Response to Reviewer #2

We thank the reviewer for the careful evaluation of our manuscript and for the constructive comments and suggestions. Our point-by-point responses are provided below, with reviewer comments shown in blue and our replies in black.

Dear authors,

I would like to thank you for a very interesting read! I have a couple of questions that I'd like you to reflect on, but overall I am very happy with the quality of the manuscript and the described research.

Eq.1 - Looking back at Figure 2, the assumption that *b* can be assumed constant along the perimeter of the city seems a bit optimistic? There is a factor of six difference in the concentration along the border in the north and the south of the city in March 2019, which suggests that a westerly or easterly wind would cause a much higher flux across the border in the south than in the north.

We agree with the reviewer that the assumption of a constant background will not describe well the influx of NO_2 along the perimeter when nearby and upwind emission sources are unevenly distributed along the perimeter. However, Figure 2 might give a misleading impression of the distribution of background concentrations due to the spatial interpolation of the coarse CAMS grid and averaging over time. For example, Table 4 shows that the suburban background concentration in the north (at ES1945) is 15.4 $\mu g/m^3$, while in the south (at ES1193) it is 21.6 $\mu g/m^3$—a difference that is significant, but far smaller than the factor suggested by the reviewer. A validation of our method is now included as Section S1 in the Supplemental Material. Nevertheless, we agree that a location-dependent background field would provide a more realistic representation. This limitation is acknowledged in Section 4.1 and will be addressed in future versions of the algorithm.

And relatedly, you write (L179): "Other sectoral emission, e.g. from industry, will be accounted for indirectly in either an increased background field or in additional residential emissions." Such industrial sites are likely not equally distributed along the border, further increasing inhomogeneities in the background border flux. So my question is the following: Why not instead discretize the border along the *l* and *z*, and apply the same dispersion kernel that is used inside the city?

This is an interesting suggestion. However, application of the dispersion kernel will not be straightforward. The dispersion kernel describes the evolution of a plume originating from a point source at a defined injection height. In contrast, the background concentration entering the domain is assumed to be vertically mixed within the boundary layer, representing a vertically uniform concentration column rather than a plume from a discrete source.

To apply the kernel in this context, we would need to represent this vertically mixed background column as a distribution of effective sources across both height and location—a transformation that would require assumptions about source strength, vertical injection profiles, and transport history outside the domain. At present, this is beyond the scope of our implementation.

Nonetheless, we agree that more detailed spatial structuring of the boundary inflow is desirable, and we will explore this in future work, particularly in cases where background contributions are expected to be highly anisotropic or dominated by near-boundary sources.

L204 - Is it reasonable to assume that residential emissions are similar during weekdays and in the weekend?

We acknowledge that that residential emissions from activities like cooking and heating may vary between weekdays and weekends. However, quantifying this weekly cycle from literature is challenging due to limited data. For example, the CAMS-TEMPO emissions inventory (Guevara et al., 2021) does not include a weekly cycle for residential and commercial combustion. Moreover, the influence of weekly variability in residential emissions is likely small relative to other sources of model uncertainty, as residential emissions account for only 16% of total NO_X emissions in the Madrid municipal area, according to the CAMS global inventory (Soulie et al., 2024). Therefore, slight differences between weekdays and weekends in this sector are unlikely to significantly impact the overall model results.

The Retina algorithm could provide further insight into the relevance of prescribing weekly residential emission profiles, but we feel that this lies beyond the scope of the current study. We added to Section 4.1: “Finding a better a priori diurnal *and weekly* emission cycle is subject to further investigation.”

It is worth noting that the algorithm does capture *seasonal* variations in residential emissions (due to i.e. heating demand), which are more pronounced and impactful than weekly fluctuations.

L230-L234 - Can you explain a bit about how AERMOD treats dispersion through street canyons? If only one dispersion kernel is calculated for each combination of wind speed and direction, stability and boundary layer height, the model cannot deal with variations in the built environment or even in roughness length (which I imagine can vary a lot from the sparsely populated northern area to downtown Madrid), correct? Do you expect this to result in large errors?

We chose to use a uniform surface roughness length as a pragmatic compromise between computational efficiency and physical representativeness. The reviewer rightly points out it limits the model’s ability to capture the complex flow and dispersion processes within street canyons, such as recirculation zones or pollutant entrapment.

We realize that this simplification introduces uncertainty at the street scale, particularly under low wind conditions or in deep canyons, and we have identified it as a source of model error in Section 4.1. We are planning to introduce a parametrized version of the street canyon effect (for instance, following Vardoulakis et al., 2003) in the next algorithm version. Combined with a more realistic traffic model we expect to substantially reduce the model uncertainty.

L269-L270 - The dispersion model calculates concentrations of NOx but rather than assimilating NOx measurements, you assimilate NO2 measurements. You get these from the

XGBoost algorithm, which introduces non-linearity to the system. While you explicitly mention that you ignore the dependence on O3 (L85-L86 of the SI), there is also the dependence on e.g. the temperature and the SEA (L264-L264). How do you reconcile this with the fact that a Kalman Filter assumes a linear measurement operator H?

The non-linearity introduced by ozone chemistry is accounted for in the calculation of the NO_2/NO_X ratio, represented by r_i in Equation (S1). The summation describes the contribution to the local NO_X concentration by sector, and does not contain any ozone dependence. The ratio r_i is evaluated with the XGBoost model based on local values of the predictors. It is assumed to remain approximately constant for small changes in NOx due to small perturbations in emission factors x_j. Consequently, Equation (S4) represents a local linearization of the model state. This linear approximation may become inaccurate if updates in emission factors are too large. However, since emission factors typically change slowly over time relative to the daily update frequency, the Kalman filter is expected to iteratively converge to stable and consistent emission estimates.

Minor comments

L199: "See 0?"

L575: Mijling (2000) should be Mijling (2020)

Mult: Please also check the manuscript for many different occurrences of Sec. 0.

Thank you for pointing this out. The references to sections and papers have been corrected.

References

Vardoulakis, S., Fisher, B.E., Pericleous, K. and Gonzalez-Flesca, N., 2003. Modelling air quality in street canyons: a review. *Atmospheric environment*, *37*(2), pp.155-182.