



FEATURE SELECTION FOR THE PREDICTION OF TROPOSPHERIC OZONE CONCENTRATION USING A WRAPPER METHOD

**C.OKAN SAKAR⁽¹⁾, GOKSEL DEMIR⁽²⁾, OLCAY KURSUN⁽³⁾, HUSEYIN OZDEMIR⁽²⁾,
GOKMEN ALTAY⁽⁴⁾, AND SENAY YALCIN⁽⁵⁾**

⁽¹⁾*Department of Computer Engineering, Bahcesehir University, Besiktas, Istanbul, Turkey*

⁽²⁾*Department of Environmental Engineering, Bahcesehir University, Besiktas, Istanbul, Turkey*

⁽³⁾*Department of Computer Engineering, Istanbul University, Avcilar 34320, Istanbul, Turkey*

⁽⁴⁾*Centre for Cancer Research & Cell Biology, Queen's University, Belfast, BT9 7BL*

⁽⁵⁾*Beykoz Logistics School of Higher Education, Beykoz, 34805, Istanbul, Turkey*

ABSTRACT—High concentrations of ozone (O₃) in the lower troposphere increase global warming, and thus affect climatic conditions and human health. Especially in metropolitan cities like Istanbul, ozone level approximates to security levels that may threaten human health. Therefore, there are many research efforts on building accurate ozone prediction models to develop public warning strategies. The goal of this study is to construct a tropospheric (ground) ozone prediction model and analyze the effectiveness of air pollutant and meteorological variables in ozone prediction using artificial neural networks (ANNs). The air pollutant and meteorological variables used in ANN modeling are taken from monitoring stations located in Istanbul. The effectiveness of each input feature is determined by using backward elimination method which utilizes the constructed ANN model as an evaluation function. The obtained results point out that outdoor temperature (OT) and solar irradiation (SI) are the most important input features of meteorological variables, and total hydrocarbons (THC), nitrogen dioxide (NO₂) and nitric oxide (NO) are those of air pollutant variables. The subset of parameters found by backward elimination feature selection method that provides the maximum prediction accuracy is obtained with six input features which are OT, SI, NO₂, THC, NO, and sulfur dioxide (SO₂) for both validation and test sets.

Key Words: Air pollution forecasting, variable sensitivity analysis, backward elimination, meteorological factors, artificial neural networks, Istanbul, Turkey.