Appendix

Data sources, retrieval and processing

The land-use/cover data were extracted from the pan-European component of the Copernicus land monitoring service (CLMS, 2019). In particular, the CORINE land cover (CLC) dataset (v.18_5) for the year 2012 was used (Corine Land Cover , 2012). To better understand the urban surface characteristics surrounding each monitoring station, a squared buffer zone of 1 km^2 spatial resolution was created and the dominant CLC category within each buffer zone was computed and assigned to the respective site. The 45 land classes available in CLC were aggregated to form the following 4 main categories: (i) continuous urban fabric - road and rail networks and associated land - port areas (LC1); (ii) discontinuous urban fabric - industrial or commercial units - mine, dump and construction sites - artificial, non-agricultural vegetated areas (LC2); (iii) agricultural areas - wetlands - water bodies (LC3); and (iv) forest and semi natural areas (LC4). Additionally, the high resolution layers of tree cover density (TCD) (Tree Cover Density, 2015) and imperviousness (IMP) (Imperviousness Density, 2015) were accessed from the same source (CLMS, 2019). The digital elevation model (DEM) product (DEM, 2012) was downloaded from the EEA website.

The land surface temperature (LST, 2016) and the normalized difference vegetation index (NDVI, 2016) generated from the MODIS Aqua and Terra platforms, the night time lights (NTL, 2013) product from the National Oceanic and Atmospheric Administration (NOAA), as well as the climatic data from the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFSv2) (Saha et al., 2011) were pre-processed using

the Google Earth Engine (GEE) API (Google Earth Engine Team, 2015). GEE makes it possible to rapidly process vast amount of satellite imagery on global scale with the power of Google's cloud computing. The road density (RD) raster was computed using the OpenStreet Maps project's collection of road shapefiles covering the continent. Particularly, the major roads comprising the motorway, trunk, primary and secondary road categories, as well as the links between them were taken into consideration. The same dataset was used to compute the distance to roads (DISR) covariate applying simple GIS (Geographic Information System) techniques. Distance to sea (DISS) was calculated using the Europe coastline shapefile (ECS, 2015) downloaded from the EEA website.

The population data were obtained from the Gridded Population of the World, Version 4 (GPWv4) database. Particularly, the population density (adjusted to match 2015 revision UN WPP country totals) dataset (SEDAC, 2016), available at 30 arc-second ($\sim 1\ km^2$) spatial resolution for the years 2000, 2005, 2010, 2015, and 2020 was used to estimate the population in 2016 at 1 km^2 pixel level applying cubic splines interpolation. The resulting estimates were aggregated at country level using the European administrative country boundaries shapefile from Eurostat's GISCO service (EuroStat, 2019).

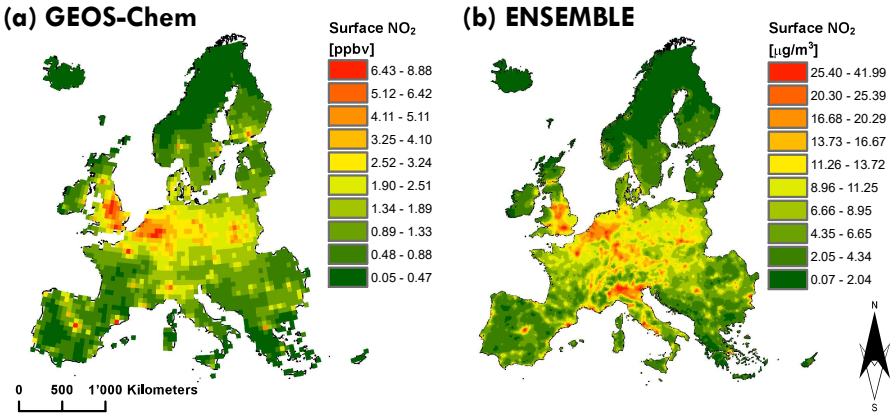

Figure S1: **Surface NO_2 simulations in 2016 from the chemical transport models (CTMs)**. **a**: GEOS-Chem at 0.5˚× 0.625˚ spatial resolution. **b**: ENSEMBLE at 0.1˚× 0.1˚ spatial resolution (generated using Copernicus Atmosphere Monitoring Service Information [2016])

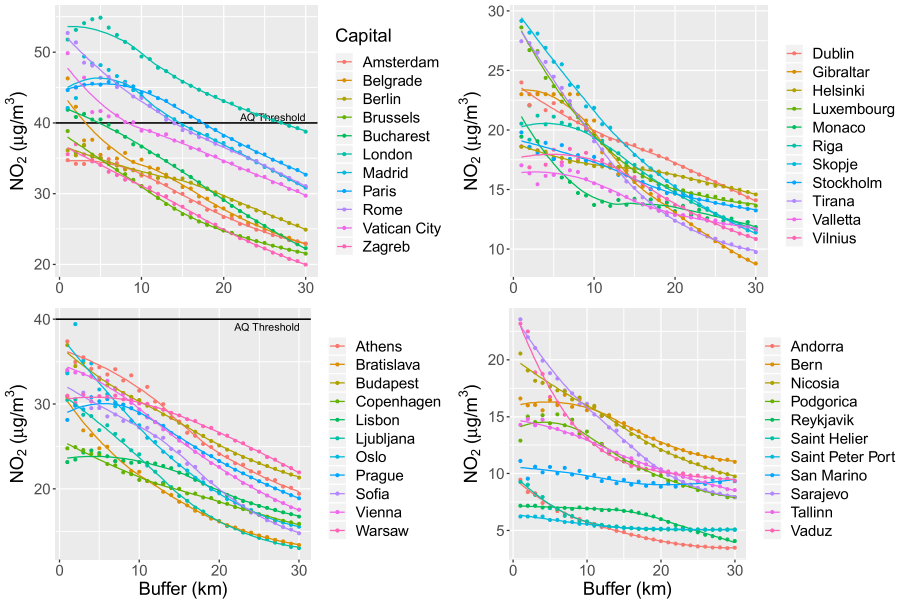

Figure S2: **Surface NO_2 estimates in 44 European capitals in 2016.** The air pollution profiles of NO_2 within 30 km buffer zone from the center of each capital. The black horizontal line corresponds to EU and WHO air quality (AQ) threshold of 40 $\mu g/m^3$.

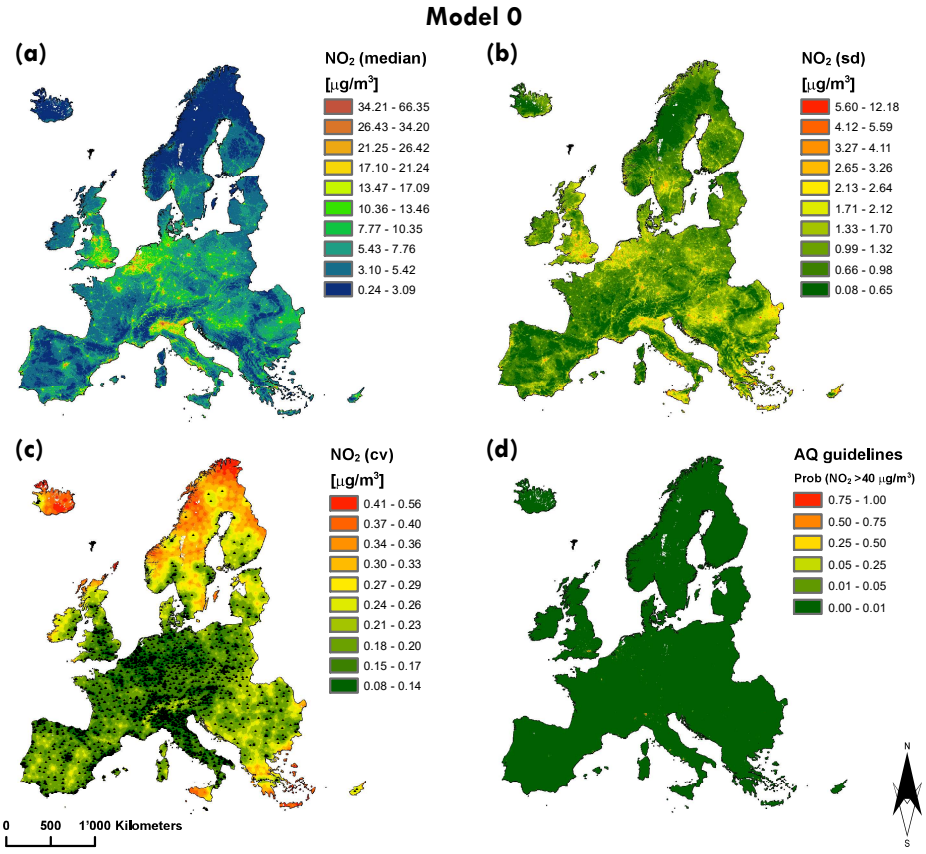

Figure S3: **Surface NO_2 estimates in Europe in 2016 predicted using Model 0** **a**: Predicted NO_2 concentrations (i.e. median of the posterior predictive distribution) at 1 km^2 spatial resolution. **b**: Prediction uncertainty in terms of standard deviation (sd) of the posterior predictive distribution of NO_2. **c**: Prediction uncertainty in terms of coefficient of variation (cv) of the posterior predictive distribution of NO_2 with the overlayed locations of NO_2 monitoring stations. **d**: Probability that NO_2 concentration exceeds the EU Directive and WHO air quality threshold of 40 $\mu g/m^3$.

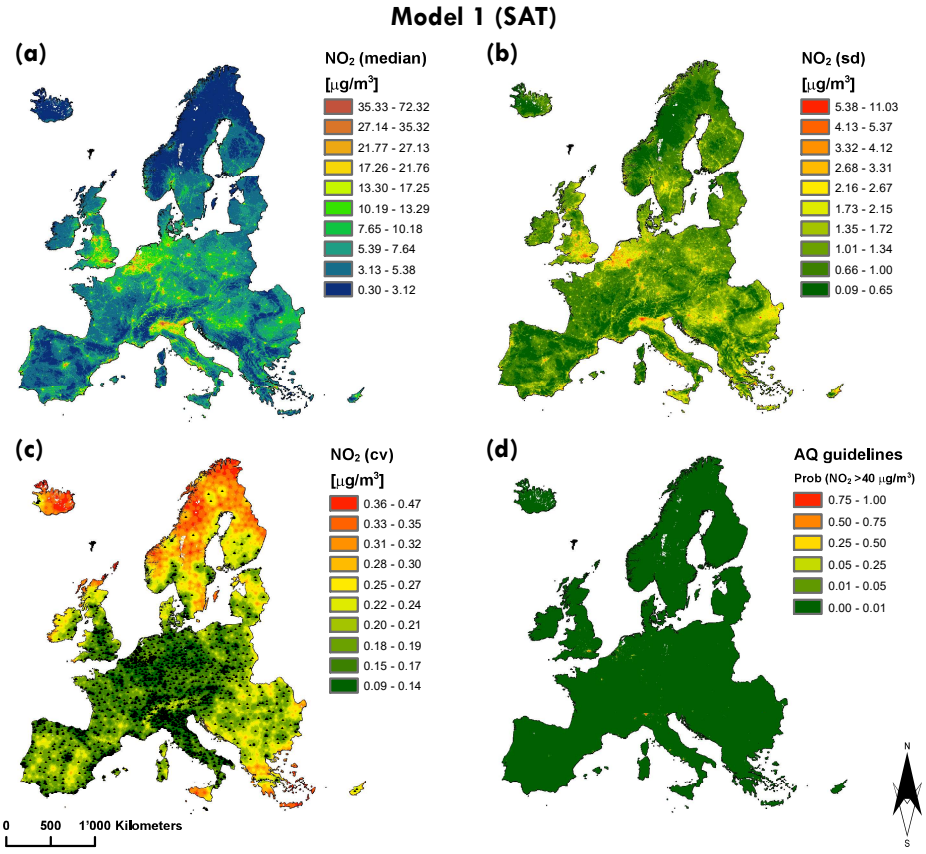

Figure S4: **Surface NO_2 estimates in Europe in 2016 predicted using Model 1 (SAT)**. **a**: Predicted NO_2 concentrations (i.e. median of the posterior predictive distribution) at 1 km^2 spatial resolution. **b**: Prediction uncertainty in terms of standard deviation (sd) of the posterior predictive distribution of NO_2. **c**: Prediction uncertainty in terms of coefficient of variation (cv) of the posterior predictive distribution of NO_2 with the overlayed locations of NO_2 monitoring stations. **d**: Probability that NO_2 concentration exceeds the EU Directive and WHO air quality threshold of 40 $\mu g/m^3$.

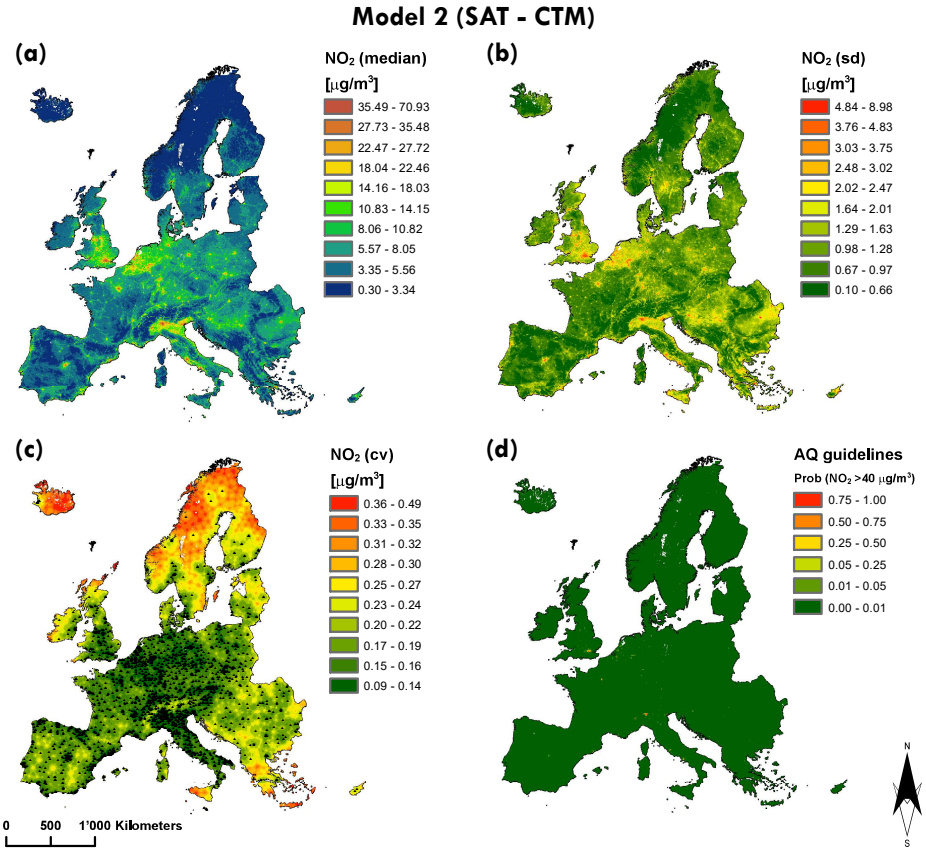

Figure S5: **Surface NO_2 estimates in Europe in 2016 predicted using Model 2 (SAT-CTM)**. **a**: Predicted NO_2 concentrations (i.e. median of the posterior predictive distribution) at 1 km^2 spatial resolution. **b**: Prediction uncertainty in terms of standard deviation (sd) of the posterior predictive distribution of NO_2. **c**: Prediction uncertainty in terms of coefficient of variation (cv) of the posterior predictive distribution of NO_2 with the overlayed locations of NO_2 monitoring stations. **d**: Probability that NO_2 concentration exceeds the EU Directive and WHO air quality threshold of 40 $\mu g/m^3$.

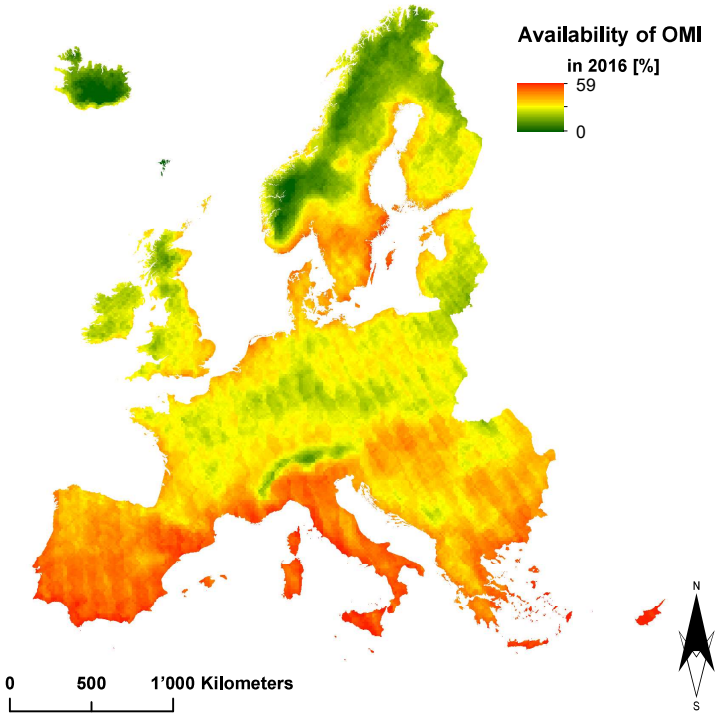

Figure S6: **Availability of OMI.** Percentage of non-missing OMI tropospheric cloud-screened days in Europe in 2016.

Table A1: Comparison between the models with no covariates (i.e. only spatial process w_i), model with only OMI_NO2 or OMI_NO2_GC (with and without w_i), as well as the ones only with ENSEMBLE data.

Model	**logscore**	**MAE**	**MAPE**	**RMSE**	**R^2**
No covariates (only spatial process w_i)	3.704	0.419	0.307	0.586	0.255
Only OMI_NO2	3.863	0.472	0.373	0.629	0.131
OMI_NO2 + w_i	3.698	0.418	0.296	0.577	0.277
Only OMI_NO2_GC	3.851	0.463	0.370	0.622	0.152
OMI_NO2_GC + w_i	3.698	0.418	0.296	0.576	0.277
Only ENSEMBLE	3.676	0.383	0.298	0.521	0.404
ENSEMBLE + w_i	3.618	0.371	0.262	0.507	0.436

Table A2: Summary statistics of the measured NO_2 concentration at the stations, the ENSEMBLE simulations (at 10 km^2 spatial resolution) and the Bayesian GR (based on Model 2) estimates (at 1 and 10 km^2 spatial resolution).

Data	**Mean**	**Median**	**SD**	**Min**	**$q_{0.25}$**	**$q_{0.75}$**	**Max**
Stations NO_2 (measured)	22.07	20.05	12.39	0.18	13.31	28.50	89.12
ENSEMBLE NO_2 (simulations at 10 km^2)	14.07	13.26	6.60	0.68	9.34	17.80	40.73
Model 2 NO_2 (estimates at 10 km^2)	13.13	11.49	7.26	0.57	7.75	16.96	44.74
Model 2 NO_2 (estimates at 1 km^2)	19.12	18.35	9.31	0.57	12.51	24.84	61.18

Table A3: The logscore and cross-validation results of the Models 1 and 2 fitted to data from monitoring stations located in areas with large (Scenario 1) and small (Scenario 2) variability in the surface-to-column ratios.

Scenario	**Model**	**logscore**	**MAE**	**MAPE**	**RMSE**	**R^2**
Scenario 1	M1 (SAT)	3.354	0.294	0.241	0.399	0.694
	M2 (SAT - CTM)	3.352	0.292	0.240	0.398	0.696
Scenario 2	M1 (SAT)	3.193	0.244	0.078	0.322	0.722
	M2 (SAT - CTM)	3.193	0.244	0.078	0.322	0.722

Table A4: The logscore and cross-validation results of the models M0-M2 fitted to data from monitoring stations located in areas with high (Scenario 1) and low (Scenario 2) OMI availability (i.e. days within 2016 with non-missing OMI observations)

Scenario	**Model**	**logscore**	**MAE**	**MAPE**	**RMSE**	**R^2**
Scenario 1 (> 39% of days with data)	M0	3.267	0.279	0.136	0.365	0.711
	M1 (SAT)	3.258	0.275	0.136	0.361	0.718
	M2 (SAT - CTM)	3.258	0.275	0.136	0.360	0.720
Scenario 2 (< 39% of days with data)	M0	3.199	0.261	0.161	0.360	0.713
	M1 (SAT)	3.193	0.260	0.160	0.360	0.713
	M2 (SAT - CTM)	3.193	0.260	0.159	0.359	0.714

References

Copernicus Land Monitoring Services (CLMS), 2019. Pan-European data products. http://land.copernicus.eu/pan-european (accessed 01-03-2019).

[Corine Land Cover 2012] Copernicus Land Monitoring Services, CORINE Land Cover 2012. http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view (accessed 01-12-2018).

[DEM 2012] European Environment Agency, 2012. Digital Elevation Model over Europe (EU-DEM). http://www.eea.europa.eu/data-and-maps/data/eu-dem (accessed 1 March 2019).

[ECS 2015] European Environment Agency, 2015. Europe coastline shapefile. http://www.eea.europa.eu/data-and-maps/data/eea-coastline-for-analysis-1/gis-data/europe-coastline-shapefile (accessed 1 March 2019).

EuroStat–GISCO service (2019). Eurogeographics for the administrative boundaries. http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units (accessed 1 March 2019).

Google Earth Engine Team, 2015. Google Earth Engine: A planetary-scale geospatial analysis platform. https://earthengine.google.com

[Imperviousness Density 2015] Copernicus Land Monitoring Services, High Resolution Layer: Imperviousness Degree 2015. https:

//land.copernicus.eu/pan-european/high-resolution-layers/imperviousness/status-maps/2015 (accessed 1 March 2019).

[LST 2016] NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (2016). MYD11A1.005 (Aqua) and MOD11A1.005 (Terra) Land Surface Temperature and Emissivity Daily Global 1 km Grid SIN. https://lpdaac.usgs.gov (accessed 1 March 2019).

[NDVI 2016] MODIS Aqua and Terra Daily Normalized Difference Vegetation Index (NDVI), Google

[NTL 2013] NOAA's National Geophysical Data Center. DMSP-OLS Nighttime Lights Time Series Version 4. https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed 1 March 2019).

Saha, S., et al., 2011. NCEP Climate Forecast System Version 2 (CFSv2) 6-hourly Products, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://dx.doi.org/10.5065/D61C1TXF

Socioeconomic Data and Applications Center (SEDAC), 2016. Center for International Earth Science Information Network - CIESIN - Columbia University. Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision UN WPP Country Totals. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4HX19NJ

[Tree Cover Density 2015] Copernicus Land Monitoring Services, High Resolution Layer: Tree Cover Density 2015. https://land.copernicus.eu/pan-european/high-resolution-layers/forests/tree-cover-density/status-maps/2015 (accessed 1 March 2019).