Geophysical Research Abstracts
Vol. 20, EGU2018-19766, 2018
EGU General Assembly 2018

Processing EUMETSAT Big Datasets to Monitor Air Pollution

Chaker El Amrani (1), Mohamed Akram Zaytar (1), Gilbert L. Rochon (2), and Tarek El-Ghazawi (3)

(1) Abdelmalek Essaadi University, Tangier, Morocco (ch.elamrani@fstt.ac.ma, mohamedakramzaytar@gmail.com), (2) MSF Global Solutions, LLC-New Orleans, LA, USA (rochon.gilbert@gmail.com), (3) George Washington University- Washington, DC, USA (tarek@gwu.edu)

Air pollution is one of the major problems in the world. According to the World Health Organization (WHO), exposure to high levels of such emissions can cause serious diseases such as asthma, bronchitis and lung cancer.
Several cities in Morocco like Casablanca, Safi and Tangier are suffering from harmful air pollution, due to industrial and farming activities, intensive roads and maritime traffic and thermal power plants carbon emissions. To mitigate pollution effects, Morocco launched several action plans to reduce gas emissions. Within this context, a real-time remote sensing ground station of the Mediterranean Dialogue Earth Observatory (MDEO) was established at Abdelmalek Essaadi University in Tangier with support from the North Atlantic Treaty Organization (NATO) Science for Peace Division for early warning of biogenic and anthropogenic disasters. The ground station includes ingestion from both geostationary and polar orbit tracking satellites.
In this study, we worked with data collected by IASI instrument on MetOp A and B satellites, and we built a data processing system, based on multiple stacked layers of computational processes that transform the Raw Binary Pollution Data coming directly to the MDEO cluster, into ready to interpret and visualize continuous data stream in near real time, using techniques varying from task automation, data pre-processing and data analysis to finally using a machine learning based feed forward artificial neural network architecture to interpolate missing pollution values.
We collected spatio-temporal values of methane, carbon dioxide, and nitrous oxide with location features and pollutant densities.
After pre-processing the native values, we trained a feed forward neural network architecture to interpolate missing pollution points using the neighboring 50 points in our areas of interest as an input, with 50 neurons in the input layer, three hidden layers with 100, 50, and 25 neurons respectively, and one neuron in the output layer for the prediction we need for the missing points.
We trained, tested and validated the model using a 150 Gigabyte datasets, which is the equivalent of around 800 million measurements.
We compared the performance of our trained model against the state of the art interpolation algorithms currently used (kernel smoothing and Kriging), the results were interesting as our model was able to be competitive in regard of all selected pollutants, performing even better in some cases.
Another study ongoing is to investigate GOME-2 instrument related pollution products, nitrogen dioxide, sulphur dioxide, bromine oxide and other trace gases. Collection and validation of pollutant values require using EUMETSAT IDL modules, to transform native satellite measurements into human interpretable values.