



Título de la Tesis Doctoral: Sistemas de computación para el estudio de biometeorología centrado en el almacenamiento y procesamiento de datos espaciales

Directores:

Ángel Marqués Mateu. Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría. Universitat Politècnica de València.

Laura Sebastiá Tarín. Departamento de Sistemas Informáticos y Computación. Universitat Politècnica de València.

Pablo Fernández de Arróyabe Hernáez. Departamento de Geografía, Urbanismo y Ordenación del Territorio. Universidad de Cantabria.

Resumen:

Los datos son parte fundamental para cualquier tipo de estudio o base de conocimiento. Hoy en día existen distintos paradigmas en el uso, tratamiento y análisis de los datos. La irrupción de diversas tecnologías para los sistemas de almacenamiento, el procesamiento y análisis de datos, dan respuesta a distintas problemáticas relacionadas con la variedad, volumen y velocidad de producción de los datos. En la actualidad, las fuentes de información son innumerables. Tecnologías relacionadas con la sensorización o la medición de fenómenos generan constantemente enormes cantidades de datos, que en muchos casos son publicados en línea para su explotación mediante *Web Service*. El concepto de *Open Data* es un ejemplo reciente de fuente de información de amplio espectro, que abarca múltiples tipologías de fenómenos, formatos y producción constante de datos, pero con una característica que lo distingue de otras fuentes de datos masivos: la posibilidad de reutilización por parte de terceros con “medios simples”. En este sentido, el origen de los datos es un aspecto clave para la calidad de cualquier trabajo y el éxito en la consecución de los objetivos planteados en base al uso de estos datos. La gestión de los datos es otro aspecto fundamental en proyectos cuyo objetivo depende de la explotación y análisis de dichos datos. La definición del sistema encargado del almacenamiento de los datos es necesaria para poder resolver los problemas que se deriven de un proyecto. Los datos por sí solos no generan conocimiento, y para llegar a extraer valor, es importante realizar una gestión adecuada para poder procesarlos y analizarlos. La tipología y la variabilidad de los datos suponen un gran reto en la gestión de la información. Por ello, en el presente estudio se establecerá un esquema de gestión de datos teniendo en cuenta la variabilidad, el volumen y la velocidad de producción de los mismos, aplicando técnicas de *Big Data*. El objetivo principal es realizar un estudio del comportamiento global de un índice Biometreológico, en particular la Diferencia de Oxígeno Atmosférico (conocida comúnmente por las siglas DOA), que tiene importantes repercusiones en distintos indicadores relacionados con la salud pública y el medio ambiente. En el trabajo se procesarán grandes cantidades de datos, con distintas tipologías y procedentes de múltiples fuentes de información para obtener un conocimiento profundo de la variable de interés, definiendo en un primer momento conjuntos de datos relacionados con variables meteorológicas y posteriormente otro grupo de conjuntos de datos relativos a la población. Durante la tesis se definirán las fuentes de información apropiadas para conseguir datos de calidad y aptos para el estudio. Una de las características de los datos utilizados es su naturaleza espacial, por lo cual hay que prestar especial atención a todos los aspectos relacionados con la georreferenciación, en particular los sistemas de referencia de cada conjunto de datos, la transformación entre sistemas de referencia y los formatos de datos



geoespaciales. También se definirá un flujo de trabajo para procesar y analizar los datos con técnicas procedentes de la Minería de Datos, buscando establecer un procedimiento óptimo y adaptado al problema, con el fin de obtener una metodología de estudio innovadora y que tenga el mayor impacto posible tanto en la comunidad científica como en la sociedad actual.

Medios disponibles:

Para la realización de esta tesis doctoral se dispone de un PC para el desarrollo de software en el Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría de la Universitat Politècnica de València. Asimismo, se dispone de un servidor instalado en la Universidad de Cantabria para el despliegue de algunos componentes del sistema. El tema de la presente tesis doctoral es multidisciplinar, los directores pertenecen a tres departamentos de dos universidades españolas que proporcionarán apoyo en las distintas fases del proyecto de tesis. En particular, uno de los coautores es el presidente de la International Society of Biometeorology (ISB) que dará apoyo en las fases iniciales.

Bibliografía:

- Fdez-Arroyabe, P. (2015). Climate Change, Local Weather and Customized Early Warning Systems based on Biometeorological Indexes. *Journal of Earth Science and Engineering*, 5, 173-181. <http://doi.org/10.17265/2159-581X/2015.03.002>.
- Fdez-Arroyabe, P., Lecha Estela, L., & Schimt, F. (2018). Digital divide, biometeorological data infrastructures and human vulnerability definition. *International Journal of Biometeorology*, 62(5), 733-740. <http://doi.org/10.1007/s00484-017-1398-x>.
- Lecha Estela, L.B. (2018). Biometeorological forecasts for health surveillance and prevention of meteor-tropic effects. *International Journal of Biometeorology*, 62(5), 741-771. <http://doi.org/10.1007/s00484-017-1405-2>.
- Marqués-Mateu, A., Moreno-Ramón, H., Balasch, S., & Ibáñez-Asensio, S. (2018). Quantifying the uncertainty of soil colour measurements with Munsell charts using a modified attribute agreement analysis. *Catena*, 171, 44-53. <https://doi.org/10.1016/j.catena.2018.06.027>.
- Menk, A., & Sebastiá, L. (2017). Are you Curious? Predicting the Human Curiosity from Facebook. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems* 25(2):79-95. <https://doi.org/10.1142/S0218488517400128>.