

P-5

Los SIG en la caracterización de la exposición de la población residente a contaminantes en el aire

Urko Elozegui

Geógrafo Consultor SIG. Especialista en análisis de riesgos ambientales
umapak@gmail.com

INTRODUCCIÓN

La esencia de la salud ambiental es caracterizar la exposición a la que se ve sometida la población por factores de riesgo ambiental vehiculados por el aire, agua, suelo, alimentos y entorno culturizado (radiaciones electromagnéticas, ruido, ...)

Este es un reto que, lamentablemente, no se considera con todo el rigor que requiere, toda vez que, incluso existen mandatos legislativos al respecto. Sería el caso de la Red de Vigilancia Epidemiológica de la Comunidad de Madrid, constituida en el año 1996, que recoge entre los sistemas que componen la Red, los Sistemas de Información Medioambiental.

Estos sistemas, cuya gestión responde a las competencias de los departamentos medioambientales, tienen una gran vocación sanitaria si se es capaz de darle ese valor añadido que supone la caracterización de la exposición, en el tiempo y en el espacio: definir dónde, cuándo y qué grupos etarios se han visto expuestos a niveles superiores de los establecidos en la legislación o en las guías definidas por organismos internacionales, como puede ser la Organización Mundial de la Salud (OMS).

Se presenta cómo se han implementado las técnicas necesarias para monitorizar y analizar la calidad del aire desde el Área de Vigilancia de Riesgos Ambientales en Salud, de la Dirección General de Salud Pública, de la Comunidad de Madrid (en adelante AVRAS-CM), utilizando Sistema de Información Geográfica para crear indicadores de contaminación atmosférica y evaluar la exposición de la población a los contaminantes atmosféricos.

Este resumen tiene como objetivo proporcionar una visión general de los recursos, métodos y resultados que pueden utilizarse como sólida base de capacitación para futuras investigaciones y políticas de salud pública relacionadas con la calidad del aire en España.

METODOLOGÍA

La forma en que se obtiene la información y se gestionan las políticas de salud pública relacionadas con la calidad del aire ha evolucionado significativamente en los últimos diez años a pesar de la paralización de dichos estudios debido a la pandemia de COVID-19.

Estimar la exposición a contaminantes del aire ha sido una tarea cuyo condicionante principal desde su inicio ha sido la disponibilidad del dato y en segundo lugar, la capacidad técnica para analizar dichos datos.

En el ámbito de la ciencia de datos, la disponibilidad ha ido en paralelo al desarrollo de infraestructuras de vigilancia de calidad, sin las cuales sería imposible realizar un análisis significativo o extraer información valiosa. En Madrid existen al menos 3 infraestructuras que proveen datos de base para caracterizar la exposición:

Las Redes de Calidad del Aire, impulsadas merced a las políticas implementadas desde la Unión Europea, desempeñan un papel fundamental liderando la monitorización y promoviendo la mejora de la calidad del aire en nuestras ciudades.

- La **Red de Calidad del Aire de la Comunidad de Madrid** cuenta en 2024 con 28 estaciones fijas y dos móviles.
- El **Sistema Integral de Vigilancia, Predicción e Información de la Contaminación Atmosférica de la Ciudad de Madrid** cuenta con una red de 25 estaciones fijas de control de calidad del aire.

La infraestructura sobre Aerobiología, Polen y Esporas de la CM, con la **Red Palinológica de la Comunidad de Madrid (Red Palinocam)** como parte integral de esta infraestructura, se dedica al estudio de las partículas biológicas presentes en el aire, pólenes y esporas, entre otras. La particularidad de esta red radica en el meticuloso análisis que se somete a los granos de polen al microscopio, con recuento e identificación individual

de cada uno de los tipos polínicos. Este servicio es de gran utilidad para las personas alérgicas, permitiéndoles tomar medidas preventivas adecuadas y para el sistema sanitario asistencial.

Además, la exposición de las emisiones de las actividades industriales en la salud pública se caracteriza con un enfoque particular. El **Registro Estatal de Emisiones y Fuentes Contaminantes (PRTR-España)** es una herramienta que obliga a las industrias a registrar anualmente todas las emisiones y recoge información, entre otras, sobre las sustancias contaminantes a la atmósfera, clave en este contexto. Desde AVRAS-CM se identifican las principales áreas industriales, los residuos generados y las emisiones al aire poniendo especial énfasis en las zonas de salud más cercanas con el objetivo de proporcionar a los profesionales sanitarios la máxima información posible sobre los factores que podrían estar afectando la salud de sus pacientes.

La ciencia de datos combina estadísticas, programación, modelización y conocimiento del dominio, permitiendo tomar decisiones informadas basadas en datos. En la actualidad la capacitación en esta materia se ha vuelto esencial, ya que se genera información a un ritmo sin precedentes y la habilidad para interpretar y utilizar estos datos de manera efectiva proporciona una ventaja competitiva significativa en todos los ámbitos.

El equipo humano del AVRAS-CM ha combinado el conocimiento de perfiles expertos en Contaminación Atmosférica y Aerobiología con la experiencia en Sistemas de Información Geográfica que ha permitido operar distintos sistemas de monitorización que han satisfecho una creciente demanda de datos, información y conocimiento.

La construcción de indicadores requiere, además de datos provenientes de las infraestructuras mencionadas, valores de población por grupos de edad para el año de referencia, tanto a nivel municipal como por sección censal¹. Los valores límites se definen para la protección de la salud humana en las normativas vigentes. También se requiere información

geográfica, es decir, geometrías de ámbitos como Municipios o Zonas homogéneas *ad-hoc* definidas según la legislación vigente y cuadrículas estadísticas homogéneas de alcance europeo (GISCO-INSPIRE).

A partir de los datos observados o declarados, se predice el comportamiento utilizando procedimientos de interpolación mediante dos métodos: la media ponderada por el inverso de la distancia o IDW, utilizada para la representación cartográfica, y el método estadístico de *krigeado* ordinario, que tiene en cuenta la autocorrelación espacial de la variable a interpolar y minimiza el error estándar. Este último se utiliza para el cálculo de la población residente en las zonas con niveles elevados de contaminación. Ambos modelos producen mapas de isoconcentraciones.

Con los datos proporcionados por el segundo modelo de predicción y mediante la superposición de las capas geográficas de concentraciones esperadas y de distribución de población por área pequeña y grupos de edad, se obtienen estimaciones de población residente. Estas estimaciones representan la cantidad total de población expuesta desagregada en grupos etarios, con énfasis en los especialmente vulnerables.

Como limitaciones del procedimiento, además del error del modelo predictivo de los niveles de inmisión, hay que señalar que la medida de la exposición se realiza solo en la localización y no se consideran variables como los ritmos de exposición (periodos, duración, frecuencia, etc.).

EVOLUCIÓN, RESULTADOS Y DISCUSIÓN

La evolución de las caracterizaciones de exposición, estrechamente ligadas a la creciente disponibilidad de datos de alta calidad, es cuantificable en el número de observaciones.

El primer informe de calidad del aire (ICA) se elaboró en 2004 y se replicó en 2007; a partir de 2009 se comenzó a elaborar información que incluía gráficas de evolución de

Tabla 1. Número de observaciones anuales de contaminantes en el aire disponibles para las caracterizaciones desarrolladas por AVRAS-CM en el periodo 2004-2022

Número de observaciones anuales de contaminantes en el aire disponibles para las caracterizaciones desarrolladas por AVRAS-CM en el periodo 2004-2022																		
2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
30660	NA	NA	30660	30660	30660	108974	108974	108974	108974	108974	130906	130906	130906	130906	133220	133220	133220	133220
NA	NA	NA	NA	NA	NA	108378	106704	108473	105820	105032	101699	99827	103490	107325	106345	101395	107406	108351
1138	1061	1155	2837	2830	2632	2548	2504	2525	2608	2789	2873	2856	24855	24623	24670	26607	26832	27933

Fuentes: www.comunidad.madrid/gobierno/datos-abiertos (CALAIRE, POLEN), datos.madrid.es (CALAIRE) y prtr-es.es (PRTR). Elaboración propia.

los contaminantes, se informaba de las superaciones de ozono en verano y de NO₂ en invierno, y en 2010 se inició la elaboración de mapas de exposición de la población a los contaminantes, adquiriendo una gran solidez a partir de 2015.

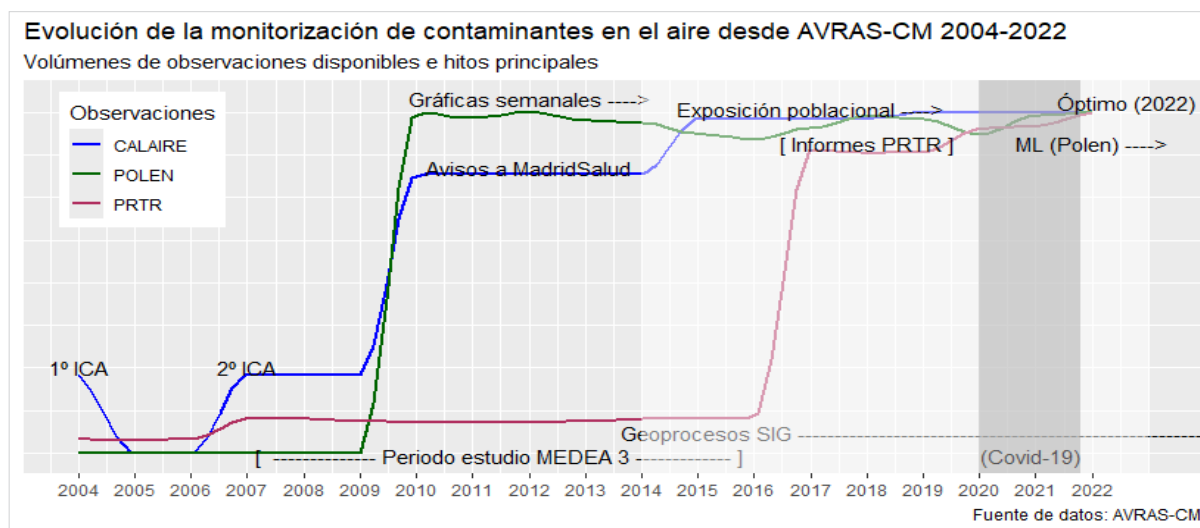
La Red Palinocam, creada en 1993, informa a través de varios medios, incluyendo correo electrónico y SMS, elabora mapas de isoconcentraciones para conocer e identificar las características de la población expuesta y recientemente ha incorporado *Machine Learning* (ML) para modelizar la predicción.

Por último, la base de datos del PRTR se utilizó inicialmente en el marco del proyecto MEDEA 3

posibilitando la consolidación de Informes anuales sistematizados entre 2016 y 2020.

Los impulsos de proyectos de investigación como MEDEA 3² y las gestiones para colaborar tanto con el ámbito asistencial de Madrid Salud como con la Agencia de Seguridad y Emergencias Madrid 112 (ASEM 112), además de la valiosa colaboración académica y ciudadana, han permitido evolucionar las capacidades de caracterización hasta un óptimo alcanzado una vez superada la pandemia del COVID-19. En resumen, la actividad del AVRAS ha logrado, con los medios obrantes y gracias al equipo de profesionales, extraer la información disponible según el estado del arte, la disponibilidad de datos y la legislación vigente en cada momento desde la puesta en marcha de las caracterizaciones.

Figura 1: Evolución de la monitorización de contaminantes en el aire desde AVRAS-CM 2004-2022. Volúmenes de observaciones disponibles e hitos principales



Fuente: AVRAS-CM. Elaboración propia.

Sin embargo, en la última década hemos sido testigos de una revolución en el manejo y análisis de datos. Conceptos como *Business Intelligence* (BI) y *Big Data* han evolucionado significativamente, pasando de ser herramientas utilizadas por empresas punteras a ser esenciales para todo tipo de disciplinas, entre las que no es ajena la Salud Pública. Al mismo tiempo, hemos visto el desarrollo de disciplinas como el *Data Mining* y el *Machine Learning*. Estas técnicas permiten extraer información valiosa traducible en conocimiento a partir de grandes conjuntos de datos y utilizarla para mejorar procesos y servicios.

Las nuevas políticas de Datos Abiertos (*Open Data*) favorecen enormemente la integración de todos los datos necesarios en la caracterización, democratizándolos y garantizando la calidad de estos.

El Internet de las Cosas (*IoT*), aunque sin la relevancia mediática que cabía esperar hace un lustro, también ha tenido un impacto significativo, permitiendo la recopilación de datos en tiempo real de una variedad de dispositivos y sensores. IoT aporta como claves futuras oportunidades para la captura de datos, como veremos más adelante.

No obstante, es la Inteligencia Artificial (IA) en particular la que ha experimentado un crecimiento explosivo, con sorprendentes aplicaciones. Los modelos de lenguaje grande o *LLM* (*Large Language Models*, en inglés), como el archiconocido Chat GPT, han revolucionado la forma en que interactuamos con la tecnología. Se trata de tecnologías que permiten altas capacidades humanas a partir de conocimientos básicos ya que estos modelos son capaces de entender

y generar texto de manera coherente, lo que los hace enormemente útiles en una variedad de aplicaciones, como la interacción con múltiples fuentes de información en modelos complejos, más allá de los desafíos éticos y de privacidad asociados con su uso, aspecto que este resumen no pretende abordar.

CLAVES FUTURAS

La teledetección posibilita recopilar información sobre un objeto sin estar en contacto físico directo con él. Una sola imagen de satélite provee millones de puntos de datos, en forma de observaciones a nivel de píxel, de una determinada resolución, de la radiancia espectral de la columna troposférica.

El valor espectral es sometido a cálculos teóricos, suposiciones sobre el sistema tierra-atmósfera y algoritmos informáticos que permiten inferir valores de parámetros geofísicos de los contaminantes en superficie.

Las observaciones satelitales posibilitan la única cobertura verdaderamente global de muchos contaminantes atmosféricos que se pueden medir desde el espacio, permitiendo un análisis de cualquier área de interés, incluso donde no existen redes de calidad de aire.

Desde agosto de 2023 la nueva misión estadounidense *TEMPO*, (*Tropospheric Emissions: Monitoring of Pollution*), ha logrado mediciones horarias de dióxido de nitrógeno con una resolución inferior a 10 km² para toda América del Norte³, confiriendo a la teledetección la capacidad de competir de cara a cara con las redes de calidad del aire.

El acceso a los datos en abierto es más fácil año tras año, gracias a los respectivos portales online *NASA Giovanni* y *Copernicus Data Space Ecosystem*, entre otros, que además ofrecen herramientas de análisis y capacitación.

El Internet de las cosas (IoT) enfocado en el monitoreo ambiental tiene un potencial real de revolucionar la medición de la contaminación al conectar dispositivos y sensores a través de Internet. El IoT marca la diferencia incluyendo:

- sensores de coste reducido que miden niveles de contaminantes PM_{2.5}, NO_x y SO₂.
- monitorización en tiempo real que proporciona datos actualizados sobre la calidad del aire potencialmente en tantas áreas como despliegues se realicen.
- mediciones que pueden apoyar, mediante algoritmos avanzados, las observaciones de las redes de calidad del aire y la teledetección desde el ámbito de la

participación ciudadana, convirtiendo al ámbito privado y a la ciudadanía en recopiladores activos de datos mediante aplicaciones móviles y dispositivos personales, fomentando además la conciencia ambiental y la colaboración en la lucha contra la contaminación.

REFERENCIAS

1. Aránguez Ruiz E, Arribas García M, Aránguez Gilarranz J, Ordoñez Iriarte JM. Salud y territorio. Aplicaciones prácticas de los sistemas de información geográfica a la salud ambiental. Madrid. Sociedad Española de Sanidad Ambiental. Serie De aeribus, aquis et locis nº2. 2012.
2. Gandarillas A, Domínguez MF, Duque I, Cebrecos A, Aránguez E, Ordóñez JM, Cervigón P, Elosegí U. Tercer atlas de mortalidad y desigualdades socioeconómicas en la Comunidad de Madrid, 2008-2015. Consejería de Sanidad, Madrid, 2021.[actualizado en 2023; citado el 01/03/2024] Disponible en: <https://www.comunidad.madrid/publicacion/ref/50460>.
3. Center for Astrophysics | Harvard & Smithsonian News Release. TEMPO Instrument Captures Its First Images of Air Pollution Over Greater North America. [actualizado en 2023; citado el 01/03/2024] Disponible en: <https://www.cfa.harvard.edu/news/tempo-instrument-captures-its-first-images-air-pollution-over-greater-north-america>.