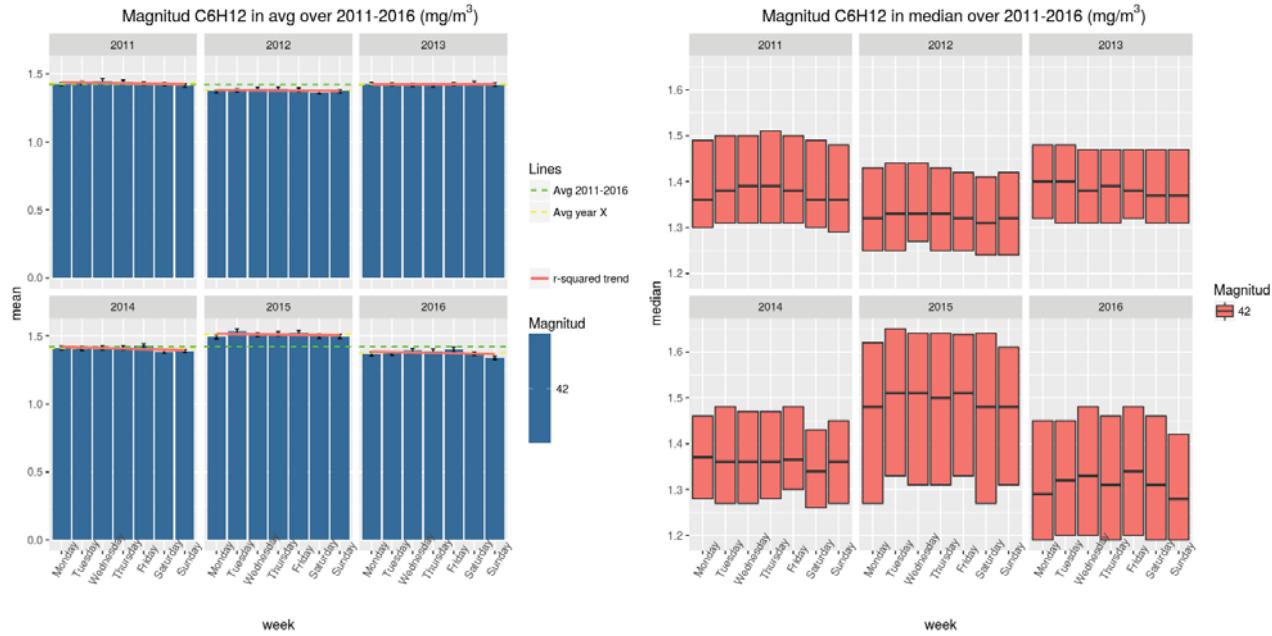


Figura 16.2. Agregado semanal C₆H₁₂ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos de los agregados semanales [F. 16.2], no se aprecia estacionalidad en los de la semana.

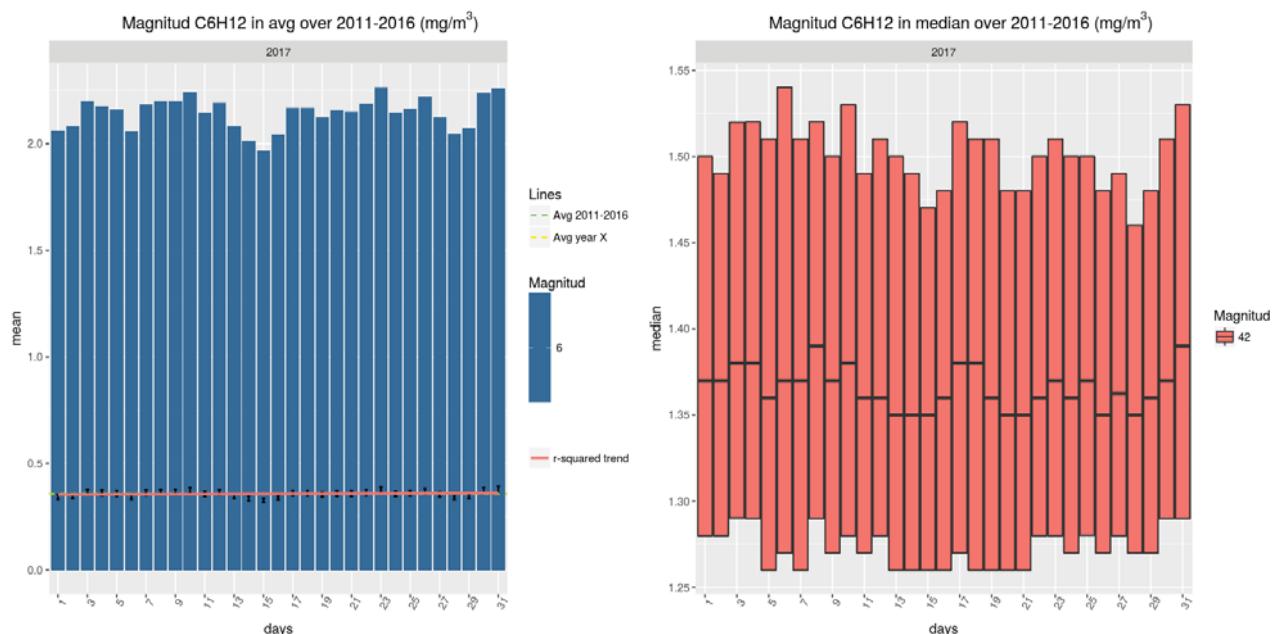


Figura 16.3. Agregado diario C₆H₁₂ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 16.3], al igual que sucede con los niveles de SO₂ existen mínimos a mediados de mes y máximos el último día del mes.

5.11 Variable Hidrocarburos totales (hexeno) (C6H12)

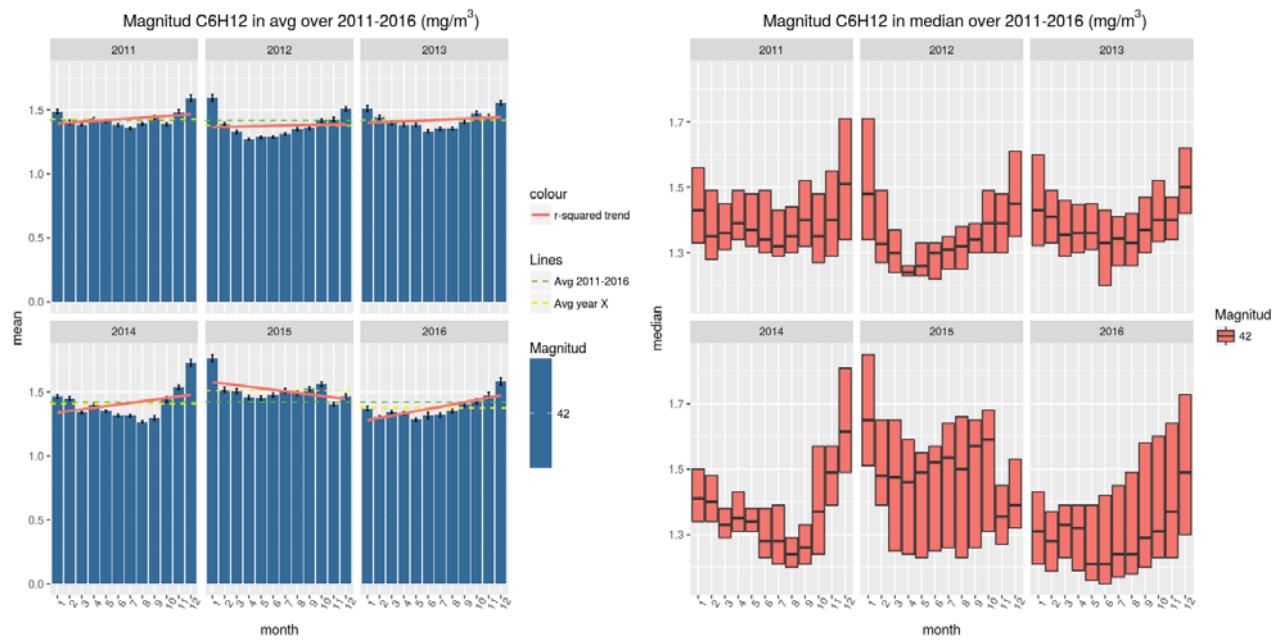


Figura 16.4. Agregado mensual C₆H₁₂ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 16.4], se puede apreciar una ligera estacionalidad en el último mes de año, donde las emisiones tienden a ser más altas.

Se puede concluir que no hay riesgo alguno para la salud según los límites indicados por la OSHA.

5.12 Variable Hidrocarburos totales no metálicos (hexano) (C6H14)

El hexano es una sustancia química elaborada a partir del petróleo crudo. Es utilizado como disolvente para algunas pinturas y procesos químicos y para quitar etiquetas de precios ya que disuelve el pegamento con que se adhieren.

El hexano puede causar efectos sobre: el sistema nervioso, irritación en las vías respiratorias y en los ojos, producir vértigo, somnolencia, dolor de cabeza, dificultad al respirar, pérdida de conocimiento, náuseas, fatiga.

La American Conference of Governmental Industrial Hygienists (ACGIH)¹⁸ recomienda un valor umbral de exposición (TLV, por sus siglas en inglés) de 50mg/m³ (50ppm) en media móvil octohoraria.

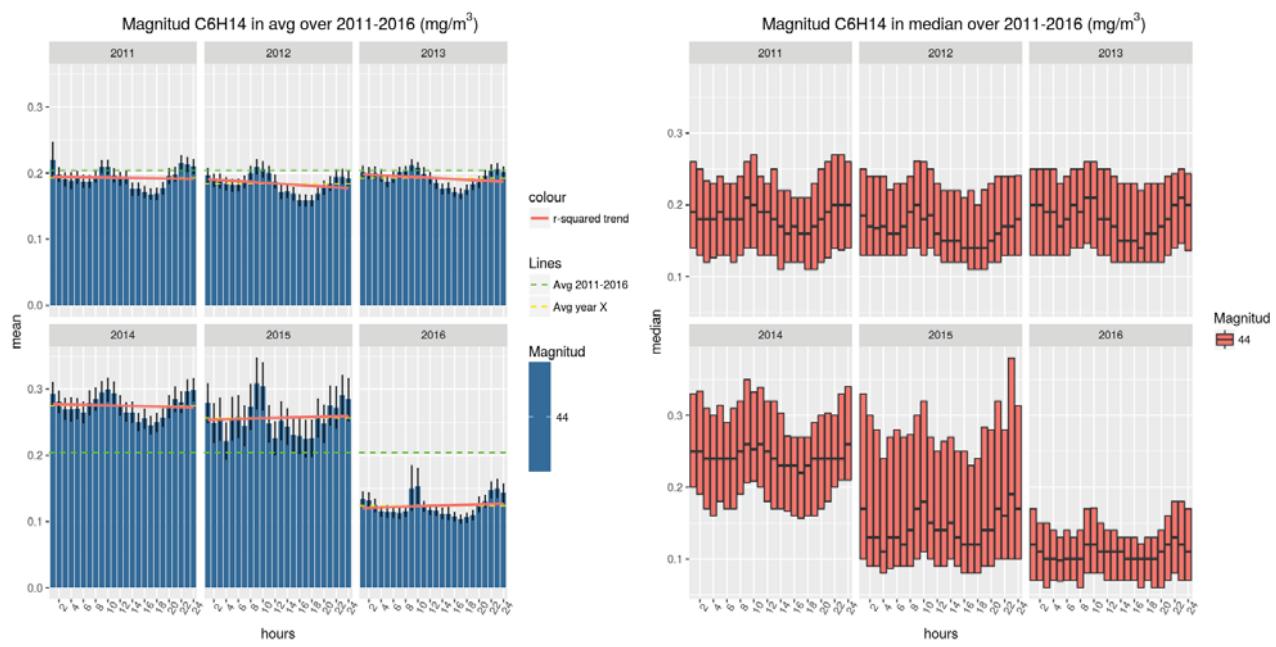


Figura 17.1. Agregado horario C₆H₁₄ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 17.1], se puede apreciar una ligera estacionalidad en las horas centrales y finales del día salvo en el año 2015. En 2015 se aprecia como el tercer cuartil está bastante lejos de la mediana, por tanto con una fuerte asimetría positiva.

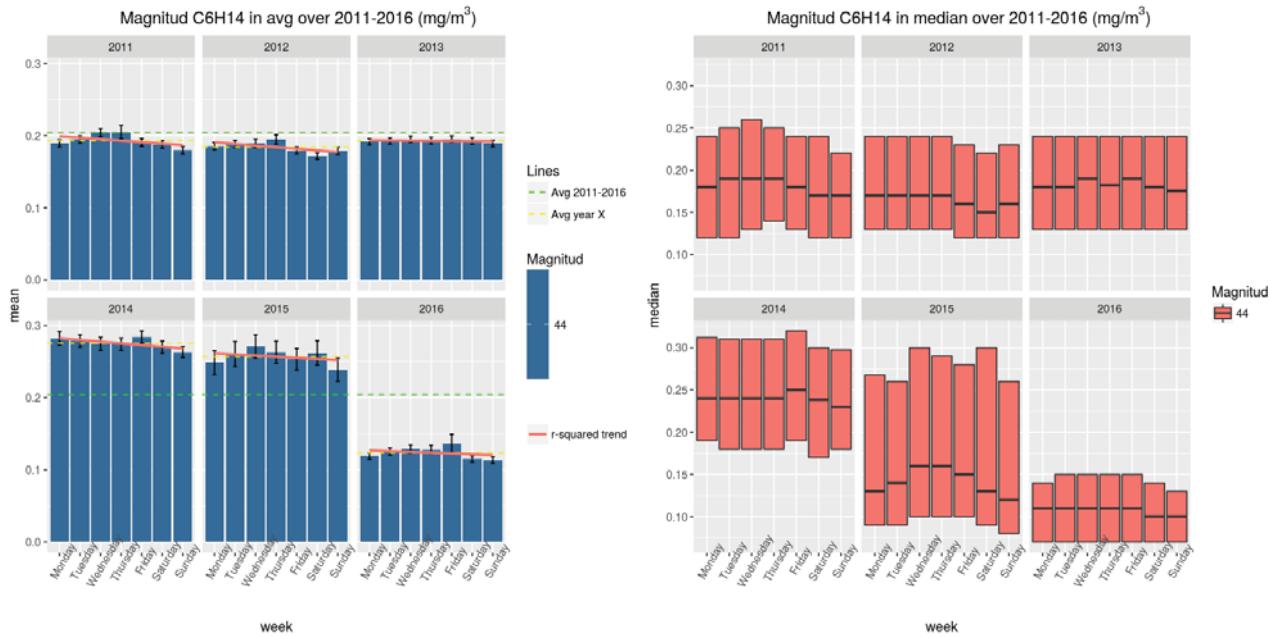


Figura 17.2. Agregado semanal C₆H₁₄ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 17.2], no se puede apreciar estacionalidad alguna. Cabe destacar en 2015 como el tercer cuartil está bastante lejos de la mediana, por tanto con una fuerte asimetría positiva.

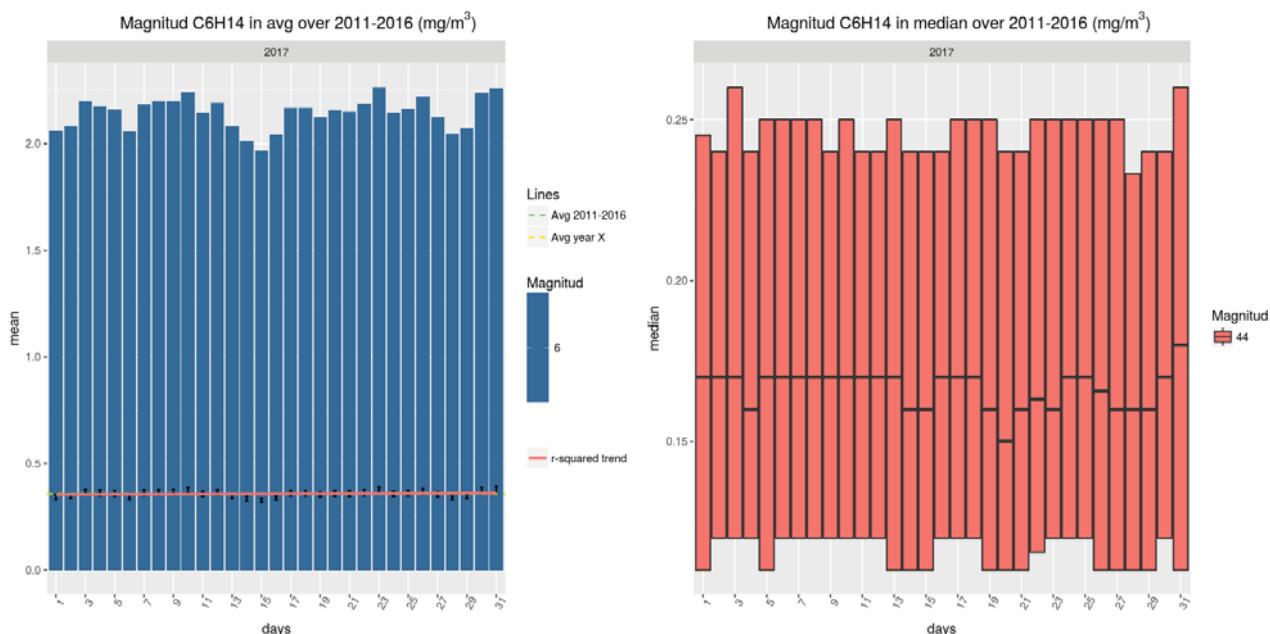


Figura 17.3. Agregado diario C₆H₁₄ utilizando la media y mediana como estadísticos

A la vista de ambos estadísticos [F. 17.3], al igual que sucede con los niveles de SO₂ se aprecia un mínimo a mediados de mes y un máximo el último día del mes.

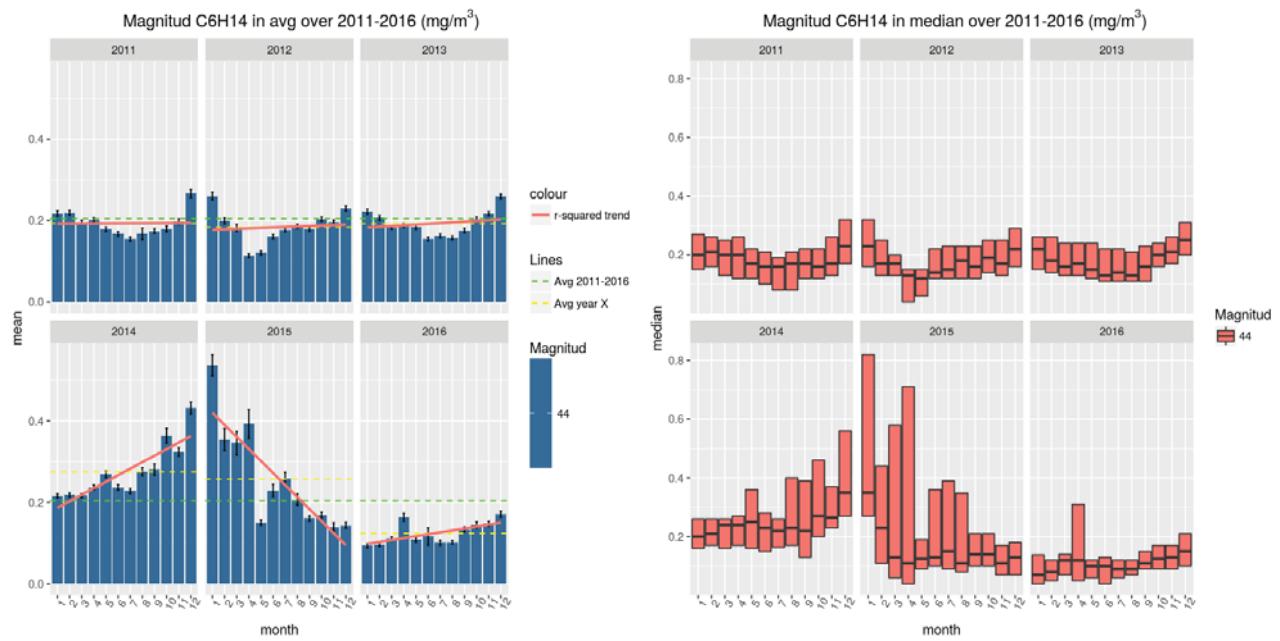


Figura 17.4. Agregado mensual C₆H₁₄ utilizando la media y mediana como estadísticos

En ambos agregados mensuales [F. 17.4], no se puede apreciar estacionalidad alguna.

Se puede concluir que los valores recogidos de hidrocarburos totales no metálicos no suponen riesgo alguno para la salud según la recomendación de la ACGIH.

5.13 Variables atmosféricas

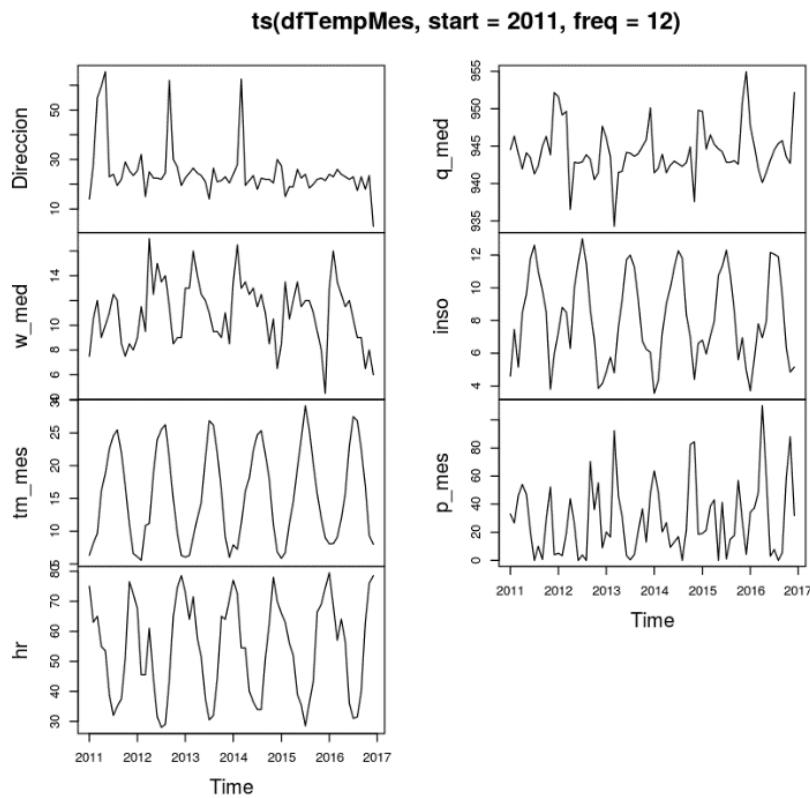


Figura 18. Plots de las variables atmosféricas

Como se puede apreciar [F. 18], todas son estacionarias a excepción de la dirección del viento. Se puede apreciar como:

- La variable “tm_mes”, que hace referencia a la temperatura media del mes, toma valores bajos a principios y fin de año, y alcanza máximos a mediados de año, además de como con el paso de los años los mínimos son algo menos mínimos. Está fuertemente correlacionada positivamente con la variable “inso” que mide la insolación mensual diaria. Y a su vez correlacionada negativamente con la variable “hr” por su parte, que alcanza máximos en invierno y mínimos en verano, y que se refiere a los valores de humedad relativa.
- La variable “p_mes” mide la precipitación media mensual, tiende a tener máximos en primavera y a finales de año, a excepción de 2014.
- La variable “w_med”, que mide la velocidad del viento mensual, alcanza máximos un par de meses antes de alcanzarse máximos en “tm_mes”.
- La variable “q_med”, que mide la presión atmosférica, por su parte alcanza máximos en otoño, y se puede apreciar fácilmente como está negativamente correlacionada con la variable “inso”.

6. Análisis de dependencias

Para determinar la dependencia existente entre las variables se ha realizado un análisis de correlación.

El análisis de correlación es una herramienta estadística que permite medir y describir el grado o intensidad de asociación entre dos variables.

Es la herramienta estadística que podemos usar para describir el grado hasta el cual una variable está linealmente relacionada con otra. Existen dos medidas para definir la correlación entre dos variables:

- El coeficiente de determinación.
- El coeficiente de correlación.

Para la realización del análisis se ha utilizado el coeficiente de correlación de Pearson, es un índice que mide la relación lineal entre dos variables aleatorias X e Y cuantitativas.

El valor del índice de correlación varía entre -1 y +1:

- Si $r_{XY} = 0$, no existe relación lineal, pero pueden existir relaciones no lineales entre las dos variables.
- Si $r_{XY} = 1$, existe una correlación positiva perfecta. Existe una dependencia directa total entre las dos variables, cuando una de ellas aumenta, la otra también lo hace en proporción constante.
- Si $r_{XY} = -1$, existe una correlación negativa perfecta. Existe una dependencia inversa total, cuando una de ellas aumenta, la otra disminuye en proporción constante.
- Si $0 < r_{XY} < 1$, existe una correlación positiva.
- Si $-1 < r_{XY} < 0$, existe una correlación negativa.

6.1 Dependencias de variables en el mismo instante de tiempo

La siguiente tabla muestra las correlaciones obtenidas de los valores horarios de cada magnitud:

	Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	1	0.26799508	0.34069837	0.32431704	0.21328039	0.10522926	-0.2321822	0.23448248	0.24202269	0.12658038	0.12437425	0.02859319
Magnitud6	0.26799508	1	0.51520518	0.41170521	0.31763579	0.22253905	-0.2480546	0.38456582	0.35894517	0.3071535	0.17382773	0.04990747
Magnitud7	0.34069837	0.51520518	1	0.53234066	0.41429464	0.26707886	-0.393477	0.47424208	0.44017064	0.3443705	0.2486493	0.06461634
Magnitud8	0.32431704	0.41170521	0.53234066	1	0.46470319	0.33185535	-0.5053096	0.4873395	0.42558729	0.32319978	0.19978052	0.06731386
Magnitud9	0.21328039	0.31763579	0.41429464	0.46470319	1	0.64613944	-0.24093	0.40593497	0.34857412	0.28171113	0.17195503	0.02373504
Magnitud10	0.10522926	0.22253905	0.26707886	0.33185535	0.64613944	1	-0.0941589	0.29870589	0.22430842	0.21438854	0.09592161	0.01003519
Magnitud14	-0.23218217	-0.24805457	-0.39347702	-0.50530962	-0.24092996	-0.0941589	1	-0.34189481	-0.29635269	-0.19899679	-0.19416413	-0.0424081
Magnitud20	0.23448248	0.38456582	0.47424208	0.4873395	0.40593497	0.29870589	-0.3418948	1	0.41427666	0.37530376	0.1968997	0.04027804
Magnitud30	0.24202269	0.35894517	0.44017064	0.42558729	0.34857412	0.22430842	-0.2963527	0.41427666	1	0.2948892	0.1799827	0.03339406
Magnitud35	0.12658038	0.3071535	0.3443705	0.32319978	0.28171113	0.21438854	-0.1989968	0.37530376	0.2948892	1	0.13349399	0.02451349
Magnitud42	0.12437425	0.17382773	0.2486493	0.19978052	0.17195503	0.09592161	-0.1941641	0.1968997	0.1799827	0.13349399	1	0.09340151
Magnitud44	0.02859319	0.04990747	0.06461634	0.06731386	0.02373504	0.01003519	-0.0424081	0.04027804	0.03339406	0.02451349	0.09340151	1

Tabla 12. Matriz de correlación de variables contaminantes en el mismo instante

Las correlaciones más altas son entre las mismas magnitudes, aunque cabe destacar como por ejemplo, la Magnitud9 (PM2.5) y la Magnitud10 (PM10) tienen un alto grado de correlación con un 64%. En contraste la Magnitud10 (PM10) y la Magnitud44 (Hexano) tienen una correlación cercana al 1%.

6.2 Dependencias de variables en instantes de tiempo anteriores

Para ver la correlación entre las diferentes variables contaminantes y las variables atmosféricas debemos trabajar con datos medios mensuales, obteniendo la siguiente tabla mediante el coeficiente de correlación de Pearson:

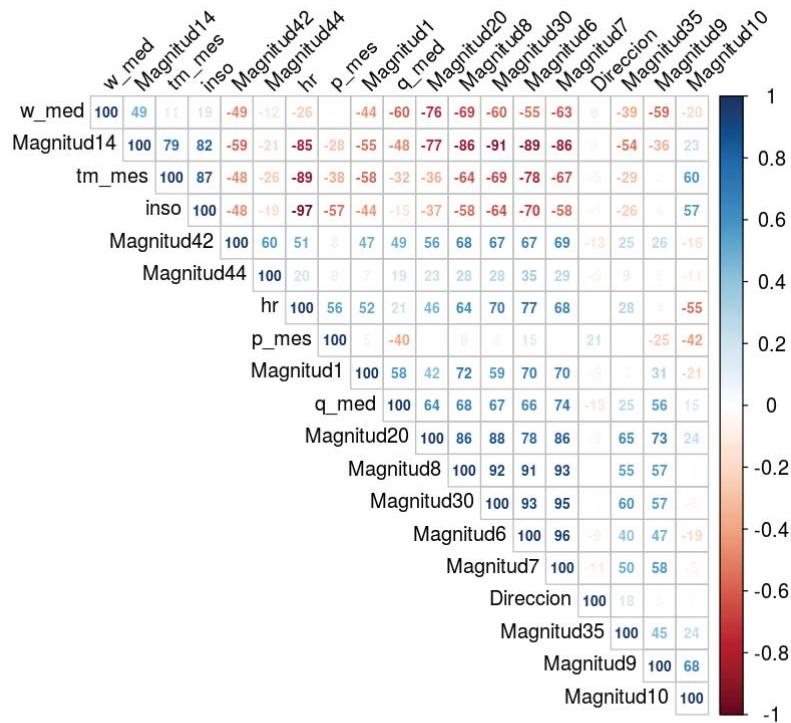


Figura 19. Matriz de correlación mensual mediante el coeficiente de correlación de Pearson

Ahora se aprecia una mayor correlación entre gases. Por ejemplo, la variable Magnitud30 (Benceno) con la Magnitud6 (CO) y Magnitud7(NO) alcanzando más de un 90% de correlación. O la Magnitud14 (Ozono) con el Benceno, el NO₂, el CO y el NO₂, alcanzando valores cercanos al 90% de correlación negativa en este caso. También cabe destacar como tm_mes (temperatura media mensual) está fuertemente relacionada con inso (media mensual de la insolación diaria) y con la Magnitud6 (CO).

6.2 Dependencias de variables en instantes de tiempo anteriores

También nos puede ser de utilidad ver la correlación de los retrasos temporales entre las variables:

6.2 Dependencias de variables en instantes de tiempo anteriores

- Correlaciones a 1 hora de diferencia entre la variable X e Y:

	Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.38266777	0.23868748	0.3105013	0.29625823	0.20692645	0.096469367	-0.21629375	0.2275329	0.22862923	0.12598946	0.12227317	0.028925744
Magnitud6	0.25689145	0.43141794	0.46526437	0.37733787	0.31889895	0.202418308	-0.23942612	0.391908	0.35960185	0.28781899	0.1668816	0.040135596
Magnitud7	0.33559447	0.46658369	0.59085958	0.48421084	0.42030652	0.254045827	-0.37279268	0.4781589	0.44055479	0.33873651	0.2465864	0.051285069
Magnitud8	0.32094408	0.3918481	0.50609282	0.6586989	0.4678306	0.323894013	-0.48708853	0.504501	0.43813461	0.33450093	0.20765973	0.062395609
Magnitud9	0.19703935	0.27372599	0.3676535	0.41526235	0.69863368	0.585722662	-0.22163481	0.3865371	0.32963218	0.27893418	0.16518646	0.01480323
Magnitud10	0.10000122	0.20373652	0.23947934	0.29931774	0.60462482	0.697255195	-0.07718082	0.2911497	0.22461084	0.22538636	0.09352608	0.000808633
Magnitud14	-0.23675243	-0.24080346	-0.38747774	-0.46013768	-0.23497736	-0.086086255	0.84172801	-0.349671	-0.2998705	-0.20602367	-0.19832945	-0.041443348
Magnitud20	0.20928539	0.33417428	0.42567634	0.41858575	0.37979917	0.265411967	-0.31287453	0.4898131	0.37772287	0.34617743	0.18341368	0.036127443
Magnitud30	0.22309487	0.31039292	0.38486165	0.36881445	0.32790321	0.19820158	-0.27775104	0.3722889	0.35339907	0.27817882	0.16921625	0.040101456
Magnitud35	0.10498761	0.25589834	0.29274136	0.28037298	0.25392811	0.189348802	-0.18508119	0.3459825	0.27403333	0.36265936	0.11604772	0.013777306
Magnitud42	0.11938375	0.16655972	0.23425529	0.18066425	0.16347125	0.090472939	-0.18392757	0.1918786	0.17187174	0.12716735	0.36654982	0.098319783
Magnitud44	0.03672731	0.04852083	0.05055848	0.05813178	0.02240035	0.003430817	-0.04300413	0.0410109	0.03985614	0.02154978	0.08435991	0.243062313

Tabla 13.1. Matriz de correlación de variables contaminantes 1 hora antes

- Correlaciones a 4 horas de diferencia entre la variable X e Y:

	Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.31039658	0.13100719	0.19756724	0.20103582	0.137723286	0.034610353	-0.16880985	0.1452357	0.14627611	0.068718247	0.09615128	0.0207248116
Magnitud6	0.14234034	0.12438797	0.19362485	0.16967355	0.170118846	0.078449198	-0.19127365	0.20784164	0.1738639	0.14119529	0.12698553	0.0137099571
Magnitud7	0.21756877	0.17651045	0.30552118	0.25588545	0.253421197	0.11183623	-0.27211447	0.29865467	0.25032917	0.184481095	0.18111022	0.0297658529
Magnitud8	0.21651147	0.18956518	0.2780526	0.31884123	0.278660733	0.160144996	-0.31503583	0.3301731	0.25759603	0.206606515	0.16229613	0.0357584201
Magnitud9	0.11505324	0.12196104	0.20099561	0.21133141	0.465432362	0.392353835	-0.1368139	0.25822636	0.18947737	0.169374483	0.12750491	0.0021119487
Magnitud10	0.03152871	0.07415644	0.09918847	0.13140293	0.424600343	0.502726943	-0.00130504	0.19010573	0.11232489	0.143590027	0.06247435	-0.0060713689
Magnitud14	-0.19955452	-0.14735155	-0.28316988	-0.21588684	-0.126606553	0.014321433	0.57110161	-0.24846001	-0.20866462	-0.140020501	-0.16195712	-0.026104948
Magnitud20	0.1292272	0.14440466	0.24255303	0.22689331	0.235941392	0.15159357	-0.23236189	0.29488629	0.21361637	0.1817369	0.13474046	0.0184486223
Magnitud30	0.13356619	0.11104059	0.20017167	0.195477	0.189435738	0.085633459	-0.21769639	0.22415991	0.19433358	0.144211266	0.1253433	0.0156671226
Magnitud35	0.04401958	0.08445524	0.14893654	0.13254977	0.148596632	0.101485095	-0.14520749	0.20021676	0.13472123	0.201967021	0.0805086	-0.0002961624
Magnitud42	0.09371032	0.09321664	0.16865348	0.11507512	0.118674334	0.055427296	-0.14530085	0.12804288	0.12411121	0.066738927	0.21365889	0.0552040013
Magnitud44	0.02737692	0.02025329	0.0387775	0.03596223	0.009664081	-0.004900489	-0.02749444	0.02374167	0.02919066	0.004063364	0.07568736	0.2181958849

Tabla 13.2. Matriz de correlación de variables contaminantes 4 horas antes

- Correlaciones a 8 horas de diferencia entre la variable X e Y:

	Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.275032646	0.14031724	0.19657057	0.18911147	0.1109585384	0.018879213	-0.1848925	0.11722309	0.12281599	0.0407092875	0.0969826	0.032469415
Magnitud6	0.080984724	0.06978937	0.12965061	0.0978998	0.0790725299	0.020882159	-0.18264268	0.10229271	0.07885133	0.0430307186	0.08472559	0.0183253
Magnitud7	0.157409705	0.12520804	0.23020589	0.18507109	0.1603191446	0.055193794	-0.25407315	0.19117215	0.15717075	0.0998515641	0.13987559	0.021719211
Magnitud8	0.151475848	0.11678476	0.21592733	0.20881236	0.1829847311	0.093413749	-0.27392976	0.19579777	0.14791143	0.0993089669	0.12517272	0.035414759
Magnitud9	0.084023307	0.10668108	0.17630846	0.17578138	0.3639175344	0.322090498	-0.15782796	0.18634039	0.12962645	0.1045780408	0.10405974	-0.000140482
Magnitud10	-0.001952709	0.04806101	0.06903914	0.08447801	0.3289610448	0.413992748	-0.02258131	0.11958085	0.05089438	0.0819440593	0.04993501	-0.008628054
Magnitud14	-0.179012991	-0.11935809	-0.24337306	-0.1509014	-0.0897265077	0.02937456	0.31948705	-0.16911363	-0.17091741	-0.0938959256	-0.11540868	-0.01529122
Magnitud20	0.097811113	0.12526532	0.20892135	0.19725312	0.1968007873	0.137793147	-0.22712361	0.22390226	0.14999476	0.1246766274	0.10851954	0.00979172
Magnitud30	0.101701555	0.10130959	0.18296217	0.17052805	0.1435053297	0.065980513	-0.24232483	0.15501189	0.13875736	0.0865693232	0.10792687	0.023646627
Magnitud35	0.021684174	0.06231369	0.11485319	0.10734187	0.1119918522	0.081451197	-0.15073684	0.13534183	0.08303931	0.1378919267	0.05372186	-0.00176085
Magnitud42	0.077179743	0.06993123	0.12495206	0.08849113	0.0910075533	0.03176541	-0.1106836	0.09446	0.0946966	0.0450689473	0.13695559	0.037445135
Magnitud44	0.023025453	0.01158998	0.02239622	0.01978212	0.0002516885	-0.007892065	-0.01658434	0.01948475	0.02019388	0.0003872223	0.04809412	0.197916711

Tabla 13.3. Matriz de correlación de variables contaminantes 8 horas antes

En cuanto a las dependencias en instantes de tiempo anteriores [T. 13.1, 13.2, 13.3], se puede concluir como a medida que se atrasa el tiempo menos grado de correlación existe entre todas las variables, y donde los grados de correlación más grandes son normalmente los de la misma variable atrasada en el tiempo.

6.3 Dependencias de variables en instantes de tiempo posteriores

6.3 Dependencias de variables en instantes de tiempo posteriores

- Correlaciones a 1 hora de diferencia entre la variable Y e X:

Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.38266777	0.2568915	0.33559447	0.32094408	0.19703935	0.100001219	-0.23675243	0.20928539	0.22309487	0.10498761	0.11938375
Magnitud6	0.23868748	0.4314179	0.46658369	0.3918481	0.27372599	0.203736522	-0.24080346	0.33417428	0.31039292	0.25589834	0.16655972
Magnitud7	0.3105013	0.4652644	0.59085958	0.50609282	0.3676535	0.239479339	-0.38747774	0.42567634	0.38486165	0.29274136	0.23425529
Magnitud8	0.29625823	0.3773379	0.48421084	0.6586989	0.41526235	0.299317739	-0.46013768	0.41858575	0.36881445	0.28037298	0.18066425
Magnitud9	0.20692645	0.3188989	0.42030652	0.4678306	0.69863368	0.604624818	-0.23497736	0.37979917	0.32790321	0.25392811	0.16347125
Magnitud10	0.09646937	0.2024183	0.25404583	0.32389401	0.58572266	0.697255195	-0.08608625	0.26541197	0.19820158	0.1893488	0.09047294
Magnitud14	-0.21629375	-0.2394261	-0.37279268	-0.48708853	-0.22163481	-0.077180819	0.84172801	-0.31287453	-0.27775104	-0.18508119	-0.18392757
Magnitud20	0.2275329	0.391908	0.47815892	0.50450103	0.38653711	0.291149664	-0.34967102	0.48981314	0.37228892	0.34598252	0.19187858
Magnitud30	0.22862923	0.3596019	0.44055479	0.43813461	0.32963218	0.224610843	-0.2998705	0.37772287	0.35339907	0.27403333	0.17187174
Magnitud35	0.12598946	0.287819	0.33873651	0.33450093	0.27893418	0.225386355	-0.20602367	0.34617743	0.27817882	0.36265936	0.12716735
Magnitud42	0.12227317	0.1668816	0.2465864	0.20765973	0.16518646	0.093526081	-0.19832945	0.18341368	0.16921625	0.11604772	0.36654982
Magnitud44	0.02892574	0.0401356	0.05128507	0.06239561	0.01480323	0.000808633	-0.04144335	0.03612744	0.04010146	0.01377731	0.09831978

Tabla 14.1. Matriz de correlación de variables contaminantes 1 hora después

- Correlaciones a 4 horas de diferencia entre la variable Y e X:

Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.31039658	0.14234034	0.21756877	0.21651147	0.115053242	0.031528709	-0.19955452	0.1292272	0.13356619	0.0440195758	0.09371032
Magnitud6	0.13100719	0.12438797	0.17651045	0.18956518	0.121961037	0.074156435	-0.14735155	0.14440466	0.11104059	0.0844552439	0.09321664
Magnitud7	0.19756724	0.19362485	0.30552118	0.2780526	0.20099561	0.099188472	-0.28316988	0.24255303	0.20017167	0.148936537	0.16865348
Magnitud8	0.20105382	0.16967355	0.25588545	0.31884123	0.211331412	0.131402931	-0.21588684	0.22689331	0.195477	0.1325497713	0.11507512
Magnitud9	0.13772329	0.17011885	0.2534212	0.27866073	0.465432362	0.424600343	-0.12660655	0.23594139	0.18943574	0.1485966325	0.11867433
Magnitud10	0.03461035	0.0784492	0.111183623	0.160145	0.392353835	0.502726943	0.01432143	0.15159357	0.08563346	0.1014850949	0.0554273
Magnitud14	-0.16880985	-0.19127365	-0.27211447	-0.31503583	-0.1368139	-0.00130504	0.57110161	-0.23236189	-0.21769639	-0.1452074873	-0.14530085
Magnitud20	0.1452357	0.20784164	0.29865467	0.3301731	0.258226359	0.190105729	-0.24846001	0.29488629	0.22415991	0.200216759	0.12804288
Magnitud30	0.14627611	0.1738639	0.25032917	0.25759603	0.189477374	0.112324887	-0.20866462	0.21361637	0.19433358	0.1347212252	0.12411121
Magnitud35	0.06871825	0.14119529	0.18448109	0.20660652	0.169374483	0.143590027	-0.1400205	0.1817369	0.14421127	0.2019670211	0.06673893
Magnitud42	0.09615128	0.12698553	0.18111022	0.16229613	0.12750491	0.062474351	-0.16195712	0.13474046	0.1253433	0.0805086037	0.21365889
Magnitud44	0.02072481	0.01370996	0.02976585	0.03575842	0.002111949	-0.006071369	-0.02610495	0.01844862	0.01566712	-0.0002961624	0.055204

Tabla 14.2. Matriz de correlación de variables contaminantes 4 horas después

- Correlaciones a 8 horas de diferencia entre la variable Y e X:

Magnitud1	Magnitud6	Magnitud7	Magnitud8	Magnitud9	Magnitud10	Magnitud14	Magnitud20	Magnitud30	Magnitud35	Magnitud42	Magnitud44
Magnitud1	0.27503265	0.08098472	0.15740971	0.15147585	0.084023307	-0.001952709	-0.17901299	0.09781111	0.10170155	0.02168417	0.07717974
Magnitud6	0.14031724	0.06978937	0.12520804	0.11678476	0.106681084	0.048061006	-0.11585809	0.12526532	0.10130959	0.06231369	0.06993123
Magnitud7	0.19657057	0.12965061	0.23020589	0.21592733	0.176308462	0.069039139	-0.24337306	0.20892135	0.18296217	0.11485319	0.12495206
Magnitud8	0.18911147	0.0978989	0.18507109	0.20881236	0.175781384	0.084478006	-0.15090114	0.19725312	0.17052805	0.10734187	0.08849113
Magnitud9	0.11095854	0.07907253	0.16031914	0.18298473	0.363917534	0.328961045	-0.08972651	0.19680079	0.14350533	0.11199185	0.09100755
Magnitud10	0.01887921	0.02088216	0.05519379	0.09341375	0.322090498	0.413992748	0.02937456	0.13779315	0.06598051	0.0814512	0.03176541
Magnitud14	-0.1848925	-0.18264268	-0.25407315	-0.27392976	-0.157827964	-0.022581313	0.31948705	-0.22712361	-0.24232483	-0.15073684	-0.0165843352
Magnitud20	0.11722309	0.10229271	0.19117215	0.19579777	0.186340386	0.119580851	-0.16911363	0.22390226	0.15501189	0.13534183	0.09446
Magnitud30	0.12281599	0.07885133	0.15717075	0.14791143	0.129626446	0.050894381	-0.17091741	0.14999476	0.13875736	0.08303931	0.0946966
Magnitud35	0.04070929	0.04303072	0.09985156	0.09930897	0.104578041	0.081944059	-0.09389593	0.12467663	0.08656932	0.13789193	0.04506895
Magnitud42	0.0969826	0.08472559	0.13987559	0.12517272	0.104059736	0.049935005	-0.11540868	0.10851954	0.10792687	0.05372186	0.13695559
Magnitud44	0.03246941	0.0183253	0.02171921	0.03541476	-0.000140482	-0.008628054	-0.01529122	0.00979172	0.02364663	-0.00176085	0.03744513

Tabla 14.3. Matriz de correlación de variables contaminantes 8 horas después

En cuanto a las dependencias en instantes de tiempo posteriores [T. 14.1, 14.2, 14.3], ocurre de manera similar a las correlaciones [T. 13.1, 13.2, 13.3], en este caso los grados de correlación más grandes son normalmente los de la misma variable adelantada en el tiempo, y a medida que pasa el tiempo menor grado de correlación se obtiene entre las variables.

7. Predicción de Variables

Para la predicción de variables se ha procedido a la obtención de un modelo VAR (Vector Autoregressive Model, en inglés)¹⁹ basado en el AIC más bajo mediante la selección de criterio que describe el algoritmo VARSelect²⁰. No obstante, previamente se ha realizado un análisis de reducción de dimensiones PCA. Con este análisis se pueden obtener las variables más significativas²¹ para nuestro VAR. Nuestro modelo VAR además de crearse mediante variables linealmente independientes, éstas han de ser estacionarias, para lo cual es posible que no lo sean y se requiera de una diferenciación de la serie en cuestión. Haciendo uso del test de Granger, se puede comprobar cualquier relación unidireccional entre las dos variables.

7.1 Análisis PCA

Se han obtenido los siguientes resultados mediante el análisis de reducción de dimensiones:

- Para un PCA que considera únicamente las variables contaminantes:

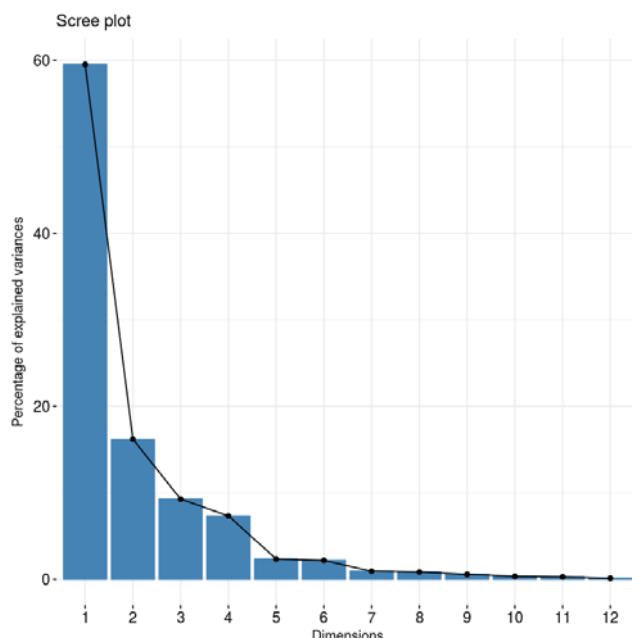


Figura 20.1. Significación de dimensiones para variables contaminantes

Donde más del 77% es representado por dos dimensiones, y la contribución obtenida de cada variable a cada dimensión:

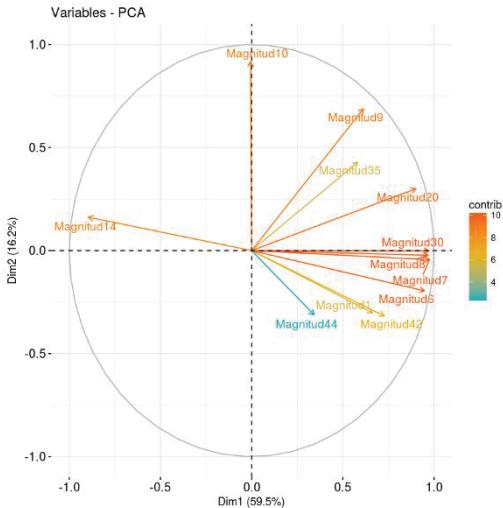


Figura 20.2. Contribución de las variables contaminantes a dos dimensiones

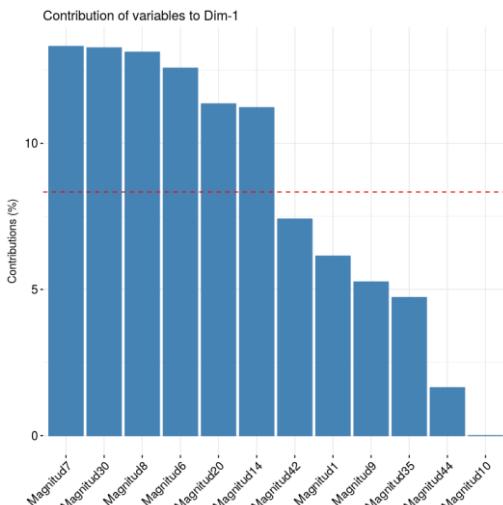


Figura 20.3. Contribución de las variables contaminantes a la dimensión 1

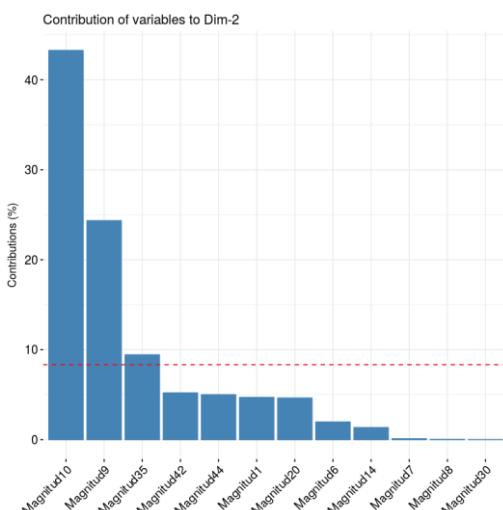


Figura 20.4. Contribución de las variables contaminantes a la dimensión 2

A la vista de los resultados [F. 20.2, 20.3, 20.4] se puede apreciar como la Magnitud44 (Hexano) es la menos relevante, en ambas dimensiones. Y como el resto de variables son bastante relevantes.

- Para un PCA que considera además las variables atmosféricas

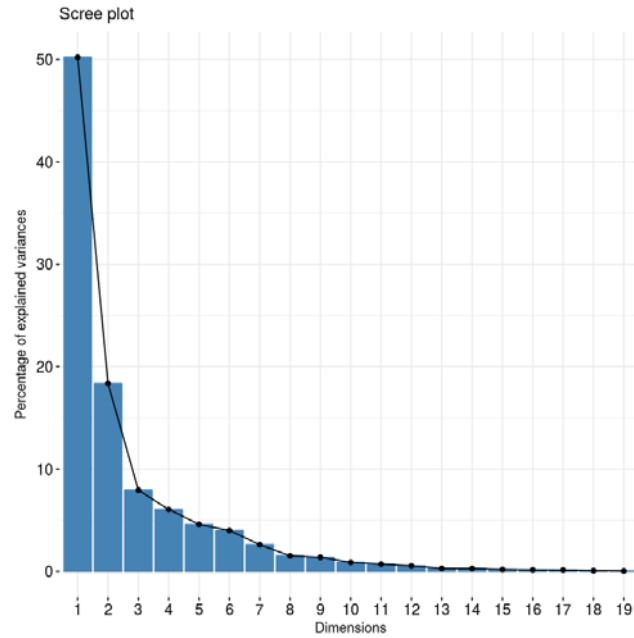


Figura 21.1. Significación de dimensiones para variables atmosféricas y contaminantes

Donde alrededor de un 68% es representado por las dos principales dimensiones, y la contribución obtenida de cada variable a cada dimensión:

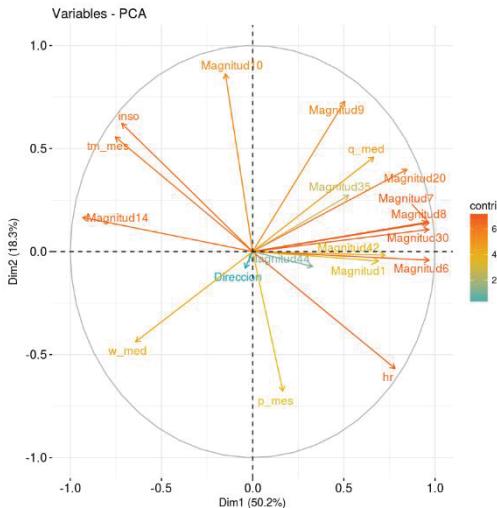


Figura 21.2. Contribución de las variables a dos dimensiones

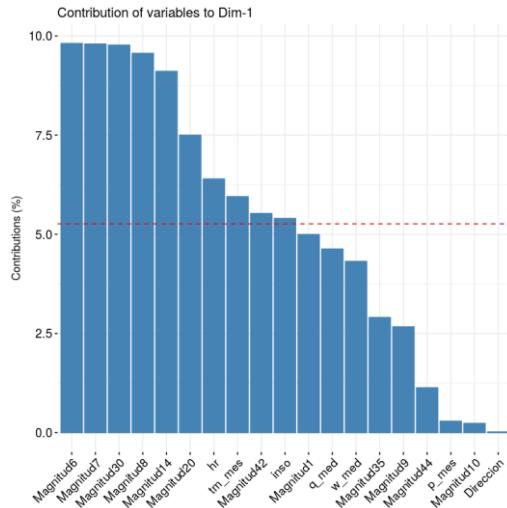


Figura 21.3. Contribución de las variables a la dimensión 1

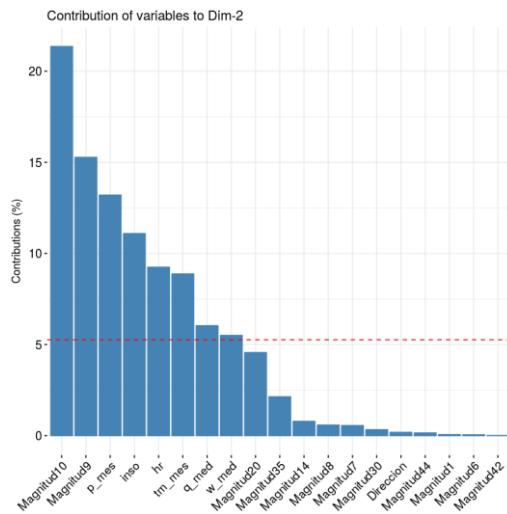


Figura 21.4. Contribución de las variables contaminantes a la dimensión 2

Mediante este análisis se observa que en el segundo PCA [F. 21.2, 21.3, 21.4], hay más variables significativas, pero tenemos que tener en cuenta que las variables atmosféricas son exógenas, es decir, no podemos controlarlas. En la creación del modelo VAR, si se precisan como variables exógenas, es necesario introducir los valores futuros de estas variables. No obstante, se obtendrían mejores resultados que considerando únicamente las variables contaminantes del primer PCA en la predicción de los gases.

7.2 Modelado VAR

Para la obtención del modelo se ha de garantizar que las variables son estacionarias, para lo cual se han usado los tests:

- Dickey-Fulley Aumentado (ADF en inglés)
- Phillips-Perron (PP)
- Zivot-Andrew
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

El test Philips-Perron toma como hipótesis nula que la serie es estacionaria. Este test a diferencia del test Augmented-Dickey-Fulley, es no paramétrico, sólo necesita que los residuos sean estacionarios. En los test adf, los residuos deben ser independientes y estar indécticamente distribuidos. Por tanto, el test ADF no es robusto cuando la serie es autocorrelacionada.

El test de Zivot Andrews toma la misma hipótesis que el test ADF. Sin embargo tiene en cuenta la presencia de posibles cambios estructurales que puedan aparecer. Este test ha sido de gran utilidad, pues tanto el test ADF como el test PP, han devuelto resultados contradictorios en presencia de un cambio estructural en algunas de las series. Por tanto, los test de Perron y Dickey-Fulley aumentado podemos decir que no diferencian una trayectoria de raíz unitaria de una estacionaria en presencia de un cambio estructural, confirmando así los resultados arrojados por el test KPSS, el cual toma como hipótesis nula la estacionalidad de las variables.

Se han obtenido que las magnitudes dióxido de carbono (SO_2) y etilbenceno (C_8H_{10}), tienen raíces unitarias y por tanto requieren de diferenciación para llegar a ser estacionarias [F. 32, 33]. En cambio, el resto de variables contaminantes son estacionarias.

Por consiguiente, se han descartado estas dos magnitudes para la obtención del modelo VAR.

Una vez obtenidas las variables estacionarias, mediante el uso del método VARselect se han obtenido los siguientes criterios de selección para la especificación de retrasos temporales en nuestro modelo [T.15]:

	1	2	3	4	5	6	7	8	9	10	11	12
AIC(n)	-11.68454	-13.31917	-14.29425	-17.60871	NaN	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
HQ(n)	-10.04612	-10.31539	-9.925111	-11.87421	NaN	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
SC(n)	-7.495854	-5.639908	-3.12441	-2.948296	NaN	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
FPE(n)	8.896148E-06	2.349815E-06	2.117924E-06	6.394247E-07	-1.368652E-37	0	0	0	0	0	0	0

Tabla 15. Estimadores de exactitud obtenidos mediante VARSelect

Tomando el valor 5, como el índice positivo más cercano a 0 definido por el FPE.

Las predicciones de las variables endógenas obtenidas por el modelo para el 2017 han sido:

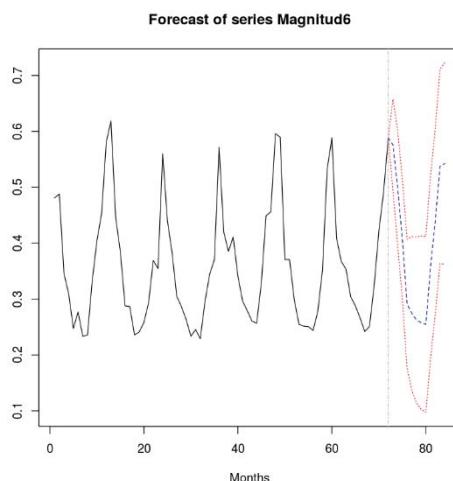


Figura 22. Predicción CO en 2017

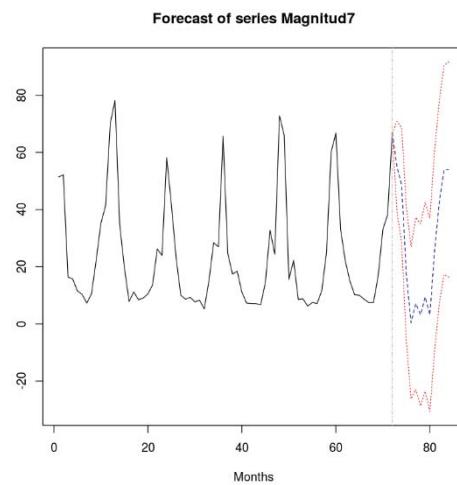


Figura 23. Predicción NO en 2017

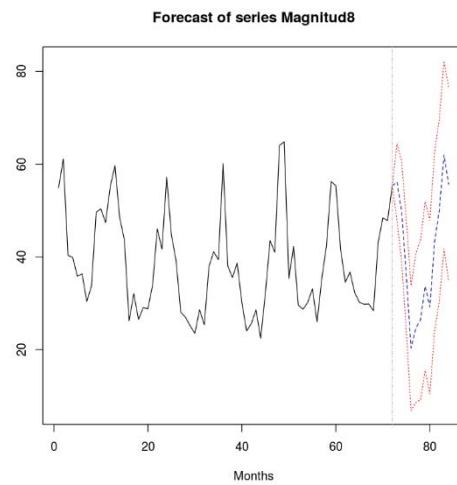


Figura 24. Predicción NO₂ en 2017

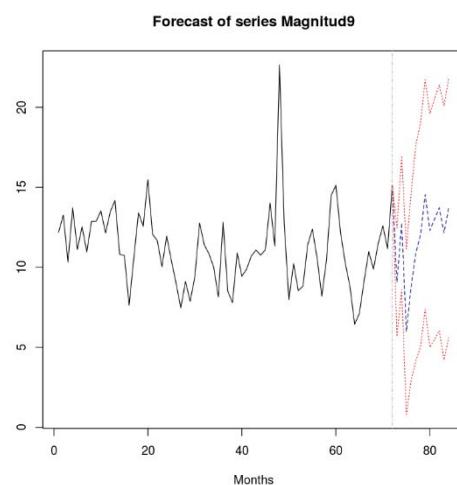


Figura 25. Predicción PM2.5 en 2017

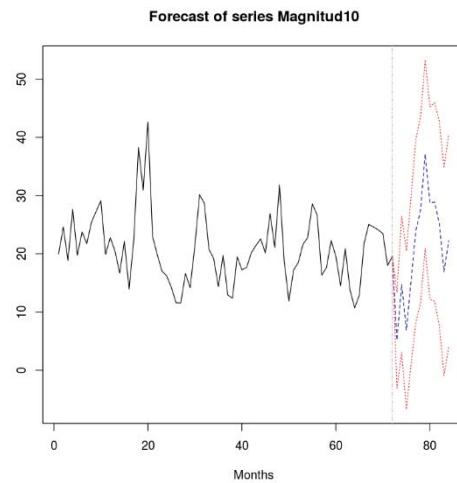


Figura 26. Predicción PM10 en 2017

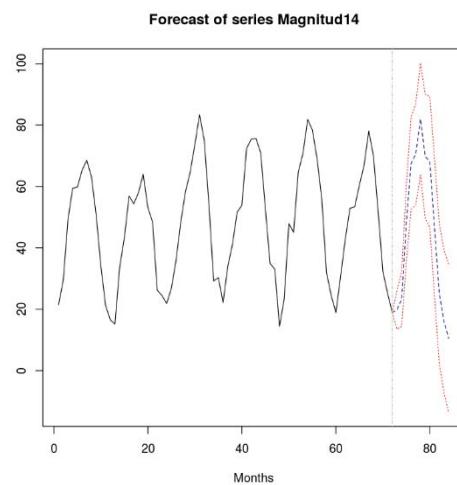


Figura 27. Predicción O₃ en 2017

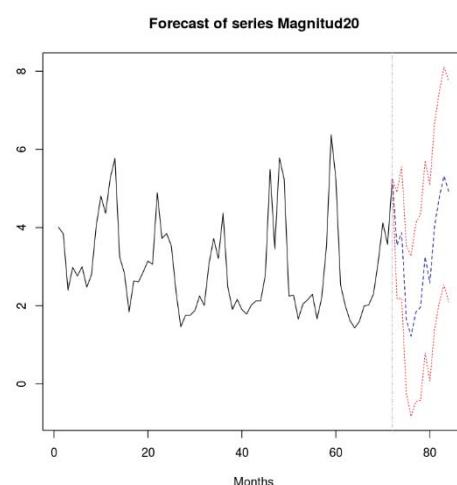


Figura 28. Predicción C₆H₅-CH₃ en 2017

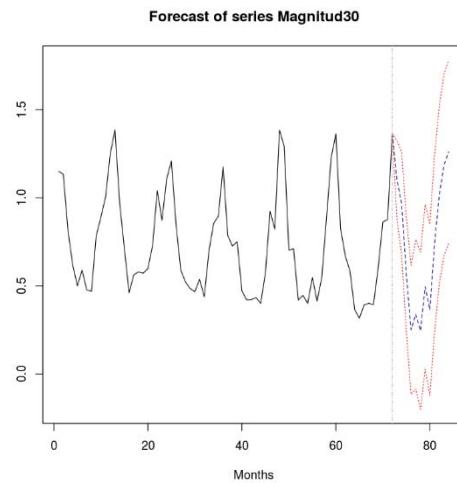


Figura 29. Predicción C₆H₆ en 2017

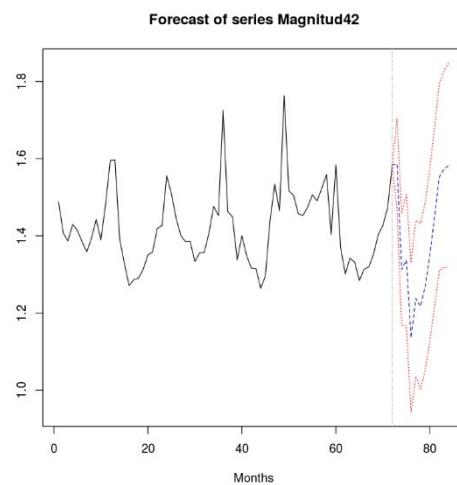


Figura 30. Predicción C₆H₁₂ en 2017

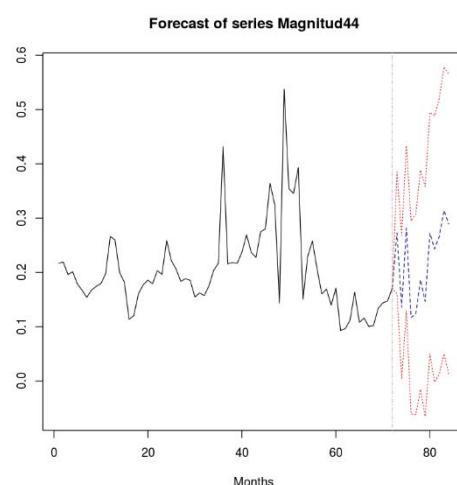


Figura 31. Predicción C₆H₁₄ en 2017

7.2 Modelado VAR

Mediante una regresión lineal y los valores predecidos de las variables endógenas por el modelo, se han obtenido las predicciones de dióxido de carbono (SO_2) y etilbenceno (C_8H_{10}):

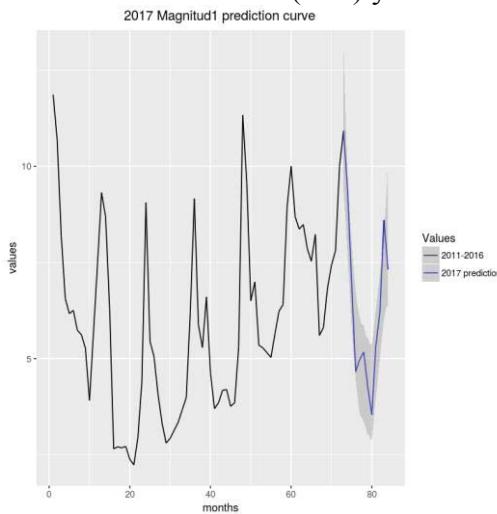


Figura 32. Predicción SO_2 en 2017

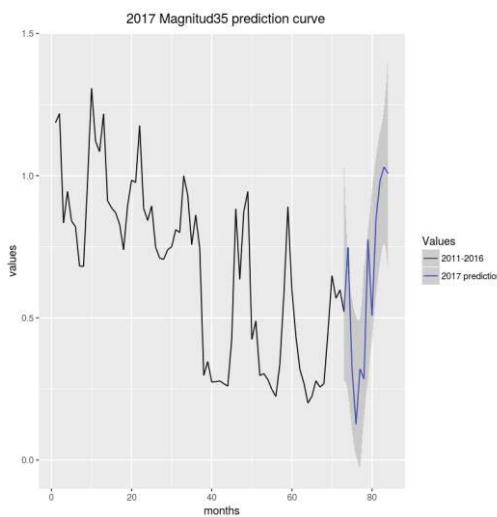


Figura 33. Predicción C_8H_{10} en 2017

A modo de resumen se puede concluir de las predicciones obtenidas [F. 22-33] toman valores cercanos a los que se puede apreciar durante los años 2011-2016.

Finalmente, se han comparado dichos resultados de este modelo con las predicciones obtenidas de los modelos univariantes ARIMA, NNAR y ETS, escogiendo como medidas de exactitud por ejemplo el MAE, RMSE, MAPE y MASE^{22,23}. Tomando como periodo de entrenamiento de 2011 a 2015, y el 2016 como periodo de prueba se obtienen:

Holts-Winter	Item	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil.s.U
Training set	In Sample Error	-0.07424049	4.584659	3.744081	-3.581464	17.76901	0.6484913	0.3011323	NA
Test set	Out Sample Error	0.81560883	3.724697	2.919659	1.250769	16.39593	0.5056978	0.1368247	0.8676932

Tabla 16.1. Estimadores de exactitud (accuracy) mediante Holts-Winter

ARIMA	Item	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil.s.U
Training set	In Sample Error	-0.05379947	5.259111	4.075943	-6.214048	20.2342	0.7059712	-0.03463588	NA
Test set	Out Sample Error	-1.61407213	4.815536	3.976706	-16.428093	26.47904	0.688783	0.44392197	1.17025

Tabla 16.2. Estimadores de exactitud (accuracy) mediante ARIMA

NNAR	Item	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil.s.U
Training set	In Sample Error	-0.00289854	1.361438	1.038751	-1.333427	5.507728	0.1799162	-0.02821042	NA
Test set	Out Sample Error	-4.62843185	6.68225	5.428868	-31.633014	35.33201	0.9403038	0.31576234	1.531417

Tabla 16.3. Estimadores de exactitud (accuracy) mediante NNAR

A la vista de las medidas de exactitud, modelo que mejores resultados ofrece es el Holts-Winter.

Una comparativa de las predicciones de estos 3 modelos para la magnitud PM10:

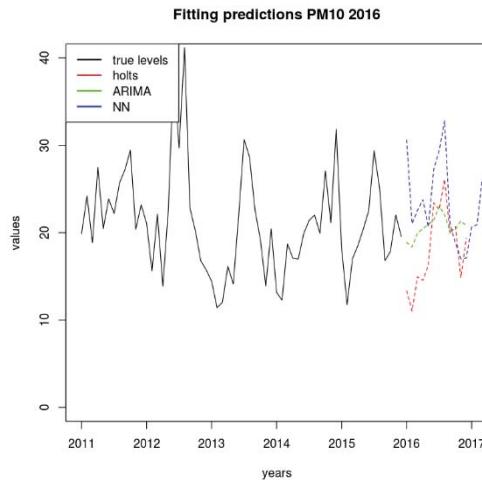


Figura 34. Comparativa de ajuste predictivo 2016 entre Holts, ARIMA, y NNAR

Escogiendo holts-winter, la predicción de la variable PM10 para 2017, se obtiene:

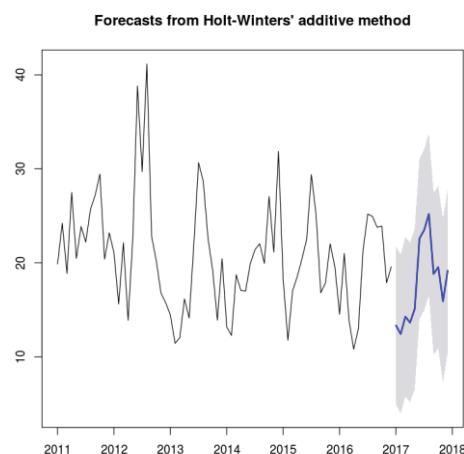


Figura 35. Predicción PM10 mediante Holts-Winter

En contraste con la predicción obtenida por el modelo VAR:

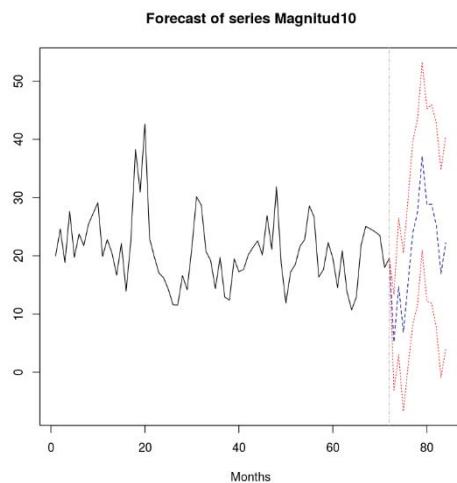


Figura 36. Predicción PM10 obtenida del modelo VAR

Lo cual podemos concluir que no es un mal resultado, aunque el error estándar es mucho más grande.

De acuerdo a las predicciones obtenidas para el 2017, los únicos contaminantes a tomar especial atención al estar muy próximos a los límites establecidos son: NO, NO₂, PM2.5 y PM10.

Los valores obtenidos en 2017 han sido:

$\mu\text{g}/\text{m}^3$	Monóxido de Nitrógeno (NO)	Dióxido de nitrógeno (NO ₂)	PM2.5	PM10
1	55.0702142	56.10909	9.0709	5.163611
2	48.4343581	49.75803	12.783204	14.756718
3	17.2053526	35.70384	5.97967	6.868104
5	0.4049149	20.33956	8.777476	15.399318
5	7.1570764	24.61027	10.884786	23.91558
6	3.1810094	26.43759	12.009499	27.431286
7	9.4417111	33.70371	14.545305	37.094799
8	3.1252052	29.18189	12.301108	28.71409
9	25.6569219	43.2883	13.01692	28.884996
10	42.5099948	50.01006	13.72132	25.24538
11	53.8444277	61.97956	12.150506	16.895407
12	54.0189405	55.66245553	13.693518	22.211626

Tabla 17.1. Valores mensuales de los contaminantes cerca del límite en 2017

Promediando los valores de 2016 y las predicciones de 2017 se ha obtenido:

	Monóxido de Nitrógeno (NO)	Dióxido de nitrógeno (NO ₂)	PM2.5	PM10
Valores límite anuales	30µg/m ³	40µg/m ³	10µg/m ³ (OMS) 25µg/m ³ (R.D.)	20µg/m ³
Promedio valores 2016	22.383µg/m ³	38.185µg/m ³	10.427µg/m ³	19.113µg/m ³
Promedio predicción 2017	26.671µg/m ³	40.565µg/m ³	11.578µg/m ³	21.048µg/m ³

Tabla 17.2 Valores contaminantes de los contaminantes cerca del límite del 2017 en promedio

Como se puede apreciar los valores de NO₂ y PM10 exceden ligeramente los índices del R.D. 102/2011 por lo que se deberían tomar medidas correctivas para evitar alcanzar estos niveles a final de año, con especial atención a los meses de verano para las variables PM y los meses invernales NO_x. Además también se volverían a exceder ligeramente los niveles de PM2.5 establecidos por la OMS.

Se puede concluir que si no se toman medidas al respecto los niveles de contaminación terminan siendo más elevados a la vista de los resultados. Cabe mencionar que las variables NO_x afectan a la acelerar la reacción de O₃, elevando así sus índices.

8. Despliegue de la aplicación

La aplicación además de desplegarse en shiny, y poderse ejecutar localmente, ha sido desplegada en la comunidad shinyapps.io.

Dicha aplicación tiene como objeto explorar los principales modelos predictivos realizados, suavizamiento exponencial, ARIMA y redes neuronales como modelos univariantes. Y el modelo VAR. A pesar de que puede servir para cualquier data frame, el modelo VAR ha sido llamado con los mismos parámetros que los ejecutados para este proyecto. Por lo que para el modelado VAR no tendría sentido ejecutarlo con otro data frame distinto del que se ha sido utilizado.

El objeto de dicha aplicación es obtener gráficos y estadísticos de las distintas variables de forma dinámica tanto analíticos como predictivos, de modo que reciba una hoja de cálculo sea .csv o .xls y mediante el menú de la pantalla principal se establezcan los delimitadores necesarios para obtener la información de manera conforme a como se haya guardado tal información, como pueden ser las opciones del header, separadores de coma, punto y coma o tabulado. Para simplificar el proceso se trabajará con datos mensuales, por lo que una fila representará datos referentes a un mes.

Una vez los datos sean legibles y debidamente representados en la tabla de la derecha ya se puede hacer uso completo de la aplicación. Si nos interesase obtener gráficos analíticos de alguna variable basta con poner su indicador en la tabla en la opción Column To Analysis, tomando el valor 1 la primera columna. A continuación se debe establecer el año de comienzo de nuestros datos obtenidos, y establecer un periodo de entrenamiento de nuestra serie, para obtener los predictores. Al tratarse de series temporales y con carácter estacional, es aconsejable establecer un periodo reciente de un año semestre o cuatrimestre, en nuestro caso tomaremos como periodo de entrenamiento el último año. Este paso es importante para obtener buenos estimadores en nuestras predicciones, ya que

delimitando nuestro periodo de entrenamiento nos dará las medidas de precisión de tales estimadores.

Una vez configurados los parámetros podemos proceder a usar todos los algoritmos ofrecidos por la aplicación.

A continuación se muestran los pasos a seguir para obtener los resultados en la aplicación:

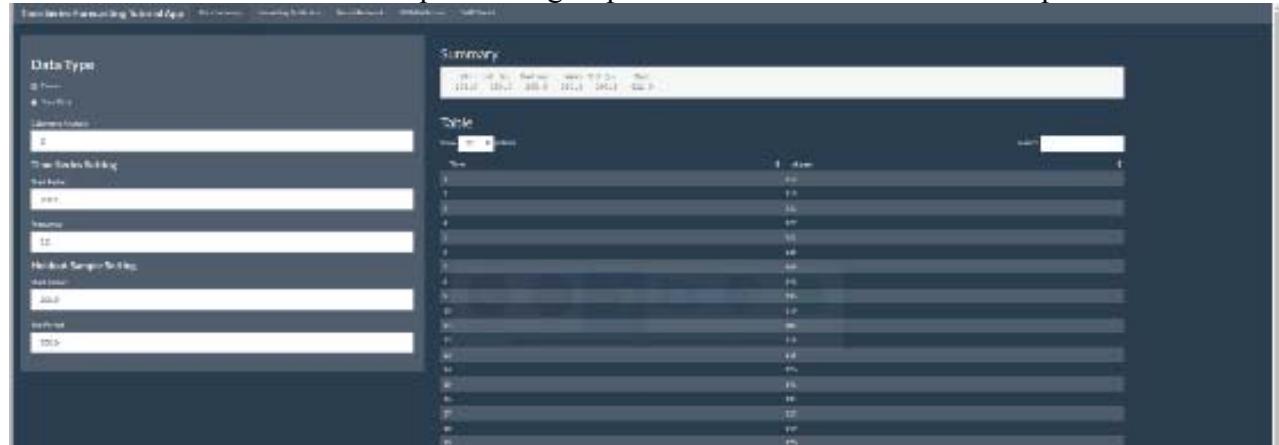


Figura 37.1. Pantalla principal de la aplicación ejecutada desde shinyapps



Figura 37.2. Hacemos click en Your Data

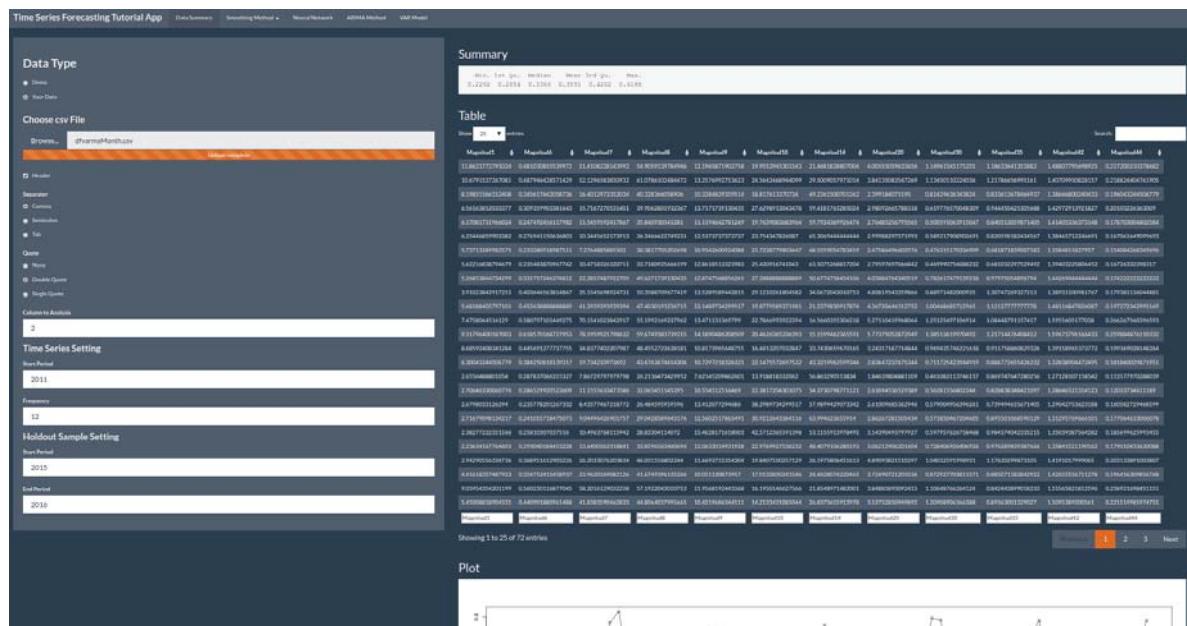


Figura 37.3. Seleccionamos nuestro fichero 'dfvarmaMonth.csv'

Despliegue de la aplicación

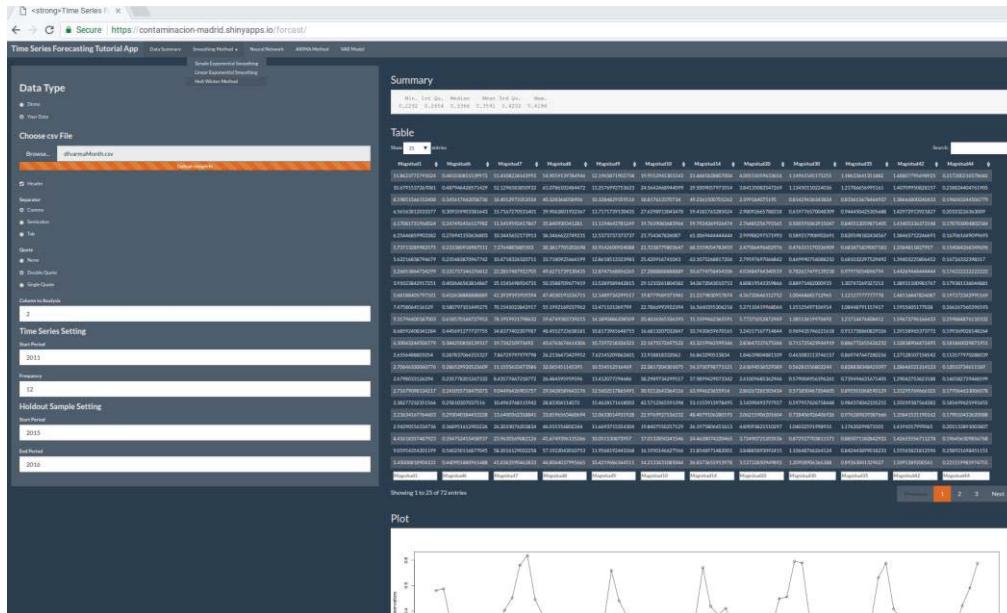


Figura 38. Seleccionamos la columna a analizar y procedemos a la elección del algoritmo

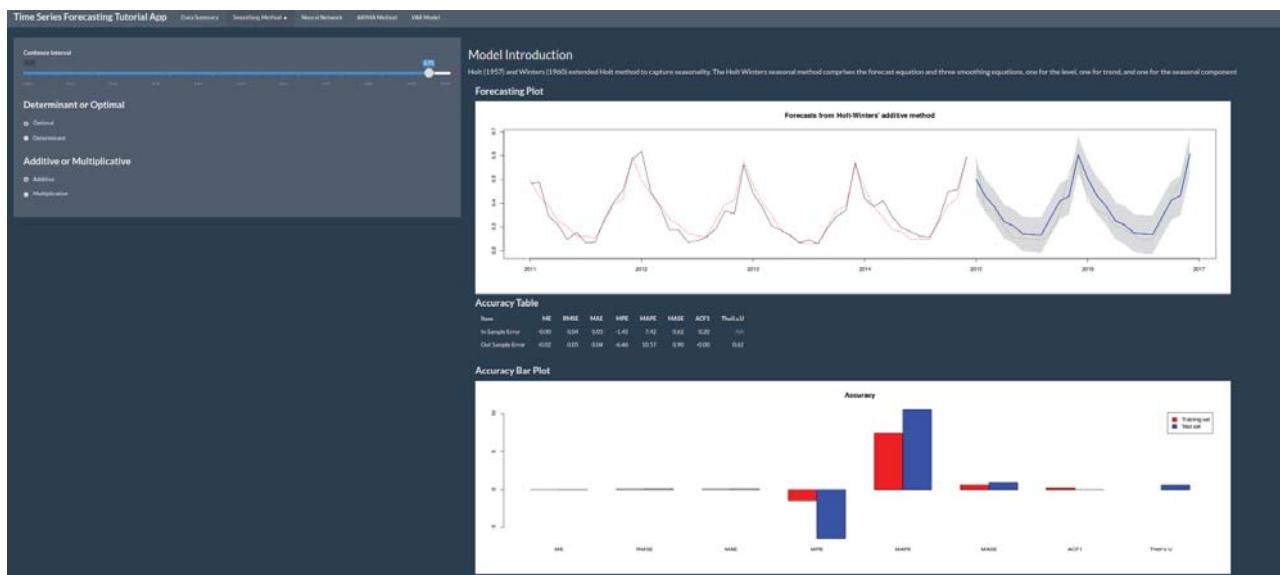


Figura 39. Seleccionando Smoothing Method, Holts-Winter obtenemos el ajuste de predicción del 2016

Despliegue de la aplicación

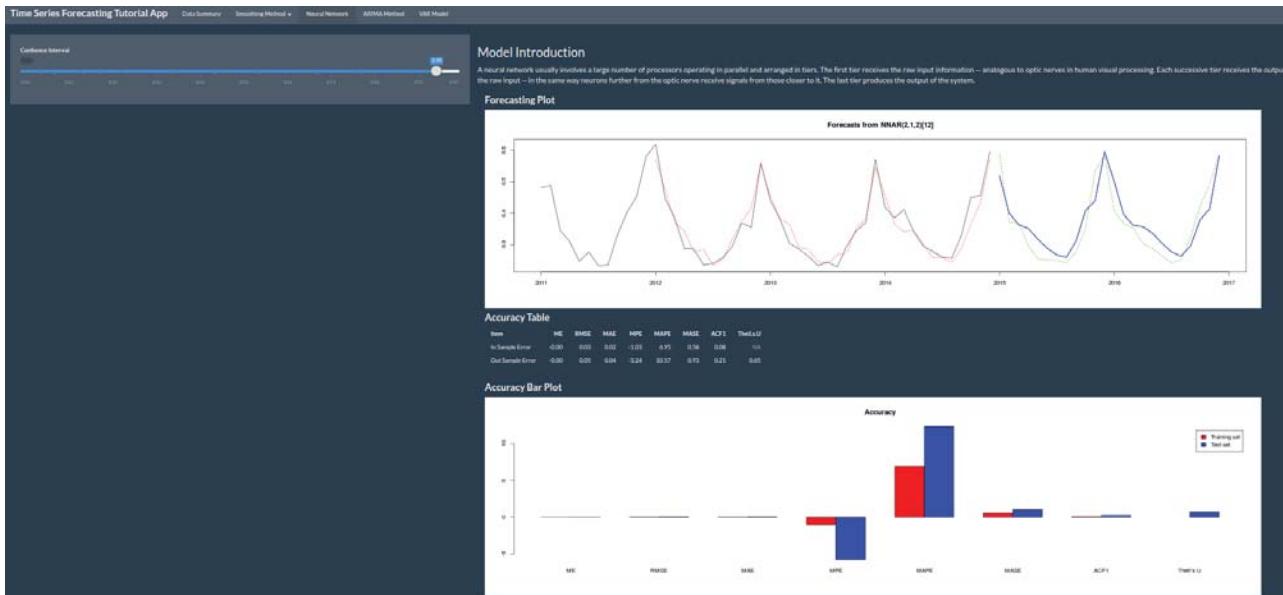


Figura 40. Seleccionando Neural Network obtenemos el ajuste de predicción del 2016

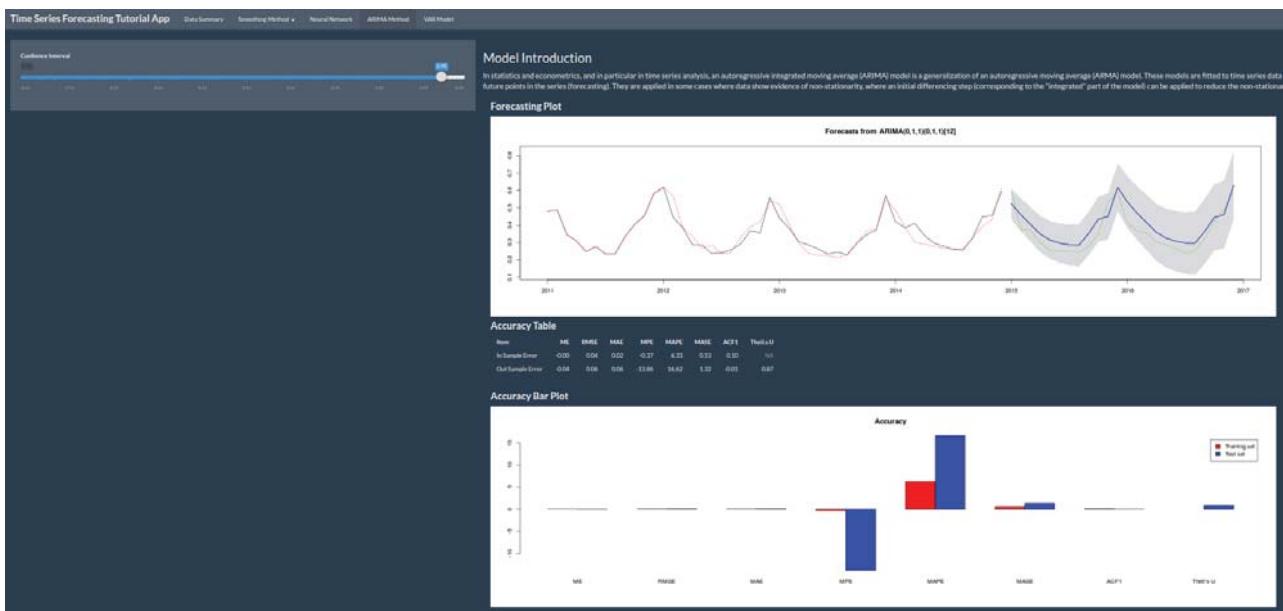


Figura 41. Seleccionando ARIMA Method, obtenemos el ajuste de predicción del 2016



Figura 42. Seleccionando VAR Model, obtenemos las predicciones de todas las variables en el 2017

9. Sistema de recomendaciones

El Ayto. de Madrid tiene su propio Plan de Calidad del Aire y Cambio Climático de la Ciudad de Madrid²⁹ conocido como Plan A, para prevenir el cambio climático y definir estrategias de adaptación. La estructura de este Plan A se inicia respecto a una evaluación de la situación actual en cuanto a contaminantes, marco normativo, características físicas y socioeconómicas, etc. Evaluando los resultados del Plan de Calidad del Aire 2011-2015^{27,28}, el cual se estructuró a través de 70 medidas, 42 de ellas dirigidas al tráfico rodado, que es el principal sector productor de emisiones contaminantes en la ciudad. Estimando futuros escenarios bajo diversas hipótesis.

El grado de cumplimiento del plan anterior fue del 64,3%, finalizándose 45 de las 70 medidas previstas. En lo que respecta a los 9 sectores del Plan, destaca el sector de movilidad y transporte, con el grueso de las acciones contempladas, donde se ha alcanzado un 69%.

Los objetivos específicos del Plan eran:

- Consolidar los logros alcanzados en la mejora de la calidad del aire, impulsando la mejora continuada de la misma de un modo integral.
- Asegurar el cumplimiento de todos los objetivos establecidos en la normativa vigente en el plazo legalmente previsto y en particular, el del dióxido de nitrógeno.

Las medidas de este nuevo plan están estructuradas en distintos programas de acción: movilidad sostenible, regeneración urbana, adaptación al cambio climático y sensibilización ciudadana y colaboración con otras administraciones.

Por último, el Plan A contempla un análisis del impacto y un plan de seguimiento y evaluación.

El Plan A marca 2020 como año límite para la consecución de los objetivos de calidad del aire exigidos por la normativa y un horizonte a más largo plazo, y 2030, para la necesaria transición energética y consolidación de un modelo de ciudad de bajas emisiones.

En el punto 3.4 Análisis de emisiones y contaminantes y GEI. Fuentes, del Plan A, se analizan los resultados del “Inventario de Emisiones Contaminantes a la Atmósfera en la Ciudad de Madrid 1999-2014”, elaborado siguiendo la metodología EMEP/CORINAIR, de la Agencia Europea de Medio Ambiente y que considera tanto las emisiones directas (fuentes fijas y móviles) como las indirectas (asociadas al consumo eléctrico).

Obteniendo así información suficiente para definir medidas adecuadas y eficaces para mejorar la calidad del aire y disminuir el cambio climático estimando la cantidad de contaminante vertido a la atmósfera y sus principales fuentes.

Las principales fuentes de contaminantes contempladas en el inventario agrupadas bajo la nomenclatura SNAP son las siguientes:

- 01 - Combustión en la producción y transformación de energía
- 02 - Plantas de combustión no industrial
- 03 - Plantas de combustión industrial
- 04 - Procesos industriales sin combustión
- 05 - Extracción y distribución de combustibles fósiles y energía geotérmica
- 06 - Uso de disolventes y otros productos
- 07 - Transporte por carretera
- 08 - Otros modos de transporte y maquinaria móvil
- 09 - Tratamiento y eliminación de residuos
- 10 - Agricultura
- 11 - Naturaleza

Respecto a los contaminantes evaluados, en la tabla siguiente se muestran los que se incluyen en el inventario, y que se analizan en el presente documento, y en la misma se distingue entre gases de efecto invernadero (GEI), acidificadores y precursores de ozono, y material particulado:

También se distinguen los distintos gases contaminantes incluidos en el inventario de acuerdo a sus efectos:

Contaminantes	
Gases de Efecto Invernadero (GEI)	Metano
	SO ₂
	Hidrocarburos
	N ₂ O
	Perfluorocarburos
	Hexafluoruro de Azufre
Acidificadores y precursores de Ozono	CO
	COV no Metánicos
	Amoniaco
	NO _x
	SO _x
Material Particulado	PM2.5
	PM10
	Partículas sólidas totales

Tabla 18. Lista de gases respecto a contaminantes evaluados

De acuerdo al análisis y la predicción obtenida, se precisa prestar especial atención son NO_x y PM, ya que son los que se encuentran más cercanos a los valores límites establecidos.

En el siguiente gráfico se expresa la evolución de la contribución porcentual de cada grupo SNAP al total de las emisiones totales de NOx. El grupo de mayor importancia relativa es el del transporte por carretera (07) que contribuye en más de un 50 % a las emisiones totales de NOx en el periodo 1999-2014. Le siguen en importancia los grupos 08 y 02, correspondientes a otros modos de transporte y maquinaria móvil y plantas de combustión no industrial.

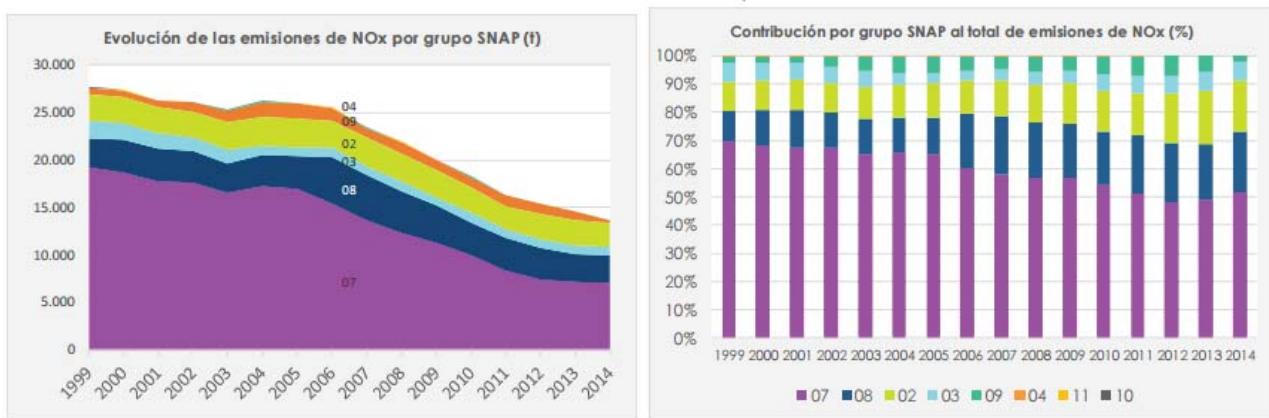


Figura 43. Fuentes contaminantes y contribución de emisiones de NOx por grupo SNAP.

En cuanto a las partículas sólidas, como ya se ha visto antes, es el grupo 07 del transporte por carretera el que contribuye a las emisiones de PM_{2.5}, PM₁₀ y PST, de forma mayoritaria, con más de un 50 % de las emisiones totales. El siguiente grupo en importancia es el SNAP 02, que supone, más de un 20% de las emisiones totales, que hace referencia a las plantas de combustión no industrial.



Figura 44. Fuentes contaminantes y contribución de emisiones de PM por grupo SNAP.

Volviendo al Plan A, los objetivos específicos son:

- Cumplir la legislación europea y nacional en materia de calidad del aire.
- Alcanzar niveles de calidad del aire para partículas en suspensión acordes con el valor guía de la Organización Mundial de la Salud (OMS).
- Conseguir una reducción en el año 2030 superior al 40% de las emisiones totales de GEI del municipio de Madrid respecto al año 1990, contribuyendo a los objetivos del Acuerdo de París y la Agenda Climática de la UE y en línea como la nueva Alianza de Alcaldes para el Clima y la Energía.
- Cumplir el compromiso de reducción del 50% de las emisiones causadas por la movilidad urbana en 2030, con respecto a 2012.
- Desarrollar una estrategia de adaptación frente a los efectos del cambio climático, disminuyendo la vulnerabilidad urbana frente a los riesgos asociados al calentamiento global.

Para alcanzar estos ambiciosos objetivos el Plan A promueve el desarrollo de un conjunto de medidas organizadas en cuatro líneas de acción: movilidad sostenible, regeneración urbana, adaptación al cambio climático y sensibilización y cooperación con otras Administraciones.

El marco temporal del Plan A contempla dos horizontes, 2020 para la implementación de medidas estructurales y tecnológicas concretas que resulten en una significativa reducción de emisiones exigida por la normativa de calidad de aire y un horizonte a más largo plazo, 2030, para la necesaria regeneración urbana, transición energética, renovación del parque de vehículos y consolidación de un modelo de ciudad de bajas emisiones que permita la consecución con garantías del conjunto de objetivos del Plan A.

Las líneas de actuación del Plan A están diseñadas para dar respuesta conjunta a dos importantes retos que tiene la ciudad: respuesta ante el cambio climático y mejora de la calidad del aire.

Las dos primeras líneas de actuación son claves para conseguir la necesaria reducción de las emisiones contaminantes tanto de los gases y partículas tóxicas para la salud, como de los gases de efecto invernadero que producen el cambio climático. La tercera línea se centra en procurar las condiciones necesarias para garantizar la mejor adaptación de la ciudad al cambio climático a través de soluciones basadas en la naturaleza. La cuarta y última línea es de carácter transversal, sirviendo tanto a los objetivos de cambio climático como a los de calidad del aire mediante la concienciación ciudadana, la colaboración con otras administraciones y la definición de incentivos para el cambio.

Movilidad más sostenible

Reducción de la intensidad de tráfico privado motorizado: actuaciones sobre la red viaria y el espacio público para la promoción de los modos activos de movilidad (peatonal y ciclista) y del transporte público.

- Medidas de fomento de tecnología de bajas emisiones:
 - impulso a la movilidad eléctrica
 - actuaciones sobre las emisiones de flotas estratégicas (autobuses, taxis, flotas de servicios municipales) y distribución urbana de mercancías
- Medidas sobre vehículos privados motorizados: incentivos fiscales, restricción gradual de acceso, aparcamiento y circulación a los vehículos más contaminantes.
- Elaboración de un plan de movilidad sostenible municipal.

Gestión urbana baja en emisiones y mayor eficiencia energética

- Fomento de la sustitución de combustibles de calefacción contaminantes: prohibición del uso de carbón y regulación del uso de la biomasa
- Promoción del uso de energías renovables

Adaptación al cambio climático: Madrid más natural

Impulso de proyectos e implementación de soluciones basadas en la naturaleza con objetivo de adaptar la ciudad a las amenazas ambientales derivadas de las alteraciones climáticas. Intervenciones en edificios, barrios y en las grandes infraestructuras de la ciudad.

Sensibilización ciudadana y colaboración con otras administraciones

Tiene como objetivo principal aumentar la conciencia ciudadana sobre las consecuencias que tienen a corto, medio y largo plazo, tanto una mala calidad del aire como el aumento de gases de efecto

invernadero, y buscar las necesarias vías de coordinación con las administraciones central y regional, que permitan abordar estos retos cuyo ámbito supera al municipio.

Las 30 medidas clave del Plan A, vienen descritas en el apartado 5 del Plan de Acción:

Movilidad sostenible: que a su vez se divide en 2 bloques de medidas: las **Actuaciones sobre la red viaria y el espacio público dirigidas a reducir la intensidad del tráfico privado y a la promoción del transporte público y de modos activos de movilidad**, de 1 a la 11, y **Actuaciones sobre el parque móvil de sectores clave con alto impacto en los patrones de movilidad en la calidad del aire (EMT, taxi, distribución urbana de mercancías y flota municipal y movilidad laboral)**.

Fomento de la movilidad compartida, de la 12 a la 21:

1.- Área Central Cero Emisiones:

Delimitación de un Área Central de perímetro continuo y acceso restringido en la que se eliminará el tráfico de paso, con el objetivo de impulsar un nuevo modelo de movilidad bajo en emisiones que priorice al peatón, a la bicicleta, al transporte público y, en general, a los vehículos menos contaminantes.

2.- Rediseño de las vías principales de distribución del tráfico y conexión periferia-centro:

Actuación sobre los ejes urbanos principales reduciendo el espacio dedicado al automóvil a favor de medios más sostenibles, potenciando las infraestructuras específicas de soporte al transporte público, la movilidad ciclista y la recuperación de espacio para los peatones.

3.- Priorización de la movilidad peatonal:

Intervenciones en espacio público ligadas a la regeneración urbana con incidencia directa en la movilidad de proximidad en los barrios, mejorando la calidad del espacio público, reduciendo la contaminación acústica y atmosférica y la siniestralidad, todo ello encaminado a crear entornos adaptados al peatón. Identificación y caracterización de la red peatonal y creación de zonas 30.

4.- Mejora y ampliación de la red y la movilidad ciclista:

Revisión y ampliación mediante la creación de una red de itinerarios ciclistas segura, eficaz e integrada con los demás medios de transporte con el objetivo de alcanzar una cuota en la distribución modal del 5 % de los desplazamientos en 2025. En una primera fase (2017) se ejecutarán 30 Km de nuevos carriles bici en vías principales de la ciudad.

5.- Ampliación del sistema de bicicleta pública y coordinación con el sistema del Consorcio Regional de Transportes de Madrid (CRTM):

Ampliar el número de bicicletas y el ámbito de actuación de BiciMAD a zonas de gran demanda, que permitan su uso de forma generalizada como modo de transporte diario y recurrente, a la vez que se renuevan las unidades y se desarrollan medios tecnológicos que posibiliten la integración completa de BiciMAD en la red madrileña de transporte público.

6.- Regulación del aparcamiento con criterios de calidad del aire:

Ampliar el número de bicicletas y el ámbito de actuación de BiciMAD a zonas de gran demanda, que permitan su uso de forma generalizada como modo de transporte diario y recurrente, a la vez que se renuevan las unidades y se desarrollan medios tecnológicos que posibiliten la integración completa de BiciMAD en la red madrileña de transporte público.

7.- Limitación de la velocidad en accesos metropolitanos y M-30:

Con objeto de canalizar el tráfico con origen destino entre sectores de la periferia urbana y metropolitana por los anillos exteriores de circunvalación de la ciudad: M-40, M-45, se establece como velocidad máxima de circulación 70 km/h en las vías de acceso a la ciudad, desde su intersección con la M-40, y se extiende la limitación de velocidad a 70 km/h a todo el recorrido del anillo urbano M-30 (subterráneo + superficie).

8.- Red de aparcamientos intermodales en la corona metropolitana:

Constitución de una red de aparcamientos en la corona metropolitana, conectada con nodos de transporte público, que facilite la transferencia del vehículo privado al transporte colectivo, constituida por doce nuevos aparcamientos en la ciudad de Madrid que aportarán 9.570 plazas. Los aparcamientos se ubicarán a una distancia máxima de 200 metros al punto de acceso del transporte colectivo y ofrecerán sistemas tarifarios adaptados a la intermodalidad.

9.- Vías preferentes y priorización semafórica para autobuses EMT:

Los aparcamientos se ubicarán a una distancia máxima de 200 metros al punto de acceso del transporte colectivo y ofrecerán sistemas tarifarios adaptados a la intermodalidad.

10.- Infraestructuras reservadas para transporte público:

Los aparcamientos se ubicarán a una distancia máxima de 200 metros al punto de acceso del transporte colectivo y ofrecerán sistemas tarifarios adaptados a la intermodalidad.

Creación de corredores para autobuses de alta capacidad o BRT (bus rapid transit) de conexión entre distritos.

11.- Ampliación y renovación de la flota de la EMT: hacia una flota de bajas emisiones:

Renovación de la flota de la EMT mediante un programa 2017-2020 de inversiones con adquisición de 750 nuevos autobuses (Gas natural, híbridos y eléctricos). Incorporación a gran escala de autobuses eléctricos (Base de operaciones Elipa).

12.- Taxi: incentivos para transformación a vehículos de bajas emisiones:

Subvenciones anuales municipales (1,5 millones € en 2017) para la adquisición de vehículos que cuenten con la etiqueta CERO o ECO. A partir de 2018 sólo se autorizará la sustitución con este tipo de vehículos.

13.- Optimización del servicio del taxi con criterios ambientales:

Incremento de la eficiencia y sostenibilidad mediante la optimización de recorridos y otras actuaciones estructurales que permitan mejorar el servicio y reducir los viajes en vacío de la flota de taxis.

14.- Distribución urbana de mercancías: optimización de la gestión de las reservas de carga y descarga en la vía pública:

Creación de un sistema de gestión y control de las plazas en vía pública reservadas para la carga y descarga en la ciudad de Madrid mediante una aplicación específica y un registro municipal de operadores de distribución urbana.

15.- Distribución urbana de mercancías con vehículos de bajas emisiones

Aplicación de medidas preferentes de acceso y horarios en el Área Central y en la zona SER para vehículos de bajas emisiones.

16.- Innovación y eficiencia en los procesos logísticos urbanos:

Fórmulas de colaboración público-privadas para la logística urbana sostenible e innovadora.

17.- Renovación del parque circulante:

Acciones encaminadas a la progresiva sustitución de vehículos motorizados con el objetivo de limitar la circulación en 2025 de los más contaminantes (sin distintivo DGT) en todo el término municipal de Madrid.

18.- Flotas de servicios municipales de bajas emisiones:

Incremento de los vehículos CERO y ECO en la flota del Ayuntamiento de Madrid hasta alcanzar el 90 % en 2030, para las flotas adscritas a contratos de servicios y las gestionadas directamente por las distintas Áreas.

19.- Planes de movilidad laboral sostenible:

Impulsar el desarrollo de planes de movilidad sostenible en empresas, así como en las administraciones públicas aprovechando su papel ejemplarizante, comenzando por la redacción de un Plan de Movilidad Sostenible Municipal. De manera complementaria instar al Gobierno Regional de la Comunidad de Madrid a la aprobación de una ley de movilidad sostenible.

20.- Red de recarga para vehículos eléctricos y suministro de combustibles alternativos: implantación de una infraestructura de combustibles alternativos conforme a la Directiva 2014/94. Promoción de una red de recarga eléctrica en espacios de acceso público, aparcamientos municipales y de residentes, lugares de trabajo o instalaciones municipales. Extensión de redes de acceso público de otros combustibles alternativos.

21.- Impulso a las iniciativas de movilidad compartida

Promoción y apoyo a las iniciativas de movilidad compartida, para mejorar y diversificar la oferta de transportes de la ciudad de Madrid.

Regeneración urbana, la cual se divide en dos bloques de medidas: **Gestión urbana baja en emisiones y eficiencia energética**, de la 22 a la 25, y la **Gestión energética en edificios e instalaciones municipales**, de la 26 a la 28:

22.- Regeneración y rehabilitación de barrios:

Desarrollo de la estrategia de regeneración urbana “Madrid Regenera” que abarca la rehabilitación del parque edificado (Plan MADRE), remodelación de espacios públicos, la producción energética local, la movilidad verde y de proximidad, la gestión de agua y materiales y la renaturalización de la ciudad.

23.- Fomento de sistemas de climatización eficiente de bajas emisiones:

Impulso de la implementación de mejoras tecnológicas de instalaciones de calor, climatización y agua caliente sanitaria. Algunas actuaciones concretas son la eliminación del uso del carbón en 2020, el fomento de los sistemas de climatización eficientes, la regulación del uso de la biomasa en todo el municipio y el desarrollo de redes de calor y frío de alta eficiencia.

24.- Generación distribuida y uso de energías renovables

Incorporación de tecnologías menos contaminantes, reducción de la demanda a través de la eficiencia energética e incremento de la autosuficiencia mediante una generación distribuida que acerque la producción de energía con renovables a los centros de consumo.

Establecimiento de una hoja de ruta para el desarrollo de las Energías Renovables, revisión de las bonificaciones del Impuesto de Bienes Inmuebles por instalaciones de energía solar, e impulso a la exploración y explotación del potencial energético del subsuelo.

25.- Reducción de emisiones en la gestión de residuos:

Implementación y optimización de los procesos de gestión de residuos en el complejo tecnológico de Valdemingómez con el propósito de reducir las emisiones derivadas de su funcionamiento.

Incremento de la recuperación de materiales con nuevas líneas de selección, puesta en marcha de instalaciones de compostaje complementadas con la progresiva implementación de la recogida selectiva de la fracción orgánica en el municipio. Mejora del rendimiento de la Planta de Tratamiento de Biogás (PTB) incrementando la producción de biometano, y aprovechamiento energético con generación eléctrica del biogás no destinado a inyección a red.

26.- Monitorización energética y gestión de consumos de instalaciones municipales:

Impulsar la transición hacia un modelo energético basado en la eficiencia y la transparencia de información promoviendo sistemas para la monitorización remota de consumos energéticos. Gestión de información asociada al control y medición de los consumos, seguimiento de medidas, inversión de ahorro y eficiencia energética, así como a acciones de divulgación y formación. Creación de una comisión municipal de eficiencia energética y energías renovables, de una Plataforma abierta de consumos en edificios municipales (objetivo 80 % del consumo monitorizado) y seguimiento específico de servicios o instalaciones municipales de alto consumo energético.

27.- Intervenciones en edificios e instalaciones municipales:

Impulso hacia un modelo de emisiones cero a través de la eficiencia energética y el uso de energías renovables actuando en los edificios de titularidad municipal y en instalaciones con consumo energético en vía pública. Desarrollo de programas de mejora de instalaciones de climatización, implantación de sistemas fotovoltaicos, plataformas tecnológicas de comunicación, implantación de sistemas de gestión ambiental y gestión energética y cumplimiento del compromiso municipal de edificios de consumo casi nulo-energía positiva.

28.- Incorporación de criterios de sostenibilidad en la contratación municipal:

Creación de un marco jurídico en la administración municipal que permita a los órganos de contratación incluir criterios orientados a la reducción de las emisiones contaminantes, la mitigación y adaptación al cambio climático y promuevan el desarrollo sostenible en general.

Elaboración de Instrucciones Técnicas y de Contratación sostenible, inclusión de cláusulas en los pliegos de contratación, e implantación de medidas de seguimiento del cumplimiento de estas medidas.

Adaptación al cambio climático, cuyo fin son planificar estrategias de adaptación y soluciones basadas en la naturaleza (vulnerabilidad y resiliencia frente al cambio climático):

29.- Adaptación al Cambio Climático. Madrid + Natural. Soluciones basadas en la naturaleza:

Desarrollo del programa Madrid + Natural como instrumento para incrementar la resiliencia urbana frente al Cambio Climático, Impulso de proyectos e implementación de soluciones basadas en la naturaleza con objetivo de adaptar la ciudad a las amenazas ambientales derivadas de las alteraciones climáticas. Intervenciones en edificios, barrios y en las grandes infraestructuras verdes y azules de la ciudad.

Sensibilización ciudadana y colaboración con otras administraciones, sobre los problemas de contaminación atmosférica y el cambio climático y creación de canales y fórmulas de colaboración entre administraciones con el fin de construir entre todos una ciudad sostenible y saludable:

30.- Sensibilización ambiental y cooperación con otras administraciones:

Las acciones de sensibilización y comunicación recorren de manera transversal las diversas líneas de actuación del Plan, con el objetivo de aumentar la conciencia sobre la necesidad de cambiar las pautas a nivel ciudadano para contribuir a la mejora de la calidad del aire y la lucha contra el cambio climático en la ciudad. Al tiempo, se requiere la colaboración de otras administraciones, tanto de los Ayuntamientos vecinos, como de la Comunidad de Madrid y del Gobierno del Estado, para hacer las modificaciones legales y adoptar las medidas de incentivo precisas para promover entre todos de forma coordinada y sinérgica, el necesario cambio de hábitos de los ciudadanía. La colaboración de otras Administraciones es imprescindible para conseguir los objetivos del Plan A de Madrid, de la misma forma que la implementación de las políticas ambientales y la consecución de los objetivos nacionales de reducción de emisiones de gases de efecto invernadero y de reducción de contaminantes atmosféricos, requiere la cooperación de los pueblos y ciudades de España. Como ha puesto de manifiesto la Comisión Europea en el Dictamen motivado por incumplimiento de la legislación de calidad del aire, hacen falta acciones eficaces a escala nacional. Es urgente la modificación del ordenamiento jurídico español que favorece fiscalmente a los vehículos diesel (tanto el impuesto de matriculación como el impuesto sobre hidrocarburos) así como la puesta en marcha de una ambiciosa línea de ayudas para la renovación del parque circulante, especialmente dirigida al sector de la logística urbana. Igualmente, la aplicación de medidas efectivas sobre el acceso de vehículos en vías de titularidad estatal (como la implantación de carril BUS-VAO-ECO), el control de las emisiones contaminantes de vehículos (ITV) y el desarrollo de energías renovables requieren actuaciones inmediatas y un marco legislativo nacional y regional que permita que las ciudades, y en consecuencia el país, puedan alcanzar sus objetivos de calidad del aire y cambio climático.

Para realizar el seguimiento del Plan se definen, tres tipos de indicadores: de impacto, asociados a los objetivos generales, indicadores específicos, que evalúan la incidencia real de las acciones ejecutadas en su conjunto y de proceso, que valoran el grado de ejecución de las medidas. El cumplimiento del Plan, sobre la información obtenida de los indicadores citados, se analizará y evaluará periódicamente para conocer tanto el grado de progreso en la implementación de las medidas, como el efecto de las mismas sobre la calidad del aire, la mitigación de emisiones de gases de efecto invernadero y la adopción de medidas de adaptación al cambio climático.

10. Conclusiones finales y líneas futuras

Se ha cumplido con la lista de objetivos del proyecto aplicadas a la metodología CRISP-DM para proyectos de data science, comprendiendo el estudio del arte y los datos ofrecidos, y preparándolos para la obtención de un modelo predictivo, el cual ha sido evaluado y desplegado. Finalmente, tomando un sistema de recomendaciones externo se han evaluado si las medidas de alerta son lo suficientemente robustas para garantizar la lógica del negocio.

El análisis de los datos ha permitido observar claramente una tendencia al descenso de los contaminantes a partir de 2012, que fue el primer año en el que se recogieron los resultados del Plan de Calidad del Aire de 2011. Los resultados son muy similares a los evaluados en el Plan A en referencia al plan anterior. Sin embargo, las medidas de acción actuales parecen no alcanzar los valores esperados para el 2017 para las variables NO_x y PM, por lo que es necesario modificar o establecer nuevas medidas restrictivas para evitar que esto suceda. No obstante, el nuevo Plan A que entró en vigor en 2016 parece suficientemente robusto en forma de mejora continua como para modificar esta tendencia.

Ha sido de gran dificultad la manipulación y creación de las estructuras necesarias para el análisis de las variables, pues en ocasiones dichos ficheros no tenían una misma estructura definida o faltaban algunos datos. Perdiendo una gran cantidad de tiempo en obtener información coherente y bien estructurada. También en la búsqueda, lectura y comprensión de los diferentes decretos, protocolos y planes de medida.

Hubiera sido de gran utilidad estudiar hasta qué punto la meteorología afecta a estos resultados así como la densidad de tráfico dentro de la ciudad. Con lo que se podrían valorar medidas también cuando el tráfico sea excesivo o el tiempo sea perjudicial. Así como indagar más en el estudio de modelos autoregresivos y sus aplicaciones.

Recientemente se ha publicado en el portal de datos de Ayto. de Madrid las mediciones correspondientes al primer cuatrimestre. Debido a que se ha publicado en el cierre de este trabajo no ha sido posible valorar dicha información. No obstante sería recomendable contrastar los resultados obtenidos con los reales y valorar la relevancia de este trabajo.

11. Referencias y bibliografía

- [1] [Aplicación de la metodología CRISP-DM a un proyecto de minería de datos en el entorno universitario](http://e-archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf?sequence=1), Víctor Galán, http://e-archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf?sequence=1, Último acceso: 06-07-2017
- [2] [Datos Abiertos, Ayuntamiento de Madrid](http://datos.madrid.es/portal/site/egob), <http://datos.madrid.es/portal/site/egob>, Último acceso: 06-07-2017
- [3] [Datos Abiertos, Agencia Estatal de Meteorología – AEMET Gobierno de España](http://www.aemet.es/es/datos_abiertos/AEMET_OpenData), http://www.aemet.es/es/datos_abiertos/AEMET_OpenData, Último acceso: 06-07-2017
- [4] [R Core Team \(2015\). R: A language and environment for statistical computing](https://www.r-project.org/), R Foundation for Statistical Computing, Vienna, Austria, <https://www.r-project.org/>, Último acceso: 06-07-2017
- [5] [Tutorial de programación en R, Universidad de Granada](http://www.ugr.es/~javiepp/master_files/Tutorial%20de%20R.pdf), http://www.ugr.es/~javiepp/master_files/Tutorial%20de%20R.pdf, Último acceso: 06-07-2017
- [6] [Manuales de R en CRAN](https://cran.r-project.org/manuals.html), <https://cran.r-project.org/manuals.html>, Último acceso: 06-07-2017
- [7] [Anaconda, Continuum Analytics](https://www.continuum.io/Downloads), <https://www.continuum.io/Downloads>, Último acceso: 06-07-2017
- [8] [Jupyter Documentation, Release 4.1.1 alpha](https://media.readthedocs.org/pdf/jupyter/latest/jupyter.pdf), <https://media.readthedocs.org/pdf/jupyter/latest/jupyter.pdf>, Último acceso: 06-07-2017
- [9] [Shinyapps, RStudio](https://shiny.rstudio.com/articles/shinyapps.html), <https://shiny.rstudio.com/articles/shinyapps.html>, Último acceso: 06-07-2017
- [10] [Aprendiendo a usar github](http://conociendogithub.readthedocs.io/en/latest/data/dinamica-de-uso/), <http://conociendogithub.readthedocs.io/en/latest/data/dinamica-de-uso/>, Último acceso: 06-07-2017
- [11] [Repositorio github Análisis de la Contaminación en la Ciudad de Madrid](https://github.com/alvarogg777/Contaminacion-Madrid), Álvaro Gallego, <https://github.com/alvarogg777/Contaminacion-Madrid>, Último acceso: 06-07-2017
- [12] [Docker docs](https://docs.docker.com/), <https://docs.docker.com/>, Último acceso: 06-07-2017
- [13] [Container Contaminacion-Madrid](https://hub.docker.com/r/alvarogg777/docker-anaconda-r-tfg/), Álvaro Gallego, <https://hub.docker.com/r/alvarogg777/docker-anaconda-r-tfg/>, Último acceso: 06-07-2017
- [14] [Directiva 2008-50-CE Calidad del Aire](http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/Directiva_2008-50-CE_calidad_del_aire_tcm7-439328.pdf), http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/Directiva_2008-50-CE_calidad_del_aire_tcm7-439328.pdf, Último acceso: 06-07-2017

- [15] [Real Decreto 1073/2002](http://noticias.juridicas.com/base_datos/Admin/rd1073-2002.html), http://noticias.juridicas.com/base_datos/Admin/rd1073-2002.html, Último acceso: 06-07-2017
- [16] [Agencia para Sustancias Tóxicas y el Registro de Enfermedades. \(ATSDR\). 2016. Reseña Toxicológica del Etilbenceno \(versión actualizada\) . Atlanta, GA: Departamento de Salud y Servicios Humanos de los EE.UU., Servicio de Salud Pública.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307733/), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307733/>, Último acceso: 06-07-2017
- [17] [Fichas internacionales de seguridad química, INSHT](http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/FISQ/Ficheros/401a500/nspn0490.pdf), <http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/FISQ/Ficheros/401a500/nspn0490.pdf>, Último acceso: 06-07-2017
- [18] [Agencia para Sustancias Tóxicas y el Registro de Enfermedades. \(ATSDR\). 2016. Reseña Toxicológica del n-hexano. Atlanta, GA: Departamento de Salud y Servicios Humanos de los EE.UU., Servicio de Salud Pública.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307733/), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307733/>, Último acceso: 06-07-2017
- [19] [Vector Autorregresivo Models, EUI Working Papers, Department of Economics, ECO 2011/30](http://cadmus.eui.eu/bitstream/handle/1814/19354/ECO_2011_30.pdf?sequence=1&isAllowed=y), http://cadmus.eui.eu/bitstream/handle/1814/19354/ECO_2011_30.pdf?sequence=1&isAllowed=y, Último acceso: 06-07-2017
- [20] [Package 'vars', CRAN R Project](https://cran.r-project.org/web/packages/vars/vars.pdf), <https://cran.r-project.org/web/packages/vars/vars.pdf>, Último acceso: 06-07-2017
- [21] [Reducing Data Dimension, April 2005, Tom M. Mitchell, Carnegie Mellon University](http://www.cs.cmu.edu/~guestrin/Class/10701-S05/slides/dimensionality.pdf), <http://www.cs.cmu.edu/~guestrin/Class/10701-S05/slides/dimensionality.pdf>, Último acceso: 06-07-2017
- [22] [An Introductory Study on Time Series Modeling and Forecasting, Ratnadip Adhikari, R. K. Agrawal](https://gejza.nipax.cz/_media/stochasticke_procesy/1302.6613.pdf), https://gejza.nipax.cz/_media/stochasticke_procesy/1302.6613.pdf, Último acceso: 06-07-2017
- [23] [Forecasting Accuracy Evaluation of Tourist Arrivals: Evidence from Parametric and Non-Parametric Techniques](http://www.up.ac.za/media/shared/61/WP/wp_2015_52.zp61700.pdf), July 2015, University of Pretoria, http://www.up.ac.za/media/shared/61/WP/wp_2015_52.zp61700.pdf, Último acceso: 06-07-2017
- [24] [La Directiva 2001/81/CE, del Parlamento Europeo y del Consejo, de 23 de octubre de 2001, sobre techos nacionales de emisión de determinados contaminantes atmosféricos](http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/emisiones/pol-med/iniciativas_europeas.aspx), http://www.mapama.gob.es/es/calidad-y-evaluacion-ambiental/temas/atmosfera-y-calidad-del-aire/emisiones/pol-med/iniciativas_europeas.aspx, Último acceso: 06-07-2017
- [25] [Real Decreto 102/2011 del 28 de enero](https://www.boe.es/boe/dias/2011/01/29/pdfs/BOE-A-2011-1645.pdf), <https://www.boe.es/boe/dias/2011/01/29/pdfs/BOE-A-2011-1645.pdf>, Último acceso: 06-07-2017
- [26] [Calidad del Aire, Madrid 2016, Dirección General de Sostenibilidad y Control Ambiental, Ayto. de Madrid](http://www.mambiente.munimadrid.es/opencms/export/sites/default/calaire/Anexos/Memoria2016.pdf), <http://www.mambiente.munimadrid.es/opencms/export/sites/default/calaire/Anexos/Memoria2016.pdf>, Último acceso: 06-07-2017
- [27] [Plan de Calidad del Aire 2011-2015, Dirección General de Sostenibilidad y Control Ambiental, Ayto. de Madrid](http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20(1).ppt), [http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20\(1\).ppt](http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20(1).ppt), Último acceso: 06-07-2017
- [28] [Plan de Calidad del Aire 2011-2015, Dirección General de Sostenibilidad y Control Ambiental, Ayto. de Madrid](http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20(2).ppt), [http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20\(2\).ppt](http://www.madrid.es/UnidadesDescentralizadas/UDCMedios/noticias/2011/12Diciembre/15Jueves/NotasPrensa/AlcaldeBotellaCalidadAire/ficheros/Plan%20de%20Calidad%20del%20Aire%20(2).ppt), Último acceso: 06-07-2017
- [29] [Plan A de Calidad del Aire y Cambio Climático de la Ciudad de Madrid, Abril 2017, Dirección General de Sostenibilidad y Control Ambiental, Ayto. de Madrid](http://transparencia.madrid.es/UnidadesDescentralizadas/Sostenibilidad/CalidadAire/Ficheros/PlanA), <http://transparencia.madrid.es/UnidadesDescentralizadas/Sostenibilidad/CalidadAire/Ficheros/PlanA>

- [CalidadAire042017.pdf](#), Último acceso: 06-07-2017
- [30] Hyndman RJ (2016). *_forecast: Forecasting functions for time series and linear models_*. R package version 7.3, <URL: <http://github.com/robjhyndman/forecast>>, Último acceso: 06-07-2017
- [31] Hyndman RJ and Khandakar Y (2008). Automatic time series forecasting: the forecast package for R. *_Journal of Statistical Software_*, *26*(3), pp. 1-22. <URL: <http://www.jstatsoft.org/article/view/v027i03>>
- [32] Max Kuhn. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2017). *caret: Classification and Regression Training*. R package version 6.0-76. <https://CRAN.R-project.org/package=caret>, Último acceso: 06-07-2017
- [33] Rob J Hyndman (2017). *fma: Data Sets from "Forecasting: Methods and Applications"* by Makridakis, Wheelwright & Hyndman (1998). R package version 2.3. <https://CRAN.R-project.org/package=fma>, Último acceso: 06-07-2017
- [34] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009, Último acceso: 06-07-2017
- [35] Hadley Wickham (2017). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.2.0. <https://CRAN.R-project.org/package=stringr>, Último acceso: 06-07-2017
- [36] Hadley Wickham (2011). The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software*, 40(1), 1-29. URL <http://www.jstatsoft.org/v40/i01/>
- [37] Hadley Wickham (2015). *pryr: Tools for Computing on the Language*. R package version 0.1.2. <https://CRAN.R-project.org/package=pryr>, Último acceso: 06-07-2017
- [38] Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2017). *shiny: Web Application Framework for R*. R package version 1.0.3. <https://CRAN.R-project.org/package=shiny>, Último acceso: 06-07-2017
- [39] Winston Chang (2016). *shinythemes: Themes for Shiny*. R package version 1.1.1. <https://CRAN.R-project.org/package=shinythemes>, Último acceso: 06-07-2017
- [40] Simon Urbanek (2017). *rJava: Low-Level R to Java Interface*. R package version 0.9-9. <http://www.rforge.net/rJava/>, Último acceso: 06-07-2017
- [41] Hornik K, Buchta C and Zeileis A (2009). Open-Source Machine Learning: R Meets Weka. *_Computational Statistics_*, *24*(2), pp. 225-232. doi: 10.1007/s00180-008-0119-7 (URL: <http://doi.org/10.1007/s00180-008-0119-7>)
- [42] Witten IH and Frank E (2005). *_Data Mining: Practical Machine Learning Tools and Techniques_*, 2nd edition. Morgan Kaufmann, San Francisco, Último acceso: 06-07-2017
- [43] Kurt Hornik (2017). *RWekajars: R/Weka Interface Jars*. R package version 3.9.1-3. <https://CRAN.R-project.org/package=RWekajars>, Último acceso: 06-07-2017
- [44] Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. *R News* 2(3), 7-10. URL <https://CRAN.R-project.org/doc/Rnews/>, Último acceso: 06-07-2017
- [45] Peter Ellis (2016). *ggseas: 'stats' for Seasonal Adjustment on the Fly with 'ggplot2'*. R package version 0.5.1. <https://CRAN.R-project.org/package=ggseas>, Último acceso: 06-07-2017
- [46] Adrian Trapletti and Kurt Hornik (2017). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-41, Último acceso: 06-07-2017
- [47] David Stoffer (2016). *astsa: Applied Statistical Time Series Analysis*. R package version 1.7. <https://CRAN.R-project.org/package=astsa>, Último acceso: 06-07-2017
- [48] Frank E Harrell Jr, with contributions from Charles Dupont and many others. (2017). *Hmisc: Harrell Miscellaneous*. R package version 4.0-3. <https://CRAN.R-project.org/package=Hmisc>, Último acceso: 06-07-2017
- [49] JJ Allaire (2017). *rsconnect: Deployment Interface for R Markdown Documents and Shiny Applications*. R package version 0.8. <https://CRAN.R-project.org/package=rsconnect>, Último acceso: 06-07-2017

- [50] Sébastien Le, Julie Josse, François Husson (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1), 1-18. 10.18637/jss.v025.i01
- [51] Alboukadel Kassambara and Fabian Mundt (2017). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.4. <https://CRAN.R-project.org/package=factoextra>, Último acceso: 06-07-2017
- [52] Bernhard Pfaff (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software* 27(4). URL <http://www.jstatsoft.org/v27/i04/>
- [53] Pfaff, B. (2008) Analysis of Integrated and Cointegrated Time Series with R. Second Edition. Springer, New York. ISBN 0-387-27960-1
- [54] Pfaff, B. (2008) Analysis of Integrated and Cointegrated Time Series with R. Second Edition. Springer, New York. ISBN 0-387-27960-1
- [55] Taiyun Wei and Viliam Simko (2016). corrplot: Visualization of a Correlation Matrix. R package version 0.77. <https://CRAN.R-project.org/package=corrplot>, Último acceso: 06-07-2017
- [56] Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5
- [57] Douglas Bates and Martin Maechler (2017). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.2-10. <https://CRAN.R-project.org/package=Matrix>, Último acceso: 06-07-2017
- [58] Rmetrics Core Team, Diethelm Wuertz, Tobias Setz, Yohan Chalabi, Martin Maechler and Joe W. Byers (2015). timeDate: Rmetrics - Chronological and Calendar Objects. R package version 3012.100. <https://CRAN.R-project.org/package=timeDate>, Último acceso: 06-07-2017
- [59] Achim Zeileis, Friedrich Leisch, Kurt Hornik and Christian Kleiber (2002). strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *Journal of Statistical Software*, 7(2), 1-38. URL <http://www.jstatsoft.org/v07/i02/>
- [60] Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
- [61] Achim Zeileis, Yves Croissant (2010). Extended Model Formulas in R: Multiple Parts and Multiple Responses. *Journal of Statistical Software* 34(1), 1-13. URL <http://www.jstatsoft.org/v34/i01/>
- [62] Achim Zeileis (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software* 11(10), 1-17. URL <http://www.jstatsoft.org/v11/i10/>
- [63] Ruben Thoplan, Simple v/s Sophisticated Methods of Forecasting for Mauritius Monthly Tourist Arrival Data, *International Journal of Statistics and Applications*, Vol. 4 No. 5, 2014, pp. 217-223. doi: 10.5923/j.statistics.20140405.01

12. Índice de tablas

Tabla 1. Identificación de las estaciones meteorológicas.....	12
Tabla 2. Identificación de las estaciones de medición de gases.....	12
Tabla 3. Listado de las variables atmosféricas consideradas.....	15
Tabla 4. Valores límites para la protección de la salud humana.....	15
Tabla 5.1 Valores límites para la protección de la vegetación.....	16
Tabla 5.2 Valores límites para la protección de la salud humana R.D 102/2011.....	16
Tabla 6.1 Umbrales de información y de alerta.....	17
Tabla 6.2 Umbrales de alerta R.D 102/2011.....	17
Tabla 7. Valores límite SO₂.....	18
Tabla 8. Valores límite CO.....	20
Tabla 9. Valores límite PM2.5.....	28
Tabla 10. Valores límite PM10.....	31
Tabla 11.1. Valores límite de contaminantes que producen O₃.....	34

Tabla 11.2. Umbrales de O₃	
Tabla 12. Matriz de correlación de variables contaminantes en el mismo instante.....	52
Tabla 13.1. Matriz de correlación de variables contaminantes 1 hora antes.....	54
Tabla 13.2. Matriz de correlación de variables contaminantes 4 horas antes	
Tabla 13.3. Matriz de correlación de variables contaminantes 8 horas antes	
Tabla 14.1. Matriz de correlación de variables contaminantes 1 hora después.....	55
Tabla 14.2. Matriz de correlación de variables contaminantes 4 hora antes	
Tabla 14.3. Matriz de correlación de variables contaminantes 8 hora antes	
Tabla 15. Estimadores de exactitud para la elección de lags, obtenidos mediante VARSelect.....	60
Tabla 16.1. Estimadores de exactitud (accuracy) mediante Holts-Winter.....	64
Tabla 16.2. Estimadores de exactitud (accuracy) mediante ARIMA	
Tabla 16.3. Estimadores de exactitud (accuracy) mediante NNAR	
Tabla 17.1 Valores contaminantes mensuales de los contaminantes cerca del límite del 2017..	66
Tabla 17.2 Valores contaminantes de los contaminantes cerca del límite del 2017 en promedio	
Tabla 18. Lista de gases respecto a contaminantes evaluados.....	72

14. Índice de figuras

Figura 1. Metodología CRISP-DM.....	8
Figura 2. Ejemplo representativo de la aplicación PCA y k-means.....	11
Figura 3. Boletín diario. Portal de Calidad del Aire.....	13
Figura 4. Consulta de datos. Portal de Calidad del Aire	13
Figura 5.1. Listado de las variables contaminantes monitorizadas	14
Figura 5.2. Representación gráfica de las mediciones en las estaciones no meteorológicas	
Figura 6.1. Agregado horario SO ₂ utilizando la media y mediana como estadísticos.....	18
Figura 6.2. Agregado semanal SO ₂ utilizando la media y mediana como estadísticos	
Figura 6.3. Agregado diario SO ₂ utilizando la media y mediana como estadísticos	
Figura 6.4. Agregado mensual SO ₂ utilizando la media y mediana como estadísticos	
Figura 7.1. Agregado horario CO utilizando la media y mediana como estadísticos.....	21
Figura 7.2. Agregado semanal CO utilizando la media y mediana como estadísticos	
Figura 7.3. Agregado diario CO utilizando la media y mediana como estadísticos	
Figura 7.4. Agregado mensual CO utilizando la media y mediana como estadísticos	
Figura 8.1. Agregado horario NO utilizando la media y mediana como estadísticos.....	23
Figura 8.2. Agregado semanal NO utilizando la media y mediana como estadísticos	
Figura 8.3. Agregado diario NO utilizando la media y mediana como estadísticos	
Figura 8.4. Agregado mensual NO utilizando la media y mediana como estadísticos	
Figura 9.1. Agregado horario NO ₂ utilizando la media y mediana como estadísticos.....	26
Figura 9.2. Agregado semanal NO ₂ utilizando la media y mediana como estadísticos	
Figura 9.3. Agregado diario NO ₂ utilizando la media y mediana como estadísticos	
Figura 9.4. Agregado mensual NO ₂ utilizando la media y mediana como estadísticos	
Figura 10.1. Agregado horario PM2.5 utilizando la media y mediana como estadísticos....	29
Figura 10.2. Agregado semanal PM2.5 utilizando la media y mediana como estadísticos	
Figura 10.3. Agregado diario PM2.5 utilizando la media y mediana como estadísticos	
Figura 10.4. Agregado mensual PM2.5 utilizando la media y mediana como estadísticos	
Figura 11.1. Agregado horario PM10 utilizando la media y mediana como estadísticos.....	31
Figura 11.2. Agregado semanal PM10 utilizando la media y mediana como estadísticos	

Figura 11.3. Agregado diario PM10 utilizando la media y mediana como estadísticos	
Figura 11.4. Agregado mensual PM10 utilizando la media y mediana como estadísticos	
Figura 12.1. Agregado horario O ₃ utilizando la media y mediana como estadísticos.....	34
Figura 12.2. Agregado semanal O ₃ utilizando la media y mediana como estadísticos	
Figura 12.3. Agregado diario O ₃ utilizando la media y mediana como estadísticos	
Figura 12.4. Agregado mensual O ₃ utilizando la media y mediana como estadísticos	
Figura 13.1. Agregado horario C ₆ H ₅ -CH ₃ utilizando la media y mediana como estadísticos..	37
Figura 13.2. Agregado semanal C ₆ H ₅ -CH ₃ utilizando la media y mediana como estadísticos	
Figura 13.3. Agregado diario C ₆ H ₅ -CH ₃ utilizando la media y mediana como estadísticos	
Figura 13.4. Agregado mensual C ₆ H ₅ -CH ₃ utilizando la media y mediana como estadísticos	
Figura 14.1. Agregado horario C ₆ H ₆ utilizando la media y mediana como estadísticos.....	40
Figura 14.2. Agregado semanal C ₆ H ₆ utilizando la media y mediana como estadísticos	
Figura 14.3. Agregado diario C ₆ H ₆ utilizando la media y mediana como estadísticos	
Figura 14.4. Agregado mensual C ₆ H ₆ utilizando la media y mediana como estadísticos	
Figura 15.1. Agregado horario C ₈ H ₁₀ utilizando la media y mediana como estadísticos.....	43
Figura 15.2. Agregado semanal C ₈ H ₁₀ utilizando la media y mediana como estadísticos	
Figura 15.3. Agregado diario C ₈ H ₁₀ utilizando la media y mediana como estadísticos	
Figura 15.4. Agregado mensual C ₈ H ₁₀ utilizando la media y mediana como estadísticos	
Figura 16.1. Agregado horario C ₆ H ₁₂ utilizando la media y mediana como estadísticos.....	45
Figura 16.2. Agregado semanal C ₆ H ₁₂ utilizando la media y mediana como estadísticos	
Figura 16.3. Agregado diario C ₆ H ₁₂ utilizando la media y mediana como estadísticos	
Figura 16.4. Agregado mensual C ₆ H ₁₂ utilizando la media y mediana como estadísticos	
Figura 17.1. Agregado horario C ₆ H ₁₄ utilizando la media y mediana como estadísticos.....	48
Figura 17.2. Agregado semanal C ₆ H ₁₄ utilizando la media y mediana como estadísticos	
Figura 17.3. Agregado diario C ₆ H ₁₄ utilizando la media y mediana como estadísticos	
Figura 17.3. Agregado mensual C ₆ H ₁₄ utilizando la media y mediana como estadísticos	
Figura 18. Plots de las variables atmosféricas.....	51
Figura 19. Matriz de correlación mensual mediante el coeficiente de correlación de Pearson....	52
Figura 20.1. Significación de dimensiones para variables contaminantes.....	56
Figura 20.2. Contribución de las variables contaminantes a dos dimensiones	
Figura 20.3. Contribución de las variables contaminantes a la dimensión 1	
Figura 20.4. Contribución de las variables contaminantes a la dimensión 2	
Figura 21.1. Significación de dimensiones para variables contaminantes y atmosféricas.....	58
Figura 21.2. Contribución de las variables a dos dimensiones	
Figura 21.3. Contribución de las variables a la dimensión 1	
Figura 21.4. Contribución de las variables a la dimensión 2	
Figura 22. Predicción CO en 2017.....	60
Figura 23. Predicción NO en 2017.....	61
Figura 24. Predicción NO ₂ en 2017.....	61
Figura 25. Predicción PM2.5 en 2017.....	61
Figura 26. Predicción PM10 en 2017.....	62
Figura 27. Predicción O ₃ en 2017.....	62
Figura 28. Predicción C ₆ H ₅ -CH ₃ en 2017.....	62
Figura 29. Predicción C ₆ H ₆ en 2017.....	63
Figura 30. Predicción C ₆ H ₁₂ en 2017.....	63
Figura 31. Predicción C ₆ H ₁₄ en 2017.....	63
Figura 32. Predicción SO ₂ en 2017.....	64

Figura 33. Predicción C₈H₁₀ en 2017.....	64
Figura 34. Comparativa de ajuste predictivo 2016 entre Holts, ARIMA, y NNAR.....	65
Figura 35. Predicción PM10 mediante Holts-Winter.....	65
Figura 36. Predicción PM10 obtenida del modelo VAR.....	66
Figura 37.1. Pantalla principal de la aplicación desde shinyapps.....	68
Figura 37.2. Hacemos click en Your Data	
Figura 37.2. Seleccionamos nuestro fichero "dfvarmaMonth.csv"	
Figura 38. Seleccionamos la columna a analizar y procedemos a la elección del algoritmo.....	69
Figura 39. Seleccionando Smoothing Method, Holts-Winter obtenemos el ajuste de predicción del año 2016.....	69
Figura 40. Seleccionando Neural Network obtenemos el ajuste de predicción del 2016.....	70
Figura 41. Seleccionando ARIMA Method, obtenemos el ajuste de predicción del 2016.....	70
Figura 42. Seleccionando VAR Model, obtenemos las predicciones de todas las variables en el 2017.....	71
Figura 43. Fuentes contaminantes y contribución de emisiones de NO_x por grupo SNAP.....	73
Figura 44. Fuentes contaminantes y contribución de emisiones de PM por grupo SNAP.....	74