



Designing Air Quality Monitoring Systems for Measurement and Environmental Policy Evaluation in Data-Scarce Environments

Fourth Submission: PP_SSA_2024_193

Nihaa Sajid, Luis Eduardo San Martin

reproducibility@worldbank.org

October 31, 2024

This review verifies the reproducibility of the exhibits included in the paper “*Designing Air Quality Monitoring Systems for Measurement and Environmental Policy Evaluation in Data-Scarce Environments*”.

Contents in this review:

1. Main findings
2. List of exhibits and reproducibility status
3. Reproduction Environment

Main findings

- The code was successfully executed on a new computer after:
 1. Changing the file path in line 26 of the main do-file.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two consecutive times yielded identical results.
- The code takes approximately 1 hour to run.
- We conducted our reproducibility analysis based on the paper shared by the authors via OneDrive on October 31.
- All outputs reproduced as shown in the paper.
- **Reproducibility Summary:**
 - **Data:** Some data is restricted and has not been included in the reproducibility package. For more details, please refer to the README file.
 - **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.
 - **Outputs:** Some outputs are not generated by code (created manually in QGIS) but instructions for producing these outputs are included in the README file.
 - **Reproducibility verification:** Reviewers used data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package. The reviewers did not verify if publicly available data matches the data provided by the authors.
 - **Dependencies environment:**

- * For Stata: The reviewers created a new environment for dependencies using the latest versions available for each dependency at the moment of the review
- * For Python: The reviewers reproduced an existing environment for dependencies using an environment metadata file provided by the authors.
- * For QGIS: The reproducibility package does not use external dependencies.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Figure 1** Reproduced. The exhibit was compared against *map_locations_PA_Regulatory.html*.
- **Figure 2** Reproduced. The exhibit was compared against *pm25_daily_values.png*.
- **Figure 3a** Reproduced. The exhibit was compared against *daily_avg_across_all_monitors.png*.
- **Figure 3b** Reproduced. The exhibit was compared against *number_active_monitors_daily.png*.
- **Table 1** Reproduced. The exhibit was compared against *calibrations_hourly.tex*.
- **Figure 4a** Reproduced. The exhibit was compared against *Satellite_and_Regulatory_from_2012-01-10_to_2020-12-31_monthly.png*.
- **Figure 4b** Reproduced. The exhibit was compared against *Satellite_and_Regulatoryfrom_2019-01-10_to_2020-12-31_weekly.png*.
- **Figure 4c** Reproduced. The exhibit was compared against *Satellite_and_Regulatoryfrom_2019-01-10_to_2019-12-31_daily.png*.
- **Figure 4d** Reproduced. The exhibit was compared against *Satellite_and_Regulatoryfrom_2020-01-10_to_2020-12-31_daily.png*.
- **Figure 5a** Reproduced. The exhibit was compared against *Satellite_and_Regulatoryfrom_2020-03-03_to_2020-04-04.png*.
- **Figure 5b** Reproduced. The exhibit was compared against *Low-cost_and_Regulatoryfrom_2020-03-03_to_2020-04-04.png*.
- **Figure 6a** Reproduced. The exhibit was compared against *PM2_5__g_m__from_Low-cost__Unadjusted__and_pm25_reliable__g_m__from_Regulatory_from_2023-05-10_to_2023-11-30_daily.png*.
- **Figure 6b** Reproduced. The exhibit was compared against *Jaffe_pm1_10__from_Low-cost__Jaffe_Calibration__and_pm25_reliable__g_m__from_Regulatory_from_2023-05-10_to_2023-11-30_daily.png*.
- **Figure 6c** Reproduced. The exhibit was compared against *Local_Calibration__Hourly__from_Low-cost__Local_Calibration__and_pm25_reliable__g_m__from_Regulatory_from_2023-05-10_to_2023-11-30_daily.png*.
- **Figure 6d** Reproduced. The exhibit was compared against *Local_Calibration__Hourly__from_Low-cost__Local_Calibration__and_pm25_reliable__g_m__from_Regulatory_from_2023-05-10_to_2023-11-30_weekly.png*.

- **Figure 7a** Reproduced. The exhibit was compared against *avg_pm25_cf1_from_Low-cost_and_PM2_5___g_m___from_Satellite_from_2020-01-10_to_2021-12-31__daily.png*.
- **Figure 7b** Reproduced. The exhibit was compared against *avg_pm25_cf1_from_Low-cost_and_PM2_5__no_dust_from_Satellite_from_2020-01-10_to_2021-12-31__daily.png*.
- **Figure 7c** Reproduced. The exhibit was compared against *Jaffe_pm1_10__from_Low-cost_and_PM2_5___g_m___from_Satellite_from_2020-01-10_to_2021-12-31__daily.png*.
- **Figure 7d** Reproduced. The exhibit was compared against *Local_Calibration__Hourly__from_Low-cost_and_PM2_5___g_m___from_Satellite_from_2020-01-10_to_2021-12-31__daily.png*.
- **Table 2 - Panel A** Reproduced. The exhibit was compared against *correlation.tex*.
- **Table 2 - Panel B** Reproduced. The exhibit was compared against *rmse.tex*.
- **Figure 8a** Reproduced with manual steps following the instructions of the README.
- **Figure 8b** Reproduced with manual steps following the instructions of the README.
- **Figure 8c** Reproduced with manual steps following the instructions of the README.
- **Figure 8d** Reproduced with manual steps following the instructions of the README.
- **Figure 8e** Reproduced with manual steps following the instructions of the README.
- **Figure 9a** Reproduced. The exhibit was compared against *Low-cost_vs_Satellite_2020_covid_event_D.png*.
- **Figure 9b** Reproduced. The exhibit was compared against *Regulatory_vs_Satellite_2020_covid_event_D.png*.
- **Figure 10a** Reproduced. The exhibit was compared against *Low-cost_covid_event_W.png*.
- **Figure 10b** Reproduced. The exhibit was compared against *Satellite_covid_event_W.png*.
- **Figure 10c** Reproduced. The exhibit was compared against *Regulatory_covid_event_W.png*.
- **Table 3** Reproduced. The exhibit was compared against *reg_hourly_2020+2023.tex*.
- **Table 4** Reproduced. The exhibit was compared against *local_calibration.tex*
- **Table 5** Reproduced. The exhibit was compared against *coefficient_comparison_hourly.tex*.

Results in the Annex

- **Figure A1** Reproduced. The exhibit was compared against *daily_avg_overtime.html*
- **Figure A2** Reproduced. The exhibit was compared against *PM2_5___g_m___from_Satellite_and_PM2_5__no_dust_from_Satellite__excluding_dust__from_2020-01-10_to_2021-12-31__daily*.
- **Figure A3** Does not apply: the exhibit was created using manual steps entirely, see the README for the reproduction instructions.
- **Figure A4a** Reproduced. The exhibit was compared against *Low-cost_covid_event_D.png*.

- **Figure A4b** Reproduced. The exhibit was compared against *Satellite_covid_event_D.png*.
- **Figure A4c** Reproduced. The exhibit was compared against *Regulatory_covid_event_D.png*.
- **Table A1** Does not apply: the exhibit does not show analysis results.
- **Table A2** Reproduced. The exhibit was compared against *calibrations_daily.tex*.
- **Table A3 - Panel A** Reproduced. The exhibit was compared against *max_daily_2020+2023.tex*.
- **Table A3 - Panel B** Reproduced. The exhibit was compared against *reg_daily_2020+2023.tex*.
- **Table A3 - Panel C** Reproduced. The exhibit was compared against *reg_hourly_2017.tex*.
- **Table A3 - Panel D** Reproduced. The exhibit was compared against *reg_daily_2017.tex*.

Reproduction Environment

Paper exhibits were reproduced on a computer with the following specifications:

- OS: Windows 11 Enterprise, version 23H2
- Processor: 11th Gen Intel(R) Core(TM) i5-1145G7 @ 2.60GHz 1.50 GHz
- Memory available: 15.7 GB
- Software version: Stata 18.0 MP, Python 3.12.4, QGIS 3.38.1