Geophysical Research Abstracts
Vol. 14, EGU2012-13448, 2012
EGU General Assembly 2012


# MLP based models to predict PM10, $O_3$ concentrations, in Sines industrial area

R. Durao (1) and M. J. Pereira (2)

(1) Centre for Natural Resources and the Environment (CERENA), Instituto Superior Técnico, Lisboa, Portugal (rmdurao@ist.utl.pt), (2) Centre for Natural Resources and the Environment (CERENA), Instituto Superior Técnico, Lisboa, Portugal (maria.pereira@ist.utl.pt)

Sines is an important Portuguese industrial area located southwest cost of Portugal with important nearby protected natural areas. The main economical activities are related with this industrial area, the deep-water port, petrochemical and thermo-electric industry. Nevertheless, tourism is also an important economic activity especially in summer time with potential to grow.
The aim of this study is to develop prediction models of pollutant concentration categories (e.g. low concentration and high concentration) in order to provide early warnings to the competent authorities who are responsible for the air quality management. The knowledge in advanced of pollutant high concentrations occurrence will allow the implementation of mitigation actions and the release of precautionary alerts to population.

The regional air quality monitoring network consists in three monitoring stations where a set of pollutants' concentrations are registered on a continuous basis. From this set stands out the tropospheric ozone ($O_3$) and particulate matter (PM10) due to the high concentrations occurring in the region and their adverse effects on human health. Moreover, the major industrial plants of the region monitor $SO_2$, $NO_2$ and particles emitted flows at the principal chimneys (point sources), also on a continuous basis,.
Therefore
Artificial neuronal networks (ANN) were the applied methodology to predict next day pollutant concentrations; due to the ANNs structure they have the ability to capture the non-linear relationships between predictor variables.

Hence the first step of this study was to apply multivariate exploratory techniques to select the best predictor variables. The classification trees methodology (CART) was revealed to be the most appropriate in this case.. Results shown that pollutants atmospheric concentrations are mainly dependent on industrial emissions and a complex combination of meteorological factors and the time of the year.

In the second step, the Multi-layer perceptron (MLP) have shown to be able to learn the existent complex relationships using different combination of meteorological and emissions variables.

Furthermore, MLP models identified what are the meteorological conditions that most affect $O_3$ and PM10 concentrations in the region, namely wind speed and direction, boundary layer height, temperature, sunshine duration, relative humidity and the weather type.
The developed MLP models showed good predictive success with model performances between 0.66 and 0.87, indicating a reasonable accuracy for models development and generalization capability. These performance values are obtained using cross entropy error functions. This error functions are only available for classification problems and ensure that the network outputs are true class membership probabilities, which is known to enhance the performance of classification neural networks.