



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-1147-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-1147

Anonymous Referee #1

Referee comment on "Data fusion uncertainty-enabled methods to map street-scale hourly NO₂ in Barcelona: a case study with CALIOPE-Urban v1.0" by Alvaro Criado et al., EGU Sphere, <https://doi.org/10.5194/egusphere-2022-1147-RC1>, 2022

Review of "Data fusion uncertainty-enabled methods to map street-scale hourly NO₂ in Barcelona city: a case study with CALIOPE-Urban v1.0" by Criado et al.

Criado et al. present a data-fusion workflow that uses Universal Kriging (UK) to merge dispersion model output (from CALIOPE-Urban) to hourly observations and microscale land-use regression (LUR) models. The authors' workflow is able to create high-resolution street-scale data of NO₂ to compute exceedance probabilities, with uncertainty calculations based on the UK technique estimated error variance. This work is comprehensive and appears to have good improvement in correlation and error metrics. I will be happy to recommend this manuscript for publication after my (mostly minor) comments below are addressed.

Major comments:

1. I want to note that the code isn't available and thus cannot be reviewed in its current form. The authors state that "So, at this moment only reviewers can access these relevant sources under a previous mail in the request form." But the request form on Zenodo requires the full name, e-mail address, and affiliation of the requester, compromising the anonymity of the referees. Thus, I was not able to review the code that is associated with this work. I would request that the authors provide the code used in this work for review, either through the Editor or the GMD portal. While I appreciate that the authors have archived the code in a repository with a DOI, access during review is important. Many other authors have publicly archived their code when submitting to GMD, despite the manuscript being under review.
2. The authors have performed data-fusion using data from two LCS campaigns, one from 2018, and one from 2017. Is there an impact on the quality of the data-fusion technique if LCS data are provided in different years? Similarly, if only one LCS campaign data set is used, how would it impact the quality of the results? A brief assessment of how much data is necessary and the applicability of the methods shown in this work will help future readers to apply this technique in the future to other major cities where urban pollution is also a major health issue.

Specific comments:

1. L57: "while the time-dependent LCS network explains the temporal behavior". Is it implied that the temporal behavior is short-term here in contrast with the long-term spatial distribution provided by the urban model?
2. L134: An adjustment factor is computed as the ratio between the observed 2017 annual mean and the average over the period of the experimental campaign. The LCS campaigns span only a few weeks (February 16th to March 15th, 2018; and February and March, 2017) – why is the 2017 annual mean used here instead of, for example, February-March mean?
3. L136-137: The authors say that this processing adds some noise to experimental results but corrects the "environmental conditions influence". What environmental conditions are referred to here? My impression is that this would mainly correct for bias in the low cost sensors' instruments.
4. Figure 3: Useful to label inset in the figure "Combined data" "CSIC" "xAire".
5. L247-248: The authors indicate that with the criteria (covariate slopes must be positive, less than four observations available in the hour) 14% and 2% of the hours in UK-DM, UK-DM-LUR are not corrected. How much percentage of these are due to negative covariate slopes? And how much are due to too few observations? If there is a significant percentage of nonphysical negative covariates, is there a common pattern to the conditions causing these?
6. L273: "We attribute this behavior ... also to the already poor predictive skills of CALIOPE-Urban in this concentration range." A citation will be useful for CALIOPE-Urban's underperformance in high-NOx conditions.

Technical corrections:

1. L204: "back-transformed" -> "back-transformation"
2. L210: "exceedance (P) a certain..." -> "exceedance (P) of a certain..."
3. Lines 233, 234: middle dot -> cross sign for scientific notation.
4. Figure 7: Units for MB, RMSE are missing the "³" (shows as micrograms/m)