

Análisis de contaminantes atmosféricos con R

Palabras clave: contaminación atmosférica, análisis de datos, series temporales, material particulado fino, variables meteorológicas, modelo predictivo

El crecimiento desmedido de la población y la ambición de la sociedad por mejorar su calidad de vida, han resultado en un creciente deterioro del ambiente. La atmósfera no es ajena a este deterioro que acaba empeorando la calidad del aire, impactando en la calidad de vida de las personas. Uno de los contaminantes más preocupantes en atmósferas urbanas son los aerosoles atmosféricos, suspensiones relativamente estables de partículas sólidas o líquidas en un gas y que presentan un diámetro menor a 10 μm . Distinto es a lo que refiere el término material particulado (PM), el cual define un conjunto de materia en estado sólido y/o líquido presente en suspensión en la atmósfera, excluyendo el agua pura. La presencia de este contaminante en la atmósfera tiene un impacto significativo en el desarrollo de la sociedad dado que afecta múltiples componentes con los que interactúa el ser humano, además de su salud, se destaca el efecto que tiene en los ecosistemas y en clima. El objetivo de este trabajo es demostrar la aplicabilidad de R en el análisis de datos de contaminantes atmosféricos, específicamente de material particulado fino ($\text{PM}_{2.5}$) a través de mi propia experiencia en el desarrollo de mi tesis de doctorado. Mi investigación se centró en el estudio de series temporales de concentración de $\text{PM}_{2.5}$, explorar su relación con variables meteorológicas y satelitales, y la creación de modelos predictivos para la concentración de $\text{PM}_{2.5}$. R fue una herramienta transversal y fundamental a lo largo de mi trabajo de investigación, abarcando desde la etapa de descarga de variables satelitales hasta la redacción de la tesis. Esta potencialidad se ve sustentada no solo en el hecho de que se trata de un software libre, sino también de que cuenta con una numerosa comunidad activa que desarrolla código que permite resolver problemáticas comunes. En primera instancia la librería KrigR me permitió descargar y preprocesar variables satelitales almacenadas en el Servicio de Cambio Climático de Copernicus (C3S),

específicamente obtenidas a partir del procedimiento de reanálisis (ERA5) del Centro Europeo de Pronósticos Meteorológicos a Plazo Medio (ECMWF). Una vez descargadas las variables, aproveche las funcionalidades de los paquetes tidyverse y lubridate para importar, limpiar, manipular y visualizar los datos (ggplot2). Además, exploré visualizaciones más avanzadas mediante el uso del paquete openair, cuyo propósito principal es el de analizar datos de calidad del aire, brindando funcionalidades específicas para el análisis de los mismos. Es pertinente añadir que muchos de los análisis de series temporales desarrollados pudieron ser ejecutados gracias a la aplicabilidad de múltiples paquetes, entre los cuales se destacan TSstudio, tseries, uroot, forecast y xts. Finalmente, el modelado predictivo lo lleve a cabo a través de la aplicación de los frameworks de modelado predictivo de tidymodels y caret, permitiéndome construir y comparar modelos predictivos. Cabe destacar que también aproveche las funcionalidades provistas por la librería thesisdown para redactar el documento final de la tesis. De esta forma busco exponer y comunicar la aplicabilidad de R para resolver múltiples problemas transversales en mi tesis doctoral, para incentivar y contagiar a otros investigadores para resolver sus problemas a partir de la aplicación de R.