Description of the layout

1. Black: Referee's comments
2. Blue: Author's response
3. *Blue, italic*: The revised content in the manuscript and the supplement.
4. Red: line numbers of revised content; revised tables; complemented figures.

Response to Reviewer #1

Review on Technical Note: DACNO$_2$ – A Multi-Constraint Deep Learning Framework for High- Resolution 3D NO$_2$ Field Estimation

In this manuscript, Sun et al. present a new machine learning (ML) model that estimates the daily average 3D distribution of NO$_2$ in central Europe. The model is trained on data from the CAMS ensemble CTM and in situ observations at the surface. Because NO$_2$ plays an important role in atmospheric chemistry and air pollution, and the model proposed by Sun et al. outperforms CAMS, the results of this manuscript are definitely worth publishing and are well in scope of ACP.

Before listing my comments below, I would like to emphasize the strengths of the manuscript in its current form. The manuscript contains several clever ideas, e.g., using flight data as a proxy for exceptional periods, such as holidays. It is apparent from the detailed explanations of the training procedure that the presented work was overall well conceptualized and executed. The model is quite complex and represents a significant body of work. The model's evaluation is detailed.

I recommend to address the following comments.

Thank you for your positive assessment and constructive comments. We have revised the manuscript accordingly and provide point-by-point responses below.

General comments

1) Why is the manuscript submitted as a technical note? The novelty and quality of the proposed model is surely sufficient for this manuscript to be a regular research article at ACP.

We thank Reviewer#1 for the suggestion. We submitted the manuscript as a Technical Note because its primary contribution is methodological. The study introduces and validates DACNO$_2$, a multi-constraint deep learning framework and workflow to generate daily 2 km × 2 km resolution 3D NO$_2$ fields over Western Europe, and demonstrates its use as a priori profiles for satellite retrievals. ACP defines Technical Notes as papers reporting new developments in methods and techniques, including numerical algorithms for interpreting atmospheric data, such as statistical and machine learning methods, which aligns with the central aim and structure of our manuscript. In contrast, ACP Research Articles are framed around new results and conclusions from scientific investigations of atmospheric properties and processes. While we discuss physical consistency and demonstrate application relevance, we do not further analyze the underlying physical and chemical processes in this manuscript.

The definition of ACP manuscript types can be found via this link (https://www.atmospheric-chemistry-and-physics.net/about/manuscript_types.html). We are happy to follow the handling editor's recommendation if reclassification is preferred.

2) It is interesting to see that satellite data were deliberately not used as input to DACNO$_2$. Nonetheless, satellite data would have resolved a current weak point of the model by capturing the temporal variability of the NO$_2$ burden. Currently, DACNO$_2$ appears to use static emission data (average of the year 2018). Looking at the other input data from Table 1, I cannot identify any input variable from which the model could determine trends across seasons and years. Re-training the model (which apparently took > 3 weeks) is surely too much to ask, but the fact that DACNO$_2$ is essentially barren of any temporal dynamic on the seasonal/yearly scale should be discussed.

We agree that satellite NO$_2$ data is valuable because it provides timely, direct observations of NO$_2$ spatiotemporal variations, and many published machine-learning-based NO$_2$ estimation studies use it as an important input feature (Sun et al., 2024; Kim et al., 2021; Li and Wu, 2021). However, satellite NO$_2$ data often have missing values, introducing additional uncertainty into the model, even though some works have investigated gap-filling methods for this dataset. In our study region, the TROPOMI NO$_2$ data missing ratio can reach around 50% (Sun et al., 2024), making it challenging to use satellite data as a reliable input to the model. Therefore, we would prefer to use the satellite NO$_2$ data as a training constraint in future work, which would allow the use of only the valid part of the data. To assess whether the DACNO$_2$ model still retains the ability to capture temporal variations of NO$_2$, we conducted additional experiments to specifically examine the DACNO$_2$ model's temporal agreement with EEA evaluation stations and CAMS NO$_2$ during the test year (2023) and the COVID-19 year (2020). Results are shown in three new supplementary figures (Figs. S4, S5, and S6).

It is found that DACNO$_2$ has captured the temporal variability of NO$_2$ since Phase-1 and has maintained it until Phase-3 (Figs. S4 and S6). This means that coupling the meteorological drivers, temporal indicators, and the CAMS constraint enables the DACNO$_2$ model to learn the temporal dynamics of NO$_2$, while static inputs and EEA constraints primarily adjust magnitude and map high-resolution spatial patterns. Nevertheless, such coupling cannot fully capture the NO$_2$ change in the anomaly period, as shown in Fig. S5 for the COVID-19 year (2020). The DACNO$_2$-Phase-2 exhibits a systematic overestimation for March to May when the COVID-19 control rules take place, while it still captures the overall seasonal NO$_2$ variations. Such bias has been mitigated through adaptive fine-tuning using EEA station data, as shown in the DACNO$_2$-Phase-3-2020 panel.

In the manuscript, we have added "*The average time-series consistency between models and EEA NO$_2$ is shown in Fig. S4*" in line 390, and "*The layer-wise temporal correlations at the regional average and grid scales are illustrated in Fig. S6*" in lines 392-393.

We have added "*Fig. S4 shows that the DACNO$_2$ model learns reliable temporal correlations with EEA NO$_2$ at the daily and seasonal scales since Phase-1 (r = 0.94), and these correlations are further enhanced in Phase-2 (r = 0.95) and Phase-3 (r = 0.98). This indicates that the model can represent temporal variability without using satellite NO$_2$ as an input, relying instead on meteorological and temporal indicators*" in lines 404-407.

We have added "*Fig. S5 shows the temporal trend between model estimations and EEA measurements. It is observed that DACNO$_2$-Phase-2 can still capture the temporal trend of NO$_2$ in this unknown and anomalous year, while a positive bias appears during March and May when COVID-19 control measures took place. DACNO$_2$-Phase-3-2020 has successfully reduced the remaining bias with the adaptive fine-tuning. This demonstrates the robustness of the DACNO$_2$ model and the necessity of adaptive fine-tuning to capture anomalous events. Additionally, CAMS maintains overall consistency across measurements but exhibits a pronounced negative bias, primarily in urban areas*" in lines 687-693.

The added figures in the Supplementary Materials are shown below:

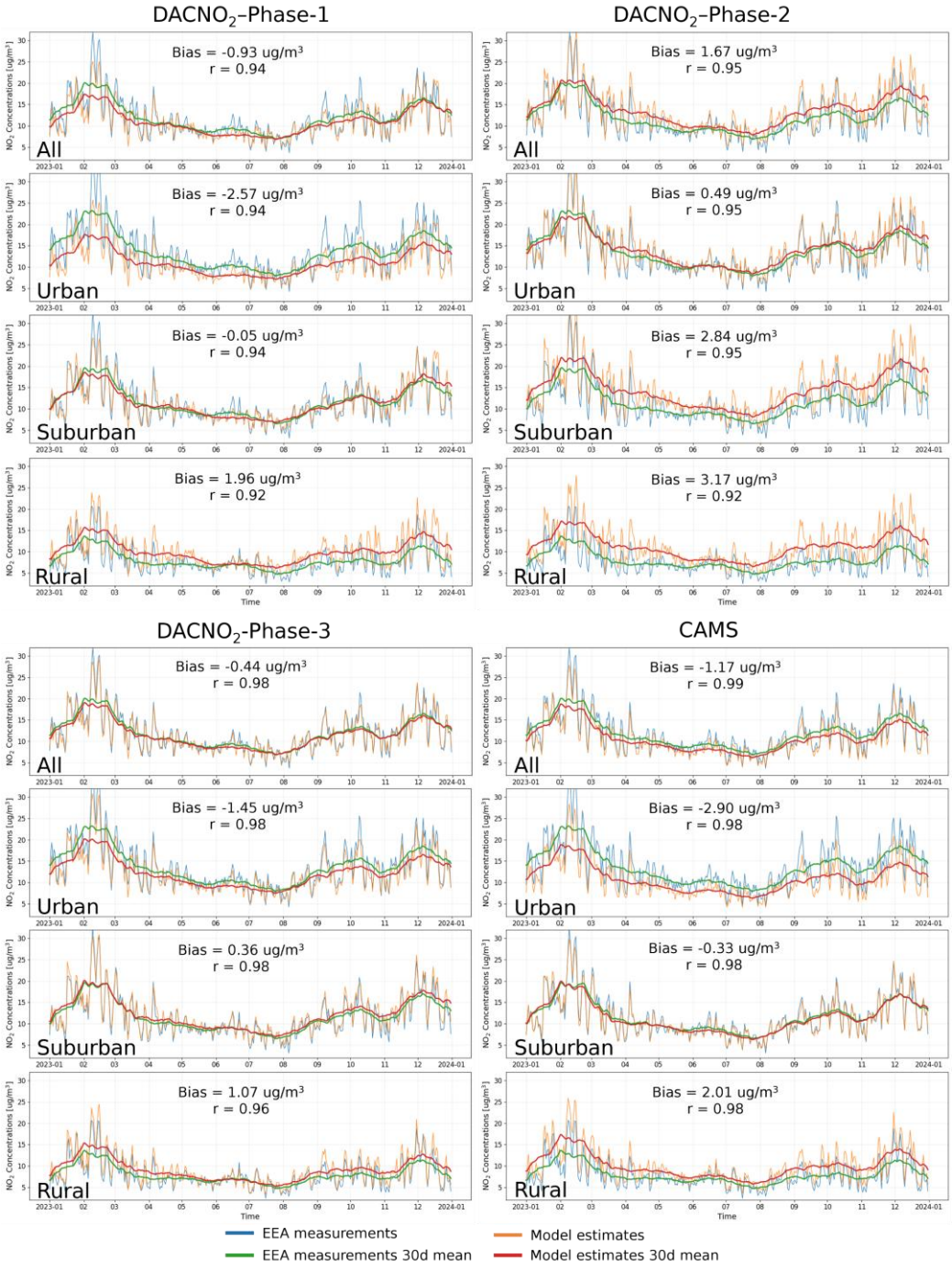

Figure S4. Daily comparison of surface NO_2 measurements and model estimates in 2023. Daily mean surface NO_2 time series for 2023 are shown for EEA measurements and for model estimates from CAMS and the three DACNO$_2$ phases. Results are presented for all stations and separately for urban, suburban, 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.

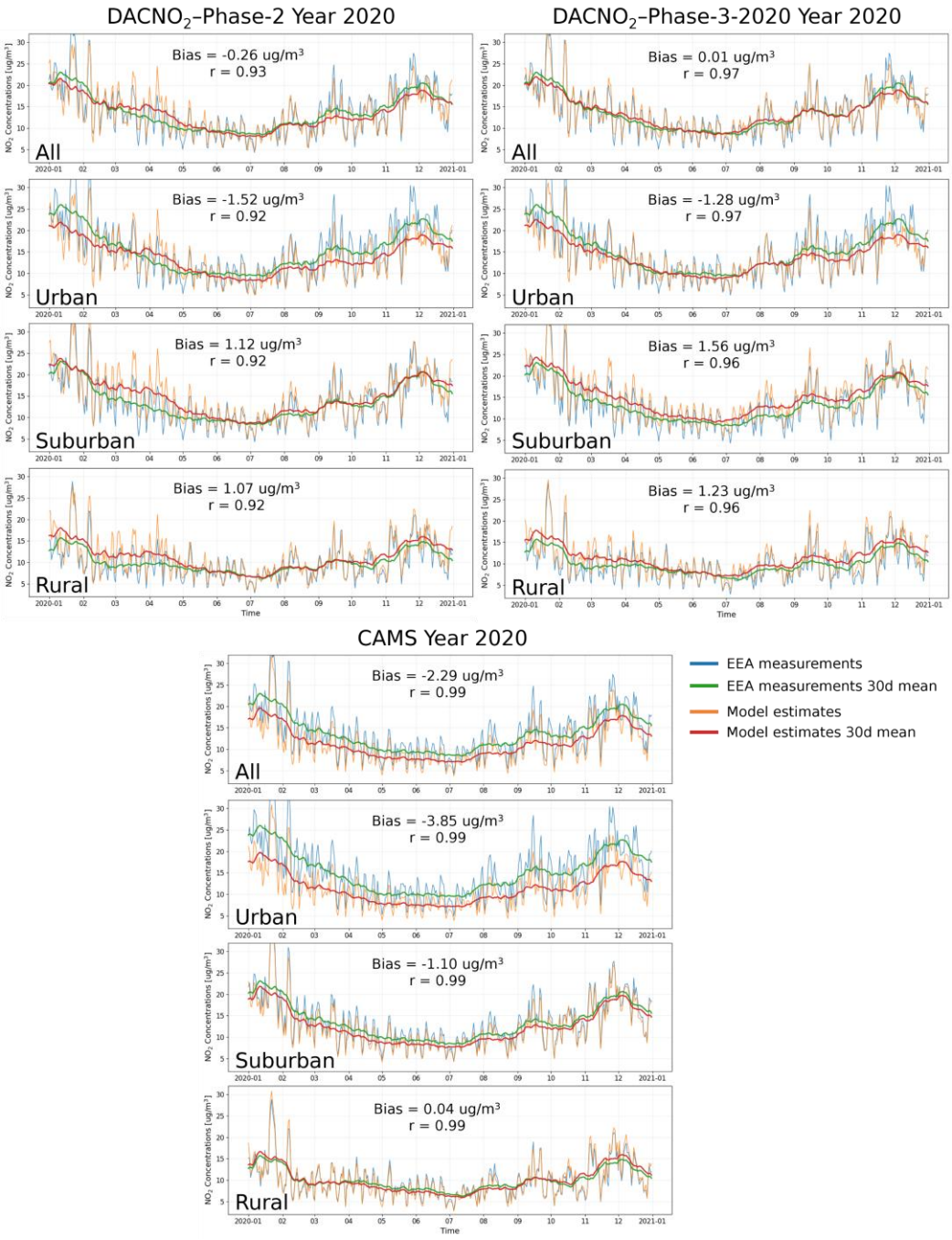

Figure S5. Daily comparison of surface NO_2 measurements and model estimates in 2020. Daily mean surface NO_2 time series for 2020 are shown for EEA measurements and for model estimates from CAMS and DACNO$_2$ (Phase-2 and Phase-3-2020). Results are presented for all stations and separately for urban, 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.

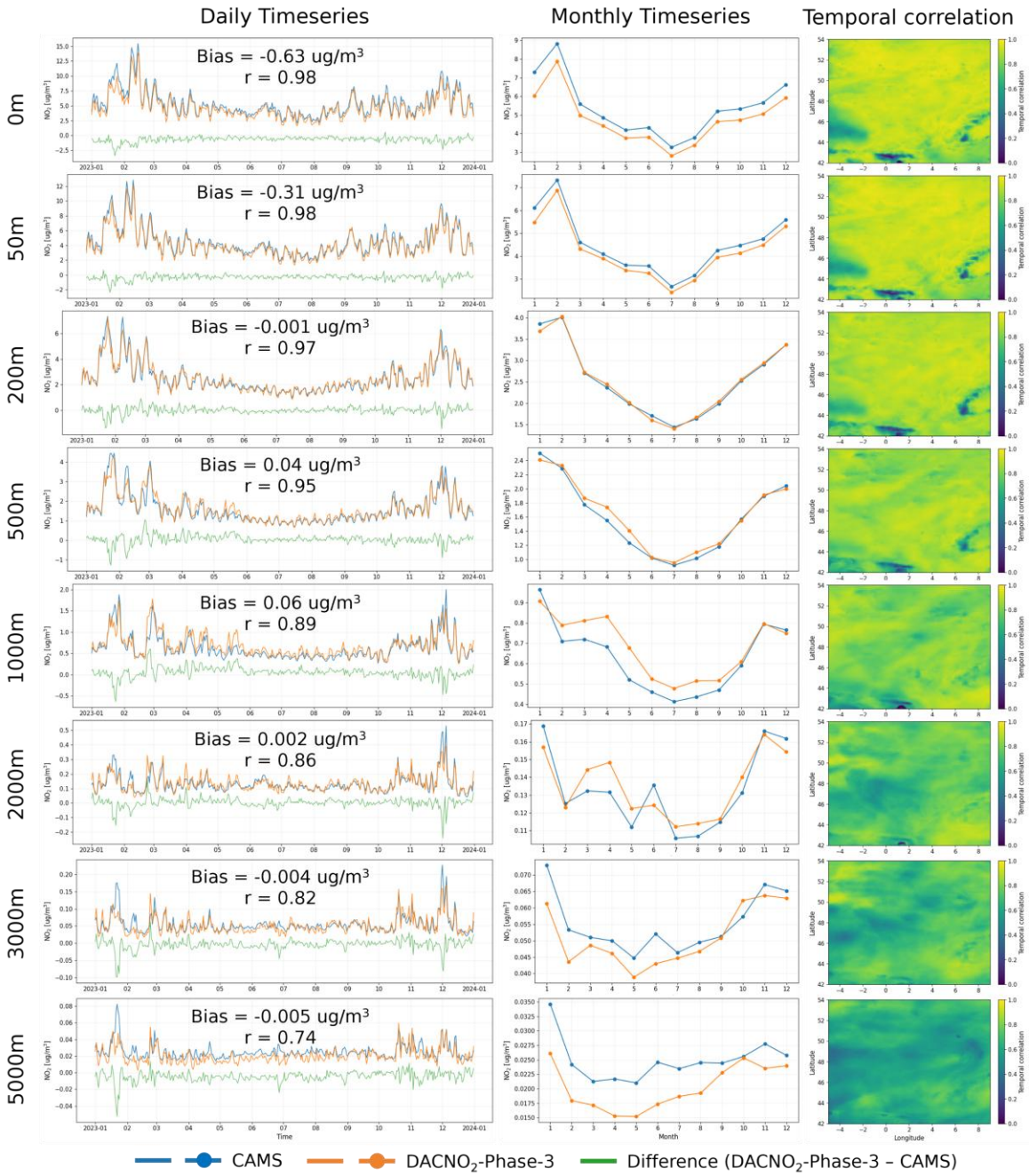

Figure S6. Layer-wise comparison between DACNO$_2$-Phase-3 and CAMS at 10 km resolution for 2023. DACNO$_2$-Phase-3 fields produced on the 2 km grid are aggregated to the 10 km grid and compared with CAMS for each vertical layer. The left column shows domain-mean daily time series of CAMS and DACNO$_2$-Phase-3 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.

3) A significant drawback of training on CAMS data is that they cannot account for one of the most significant measurement biases of in situ instrument: The NO_y bias of molybdenum-based photolysis converters that many instruments in Europe use. This has been

discussed elsewhere, see e.g. Lamsal et al. (2008) and Villena et al. (2012). Although the problem cannot be resolved based on CAMS data, it should be acknowledged in the paper.

We are grateful to the reviewer for mentioning the bias in EEA measurements. We have added "*It is important to note that some in-situ NO_2 measurements can be biased positively. This occurs because chemiluminescence instruments equipped with heated molybdenum converters can partially convert other reactive nitrogen species (NO_z, such as peroxyacetyl nitrate (PAN) and HNO_3) and misreport them as NO_2. This introduces NO_y bias into the EEA measurements (Lamsal et al., 2008; Villena et al., 2012). To address this issue in future research, one potential approach is to use chemical model simulations, such as WRF-Chem, to estimate this interference and adjust the affected monitoring stations (Kuhn et al., 2024)*" in lines 203-208 to acknowledge this potential bias.

4) In many places, the manuscript assumes a fair amount of ML terminology/knowledge from the reader. Terms such as LSTM, inception modules, latent space, residual connections, max-pooling branch, dropout, and batch norm are often used with very brief or no references/explanations. I am afraid that non-ML readers might be overwhelmed by this.

We appreciate the reviewer's reminder and acknowledge that additional clarification of ML terminology and concepts is necessary for ACP readers from diverse backgrounds. In Section 2.3 (Model Architecture and Design), we have conducted the following update:

Added "*The residual connections pass intermediate feature maps from the encoder directly to matching decoder stages, which helps retain fine-scale information across the upscaling path and improves training stability*" in lines 241-243.

Added "*LSTM units use gated memory to retain information from earlier time steps (Hochreiter and Schmidhuber, 1997), so the ConvLSTM branches can learn day-scale meteorological evolution rather than treating each hour independently*" in lines 254-256.

Added "*The latent space represents a compressed internal representation where all encoder branches are mapped onto a common 3D tensor before the decoder reconstructs the 2 km × 2 km output fields*" in lines 266-268.

Added "*Each inception module runs multiple convolutional paths with different kernel sizes in parallel and concatenates their outputs, so $DACNO_2$ can capture both local gradients and broader regional structure within the same layer*" in lines 284-286.

Added "*The max-pooling branch performs spatial downsampling by taking local maxima, which provides a coarse-scale summary that complements the convolution branches and improves multi-scale feature extraction*" in lines 289-291.

Added "*Batch normalization normalizes intermediate activations within each mini-batch, reducing sensitivity to feature scaling and often improving optimization behavior (Ioffe and Szegedy, 2015). Dropout randomly masks a fraction of activations during training, which reduces overfitting and helps generalization when training on heterogeneous inputs (Srivastava et al., 2014)*" in lines 296-299.

5) Some parts of sect. 2.2–3.3 are very detailed in their technical elaborations, but still leave questions wrt. the broader methodology. For example, sect. 2.2.3: I am sure that the readers will understand what was done, but not necessarily why. Why is it necessary for the patches to overlap? Why 12 patches per day? How does all of this help to "balance the model's receptive field and computational efficiency", as the authors claim?

We agree with Reviewer #1's suggestion to provide more explanation for the practical implementation. In Section 2.2.3, which is about Patch-Based Data Processing and Reconstruction, we have expanded the first paragraph, and it now reads: "*To balance the model's receptive field and computational efficiency, we used a patching method because training and inference on the full domain on the 2 km grid as a single sample is not feasible, given the multi-branch 2D and 3D inputs and the 3D decoder. Specifically, all datasets except the temporal indicator were divided into patches of 512 km × 512 km with partial overlap. This produced grid sizes of 32 × 32 for ERA5 meteorological data, 64 × 64 for CAMS NO_2 data, and 256 × 256 for emission inventories and proxies, geographic data, and EEA NO_2 data. The patch size retains regional spatial context relevant for NO_2 variability while remaining compatible with the factor-of-two scaling scheme used in the encoder–decoder. Partial overlapping is used to reduce boundary effects because predictions near patch edges have reduced spatial context, and overlapping patches ensure each grid cell is predicted from at least one patch interior. In this study, each patch was treated as a single input sample, and the stride was set to generate 12 overlapping samples per day, covering the full domain while keeping the daily sample count computationally manageable. More samples can be generated as needed by reducing the stride of the sliding window. Additionally, if targeting higher resolution (e.g., 1 km × 1 km or 500 m × 500 m), larger patches are required, resulting in an exponential increase in computational cost.*" (Lines 211-223).

6) I do not think it is correct to say that the use of a priori profiles enhances the sensitivity of the satellite measurements (l. 55). Their purpose is rather to "fill" the gap of missing NO_2 that arises from a lack of measurement sensitivity. It should also be mentioned that the true NO_2 profiles extend beyond 5 km, and often show raised concentrations at 8000–12000 m, see e.g. Douros et al. (2023). On one hand, the effect of this "tail" on the resulting NO_2 VCD is expected to be rather small. On the other hand, the difference of DACNO$_2$-S5P to TROPOMI CAMS is only a few percent, hence the authors are arguably in a territory where the discussion of small effects is relevant.

We thank Reviewer #1 for suggesting a more rigorous expression of the benefits of high-resolution a-priori profiles. In the introduction, we have revised the original content of "… enhances the sensitivity of the satellite NO_2 product" to "*This progress has increased demand for high-resolution a-priori profiles, which can better account for near-surface NO_2 enhancements and strong spatial gradients over emission hotspots in satellite NO_2 retrieval products*" in lines 55-57.

In Section 4.3, we have revised the sentence "… which enhances the sensitivity of TROPOMI retrievals to NO_2 hotspots …" to "*The increase associated with CAMS-S5P is consistent with previous findings (Douros et al., 2023) and is primarily attributable to the improved spatial resolution of the a-priori profile, which better represents near-surface NO_2 enhancements and fine-scale spatial gradients, resulting in larger retrieved tropospheric columns over emission hotspots (Tack et al., 2021; Ialongo et al., 2020)*" in lines 646-649.

In Conclusions and Outlook, we have revised "… using DACNO$_2$-generated a-priori profiles enhances the sensitivity of the TROPOMI NO_2 products to …" as "*Application to satellite NO_2 retrievals demonstrates that using DACNO$_2$-generated a-priori profiles makes that the TROPOMI NO_2 products better account for near-surface concentrations and emission hotspots, particularly for small-scale emission sources and complex geographic regions*" in lines 735-737.

Although the difference of DACNO$_2$-S5P to TROPOMI CAMS is 3% on average, the change is accompanied by a clear spatial structure in the differences. We have added "*This change is accompanied by a clear spatial structure in the differences, with localized increases over small-scale emission hotspots and decreases over low-emission regions*" in lines 653-654.

We agree that the limitation of profile missing above 5000 m should be acknowledged and have added "*In addition, DACNO$_2$ provides a-priori NO_2 profiles up to 5000 m, while NO_2 signals at roughly 8 – 12 km show a slight enhancement, possibly linked to aviation and lightning (Douros et al., 2023; Kuhn et al., 2024; Dahlmann et al., 2011; Richter, 2009). This should also be considered in future DACNO$_2$ development*" in lines 669-672.

7) I would recommend writing out (horizontal) spatial resolutions more carefully. Consider writing "2 km × 2 km" instead of just "2 km", and make sure to avoid unneccesary conversions. For example: Instead of specifying the TM5 resolution as "about 100 km", give the actual resolution: 1° × 1°.

We revised the manuscript to report horizontal spatial resolutions more explicitly and consistently. We now write the DACNO$_2$ horizontal resolution as 2 km × 2 km and have added "*2 km × 2 km horizontal grid (hereafter referred to as the 2 km grid)*" in lines 57–58, and subsequently use "2 km × 2 km" and "the 2 km grid" throughout. We have also revised the TM5 resolution description to "*1° × 1°, approximately 100 km × 100 km*" in line 641, where the conversion is retained only to facilitate direct comparison with CAMS and

DACNO$_2$ (both reported in km). Similar edits were applied to other resolutions to ensure a uniform style across the manuscript.

Specific comments

• l. 19: To find "physically interpretable relationships" is a very high standard, and the manuscript does not come back to this anywhere; consider lowering the expectations a little bit.

We have revised the original sentence in the Abstract, and it now reads "*The model resolves spatial details and exhibits physically plausible behavior*" in line 19.

• l. 64: What is the difference between "NO$_2$ fields generated by CTMs" and "process-based CTM outputs"?

They are the same, and we have changed "process-based CTM outputs" to "*process-based 3D NO$_2$ fields*" in line 67 to avoid confusion.

• l. 178: Multiplication by an inverse = division.

The original sentence has been modified as "*In addition, CAMS NO$_2$ concentrations at each vertical layer were rescaled by dividing them by the ratio of the mean NO$_2$ concentration at that layer to the mean surface-layer NO$_2$ concentration, where this ratio was calculated from the training dataset.*" (lines 183-185)

• l. 311: The best model out of what ensemble? This suggests that the authors tested multiple variants of the model, but the differences between the variants are not explained.

We save the model checkpoint at the end of each epoch during training. The "best model" refers to the checkpoint with the best validation performance, rather than an ensemble or multiple model variants. We have revised "the model with the best validation…" to "*the model checkpoint with the best validation …*" in lines 347 and 364.

• Sect. 4.1: I would suggest to shorten the discussion on the feature importance, because Fig. 5 already speaks for itself.

We have refined Section 4.1 by shortening the figure description and emphasizing the key insights from the feature importance analysis. Section 4.1 (lines 532-568) now reads:

"*We assessed the relative importance of input feature groups in DACNO$_2$ using the integrated gradients (IG) method (Sundararajan et al., 2017) implemented via the Captum interpretability library (Kokhlikyan et al., 2020). IG quantifies the effect of varying each input feature from a zero baseline to its actual value on a selected target function. In this analysis, we computed IG at two targets: (1) the RMSE between DACNO$_2$ predictions and 2023 EEA NO$_2$ training measurements at the surface, and (2) the RMSE between DACNO$_2$ predictions and 2023 CAMS NO$_2$ at multiple vertical layers. Feature group results are shown in Fig. 5, and results for individual features are provided in Fig. S9.*

For surface NO_2 predictions evaluated against EEA measurements, DACNO$_2$ relies primarily on emission proxies, geographic features, and multi-level meteorological variables, while temporal indicators and single-level meteorological features play a lesser role. The addition of the EEA NO_2 constraint in Phase-2 and Phase-3 increases the importance of geographic data, highlighting its value for high-resolution surface NO_2 estimation. As shown in Fig. S9, land cover emerges as the most influential single feature (36.6%) in DACNO$_2$-Phase-3. Multi-level meteorological variables dominate the meteorological contribution, suggesting partial redundancy between single-level and multi-level meteorological inputs.

For NO_2 estimates by layer evaluated against CAMS, the distribution of input feature importance at lower layers (up to 1000 m) is similar to that for surface NO_2 evaluated against EEA, suggesting that DACNO$_2$ remains relatively stable across training phases with different constraints. Differences between the three-phase models are most apparent near the surface but gradually diminish with height. The importance of geographic features steadily decreases with height, whereas emission features reach their strongest influence at approximately 500 m before declining. Above 3000 m, both become negligible, reflecting the transition from the Planetary Boundary Layer (PBL), which is influenced by local surface features, into the free troposphere, which is dominated by broad-scale processes. In contrast, temporal indicators, single-level meteorological features, and especially multi-level meteorological features become increasingly important with height. This shift highlights the greater reliance on temporal and large-scale atmospheric information for NO_2 estimates at higher layers. Among these features, radiation flux is the most important single-level meteorological variable, and wind is the dominant variable among all meteorological features (Fig. S9). Given the consistently low overall contribution of single-level meteorological variables, future model development may consider reducing or refining the use of this feature group to streamline the input space.

Overall, the DACNO$_2$ model is developed by combining multi-scale inputs and multi-source constraints. The fine-scale spatial structure on the 2 km grid is primarily informed by high-resolution emission-related proxies and geographic features, whereas large-scale spatiotemporal variation and vertical structure are driven by meteorological variables and temporal indicators. Through the phased training strategy, the CAMS constraint transfers large-scale spatiotemporal variation to the DACNO$_2$ model, and the EEA constraint guides the model to use fine-scale static inputs to shape this variation on the 2 km grid spatially."

References

Dahlmann, K., Grewe, V., Ponater, M., and Matthes, S.: Quantifying the contributions of individual NO_x sources to the trend in ozone radiative forcing, Atmospheric Environment, 45, 2860-2868, 10.1016/j.atmosenv.2011.02.071, 2011.

Douros, J., Eskes, H., van Geffen, J., Boersma, K. F., Compernolle, S., Pinardi, G., Blechschmidt, A. M., Peuch, V. H., Colette, A., and Veefkind, P.: Comparing Sentinel-5P TROPOMI NO_2 column observations with the CAMS regional air quality ensemble, Geoscientific Model Development, 16, 509-534, 10.5194/gmd-16-509-2023, 2023.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, 9, 1735-1780, 10.1162/neco.1997.9.8.1735, 1997.

Ialongo, I., Virta, H., Eskes, H., Hovila, J., and Douros, J.: Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki, Atmospheric Measurement Techniques, 13, 205-218, 10.5194/amt-13-205-2020, 2020.

Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, Proceedings of the 32nd International Conference on Machine Learning, Proceedings of Machine Learning Research, 10.5555/3045118.3045167, 2015.

Kim, M., Brunner, D., and Kuhlmann, G.: Importance of satellite observations for high-resolution mapping of near-surface NO_2 by machine learning, Remote Sensing of Environment, 264, 112573, 10.1016/j.rse.2021.112573, 2021.

Kokhlikyan, N., Miglani, V., Martin, M., Wang, E., Alsallakh, B., Reynolds, J., Melnikov, A., Kliushkina, N., Araya, C., Yan, S., and Reblitz-Richardson, O.: Captum: A unified and generic model interpretability library for PyTorch, arXiv:2009.07896, 10.48550/arXiv.2009.07896, 2020.

Kuhn, L., Beirle, S., Osipov, S., Pozzer, A., and Wagner, T.: NitroNet – a machine learning model for the prediction of tropospheric NO_2 profiles from TROPOMI observations, Atmospheric Measurement Techniques, 17, 6485-6516, 10.5194/amt-17-6485-2024, 2024.

Lamsal, L. N., Martin, R. V., van Donkelaar, A., Steinbacher, M., Celarier, E. A., Bucsela, E., Dunlea, E. J., and Pinto, J. P.: Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument, Journal of Geophysical Research: Atmospheres, 113, 10.1029/2007JD009235, 2008.

Li, L. and Wu, J.: Spatiotemporal estimation of satellite-borne and ground-level NO_2 using full residual deep networks, Remote Sensing of Environment, 254, 112257, 10.1016/j.rse.2020.112257, 2021.

Richter, A.: Nitrogen oxides in the troposphere – What have we learned from satellite measurements?, EPJ Web of Conferences, 1, 149-156, 10.1140/epjconf/e2009-00916-9, 2009.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting, Journal of Machine Learning Research, 15, 1929–1958, 10.5555/2627435.2670313, 2014.

Sun, W., Tack, F., Clarisse, L., Schneider, R., Stavrakou, T., and Van Roozendael, M.: Inferring Surface NO_2 Over Western Europe: A Machine Learning Approach With Uncertainty Quantification, Journal of Geophysical Research: Atmospheres, 129, e2023JD040676, 10.1029/2023JD040676, 2024.

Sundararajan, M., Taly, A., and Yan, Q.: Axiomatic Attribution for Deep Networks, 10.48550/arXiv.1703.01365, 2017.

Tack, F., Merlaud, A., Iordache, M. D., Pinardi, G., Dimitropoulou, E., Eskes, H., Bomans, B., Veefkind, P., and Van Roozendael, M.: Assessment of the TROPOMI tropospheric NO_2 product based on airborne APEX observations, Atmospheric Measurement Techniques, 14, 615-646, 10.5194/amt-14-615-2021, 2021.

Villena, G., Bejan, I., Kurtenbach, R., Wiesen, P., and Kleffmann, J.: Interferences of commercial NO₂ instruments in the urban atmosphere and in a smog chamber, Atmospheric Measurement Techniques, 5, 149-159, 10.5194/amt-5-149-2012, 2012.