

---

**ESTIMACIÓN DE LA CALIDAD DEL AIRE EN LA CIUDAD DE TALCA  
UTILIZANDO ALGORITMOS DE APRENDIZAJE AUTOMÁTICO**

**EDUARDO ALEJANDRO ZAPATA GONZÁLEZ  
MAGÍSTER EN GESTIÓN DE OPERACIONES**

**RESUMEN**

Actualmente, muchas ciudades del centro y sur de Chile presentan un problema estacional (otoño-invierno) de contaminación del aire por material particulado, asociado a las características meteorológicas de los emplazamientos, y por otra parte, a altas emisiones de calefactores domiciliarios a leña. Sin embargo, los estudios de predicción de la calidad del aire se han concentrado en la ciudad de Santiago de Chile, observándose sólo dos estudios en ciudades de tamaño medio del centro y sur (Saide, Mena-Carrasco, et al. 2016). El presente trabajo se ha enfocado en generar un modelo de predicción de la calidad del aire en la ciudad de Talca ( $35^{\circ}26'S$ ;  $71^{\circ}44'W$ ), ciudad de tamaño medio, localizada en la zona central de Chile. El objetivo del estudio es predecir con un día de anticipación el material particulado de diámetro igual o menor a 2.5 micrómetros (MP<sub>2,5</sub>), para lo cual hemos utilizado tres modelos de aprendizaje automático: Perceptrón Multicapa (MLP), redes neuronales semisupervisadas basado en Selforganizing map (SOM) y Random Forest (RF). Los modelos utilizan una base de datos construida a partir de los registros históricos de calidad del aire, además de registros meteorológicos de velocidad del viento y temperatura en tres estaciones de medición de la ciudad. En tanto, los procesos de modelación consideran una búsqueda exhaustiva de parámetros. Nuestros resultados comprueban la alta capacidad de Random Forest como algoritmo predictor de episodios de emergencia ambiental en las tres estaciones de monitoreo de la ciudad, superando a los algoritmos de redes neuronales artificiales, reportados con los mejores resultados en la bibliografía. El modelo es una alternativa para las autoridades locales de la ciudad, ya que permitiría mejorar el sistema actual de pronósticos de calidad del aire.

## ABSTRACT

At present, many cities in central and southern Chile have a seasonal (autumn-winter) problem of air pollution by particulate matter, associated with the meteorological characteristics of the sites and high emissions of domestic wood-burning stoves. However, air quality forecasts studies have been concentrated in Santiago-Chile, observing only two studies in medium-sized cities in the center and south from Chile (Saide, Mena-Carrasco, et al. 2016). The present paper has focused on generating a model of air quality forecasts in the city of Talca ( $35^{\circ}26'S$ ;  $71^{\circ}44'W$ ), medium-sized city, located in the central zone from Chile. The aim of the study is to predict one-day-ahead particulate matter with a diameter equal to or less than 2.5 micrometers ( $PM_{2.5}$ ), for which we have used three models of machine learning: Multi-layer Perceptron (MLP), semisupervised neural networks based on Self-organizing map (SOM) and Random Forest (RF). The models use a database constructed from the historical air quality records and meteorological records (wind speed and temperature) in three measurement stations of the city. Modeling processes consider a comprehensive search for parameters. Our results verify the high capacity of Random Forest as an algorithm predictor of environmental emergency episodes in the three monitoring stations of the city, surpassing the algorithms artificial neural networks, reported with the best results in the literature. The model is an alternative for local authorities in the city, as it would improve the current system of air quality forecasts.