
Estimation of Air Pollution with Remote Sensing Data: Revealing Greenhouse Gas Emissions from Space

Linus Scheibenreif¹ Michael Mommert¹ Damian Borth¹

Abstract

Air pollution is a major driver of climate change. Anthropogenic emissions from the burning of fossil fuels for transportation and power generation emit large amounts of problematic air pollutants, including Greenhouse Gases (GHGs). Despite the importance of limiting GHG emissions to mitigate climate change, detailed information about the spatial and temporal distribution of GHG and other air pollutants is difficult to obtain. Existing models for surface-level air pollution rely on extensive land-use datasets which are often locally restricted and temporally static. This work proposes a deep learning approach for the prediction of ambient air pollution that only relies on remote sensing data that is globally available and frequently updated. Combining optical satellite imagery with satellite-based atmospheric column density air pollution measurements enables the scaling of air pollution estimates (in this case NO_2) to high spatial resolution (up to $\sim 10\text{m}$) at arbitrary locations and adds a temporal component to these estimates. The proposed model performs with high accuracy when evaluated against air quality measurements from ground stations (mean absolute error $< 6 \mu\text{g}/\text{m}^3$). Our results enable the identification and temporal monitoring of major sources of air pollution and GHGs.

1. Introduction

Air pollution and the emission of GHGs is the main cause of climate change with annual global emission levels still on the rise (Friedlingstein et al., 2019). In particular, anthropogenic GHG emissions from the combustion of fossil fuels in industrial plants or for transportation are harmful

¹Institute of Computer Science, University of St. Gallen, Switzerland. Correspondence to: Linus Scheibenreif <linus.scheibenreif@unisg.ch>.

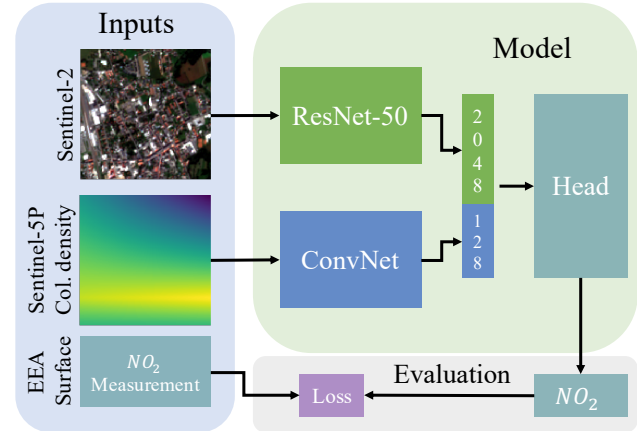


Figure 1. Overview of the proposed air pollution prediction system.

to the environment and contribute to global warming trends (Ledley et al., 1999). Besides the primary greenhouse gas, CO_2 , the burning of fossil fuels also emits molecules like NO_2 and CO , which have been used as proxy for the estimation of CO_2 emissions (Berezin et al., 2013). Detailed information about sources and distribution of air pollutants within the atmosphere is of high relevance for a number of applications with climate change impact, including the compilation of emission inventories (Eggleston et al., 2006), the design and implementation of pollution limits (Bollen & Brink, 2014), and the quantification of large anthropogenic emissions (Liu et al., 2020).

At present, continual data on air pollution concentrations in the atmosphere are primarily collected through two different approaches with distinct drawbacks. On the Earth’s surface, networks of measurement stations record the concentration of various chemicals at select locations (Guerreiro et al., 2014). Such networks are commonly run by environmental agencies and provide frequent measurements while often lacking in spatial coverage. This drawback can be partly addressed by space-borne air pollution monitoring: satellites equipped with spectrometers measure the abundance of select molecules in the form of atmospheric column densities (Gupta et al., 2006). While their position in Earth’s orbit

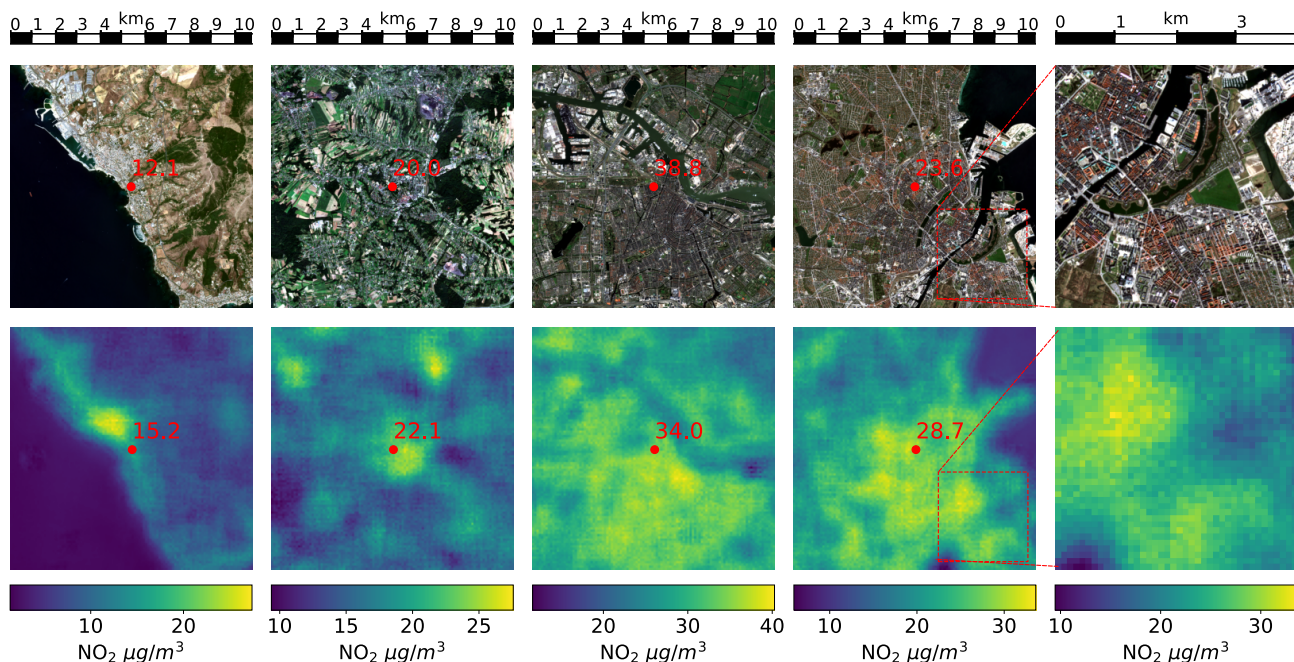


Figure 2. Exemplary NO_2 predictions based on Sentinel-2 and Sentinel-5P input data. **Top:** RGB bands of the Sentinel-2 image, red dots mark locations of air quality stations, red text indicates the average NO_2 concentration measured on the ground during the 2018-2020 timespan. **Bottom:** Predicted NO_2 concentrations for the locations above (not seen during training) with predictions at the exact position of air quality stations in red. The heatmaps are constructed from individual predictions for overlapping 120×120 pixel tiles of the top image and corresponding Sentinel-5P data, resulting in an effective spatial resolution of 100m. This approach is equally applicable to locations without air quality stations, providing a means to map air pollution on the surface level to identify sources of air pollution and GHG emissions (see Fig. A1 for more examples).

allows satellites to frequently map most locations on Earth, remote sensing spectrometers currently only provide spatial resolutions in the kilometer range and with little information about the pollutant’s vertical distribution. Specifically, the estimation of concentrations near the surface, where these pollutants originate from, is a non-trivial task (Scheibenreif et al., 2021). One of the primary anthropogenic air pollutants is Nitrogen Dioxide (NO_2). Elevated levels of NO_2 harm the vegetation, contribute to acid rain, and act as a precursor of potent GHGs like Ozone (Montzka et al., 2011). Additionally, NO_2 is jointly emitted with CO_2 during the combustion of fossil fuels at high temperatures, making it a suitable proxy to identify CO_2 emission sources (Konovalov et al., 2016; Goldberg et al., 2019). This work leverages a large body of publicly available NO_2 concentration measurements on the ground by the European Environment Agency’s¹ (EEA) network of air quality stations and satellite measurements from the European Space Agency’s (ESA) Copernicus program to investigate the distribution of air pollutants through a deep learning approach. The results of this work enable the identification of major sources of GHG emissions and their temporal monitoring on a global scale.

¹eea.europa.eu

2. Background

Various prediction and interpolation techniques have been used to derive detailed information about the spatial distribution of air-borne pollutants such as GHGs. Typically, these models are based on point measurements from air quality monitoring stations that are spatially limited to specific locations. Beyond interpolation with geostatistical approaches like kriging (Janssen et al., 2008), land-use-regression (LUR) is commonly applied to incorporate covariates such as population density or traffic data into the models (see Hoek et al., 2008, for a review). LUR models often involve variable selection procedures to identify predictive inputs over large sets of candidate variables, making it difficult to scale to regions not covered by detailed datasets, even if some air quality measurements are available. Building on existing work that incorporates satellite measurements into LUR frameworks (Novotny et al., 2011), we extend this approach to model air pollution at high spatial resolution solely from satellite data. Our work is based on NO_2 concentration measurements by the EEA. We consider NO_2 as pollutant of interest due to its relevance as major anthropogenic air pollutant and chemical properties that facilitate its detection from space with high accuracy (opposed to GHGs like CO_2). Additionally, it is co-emitted with CO_2

Table 1. Performance metrics for NO₂ estimation with various model architectures and datasets on unseen observations, averaged over 10 training runs with varying random seeds (PT: Pre-trained model, *-T10: Performance of top model out of 10 training runs).

DATA	TIME	N-OBS.	PT	R2	R2-T10	MAE	MAE-T10	MSE	MSE-T10
SEN.-2	2018-20	3.2K	×	0.25±0.05	0.28	8.06±0.49	7.31	105.7±10.29	91.72
SEN.-2	2018-20	3.2K	✓	0.45±0.03	0.49	6.62±0.17	6.23	77.03±3.64	65.81
SEN.-2,5P	2018-20	3.1K	×	0.38±0.03	0.43	7.06±0.35	6.68	83.72±4.14	78.4
SEN.-2,5P	2018-20	3.1K	✓	0.54±0.04	0.59	5.92±0.44	5.42	62.52±5.47	56.28
SEN.-2,5P	QUART.	19.6K	✓	0.52±0.05	0.57	6.24±0.22	5.98	73.1±6.88	66.12
SEN.-2,5P	MONTH.	59.6K	✓	0.51±0.01	0.53	6.54±0.15	6.31	78.96±4.2	73.74

in the burning of fossil fuels, which makes it possible to constrain CO₂ emissions from NO₂ measurements (Berezin et al., 2013). To facilitate the identification of air pollutant sources, which are commonly located on the ground, we model surface-level concentrations (rather than e.g. atmospheric column densities). The EEA network of air quality stations provides frequent (mostly hourly) measurements of NO₂ concentrations at more than 3,000 locations in Europe. Additionally, remote sensing data from ESA’s Sentinel-2 and Sentinel-5P satellites is utilized to model air quality. Sentinel-2 is a constellation of two satellites carrying the Multi Spectral Instrument, a spectrometer covering the visible, near-infrared and shortwave-infrared wavelengths with imaging resolutions between 10 and 60 meters (Drusch et al., 2012). Sentinel-2 data is widely used in applications like land cover classification or crop monitoring (Helber et al., 2019) but also for the monitoring of GHGs at locations of interest (e.g., based on the presence of smoke plumes, Mommert et al., 2020). In our work, globally available and continually updated Sentinel-2 images replace conventional LUR predictor variables such as street networks, population density or vegetation information. The Sentinel-5P satellite observes trace-gases and aerosols in the atmosphere through differential optical absorption spectroscopy (Veeffkind et al., 2012). It provides daily global coverage for gases including NO₂, O₃, CO or CH₄ with a spatial resolution of 5 × 3.5 km. We utilize the NO₂ tropospheric column density product of Sentinel-5P to model the temporal variation in surface NO₂ levels.

3. Methods

This work approaches the estimation of air pollution as a supervised computer vision problem. We collect a dataset of harmonized remote sensing data from Sentinel-2 and Sentinel-5P, spatially and temporally aligned with measurements from air quality monitoring stations. The proposed model is trained on pairs of remote sensing input and air quality target values (see Fig. 1), which yields a system that predicts air pollution levels solely from globally available

remote sensing data².

3.1. Data Processing

We consider the 2018-2020 timespan, historically limited by the start of the Sentinel-5P nominal mission. NO₂ measurements by EEA air quality stations are filtered to remove values with insufficient quality (`validity` or `verification` value $\neq 1$). Besides modelling the entire 2018-2020 timespan, we also investigate the possibility to estimate NO₂ concentrations at quarterly and monthly frequencies. To that end, the mean of NO₂ measurements for each frequency is used as prediction target. To build the dataset, we downloaded Sentinel-2 Level-2A data (i.e. corrected for atmospheric effects and enriched with cloud masks) with low cloud-coverage at the locations of air quality stations, containing 12 different bands (band 10 is empty in the case of Level-2A data). The images were then cropped to 120×120 pixel size ($\sim 1.2 \times 1.2$ km) centered at the location of interest, and all bands were upsampled to 10 m resolution with bilinear upsampling. Additionally, we visually inspected the RGB bands of all images to ensure that no clouds or artifacts are present. Similarly, Sentinel-5P data over Europe was downloaded for the 2018-2020 timespan (5449 Level-2 products) and mapped to a common rectangular grid of 0.05×0.05 ° ($\sim 5 \times 5$ km) resolution after removing invalid measurements (`qa_value` < 75). The resulting dataset was averaged at the different temporal frequencies and 20×20 km regions at the locations of air quality stations were extracted. To facilitate processing despite the coarse resolution ($\sim 500 \times$ lower than Sentinel-2), we linearly interpolated the Sentinel-5P data to 10 m resolution and cropped to 120×120 pixel centered at the locations of interest.

3.2. Model Architecture

Our core model for NO₂ prediction from imaging data is based on the ResNet-50 architecture (He et al., 2016) (see Fig. 1). The input layer is modified to accommodate the 12-band Sentinel-2 input data and the final layer is replaced by

²code available at github.com/HSG-AIML/RemoteSensingNO2Estimation

two dense layers with ReLU activation (named *head*) that map the 2048-dimensional feature vector to a scalar value. We employ transfer learning by pretraining the model on a land-cover classification (LCC) task with the BigEarthNet dataset (Sumbul et al., 2019). After pretraining, the final classification layer is replaced by the *head*, i.e., only the trained convolutional backbone of the ResNet is retained. Intuitively, learned features that are informative for LCC (e.g., distinguishing industrial areas from forests) will also be useful when estimating emission profiles of different areas. To handle additional input data from Sentinel-5P, the model architecture is extended with a small sub-network, consisting of two convolutional layers (with 10, 15 channels and kernel sizes 3, 5, respectively), each followed by ReLU activation functions and max-pooling (kernel size 3), and a final linear layer. This sub-network is much smaller than the ResNet-50 used to process the Sentinel-2 input stream to reflect the lower native resolution and single band nature of the Sentinel-5P data. It learns a 128 dimensional latent vector from the Sentinel-5P input image. To obtain an NO₂ prediction, the latent vectors of both input-streams are concatenated and again processed by the *head* with adjusted input dimensions (2048+128). All presented models were trained 10 times with varying seeds, mean-squared-error loss function and random train/test/validation split of 60:20:20. To limit overfitting, training is stopped once the loss on the validation set stops decreasing. Additionally, we employ random flipping and rotation of the inputs as augmentation during training.

4. Experiments

To assess the predictive power of Sentinel-2 images for air pollution prediction we initially train a model on Sentinel-2 images as inputs with air quality station measurements as target. Using only Sentinel-2 images forces the model to associate features that are apparent in medium-resolution satellite imagery, like built-up areas, forests or streets, with representative NO₂ levels. Training this model from scratch leads to a mean-absolute-error (MAE) of $8.06 \pm 0.49 \mu\text{g}/\text{m}^3$ and R2-Score of $0.25 \pm 0.05 \mu\text{g}/\text{m}^3$ (see Table 1), presumably limited by the dataset size of only 3,227 images. Following the intuition that LCC shares predictive features with air pollution prediction, we then investigated a transfer learning approach by pre-training the ResNet backend on BigEarthNet (590,326 images with multi-label annotations, Sumbul et al., 2019). Using the pretrained backend in the NO₂ prediction model and fine-tuning on the Sentinel-2 images at air quality stations, we obtain a significantly better performance. The MAE drops to $6.62 \pm 0.17 \mu\text{g}/\text{m}^3$ with an R2-Score of 0.45 ± 0.03 . This first result supports our hypothesis that medium-resolution satellite imagery is valuable for the estimation of ambient air pollution. We then investigated ways of incorporating tropospheric column den-

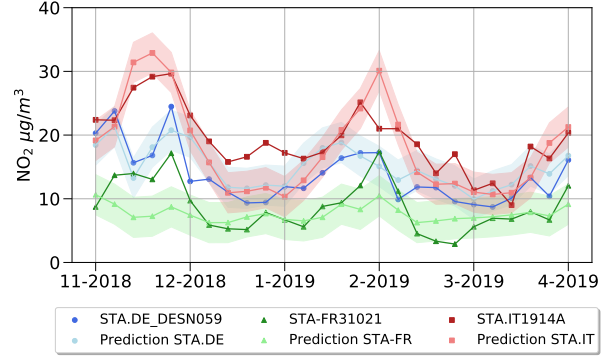


Figure 3. Monthly average NO₂ measurements from three EEA air quality stations in Germany, France and Italy (dark colors) and monthly NO₂ predictions based on Sentinel-2 and Sentinel-5P measurements at the same locations (not seen during training). The shaded area indicates the model’s MAE envelope centered at the nominal predictions.

sity measurements of NO₂ from Sentinel-5P into the model using a second input stream. The additional satellite data results in a further performance increase with an MAE of 5.92 ± 0.44 and R2-Score of 0.54 ± 0.04 and allows us to derive detailed pollution maps for any location of interest (see Fig. 2). Inclusion of Sentinel-5P data, which is updated daily, also provides us with a way of modeling temporal variations in NO₂ levels. Aggregating the data at higher frequency significantly increases the number of observations (from 3.1k to 19.6k quarterly and 59.6k monthly samples), which enables the model to maintain a performance comparable to the static predictions (MAE of 6.24 ± 0.22 and $6.52 \pm 0.15 \mu\text{g}/\text{m}^3$ for quarterly and monthly predictions, respectively). Similarly, the R2-Scores remain at 0.52 ± 0.05 (quarterly) and 0.51 ± 0.01 (monthly) despite the increase in prediction frequency. This makes it possible to model seasonal changes in NO₂ concentrations with good accuracy (see Fig. 3).

5. Conclusion

We present an end-to-end approach for the estimation of surface NO₂ concentrations with deep learning. Utilizing only remote sensing data as inputs, it is possible to model arbitrary regions on Earth, independent of the availability of detailed datasets as commonly used in the prediction of air pollutant distributions. Qualitative evaluation shows that our models are robust across most regions of Europe, except for rare atypical locations that are badly represented in our dataset, e.g., snowy mountain peaks. In future work, measurements from air quality networks outside of Europe can be incorporated into model training to improve model generalization.

The focus of this work on NO₂ allows us to leverage a large corpus of pollutant measurements from air quality stations and from space to better localize the sources of air pollution

and GHG emitters. This information enables an approximate analysis of the spatial and temporal distribution of air pollutants and GHG emissions alike, providing constraints that are vital for our effort to reduce GHG emissions and reaching the net-zero emission target.

Acknowledgements

We thank the ESA Copernicus programme and the European Environment Agency for providing the data used in this work.

References

- Berezin, E., Konovalov, I., Ciais, P., Richter, A., Tao, S., Janssens-Maenhout, G., Beekmann, M., and Schulze, E.-D. Multiannual changes of CO₂ emissions in China: indirect estimates derived from satellite measurements of tropospheric NO₂ columns. *Atmospheric Chemistry and Physics*, 13(18):9415–9438, 2013.
- Bollen, J. and Brink, C. Air pollution policy in Europe: Quantifying the interaction with greenhouse gases and climate change policies. *Energy Economics*, 46:202–215, 2014.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120:25–36, 2012.
- Eggleston, H., Buendia, L., Miwa, K., Ngara, T., and Tanabe, K. 2006 IPCC guidelines for national greenhouse gas inventories. 2006.
- Friedlingstein, P., Jones, M. W., O'sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Quéré, C. L., et al. Global carbon budget 2019. *Earth System Science Data*, 11(4):1783–1838, 2019.
- Goldberg, D. L., Lu, Z., Oda, T., Lamsal, L. N., Liu, F., Griffin, D., McLinden, C. A., Krotkov, N. A., Duncan, B. N., and Streets, D. G. Exploiting OMI NO₂ satellite observations to infer fossil-fuel CO₂ emissions from US megacities. *Science of The Total Environment*, 695: 133805, 2019.
- Guerreiro, C. B., Foltescu, V., and De Leeuw, F. Air quality status and trends in Europe. *Atmospheric Environment*, 98:376–384, 2014.
- Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y., and Kumar, N. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment*, 40(30):5880–5892, 2006.
- He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.
- Helber, P., Bischke, B., Dengel, A., and Borth, D. Eurosat: A novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2217–2226, 2019.
- Hoek, G., Beelen, R., De Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., and Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33):7561–7578, 2008.
- Janssen, S., Dumont, G., Fierens, F., and Mensink, C. Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmospheric Environment*, 42 (20):4884–4903, 2008.
- Konovalov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Zhuravlev, R. V., and Janssens-Maenhout, G. Estimation of fossil-fuel CO₂ emissions using satellite measurements of "proxy" species. *Atmospheric Chemistry and Physics*, 16(21):13509–13540, 2016.
- Ledley, T. S., Sundquist, E. T., Schwartz, S. E., Hall, D. K., Fellows, J. D., and Killeen, T. L. Climate change and greenhouse gases. *Eos, Transactions American Geophysical Union*, 80(39):453–458, 1999.
- Liu, F., Duncan, B. N., Krotkov, N. A., Lamsal, L. N., Beirle, S., Griffin, D., McLinden, C. A., Goldberg, D. L., and Lu, Z. A methodology to constrain carbon dioxide emissions from coal-fired power plants using satellite observations of co-emitted nitrogen dioxide. *Atmospheric Chemistry and Physics*, 20(1):99–116, 2020.
- Mommert, M., Sigel, M., Neuhausler, M., Scheibenreif, L., and Borth, D. Characterization of industrial smoke plumes from remote sensing data. "Tackling Climate Change with Machine Learning" workshop at NeurIPS 2020, 2020.
- Montzka, S. A., Dlugokencky, E. J., and Butler, J. H. Non-CO₂ greenhouse gases and climate change. *Nature*, 476 (7358):43–50, 2011.
- Novotny, E. V., Bechle, M. J., Millet, D. B., and Marshall, J. D. National satellite-based land-use regression: NO₂ in the United States. *Environmental Science & Technology*, 45(10):4407–4414, 2011.
- Scheibenreif, L., Mommert, M., and Borth, D. A Novel Dataset and Benchmark for Surface NO₂ Prediction from

Remote Sensing Data Including COVID Lockdown Measures. In *IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2021.

Sumbul, G., Charfuelan, M., Demir, B., and Markl, V. Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5901–5904. IEEE, 2019.

Veefkind, J., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H., De Haan, J., Kleipool, Q., et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120: 70–83, 2012.

A. Appendix

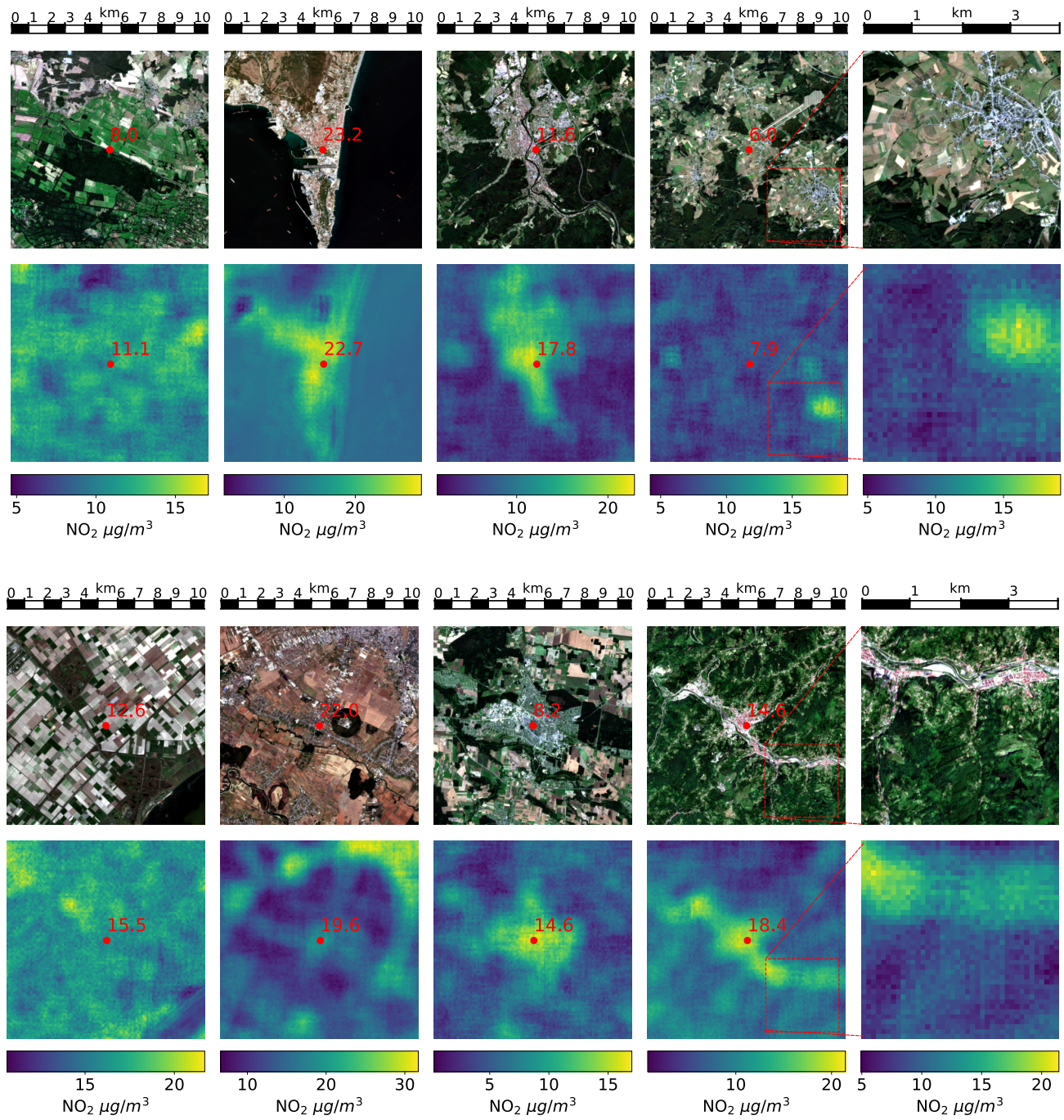


Figure A1. Additional examples of surface NO_2 predictions from Sentinel-2 and Sentinel-5P data across Europe. Pictures are centered at locations of EEA air quality stations (red dots). The red text indicates average NO_2 measurements (RGB images) and corresponding NO_2 estimates (heatmaps).