Supplement of Atmos. Chem. Phys., 26, 7741–7764, 2026
https://doi.org/10.5194/acp-26-7741-2026-supplement

Atmospheric Chemistry and Physics Open Access

Supplement of

Technical note: DACNO$_2$ – a multi-constraint deep learning framework for high-resolution 3D NO$_2$ field estimation

Wenfu Sun et al.

Correspondence to: Wenfu Sun (wenfu.sun@aeronomie.be)

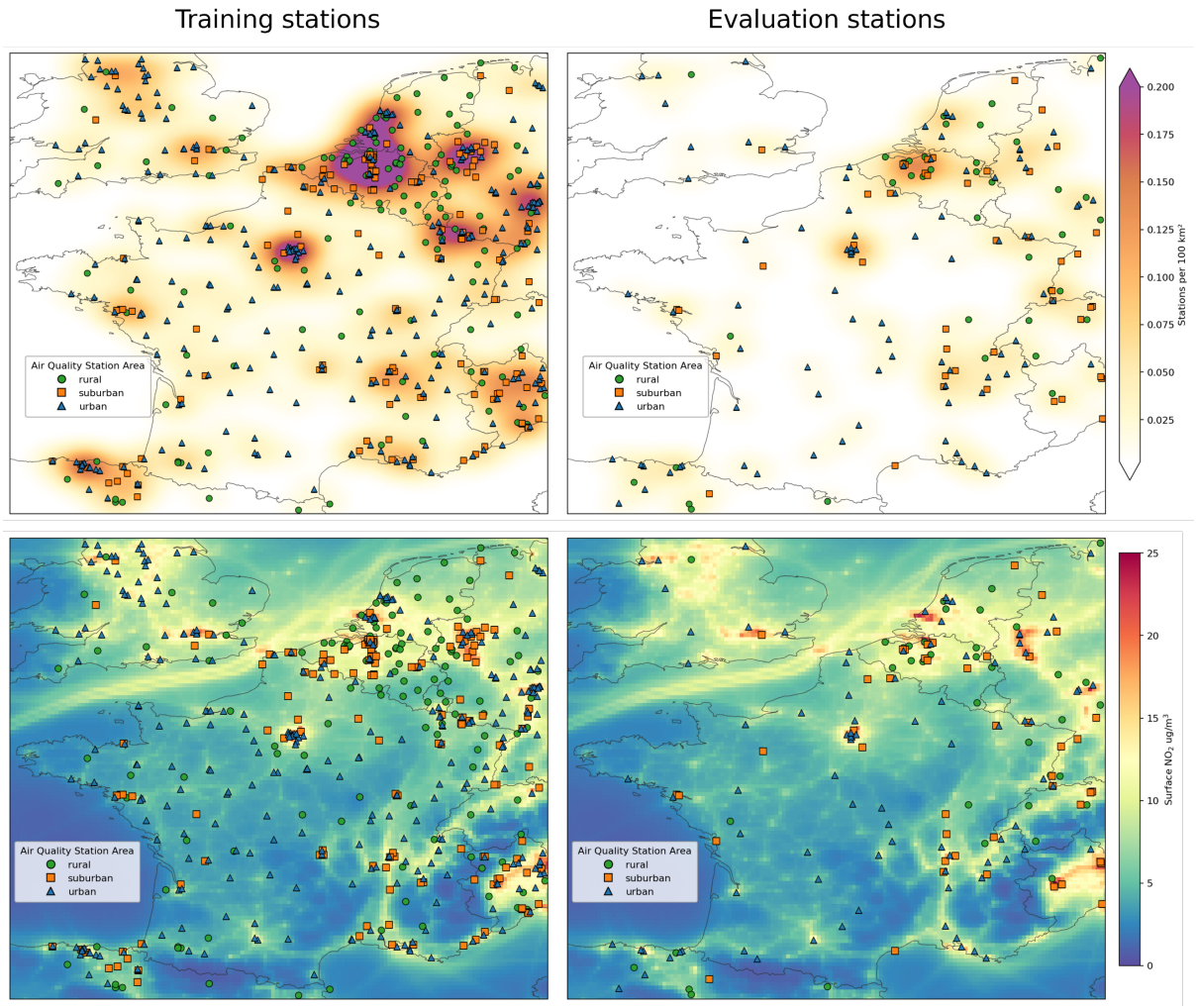


Figure S1. Spatial distribution of EEA surface NO_2 monitoring stations over the study domain, with station density and background surface NO_2 levels for context. EEA air-quality monitoring stations used in this work are shown by station area type (rural, suburban, urban). The upper panels show the spatial station density (stations per 100 km^2), estimated by smoothing station locations using a Gaussian kernel density estimator (KDE) on a 2 km grid with a bandwidth of 35 km. The lower panels show the CAMS 2023 mean surface NO_2 concentration as background context, allowing the station distribution to be interpreted relative to typical pollution levels across the domain.

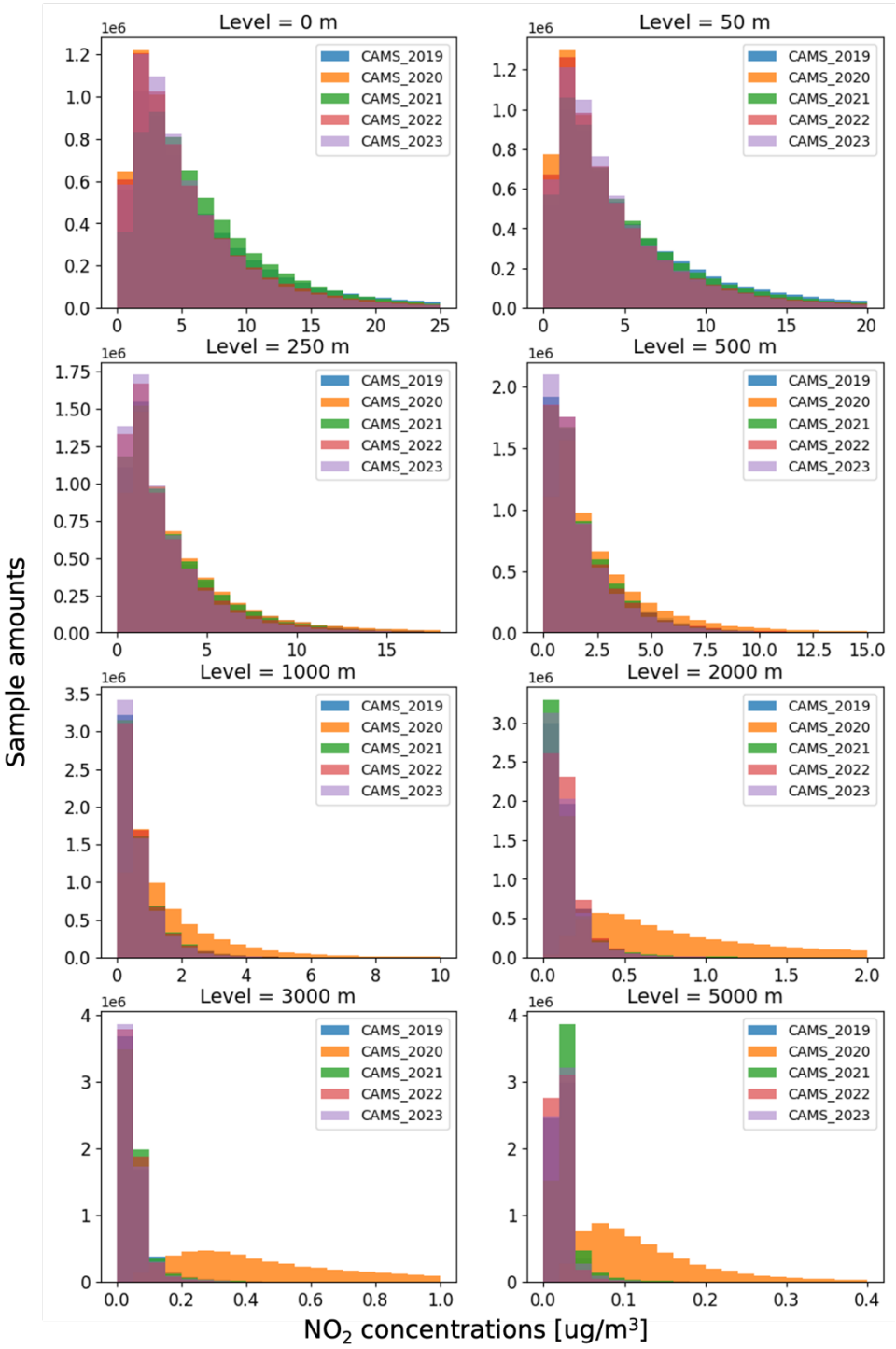
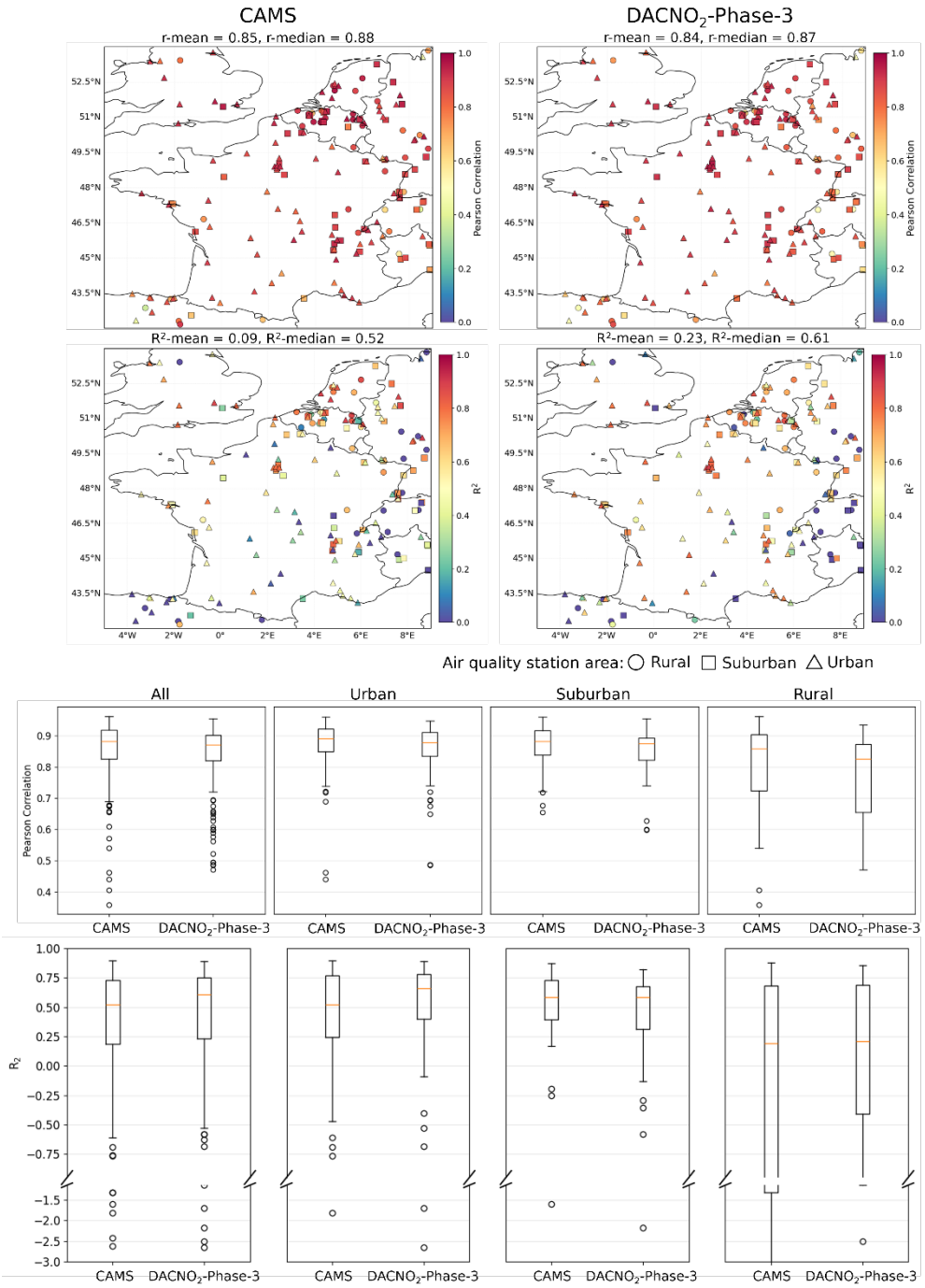


Figure S2. Distributions of CAMS NO_2 concentrations for 2019–2023 at each vertical level. Each panel displays the frequency histogram of daily grid cell NO_2 values (μg/m^3) at various heights (0, 50, 250, 500, 1000, 2000, 3000, and 5000 m) for each year. Notably, 2020 exhibits a marked shift toward higher NO_2 values above 1000 m compared to other years, indicating anomalies in the data for that year.



25 **Figure S3.** Station-specific performance of CAMS and DACNO2-Phase-3 against EEA measurements in 2023. Model performance is evaluated at individual EEA stations using daily surface NO_2 measurements. For each station, statistics are calculated from the paired daily time series of measurements and model estimates over 2023. The upper panels show the spatial distribution of Pearson correlation coefficient (r) and coefficient of determination (R^2) for CAMS and DACNO2-Phase-3. Station types (urban, suburban, rural) are indicated by
30 different marker shapes. The lower panels summarize the same station-specific statistics using boxplots, allowing comparison of the distribution of performance metrics across different station groups. In each boxplot, the horizontal line indicates the median value. This figure characterizes how model skill varies geographically and across station environments at the individual-station scale.

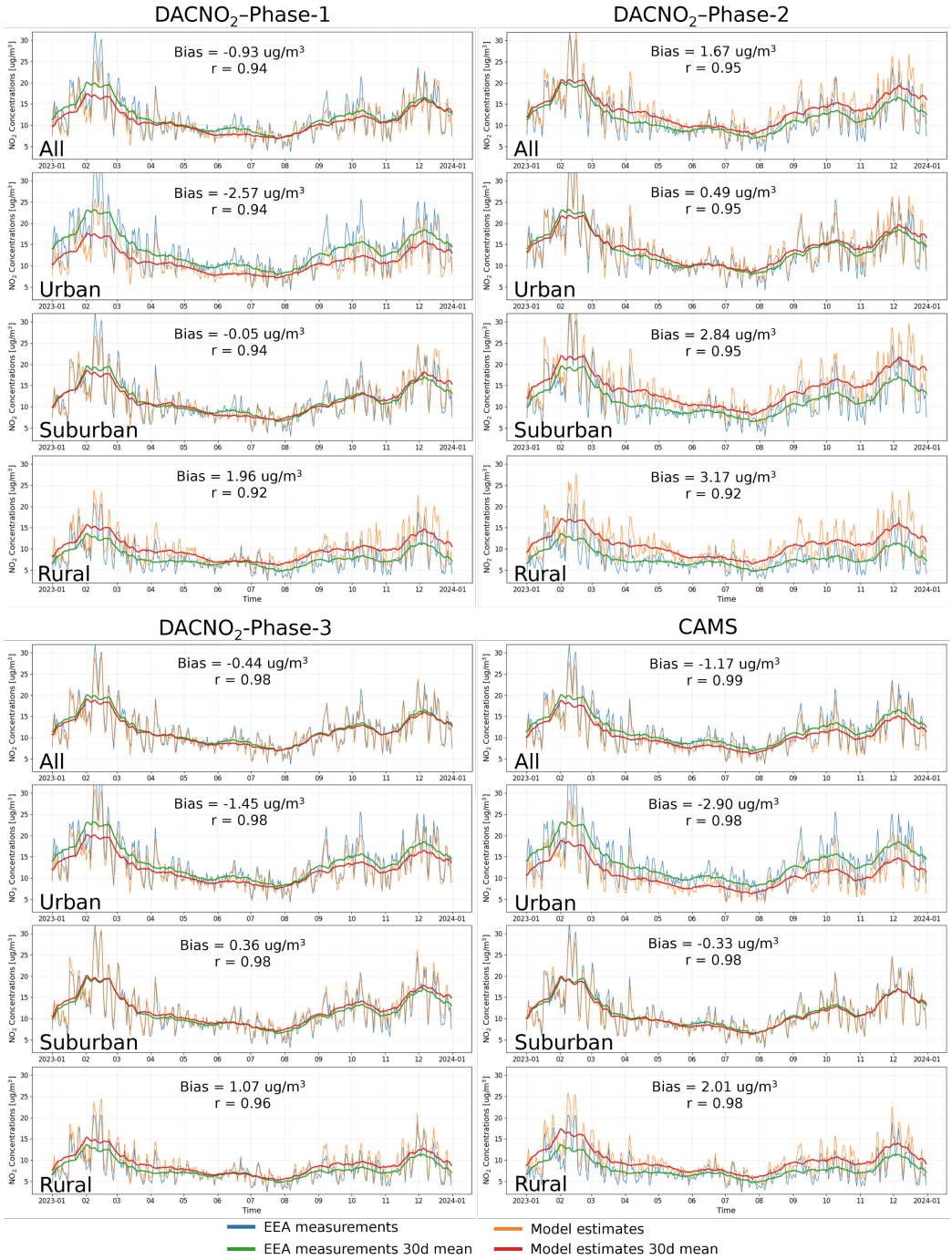


Figure S4. Daily comparison of surface NO_2 measurements and model estimates in 2023. Daily mean surface NO_2 time series for 2023 are shown for EEA measurements and for model estimates from CAMS and the three $DACNO_2$ phases. Results are presented for all stations and separately for urban, suburban, and rural stations according to EEA metadata. For each day, two station-network means are computed independently for measurements and model estimates by averaging over the stations where both values are available on that day. Thin lines show the resulting daily network means, and thick lines indicate the 30-day moving average to emphasize seasonal variability. Performance metrics (bias and Pearson correlation coefficient r) are calculated from the paired daily network-mean time series over the full year.

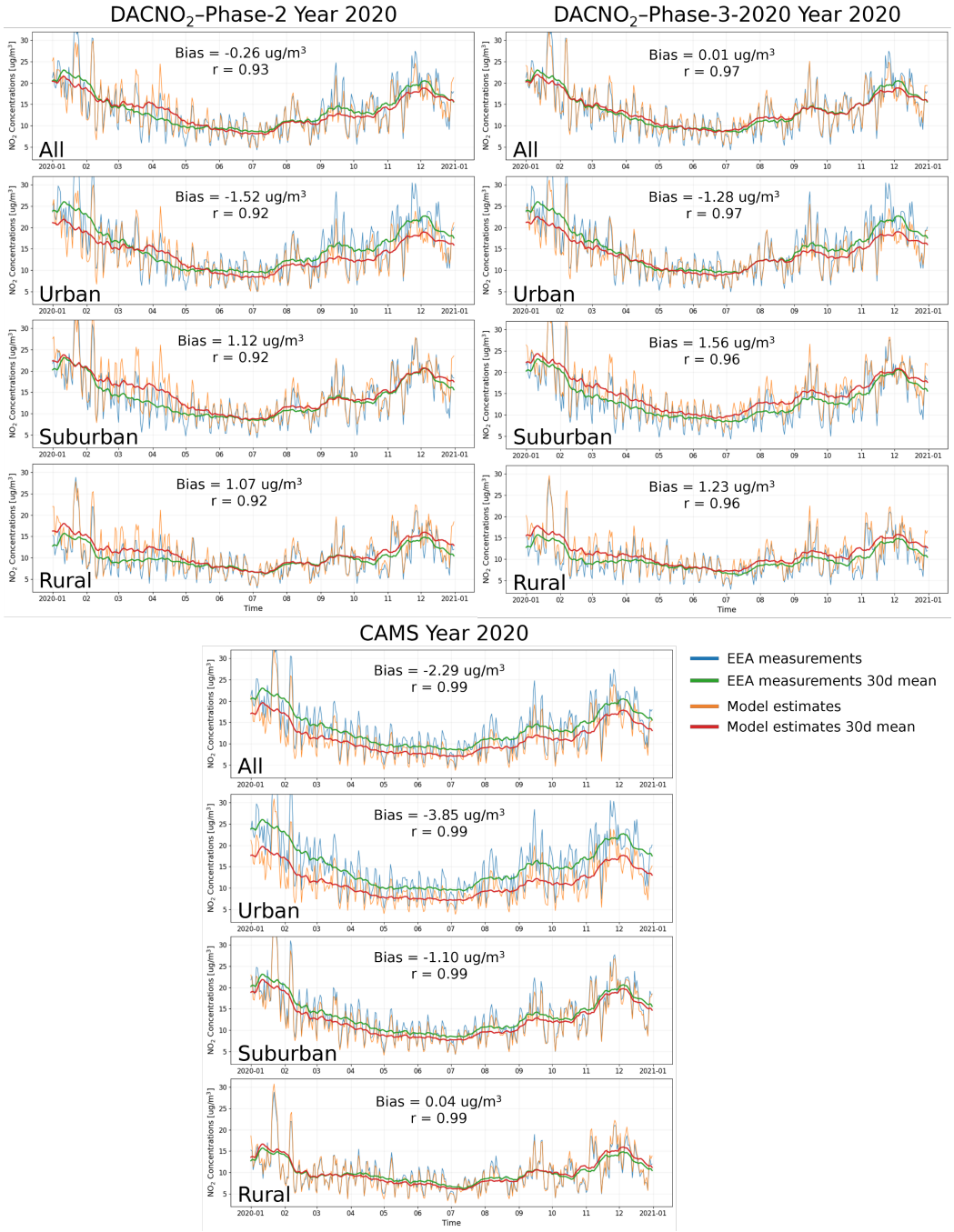


Figure S5. Daily comparison of surface NO_2 measurements and model estimates in 2020. Daily mean surface NO_2 time series for 2020 are shown for EEA measurements and for model estimates from CAMS and DACNO$_2$ (Phase-2 and Phase-3-2020). Results are presented for all stations and separately for urban, suburban, and rural stations according to EEA metadata. For each day, two station-network means are computed independently for measurements and model estimates by averaging over the stations where both values are available on that day. Thin lines show the daily network means, and thick lines indicate the 30-day moving average to highlight seasonal variability. Performance metrics (bias and Pearson correlation coefficient r) are calculated from the paired daily network-mean time series over the full year. In this figure, the model of DACNO$_2$-Phase-2 is identical to that shown in Fig. S4, while DACNO$_2$-Phase-3-2020 is derived by fine-tuning DACNO$_2$-Phase-2 with EEA training stations for 2020.

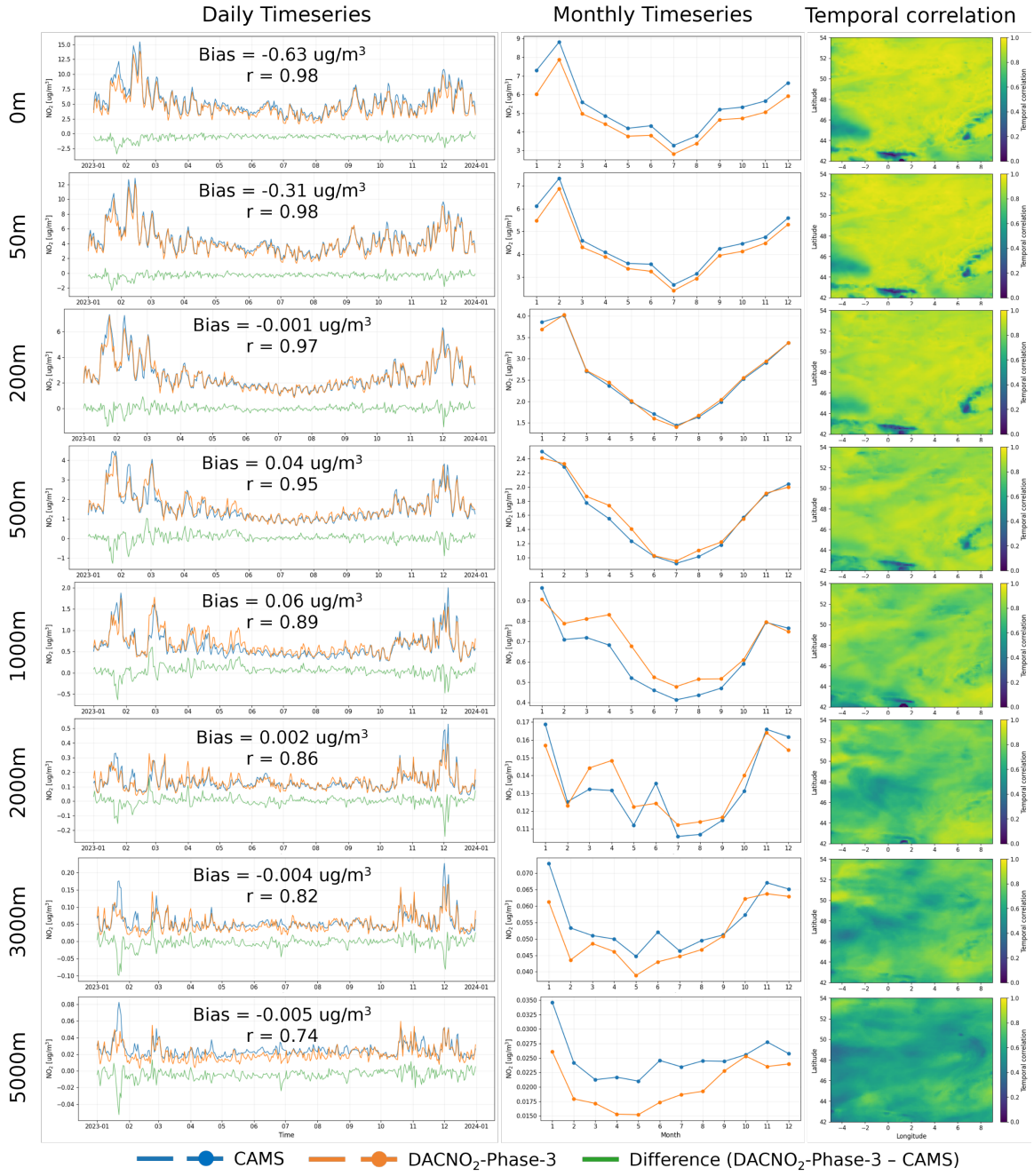


Figure S6. Layer-wise comparison between DACNO$_2$-Phase-3 and CAMS at 10 km resolution for 2023. DACNO$_2$-Phase-3 fields produced on the 2 km grid are aggregated to the 10 km grid and compared with CAMS for each vertical layer. The left column shows domain-mean daily time series of CAMS and DACNO$_2$-Phase-3 together with their difference (DACNO$_2$-Phase-3 minus CAMS). The middle column shows the corresponding monthly means. Reported statistics (bias and Pearson correlation coefficient) are calculated from the paired daily domain-mean series. The right column shows the spatial distribution of the temporal Pearson correlation between DACNO$_2$-Phase-3 and CAMS at each grid cell derived from the full-year daily time series.

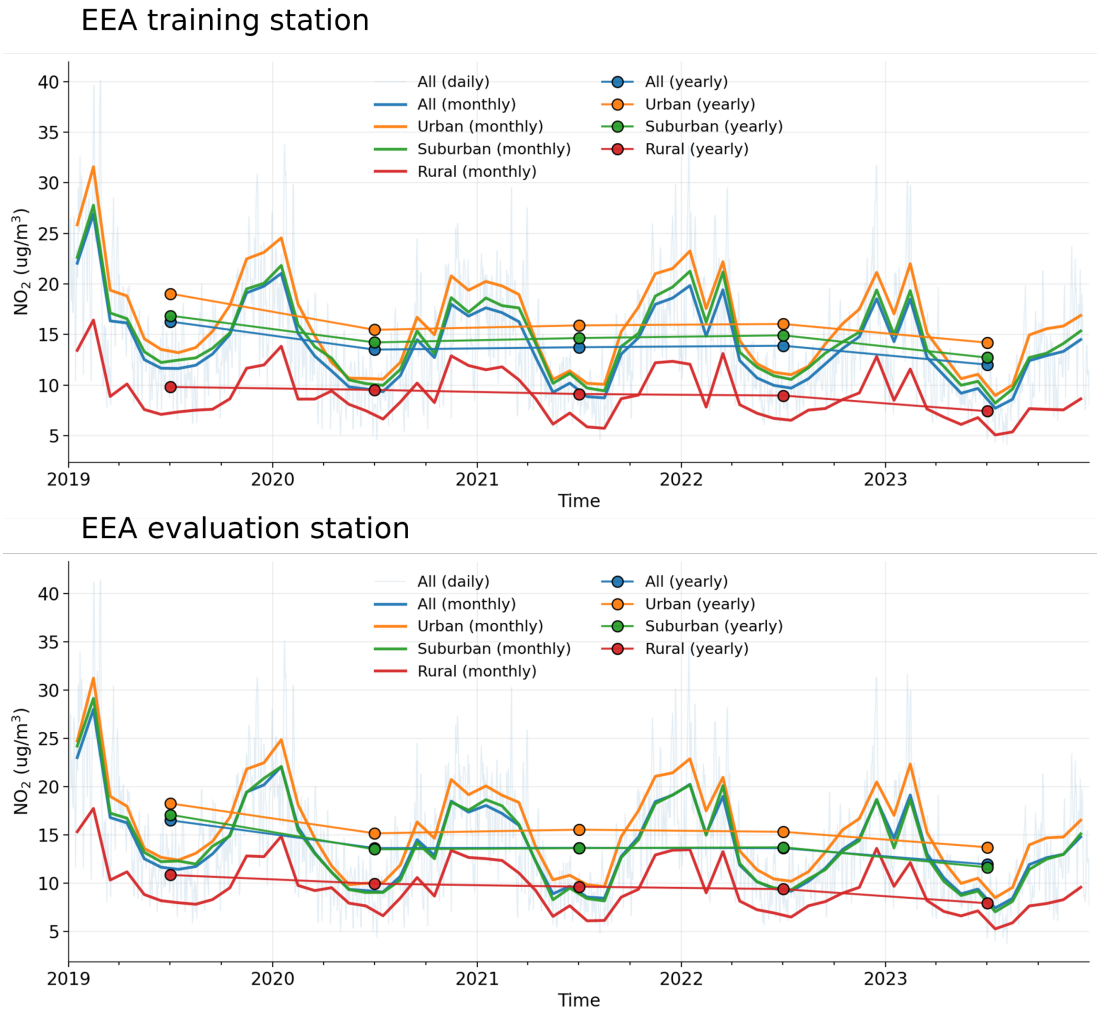


Figure S7. Long-term surface NO_2 measurements at EEA training and evaluation stations (2019–2023). Time series of measured NO_2 concentrations from EEA monitoring stations between 2019 and 2023. The upper panel shows stations used for model training, and the lower panel shows independent evaluation stations. The light blue background line represents the daily network-mean concentration, computed as the average across all available stations for each day. Colored curves indicate monthly averages for all, urban, suburban, and rural stations, while filled markers denote the corresponding yearly average. This figure characterizes the temporal variability and relative concentration levels across station types used in the study.

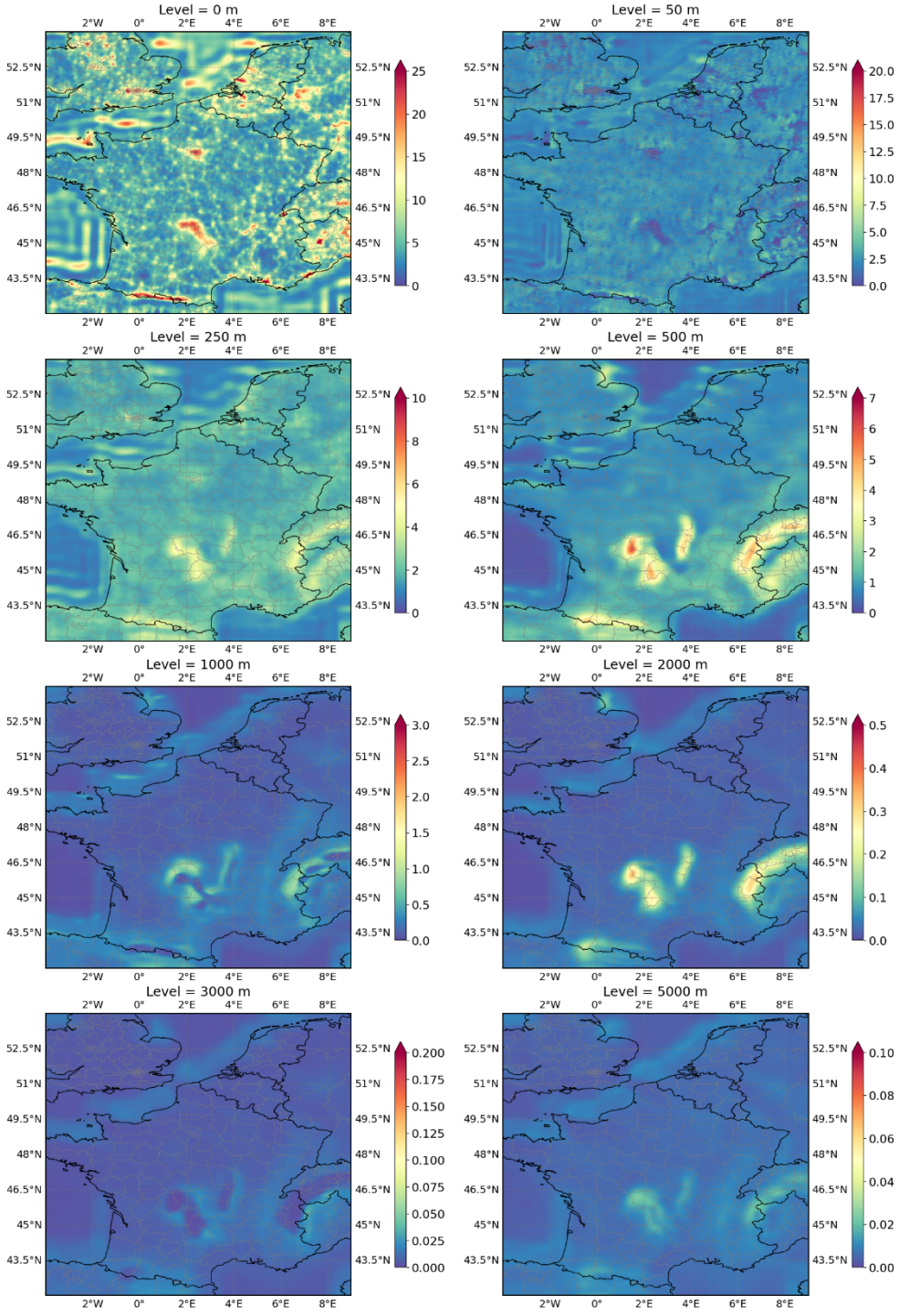


Figure S8. Annual mean NO_2 fields for 2023 from the DACNO₂-onlyobs model at eight vertical levels (0, 50, 250, 500, 1000, 2000, 3000, and 5000 m). This test model was trained exclusively with EEA surface NO_2 measurements, following a similar Phase-2 and Phase-3 process but without CAMS NO_2 constraints. Results show that effective NO_2 estimates are produced primarily near the land surface, while artifacts appear at higher levels and over the ocean, highlighting the need for physically consistent simulated data to support spatial and vertical constraints in high-resolution 3D NO_2 modeling.

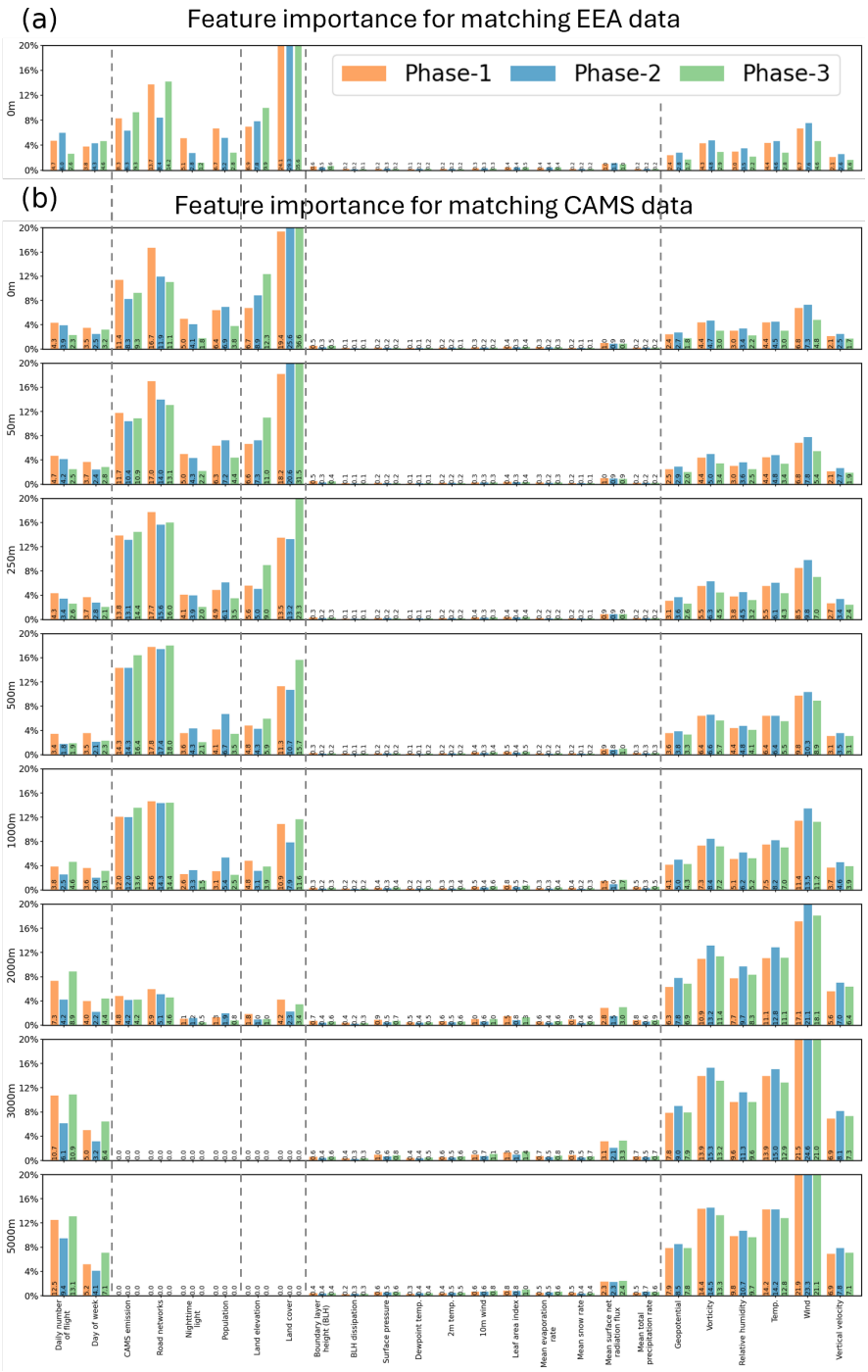


Figure S9. Relative importance of each individual input feature for DACNO$_2$ model predictions, evaluated using the integrated gradients (IG) method. (a) Feature contributions to RMSE between DACNO$_2$ surface NO$_2$ estimates and EEA ground-based measurements for 2023. (b) Feature contributions to RMSE between DACNO$_2$ and CAMS NO$_2$ estimates at different vertical levels for 2023. Features are grouped (separated by dashed lines) into temporal indicators, emission inventories and proxies, geography, ERA5 single-level meteorology, and ERA5 multi-level meteorology. Results are shown for each model training phase (i.e., Phase-1, Phase-2, and Phase-3), illustrating how the relative influence of individual input features varies with training constraints and altitude. Feature definitions and sources are provided in Table 1.

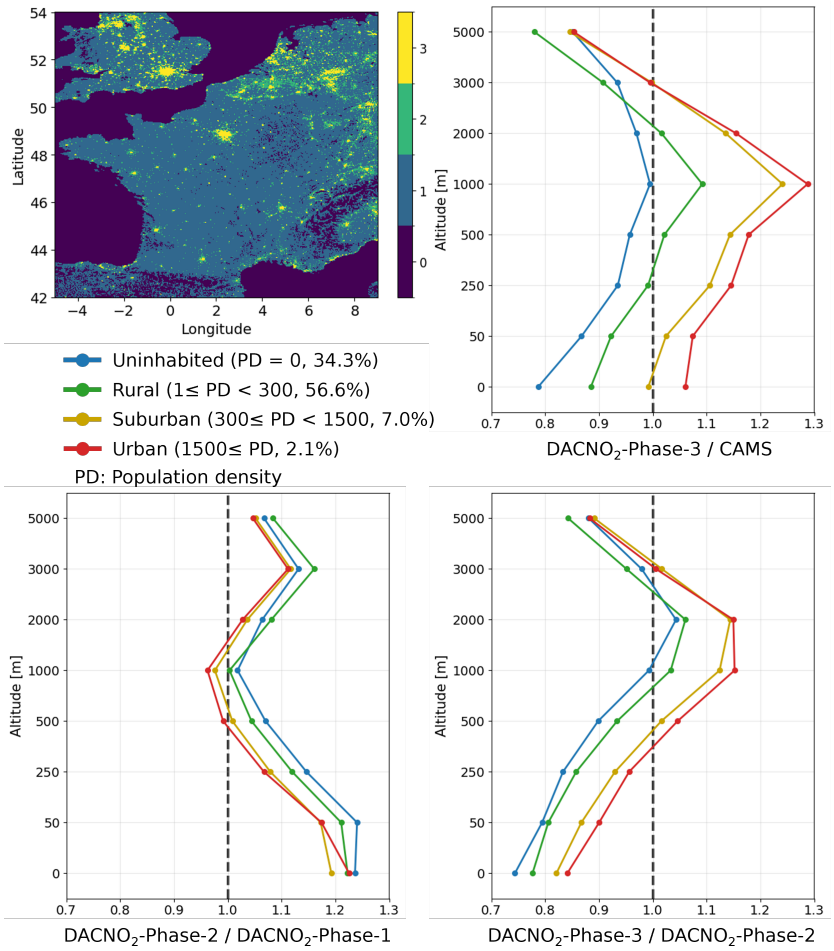


Figure S10. Vertical profile relative changes across urban, suburban, rural, and uninhabited environments. Relative changes of vertical NO_2 profiles between $DACNO_2$ model phases and CAMS across different environments. The upper-left panel shows the spatial distribution of four environment types derived from population density (0-uninhabited, 1-rural, 2-suburban, 3-urban) based on the JRC-GEOSTAT 2018 gridded population dataset (Silva et al., 2021) and the urbanization definition of Dijkstra et al. (2021). The remaining panels present mean vertical profile ratios for each environment type: $DACNO_2$-Phase-3 relative to CAMS (upper right), $DACNO_2$-Phase-2 relative to $DACNO_2$-Phase-1 (lower left), and $DACNO_2$-Phase-3 relative to $DACNO_2$-Phase-2 (lower right). Profiles are averaged across all grid cells within each class. The dashed vertical line denotes a ratio of 1.0, indicating no change between the two compared datasets.

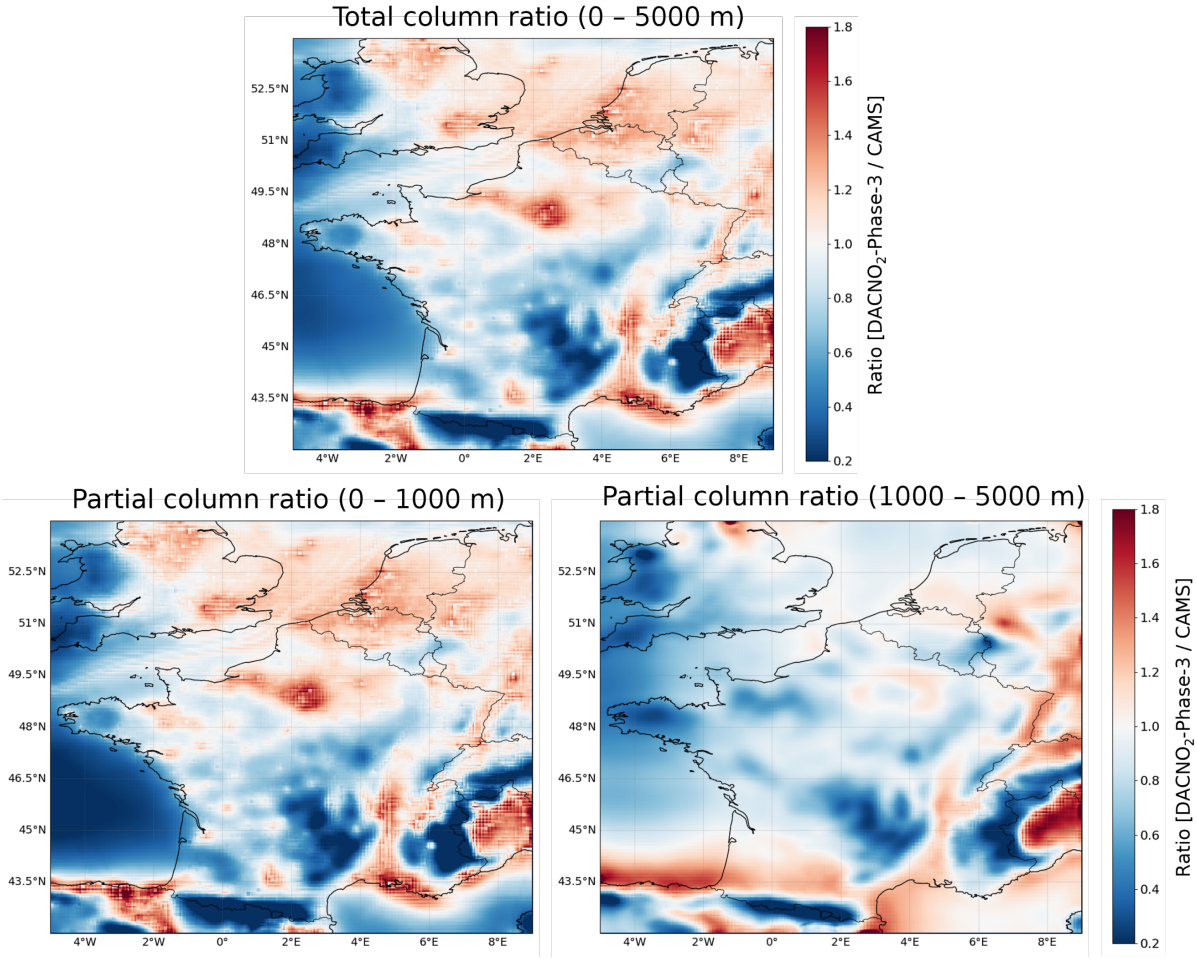


Figure S11. Comparison of vertically integrated NO_2 between DACNO$_2$-Phase-3 and CAMS. Ratios between DACNO$_2$-Phase-3 and CAMS for vertically integrated NO_2 over the study domain. The upper panel shows the full-column layer-integrated proxy, while the lower panels show integrations over 0–1000 m and 1000–5000 m. The integrated quantity is obtained by summing concentration multiplied by layer thickness on the common model levels, providing a consistent vertical mass proxy for relative comparison between the two models. A diverging color scale centered at 1 highlights relative enhancement and reduction. Area-weighted domain-mean ratios are 0.995 for the full column, 1.000 for 0–1000 m, and 0.949 for 1000–5000 m, indicating that the total column is largely preserved despite redistribution.

Table S1. Evaluation performance of DACNO$_2$-S5P

Year 2023 UTC11-13 (2 km × 2 km)	DACNO$_2$-S5P				CAMS-S5P-2km			
EEA-S5P	RMSE (ug/m^3)	R	R^2	Bias (ug/m^3)	RMSE (ug/m^3)	r	R^2	Bias (ug/m^3)
Total	5.07*	0.77*	0.59*	0.05*	5.27	0.76	0.55	-0.94
Urban	5.41*	0.74	0.54*	-0.63*	5.60	0.76*	0.51	-1.98
Suburban	5.08*	0.79*	0.62*	0.63*	5.24	0.78	0.60	-0.67
Rural	4.08*	0.80*	0.61*	1.01*	4.40	0.78	0.55	1.27

(10 km × 10 km)	DACNO$_2$-S5P-10km			
CAMS-S5P	RMSE (ug/m^3)	r	R^2	Bias (ug/m^3)
All layers	0.98	0.94	0.88	0.03
0 m	1.82	0.91	0.82	0.06
50 m	1.52	0.92	0.85	-0.06
250 m	1.11	0.92	0.84	0.14
500 m	0.78	0.89	0.80	0.06
1000 m	0.36	0.83	0.68	0.03
2000 m	0.08	0.70	0.49	-0.01
3000 m	0.03	0.63	0.39	0.00
5000 m	0.01	0.53	0.22	0.00

Note: This table is similar to Tables 2 and 3. The DACNO$_2$-S5P model is a Phase-3 model developed for the TROPOMI overpass time, predicting a 3-hour average NO$_2$ (11:00–13:00 UTC) using the three-phase strategy. CAMS-S5P and EEA-S5P represent process-based and measured NO$_2$ data during overpass time. CAMS-S5P-2km is derived by bilinearly interpolating the CAMS-S5P data. For the comparison against EEA NO$_2$ (shown in the upper panel), both DACNO$_2$ outputs and CAMS NO$_2$ were evaluated on the 2 km grid. For evaluating DACNO$_2$ using CAMS NO$_2$ (shown in the lower panel), DACNO$_2$ outputs were downsampled and evaluated at the original 10 km × 10 km resolution of CAMS across all vertical layers as well as for individual layers. Best values within each row are marked with an asterisk (*).

References

Silva, F. B. E., Poelman, H., and Dijkstra, L.: JRC-GEOSTAT 2018 population grid, version 2021.02.10, Eurostat GISCO [data set], https://ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/population-grids (last access: 1 June 2024), 2021.

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., and Schiavina, M.: Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation, Journal of Urban Economics, 125, 103312, 10.1016/j.jue.2020.103312, 2021.