Project: **1234**

Project title: **Artificial Intelligence based system for sub-urban scale air quality prediction (QUALARIA): High-resolution data-fusion air quality forecast system for Brazil and Latin America**

Principal investigator: **Andrea Orfanoz-Cheuquelaf (Former PI: Cathy Li)**

Report period: **2024-11-01 to 2025-10-31**

The main objective of the QUALARIA project is to develop an AI model that can produce high-resolution (ΔX = 100 m) pollutant concentration data. This model requires input data such as pollutant concentrations from São Paulo CETESB Air Quality Stations, among others the target data. During this period, activities focused on finding, creating, and processing high-resolution data, as well as configuring and tuning the AI model. The first version of the AI model considered only CAMS input (ΔX ~40 km) (Inness *et al.*, 2019) to reproduce concentrations, with additional inputs of topography and land use. A second version of the AI model includes vehicular emissions and building heights with 100-meter spatial resolution (European Commission, 2023). Figure 1 shows maps of São Paulo for nitrogen dioxide (NO_2, left) and ozone (O_3, right). The top panels correspond to the first version of the AI model, and the bottom panels correspond to the second iteration. The second iteration improved the representation of NO_2 and O_3 concentrations. The results showed an increase in the linear correlation coefficient and a reduction in mean absolute error values. Our last iteration includes population data (Carioli, Alessandra Schiavina, Marcello MacManus and Freire, 2023) and data from air quality networks beyond São Paulo, including Monitor Air from Rio de Janeiro and IEMA data from Minas Gerais, Paraná, and Espirito

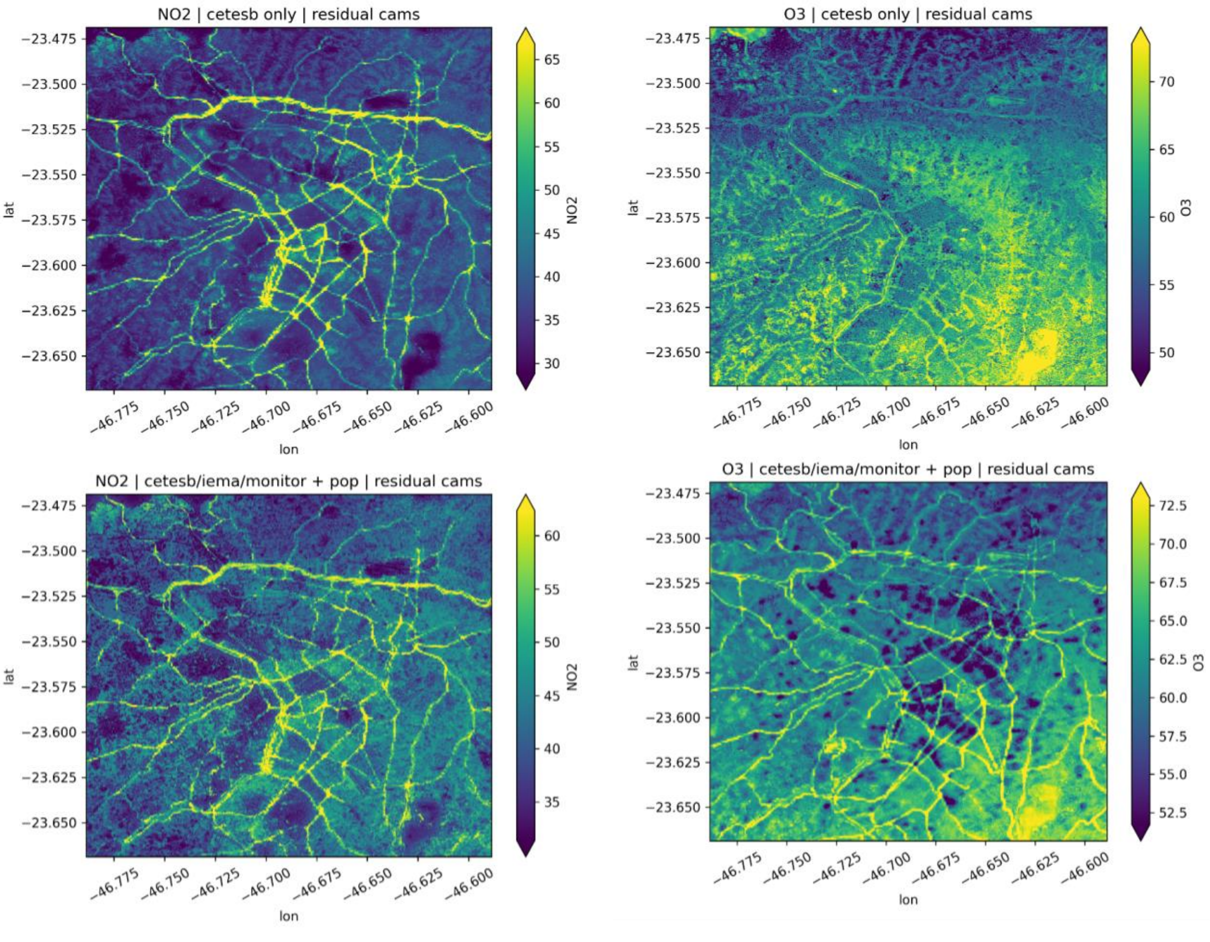

Figure 1. Maps of the outputs from the first two iterations of the AI model for the QUALARIA system in the metropolitan area of São Paulo, Brazil. The panels on the left show NO_2 and the panels on the right show O_3. The top panels show the outputs using only input data from CAMS, while the bottom panels show the outputs with extended input datasets.

Santo. Data from BERNARD low-cost sensors, which were recently deployed from June to August, will increase the coverage of the target data. Next, we will include additional input in the form of a WRF-Chem year simulation, which is being processed in Levante. We will also incorporate the low-cost sensor data into the pipeline and perform sensitivity tests to select the most important input data for the AI model. Finally, we will obtain the prototype of the QUALARIA system.

In collaboration with the Clean Air Catalyst (CAC) partnership from the World Resources Institute, source attribution simulations were performed for Indore, India; Nairobi, Kenya; and Jakarta, Indonesia. Air quality simulations were provided using WRF-Chem at a resolution of 3 km. Source attribution simulations were performed in 2019 for the three cities, comprising four rounds of simulations. The first round of simulations used inert tracers to analyze the influence of local and external emissions and estimate the contribution of different source sectors and regions around the three cities. The second round provided insight into the model's performance by using full chemistry and including locally produced emissions. The results were then compared with existing observations. A framework was then developed to incorporate local emission inventories into the WRF-Chem system, enhancing the model's ability to reproduce and predict local air quality conditions. The third round of simulations, the zero-out run, investigated the potential impact of eliminating anthropogenic emissions within the cities entirely. The fourth and final round employed a policy scenario that reduced local transport emissions, assessing the potential impact of mitigation policies. Several reports were produced, one for each round of simulations and city. Figure 2 shows the emissions treatment for the zero-out simulations in Indore, India.

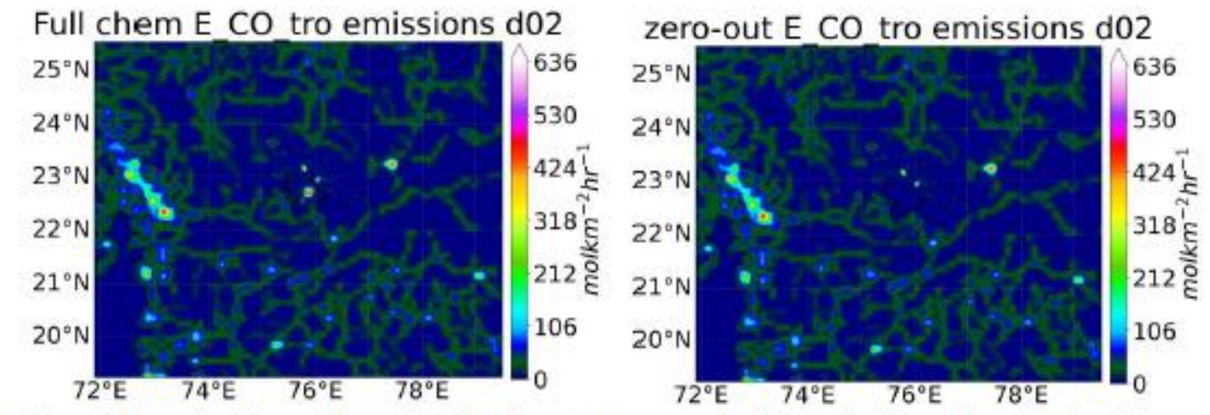

Figure 2. Maps of CO emissions for Indore, India; for the baseline scenario (left) and the zero-out sensitivity run (right).

References:

Carioli, Alessandra Schiavina, Marcello MacManus, K.J. and Freire, S. (2023) ‘GHS-POP R2023A - GHS population grid multitemporal (1975-2030)’. European Commission, Joint Research Centre (JRC). Available at: https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE.

European Commission (2023) *GHSL Data Package 2023*. Luxembourg: Publications Office of the European Union. Available at: https://doi.org/10.2760/098587.

Inness, A. *et al.* (2019) ‘The CAMS reanalysis of atmospheric composition’, *Atmospheric Chemistry and Physics*, 19(6), pp. 3515–3556. Available at: https://doi.org/10.5194/acp-19-3515-2019.