Point by point responses to the Anonymous Referee #1

We are very thankful to the reviewer for providing a detailed revision of our manuscript. The comments of the reviewer are indicated point-by-point in the following text. We explain how we have carefully addressed each of them (our answers in blue text). Modifications and new sections are highlighted with track changes in the manuscript and the Supporting Information.

General comments

1. Reviewer #1. The analysis only considers NO_2 concentrations. NO_2 is a pollutant of active current interest for regulation and health effects, however it is challenging to model due to the interaction of dispersion and chemistry. Hence it is valuable to analyse modelled and measured NO_X concentrations before considering NO_2, to assist with distinguishing between uncertainties in emissions, dispersion and chemistry. The consideration of chemistry effects should also be included in the associated discussions of NO_2 results.

Authors:

With regard to the chemistry used in the street-scale model, Valencia et al. (2018) evaluated multiple NO-NO_2-O_3 chemical mechanisms in R-LINE. They compared the GRS mechanism, used by ADMS-Urban (Malkin et al., 2016), and two other algorithms for NO_X chemistry with near-road data in Michigan. Their results indicate that the GRS mechanism was the most consistent in predicting NO_2 for near-roadway environments. We believe, based on the Valencia et. al. (2018) evaluation of GRS in R-LINE and the use of GRS in the ADMS-Urban model, that this is an appropriate way to model NO_2 chemistry without the need to also evaluate NO_X.

However, following the advice of the reviewer we show below the comparison of NO_X average daily cycle concentrations for each model (equivalent figure as Figure 8 of the initial submitted manuscript). We can see that the errors for NO_2 and NO_X are very similar for all the stations, indicating that the chemistry of GRS is performing well in our simulation experiments. Thus, we believe that it is reasonable to evaluate modelled NO_2 directly as suggested by Valencia et. al. (2018) and Malkin et al. (2016). Given that the paper is focused on NO_2 due to its interest for regulation and health effects we decided not to include the NO_X discussion in the manuscript and keep the focus on the NO_2 exclusively.

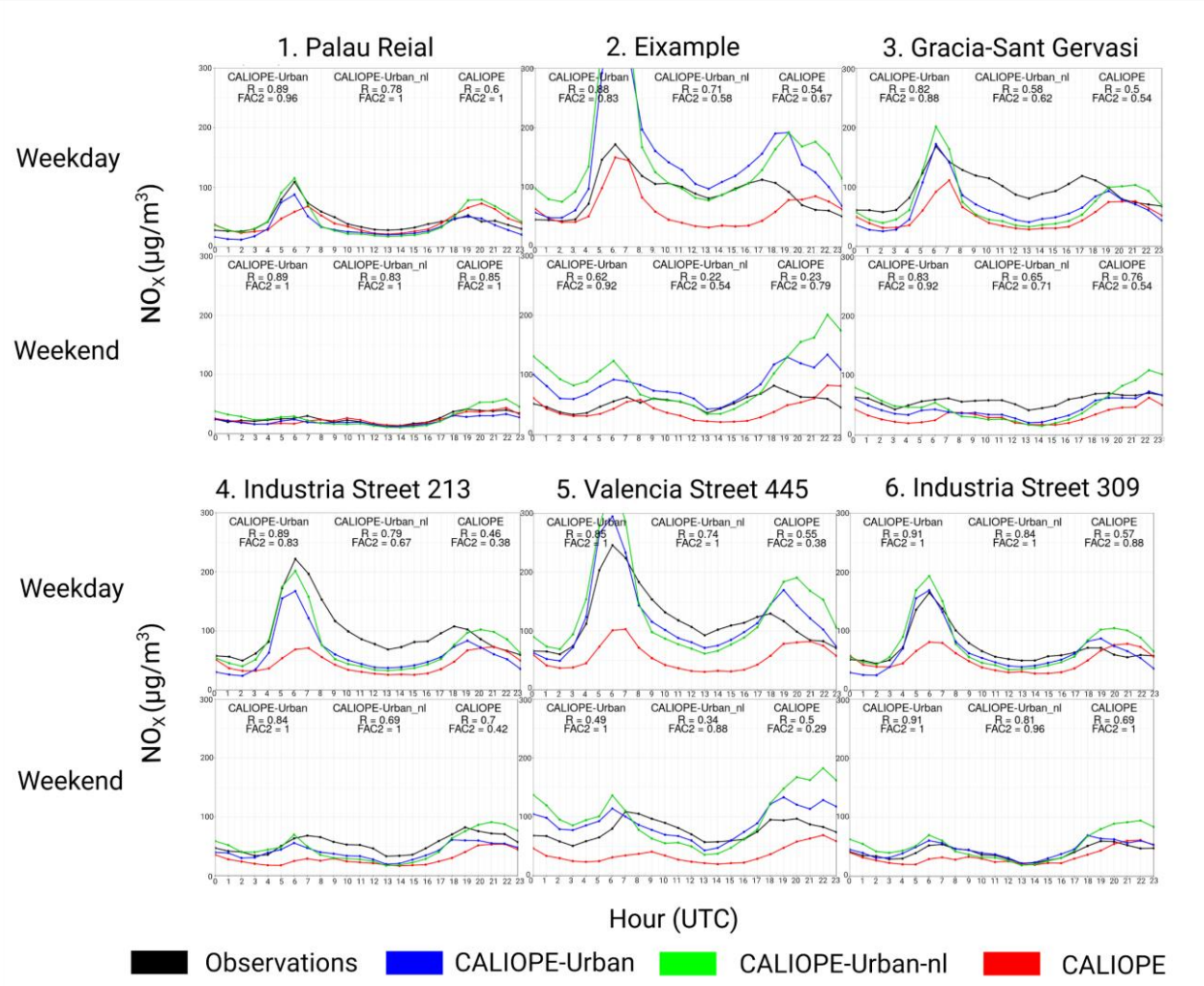

Figure 1. NO_X average daily cycle at all sites described in Sect. 3.1 during April and May 2013 for weekday and weekend. Observations are represented in black coloured lines, red lines are CALIOPE, blue lines are CALIOPE-Urban and green lines represent CALIOPE-Urban without local developments (CALIOPE-Urban-nl).

2. Reviewer #1. The method for taking into account the effects of a specific street canyon on dispersion described in Section 2.3.1 only considers flow channelling along the canyon. However, canyons are also known to cause recirculating flow across the canyon, which significantly alters the dispersion of road traffic emissions and hence the concentration variation with wind direction for receptors within the canyon. No analysis of the modelled or measured variation of concentrations with wind direction is presented, so it is difficult to assess the effectiveness of this formulation.

Authors:

We agree with the reviewer on the importance of modelling the recirculating flow across the canyon and its impact on pollutant dispersion. It is widely known that canyons with sufficient aspect ratio (i.e. more than 0.15) may cause recirculating flow across the canyon (Oke, 1988; Yamartino and Wiegand., 1986; Dobre et al., 2005) when over roof winds are close to perpendicular to street direction. This recirculation may affect significantly the dispersion of road traffic emissions. However, in this work we assume that recirculation process is negligible for multiple reasons: first, dispersion models, such as R-LINE, are not designed to model extremely detailed urban flows (e.g. CFD models), but rather are based on representative flows that are influenced by average urban attributes near the source location; second, dispersion models are designed to give accurate concentrations averaged over a long a time period (usually one hour) where variability in wind speeds occur and thus recirculation may not be longer consistent; third, there is a recirculation and a vehicle induced turbulence occurring within the street canyon, both are contributing to a well mixed more homogeneous air mass within the street canyon, especially over the long averaging time and variable wind conditions; and lastly evaluation of the potential impact (positive or negative) of including recirculating flows across the canyon is not possible without multiple simultaneous meteorological and pollutant measurements within a street canyon at a fine temporal scale, which was not part of the experimental design and thus are not available.

We have added an explicit note about this limitation in the revised manuscript in in Section 2.3.1, page 8 lines 18-22 as “In this work, we assume that recirculation flows within street canyons are negligible because R-LINE computes concentrations averaged over an hour, when recirculation and vehicle induced turbulence are assumed to contribute to a well mixed more homogeneous air mass driven by variable wind conditions. Additionally, evaluation of the potential impact of including recirculating flows across the canyon is not possible without multiple simultaneous meteorological and pollutant measurements at a fine temporal scale, which are not available.”.

Regarding flow channelling along canyon, we didn’t assess the effectiveness of this formulation because we don’t have access to a complete dataset of measured wind conditions within a diverse range of streets in the city. However, the positive results of the work indicate that our formulation is appropriate for the objective of the study.

Specific comments

3. Reviewer #1. Section 2.1: Please state explicitly the depth of the lowest model layer in WRF and CMAQ, which is alluded to in Section 2.3.1.

Authors:

We have added a comment in Section 2.3.1, page 6 lines 19-21 as “Most buildings in Barcelona have lower heights than the WRF bottom layer (40.6 m depth). WRF results are assumed to represent over roof wind and stability conditions because its mid-point height (20.3 m) is similar to average building height (bh) in a typical neighbourhood of Barcelona (e.g. Eixample district; 20.7 m).”

4. Reviewer #1. Section 2.3.2 and Figure 3: Please comment on the negative value of intercept, which may indicate that the Ciutadella site does not fully represent an appropriate urban background concentration for the Eixample traffic site.

Authors:

We believe that Ciutadella is a reasonable background site due to its upwind location in the predominant wind direction during the day and its location within the main park of the city (see Figure 2 in the manuscript). In addition, during the two-week period of the passive dosimeters campaign its mean NO_2 concentrations were 40.2 $\mu g/m^3$, which is very close to the observed mean concentrations of 42.1 $\mu g/m^3$ using the passive dosimeters, suggesting that it is a reasonable background site (Amato, personal communication, April 24, 2019). The value of the intercept is very close to zero (i.e. the remaining background influence), which means that the regional and urban background contribution have been taken out reasonably well.

5. Reviewer #1. Section 2.3.3: Is the background mixing correction applied uniformly to all pollutants? In particular, O_3 usually shows opposite behaviour to NOx and NO_2, so this formulation may distort background concentrations used for chemistry calculations. Please also clarify how the background concentration is used within R-LINE, especially in regard to the chemistry calculations.

Authors:

Yes, the background vertical mixing is applied uniformly as done in all split-operator models. The vertical distribution of pollutants are solved first with a vertical diffusion following similarity theory, applied uniformly to all pollutants, and then we solve the chemistry. With this approach, O_3 shows opposite profile to NO_X and NO_2 as noted by the reviewer. This is consistent with the chemical reaction of emitted NO with ambient O_3 to form NO_2. We have included a clarifying note in the revised manuscript in Section 2.3.3, page 11 lines 10-12 as “To calculate street-level NO_2 concentrations, the vertical distribution of pollutants are solved first using the background decay method, applied uniformly to all pollutants, and then the GRS chemical mechanism is solved.”

6. Reviewer #1. Section 2.4: Although the analytical model shows a significant reduction in execution time relative to the numerical local model, is 44 minutes execution time for 1 modelled hour realistic for use in an operational forecasting system?

Authors:

The current model design and methodology was explored as a potential way to forecast pollutant levels in the future, therefore we are using this initial evaluation to determine if this is possible and with what level of accuracy. It is important to explore both the analytical and numerical calculations in R-LINE to determine strengths and weaknesses of both. Here we present that the analytical solution is much faster, however the numerical solution is more accurate in some instances, so a final forecasting model would need to balance speed and accuracy. Once, we determine the validity and accuracy of our method we will begin the process of code optimization. For instance, the R-LINE code is not currently parallelized. Parallelization could be done at the road-segment level, which will speed-up the code by several orders of magnitude, making this an extremely cheap and valuable tool for a forecasting system.

7. Reviewer #1. Section 3.1: Please state the measurement height(s) for the official network sites.

Authors:

We added a note in Section 3.1 page 15 Table 3 to clarify this: “The measurement height of the official network sites and the mobile sites is 3 meters”.

8. Reviewer #1. Section 4: Please add an initial assessment of NO_X modelled and measured concentrations.

Authors:

As explained in response number 1, R-LINE has been evaluated for roadways in urban areas with inert pollutants (such as NO_X) and reactive pollutants (such as NO_2), therefore there is no value added to present a full evaluation of NO_X in this instance. In addition, chemical transport models such as CMAQ have been evaluated for use in urban areas and have been previously used to provide background concentrations (Beevers et al., 2012; Isakov et al., 2014). We are using these previous models as is, by coupling R-LINE and CMAQ, and making adjustments based on the data to evaluate the additions of the street canyon and background adjustments. The street canyon adjustments are evaluated using a variety of street canyons throughout Barcelona. The background adjustments are evaluated at background sites throughout the city.

Considering that the scope of the present manuscript is the modelling of NO_2 concentrations at street level, we prefer not to add the discussion of NO_X.

9. Reviewer #1. Section 4.1: The discussion relating to Figure 8 does not mention the varying influence of chemistry processes through the day, which can lead to inaccuracies in diurnal profiles.

Authors:

R-LINE first calculates the dispersion of pollutants from the road source, and then it resolves the parameterised equations of the chemical reactions for the pollutant transportation time interval. The GRS chemistry mechanism solves the photochemistry of NO_2 assuming clear-sky conditions. Hence, it does not consider cloud effects on photolysis, representing one of its major limitations.

Valencia et. al. (2018) also show that the GRS method in R-LINE has less than a 15% bias of the results even though they do not account for cloud cover. From our experience, the processes that may have a greater influence on the results in this modelling system are the correct wind and stability conditions and the accuracy of emission sources within the street canyons. These may have more influence on the concentrations than the photolysis in the chemistry scheme. For example, in Figure 2 we show the weekday average daily cycle for Valencia Street 445 to compare the effect of setting the GRS mechanism photolysis rates to zero (caliope_urban_no_photo) with the effect of setting atmospheric conditions to stable (caliope_urban_stable). The stable conditions are set using the following parameter values from Snyder et al. (2013): Lmon (Monin-Obukhov length) equals 11.1; ustar (friction velocity) equals 0.12; Wsrefh (wind speed at roof-top level) equals 2.0. The impact of neglecting completely the photochemical reaction of the GRS chemical mechanism results in an overestimation of NO2 concentrations during daytime. Although we see a negative impact of not using the photochemical reaction in the GRS chemical mechanism (purple line), it is clear that setting stable atmospheric conditions dramatically changes the modelled concentration levels, producing a greater overestimation (green line). These results confirm our initial hypothesis that atmospheric stability has a much greater influence on the NO_2 concentration than neglecting clouds in the calculation of the NO_2 photolysis rate applied in the GRS chemical mechanism.

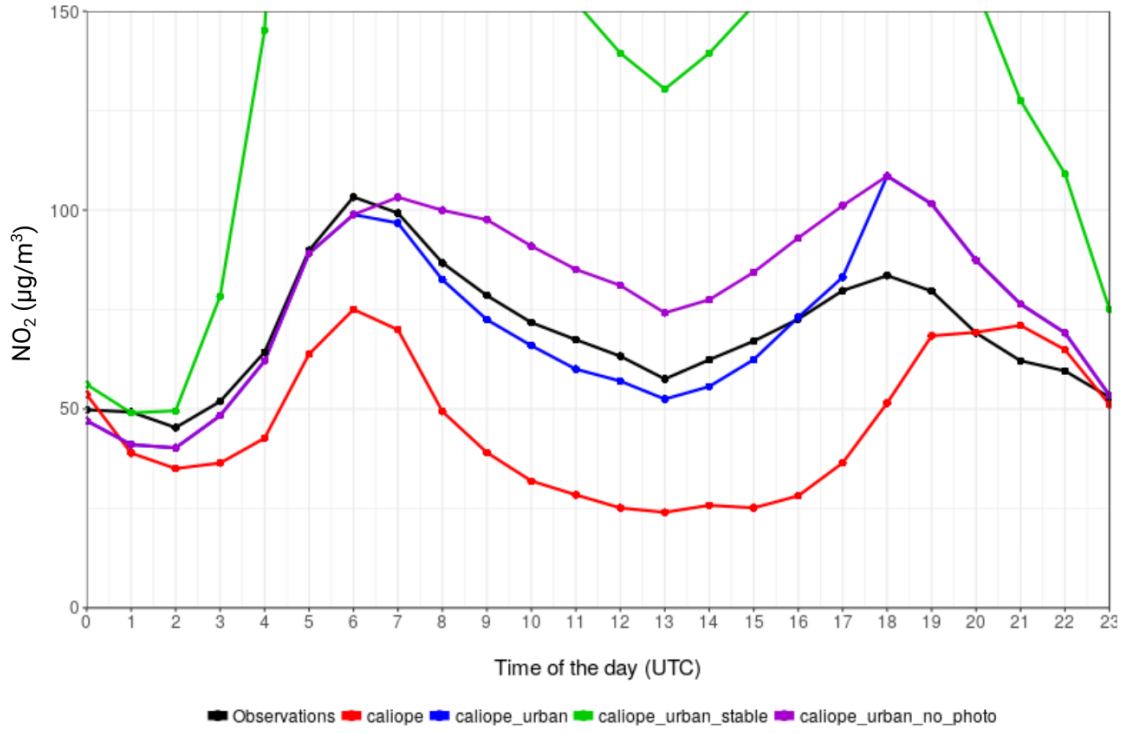

Figure 2. NO_2 average daily cycle at Valencia Street 445 site described in Sect. 3.1 during April and May 2013 for weekday. Observations are represented in black coloured lines, red lines are CALIOPE, blue lines are CALIOPE-Urban, green lines represent CALIOPE-Urban with stable atmospheric conditions (caliope_urban_stable) and purple lines depict CALIOPE-Urban with photolysis rates set to zero (caliope_urban_no_photo).

In the revised manuscript, we have added a note about neglecting the effect of clouds in the R-LINE photolysis rate (Section 2.2, page 6 lines 3-4): "GRS chemistry mechanism solves the photochemistry of NO_2 assuming clear-sky conditions. Thus, it does not consider cloud effects on the NO_2 photolysis rate, representing this one of its major limitations."

10. Reviewer #1. Section 4.3: It is common for Gaussian-type models such as R-LINE to perform poorly in low wind speed conditions due to uncertainties about associated wind directions. They also do not take into account possible accumulation of pollutants between hours in low wind speed conditions, which is in contrast to the assumption in the background adjustment in this work of reduced mixing causing reduced surface background concentrations. Figure 10f) suggests that the unadjusted regional background could be more appropriate than the adjusted in the early morning hours, though not in the evening. Are there other differences (eg. wind direction) between these two periods?

Authors:

Regarding the accumulation of pollutants, it's true that we can't consider it from one hour to the next within the street canyon. As the reviewer correctly identifies, this is a limitation in dispersion models and in our implementation.

Regarding the reviewer comment on Figure 10 panel f) that the unadjusted regional background could be more appropriate than the adjusted in the early morning hours, we are aware that the result from the upwind background scheme (assumed as roof-level background concentration provider) gives a more precise result from 0 to 7 UTC under calm winds in Figure 10 panel f). In contrast, at night from 18 to 23 UTC the opposite happens, background concentrations at surface level are more accurate than at roof level in comparison with observations. We see this result as positive because it suggests that our method can reduce the overestimation of night surface level concentrations. The contradictory result at surface level between morning and night hours for stable hours with calm wind conditions is dependent on the results of the mesoscale system and we would need more observational data to further investigate this issue.

Concerning the uncertainties associated to wind direction under low wind conditions, we show in the following figure the difference in wind direction at the Barcelona airport for the two periods discussed in Section 4.3:

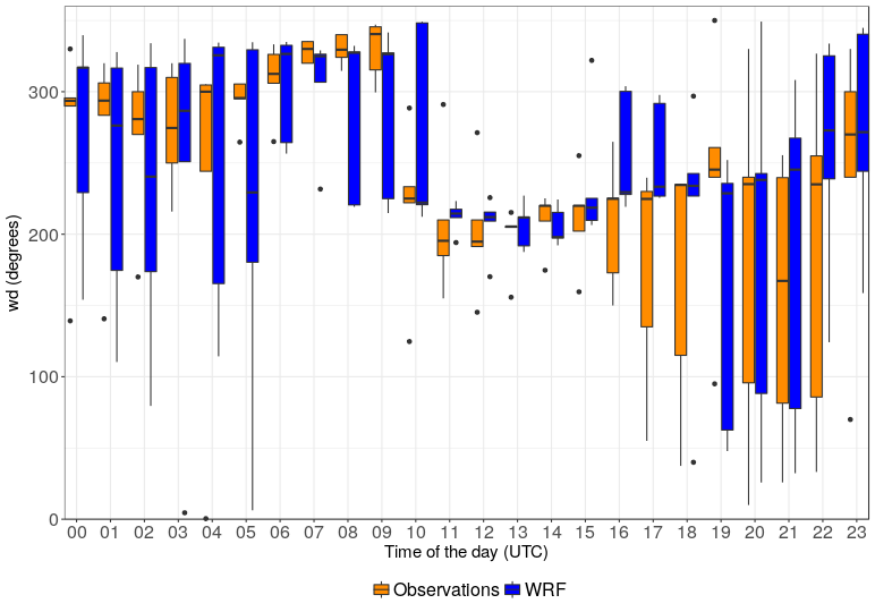

Figure 3. Boxplots by time of the day of good performance days for WRF (blue) and observed (orange) wind directions at Barcelona airport (10m height) with dots representing outliers.

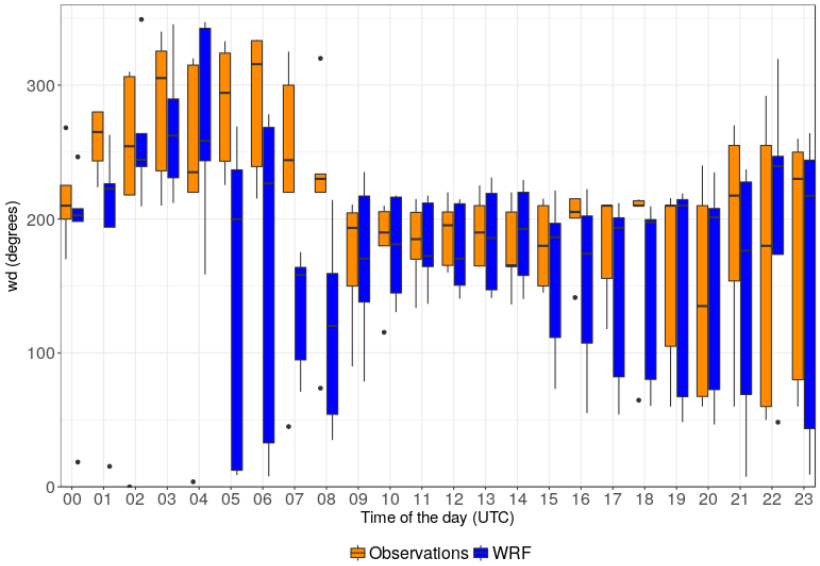

Figure 4. Boxplots by time of the day of bad performance days for WRF (blue) and observed (orange) wind directions at Barcelona airport (10m height) with dots representing outliers.

From the two figures, we see a similar pattern in both periods in general, being the exception the morning transition from stable to unstable atmospheric conditions (5-8 UTC) in bad performance days. The wind direction difference between WRF and observations is approximately 90 degrees on average. In contrast, looking at Figure 10 panel d) (surface NO_2 concentrations) and f) (background NO_2 concentrations), we do not see a clear impact of the mentioned error of wind direction from 5 to 8 UTC. In our system, wind direction may not play a crucial role when wind speed is very low (i.e. for bad performance days approx. 1.5 m/s) for two reasons: (1) the upwind background scheme will take CMAQ concentrations from nearby grid cells (i.e. tending to give similar concentration levels in different wind directions), typically within city boundaries, and (2) R-LINE meandering partial contribution, which disperses radially in all directions, is greater under low wind speeds reducing the potential impact of the wind direction error.

11. Reviewer #1. Figure 10: for panels e) and f) why is the model background concentration an average over six sites, not also taken from the Ciutadella site?

Authors:

As stated in response to reviewer comment 4, the Ciutadella Park monitoring station is the only reference site available for urban background in Barcelona. Our aim in Figure 10 for panels e) and f) is to compare the modelled background concentrations (i.e.

excluding local vehicular traffic contribution) with the most suitable observed urban background, which in our case is given by Ciutadella site. The model background concentration is taken to be an average over six sites because the interest is to represent a summary of the variability of the modelled background at the six sites in comparison with the observed background. We added a clarification in the revised manuscript in Section 4.3 page 25 lines 3-4: "We aimed to compare the modelled background concentrations (i.e. excluding local vehicular traffic contribution) with the most representative urban background observation, which in our case is the Ciutadella site."

12. Reviewer #1. Section 5: Again, the uncertainty in NO_2 resulting from chemistry processes should form part of the discussion.

Authors:

We added a comment about the uncertainty of chemistry processes used in our solution in Section 5, page 28 lines 6-7: "Finally, we consider an additional source of uncertainty the assumption of clear-sky conditions in the photolysis rate calculation of the GRS chemistry mechanism."

Technical corrections

13. Reviewer #1. Abstract, p1 line 15: In this case, the coupled system **also** shows

Authors: Amended

Section 1, p3 line 8: subtract its result **from** the mesoscale model

Authors: Amended

Section 1, p3 line 30: please re-phrase 'over background roof-level concentrations' as the meaning is unclear

Authors: Amended

Section 1, p3 line 35 – p4 line 1: 5 traffic site**s** and 1 background

Authors: Amended

Section 1, p4 line 2: campaign that **deployed** 182 NO2 passive dosimeters **across Barcelona** for two weeks..

Authors: Amended

Section 2.3.3, p12 line 8: please re-phrase 'ends when the surface background gets over roof value for bd equals 0' as the meaning is unclear

Authors: Amended

Section 3.1, p15 line 8: centred **on** the measurement site

Authors: Amended

Section 4.1, p18 lines 9 and 11: unnecessary **the** before Appendix B

Authors: Amended

Figure 7: these plots look vertically distorted, as the target area is usually viewed as circular.

We agree with the reviewer that the image looks vertically distorted but the current version of the Delta Tool for Windows is producing this kind of plot and as far as we know we can't do anything to change it. We downloaded the most updated version and it produced the same kind of plot. In the informational website of Delta Tools it is shown as vertically distorted, too: https://ec.europa.eu/jrc/en/scientific-tool/fairmode-delta-tool

Figure 8: The vertical and horizontal axis scale labels are too small to read.

Authors: Amended

Section 4.1, p20 line 3: higher modelled traffic emissions, **resulting** in higher local pollutant concentrations...

Authors: Amended

Section 4.3, p23 lines 15-18: The first sentence says ten days of highest RMSE and ten days of lowest RMSE, whereas the following sentences suggest five days of high RMSE and five days of low RMSE. Please clarify how many days were selected and analysed.

Authors: Amended

Section 5, p27 line 11: ... gives surface concentrations **by** applying a vertical…

Authors: Amended

References

Beevers, S., Kitwiroon, N., Williams, M.L., and Carslaw, D.C. "One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions". Atmospheric Environment 59, pp. 47–58. DOI: https://doi.org/10.1016/j.atmosenv.2012.05.034. 2012.

Dobre, A., Arnold, S. J.Smalley, R. J., Boddy, J. W D, Barlow, J. F. "Flow field measurements in the proximity of an urban intersection in London, UK". Atmospheric Environment 39, pp. 4647-4657. DOI: https://doi.org/10.1016/j.atmosenv.2005.04.015. 2005.

Isakov, V. et al. "Air Quality Modeling in Support of the Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS)". International Journal of Environmental Research and Public Health 11.9, pp.

8777–8793. DOI: https://doi.org/10.3390/ijerph110908777. 2014.

Malkin, T.L., Heard, D.E., Hood, C., Stocker, J., Carruthers, D. “Assessing chemistry schemes and constraints in air quality models used to predict ozone in London against the detailed Master Chemical Mechanism”. Faraday Discussions 189, pp. 589-616. DOI: https://doi.org/10.1039/C5FD00218D. 2016.

Oke, T.R. “Street design and urban canopy layer climate”. Energy and Buildings 11.1-3, pp. 103–113. DOI: https://10.1016/0378-7788(88)90026-6. 1988.

Valencia, A., Venkatram, A., Heist, D., Carruthers, D., and Arunachalam, S. “Development and evaluation of the R-LINE model algorithms to account for chemical transformation in the near-road environment”. Transportation Research Part D: Transport and Environment 59.2, pp. 464–477. DOI: https://doi.org/10.1016/j.trd.2018.01.028. 2018.

Yamartino, R.J., Wiegand, G. “Development and evaluation of simple models for flow, turbulence and pollutant concentration fields within an urban street canyon”. Atmospheric Environment 20, pp. 2137-2156. 1986.

Point by point responses to the Anonymous Referee #2

We are very thankful to the reviewer for providing a detailed revision of our manuscript. The comments of the reviewer are indicated point-by-point in the following text. We explain how we have carefully addressed each of them (our answers in blue text). Modifications and new sections are highlighted with track changes in the manuscript and the Supporting Information.

General comments

1. Reviewer #2. First, it is claimed that atmospheric stability influences the relation between rooftop concentrations and the concentrations at street level (e.g. p. 11, lines 8-11). While this well may be the case, no independent evaluation of the influence of atmospheric stability on the vertical concentration profile in street canyons is given in the paper. The authors should either refer to previous studies in Barcelona or show an evaluation based on own measurement data.

Authors:

Several dispersion models integrate in their formulation the concept of street and over-roof concentrations exchange dependent on atmospheric stability (Hotchkiss and Harlow., 1973; Berkowicz et al., 2000; Soulhac et al., 2011; Kim et al., 2018). In addition, this influence has been demonstrated using wind tunnel experiments (Salizzoni et al., 2009). The influence of atmospheric stability on street wind speed and over roof winds has been shown using experimental measurements, too (Rotach,. 1995). For these reasons, we consider scientifically robust to assume that atmospheric stability influences the relation between rooftop concentrations and the concentrations at street level.

The methodology in this manuscript is a first attempt at coupling the urban-scale model, R-LINE, with a mesoscale model, CMAQ, to obtain concentrations throughout the city. We do present multiple street level measurements of varying degrees of urban structural influence. The results indicate that using one consistent background methodology in all these instances provides validity in our approach. Our methodology is still under refinement and will need further evaluation based on additional datasets. We are currently working on the analysis of measured vertical profiles of Black Carbon (BC) within a few street canyons of the Barcelona city. Unfortunately, we don't have access to high frequency vertical profiles of NO_2 concentrations and wind conditions within street canyons in Barcelona. From BC vertical profiles results, it seems that our hypothesis is well-oriented as the reviewer can see in the figure below, where we show the modelled contributions compared with the hourly averaged observed BC at different heights in a very narrow Street in Barcelona. We show 12 UTC (13 hour local time), an

hour that it is expected to have a low contribution from local traffic (i.e. more signal from background). In addition, we expect a well mixed vertical BC column due to the convective conditions typically occurring at this period of the day. We see in the figure that the overall dynamic of the modelled vertical profile is in agreement with the observed profile.

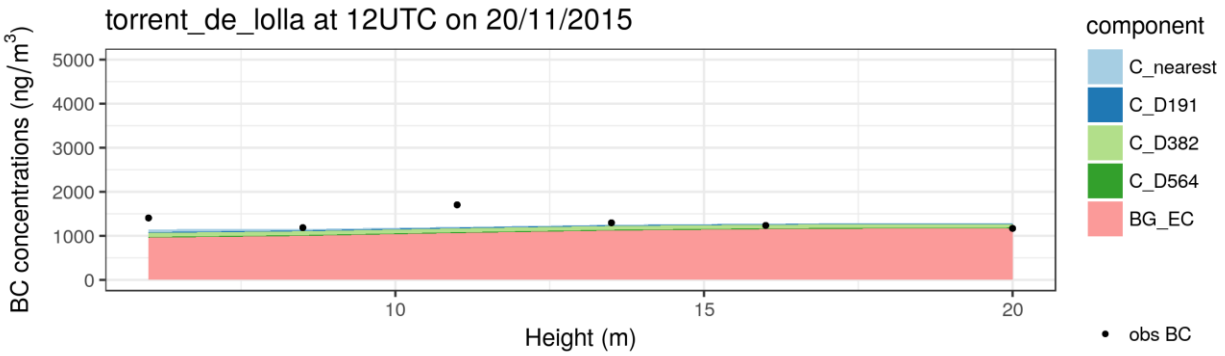

Figure 1. Local traffic and background contributions to BC vertical profile at Torrent de l'Olla Street at 12 UTC on 20th November 2015. Observations are depicted as dots and coloured levels represent each local traffic contribution: nearest roads is light blue, roads within 191 m (excluding the nearest roads) is dark blue, roads in 392 m (excl. roads within 191 m) is light green, dark green is for roads in 1 km^2 (excl. roads within 392 m) and pink is for background.

Lastly, it is also important to note that there may not be datasets available that specifically address all possible aspects of our approach, and thus we must rely on the datasets available and on the results of our methodology within the modeling system.

We added a note in Section 2.3.3 page 11 lines 14-16 to support the incorporation of atmospheric stability influence on vertical mixing: "In the research literature, the influence of atmospheric stability on vertical mixing within a street canyon has been demonstrated using experimental measurements (Rotach, 1995), wind tunnel experiments (Salizzoni et al., 2009), and it has been implemented in some dispersion models (e.g. Soulhac et al., 2011; Kim et al., 2018)".

2. Reviewer #2. Second, the model system shows poor skill when the observed wind speed is low. I would expect that the traffic-induced turbulence dominates the turbulence in street canyons at low wind speeds. However, it seems that turbulence generated by the moving traffic is not included in the parametrization. Calm winds potentially lead to highest concentrations and can cause severe pollution episodes. Hence, it would be crucial for a street canyon model to cope with low wind situations.

Authors:

In our approach, we have not considered the traffic-induced turbulence directly, however we do have an initial vertical dispersion of roadway emissions which somewhat models traffic-induced turbulence. In previous roadway studies (Snyder et al., 2013; Heist et al., 2013) this same approach has been used in the median of the roadway for Caltrans Highway 99 (Benson, 1989) and results are accurate when compared to near-road measurements, therefore not explicitly modeling vehicle-induced turbulence is not believed to have a large impact on the results.

Specific Comments

3. Reviewer #2. 1.) P. 2 line 9-21: In this part of the Introduction, several systems coupling regional and urban scale models are described. It would be better to divide this presentation into (1) systems that apply nesting of an urban scale model within a regional scale model and (2) regional scale models that apply downscaling (using a dispersion kernel). The given examples from literature are not exhaustive. Also mention a second method for downscaling, by embedding Gaussian dispersion models within the grid.

Authors:

In the revised manuscript we present a more complete list of systems that couple off-line regional and urban scale models by downscaling the regional model using a dispersion kernel. We consider adequate to uniquely present downscaling methods because our system belongs to this category, which as far as we know is the most extended methodology to couple regional and urban scales. The revised paragraph in the manuscript is as follows (Section 1 Page 2 Line 9 to 27),

"In order to overcome these limitations, coupling off-line the regional and urban scales by downscaling the regional model using a dispersion kernel has been successfully applied in some cities (Beevers et al., 2012; Moussafir et al., 2014; Isakov et al., 2014; Jensen et al., 2017; Maiheu et al., 2017; Kim et al., 2018; Hood et al., 2018, Fagerli et al., 2019). For instance, Hood et al. (2018) coupled a regional climate-chemistry model with 5 km horizontal resolution (EMEP4UK) with the fine-scale model ADMS-URBAN to simulate air quality over London in 2012. They compared the coupled system results with the regional and the fine-scale models run separately. Authors found that both the fine-scale model and the coupled system performed better than the regional for NO_2 at both annual mean and hourly concentration levels due to the explicit treatment of traffic emissions within the city. In addition, Jensen et al. (2017) estimated annual NO_2 concentrations at 2.4 million addresses in Denmark using the street canyon model OSPM coupled with DEHM for regional background concentrations and UBM for urban background obtaining a good correlation in Copenhaguen (r^2 = 0.70) against 98 measurement sites for NO_2 in the year 2012. Maiheu et al. (2017) covered a broader

spatial context, estimating EU-wide NO_2 annual average levels at 100 meter resolution with a regional model coupled with a dispersion kernel-based method. The approach does not produce hourly concentration levels and approximates road-link level traffic emissions by distributing the regional grid cell traffic emissions to each road-link based on road capacity. Hence, it provides more spatial detail than previous EU scale NO_2 assessment studies, but more specific methods are required to resolve air quality in cities. In this sense, there is a lack of air quality urban forecasting methodologies that can be applied to a diverse range of cities and that consistently resolve at least some of the major challenges already identified by the community, i.e., 1) downscaling regional meteorology to street level as required to drive pollutant dispersion; 2) obtaining background concentrations from the mesoscale system avoiding the double counting of traffic emissions. Additionally, we consider vertical mixing with background air a key process to be resolved when coupling the regional and urban scales."

4. Reviewer #2. 2.) P.4 line 1: How representative is this period (April and May 2013) for the season? Why was such a short period chosen?

Authors:

NO_2 exceedances in BCN are chronic along the year. April and May are months with a reduced amount of holidays and vehicular traffic behaviour is representative of the pulse of the city. We are aware that a longer period would be of interest to evaluate the skills of the model across different seasons. This will be presented in a future work. In this work, we focus on the experimental campaign with multiple simultaneous measurements along trafficked street canyons. We find this dataset relevant because we can evaluate CALIOPE-Urban close to road sources.

5. Reviewer #2. 3.) P.5 line 26-27: Why were the 38 vertical layers from WRF collapsed to 15 layers for the CMAQ computation? With only six layers in the PBL, this leads to a rather crude treatment of the near-ground chemistry and boundary layer mixing processes.

Authors:

In Europe, there are a wide range of mesoscale air quality models that work with low vertical resolution for computational reasons (e.g., LOTOS-EUROS, CHIMERE), and the skills of those models have been shown to be in the same order as other systems with higher vertical resolution. We use the default CALIOPE forecast configuration, which aims to reduce computational time to allow for rapid forecasting. CALIOPE skills are within the state-of-the-art forecasting systems (e.g. Pay et al., 2014). Since we are

most concerned with NO_2 here, which has a rapid chemical transition from emitted NO to ambient NO_2, the most important chemistry is the near-road chemistry that is simulated in the fine-scale dispersion model.

6. Reviewer #2.4.) P.4 line 1: Please provide a list of the chemical reactions in the GRS here.

Authors:

We have included a list of the chemical reactions in the GRS in Section 2.2 page 6 Table 1 and a note referencing the table in page 6 lines 1-3: "In order to estimate NO_2 concentrations, R-LINE incorporates a chemistry module to resolve simple NO to NO_2 chemistry with the Generic Reaction Set (GRS; Valencia et al., 2018) considering the chemical reactions in Table 1."

7. Reviewer #2. 5.) P. 7 line 3 and P.12 line 4: Several of the empirical parametrizations in this paper have been calibrated with NO_2 measurements (parameters C and m). This raises the question about independence of the calibration data. Was the calibration done with an independent NO_2 dataset, not used in the presented model evaluation?

Authors:

The scarcity of measurements didn't allow us to separate the observations for an independent calibration and validation process. We have used the whole set of observations to calibrate and evaluate the model. As responded in reviewer comment 4, we are aware that a longer period with an independent NO_2 dataset would be of interest to evaluate the skills of the model across different seasons. This will be presented in a future work.

8. Reviewer #2. 6.) P.7 line 25: Wind channelling may not occur in streets that are relatively short. The validity of the channelling effect should be analysed for street network of Barcelona.

Authors:

We tried to apply a common simple approach for the entire city. We agree with the reviewer that a more refined implementation of the channelling will be needed in the future but it is out of the scope of the present paper. We didn't assess the effectiveness of the channelling effect formulation because we don't have access to a complete dataset of measured wind conditions within a diverse range of streets in the city.

9. Reviewer #2. 7.) Section 2.3.3: The large scale model grids are step wise in nature. This could lead to significant edge effects caused by the concentration steps between the CMAQ grid cells. How is this considered in the UBS when applying bilinear interpolation to provide background concentrations at the receptors? The error of the background concentrations at low wind speeds should be estimated.

Authors:

We estimate background concentrations at roof level using the urban background scheme in two steps. First, our method selects CMAQ cells as background concentration providers depending on the wind speed and direction. Second, for each receptor we apply a bilinear interpolation method to provide a background at very high resolution calculating weights at each receptor and computing weighted data.

With regard to the comment "The Error of background concentrations at low wind speeds should be estimated", this has been discussed in Section 4.3 Figure 10 panel f) page 25 lines 8-12 as "background concentrations are underestimated at the beginning of the day (1-4 UTC) and are overestimated during nighttime (19-22 UTC) in days with calm conditions."

10. Reviewer #2. 8.) P. 11 line 10: Does wind channelling affect the ratio ws_sfc/ws_bh?

Authors:

No. Wind channelling does not affect the ratio ws_sfc/ws_bh because in the formulation we consider that channelling would affect equally winds at surface and rooftop level. Hence, dividing the channelling effect by itself would give 1 and it would be omitted.

We added a note to clarify this in the revised manuscript in Section 2.3.3 page 12 lines 2-4 as follows, "wind channelling does not affect the ratio ws_sfc/ws_bh because we assume that channelling affects equally winds at surface and rooftop level".

11. Reviewer #2. 9.) P.12 line 5 - 10: Determination of the surface background concentrations needs more explanation. An illustration of the surface background concentration as function of building density would be helpful for understanding how it is derived from the rooftop background under different stability conditions.

Authors:

We added the illustration below to the revised manuscript showing the adimensional vertical mixing variable (fac_{bg}) that is multiplied to rooftop background to obtain surface background concentration as a function of building density and atmospheric stability.

Under low building density (i.e. *bd* below 0.1), background concentrations at surface level and over-roof level are assumed to be similar because there are almost no buildings acting as barriers. When building density increases, the difficulty of the overlying air masses to penetrate the street cavities (building height is 20 m in the illustration) is assumed to increase and more difference in concentrations is expected as a consequence. Under convective conditions, we expect more air mixing between street air and overlying air. Hence, for those conditions we assign a background within the street that is higher compared to stable atmospheric cases. Under stable atmosphere, we assume that a decrease of air mixing will increase air stratification bringing more difference in concentrations between over-roof and surface level concentrations. We include the image below in the revised manuscript in Section 2.3.3 Page 12 to visually support the explanation of the background decay method.

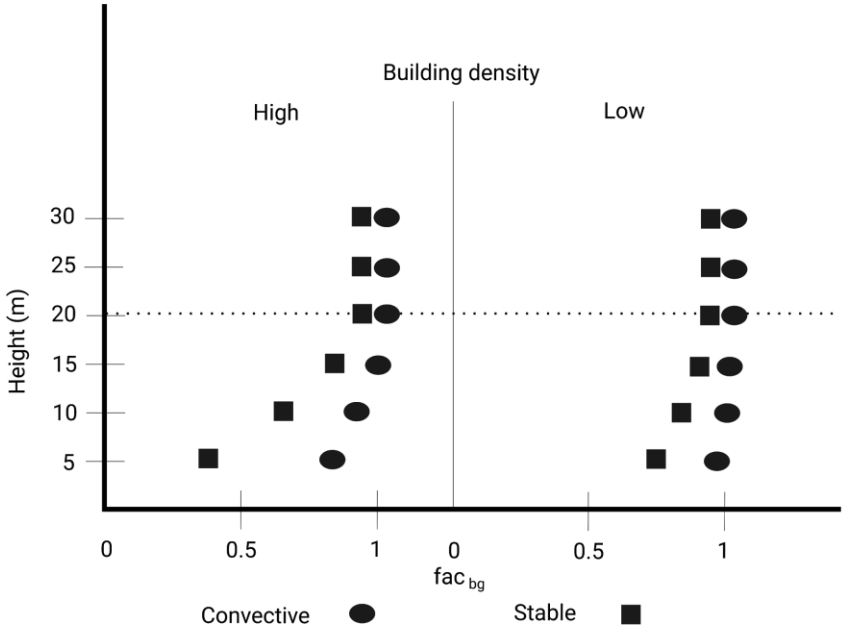

Figure 2. Illustration of the background decay method concept. Building height is approx. 20 m.

12. Reviewer #2. 10.) P.15 line 11 - 12: Which QA/QC procedure was in place for the monitoring with passive dosimeters?

Authors:

Duplicate dosimeters (reproducibility) were installed in some sites, and other dosimeters were installed in the permanent XVPCA network sites for comparison with reference

instrumentation.

The dosimeters were 7 cm diffusion tubes (Palm, GRADKO) that were sent to the laboratory once removed to obtain the concentrations at ambient conditions (nonstandard). Although the concentrations obtained with the dosimeters were ambient, the comparison with the data supplied by the XVPCA network permitted to correct the concentrations with the measures obtained using reference instrumentation at standard conditions. Therefore the concentrations corrected are equivalent to the standard conditions.

13. Reviewer #2. 11.) Table 3 and Figure 8: Measurements at station Gracia-Sant Gervasi are underestimated by all three model configurations in the daytime between morning and afternoon rush hours. Table 3 shows a positive bias for CALIOPE-urban-nl (marked as best performance for MB at this site), but this is deceiving since Figure 8 shows that overestimation at rush hours increased the bias. Obviously, the traffic increment is not correctly represented at this site. Could this be caused by the missing contribution from recirculation of traffic exhaust?

Authors:

The area of Gracia-Sant Gervasi site is a wide street area, which has a large street width compared to building height thus a low aspect ratio (i.e. approx. 0.38). According to Oke (1988), this kind of street is considered to be in the transition between isolated roughness and wake interference flows. The recirculation in that kind of geometrical settings is not as well documented as skimming flow cases (i.e. higher aspect ratio), where a stable recirculatory vortex is established in the canyon. Hence, we do not expect the missing contribution to be from traffic exhaust recirculation. In addition, in case of missing a relevant contribution from recirculation of traffic exhaust in this site we would expect to miss a similar contribution in all the other sites, specially in the street canyons. From the results analysis, we didn't miss that relevant contribution in all the other sites.

14. Reviewer #2. 12.) P.19 line 9 and Figure 8: Give the possible reason for the afternoon underestimation of NO_2 concentrations at sites with low traffic. The underestimation of NO_2 in the afternoons could also be linked to photochemical conversion. Therefore, I recommend to repeat the plots of Figure 8 for NO_X concentrations.

Authors:

As suggested by the reviewer, we repeated the plots of Figure 8 for NO_X concentrations below. We believe that the possible reason for the afternoon underestimation may be

the overestimated mixing from WRF that leads CMAQ to NO_2 underestimations over the city. We find very similar NO_2 and NO_X afternoon underestimations, indicating that photochemical conversion may not be the principal reason.

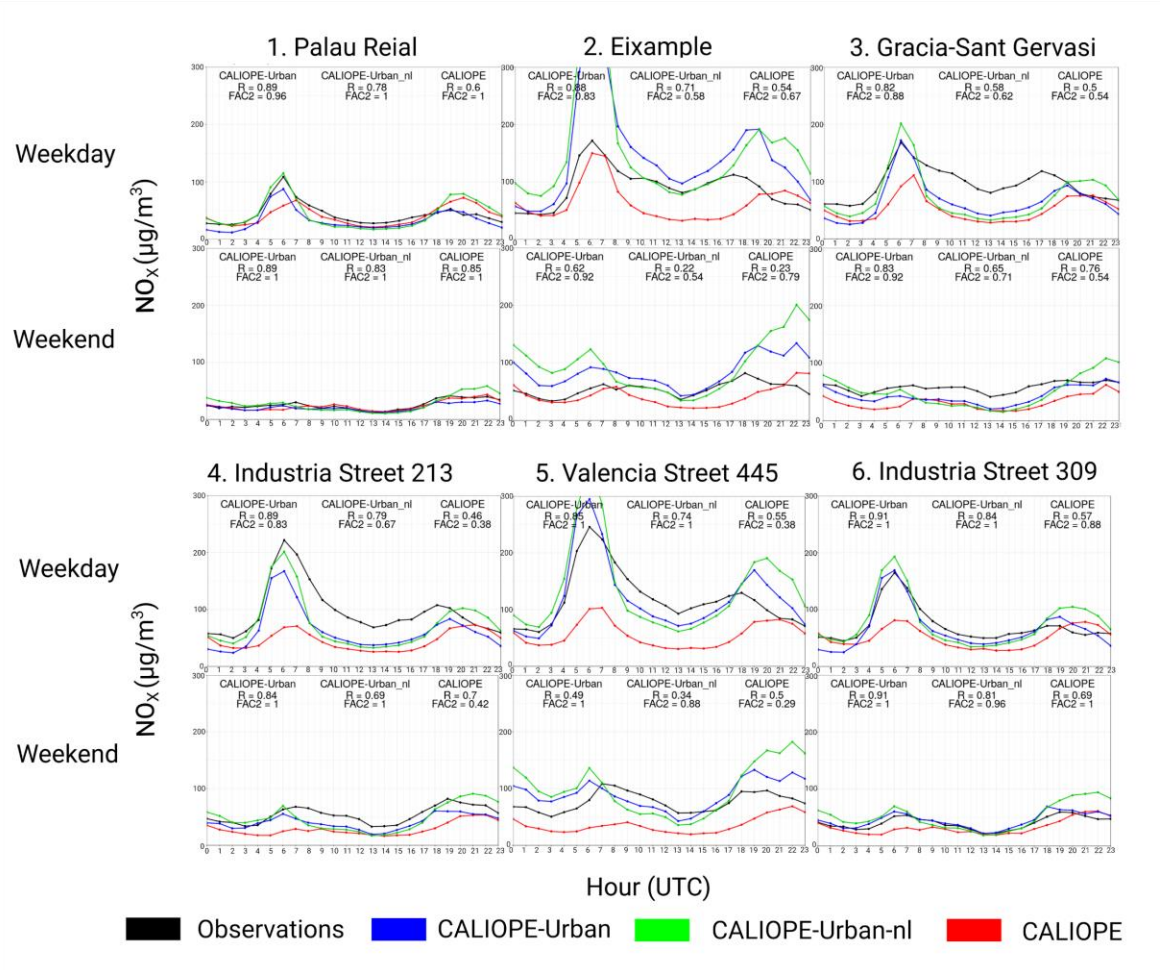

Figure 3. NO_X average daily cycle at all sites described in Sect. 3.1 during April and May 2013 for weekday and weekend. Observations are represented in black coloured lines, red lines are CALIOPE, blue lines are CALIOPE-Urban and green lines represent CALIOPE-Urban without local developments (CALIOPE-Urban-nl).

15.Reviewer#2.Technical Corrections

P. 6 line 8: “This approach addresses” fits better.

Authors: Amended

P.15 line 12: In every km2 grid cell?

Authors:

The measurements were taken independently from the model grid. Every km^2 in the manuscript refers to square kilometers of surface. A comment has been added in the manuscript in page 16 line 2 making it explicit: “In every km^2 of surface there were at least two dosimeters”.

Figure 11: It should be mentioned in the figure caption whether the resolution is 10m x 10m for the entire concentration map or only in the 250m buffers along streets.

Authors:

We added the following note in the figure caption (page 26, Figure 12 in revised manuscript): “The resolution is for the entire concentration map.”

Figure A1: What explains the zero values for the aspect ratio values in the scatterplot?

Authors:

The aspect ratio is assumed to tend to zero when there are no buildings on street segments sides within a distance of 100m. The algorithm to assign aspect ratio to a street segment follow this procedure:

- Build two rectangular buffers at each side of street segment given a rectangle side of 100 m (i.e. set as the maximum distance between road edge and buildings to be considered a street).
- Intersect the Barcelona buildings dataset with the two buffers
- If there are buildings at both sides:
  - Estimate the minimum distance between road edge (line) and buildings, which is assumed to represent the distance between road edge and buildings on the side of the street. Add distances at both sides of the road edge to obtain the street width.
  - Estimate the average building height of the buildings falling in both buffers
  - Estimate the aspect ratio by dividing average building height by street width.
- If there are no buildings at both sides: assign aspect ratio equal zero.

References

Beevers, S., Kitwiroon, N., Williams, M.L., and Carslaw, D.C. “One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions”. Atmospheric Environment 59, pp. 47–58. DOI: https://doi.org/10.1016/j.atmosenv.2012.05.034. 2012.

Benson, P.E., CALINE4-a Dispersion Model for Predicting Air Pollution Concentration Near Roadways. FHWA/CA/TL-84/15, p. 245. 1989.

Berkowicz, R. “A simple model for urban background pollution”. Environmental Monitoring and Assessment 65.1-2, pp. 259–267. DOI : https://doi.org/10.1023/A:1006466025186. 2000.

Fagerli, H., Denby, B., and Wind, P. “Assessment of LRT contribution to cities in Europe using uEMEP?”. 2019.

Heist, D., Isakov, V. Perry, S., Snyder, M., Venkatram, A.“Estimating near-road pollutant dispersion: A model inter-comparison”. Transportation Research Part D: Transport and Environment 25, 93-105. 2013.

Hood, C., Mackenzie, I., Stocker, J., Johnson, K., Carruthers, D., Vieno, M., and Doherty, R. “Air quality simulations for London using a coupled regional-to-local modelling system”. Atmospheric Chemistry and Physics Discussions February, pp. 1–44. DOI: https://doi.org/10.5194/acp-18-11221-2018. 2018.

Hotchkiss and Harlow. “Air Pollution Transport in Street Canyons”. EPA-R4-73-029. 1973.

Isakov, V. et al. “Air Quality Modeling in Support of the Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS)”. International Journal of Environmental Research and Public Health 11.9, pp. 8777–8793. DOI: https://doi.org/10.3390/ijerph110908777. 2014.

Jensen, S. S., Ketzel, M., Becker, T., Christensen, J., Brandt, J., Plejdrup, M., Winther, M., Nielsen, O.K., Hertel, O., and Ellermann, T. “High resolution multi-scale air quality modelling for all streets in Denmark”. Transportation Research Part D: Transport and Environment 52, pp. 322–339. DOI: https://doi.org/10.1016/j.trd.2017.02.019. 2017.

Kim, Y., Wu, Y., Seigneur, C., and Roustan, Y. “Multi-scale modeling of urban air pollution : development and application of a Street-in-Grid model by coupling MUNICH and”. Geoscientific Model Development Discussions September, pp. 1–24. DOI: https://doi.org/10.5194/gmd-11-611-2018. 2018.

Maiheu, B., Lefebvre, W., Walton, H., Dajnak, D., Janssen, S., Williams, M., Blyth, L., and Beevers, S. Improved Methodologies for NO2 Exposure Assessment in the EU. Tech. rep. 2. VITO, 2017. URL: http://ec.europa.eu/environment/air/publications/models.htm.

Moussafir, J, Olry, C, Nibart, M, Albergel, A, Armand, P, Duchenne, C, and Thobois, L. “Aircity: a very high resolution atmospheric dispersion modeling system for Paris”. American Society of Mechanical Engineers, Fluids Engineering Division (Publication) FEDSM 1. DOI: https://doi.org/10.1115/FEDSM2014-

21820. 2014.

Oke, T.R. “Street design and urban canopy layer climate”. Energy and Buildings 11.1-3, pp. 103–113. DOI: https://10.1016/0378-7788(88)90026-6. 1988.

Pay, M.T., Martınez, F., Guevara, M., and Baldasano, J.M. “Air quality forecasts on a kilometer-scale grid over complex Spanish terrains”. Geoscientific Model Development 7.5, pp. 1979–1999. DOI: https://doi.org/10.5194/gmd-7-1979-2014. 2014.

Rotach, M. W. “Profiles of turbulence statistics in and above an urban street canyon”. Atmospheric Environment 29.13, pp. 1473–1486. DOI : 10.1016/1352-2310(95)00084-C. 1995.

Salizzoni, P., Soulhac, L., and Mejean, P. “Street canyon ventilation and atmospheric turbulence”. Atmospheric Environment 43.32, pp. 5056–5067. DOI : 10.1016/j.atmosenv.2009.06.045. URL: http://dx.doi.org/10.1016/j.atmosenv.2009.06.045, 2009.

Snyder, M.G., Venkatram, A., Heist, D.K., Perry, S.G., Petersen, W.B., and Isakov, V. “RLINE: a line source dispersion model for near-surface releases”. Atmospheric Environment 77, pp. 748–756. DOI : https://doi.org/10.1016/j.atmosenv.2013.05.074. 2013.

Soulhac, L., Salizzoni, P., Cierco, F. X., and Perkins, R. “The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model”. Atmospheric Environment 45.39, pp. 7379–7395. DOI: https://10.1016/j.atmosenv.2011.07.008. 2011.

CALIOPE-Urban v1.0: Coupling R-LINE with a mesoscale air quality modelling system for urban air quality forecasts over Barcelona city (Spain)

Jaime Benavides[1], Michelle Snyder[2], Marc Guevara[1], Albert Soret[1], Carlos Pérez García-Pando[1], Fulvio Amato[3], Xavier Querol[3], and Oriol Jorba[1]

[1]Barcelona Supercomputing Center, Spain
[2]Institute for the Environment, University of North Carolina at Chapel Hill, USA
[3]Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Spain

Correspondence to: Oriol Jorba (oriol.jorba@bsc.es)

Abstract. The NO_2 annual air quality limit value is systematically exceeded in many European cities. In this context, understanding human exposure, improving policy and planning, and providing forecasts requires the development of accurate air quality models at urban (street–level) scale. We describe CALIOPE-Urban, a system coupling CALIOPE - an operational mesoscale air quality forecast system based on HERMES (emissions), WRF (meteorology) and CMAQ (chemistry) models - with the urban roadway dispersion model R-LINE. Our developments have focused on Barcelona city (Spain), but the methodology may be replicated for other cities in the future. WRF drives pollutant dispersion and CMAQ provides background concentrations to R-LINE. Key features of our system include the adaptation of R-LINE to street canyons, the use of a new methodology that considers upwind grid cells in CMAQ to avoid double counting traffic emissions, a new method to estimate local surface roughness within street canyons, and a vertical mixing parametrization that considers urban geometry and atmospheric stability to calculate surface level background concentrations. We show that the latter is critical to correct the nighttime overestimations in our system. Both CALIOPE and CALIOPE-Urban are evaluated using two sets of observations. The temporal variability is evaluated against measurements from five traffic sites and one urban background site for April-May 2013. While both systems show a fairly good agreement at the urban background site, CALIOPE-Urban shows a better agreement in traffic sites. The spatial variability is evaluated using 182 passive dosimeters that were distributed across Barcelona during two weeks for February-March 2017. In this case, also the coupled system shows a more realistic distribution than the mesoscale system, which systematically underpredicts NO_2 close to traffic emission sources. Overall CALIOPE-Urban improves mesoscale model results, demonstrating that the combination of both scales provides a more realistic representation of NO_2 spatio-temporal variability in Barcelona.

1 Introduction

Persistent exposure to high NO_2 atmospheric concentrations in cities causes detrimental health effects (e.g., Sunyer et al., 2015; Barone-Adesi et al., 2015). In 2016, 19 out of the 28 European Union (EU) countries reported NO_2 exceedances of

CALIOPE-Urban v1.0: Coupling R-LINE with a mesoscale air quality modelling system for urban air quality forecasts over Barcelona city (Spain)

Jaime Benavides[1], Michelle Snyder[2], Marc Guevara[1], Albert Soret[1], Carlos Pérez García-Pando[1], Fulvio Amato[3], Xavier Querol[3], and Oriol Jorba[1]

[1]Barcelona Supercomputing Center, Spain
[2]Institute for the Environment, University of North Carolina at Chapel Hill, USA
[3]Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Spain

Correspondence to: Oriol Jorba (oriol.jorba@bsc.es)

Abstract. The NO_2 annual air quality limit value is systematically exceeded in many European cities. In this context, understanding human exposure, improving policy and planning, and providing forecasts requires the development of accurate air quality models at urban (street–level) scale. We describe CALIOPE-Urban, a system coupling CALIOPE - an operational mesoscale air quality forecast system based on HERMES (emissions), WRF (meteorology) and CMAQ (chemistry) models - with the urban roadway dispersion model R-LINE. Our developments have focused on Barcelona city (Spain), but the methodology may be replicated for other cities in the future. WRF drives pollutant dispersion and CMAQ provides background concentrations to R-LINE. Key features of our system include the adaptation of R-LINE to street canyons, the use of a new methodology that considers upwind grid cells in CMAQ to avoid double counting traffic emissions, a new method to estimate local surface roughness within street canyons, and a vertical mixing parametrization that considers urban geometry and atmospheric stability to calculate surface level background concentrations. We show that the latter is critical to correct the nighttime overestimations in our system. Both CALIOPE and CALIOPE-Urban are evaluated using two sets of observations. The temporal variability is evaluated against measurements from five traffic sites and one urban background site for April-May 2013. While both systems show a fairly good agreement at the urban background site, CALIOPE-Urban shows a better agreement in traffic sites. The spatial variability is evaluated using 182 passive dosimeters that were distributed across Barcelona during two weeks for February-March 2017. In this case, the coupled system also shows a more realistic distribution than the mesoscale system, which systematically underpredicts NO_2 close to traffic emission sources. Overall CALIOPE-Urban improves mesoscale model results, demonstrating that the combination of both scales provides a more realistic representation of NO_2 spatio-temporal variability in Barcelona.

1 Introduction

Persistent exposure to high NO_2 atmospheric concentrations in cities causes detrimental health effects (e.g., Sunyer et al., 2015; Barone-Adesi et al., 2015). In 2016, 19 out of the 28 European Union (EU) countries reported NO_2 exceedances of

1

the annual air quality limit value (40 µg m^{-3}) mostly at urban traffic monitoring stations (EEA, 2018). In this context there is a need for NO_2 data at street level in urban areas that enables individuals and communities to mitigate the problem by, for example, walking in less polluted streets or reducing traffic in school areas. However, both the poor density of air quality monitoring stations and the resolution of mesoscale air quality modeling systems (on the order of 1-km grid resolution), do not adequately represent the NO_2 concentration gradients that typically occur near heavily trafficked streets (Duyzer et al., 2015; Borge et al., 2014). Urban dispersion models can estimate these gradients but their use has been typically limited to historic periods partly because the needed background concentrations and meteorological input have been approximated using observations (Vardoulakis et al., 2003).

In order to overcome these limitations, coupling regional and urban scale models has been recently found to be successful in some cities. Hood et al. (2018) coupled a regional climate-chemistry model with 5 km horizontal resolution (EMEP4UK) with the fine-scale model ADMS-URBAN to simulate air quality over London in 2012. They compared the coupled system results with the regional and the fine-scale models run separately. Authors found that both the fine-scale model and the coupled system performed better than the regional for NO_2 at both annual mean and hourly concentration levels due to the explicit treatment of traffic emissions within the city. Jensen et al. (2017) estimated annual NO_2 concentrations at 2.4 million addresses in Denmark using the street canyon model OSPM coupled with DEHM for regional background concentrations and UBM for urban background obtaining a good correlation in Copenhaguen (r^2 = 0.70) against 98 measurement sites for NO_2 in the year 2012. In addition, Maiheu et al. (2017) estimated EU-wide NO_2 annual average levels at 100 meter resolution with a regional model coupled with a dispersion kernel-based method. The approach does not produce hourly concentration levels and approximates road-link level traffic emissions by distributing the regional grid cell traffic emissions to each road-link based on road capacity. Hence, it provides more spatial detail than previous EU scale NO_2 assessment studies, but more specific methods are required to resolve air quality in cities. In this sense, there is a lack of air quality urban forecasting methodologies that can be applied to a diverse range of cities and that consistently resolve at least some of the major challenges already identified by the community, i.e., 1) downscaling regional meteorology to street level as required to drive pollutant dispersion; 2) obtaining background concentrations from the mesoscale system avoiding the double counting of traffic emissions. Additionally, we consider vertical mixing with background air a key process to be resolved when coupling the regional and urban scales.

Different approaches to downscale mesoscale meteorology are found in the research literature. Brousse et al. (2016) applied the Weather Research and Forecasting meteorological model (WRF) using the Building Effect Parametrization (Martilli et al., 2002) over Madrid considering WUDAPT Local Climate Zone data (Bechtel et al., 2015). This approach increases the mesoscale model's ability to resolve urban processes but does not reproduce the specific meteorological conditions in each street as required by dispersion models. Kochanski et al. (2015) used a simplified CFD (QUIC) in combination with WRF to estimate wind conditions at street level. Hood et al. (2018) estimate an urban canopy flow field at the same resolution of the regional model. This calculation is based on the variation of surface roughness within the city. This approach includes the variation of some atmospheric stability parameters (e.g. friction velocity) but it neglects the variation of vertical mixing with background air depending on atmospheric stability and urban geometry. On the other hand, Jensen et al. (2017) do not consider atmospheric stability within the street canyon model OSPM and within the vertical mixing with background air. The approach

the annual air quality limit value (40 µg m^{-3}) mostly at urban traffic monitoring stations (EEA, 2018). In this context there is a need for NO_2 data at street level in urban areas that enables individuals and communities to mitigate the problem by, for example, walking in less polluted streets or reducing traffic in school areas. However, both the poor density of air quality monitoring stations and the resolution of mesoscale air quality modeling systems (on the order of 1-km grid resolution), do not adequately represent the NO_2 concentration gradients that typically occur near heavily trafficked streets (Duyzer et al., 2015; Borge et al., 2014). Urban dispersion models can estimate these gradients but their use has been typically limited to historic periods partly because the needed background concentrations and meteorological input have been approximated using observations (Vardoulakis et al., 2003).

In order to overcome these limitations, coupling off-line the regional and urban scales by downscaling the regional model using a dispersion kernel has been successfully applied in some cities (Beevers et al., 2012; Moussafir et al., 2014; Isakov et al., 2014; Jensen et al., 2017; Maiheu et al., 2017; Kim et al., 2018; Hood et al., 2018; Fagerli et al., 2019). For instance, Hood et al. (2018) coupled a regional climate-chemistry model with 5 km horizontal resolution (EMEP4UK) with the fine-scale model ADMS-URBAN to simulate air quality over London in 2012. They compared the coupled system results with the regional and the fine-scale models run separately. Authors found that both the fine-scale model and the coupled system performed better than the regional for NO_2 at both annual mean and hourly concentration levels due to the explicit treatment of traffic emissions within the city. In addition, Jensen et al. (2017) estimated annual NO_2 concentrations at 2.4 million addresses in Denmark using the street canyon model OSPM coupled with DEHM for regional background concentrations and UBM for urban background obtaining a good correlation in Copenhaguen (r^2 = 0.70) against 98 measurement sites for NO_2 in the year 2012. Maiheu et al. (2017) covered a broader spatial context, estimating EU-wide NO_2 annual average levels at 100 meter resolution with a regional model coupled with a dispersion kernel-based method. The approach does not produce hourly concentration levels and approximates road-link level traffic emissions by distributing the regional grid cell traffic emissions to each road-link based on road capacity. Hence, it provides more spatial detail than previous EU scale NO_2 assessment studies, but more specific methods are required to resolve air quality in cities. In this sense, there is a lack of air quality urban forecasting methodologies that can be applied to a diverse range of cities and that consistently resolve at least some of the major challenges already identified by the community, i.e., 1) downscaling regional meteorology to street level as required to drive pollutant dispersion; 2) obtaining background concentrations from the mesoscale system avoiding the double counting of traffic emissions. Additionally, we consider vertical mixing with background air a key process to be resolved when coupling the regional and urban scales.

Different approaches to downscale mesoscale meteorology are found in the research literature. Brousse et al. (2016) applied the Weather Research and Forecasting meteorological model (WRF) using the Building Effect Parametrization (Martilli et al., 2002) over Madrid considering WUDAPT Local Climate Zone data (Bechtel et al., 2015). This approach increases the mesoscale model's ability to resolve urban processes but does not reproduce the specific meteorological conditions in each street as required by dispersion models. Kochanski et al. (2015) used a simplified CFD (QUIC) in combination with WRF to estimate wind conditions at street level. Hood et al. (2018) estimate an urban canopy flow field at the same resolution of the regional model. This calculation is based on the variation of surface roughness within the city. This approach includes the variation of some atmospheric stability parameters (e.g. friction velocity) but it neglects the variation of vertical mixing with

presented here to downscale mesoscale meteorology to street-scale describing wind conditions and atmospheric stability in each street can be a promising solution to drive dispersion models and vertical mixing.

Background concentrations can be obtained from observations or mesoscale models, which are commonly used in forecasting applications. However, coupling mesoscale and urban dispersion models can lead to a double counting of traffic emissions. To avoid double counting, Arunachalam et al. (2014) multiply urban background site observations by an estimated ratio between two mesoscale air quality simulations. The first run contains all the emission sources and the second neglects traffic emissions. Lefebvre et al. (2011) and Stocker et al. (2014) run first the urban dispersion model at mesoscale grid resolution with only traffic emissions and subtract its result to the mesoscale model simulation, which includes all the emission sources. Then, street scale model outputs are added to the result from the prior computation at finer resolution. Although, these methods avoid double counting emissions they do not explicitly account for vertical mixing, a process that occurs at the intersection of regional and street scales. Urban air quality models such as SIRANE (Soulhac et al., 2011) have already implemented vertical mixing depending on local meteorology. In this study, we will show that this process may be relevant and explain some systematic errors found in the literature: nighttime NO_2 concentration values tend to be overestimated and afternoon values tend to be underestimated in traffic areas (e.g., Hood et al., 2018). Further efforts are necessary to explicitly resolve processes happening among scales and to correct these biases in the mentioned periods of the day.

This work describes a methodology to couple the mesoscale air quality forecasting system CALIOPE (Baldasano et al., 2011; http://www.bsc.es/caliope/?language=en) with the Research LINE source dispersion model (R-LINE; Snyder et al., 2013) and its evaluation over the city of Barcelona, Spain. In Barcelona, chronic NO_2 exceedances have been recorded since the year 2000, and according to the local Public Health Agency about 68% of citizens were exposed to NO_2 levels above the annual air quality limit value in 2016 (ASPB, 2017). Barcelona has a very high vehicle density (approx. 5500 vehicles km^{-2}) and the majority of passenger cars are diesel (67%) (Barcelona City Council, 2017). Located in the north east of the Iberian Peninsula, Barcelona is surrounded by the Mediterranean sea, two rivers and a mountain range. Due to its coastal emplacement, during the warm season, transport and dispersion of air pollutants within the city are dominated by the breeze blowing in from the sea during daytime and from the land during nighttime. This pattern persists under the presence of high-pressure systems accompanied by clear skies and warm temperatures in the summer season. In contrast, the winter season is dominated by north western advections typically cleaning the atmosphere of the city (Jorba et al., 2011). Our aim is to produce more accurate NO_2 concentrations with CALIOPE-Urban, the coupled system, than with the mesoscale system alone and give a more realistic representation of NO_2 spatial distribution and temporal variability across the city. To achieve these objectives a set of system enhancements have been implemented: an adaptation of R-LINE to dense urban areas (e.g. street canyons); a background model to estimate over background roof-level concentrations; a parametrization of the vertical mixing to estimate background concentrations within the street that considers atmospheric stability and urban geometry; and a local surface roughness parametrization to estimate turbulent parameters within a street canyon. The mesoscale system has been executed using the operational forecast configuration. We compare the estimated temporal variability of NO_2 concentrations from the coupled modeling system with those derived from CALIOPE and with ambient street-level measurements (i.e. 5 traffic

background air depending on atmospheric stability and urban geometry. On the other hand, Jensen et al. (2017) do not consider atmospheric stability within the street canyon model OSPM and within the vertical mixing with background air. The approach presented here to downscale mesoscale meteorology to street-scale describing wind conditions and atmospheric stability in each street can be a promising solution to drive dispersion models and vertical mixing.

Background concentrations can be obtained from observations or mesoscale models, which are commonly used in forecasting applications. However, coupling mesoscale and urban dispersion models can lead to a double counting of traffic emissions. To avoid double counting, Arunachalam et al. (2014) multiply urban background site observations by an estimated ratio between two mesoscale air quality simulations. The first run contains all the emission sources and the second neglects traffic emissions. Lefebvre et al. (2011) and Stocker et al. (2014) run first the urban dispersion model at mesoscale grid resolution with only traffic emissions and subtract its result from the mesoscale model simulation, which includes all the emission sources. Then, street scale model outputs are added to the result from the prior computation at finer resolution. Although, these methods avoid double counting emissions they do not explicitly account for vertical mixing, a process that occurs at the intersection of regional and street scales. Urban air quality models such as SIRANE (Soulhac et al., 2011) have already implemented vertical mixing depending on local meteorology. In this study, we will show that this process may be relevant and explain some systematic errors found in the literature: nighttime NO_2 concentration values tend to be overestimated and afternoon values tend to be underestimated in traffic areas (e.g., Hood et al., 2018). Further efforts are necessary to explicitly resolve processes happening among scales and to correct these biases in the mentioned periods of the day.

This work describes a methodology to couple the mesoscale air quality forecasting system CALIOPE (Baldasano et al., 2011; http://www.bsc.es/caliope/?language=en) with the Research LINE source dispersion model (R-LINE; Snyder et al., 2013) and its evaluation over the city of Barcelona, Spain. In Barcelona, chronic NO_2 exceedances have been recorded since the year 2000, and according to the local Public Health Agency about 68% of citizens were exposed to NO_2 levels above the annual air quality limit value in 2016 (ASPB, 2017). Barcelona has a very high vehicle density (approx. 5500 vehicles km^{-2}) and the majority of passenger cars are diesel (67%) (Barcelona City Council, 2017). Located in the north east of the Iberian Peninsula, Barcelona is surrounded by the Mediterranean sea, two rivers and a mountain range. Due to its coastal emplacement, during the warm season, transport and dispersion of air pollutants within the city are dominated by the breeze blowing in from the sea during daytime and from the land during nighttime. This pattern persists under the presence of high-pressure systems accompanied by clear skies and warm temperatures in the summer season. In contrast, the winter season is dominated by north western advections typically cleaning the atmosphere of the city (Jorba et al., 2011). Our aim is to produce more accurate NO_2 concentrations with CALIOPE-Urban, the coupled system, than with the mesoscale system alone and give a more realistic representation of NO_2 spatial distribution and temporal variability across the city. To achieve these objectives a set of system enhancements have been implemented: an adaptation of R-LINE to dense urban areas (e.g. street canyons); a background model to estimate background concentrations at roof-level; a parametrization of the vertical mixing to estimate background concentrations within the street that considers atmospheric stability and urban geometry; and a local surface roughness parametrization to estimate turbulent parameters within a street canyon. The mesoscale system has been executed using the operational forecast configuration. We compare the estimated temporal variability of NO_2 concentrations

site and 1 urban background site) in April and May 2013. Its spatial variability is evaluated using a two-week measurement campaign that covered Barcelona with 182 NO_2 passive dosimeters for two weeks in February and March 2017.

2 Methods

CALIOPE-Urban estimates hourly NO_2 concentrations by coupling the CALIOPE mesoscale air quality forecasting system, providing background concentrations, meteorological data and road-link traffic emissions, with the R-LINE dispersion model adapted to street canyons. Here we introduce and describe the components of the coupled model as depicted in Fig. 1.

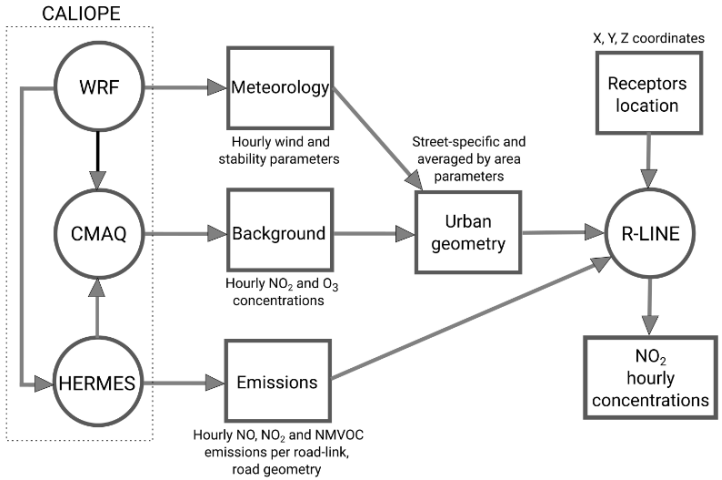

Figure 1. CALIOPE-Urban workflow. Models are represented by circles and data by rectangular shapes. CALIOPE is left untouched. Meteorology and background from WRF and CMAQ are combined with urban geometry to create inputs for R-LINE. R-LINE dispersion is left untouched, after adjusting meteorology and surface roughness for local urban geometry.

2.1 Mesoscale air quality forecasting system-CALIOPE

CALIOPE (Baldasano et al., 2011) integrates the Weather Research and Forecasting model version 3 (WRF; Skamarock and Klemp, 2008), the High-Elective Resolution Modelling Emission System (HERMESv2.0; Guevara et al., 2013), the Community Multiscale Air Quality Modeling System version 5.0.2 (CMAQ; Byun and Schere, 2006) and the mineral Dust REgional Atmospheric Model (BSC-DREAM8b; Basart et al., 2012). The mesoscale system is run over Europe at a 12 km × 12 km

from the coupled modeling system with those derived from CALIOPE and with ambient street-level measurements (i.e. 5 traffic sites and 1 urban background site) in April and May 2013. Its spatial variability is evaluated using a two-week measurement campaign that deployed 182 NO_2 passive dosimeters across Barcelona for two weeks in February and March 2017.

2 Methods

CALIOPE-Urban estimates hourly NO_2 concentrations by coupling the CALIOPE mesoscale air quality forecasting system, providing background concentrations, meteorological data and road-link traffic emissions, with the R-LINE dispersion model adapted to street canyons. Here we introduce and describe the components of the coupled model as depicted in Fig. 1.

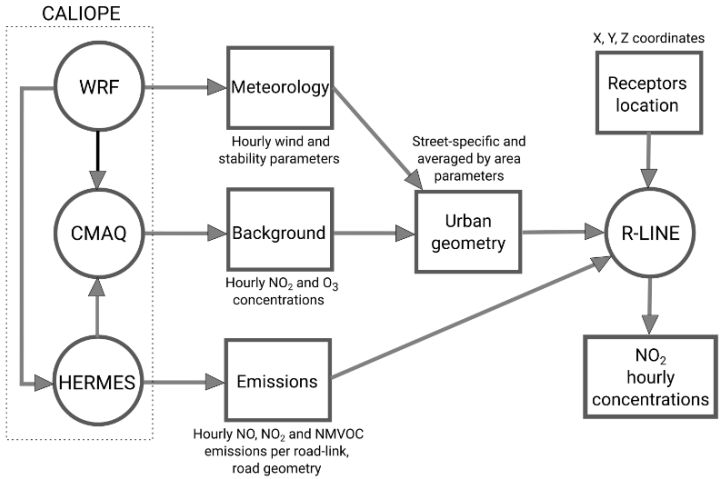

Figure 1. CALIOPE-Urban workflow. Models are represented by circles and data by rectangular shapes. CALIOPE is left untouched. Meteorology and background from WRF and CMAQ are combined with urban geometry to create inputs for R-LINE. R-LINE dispersion is left untouched, after adjusting meteorology and surface roughness for local urban geometry.

2.1 Mesoscale air quality forecasting system-CALIOPE

CALIOPE (Baldasano et al., 2011) integrates the Weather Research and Forecasting model version 3 (WRF; Skamarock and Klemp, 2008), the High-Elective Resolution Modelling Emission System (HERMESv2.0; Guevara et al., 2013), the Community Multiscale Air Quality Modeling System version 5.0.2 (CMAQ; Byun and Schere, 2006) and the mineral Dust REgional

horizontal resolution, Iberian Peninsula at 4 km × 4 km, and the Catalonian domain, including Barcelona, at 1 km × 1 km. CALIOPE results have been evaluated in detail elsewhere (e.g., Pay et al., 2014).

In our system, WRF uses the Global Forecasting System (GFS) model initial/boundary conditions from the National Centers for Environmental Prediction (NCEP) to forecast the mesoscale meteorological conditions. Three nested domains are designed to provide a final high-resolution run over Catalonian domain. In the vertical, WRF is configured with 38 sigma layers up to 50 hPa, where 11 cover the planetary boundary layer (PBL). Our WRF setup utilizes the Rapid Radiation Transfer model for long-wave radiation and Dudhia for short-wave, the Kain Fritsch cumulus parameterization, the single-moment 3-class microphysics scheme, the Yonsei University PBL scheme, and the Noah land-surface model based on the CORINE land-use data from the year 2006.

For the mesoscale model, pollutant emissions are obtained from the high resolution emission model HERMESv2.0 gridded up to 1 km × 1 km and temporal (1h) resolution. HERMESv2.0 estimates atmospheric emissions for Europe and Spain according to the Selected Nomenclature for Air Pollution (SNAP) and taking the year 2009 as the reference period. Emissions are estimated for nitrogen oxides (NO_X), non-methane volatile organic compounds (NMVOCs), sulphur dioxide, carbon monoxide, ammonia, total suspended particles, PM10 and PM2.5 fractions. The final model output consists of hourly, gridded and speciated emissions according to the CB05 chemical mechanism used by the chemical transport model CMAQ. For Europe, HERMESv2.0 implements a SNAP sector-dependent spatial, temporal and speciation treatment of the original annual EMEP gridded emissions (Ferreira et al., 2013). For Spain, the model uses a bottom-up approach for pollutant sources including point (e.g. power plants, industries), maritime (e.g. ports), air traffic (e.g. airports), agricultural machinery (e.g. tractors and harvesters) and road transport. For the rest of pollutant sources a combination of top-down approaches (i.e. residential/commercial combustion; energy consumption statistics combined with a population map) and downscaling methodologies (i.e. use of solvents, extraction and distribution of fossil fuels; specific spatial proxies and temporal profiles assigned to the Spanish National Emission Inventory by categories at third level of SNAP) is adopted. The results of the HERMESv2.0 model have been used to support several air quality evaluation and planning studies (e.g., Baldasano et al., 2014; Soret et al., 2014) as well as emission inventory intercomparison exercises (Guevara et al., 2017).

The chemical transport model used in the CALIOPE system is the CMAQv5.0.2. It uses the CB05 gas-phase chemical mechanism, the AERO5 aerosol scheme, and an in-line photolysis calculation. CMAQ vertical levels are collapsed from the 38 WRF levels to 15 layers up to 50 hPa with six layers falling within the PBL. We use as boundary conditions for the European domain MOZART-4.

2.2 Street scale dispersion model: R-LINE

R-LINE is a near-road Gaussian dispersion model (Snyder et al., 2013) that incorporates state-of-the-art Gaussian dispersion curves (Venkatram et al., 2013) to simulate dispersion of road source emissions. The model resolves either numerically or analytically the integration of the contributions of point sources along a street segment (Snyder and Heist, 2013). The first option is more accurate and the latter spends less time computing dispersion. The analytical version is best suited for near-ground level sources and receptors. In order to estimate NO_2 concentrations R-LINE incorporates a chemistry module to

Atmospheric Model (BSC-DREAM8b; Basart et al., 2012). The mesoscale system is run over Europe at a 12 km × 12 km horizontal resolution, Iberian Peninsula at 4 km × 4 km, and the Catalonian domain, including Barcelona, at 1 km × 1 km. CALIOPE results have been evaluated in detail elsewhere (e.g., Pay et al., 2014).

In our system, WRF uses the Global Forecasting System (GFS) model initial/boundary conditions from the National Centers for Environmental Prediction (NCEP) to forecast the mesoscale meteorological conditions. Three nested domains are designed to provide a final high-resolution run over Catalonian domain. In the vertical, WRF is configured with 38 sigma layers up to 50 hPa, where 11 cover the planetary boundary layer (PBL). Our WRF setup utilizes the Rapid Radiation Transfer model for long-wave radiation and Dudhia for short-wave, the Kain Fritsch cumulus parameterization, the single-moment 3-class microphysics scheme, the Yonsei University PBL scheme, and the Noah land-surface model based on the CORINE land-use data from the year 2006.

For the mesoscale model, pollutant emissions are obtained from the high resolution emission model HERMESv2.0 gridded up to 1 km × 1 km and temporal (1h) resolution. HERMESv2.0 estimates atmospheric emissions for Europe and Spain according to the Selected Nomenclature for Air Pollution (SNAP) and taking the year 2009 as the reference period. Emissions are estimated for nitrogen oxides (NO_X), non-methane volatile organic compounds (NMVOCs), sulphur dioxide, carbon monoxide, ammonia, total suspended particles, PM10 and PM2.5 fractions. The final model output consists of hourly, gridded and speciated emissions according to the CB05 chemical mechanism used by the chemical transport model CMAQ. For Europe, HERMESv2.0 implements a SNAP sector-dependent spatial, temporal and speciation treatment of the original annual EMEP gridded emissions (Ferreira et al., 2013). For Spain, the model uses a bottom-up approach for pollutant sources including point (e.g. power plants, industries), maritime (e.g. ports), air traffic (e.g. airports), agricultural machinery (e.g. tractors and harvesters) and road transport. For the rest of pollutant sources a combination of top-down approaches (i.e. residential/commercial combustion; energy consumption statistics combined with a population map) and downscaling methodologies (i.e. use of solvents, extraction and distribution of fossil fuels; specific spatial proxies and temporal profiles assigned to the Spanish National Emission Inventory by categories at third level of SNAP) is adopted. The results of the HERMESv2.0 model have been used to support several air quality evaluation and planning studies (e.g., Baldasano et al., 2014; Soret et al., 2014) as well as emission inventory intercomparison exercises (Guevara et al., 2017).

The chemical transport model used in the CALIOPE system is the CMAQv5.0.2. It uses the CB05 gas-phase chemical mechanism, the AERO5 aerosol scheme, and an in-line photolysis calculation. CMAQ vertical levels are collapsed from the 38 WRF levels to 15 layers up to 50 hPa with six layers falling within the PBL. We use as boundary conditions for the European domain MOZART-4.

2.2 Street scale dispersion model: R-LINE

R-LINE is a near-road Gaussian dispersion model (Snyder et al., 2013) that incorporates state-of-the-art Gaussian dispersion curves (Venkatram et al., 2013) to simulate dispersion of road source emissions. The model resolves either numerically or analytically the integration of the contributions of point sources along a street segment (Snyder and Heist, 2013). The first option is more accurate and the latter spends less time computing dispersion. The analytical version is best suited for near-

resolve simple NO to NO_2 chemistry with the Generic Reaction Set (GRS; Valencia et al., 2018). R-LINE has been applied to estimate exposure to traffic-related air pollutants in a large scale study in Detroit, United States (Isakov et al., 2014). However, to our knowledge it has not been applied to European cities, where street canyon morphology dominates. Hence, in order to apply R-LINE over Barcelona its meteorology has been adapted to street canyons as described in Sect. 2.3.1 and the background concentrations are obtained from CMAQ model considering local meteorology and urban geometry as described in Sect. 2.3.3.

2.3 Coupling CALIOPE with R-LINE

CALIOPE and R-LINE are coupled offline, first CALIOPE is run over Europe, Iberian Peninsula and Catalonia and then R-LINE is executed for Barcelona city. This approach presents two main challenges that have already been highlighted in the research literature: (1) downscaling regional meteorology to street scale to drive pollutant dispersion; and (2) obtaining background concentrations from the mesoscale model without double counting traffic emissions in regional and street scale models. In addition to these challenges, we consider relevant to couple meteorology and background concentrations in a consistent way, taking into account atmospheric stability and urban geometry when estimating background contribution within urban streets. Here we describe our methodology when coupling the models to mitigate these challenges.

2.3.1 Meteorology

WRF bottom layer results are assumed to represent over roof wind and stability conditions because its mid-point height (20.3 m) is similar to average building height ($\overline{bh}$) in a typical neighbourhood of Barcelona (e.g. Eixample district; 20.7 m). WRF is executed consistently with the forecasting air quality system CALIOPE, giving a constant surface roughness (z_0) equal to 1 m over the urban area. In order to apply R-LINE over Barcelona, its meteorology has been adapted to street canyons. We have developed a methodology to estimate specific z_0 based on urban geometry (e.g. building height, street width). Once z_0 is adjusted, the displacement height ($dispht$), friction velocity ($u*$), convective velocity scale ($w*$), PBL height, and Monin-Obukov length (L) are re-calculated (Cimorelli et al., 2005). The increase in z_0 generally leads to a larger $dispht$, $u*$, $w*$, and PBL height. Therefore, L is less stable and atmospheric conditions are more convective. Ultimately, these adjustments have an effect on the way the winds are profiled and on the rate of dispersion of the roadway emissions within the urban area.

The geometrical parameters used for z_0 calculation are divided into two categories: (1) averaged over an area of 250m $\times$ 250m (planar building density, bd; average building height, $\overline{bh}$; and building height standard deviation, $bhdev$); and (2) specific aspect ratio (a_r) for each street segment consisting of street-averaged building height divided by street width. The geometrical parameters are calculated from a Barcelona City Council dataset containing 2-D geometries and number of floors for each building (Barcelona City Council, 2016), assuming 3 m height for each floor.

To estimate specific z_0 for each street segment we propose a new morphometric method inspired by previous studies in the literature. z_0 is composed by the WRF's background roughness (z_{0bg}) and the one estimated locally (Eq. 1), which incorporates building height influence through the $range$ parameter scaled by two parabolic ratios based on aspect ratio (a_{rr}) and building density (bd_r). The $range$ parameter (Eq. 2) and z_0 increase with $\overline{bh}$ following most morphometric methods (e.g. Macdonald et al., 1998). In addition, $range$ and z_0 increase with an increasing $bhdev$. This assumption is based on Kent et al. (2017),

ground level sources and receptors. In order to estimate NO_2 concentrations, R-LINE incorporates a chemistry module to resolve simple NO to NO_2 chemistry with the Generic Reaction Set (GRS; Valencia et al., 2018) considering the chemical reactions in Table 1. GRS chemistry mechanism solves the photochemistry of NO_2 assuming clear-sky conditions. Thus, it does not consider cloud effects on the NO_2 photolysis rate, representing this one of its major limitations. R-LINE has been applied to estimate exposure to traffic-related air pollutants in a large scale study in Detroit, United States (Isakov et al., 2014). However, to our knowledge it has not been applied to European cities, where street canyon morphology dominates. Hence, in order to apply R-LINE over Barcelona its meteorology has been adapted to street canyons as described in Sect. 2.3.1 and the background concentrations are obtained from CMAQ model considering local meteorology and urban geometry as described in Sect. 2.3.3.

Table 1. Chemical reactions in the Generic Reaction Set (GRS). ROC are Reactive Organic Compounds, RP is meant for Radical Pool, SGN are Stable Gaseous Nitrogen products, and SNGN are Stable Non-Gaseous Nitrogen products.

$ROC + h\upsilon \rightarrow RP + ROC$
$RP + NO \rightarrow NO_2$
$NO_2 + h\upsilon \rightarrow NO + O_3$
$NO + O_3 \rightarrow NO_2$
$RP + RP \rightarrow RP$
$RP + NO_2 \rightarrow SGN$
$RP + NO_2 \rightarrow SNGN$

2.3 Coupling CALIOPE with R-LINE

CALIOPE and R-LINE are coupled offline, first CALIOPE is run over Europe, Iberian Peninsula and Catalonia and then R-LINE is executed for Barcelona city. This approach addresses two main challenges that have already been highlighted in the research literature: (1) downscaling regional meteorology to street scale to drive pollutant dispersion; and (2) obtaining background concentrations from the mesoscale model without double counting traffic emissions in regional and street scale models. In addition to these challenges, we consider relevant to couple meteorology and background concentrations in a consistent way, taking into account atmospheric stability and urban geometry when estimating background contribution within urban streets. Here we describe our methodology when coupling the models to mitigate these challenges.

2.3.1 Meteorology

Most buildings in Barcelona have lower heights than the WRF bottom layer (40.6 m depth). WRF results are assumed to represent over roof wind and stability conditions because its mid-point height (20.3 m) is similar to average building height ($\overline{bh}$) in a typical neighbourhood of Barcelona (e.g. Eixample district; 20.7 m). WRF is executed consistently with the forecasting air quality system CALIOPE, giving a constant surface roughness (z_0) equal to 1 m over the urban area. In order to apply R-LINE over Barcelona, its meteorology has been adapted to street canyons. We have developed a methodology to estimate

who compared nine methods to estimate z_0 concluding that methods considering height variability through $bhdev$ (i.e. a higher $bhdev$ brings an increase of z_0) provide better results (e.g. Kanda et al., 2013). The parameter C multiplying the equation for $range$ calculation is an empirical constant set to 1/20 after calibrating the system with the NO_2 measurements used in this work for CALIOPE-Urban evaluation. $dispht$ is calculated following R-LINE methodology given a factor of displacement height ($facdispht$) equal to 5 (Eq. 3) as suggested by Snyder and Heist (2013).

$$z_0 = a_{rr} \cdot bd_r \cdot range + z_{0bg} \quad (1)$$

$$range = C \cdot (\overline{bh} + bhdev) \quad (2)$$

$$dispht = facdispht \cdot z_0 \quad (3)$$

To model the influence of building density and aspect ratio, we use Oke (1988) finding based on wind tunnel and experimental studies. Oke concluded that over-roof air roughness and satisfactory dispersion within the street canyon are maximum under similar geometrical conditions. Specifically, showing that an a_r equals 0.65 and a bd equals 0.25 give maximum roughness for overlying air and optimal dispersion conditions in the street canyon.

In practice, z_0 increases with an increasing a_r to a maximum of a_r = 0.65 and decreases for a_r > 0.65 (Eq. 4). Additionally, an increasing bd produces higher z_0 until a maximum at $bd = 0.25$ and decreases for higher bd (Eq. 5). We model these ratios using parabolic shapes ranging from 0 to 1. Both urban characteristics are modelled using one parabola to the left of the maximum and another to the right due to the non symmetrical distribution of the parameter values within Barcelona city (see Fig. A1 in Appendix A). The parabolic ratios will be maximum (i.e. equal to 1) if the roughness effect is maximum. The ratios are prevented from having negative values by setting a minimum of 0.

$$a_{rr} = \begin{cases} 1.0 - 2.3 \cdot (a_r - 0.65)^2 & \text{if } ar \text{ is} <= 0.65 \\ max(0, 1.0 - 1.38 \cdot (a_r - 0.65)^2) & \text{if } ar \text{ is} > 0.65 \end{cases} \quad (4)$$

$$bd_r = \begin{cases} 1.0 - 16.0 \cdot (bd - 0.25)^2 & \text{if } bd \text{ is} <= 0.25 \\ max(0, 1.0 - 8.1 \cdot (bd - 0.25)^2) & \text{if } bd \text{ is} > 0.25 \end{cases} \quad (5)$$

In addition to the z_0 adjustment, we adjust the wind speed and direction to represent more closely the winds blowing down the street as constrained by the buildings, which is called "channelling" (similarly to Fisher et al., 2005). We have adapted R-LINE to incorporate the orientation of roadways (and thus the buildings) where the wind direction follows the street direction. This leads to a recalculation of the wind direction and speed for each roadway before emissions are dispersed within a city. Wind speed channelling is parametrized following Soulhac et al. (2008) who showed that mean velocity along a canyon for

specific z_0 based on urban geometry (e.g. building height, street width). Once z_0 is adjusted, the displacement height ($dispht$), friction velocity ($u*$), convective velocity scale ($w*$), PBL height, and Monin-Obukov length (L) are re-calculated (Cimorelli et al., 2005). The increase in z_0 generally leads to a larger $dispht$, $u*$, $w*$, and PBL height. Therefore, L is less stable and atmospheric conditions are more convective. Ultimately, these adjustments have an effect on the way the winds are profiled and on the rate of dispersion of the roadway emissions within the urban area.

The geometrical parameters used for z_0 calculation are divided into two categories: (1) averaged over an area of 250m × 250m (planar building density, bd; average building height, $\overline{bh}$; and building height standard deviation, $bhdev$); and (2) specific aspect ratio (a_r) for each street segment consisting of street-averaged building height divided by street width. The geometrical parameters are calculated from a Barcelona City Council dataset containing 2-D geometries and number of floors for each building (Barcelona City Council, 2016), assuming 3 m height for each floor.

To estimate specific z_0 for each street segment we propose a new morphometric method inspired by previous studies in the literature. z_0 is composed by the WRF's background roughness (z_{0bg}) and the one estimated locally (Eq. 1), which incorporates building height influence through the $range$ parameter scaled by two parabolic ratios based on aspect ratio (a_{rr}) and building density (bd_r). The $range$ parameter (Eq. 2) and z_0 increase with $\overline{bh}$ following most morphometric methods (e.g. Macdonald et al., 1998). In addition, $range$ and z_0 increase with an increasing $bhdev$. This assumption is based on Kent et al. (2017), who compared nine methods to estimate z_0 concluding that methods considering height variability through $bhdev$ (i.e. a higher $bhdev$ brings an increase of z_0) provide better results (e.g. Kanda et al., 2013). The parameter C multiplying the equation for $range$ calculation is an empirical constant set to 1/20 after calibrating the system with the NO_2 measurements used in this work for CALIOPE-Urban evaluation. $dispht$ is calculated following R-LINE methodology given a factor of displacement height ($facdispht$) equal to 5 (Eq. 3) as suggested by Snyder and Heist (2013).

$$z_0 = a_{rr} \cdot bd_r \cdot range + z_{0bg} \quad (1)$$

$$range = C \cdot (\overline{bh} + bhdev) \quad (2)$$

$$dispht = facdispht \cdot z_0 \quad (3)$$

To model the influence of building density and aspect ratio, we use Oke (1988) finding based on wind tunnel and experimental studies. Oke concluded that over-roof air roughness and satisfactory dispersion within the street canyon are maximum under similar geometrical conditions. Specifically, showing that an a_r equals 0.65 and a bd equals 0.25 give maximum roughness for overlying air and optimal dispersion conditions in the street canyon.

In practice, z_0 increases with an increasing a_r to a maximum of a_r = 0.65 and decreases for a_r > 0.65 (Eq. 4). Additionally, an increasing bd produces higher z_0 until a maximum at $bd = 0.25$ and decreases for higher bd (Eq. 5). We model these ratios using parabolic shapes ranging from 0 to 1. Both urban characteristics are modelled using one parabola to the left of the maximum and another to the right due to the non symmetrical distribution of the parameter values within Barcelona city (see

any wind direction is directly proportional to the cosine of the angle between street direction and over roof wind direction (i.e. angle of incidence).

$$ws_{ch} = ws_{bh} \cdot max(0.1, abs(cos(\theta))) \quad (6)$$

where ws_{ch} means channelled wind speed at roof level, the wind speed at roof level (ws_{bh}) is taken from the WRF bottom layer in m/s and θ is the angle of incidence. The minimum value of the right component is set to avoid an unrealistic zero value for wind speed. Its value of 0.1 is defined in line with Kastner-Klein et al. (2001), who showed that minimum longitudinal mean flow velocity component at canyon top is equivalent to 0.12 times the above canyon wind speed for perpendicular over roof winds according to their wind tunnel experiments. Then, to estimate wind speed at street level a logarithmic profile incorporated within R-LINE that is based on similarity theory (Monin and Obukhov, 1954) is used.

2.3.2 Emissions

HERMESv2.0 provides hourly NO_X and NMVOCs road transport emissions at the road link level, which are used by the R-LINE model algorithms to account for NO_2 near-road chemistry (Valencia et al., 2018). Road transport emissions (i.e. exhaust, evaporative, wear and resuspension) are estimated combining the Tier 3 method described in the EMEP/EEA air pollutant emission inventory guidebook (fully incorporated in version 5.1 of the COPERT IV software) with a digitized traffic network that contains specific information by road stretch for daily average traffic, mean speed circulation, temporal profiles and vehicular park profiles. We note that HERMESv2.0 uses COPERT IV, which does not incorporate revised emission factors of NO_X related to diesel gate. Hence, NO_X emissions from traffic may be underestimated. Input activity data is obtained by combining different datasets, including traffic data from the Barcelona Automatic Traffic Counting Equipment and vehicle composition profiles derived from a Remote Sensing Campaign performed in different areas of Barcelona during 2010 (Barcelona City Council, 2010). In Barcelona, higher levels of traffic emissions are found in the city center and in the highways surrounding the city (Fig. 2). In order to produce emissions in $gm^{-1}s^{-1}$ for straight street segments as required by R-LINE, we converted the digitized road network curved segments in HERMES to straight segments with no intermediate vertices using the Douglas-Peucker algorithm in the QGIS simplify geometries tool (QGIS Development Team, 2017).

Fig. A1 in Appendix A). The parabolic ratios will be maximum (i.e. equal to 1) if the roughness effect is maximum. The ratios are prevented from having negative values by setting a minimum of 0.

$$a_{rr} = \begin{cases} 1.0 - 2.3 \cdot (a_r - 0.65)^2 & \text{if } ar \text{ is} <= 0.65 \\ max(0, 1.0 - 1.38 \cdot (a_r - 0.65)^2) & \text{if } ar \text{ is} > 0.65 \end{cases} \quad (4)$$

$$bd_r = \begin{cases} 1.0 - 16.0 \cdot (bd - 0.25)^2 & \text{if } bd \text{ is} <= 0.25 \\ max(0, 1.0 - 8.1 \cdot (bd - 0.25)^2) & \text{if } bd \text{ is} > 0.25 \end{cases} \quad (5)$$

In addition to the z_0 adjustment, we adjust the wind speed and direction to represent more closely the winds blowing down the street as constrained by the buildings, which is called "channelling" (similarly to Fisher et al., 2005). We have adapted R-LINE to incorporate the orientation of roadways (and thus the buildings) where the wind direction follows the street direction. This leads to a recalculation of the wind direction and speed for each roadway before emissions are dispersed within a city. Wind speed channelling is parametrized following Soulhac et al. (2008) who showed that mean velocity along a canyon for any wind direction is directly proportional to the cosine of the angle between street direction and over roof wind direction (i.e. angle of incidence).

$$ws_{ch} = ws_{bh} \cdot max(0.1, abs(cos(\theta))) \quad (6)$$

where ws_{ch} means channelled wind speed at roof level, the wind speed at roof level (ws_{bh}) is taken from the WRF bottom layer in m/s and θ is the angle of incidence. The minimum value of the right component is set to avoid an unrealistic zero value for wind speed. Its value of 0.1 is defined in line with Kastner-Klein et al. (2001), who showed that minimum longitudinal mean flow velocity component at canyon top is equivalent to 0.12 times the above canyon wind speed for perpendicular over roof winds according to their wind tunnel experiments. Then, to estimate wind speed at street level a logarithmic profile incorporated within R-LINE that is based on similarity theory (Monin and Obukhov, 1954) is used. In this work, we assume that recirculation flows within street canyons are negligible because R-LINE computes concentrations averaged over an hour, when recirculation and vehicle induced turbulence are assumed to contribute to a well mixed more homogeneous air mass driven by variable wind conditions. Additionally, evaluation of the potential impact of including recirculating flows across the canyon is not possible without multiple simultaneous meteorological and pollutant measurements at a fine temporal scale, which are currently not available for the Barcelona city.

2.3.2 Emissions

HERMESv2.0 provides hourly NO_X and NMVOCs road transport emissions at the road link level, which are used by the R-LINE model algorithms to account for NO_2 near-road chemistry (Valencia et al., 2018). Road transport emissions (i.e. exhaust,

Figure 2. NO_X emissions in $gm^{-1}h^{-1}$ in Barcelona city at 7 UTC on 11/4/2013 and location of the two fixed monitoring stations used to estimate NO_2/NO_X ratio. White circles with letters inside represent the stations: *tr* is Eixample traffic station and *bg* is Ciutadella Park urban background station.

We have estimated NO_2/NO_X ratio following Carslaw and Beevers (2004), which produces an approximation to the NO_2 primary contribution. This method relates total O_X (NO_2 + O_3) to total NO_X (NO_2 + NO) in a traffic monitoring station subtracting O_X and NO_X from a background site in order to remove the effect of background and to only calculate the contribution at the traffic site. As the traffic station we used Eixample site and as the urban background station Ciutadella Park (see Fig. 2), which is located upwind of the dominant wind direction. Figure 3 compares O_X to NO_X in Eixample after subtracting the background represented by Ciutadella from the beginning of October to end of February for years 2012 to 2016. The photochemical season (April-September) is not used to avoid greater scatter than it is found in the winter months as shown by Clapp and Jenkin (2001). The O_X slope value of 18.9% is considered an estimate of the potential primary NO_2 contribution from vehicles on Eixample traffic station. This value is consistent with studies conducted in other cities with high diesel vehicle fleet (e.g., Carslaw et al., 2016; Wild et al., 2017) and is assumed to represent the NO_2/NO_X ratio in Barcelona in the present work.

evaporative, wear and resuspension) are estimated combining the Tier 3 method described in the EMEP/EEA air pollutant emission inventory guidebook (fully incorporated in version 5.1 of the COPERT IV software) with a digitized traffic network that contains specific information by road stretch for daily average traffic, mean speed circulation, temporal profiles and vehicular park profiles. We note that HERMESv2.0 uses COPERT IV, which does not incorporate revised emission factors of NO_X related to diesel gate. Hence, NO_X emissions from traffic may be underestimated. Input activity data is obtained by combining different datasets, including traffic data from the Barcelona Automatic Traffic Counting Equipment and vehicle composition profiles derived from a Remote Sensing Campaign performed in different areas of Barcelona during 2010 (Barcelona City Council, 2010). In Barcelona, higher levels of traffic emissions are found in the city center and in the highways surrounding the city (Fig. 2). In order to produce emissions in $gm^{-1}s^{-1}$ for straight street segments as required by R-LINE, we converted the digitized road network curved segments in HERMES to straight segments with no intermediate vertices using the Douglas-Peucker algorithm in the QGIS simplify geometries tool (QGIS Development Team, 2017).

Figure 2. NO_X emissions in $gm^{-1}h^{-1}$ in Barcelona city at 7 UTC on 11/4/2013 and location of the two fixed monitoring stations used to estimate NO_2/NO_X ratio. White circles with letters inside represent the stations: *tr* is Eixample traffic station and *bg* is Ciutadella Park urban background station.

We have estimated NO_2/NO_X ratio following Carslaw and Beevers (2004), which produces an approximation to the NO_2 primary contribution. This method relates total O_X (NO_2 + O_3) to total NO_X (NO_2 + NO) in a traffic monitoring station

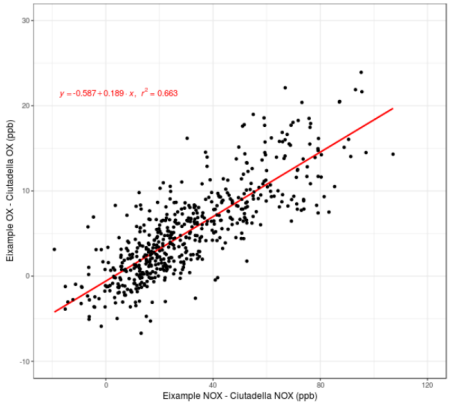

Figure 3. Scatter plot showing daylight mean O_X and NO_X relation of the difference between Eixample and Ciutadella stations from begin of October to end of February for years 2012 to 2016.

2.3.3 Background concentrations

We use the upwind Urban Background Scheme (UBS) to avoid the double counting of traffic emissions when coupling the mesoscale with the street scale model. The UBS makes a selective choice of CMAQ cells as sketched in Fig. 4 to estimate over roof background concentrations. For each hour, a polygon covering upwind air masses (white) is created. In the figure, the average distance traversed by air masses during an hour (10.8 km) is estimated for WRF's bottom layer wind speed (3 m/s in the image). Squares falling within the scheme polygon represent CMAQ cells and their color refer to cell pollutant values (e.g. NO_2 at peak traffic hours may be higher within the city than over the Mediterranean sea). Grid cell values falling over the scheme polygon are inverse-distance averaged to produce the background estimate of the scheme. Under calm conditions, only the upwind cell is chosen. This method is inspired by Berkowicz (2000) who apply a similar concept based on air masses trajectory to develop a background model.

subtracting O_X and NO_X from a background site in order to remove the effect of background and to only calculate the contribution at the traffic site. As the traffic station we used Eixample site and as the urban background station Ciutadella Park (see Fig. 2), which is located upwind of the dominant wind direction. Figure 3 compares O_X to NO_X in Eixample after subtracting the background represented by Ciutadella from the beginning of October to end of February for years 2012 to 2016. The photochemical season (April-September) is not used to avoid greater scatter than it is found in the winter months as shown by Clapp and Jenkin (2001). The O_X slope value of 18.9% is considered an estimate of the potential primary NO_2 contribution from vehicles on Eixample traffic station. This value is consistent with studies conducted in other cities with high diesel vehicle fleet (e.g., Carslaw et al., 2016; Wild et al., 2017) and is assumed to represent the NO_2/NO_X ratio in Barcelona in the present work.

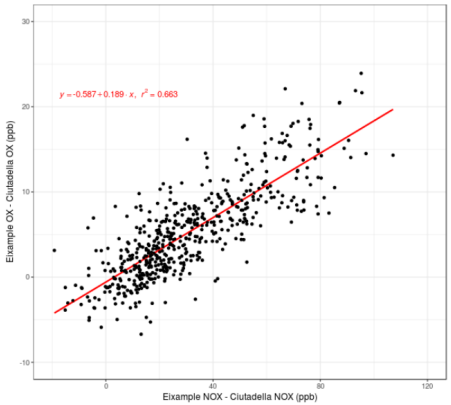

Figure 3. Scatter plot showing daylight mean O_X and NO_X relation of the difference between Eixample and Ciutadella stations from begin of October to end of February for years 2012 to 2016.

2.3.3 Background concentrations

We use the upwind Urban Background Scheme (UBS) to avoid the double counting of traffic emissions when coupling the mesoscale with the street scale model. The UBS makes a selective choice of CMAQ cells as sketched in Fig. 4 to estimate over roof background concentrations. For each hour, a polygon covering upwind air masses (white) is created. In the figure, the average distance traversed by air masses during an hour (10.8 km) is estimated for WRF's bottom layer wind speed (3 m/s in the image). Squares falling within the scheme polygon represent CMAQ cells and their color refer to cell pollutant values (e.g. NO_2 at peak traffic hours may be higher within the city than over the Mediterranean sea). Grid cell values falling over the scheme polygon are inverse-distance averaged to produce the background estimate of the scheme. Under calm conditions,

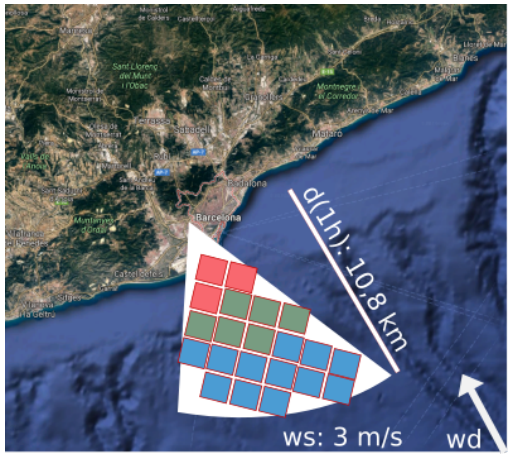

Figure 4. Upwind urban background scheme concept.

Background concentrations are required at each receptor in CALIOPE-Urban. Urban dispersion models are typically run at a very high spatial resolution (e.g. 20 m x 20 m). Running the UBS every 20 meters would have a high computational cost due to its spatial computations and background concentration values are not expected to vary substantially over tens of meters because CMAQ produces results with 1 km $\times$ 1 km spatial resolution. Hence, we first run the UBS to produce background concentration values at CMAQ grid cell centroids, then we apply a bilinear interpolation method to provide background at very high spatial resolution.

In addition to the UBS we implement a background decay method to calculate the surface level background concentrations assuming that the UBS provides the concentration at rooftop level. The relationship between rooftop and surface level concentrations is assumed to depend on atmospheric stability, localized surface roughness and urban geometry. The ratio of wind speeds at surface and rooftop levels (ws_{sfc}/ws_{bh}) estimated by R-LINE using similarity theory (Monin and Obukhov, 1954) is used as a proxy for the vertical mixing. Using this ratio, we calculate fac_{bg} that represents the adimensional vertical mixing variable that is multiplied to rooftop background concentration to obtain surface level background concentration at a given height. In order to diminish the effect of afternoon underestimations from the regional system near traffic, background levels under convective situations are enhanced. We consider the upward heat flux at the surface ($hflux$) as representing convective conditions for values higher than 0.30. This value is set to exclude slightly stable night hours with low positive $hflux$

only the upwind cell is chosen. This method is inspired by Berkowicz (2000) who apply a similar concept based on air masses trajectory to develop a background model.

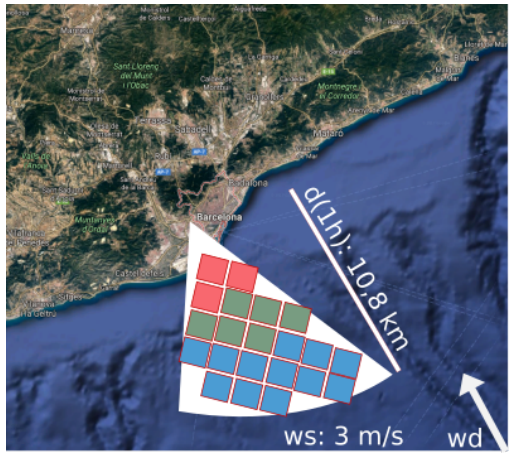

Figure 4. Upwind urban background scheme concept.

Background concentrations are required at each receptor in CALIOPE-Urban. Urban dispersion models are typically run at a very high spatial resolution (e.g. 20 m x 20 m). Running the UBS every 20 meters would have a high computational cost due to its spatial computations and background concentration values are not expected to vary substantially over tens of meters because CMAQ produces results with 1 km $\times$ 1 km spatial resolution. Hence, we first run the UBS to produce background concentration values at CMAQ grid cell centroids, then we apply a bilinear interpolation method to provide background at very high spatial resolution.

In addition to the UBS we implement a background decay method to calculate the surface level background concentrations assuming that the UBS provides the concentration at rooftop level. To calculate street-level NO_2 concentrations, the vertical distribution of pollutants are solved first using the background decay method, applied uniformly to all pollutants, and then the GRS chemical mechanism is solved. The relationship between rooftop and surface level concentrations is assumed to depend on atmospheric stability, localized surface roughness and urban geometry (see Figure 5 as an illustration of the background decay concept). In the research literature, the influence of atmospheric stability on vertical mixing within a street canyon has been demonstrated using experimental measurements (Rotach, 1995), wind tunnel experiments (Salizzoni et al., 2009), and it has been implemented in some dispersion models (e.g., Soulhac et al., 2011; Kim et al., 2018). The ratio of wind speeds at surface and rooftop levels (ws_{sfc}/ws_{bh}) estimated by R-LINE using similarity theory (Monin and Obukhov, 1954) is used as a

values mainly caused by the urban heat island (i.e. Barcelona city has been found to be 2.9 °C warmer than its periphery by Moreno-Garcia, 1994). The following parametrization is used for cases with bd higher than 0.1,

$$fac_{bg} = \begin{cases} 1 - F + ws_{sfc}/ws_{bh} \cdot F & \text{if } hflux \text{ is} > 0.30 \\ ws_{sfc}/ws_{bh} & \text{if } hflux \text{ is} <= 0.30 \end{cases} \quad (7)$$

where $F = m + abs(0.25 - bd)$, being m an empirical parameter set to 0.35 after system calibration with NO_2 measurements; $hflux$ is upward heat flux at the surface (W m^{-2}). Surface background concentrations for convective situations are maximum for bd equal 0.25 consistently with z_0 estimation in Sect. 2.3.1. On the other hand, we assume that for bd close to zero, surface background concentrations tend linearly to rooftop level background concentrations. This linear transition starts when bd equals 0.1 and ends when the surface background gets over roof value for bd equals 0. The threshold bd = 0.1 is based on Grimmond and Oke (1999), who set it as an inferior limit for real cities and show that below this value an isolated flow regime governs. Within this regime, street level and over roof air is well mixed due to the low building density. Hence, for cases with bd equal or lower than 0.1, fac_{bg} tends linearly to 1 following,

$$fac_{bg} = \begin{cases} 1 - 5 \cdot bd + ws_{sfc}/ws_{bh} \cdot (5 \cdot bd) & \text{if } hflux \text{ is} > 0.30 \\ 1 - 10 \cdot bd + ws_{sfc}/ws_{bh} \cdot (10 \cdot bd) & \text{if } hflux \text{ is} <= 0.30 \end{cases} \quad (8)$$

Equations 8 are linear variations between the point at bd = 0 and fac_{bg} = 1, and the point at bd = 0.1 with the corresponding fac_{bg} value from the Eq. 7.

2.4 Execution setup

We have run CALIOPE-Urban for receptors as far as 250 metres from roads with sufficient Annual Average Daily Traffic (AADT) (i.e. 2000 vehicles/day following Jensen et al. (2017)) and receptors further away receive directly CMAQ values interpolated. The 250 m limit is chosen as similar but less restrictive (i.e. to allow longer distances under stable hours) than the one used in Beevers et al. (2012) who used 225 m for London. To smooth out the variation between system outputs, we define a transition area (i.e. 140 m to 250 m) where receptors are given concentration values weighted by distance. For temporal and spatial evaluation runs, we locate receptors at the specific coordinates of the measurement sites.

To obtain high resolution concentration maps for the entire city, we set the spatial context as the minimum rectangle where Barcelona municipality is contained and extended it by 250 m buffers that include the highways surrounding the city. The context is covered by a regular receptor grid of 10 meter resolution. R-LINE execution loops over each hour, road and receptor to estimate the contribution from each source to each receptor.

Aiming to understand the impact on accuracy of the local parametrization for background and meteorology and the impact of using the analytical approach for dispersion, we have run CALIOPE-Urban with different configurations. In Table 1, we describe the different scenarios that have been run. As seen in the table, the CALIOPE-Urban and the CALIOPE-Urban

proxy for the vertical mixing. Using this ratio, we calculate fac_{bg} that represents the adimensional vertical mixing variable that is multiplied to rooftop background concentration to obtain surface level background concentration at a given height. Wind channelling does not affect the ratio ws_{sfc}/ws_{bh} because we assume that channelling affects equally winds at surface and rooftop level.

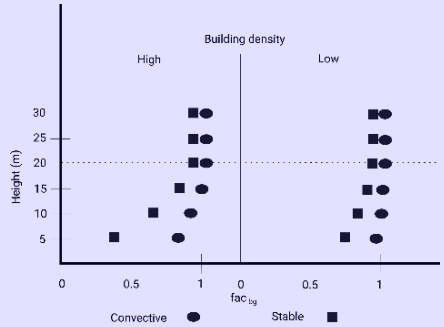

Figure 5. Illustration of the background decay method concept. Building height is approx. 20 m.

In order to diminish the effect of afternoon underestimations from the regional system near traffic, background levels under convective situations are enhanced. We consider the upward heat flux at the surface ($hflux$) as representing convective conditions for values higher than 0.30. This value is set to exclude slightly stable night hours with low positive $hflux$ values mainly caused by the urban heat island (i.e. Barcelona city has been found to be 2.9 °C warmer than its periphery by Moreno-Garcia, 1994). The following parametrization is used for cases with bd higher than 0.1,

$$fac_{bg} = \begin{cases} 1 - F + ws_{sfc}/ws_{bh} \cdot F & \text{if } hflux \text{ is} > 0.30 \\ ws_{sfc}/ws_{bh} & \text{if } hflux \text{ is} <= 0.30 \end{cases} \quad (7)$$

where $F = m + abs(0.25 - bd)$, being m an empirical parameter set to 0.35 after system calibration with NO_2 measurements; $hflux$ is upward heat flux at the surface (W m^{-2}). Surface background concentrations for convective situations are maximum for bd equal 0.25 consistently with z_0 estimation in Sect. 2.3.1. On the other hand, we assume that for bd close to zero, surface background concentrations tend linearly to rooftop level background concentrations. The threshold bd = 0.1 is based on Grimmond and Oke (1999), who set it as an inferior limit for real cities and show that below this value an isolated flow regime

governs. Within this regime, street level and over roof air is well mixed due to the low building density. Hence, for cases with bd equal or lower than 0.1, fac_{bg} tends linearly to 1 following,

$$fac_{bg} = \begin{cases} 1 - 5 \cdot bd + ws_{sfc}/ws_{bh} \cdot (5 \cdot bd) & \text{if } hflux \text{ is} > 0.30 \\ 1 - 10 \cdot bd + ws_{sfc}/ws_{bh} \cdot (10 \cdot bd) & \text{if } hflux \text{ is} <= 0.30 \end{cases} \quad (8)$$

Equations 8 are linear variations between the point at bd = 0 and fac_{bg} = 1, and the point at bd = 0.1 with the corresponding fac_{bg} value from the Eq. 7.

2.4 Execution setup

We have run CALIOPE-Urban for receptors as far as 250 metres from roads with sufficient Annual Average Daily Traffic (AADT) (i.e. 2000 vehicles/day following Jensen et al. (2017)) and receptors further away receive directly CMAQ values interpolated. The 250 m limit is chosen as similar but less restrictive (i.e. to allow longer distances under stable hours) than the one used in Beevers et al. (2012) who used 225 m for London. To smooth out the variation between system outputs, we define a transition area (i.e. 140 m to 250 m) where receptors are given concentration values weighted by distance. For temporal and spatial evaluation runs, we locate receptors at the specific coordinates of the measurement sites.

To obtain high resolution concentration maps for the entire city, we set the spatial context as the minimum rectangle where Barcelona municipality is contained and extended it by 250 m buffers that include the highways surrounding the city. The context is covered by a regular receptor grid of 10 meter resolution. R-LINE execution loops over each hour, road and receptor to estimate the contribution from each source to each receptor.

Aiming to understand the impact on accuracy of the local parametrization for background and meteorology and the impact of using the analytical approach for dispersion, we have run CALIOPE-Urban with different configurations. In Table 2, we describe the different scenarios that have been run. As seen in the table, the CALIOPE-Urban and the CALIOPE-Urban Analytical configurations make use of the developed local parametrizations for background and meteorology. In contrast, the CALIOPE-Urban-nl (Non Local) configuration does not apply the local parametrizations for background and meteorology. Instead, it uses as background the UBS output without vertical mixing and it omits the use of wind channelling and specific stability parameters for each street segment based on local z_0. We show this configuration's results in order to understand if the new implementations in this work contribute substantially to improve the system's ability to simulate NO_2 concentrations in Barcelona. R-LINE dispersion algorithm options (i.e. analytical and numerical) are described in Sect. 2.2. For meteorological options, we refer to Sect. 2.3.1. The background method is described in Sect. 2.3.3.

Analytical configurations make use of the developed local parametrizations for background and meteorology. In contrast, the CALIOPE-Urban-nl (Non Local) configuration does not apply the local parametrizations for background and meteorology. Instead, it uses as background the UBS output without vertical mixing and it omits the use of wind channelling and specific stability parameters for each street segment based on local z_0. We show this configuration's results in order to understand if the new implementations in this work contribute substantially to improve the system's ability to simulate NO_2 concentrations in Barcelona. R-LINE dispersion algorithm options (i.e. analytical and numerical) are described in Sect. 2.2. For meteorological options, we refer to Sect. 2.3.1. The background method is described in Sect. 2.3.3.

Table 1. Description of the execution setup. Execution time is for the entire city of Barcelona during one hour running CALIOPE-Urban (i.e. only the urban system, after CALIOPE run completion) over 11251 street segments and 965458 receptors at 10 m $\times$ 10 m spatial resolution.

Configuration Name	Dispersion algorithm	Meteorology	Background	Execution time
CALIOPE-Urban	Numerical	Local	Local	88 minutes
CALIOPE-Urban-nl	Numerical	Non Local	Non Local	56 minutes
CALIOPE-Urban Analytical	Analytical	Local	Local	44 minutes

3 Observational datasets

We use three datasets of observations to evaluate the performance of CALIOPE-Urban to reproduce the temporal and spatial variation of NO_2 concentrations within Barcelona city. Fig. 5 shows the locations of measurements used in this study, which are described below.

Table 2. Description of the execution setup. Execution time is for the entire city of Barcelona during one hour running CALIOPE-Urban (i.e. only the urban system, after CALIOPE run completion) over 11251 street segments and 965458 receptors at 10 m $\times$ 10 m spatial resolution.

Configuration Name	Dispersion algorithm	Meteorology	Background	Execution time
CALIOPE-Urban	Numerical	Local	Local	88 minutes
CALIOPE-Urban-nl	Numerical	Non Local	Non Local	56 minutes
CALIOPE-Urban Analytical	Analytical	Local	Local	44 minutes

3 Observational datasets

We use three datasets of observations to evaluate the performance of CALIOPE-Urban to reproduce the temporal and spatial variation of NO_2 concentrations within Barcelona city. Fig. 6 shows the locations of measurements used in this study, which are described below.

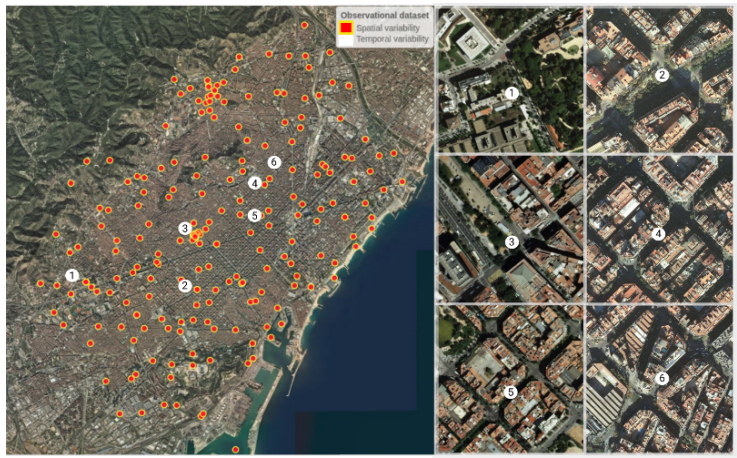

Figure 6. Passive dosimeters and monitoring sites location used in the evaluation of CALIOPE-Urban in this work. Red dots with yellow border represent passive dosimeters (spatial performance) location and white numbered dots depict monitoring site emplacements (temporal variability). White dots numbered 1 (Palau Reial), 2 (Eixample) and 3 (Gràcia-Sant Gervasi) are air quality monitoring sites and 4 (213 Industria Street), 5 (445 Valencia Street) and 6 (309 Industria Street) correspond to mobile units.

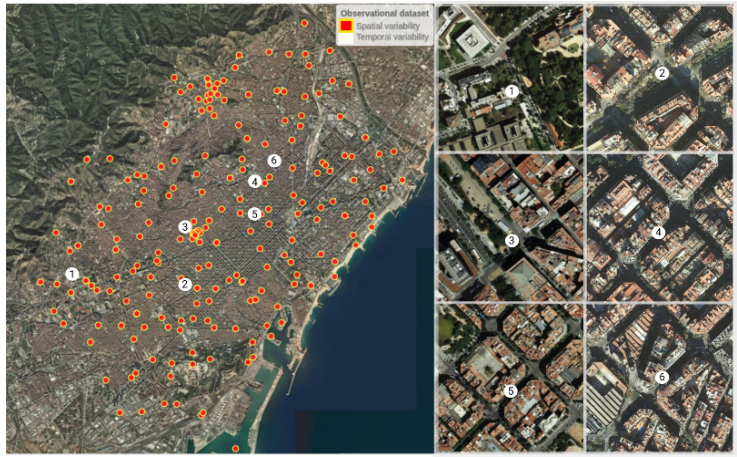

Figure 5. Passive dosimeters and monitoring sites location used in the evaluation of CALIOPE-Urban in this work. Red dots with yellow border represent passive dosimeters (spatial performance) location and white numbered dots depict monitoring site emplacements (temporal variability). White dots numbered 1 (Palau Reial), 2 (Eixample) and 3 (Gràcia-Sant Gervasi) are air quality monitoring sites and 4 (213 Industria Street), 5 (445 Valencia Street) and 6 (309 Industria Street) correspond to mobile units.

3.1 NO_2 temporal variability: Street canyon campaign and permanent XVPCA network

To evaluate the NO_2 temporal variability we use hourly NO_2 concentrations reported by the official monitoring network in Catalunya (XVPCA) and from an experimental campaign conducted using mobile units in April and May 2013 in Barcelona (Amato et al., 2014). The official monitoring network has 10 stations in Barcelona and only two of them (i.e. Gràcia-Sant Gervasi and Eixample) are considered representative of near traffic conditions and provide NO_2 hourly levels. Measured data from 3 sites of the official network are used in this study: Eixample and Gràcia-Sant Gervasi (traffic) and Palau Reial (background). Both traffic sites are located in complex wide areas where several streets intersect (see sites 2 and 3 in Fig. 5 and in description Table 2). Palau Reial station (i.e. site 1 in Fig. 5) is located in a medium *bd* area of the city, 300 metres away from a heavily trafficked street. This dataset is complemented with observations from a experimental campaign where mobile units placed at the parking lane of several street segments measured air quality parameters at 3 m height. For this study, we used data gathered every 30 minutes and aggregated to hourly levels for homogeneity at 213 Industria Street, 309 Industria Street and 445 Valencia Street. These streets present a marked canyon pattern (see sites 4, 5 and 6 in Fig. 5 and description table) where aspect ratio is approximately 1. In Barcelona, different street geometrical patterns cohabit. For example, the Eixample district, which has the highest number of inhabitants and the greatest population density (33000 inhabitants km^{-2}), is characterized by

3.1 NO_2 temporal variability: Street canyon campaign and permanent XVPCA network

To evaluate the NO_2 temporal variability we use hourly NO_2 concentrations reported by the official monitoring network in Catalunya (XVPCA) and from an experimental campaign conducted using mobile units in April and May 2013 in Barcelona (Amato et al., 2014). The official monitoring network has 10 stations in Barcelona and only two of them (i.e. Gràcia-Sant Gervasi and Eixample) are considered representative of near traffic conditions and provide NO_2 hourly levels. Measured data from 3 sites of the official network are used in this study: Eixample and Gràcia-Sant Gervasi (traffic) and Palau Reial (background). Both traffic sites are located in complex wide areas where several streets intersect (see sites 2 and 3 in Fig. 6 and in description Table 3). Palau Reial station (i.e. site 1 in Fig. 6) is located in a medium *bd* area of the city, 300 metres away from a heavily trafficked street. This dataset is complemented with observations from a experimental campaign where mobile units placed at the parking lane of several street segments measured air quality parameters at 3 m height. For this study, we used data gathered every 30 minutes and aggregated to hourly levels for homogeneity at 213 Industria Street, 309 Industria Street and 445 Valencia Street. These streets present a marked canyon pattern (see sites 4, 5 and 6 in Fig. 6 and description table) where aspect ratio is approximately 1. In Barcelona, different street geometrical patterns cohabit. For example, the Eixample district, which has the highest number of inhabitants and the greatest population density (33000 inhabitants km^{-2}), is characterized by

a marked street canyon pattern. Most of its canyons are about 20 to 25 m high and 20 m wide (i.e. a_r=1 and higher than 1). Experimental campaign sites are considered traffic sites in this work because they are exposed to similar AADT and vehicles km^{-2} compared to official traffic sites as shown in the table below. We apply Eq. (9) to obtain vehicles km^{-2}, a variable that describes traffic density in an area of 1 km^2.

$$vehicles \cdot km^{-2} = \sum_{n=1}^{st} Vehicles/s \cdot length \tag{9}$$

To obtain the amount of vehicles per second, AADT is divided by 3600 * 24 and multiplied by a temporal factor (i.e. 1.47) representing a typical factor for morning traffic peak in Barcelona. $Length$ is street length in metres. st is the number of streets over the circular area of 1 km^2 centred on the measurement site.

Table 3. Morphometric and traffic description of measurement sites used in CALIOPE-Urban evaluation. The measurement height of the official network sites and the mobile sites is 3 meters. AADT from the nearest street is considered. Vehicles km^{-2} estimated following Eq. (9). Palau Reial vehicles km^{-2} is not included because it is an urban background site not directly exposed to high traffic.

Site	a_r	$\overline{bh}$	bd	$bhdev$	z_0	AADT	vehicles km^{-2}
1. Palau Reial	0.12	14.6	0.12	6.4	1.27	3900	-
2. Eixample	0.00	21.1	0.40	8.4	1.03	41000	5666
3. Gràcia-Sant Gervasi	0.38	17.2	0.45	7.1	1.68	12700	3884
4. 213 Industria Street	1.00	18.1	0.38	8.3	1.94	15200	3003
5. 445 Valencia Street	0.86	19.5	0.32	7.2	2.20	32500	5978
6. 309 Industria Street	1.03	17.0	0.31	8.1	1.97	12900	3320

a marked street canyon pattern. Most of its canyons are about 20 to 25 m high and 20 m wide (i.e. a_r=1 and higher than 1). Experimental campaign sites are considered traffic sites in this work because they are exposed to similar AADT and vehicles km^{-2} compared to official traffic sites as shown in the table below. We apply Eq. (9) to obtain vehicles km^{-2}, a variable that describes traffic density in an area of 1 km^2.

$$vehicles \cdot km^{-2} = \sum_{n=1}^{st} Vehicles/s \cdot length \quad (9)$$

To obtain the amount of vehicles per second, AADT is divided by 3600 * 24 and multiplied by a temporal factor (i.e. 1.47) representing a typical factor for morning traffic peak in Barcelona. $Length$ is street length in metres. st is the number of streets over the circular area of 1 km^2 centered in the measurement site.

Table 2. Morphometric and traffic description of measurement sites used in CALIOPE-Urban evaluation. AADT from the nearest street is considered. Vehicles km^{-2} estimated following Eq. (9). Palau Reial vehicles km^{-2} is not included because it is an urban background site not directly exposed to high traffic.

Site	a_r	$\overline{bh}$	bd	$bhdev$	z_0	AADT	vehicles km^{-2}
1. Palau Reial	0.12	14.6	0.12	6.4	1.27	3900	-
2. Eixample	0.00	21.1	0.40	8.4	1.03	41000	5666
3. Gràcia-Sant Gervasi	0.38	17.2	0.45	7.1	1.68	12700	3884
4. 213 Industria Street	1.00	18.1	0.38	8.3	1.94	15200	3003
5. 455 Valencia Street	0.86	19.5	0.32	7.2	2.20	32500	5978
6. 309 Industria Street	1.03	17.0	0.31	8.1	1.97	12900	3320

3.2 NO_2 spatial variability: Passive dosimeters campaign across the city

With the objective of representing the NO_2 spatial variability, 212 passive dosimeters were located in Barcelona from 28/2/2017 to 15/3/2017 as depicted by red dots with yellow border in Fig. 5. In every km^2 there were at least two dosimeters, representing the background and traffic conditions, respectively, at 2.2-2.5 m height. The 100 background dosimeters were placed more than 10 m away from the road and the 112 traffic dosimeters were located at less than 3 m away from the road and at least 25 m away from intersections. To ensure the equivalence of measurements to standard conditions, these were corrected through comparison with reference equipment from several sites of the XVPQA network. After a preliminary inspection of the location of the dosimeters, we discarded data from 30 dosimeters to avoid results that could not be interpreted for several reasons (e.g. dosimeter and simulated road at different heights; highway covered by a tunnel near dosimeter location that is not considered in the emission inventory; lack of emission sources near dosimeter).

4 Results and discussion

Section 4.1 presents the temporal variability of NO_2 concentrations estimated by CALIOPE and CALIOPE-Urban compared to observations at the six sites described in Sect. 3.1. Section 4.2 describes the results in terms of the spatial variation during the

3.2 NO_2 spatial variability: Passive dosimeters campaign across the city

With the objective of representing the NO_2 spatial variability, 212 passive dosimeters were located in Barcelona from 28/2/2017 to 15/3/2017 as depicted by red dots with yellow border in Fig. 6. In every km^2 of surface there were at least two dosimeters, representing the background and traffic conditions, respectively, at 2.2-2.5 m height. The 100 background dosimeters were placed more than 10 m away from the road and the 112 traffic dosimeters were located at less than 3 m away from the road and at least 25 m away from intersections. To ensure the equivalence of measurements to standard conditions, these were corrected through comparison with reference equipment from several sites of the XVPQA network. After a preliminary inspection of the location of the dosimeters, we discarded data from 30 dosimeters to avoid results that could not be interpreted for several reasons (e.g. dosimeter and simulated road at different heights; highway covered by a tunnel near dosimeter location that is not considered in the emission inventory; lack of emission sources near dosimeter).

4 Results and discussion

Section 4.1 presents the temporal variability of NO_2 concentrations estimated by CALIOPE and CALIOPE-Urban compared to observations at the six sites described in Sect. 3.1. Section 4.2 describes the results in terms of the spatial variation during the two-week passive dosimeter campaign described in Sect. 3.2. Model performance is quantified using performance measures as described by Chang and Hanna (2004) and using assessment target plots (defined in the FAIRMODE initiative, Janssen et al., 2017). The performance statistics used here are the geometric mean bias (GeoMean), the fraction of model results within a factor of two of observations (FAC2), the geometric standard deviation (GeoSD), the correlation coefficient (R), the mean bias (MB) and the root mean square error (RMSE). The mathematical expressions of these statistics can be found in the Appendix C.

4.1 Temporal variation of NO_2 concentrations within urban streets

The scatter plots of Fig. 7 compare CALIOPE and CALIOPE-Urban outputs with observations based on hourly, daily mean and maximum modelled concentrations in the six sites described in Sect. 3.1 for April and May 2013. In general, CALIOPE-Urban shows a greater agreement for hourly, daily means and maximum concentrations but tends to underpredict daily peak concentrations in sites not exposed to very high traffic intensity (i.e. sites where urban background contribution predominates like Gràcia-Sant Gervasi). During the study period most of daily maxima (i.e. 56 %) occur at morning or evening traffic peak times (i.e. 6-7 or 18-20 UTC) when atmospheric conditions are typically stable and traffic intensity is high.

two-week passive dosimeter campaign described in Sect. 3.2. Model performance is quantified using performance measures as described by Chang and Hanna (2004) and using assessment target plots (defined in the FAIRMODE initiative, Janssen et al., 2017). The performance statistics used here are the geometric mean bias (GeoMean), the fraction of model results within a factor of two of observations (FAC2), the geometric standard deviation (GeoSD), the correlation coefficient (R), the mean bias (MB) and the root mean square error (RMSE). The mathematical expressions of these statistics can be found in the Appendix C.

4.1 Temporal variation of NO_2 concentrations within urban streets

The scatter plots of Fig. 6 compare CALIOPE and CALIOPE-Urban outputs with observations based on hourly, daily mean and maximum modelled concentrations in the six sites described in Sect. 3.1 for April and May 2013. In general, CALIOPE-Urban shows a greater agreement for hourly, daily means and maximum concentrations but tends to underpredict daily peak concentrations in sites not exposed to very high traffic intensity (i.e. sites where urban background contribution predominates like Gràcia-Sant Gervasi). During the study period most of daily maxima (i.e. 56 %) occur at morning or evening traffic peak times (i.e. 6-7 or 18-20 UTC) when atmospheric conditions are typically stable and traffic intensity is high.

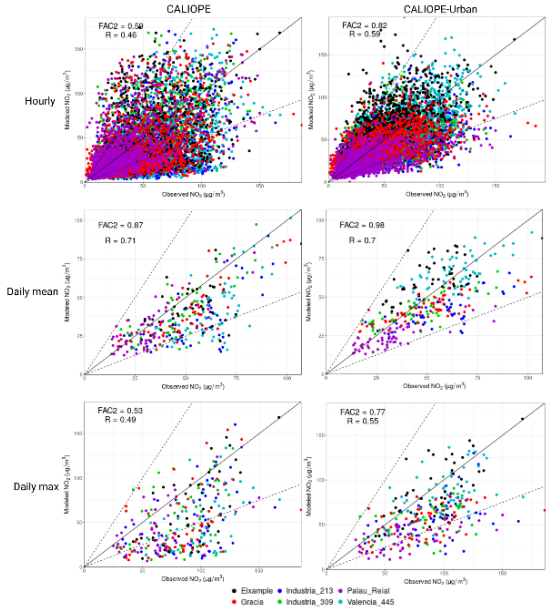

Figure 6. Scatter plot of hourly (top), daily mean (middle) and daily maximum (bottom) modelled concentrations against observed concentrations with colors representing monitoring sites for CALIOPE (left) and CALIOPE-Urban (right). Purple color represents Palau Reial, the urban background site. The other colors represent traffic sites as described in Sect. 3.1.

Table 3 shows the model performance statistics computed with hourly data, including CALIOPE-Urban-nl run. We compare CALIOPE-Urban and CALIOPE-Urban-nl to assess the difference in performance derived by the use of the local developments described in Sect. 2.3. All systems perform well at urban background sites and only CALIOPE-Urban gives good agreement with observations in traffic sites. The greatest difference between CALIOPE and CALIOPE-Urban systems performance is produced at the 455 Valencia Street site due to its street canyon morphology (a_r = 0.86). In this site, the mean transport is well resolved by the channelled winds, and its high AADT produces a high increase in traffic emissions within R-LINE. CALIOPE-Urban-nl largely overestimates NO_2 concentrations in this site for several reasons: it uses directly the output of UBS for background, instead of applying the vertical mixing that reduces background at street level specially under stable conditions; z_0 is given the WRF value (z_0 = 1.0), which is much lower than its locally estimated value (i.e. z_0 = 2.2, see Table

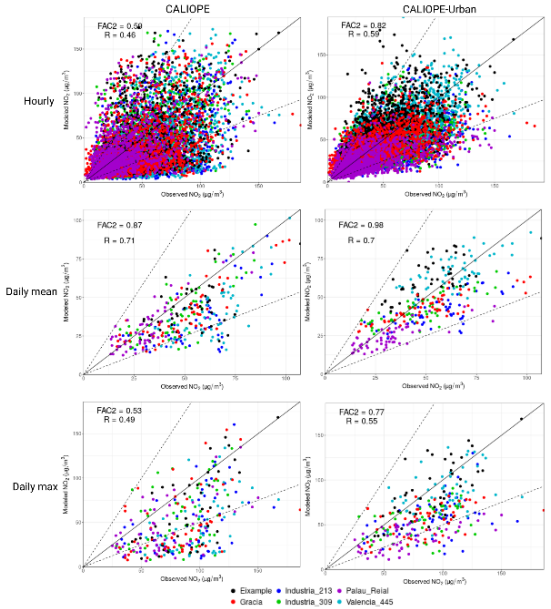

Figure 7. Scatter plot of hourly (top), daily mean (middle) and daily maximum (bottom) modelled concentrations against observed concentrations with colors representing monitoring sites for CALIOPE (left) and CALIOPE-Urban (right). Purple color represents Palau Reial, the urban background site. The other colors represent traffic sites as described in Sect. 3.1.

Table 4 shows the model performance statistics computed with hourly data, including CALIOPE-Urban-nl run. We compare CALIOPE-Urban and CALIOPE-Urban-nl to assess the difference in performance derived by the use of the local developments described in Sect. 2.3. All systems perform well at urban background sites and only CALIOPE-Urban gives good agreement with observations in traffic sites. The greatest difference between CALIOPE and CALIOPE-Urban systems performance is produced at the 445 Valencia Street site due to its street canyon morphology (a_r = 0.86). In this site, the mean transport is well resolved by the channelled winds, and its high AADT produces a high increase in traffic emissions within R-LINE. CALIOPE-Urban-nl largely overestimates NO_2 concentrations in this site for several reasons: it uses directly the output of UBS for background, instead of applying the vertical mixing that reduces background at street level specially under stable conditions; z_0 is given the WRF value (z_0 = 1.0), which is much lower than its locally estimated value (i.e. z_0 = 2.2, see Table

2) that enhances dispersion decreasing concentration levels; lastly pollutant dispersion is not channelled within the street, so higher contributions of nearby streets may be expected.

Table 3. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) for hourly concentrations during April and May 2013 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban without local developments (CALIOPE-Urban-nl). Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE	**0.73**	**-1.23**	24.11	**1.10**	1.25	0.55
	CALIOPE-Urban	0.72	-8.70	**21.57**	1.28	**1.22**	**0.57**
	CALIOPE-Urban-nl	0.67	3.61	26.34	1.28	**1.22**	0.55
2. Eixample	CALIOPE	0.60	**-8.57**	35.14	1.35	1.35	0.34
	CALIOPE-Urban	**0.86**	9.38	**26.70**	**0.83**	**1.11**	**0.55**
	CALIOPE-Urban-nl	0.61	39.53	54.92	0.57	1.38	0.45
3. Gràcia-Sant Gervasi	CALIOPE	0.55	-10.95	31.95	1.38	1.39	0.47
	CALIOPE-Urban	**0.79**	-7.39	**25.11**	**1.07**	**1.19**	**0.52**
	CALIOPE-Urban-nl	0.66	**6.00**	35.91	0.91	1.43	0.38
4. 213 Industria Street	CALIOPE	0.52	-19.13	35.13	1.79	1.54	0.44
	CALIOPE-Urban	**0.78**	-13.62	**26.55**	1.30	**1.17**	**0.57**
	CALIOPE-Urban-nl	0.75	**1.57**	31.12	**1.04**	1.26	0.54
5. 455 Valencia Street	CALIOPE	0.50	-21.94	38.31	1.85	1.53	0.43
	CALIOPE-Urban	**0.92**	**2.92**	**23.26**	**0.94**	**1.07**	**0.56**
	CALIOPE-Urban-nl	0.79	23.72	42.29	0.75	1.19	0.47
6. 309 Industria Street	CALIOPE	0.64	-7.41	28.49	1.36	1.33	0.53
	CALIOPE-Urban	**0.84**	**-4.60**	**22.72**	**1.05**	**1.13**	0.53
	CALIOPE-Urban-nl	0.78	11.60	31.13	0.83	1.24	**0.58**

On the other hand, CALIOPE-Urban underestimations at 213 and 309 Industria Street and Gràcia-Sant Gervasi may be due to an unrealistically low AADT level on the street segment close to the site. We work with AADT data that is based on the outputs of the traffic model used by Barcelona City Council that may be underestimating traffic. Another explanation may be an underestimation of local background levels within the area mostly during the afternoon. The afternoon underestimations in the mesoscale system could be caused by an overestimation of the mixing that produces a too low background NO_2 concentration level. This issue is difficult to correct because background concentrations used in the system are dependent on mesoscale concentrations, which are underestimated during daytime. In Table B1 in the Appendix B, same statistics are computed for daily mean results, finding similar results as in the hourly analysis. In addition, the analytical version of CALIOPE-Urban is shown to produce similar results for hourly concentrations to the numerical version in Table B2 in the Appendix B. This result may be interesting for forecasting applications at urban scale that require high-resolution because the analytical dispersion algorithm spends approx. half the time computing in comparison to the numerical dispersion algorithm as shown in Table 1.

Figure 7 shows NO_2 assessment target plots for CALIOPE and CALIOPE-Urban. In the plots the centred root mean square error (CRMSE) for each measurement station is plotted against the normalized bias. Distance from circle origin gives an estimate for the model quality indicator (MQI; Thunis and Cuvelier, 2016) that measures general model accuracy depending on measurement uncertainty. MQI values below 1 (i.e. green shading area) are considered to comply with the model quality

3) that enhances dispersion decreasing concentration levels; lastly pollutant dispersion is not channelled within the street, so higher contributions of nearby streets may be expected.

Table 4. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) for hourly concentrations during April and May 2013 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban without local developments (CALIOPE-Urban-nl). Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE	**0.73**	**-1.23**	24.11	**1.10**	1.25	0.55
	CALIOPE-Urban	0.72	-8.70	**21.57**	1.28	**1.22**	**0.57**
	CALIOPE-Urban-nl	0.67	3.61	26.34	1.28	**1.22**	0.55
2. Eixample	CALIOPE	0.60	**-8.57**	35.14	1.35	1.35	0.34
	CALIOPE-Urban	**0.86**	9.38	**26.70**	**0.83**	**1.11**	**0.55**
	CALIOPE-Urban-nl	0.61	39.53	54.92	0.57	1.38	0.45
3. Gràcia-Sant Gervasi	CALIOPE	0.55	-10.95	31.95	1.38	1.39	0.47
	CALIOPE-Urban	**0.79**	-7.39	**25.11**	**1.07**	**1.19**	**0.52**
	CALIOPE-Urban-nl	0.66	**6.00**	35.91	0.91	1.43	0.38
4. 213 Industria Street	CALIOPE	0.52	-19.13	35.13	1.79	1.54	0.44
	CALIOPE-Urban	**0.78**	-13.62	**26.55**	1.30	**1.17**	**0.57**
	CALIOPE-Urban-nl	0.75	**1.57**	31.12	**1.04**	1.26	0.54
5. 445 Valencia Street	CALIOPE	0.50	-21.94	38.31	1.85	1.53	0.43
	CALIOPE-Urban	**0.92**	**2.92**	**23.26**	**0.94**	**1.07**	**0.56**
	CALIOPE-Urban-nl	0.79	23.72	42.29	0.75	1.19	0.47
6. 309 Industria Street	CALIOPE	0.64	-7.41	28.49	1.36	1.33	0.53
	CALIOPE-Urban	**0.84**	**-4.60**	**22.72**	**1.05**	**1.13**	0.53
	CALIOPE-Urban-nl	0.78	11.60	31.13	0.83	1.24	**0.58**

On the other hand, CALIOPE-Urban underestimations at 213 and 309 Industria Street and Gràcia-Sant Gervasi may be due to an unrealistically low AADT level on the street segment close to the site. We work with AADT data that is based on the outputs of the traffic model used by Barcelona City Council that may be underestimating traffic. Another explanation may be an underestimation of local background levels within the area mostly during the afternoon. The afternoon underestimations in the mesoscale system could be caused by an overestimation of the mixing that produces a too low background NO_2 concentration level. This issue is difficult to correct because background concentrations used in the system are dependent on mesoscale concentrations, which are underestimated during daytime. In Table B1 in Appendix B, same statistics are computed for daily mean results, finding similar results as in the hourly analysis. In addition, the analytical version of CALIOPE-Urban is shown to produce similar results for hourly concentrations to the numerical version in Table B2 in Appendix B. This result may be interesting for forecasting applications at urban scale that require high-resolution because the analytical dispersion algorithm spends approx. half the time computing in comparison to the numerical dispersion algorithm as shown in Table 2.

Figure 8 shows NO_2 assessment target plots for CALIOPE and CALIOPE-Urban. In the plots the centred root mean square error (CRMSE) for each measurement station is plotted against the normalized bias. Distance from circle origin gives an estimate for the model quality indicator (MQI; Thunis and Cuvelier, 2016) that measures general model accuracy depending on measurement uncertainty. MQI values below 1 (i.e. green shading area) are considered to comply with the model quality

objective. All sites in CALIOPE-Urban simulation fall within the green shading area (i.e. complying with FAIRMODE's model quality objective). In contrast, four out of six in CALIOPE lie within the green shading area clearly showing the positive effect of the street scale model in the coupled system.

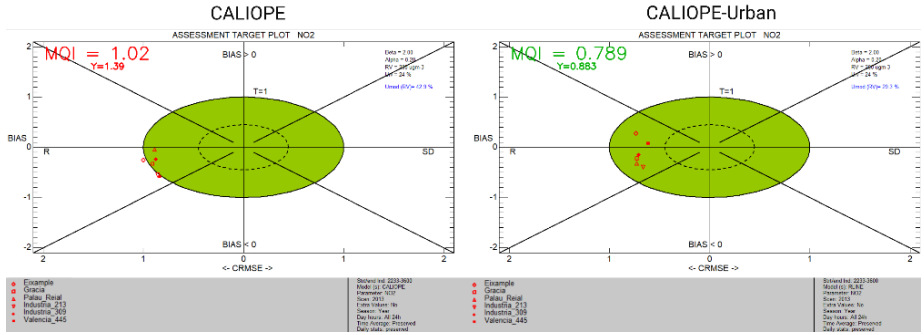

Figure 7. NO_2 model assessment target plots for CALIOPE (left) and CALIOPE-Urban (right). Symbols correspond to the six measurement sites described in Sect. 3.1 and the MQI for each site is represented by the distance between the circle origin and the site symbol.

Figure 8 shows averaged daily cycles for weekday and weekend periods for the six sites described in Sect. 3.1 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl. In general, all systems show a significant change between weekday and weekend in accordance with observations. The overall dynamic is well reproduced by all systems but CALIOPE tends to underestimate the afternoon levels and to overestimate nighttime values. CALIOPE-Urban-nl overestimates nighttime values and morning peaks. CALIOPE-Urban partly corrects CALIOPE afternoon underestimations close to high traffic (i.e. Valencia Street and Eixample stations) but still underestimates in low traffic sites. CALIOPE's tendency to overestimate the evening peak and night values may bring CALIOPE-Urban to generally overestimate on those hours as found in the literature near road sites (Hood et al., 2018). However, the vertical mixing implemented in CALIOPE-Urban decreases background concentrations mixing from aloft during night hours because under stable atmospheric conditions vertical mixing is reduced compared to daylight hours, which are more convective. This effect can be noticed in the difference between CALIOPE-Urban and CALIOPE-Urban-nl from 0 to 6 and from 18 to 23 (UTC) in traffic sites (i.e. sites 2,3,4,5 and 6 in Fig. 8), where CALIOPE-Urban concentration levels correct the night overestimations seen in CALIOPE-Urban-nl. Such result shows the benefit of considering the vertical stability in the coupling procedure of the mesoscale and the street scale dispersion model.

objective. All sites in CALIOPE-Urban simulation fall within the green shading area (i.e. complying with FAIRMODE's model quality objective). In contrast, four out of six in CALIOPE lie within the green shading area clearly showing the positive effect of the street scale model in the coupled system.

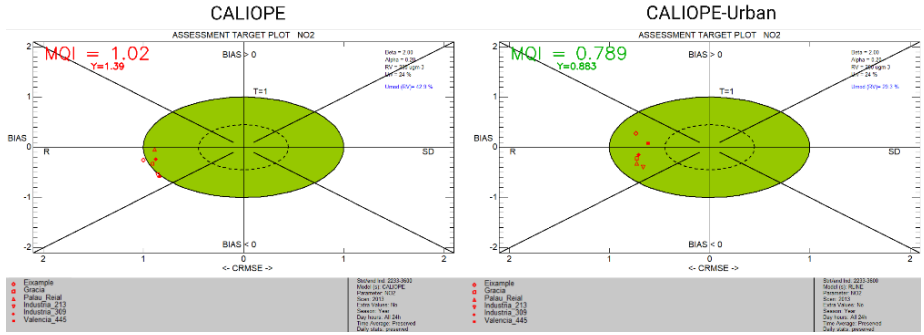

Figure 8. NO_2 model assessment target plots for CALIOPE (left) and CALIOPE-Urban (right). Symbols correspond to the six measurement sites described in Sect. 3.1 and the MQI for each site is represented by the distance between the circle origin and the site symbol.

Figure 9 shows averaged daily cycles for weekday and weekend periods for the six sites described in Sect. 3.1 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl. In general, all systems show a significant change between weekday and weekend in accordance with observations. The overall dynamic is well reproduced by all systems but CALIOPE tends to underestimate the afternoon levels and to overestimate nighttime values. CALIOPE-Urban-nl overestimates nighttime values and morning peaks. CALIOPE-Urban partly corrects CALIOPE afternoon underestimations close to high traffic (i.e. Valencia Street and Eixample stations) but still underestimates in low traffic sites. CALIOPE's tendency to overestimate the evening peak and night values may bring CALIOPE-Urban to generally overestimate on those hours as found in the literature near road sites (Hood et al., 2018). However, the vertical mixing implemented in CALIOPE-Urban decreases background concentrations mixing from aloft during night hours because under stable atmospheric conditions vertical mixing is reduced compared to daylight hours, which are more convective. This effect can be noticed in the difference between CALIOPE-Urban and CALIOPE-Urban-nl from 0 to 6 and from 18 to 23 (UTC) in traffic sites (i.e. sites 2,3,4,5 and 6 in Fig. 9), where CALIOPE-Urban concentration levels correct the night overestimations seen in CALIOPE-Urban-nl. Such result shows the benefit of considering the vertical stability in the coupling procedure of the mesoscale and the street scale dispersion model.

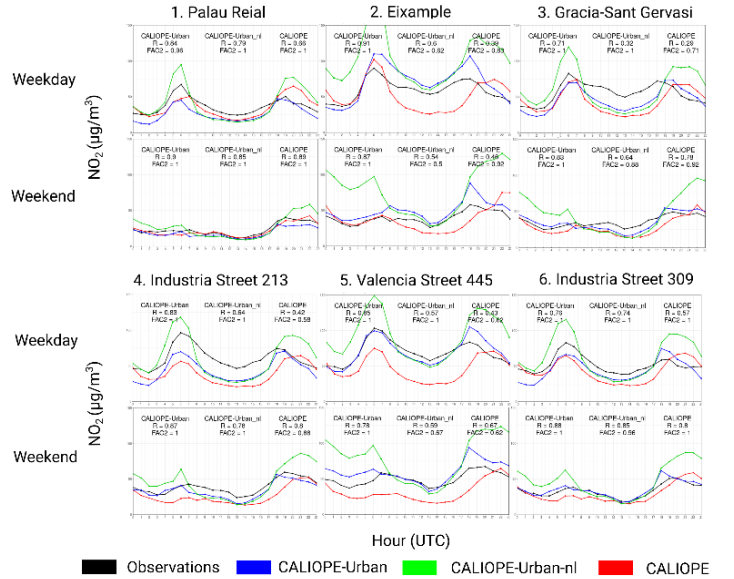

Figure 8. NO_2 average daily cycle at all sites described in Sect. 3.1 during April and May 2013 for weekday and weekend. Observations are represented in black coloured lines, red lines are CALIOPE, blue lines are CALIOPE-Urban and green lines represent CALIOPE-Urban without local developments (CALIOPE-Urban-nl).

There is a noticeable difference between CALIOPE-Urban's accuracy at 213 Industria Street and 445 Valencia Street given similar observations and CALIOPE levels at both sites. Although both sites are located in areas with considerable traffic activity, Valencia Street site has higher modelled traffic emissions, deriving in higher local pollutant concentrations, and a higher density of vehicles km^{-2} as described in Table 2. Consequently, to improve CALIOPE-Urban accuracy an increase of local simulated traffic at 213 Industria Street site could bring a model accuracy improvement. However, the lack of observational traffic count data at the monitoring sites does not permit to explore the precision of the input AADT information considered in HERMESv2.0 at those locations.

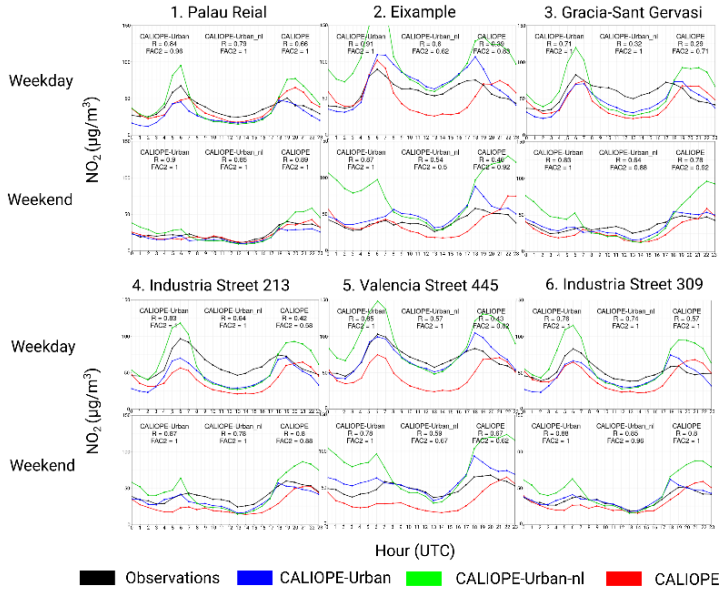

Figure 9. NO_2 average daily cycle at all sites described in Sect. 3.1 during April and May 2013 for weekday and weekend. Observations are represented in black coloured lines, red lines are CALIOPE, blue lines are CALIOPE-Urban and green lines represent CALIOPE-Urban without local developments (CALIOPE-Urban-nl).

There is a noticeable difference between CALIOPE-Urban's accuracy at 213 Industria Street and 445 Valencia Street given similar observations and CALIOPE levels at both sites. Although both sites are located in areas with considerable traffic activity, Valencia Street site has higher modelled traffic emissions, resulting in higher local pollutant concentrations, and a higher density of vehicles km^{-2} as described in Table 3. Consequently, to improve CALIOPE-Urban accuracy an increase of local simulated traffic at 213 Industria Street site could bring a model accuracy improvement. However, the lack of observational traffic count data at the monitoring sites does not permit to explore the precision of the input AADT information considered in HERMESv2.0 at those locations.

4.2 Spatial variation of NO_2 concentrations across the city

We evaluate CALIOPE and CALIOPE-Urban NO_2 in terms of spatial variations across Barcelona city using measurements from 182 valid passive dosimeters as described in Sect. 3.2. Table 4 gives statistics at the 182 sites where passive dosimeters measured NO_2 concentrations for a two week period (28 February - 15 March 2017) for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl (without local developments).

Table 4. NO_2 model evaluation statistics calculated at 182 passive dosimeter valid sites (described in Sect. 3.2) during two weeks from 28 February to 15 March 2017 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl mean concentrations. Bold numbers represent model results with better performance for each statistic and site. Results are shown for all (including traffic and urban background), only urban background and only traffic sites.

Sites	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
All	CALIOPE	0.92	-14.00	21.88	1.30	1.08	0.36
	CALIOPE-Urban	0.97	**-7.26**	**17.21**	1.20	**1.06**	**0.70**
	CALIOPE-Urban-nl	**0.98**	15.24	28.30	**0.81**	**1.06**	0.69
Background	CALIOPE	**1.00**	**-2.84**	**8.08**	**1.06**	**1.02**	0.66
	CALIOPE-Urban	0.97	-7.34	12.71	1.23	1.06	0.54
	CALIOPE-Urban-nl	**1.00**	10.09	15.03	0.82	1.04	**0.72**
Traffic	CALIOPE	0.81	-25.57	30.74	1.60	1.63	0.22
	CALIOPE-Urban	**0.97**	**-7.17**	**20.62**	**1.17**	**1.06**	**0.53**
	CALIOPE-Urban-nl	0.96	20.17	36.76	0.81	1.08	0.50

Considering all sites, CALIOPE-Urban shows a much better correlation coefficient (0.70 vs 0.36) than CALIOPE due to its good performance at traffic sites. Compared to CALIOPE-Urban-nl their correlation is similar. If we consider only urban background sites, CALIOPE shows a greater correlation coefficient than CALIOPE-Urban (0.66 vs. 0.54) and a MB closer to 0. In addition, CALIOPE-Urban-nl gives a better correlation than both systems. A potential explanation for this result is related to the error compensation shown in the temporal evaluation (Sect. 4.1). CALIOPE and CALIOPE-Urban-nl may compensate the underestimation during daytime with the overestimation during nighttime. In contrast, CALIOPE-Urban may not compensate the daytime underestimations with overestimated night values because the background is reduced due to low vertical mixing effect during nighttime (stable) hours. An enhanced daytime NO_2 background contribution would improve CALIOPE-Urban accuracy at urban background sites.

For traffic sites, CALIOPE shows a strong underestimation (MB = -25.57 µg m^{-3}) and CALIOPE-Urban gives MB levels closer to 0. CALIOPE-Urban underestimations may be influenced by afternoon underestimations and a misrepresentation of traffic emissions in some areas of the city. In contrast, CALIOPE-Urban-nl gives a high MB and the highest RMSE among the three systems. This tendency to over estimate near traffic of CALIOPE-Urban-nl may be due to the reasons stated in Sect. 4.1. In general, closer to intense traffic CALIOPE-Urban is very sensitive to emissions and its dispersion characterizes well the spatial variability for the study period. Reproducing spatial gradients near intense traffic is crucial in a city like Barcelona given its high vehicle density and NO_2 concentration levels.

4.2 Spatial variation of NO_2 concentrations across the city

We evaluate CALIOPE and CALIOPE-Urban NO_2 in terms of spatial variations across Barcelona city using measurements from 182 valid passive dosimeters as described in Sect. 3.2. Table 5 gives statistics at the 182 sites where passive dosimeters measured NO_2 concentrations for a two week period (28 February - 15 March 2017) for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl (without local developments).

Table 5. NO_2 model evaluation statistics calculated at 182 passive dosimeter valid sites (described in Sect. 3.2) during two weeks from 28 February to 15 March 2017 for CALIOPE, CALIOPE-Urban and CALIOPE-Urban-nl mean concentrations. Bold numbers represent model results with better performance for each statistic and site. Results are shown for all (including traffic and urban background), only urban background and only traffic sites.

Sites	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
All	CALIOPE	0.92	-14.00	21.88	1.30	1.08	0.36
	CALIOPE-Urban	0.97	**-7.26**	**17.21**	1.20	**1.06**	**0.70**
	CALIOPE-Urban-nl	**0.98**	15.24	28.30	**0.81**	**1.06**	0.69
Background	CALIOPE	**1.00**	**-2.84**	**8.08**	**1.06**	**1.02**	0.66
	CALIOPE-Urban	0.97	-7.34	12.71	1.23	1.06	0.54
	CALIOPE-Urban-nl	**1.00**	10.09	15.03	0.82	1.04	**0.72**
Traffic	CALIOPE	0.81	-25.57	30.74	1.60	1.63	0.22
	CALIOPE-Urban	**0.97**	**-7.17**	**20.62**	**1.17**	**1.06**	**0.53**
	CALIOPE-Urban-nl	0.96	20.17	36.76	0.81	1.08	0.50

Considering all sites, CALIOPE-Urban shows a much better correlation coefficient (0.70 vs 0.36) than CALIOPE due to its good performance at traffic sites. Compared to CALIOPE-Urban-nl their correlation is similar. If we consider only urban background sites, CALIOPE shows a greater correlation coefficient than CALIOPE-Urban (0.66 vs. 0.54) and a MB closer to 0. In addition, CALIOPE-Urban-nl gives a better correlation than both systems. A potential explanation for this result is related to the error compensation shown in the temporal evaluation (Sect. 4.1). CALIOPE and CALIOPE-Urban-nl may compensate the underestimation during daytime with the overestimation during nighttime. In contrast, CALIOPE-Urban may not compensate the daytime underestimations with overestimated night values because the background is reduced due to low vertical mixing effect during nighttime (stable) hours. An enhanced daytime NO_2 background contribution would improve CALIOPE-Urban accuracy at urban background sites.

For traffic sites, CALIOPE shows a strong underestimation (MB = -25.57 µg m^{-3}) and CALIOPE-Urban gives MB levels closer to 0. CALIOPE-Urban underestimations may be influenced by afternoon underestimations and a misrepresentation of traffic emissions in some areas of the city. In contrast, CALIOPE-Urban-nl gives a high MB and the highest RMSE among the three systems. This tendency to over estimate near traffic of CALIOPE-Urban-nl may be due to the reasons stated in Sect. 4.1. In general, closer to intense traffic CALIOPE-Urban is very sensitive to emissions and its dispersion characterizes well the spatial variability for the study period. Reproducing spatial gradients near intense traffic is crucial in a city like Barcelona given its high vehicle density and NO_2 concentration levels.

Figure 9 shows the difference between CALIOPE and CALIOPE-Urban results and measurements (top panels) and scatter plots at all sites (bottom panels) distinguished with colors by site type (e.g., traffic site, urban background site).

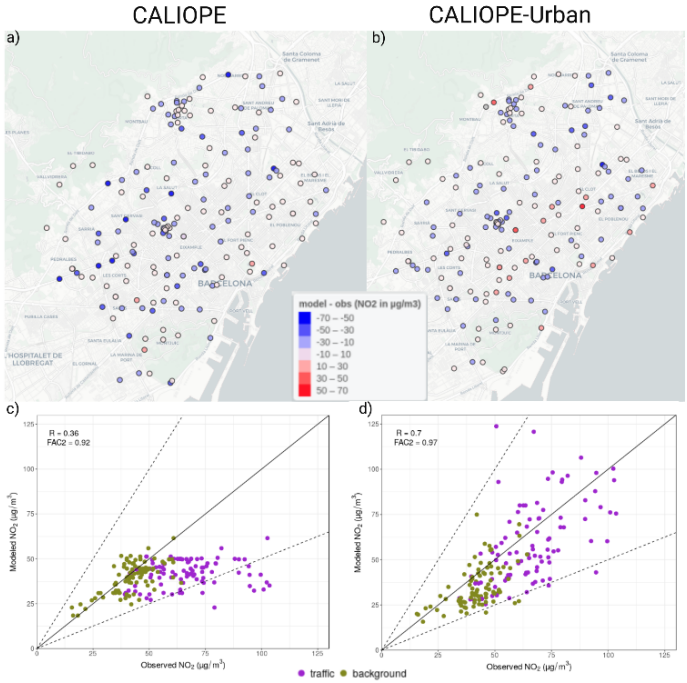

Figure 9. (Top) NO_2 concentrations difference (model - observations) for two week averaged values during dosimeters campaign in 2017, CALIOPE is left (a) and CALIOPE-Urban is right (b). (Bottom) scatter plot of modelled vs observed concentrations with colors representing site type (olive green for background and purple for traffic) for CALIOPE (c) and CALIOPE-Urban (d). Correlation (R) and agreement factor of 2 (FAC2) are computed for all sites.

In Fig. 9a the concentration difference map of CALIOPE shows an overall underestimation, represented by blue dots. This underestimation is found to be systematic in traffic sites in the scatter of Fig. 9c (purple dots), where modelled values barely exceed 50 µg m^{-3} while most of the observed values at traffic sites are above that value. In contrast, the CALIOPE-Urban difference map (Fig. 9b) shows a more mixed picture with a broader representation of white dots (bias close to 0) but also more

Figure 10 shows the difference between CALIOPE and CALIOPE-Urban results and measurements (top panels) and scatter plots at all sites (bottom panels) distinguished with colors by site type (e.g., traffic site, urban background site).

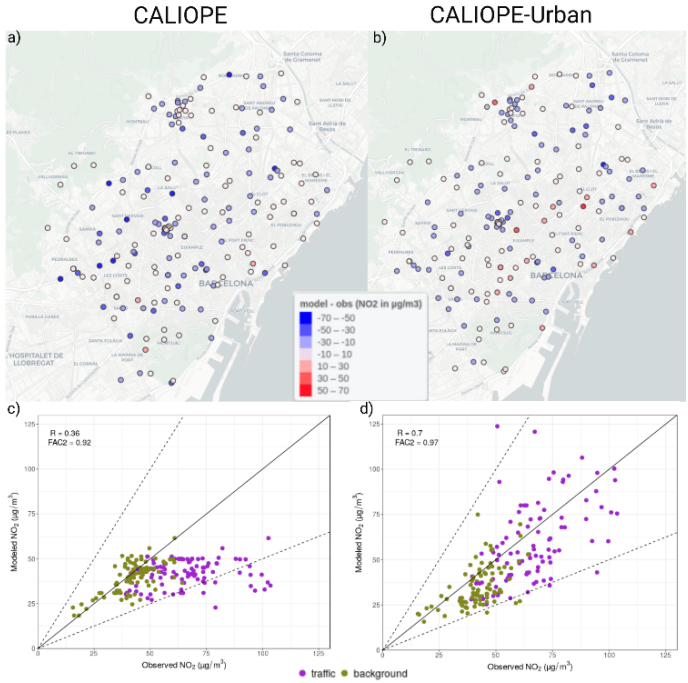

Figure 10. (Top) NO_2 concentrations difference (model - observations) for two week averaged values during dosimeters campaign in 2017, CALIOPE is left (a) and CALIOPE-Urban is right (b). (Bottom) scatter plot of modelled vs observed concentrations with colors representing site type (olive green for background and purple for traffic) for CALIOPE (c) and CALIOPE-Urban (d). Correlation (R) and agreement factor of 2 (FAC2) are computed for all sites.

In Fig. 10a the concentration difference map of CALIOPE shows an overall underestimation, represented by blue dots. This underestimation is found to be systematic in traffic sites in the scatter of Fig. 10c (purple dots), where modelled values barely exceed 50 µg m^{-3} while most of the observed values at traffic sites are above that value. In contrast, the CALIOPE-Urban difference map (Fig. 10b) shows a more mixed picture with a broader representation of white dots (bias close to 0) but

red ones in the city center and close to the highways. For CALIOPE-Urban's scatter, most of the model results at traffic sites are within the 1:2 and 1:0.5 dashed lines, showing a better agreement at traffic sites than CALIOPE (Fig. 9d). In CALIOPE-Urban's difference map, we see a spatial pattern with average bias close to 0 in the city centre, where traffic is denser and close to the highways surrounding the city. The appearance of red dots may indicate that CALIOPE-Urban overestimates close to high trafficked areas while CALIOPE underestimate in these areas. This may be due to an overestimation of traffic emissions or background concentrations in these areas. In contrast, in locations where traffic is not very intense (see Fig. 2 for NO_X emissions) CALIOPE-Urban shows systematic underestimations. This result may be derived from the systematic underestimation of midday NO_2 concentrations in low traffic areas as shown in Sect. 4.1.

4.3 Major uncertainty sources

Here we discuss potential sources of error in our model by analyzing episodes when the model was skillful compared with episodes when the model was not. Our analysis solely considers the meteorological and background concentration inputs as potential sources of error. While road traffic emission estimates may introduce large errors, we lack observations of traffic counts at the measurement site locations to properly assess them.

We calculated daily the RMSE of the hourly modelled NO_2 concentrations versus the observed values in the six sites described in Sect. 3.1 during the period April and May 2013. For each site we picked the ten days with highest RMSE as potential candidates and ten days with the lowest RMSE. We conducted this analysis for both CALIOPE and CALIOPE-Urban, finding that both systems share to a large extent the days with skill (4 out of 5 days) and without (3 out of 5). This result shows that the coupled system performance is highly dependent on the mesoscale model performance. To explore errors potentially caused by R-LINE inputs, in Fig. 9 we compare the five days with less skill (i.e. 11, 16, 17 April and 7,8 May) and the five days with more skill (i.e. 7, 20, April and 18, 19, 25 May) with observations for wind speed (ws), street level NO_2 and background NO_2.

On skillful days, winds are relatively strong and well represented in WRF (Fig. 10a). Poor skills appear when the observed wind speed is low. Because WRF largely underestimates wind speeds (Fig. 10b) and NO_2 concentrations are underestimated under calm conditions (Fig. 10d), other processes (e.g. atmospheric stability) may have a greater importance in this case. In our coupling under very stable atmospheric situations, dispersion is reduced and background injection from the overlying atmosphere is limited. This control mechanism adapts the system to specific street conditions, regulating dispersion and background injection. For these days, an extended observational dataset would be needed to better understand the model behaviour.

also more red ones in the city center and close to the highways. For CALIOPE-Urban's scatter, most of the model results at traffic sites are within the 1:2 and 1:0.5 dashed lines, showing a better agreement at traffic sites than CALIOPE (Fig. 10d). In CALIOPE-Urban's difference map, we see a spatial pattern with average bias close to 0 in the city centre, where traffic is denser and close to the highways surrounding the city. The appearance of red dots may indicate that CALIOPE-Urban overestimates close to high trafficked areas while CALIOPE underestimate in these areas. This may be due to an overestimation of traffic emissions or background concentrations in these areas. In contrast, in locations where traffic is not very intense (see Fig. 2 for NO_X emissions) CALIOPE-Urban shows systematic underestimations. This result may be derived from the systematic underestimation of midday NO_2 concentrations in low traffic areas as shown in Sect. 4.1.

4.3 Major uncertainty sources

Here we discuss potential sources of error in our model by analyzing episodes when the model was skillful compared with episodes when the model was not. Our analysis solely considers the meteorological and background concentration inputs as potential sources of error. While road traffic emission estimates may introduce large errors, we lack observations of traffic counts at the measurement site locations to properly assess them.

We calculated daily the RMSE of the hourly modelled NO_2 concentrations versus the observed values in the six sites described in Sect. 3.1 during the period April and May 2013. For each site we picked the ten days with highest RMSE as potential candidates and ten days with the lowest RMSE. Then, we put together the candidates of all sites and we chose the most frequent five days (i.e. from good and bad performance candidate days) for both CALIOPE and CALIOPE-Urban, finding that both systems share to a large extent the days with skill (4 out of 5 days) and without (3 out of 5). This result shows that the coupled system performance is highly dependent on the mesoscale model performance. To explore errors potentially caused by R-LINE inputs, in Fig. 11 we compare the five days with less skill (i.e. 11, 16, 17 April and 7,8 May) and the five days with more skill (i.e. 7, 20, April and 18, 19, 25 May) with observations for wind speed (ws), street level NO_2 and background NO_2.

On skillful days, winds are relatively strong and well represented in WRF (Fig. 11a). Poor skills appear when the observed wind speed is low. Because WRF largely underestimates wind speeds (Fig. 11b) and NO_2 concentrations are underestimated under calm conditions (Fig. 11d), other processes (e.g. atmospheric stability) may have a greater importance in this case. In our coupling under very stable atmospheric situations, dispersion is reduced and background injection from the overlying atmosphere is limited. This control mechanism adapts the system to specific street conditions, regulating dispersion and background injection. For these days, an extended observational dataset would be needed to better understand the model behaviour.

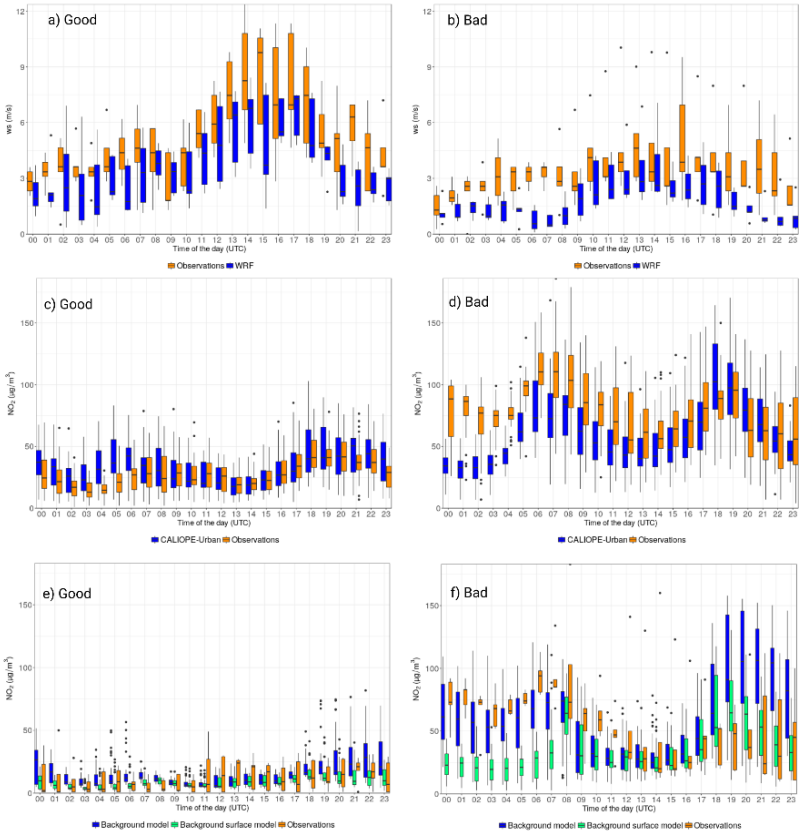

Figure 10. Boxplots by time of the day of good (left panels) and bad performance days (right panels) for CALIOPE-Urban inputs and observations with dots representing outliers. a) and b) represent WRF and observed wind speeds at Barcelona airport (10m height); c) and d) show observed and modelled NO_2 concentrations for the six sites in Sect. 3.1; e) and f) depict NO_2 observed concentrations at Ciutadella urban background station and background model averaged results at the six sites. Observed values are orange coloured and modelled results are blue. Light green represents background model results at surface level.

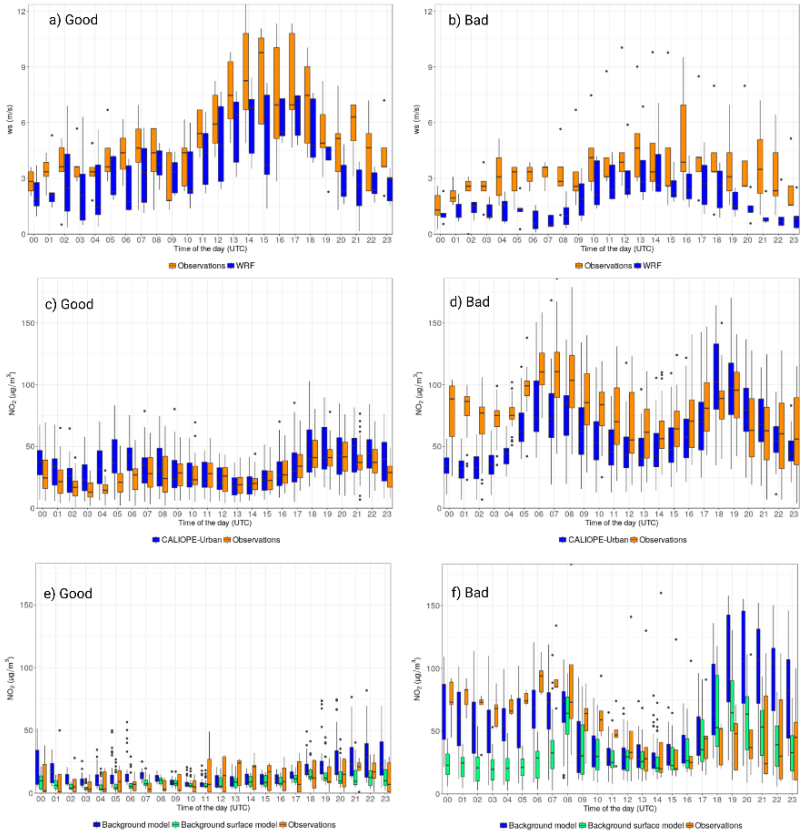

Figure 11. Boxplots by time of the day of good (left panels) and bad performance days (right panels) for CALIOPE-Urban inputs and observations with dots representing outliers. a) and b) represent WRF and observed wind speeds at Barcelona airport (10m height); c) and d) show observed and modelled NO_2 concentrations for the six sites in Sect. 3.1; e) and f) depict NO_2 observed concentrations at Ciutadella urban background station and background model averaged results at the six sites. Observed values are orange coloured and modelled results are blue. Light green represents background model results at surface level.

To analyze the background concentrations from the mesoscale simulation as a potential error source, we compared NO_2 observations from the Ciutadella urban background station with hourly modelled concentrations averaged over the six sites. The results shown in Fig. 10e,f represent concentrations provided by upwind CMAQ grid cells depending on wind speed and direction (blue) as described in Sect. 2.3.3 downscaled to surface level using the vertical decay method (green). As expected, observed NO_2 concentrations on days with calm conditions and therefore poor skill are higher than on those with enhanced ventilation and better skills. The background model reproduces well the variation during both types of days but overestimates concentrations during nighttime (19-22 UTC), particularly during days with calm conditions. This problem is partially corrected by using the background vertical decay method as seen in Fig. 10f and in Fig. 8. In addition, NO_2 concentrations are underestimated at the beginning of the day (1-4 UTC). The fact that the averaged diurnal cycle in Fig. 8 shows similar error patterns suggests that NO_2 background concentrations greatly influence NO_2 street level concentrations in agreement with Degraeuwe et al. (2017).

4.4 Hourly variation of street NO_2 concentrations

Hourly street NO_2 concentrations are expected to vary spatially and temporally with higher values close to intense traffic sites during rush hours. Figure 11 shows high resolution (10 m $\times$ 10 m) NO_2 concentration maps at four different hours on Thursday 11th of April 2013 (i.e. 0, 7, 12 and 18 UTC). This day is chosen because it shows a marked diurnal cycle with maxima consistent with the morning and evening traffic peaks (i.e. 6-7 or 18-20 UTC). Higher concentrations are found at 7 and 18 UTC where high traffic emissions are concentrated (i.e. highways surrounding the city and city center) because traffic intensity is higher at these hours of the day and the atmosphere tends to be stable, making pollutant dispersion more difficult. On the other hand, lower concentrations are found at 0 UTC due to the lower traffic intensity and at 12 UTC. At 12 UTC traffic intensity is considerably higher than at 0 UTC but the atmosphere is more convective and pollutant dispersion is enhanced.

In agreement with Duyzer et al. (2015) our modelling results show that Eixample and Gràcia-Sant Gervasi traffic stations do not represent the highest NO_2 concentrations in Barcelona. The highest levels are found in street canyons exposed to very high traffic intensity and not as well ventilated as the above-mentioned locations, and in open areas near highways surrounding the city. For example, measurements at 445 Valencia Street site show 20 % higher concentrations than in Eixample and Gràcia-Sant Gervasi traffic sites on average during the morning peak on weekdays (see Fig. 8). Hence, additional monitoring sites located within highly trafficked streets are clearly needed to better represent highest NO_2 concentration levels in Barcelona.

To analyze the background concentrations from the mesoscale simulation as a potential error source, we compared NO_2 observations from the Ciutadella urban background station with hourly modelled concentrations averaged over the six sites. We aimed to compare the modelled background concentrations (i.e. excluding local vehicular traffic contribution) with the most representative urban background observation, which in our case is the Ciutadella site. The results shown in Fig. 11e,f represent concentrations provided by upwind CMAQ grid cells depending on wind speed and direction (blue) as described in Sect. 2.3.3 downscaled to surface level using the vertical decay method (green). As expected, observed NO_2 concentrations on days with calm conditions and therefore poor skill are higher than on those with enhanced ventilation and better skills. The background model reproduces well the variation during both types of days but overestimates concentrations during nighttime (19-22 UTC), particularly during days with calm conditions. This problem is partially corrected by using the background vertical decay method as seen in Fig. 11f and in Fig. 9. In addition, NO_2 concentrations are underestimated at the beginning of the day (1-4 UTC). The fact that the averaged diurnal cycle in Fig. 9 shows similar error patterns suggests that NO_2 background concentrations greatly influence NO_2 street level concentrations in agreement with Degraeuwe et al. (2017).

4.4 Hourly variation of street NO_2 concentrations

Hourly street NO_2 concentrations are expected to vary spatially and temporally with higher values close to intense traffic sites during rush hours. Figure 12 shows high resolution (10 m $\times$ 10 m) NO_2 concentration maps at four different hours on Thursday 11th of April 2013 (i.e. 0, 7, 12 and 18 UTC). This day is chosen because it shows a marked diurnal cycle with maxima consistent with the morning and evening traffic peaks (i.e. 6-7 or 18-20 UTC). Higher concentrations are found at 7 and 18 UTC where high traffic emissions are concentrated (i.e. highways surrounding the city and city center) because traffic intensity is higher at these hours of the day and the atmosphere tends to be stable, making pollutant dispersion more difficult. On the other hand, lower concentrations are found at 0 UTC due to the lower traffic intensity and at 12 UTC. At 12 UTC traffic intensity is considerably higher than at 0 UTC but the atmosphere is more convective and pollutant dispersion is enhanced.

In agreement with Duyzer et al. (2015) our modelling results show that Eixample and Gràcia-Sant Gervasi traffic stations do not represent the highest NO_2 concentrations in Barcelona. The highest levels are found in street canyons exposed to very high traffic intensity and not as well ventilated as the above-mentioned locations, and in open areas near highways surrounding the city. For example, measurements at 445 Valencia Street site show 20 % higher concentrations than in Eixample and Gràcia-Sant Gervasi traffic sites on average during the morning peak on weekdays (see Fig. 9). Hence, additional monitoring sites located within highly trafficked streets are clearly needed to better represent highest NO_2 concentration levels in Barcelona.

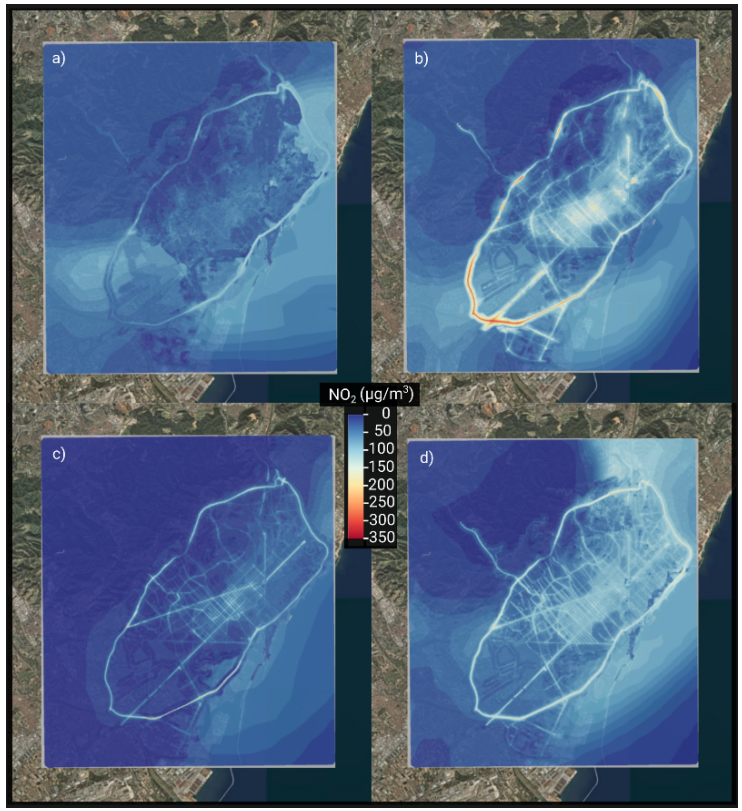

Figure 11. NO_2 high resolution (10 m × 10 m) concentration maps on 11th April 2013. a) represents concentrations at 0 UTC (2 am local time); b) is 7 UTC (9 am local time); c) is 12 UTC (2 pm local); and d) is 18 UTC (8 pm local).

5 Conclusions

This study describes the development of a coupled regional to street scale modelling system, CALIOPE-Urban, which provides high spatial and temporal resolution (up to 10 m × 10 m, hourly) NO_2 concentrations for Barcelona. It couples the mesoscale air quality forecasting system CALIOPE (WRF-HERMES-CMAQ-BSC-DREAM8b) with the urban roadway dispersion model,

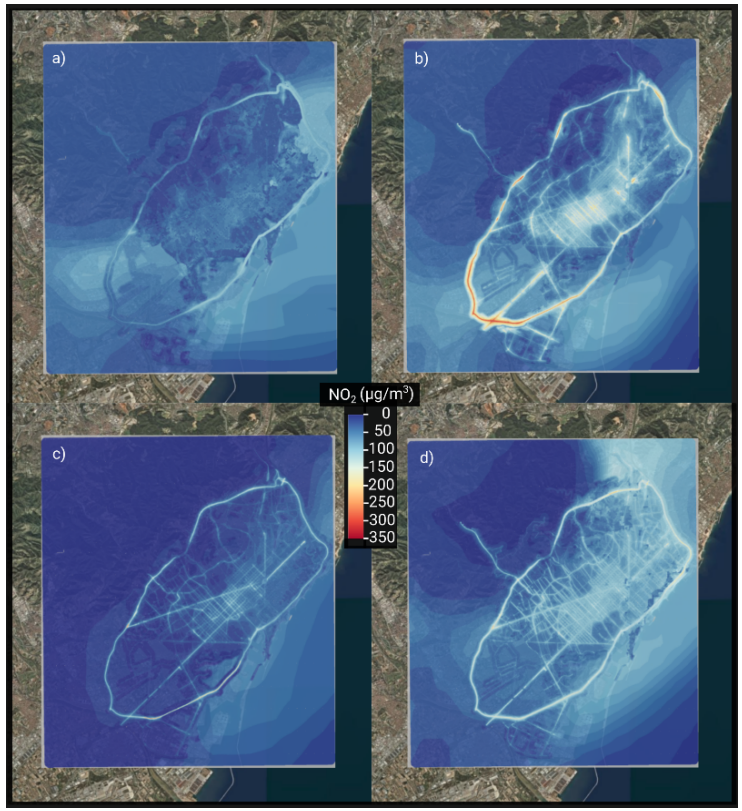

Figure 12. NO_2 high resolution (10 m × 10 m) concentration maps on 11th April 2013. The resolution is for the entire concentration map. a) represents concentrations at 0 UTC (2 am local time); b) is 7 UTC (9 am local time); c) is 12 UTC (2 pm local); and d) is 18 UTC (8 pm local).

5 Conclusions

This study describes the development of a coupled regional to street scale modelling system, CALIOPE-Urban, which provides high spatial and temporal resolution (up to 10 m × 10 m, hourly) NO_2 concentrations for Barcelona. It couples the mesoscale air

R-LINE. For each regional 1 km × 1 km grid cell, meteorological data from WRF and background concentrations from CMAQ are used as input combined with traffic emissions from the HERMES emission model at road link level. R-LINE has been adapted to Barcelona's geometrical conditions by considering specific meteorology and background concentrations for each street. CALIOPE-Urban NO_2 simulations are compared with CALIOPE and with observations for temporal evaluation, using data from five traffic sites and one urban background site during April and May 2013, and for spatial evaluation, with NO_2 concentrations measured by 182 passive dosimeters distributed across the entire city during two weeks in February-March 2017.

CALIOPE-Urban methodology adapts dynamically to street conditions by coupling the meteorology and background using street-specific surface roughness based on urban geometry. It adapts R-LINE dispersion model to compact cities using channeled winds to drive dispersion and recalculated meteorological parameters for each street. Regarding background concentrations, it estimates over roof levels using an upwind background scheme and gives surface concentrations applying a vertical mixing parametrization based on urban geometry and atmospheric stability. The upwind background scheme avoids double counting traffic emissions in regional and dispersion models by using upwind grid cells concentrations to estimate over roof background concentrations. Doing so we omit the use of the grid cell over the estimated area, where traffic emissions are considered in the dispersion model. To estimate background concentrations at surface level, the vertical mixing parametrization enhances background mixing from the overlying atmosphere under daytime convective atmospheric conditions and limits background air mixing during nighttime (stable) hours. For the transition from urban to suburban areas, CALIOPE-Urban implements a smooth variation for wind conditions, background and total concentrations.

Temporally, CALIOPE-Urban agrees better with observations than CALIOPE at the five traffic sites evaluated, where the contribution of local emissions predominates. For the urban background site of Palau Reial, both systems give similar (good) results. For traffic sites, the coupled system shows better agreement in highly trafficked areas where local dispersion plays a crucial role. Regarding the diurnal average cycle in the observation sites, both systems follow the overall daily cycle in the observations but CALIOPE-Urban predicts better morning peaks, and corrects the afternoon levels at traffic sites as well as the systematic nighttime overestimation produced by the regional system. The vertical mixing of rooftop background concentrations to surface levels based on atmospheric stability and urban geometry appears to be a good method to correct the strong positive bias of the mesoscale model under stable atmospheric conditions during the evening.

Spatially, CALIOPE-Urban performs better than CALIOPE at the dosimeters located close to traffic. This result is because R-LINE explicitly resolves road traffic emission dispersion simulating the high gradients of NO_2 observed levels that occur within a mesoscale system grid cell. CALIOPE-Urban gives more overestimation close to high trafficked areas. This behaviour may be produced by an overestimation of traffic emissions in these roads or by underestimating dispersion. For dosimeters located more than 10 m away from traffic both systems perform reasonably well. The higher the traffic in the surrounding area, the better is CALIOPE-Urban performance compared to the regional system.

When exploring the main error sources, overall both systems produce results that are either accurate or inaccurate on the same days. This fact suggests that coupled system results are highly influenced by the regional system results. Furthermore, we find that CALIOPE-Urban gives the higher errors (i.e. stronger underestimations) under stable conditions with light winds and

quality forecasting system CALIOPE (WRF-HERMES-CMAQ-BSC-DREAM8b) with the urban roadway dispersion model, R-LINE. For each regional 1 km × 1 km grid cell, meteorological data from WRF and background concentrations from CMAQ are used as input combined with traffic emissions from the HERMES emission model at road link level. R-LINE has been adapted to Barcelona's geometrical conditions by considering specific meteorology and background concentrations for each street. CALIOPE-Urban NO_2 simulations are compared with CALIOPE and with observations for temporal evaluation, using data from five traffic sites and one urban background site during April and May 2013, and for spatial evaluation, with NO_2 concentrations measured by 182 passive dosimeters distributed across the entire city during two weeks in February-March 2017.

CALIOPE-Urban methodology adapts dynamically to street conditions by coupling the meteorology and background using street-specific surface roughness based on urban geometry. It adapts R-LINE dispersion model to compact cities using channeled winds to drive dispersion and recalculated meteorological parameters for each street. Regarding background concentrations, it estimates over roof levels using an upwind background scheme and gives surface concentrations by applying a vertical mixing parametrization based on urban geometry and atmospheric stability. The upwind background scheme avoids double counting traffic emissions in regional and dispersion models by using upwind grid cells concentrations to estimate over roof background concentrations. Doing so we omit the use of the grid cell over the estimated area, where traffic emissions are considered in the dispersion model. To estimate background concentrations at surface level, the vertical mixing parametrization enhances background mixing from the overlying atmosphere under daytime convective atmospheric conditions and limits background air mixing during nighttime (stable) hours. For the transition from urban to suburban areas, CALIOPE-Urban implements a smooth variation for wind conditions, background and total concentrations.

Temporally, CALIOPE-Urban agrees better with observations than CALIOPE at the five traffic sites evaluated, where the contribution of local emissions predominates. For the urban background site of Palau Reial, both systems give similar (good) results. For traffic sites, the coupled system shows better agreement in highly trafficked areas where local dispersion plays a crucial role. Regarding the diurnal average cycle in the observation sites, both systems follow the overall daily cycle in the observations but CALIOPE-Urban predicts better morning peaks, and corrects the afternoon levels at traffic sites as well as the systematic nighttime overestimation produced by the regional system. The vertical mixing of rooftop background concentrations to surface levels based on atmospheric stability and urban geometry appears to be a good method to correct the strong positive bias of the mesoscale model under stable atmospheric conditions during the evening.

Spatially, CALIOPE-Urban performs better than CALIOPE at the dosimeters located close to traffic. This result is because R-LINE explicitly resolves road traffic emission dispersion simulating the high gradients of NO_2 observed levels that occur within a mesoscale system grid cell. CALIOPE-Urban gives more overestimation close to high trafficked areas. This behaviour may be produced by an overestimation of traffic emissions in these roads or by underestimating dispersion. For dosimeters located more than 10 m away from traffic both systems perform reasonably well. The higher the traffic in the surrounding area, the better is CALIOPE-Urban performance compared to the regional system.

When exploring the main error sources, overall both systems produce results that are either accurate or inaccurate on the same days. This fact suggests that coupled system results are highly influenced by the regional system results. Furthermore, we

low PBL height than under more convective conditions, with stronger winds and higher PBL heights. Another potential source of uncertainty is the integration within HERMESv2.0 of COPERT IV instead of COPERT V, which considers diesel NO_X exceedances derived from diesel-gate for EURO 5 and EURO 6 diesel cars (Brown et al., 2018). In a future work, we plan to update HERMESv2.0 with the new emissions factors released by COPERT V and examine the influence of traffic emissions in CALIOPE-Urban results.

For high resolution air quality forecasts, we show that CALIOPE-Urban using either the numerical or the analytical dispersion algorithm gives good results. However, an entire city system execution using the analytical configuration takes approx. half the time compared to the numerical one. Hence, the analytical dispersion algorithm may be a suitable option for forecasting applications when sources, such as roadways, and receptors are located near the ground.

We show that traffic monitoring stations in Barcelona do not represent the highest NO_2 concentrations in the city. We find the highest levels in heavily trafficked street canyons that are not well ventilated and near highways in the city surroundings. As a consequence, we consider that additional monitoring sites located in these areas may better characterize the range of NO_2 concentration levels in Barcelona and give a better representation of human exposures.

This study has demonstrated that CALIOPE-Urban improves the accuracy of model outputs estimating NO_2 concentrations in Barcelona compared to CALIOPE. The methodology is replicable in cities where a mesoscale chemistry transport model provides NO_2 simulations if urban geometrical data is available. The next step is to implement CALIOPE-Urban in the operational forecasting system for Barcelona to provide NO_2 concentrations at street level, and explore emissions impacts due to improved NO_X emissions estimates.

Code availability. CALIOPE-Urban source code is available for non-commercial use. Contact Oriol Jorba (oriol.jorba@bsc.es) and Jaime Benavides (jaime.benavides@bsc.es) for agreement details. Observational data in this work has been provided by co-authors from Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Spain. Contact them if interested on these datasets.

Appendix A: Extended urban geometry characterization

find that CALIOPE-Urban gives the higher errors (i.e. stronger underestimations) under stable conditions with light winds and low PBL height than under more convective conditions, with stronger winds and higher PBL heights. Another potential source of uncertainty is the integration within HERMESv2.0 of COPERT IV instead of COPERT V, which considers diesel NO_X exceedances derived from diesel-gate for EURO 5 and EURO 6 diesel cars (Brown et al., 2018). In a future work, we plan to update HERMESv2.0 with the new emissions factors released by COPERT V and examine the influence of traffic emissions in CALIOPE-Urban results. Finally, we consider an additional source of uncertainty the assumption of clear-sky conditions in the photolysis rate calculation of the GRS chemistry mechanism.

For high resolution air quality forecasts, we show that CALIOPE-Urban using either the numerical or the analytical dispersion algorithm gives good results. However, an entire city system execution using the analytical configuration takes approx. half the time compared to the numerical one. Hence, the analytical dispersion algorithm may be a suitable option for forecasting applications when sources, such as roadways, and receptors are located near the ground.

We show that traffic monitoring stations in Barcelona do not represent the highest NO_2 concentrations in the city. We find the highest levels in heavily trafficked street canyons that are not well ventilated and near highways in the city surroundings. As a consequence, we consider that additional monitoring sites located in these areas may better characterize the range of NO_2 concentration levels in Barcelona and give a better representation of human exposures.

This study has demonstrated that CALIOPE-Urban improves the accuracy of model outputs estimating NO_2 concentrations in Barcelona compared to CALIOPE. The methodology is replicable in cities where a mesoscale chemistry transport model provides NO_2 simulations if urban geometrical data is available. The next step is to implement CALIOPE-Urban in the operational forecasting system for Barcelona to provide NO_2 concentrations at street level, and explore emissions impacts due to improved NO_X emissions estimates.

Code availability. Copies of the code are readily available upon request from the corresponding authors. Observational data in this work has been provided by co-authors from Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Spain. Contact them if interested on these datasets.

Appendix A: Extended urban geometry characterization

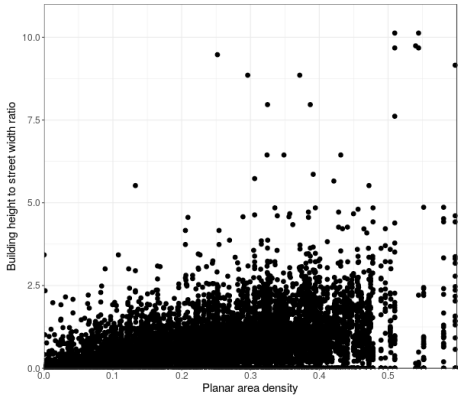

Figure A1. Scatter plot showing aspect ratio and building density relation in Barcelona city.

Appendix B: Extended performance evaluation

Table B1. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) during April and May 2013 for CALIOPE and CALIOPE-Urban daily mean concentrations. Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE	**1.00**	-10.97	15.44	1.30	1.08	**0.84**
	CALIOPE-Urban	**0.95**	**-8.72**	**13.85**	**1.29**	**1.07**	0.80
2. Eixample	CALIOPE	0.91	**-9.62**	17.37	1.24	1.08	0.52
	CALIOPE-Urban	**1.00**	9.64	**15.60**	**0.83**	**1.04**	**0.60**
3. Gràcia-Sant Gervasi	CALIOPE	0.91	-10.97	15.44	1.30	1.08	**0.82**
	CALIOPE-Urban	**1.00**	**-7.40**	**15.33**	**1.10**	**1.05**	0.81
4. 213 Industria Street	CALIOPE	0.75	-19.15	23.51	1.62	1.20	**0.68**
	CALIOPE-Urban	**0.95**	**-13.79**	**18.90**	**1.29**	**1.07**	**0.68**
5. 455 Valencia Street	CALIOPE	0.70	-22.05	26.07	1.64	1.20	**0.67**
	CALIOPE-Urban	**1.00**	**3.00**	**11.25**	**0.94**	**1.02**	0.65
6. 309 Industria Street	CALIOPE	0.95	-7.13	13.33	1.22	**1.07**	**0.78**
	CALIOPE-Urban	**0.95**	**-7.12**	**13.32**	**1.30**	**1.07**	0.68

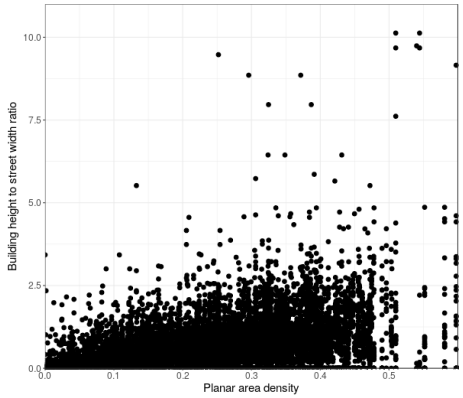

Figure A1. Scatter plot showing aspect ratio and building density relation in Barcelona city.

Appendix B: Extended performance evaluation

Table B1. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) during April and May 2013 for CALIOPE and CALIOPE-Urban daily mean concentrations. Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE	**1.00**	-10.97	15.44	1.30	1.08	**0.84**
	CALIOPE-Urban	**0.95**	**-8.72**	**13.85**	**1.29**	**1.07**	0.80
2. Eixample	CALIOPE	0.91	**-9.62**	17.37	1.24	1.08	0.52
	CALIOPE-Urban	**1.00**	9.64	**15.60**	**0.83**	**1.04**	**0.60**
3. Gràcia-Sant Gervasi	CALIOPE	0.91	-10.97	15.44	1.30	1.08	**0.82**
	CALIOPE-Urban	**1.00**	**-7.40**	**15.33**	**1.10**	**1.05**	0.81
4. 213 Industria Street	CALIOPE	0.75	-19.15	23.51	1.62	1.20	**0.68**
	CALIOPE-Urban	**0.95**	**-13.79**	**18.90**	**1.29**	**1.07**	**0.68**
5. 445 Valencia Street	CALIOPE	0.70	-22.05	26.07	1.64	1.20	**0.67**
	CALIOPE-Urban	**1.00**	**3.00**	**11.25**	**0.94**	**1.02**	0.65
6. 309 Industria Street	CALIOPE	0.95	-7.13	13.33	1.22	**1.07**	**0.78**
	CALIOPE-Urban	**0.95**	**-7.12**	**13.32**	**1.30**	**1.07**	0.68

Table B2. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) during April and May 2013 for hourly concentrations of CALIOPE-Urban, CALIOPE-Urban Analytical or CALIOPE-Urban-nl (Non Local) configurations. Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE-Urban	**0.72**	-8.70	**21.57**	1.28	**1.22**	**0.57**
	CALIOPE-Urban Analytical	0.69	-10.46	22.44	1.38	1.24	0.56
	CALIOPE-Urban-nl	0.67	**3.61**	26.34	**0.99**	1.40	0.55
2. Eixample	CALIOPE-Urban	0.86	9.38	26.70	0.83	1.12	0.55
	CALIOPE-Urban Analytical	**0.87**	**7.99**	**26.25**	**0.85**	**1.11**	**0.56**
	CALIOPE-Urban-nl	0.61	39.53	54.92	0.57	1.38	0.45
3. Gràcia-Sant Gervasi	CALIOPE-Urban	**0.79**	-7.39	**25.11**	**1.07**	**1.19**	0.52
	CALIOPE-Urban Analytical	0.78	-9.29	25.54	1.13	1.20	**0.53**
	CALIOPE-Urban-nl	0.66	**6.00**	35.91	0.91	1.43	0.38
4. 213 Industria Street	CALIOPE-Urban	**0.78**	-13.62	**26.65**	1.30	**1.17**	**0.57**
	CALIOPE-Urban Analytical	0.76	-14.85	27.22	1.35	1.18	**0.57**
	CALIOPE-Urban-nl	0.75	**1.57**	31.12	**1.05**	1.26	0.54
5. 455 Valencia Street	CALIOPE-Urban	0.92	2.92	23.26	0.94	**1.07**	0.56
	CALIOPE-Urban Analytical	**0.93**	**0.87**	**23.17**	**0.97**	**1.07**	**0.57**
	CALIOPE-Urban-nl	0.79	23.72	42.29	0.75	1.19	0.47
6. 309 Industria Street	CALIOPE-Urban	**0.84**	**-4.60**	**22.72**	**1.05**	**1.13**	0.53
	CALIOPE-Urban Analytical	0.83	-6.64	22.94	1.12	1.14	0.54
	CALIOPE-Urban-nl	0.78	11.60	31.13	0.83	1.24	**0.58**

Appendix C: Description of model evaluation statistics

Here we define the model evaluation statistics used to compare observed measurements (obs) with modelled concentrations (mod): the geometric mean bias (GeoMean), the fraction of model results within a factor of two of observations (FAC2), the geometric standard deviation (GeoSD), the correlation coefficient (R), the mean bias (MB) and the root mean square error (RMSE).

$$GeoMean = exp\left(\overline{\ln(obs)} - \overline{\ln(mod)}\right) \tag{C1}$$

$$FAC2 = 0.5 <= \frac{mod_i}{obs_i} <= 2.0 \tag{C2}$$

$$GeoSD = exp\left(\frac{\ln(F)}{\sqrt{2}\, eri^{-1}\left(A_F\right)}\right) \tag{C3}$$

$$R = \frac{\overline{\left(obs_i - \overline{obs}\right)\left(mod_i - \overline{mod}\right)}}{\sigma mod\, \sigma obs} \tag{C4}$$

$$MB = \frac{1}{n}\sum_{i=1}^{n} mod_i - obs_i \tag{C5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(mod_i - obs_i)^2}{n}} \tag{C6}$$

Table B2. NO_2 model evaluation statistics calculated at six sites (described in Sect. 3) during April and May 2013 for hourly concentrations of CALIOPE-Urban, CALIOPE-Urban Analytical or CALIOPE-Urban-nl (Non Local) configurations. Bold numbers represent model results with better performance for each statistic and site.

Site	Method	FAC2	MB	RMSE	GeoMean	GeoSD	r
1. Palau Reial	CALIOPE-Urban	**0.72**	-8.70	**21.57**	1.28	**1.22**	**0.57**
	CALIOPE-Urban Analytical	0.69	-10.46	22.44	1.38	1.24	0.56
	CALIOPE-Urban-nl	0.67	**3.61**	26.34	**0.99**	1.40	0.55
2. Eixample	CALIOPE-Urban	0.86	9.38	26.70	0.83	1.12	0.55
	CALIOPE-Urban Analytical	**0.87**	**7.99**	**26.25**	**0.85**	**1.11**	**0.56**
	CALIOPE-Urban-nl	0.61	39.53	54.92	0.57	1.38	0.45
3. Gràcia-Sant Gervasi	CALIOPE-Urban	**0.79**	-7.39	**25.11**	**1.07**	**1.19**	0.52
	CALIOPE-Urban Analytical	0.78	-9.29	25.54	1.13	1.20	**0.53**
	CALIOPE-Urban-nl	0.66	**6.00**	35.91	0.91	1.43	0.38
4. 213 Industria Street	CALIOPE-Urban	**0.78**	-13.62	**26.65**	1.30	**1.17**	**0.57**
	CALIOPE-Urban Analytical	0.76	-14.85	27.22	1.35	1.18	**0.57**
	CALIOPE-Urban-nl	0.75	**1.57**	31.12	**1.05**	1.26	0.54
5. 445 Valencia Street	CALIOPE-Urban	0.92	2.92	23.26	0.94	**1.07**	0.56
	CALIOPE-Urban Analytical	**0.93**	**0.87**	**23.17**	**0.97**	**1.07**	**0.57**
	CALIOPE-Urban-nl	0.79	23.72	42.29	0.75	1.19	0.47
6. 309 Industria Street	CALIOPE-Urban	**0.84**	**-4.60**	**22.72**	**1.05**	**1.13**	0.53
	CALIOPE-Urban Analytical	0.83	-6.64	22.94	1.12	1.14	0.54
	CALIOPE-Urban-nl	0.78	11.60	31.13	0.83	1.24	**0.58**

Appendix C: Description of model evaluation statistics

Here we define the model evaluation statistics used to compare observed measurements (obs) with modelled concentrations (mod): the geometric mean bias (GeoMean), the fraction of model results within a factor of two of observations (FAC2), the geometric standard deviation (GeoSD), the correlation coefficient (R), the mean bias (MB) and the root mean square error (RMSE).

$$GeoMean = exp\left(\overline{\ln(obs)} - \overline{\ln(mod)}\right) \tag{C1}$$

$$FAC2 = 0.5 <= \frac{mod_i}{obs_i} <= 2.0 \tag{C2}$$

$$GeoSD = exp\left(\frac{\ln(F)}{\sqrt{2}\, eri^{-1}\left(A_F\right)}\right) \tag{C3}$$

$$R = \frac{\overline{\left(obs_i - \overline{obs}\right)\left(mod_i - \overline{mod}\right)}}{\sigma mod\, \sigma obs} \tag{C4}$$

$$MB = \frac{1}{n}\sum_{i=1}^{n} mod_i - obs_i \tag{C5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(mod_i - obs_i)^2}{n}} \tag{C6}$$

where, mod are modelled concentations; obs are observed concentrations; overbar ($\overline{d}$) represents the average over a dataset d; F is considered to be 2; eri is the inverse of error function; A_F is the proportion of the ratio; σd is the standard deviation of d; n is the number of paired modelled and observed concentrations and subscripts represent a value between one and n. For further details on the evaluation statistics we refer to Chang and Hanna (2004).

Author contributions. Author contributions: JB developed the code. JB, MS, OJ designed the research. FA and XQ provided the observational data. CP, MG, FA, XQ and AS contributed to the discussion of the results. JB wrote the original paper, and all authors contributed to the review and editing of the paper.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. This work was started while Dr. Michelle Snyder was a researcher at the University of North Carolina Institute for the Environment. Dr. Snyder's efforts throughout the project were completely voluntary.

Acknowledgements. BSC researchers acknowledge the grants CGL2013-46736-R, CGL2016-75725-R and COMRDI15-1-0011-04 of the Spanish Government. J. Benavides PhD work is funded with the grant BES-2014-070637 from the FPI Programme by the Spanish Ministry of the Economy and Competitiveness. J. Benavides developed part of this work as research visitor at the Institute for the Environment at UNC funded with the mobility grant EEBB-I-17-12296 by the same Ministry. IDAEA-CSIC acknowledges the Barcelona City Council for the support to the experimental campaign. Carlos Pérez García-Pando acknowledges long-term support from the AXA Research Fund, as well as the support received through the Ramón y Cajal programme (grant RYC-2015-18690) of the Spanish Ministry of Economy and Competitiveness.

References

Amato, F., Karanasiou, A., Cordoba, P., Alastuey, A., Moreno, T., Lucarelli, F., Nava, S., Calzolai, G., and Querol, X. "Effects of Road Dust Suppressants on PM Levels in a Mediterrazean Urban Area". *Environmental Science & Technology* 48.14, pp. 8069–8077. DOI: https://doi.org/10.1021/es502496s. 2014.

Arunachalam, S., Valencia, A., Akita, Y., Serre, M., Omary, M., Garcia, V., and Isakov, V. "A Method for Estimating Urban Background Concentrations in Support of Hybrid Air Pollution Modeling for Environmental Health Studies". *International Journal of Environmental Research and Public Health* 11.10, pp. 10518–10536. ISSN: 1660-4601. DOI: https://doi.org/10.3390/ijerph111010518. 2014.

ASPB. *Avaluació de la qualitat de l'aire a la ciutat de Barcelona 2016*. Barcelona, 2017.

Baldasano, J.M., Pay, M.T., Jorba, O., Gassó, S., and Jiménez-Guerrero, P. "An annual assessment of air quality with the CALIOPE modeling system over Spain". *Science of the Total Environment* 409.11, pp. 2163–2178. DOI: https://doi.org/10.1016/j.scitotenv.2011.01.041. 2011.

where, mod are modelled concentations; obs are observed concentrations; overbar ($\overline{d}$) represents the average over a dataset d; F is considered to be 2; eri is the inverse of error function; A_F is the proportion of the ratio; σd is the standard deviation of d; n is the number of paired modelled and observed concentrations and subscripts represent a value between one and n. For further details on the evaluation statistics we refer to Chang and Hanna (2004).

Author contributions. Author contributions: JB developed the code. JB, MS, OJ designed the research. FA and XQ provided the observational data. CP, MG, FA, XQ and AS contributed to the discussion of the results. JB wrote the original paper, and all authors contributed to the review and editing of the paper.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. This work was started while Dr. Michelle Snyder was a researcher at the University of North Carolina Institute for the Environment. Dr. Snyder's efforts throughout the project were completely voluntary.

Acknowledgements. BSC researchers acknowledge the grants CGL2013-46736-R and CGL2016-75725-R of the Spanish Ministry of the Economy and Competitiveness. J. Benavides PhD work is funded with the grant BES-2014-070637 from the FPI Programme by the Spanish Ministry of the Economy and Competitiveness. J. Benavides developed part of this work as research visitor at the Institute for the Environment at UNC funded with the mobility grant EEBB-I-17-12296 by the same Ministry. IDAEA-CSIC acknowledges the Barcelona City Council for the support to the experimental campaign. Carlos Pérez García-Pando acknowledges long-term support from the AXA Research Fund, as well as the support received through the Ramón y Cajal programme (grant RYC-2015-18690) of the Spanish Ministry of Economy and Competitiveness.

References

Amato, F., Karanasiou, A., Cordoba, P., Alastuey, A., Moreno, T., Lucarelli, F., Nava, S., Calzolai, G., and Querol, X. "Effects of Road Dust Suppressants on PM Levels in a Mediterrazean Urban Area". *Environmental Science & Technology* 48.14, pp. 8069–8077. DOI: https://doi.org/10.1021/es502496s. 2014.

Arunachalam, S., Valencia, A., Akita, Y., Serre, M., Omary, M., Garcia, V., and Isakov, V. "A Method for Estimating Urban Background Concentrations in Support of Hybrid Air Pollution Modeling for Environmental Health Studies". *International Journal of Environmental Research and Public Health* 11.10, pp. 10518–10536. ISSN: 1660-4601. DOI: https://doi.org/10.3390/ijerph111010518. 2014.

ASPB. *Avaluació de la qualitat de l'aire a la ciutat de Barcelona 2016*. Barcelona, 2017.

Baldasano, J.M., Pay, M.T., Jorba, O., Gassó, S., and Jiménez-Guerrero, P. "An annual assessment of air quality with the CALIOPE modeling system over Spain". *Science of the Total Environment* 409.11, pp. 2163–2178. DOI: https://doi.org/10.1016/j.scitotenv.2011.01.041. 2011.

Baldasano, J.M., Soret, A., Guevara, M., Martınez, F., and Gassó, S. “Integrated assessment of air pollution using observations and modelling in Santa Cruz de Tenerife (Canary Islands)”. *Science of The Total Environment* 473-474, pp. 576–588. DOI: https://doi.org/10.1016/j.scitotenv.2013.12.062. 2014.

Barcelona City Council. *Caracterització dels vehicles i les seves emissions a l'àrea metropolitana de Barcelona*. Tech. rep. 2017.

– *CartoBCN v.1.1.0*. 2016. URL: http://w20.bcn.cat/cartobcn/.

– *Evaluation of the NOX and PM10 emission reductions from traffic in Barcelona city based on the characterization of the vehicle pool*. Tech. rep. 2010. URL: https://w110.bcn.cat/MediAmbient/Continguts/Documents/Documentacio/4-AvaluacioEmissionsParcMobilBCN.pdf.

Barone-Adesi, F., Dent, J. E, Dajnak, D., Beevers, S., Anderson, H.R., Kelly, Frank J, Cook, Derek G, and Whincup, Peter H. “Long-Term Exposure to Primary Traffic Pollutants and Lung Function in Children: Cross-Sectional Study and Meta-Analysis.” *PloS one* 10.11, e0142565. DOI: https://doi.org/10.1371/journal.pone.0142565. 2015.

Basart, S., Pérez, C., Nickovic, S., Cuevas, E., and Baldasano, J.M. “Development and evaluation of the BSC-DREAM8b dust regional model over northern Africa, the mediterranean and the middle east”. *Tellus, Series B: Chemical and Physical Meteorology* 64.1, pp. 1–23. DOI: https://doi.org/10.3402/tellusb.v64i0.18539. 2012.

Bechtel, B., Alexander, P., Böhner, J., Ching, J., Conrad, O., Feddema, Johannes, Mills, Gerald, See, Linda, and Stewart, Iain. “Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities”. *ISPRS International Journal of Geo-Information* 4.1, pp. 199–219. DOI: https://doi.org/10.3390/ijgi4010199. 2015.

Berkowicz, R. “A simple model for urban background pollution”. *Environmental Monitoring and Assessment* 65.1-2, pp. 259–267. DOI: https://doi.org/10.1023/A:1006466025186. 2000.

Borge, R., Lumbreras, J., Pérez, J., De la Paz, D., Vedrenne, M., Andrés, Juan Manuel de, and Rodrıguez, Mª Encarnación. “Emission inventories and modeling requirements for the development of air quality plans. Application to Madrid (Spain)”. *Science of The Total Environment* 466-467, pp. 809–819. DOI: https://doi.org/10.1016/j.scitotenv.2013.07.093. 2014.

Brousse, O., Martilli, A., Foley, M., Mills, G., and Bechtel, B. “WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid”. *Urban Climate* 17, pp. 116–134. DOI: https://doi.org/10.1016/j.uclim.2016.04.001. 2016.

Brown, P., Wakeling, D., Pang, Y., and Murrells, T. *Methodology for the UK ' s Road Transport Emissions Inventory*. Tech. rep. ED59803130. 1. United Kingdom Government. Department for Business, Energy & Industrial Strategy, 2018.

Byun, D.W. and Schere, K.L. “Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system”. *Applied Mechanics Reviews* 59, pp. 51–77. DOI: https://doi.org/10.1115/1.2128636. 2006.

Carslaw, D.C. and Beevers, S.D. “Investigating the potential importance of primary NO2emissions in a street canyon”. *Atmospheric Environment* 38.22, pp. 3585–3594. DOI: https://doi.org/10.1016/j.atmosenv.2004.03.041. 2004.

Carslaw, D.C., Murrells, Tim P, and Keenan, Matthew. “Have vehicle emissions of primary NO 2 peaked ?” *Faraday Discussions* 189, pp. 439–454. DOI: https://doi.org/10.1039/c5fd00162e. 2016.

Chang, J.C. and Hanna, S.R. “Air quality model performance evaluation”. *Meteorology and Atmospheric Physics* 87.1-3, pp. 167–196. DOI: https://doi.org/10.1007/s00703-003-0070-7. URL: http://link.springer.com/10.1007/s00703-003-0070-7, 2004.

Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., and Brode, R.W. “AERMOD : A Dispersion Model for Industrial Source Applications . Part I : General Model Formulation and Boundary Layer Characterization”. *Journal of Applied Meteorology* 44, pp. 682–693. DOI: https://doi.org/10.1175/JAM2227.1. 2005.

Baldasano, J.M., Soret, A., Guevara, M., Martınez, F., and Gassó, S. “Integrated assessment of air pollution using observations and modelling in Santa Cruz de Tenerife (Canary Islands)”. *Science of The Total Environment* 473-474, pp. 576–588. DOI: https://doi.org/10.1016/j.scitotenv.2013.12.062. 2014.

Barcelona City Council. *Caracterització dels vehicles i les seves emissions a l'àrea metropolitana de Barcelona*. Tech. rep. 2017.

– *CartoBCN v.1.1.0*. 2016. URL: http://w20.bcn.cat/cartobcn/.

– *Evaluation of the NOX and PM10 emission reductions from traffic in Barcelona city based on the characterization of the vehicle pool*. Tech. rep. 2010. URL: https://w110.bcn.cat/MediAmbient/Continguts/Documents/Documentacio/4-AvaluacioEmissionsParcMobilBCN.pdf.

Barone-Adesi, F., Dent, J. E, Dajnak, D., Beevers, S., Anderson, H.R., Kelly, Frank J, Cook, Derek G, and Whincup, Peter H. “Long-Term Exposure to Primary Traffic Pollutants and Lung Function in Children: Cross-Sectional Study and Meta-Analysis.” *PloS one* 10.11, e0142565. DOI: https://doi.org/10.1371/journal.pone.0142565. 2015.

Basart, S., Pérez, C., Nickovic, S., Cuevas, E., and Baldasano, J.M. “Development and evaluation of the BSC-DREAM8b dust regional model over northern Africa, the mediterranean and the middle east”. *Tellus, Series B: Chemical and Physical Meteorology* 64.1, pp. 1–23. DOI: https://doi.org/10.3402/tellusb.v64i0.18539. 2012.

Bechtel, B., Alexander, P., Böhner, J., Ching, J., Conrad, O., Feddema, Johannes, Mills, Gerald, See, Linda, and Stewart, Iain. “Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities”. *ISPRS International Journal of Geo-Information* 4.1, pp. 199–219. DOI: https://doi.org/10.3390/ijgi4010199. 2015.

Beevers, S., Kitwiroon, N., Williams, M.L., and Carslaw, D.C. “One way coupling of CMAQ and a road source dispersion model for fine scale air pollution predictions”. *Atmospheric Environment* 59, pp. 47–58. DOI: https://doi.org/10.1016/j.atmosenv.2012.05.034. 2012.

Berkowicz, R. “A simple model for urban background pollution”. *Environmental Monitoring and Assessment* 65.1-2, pp. 259–267. DOI: https://doi.org/10.1023/A:1006466025186. 2000.

Borge, R., Lumbreras, J., Pérez, J., De la Paz, D., Vedrenne, M., Andrés, Juan Manuel de, and Rodrıguez, Mª Encarnación. “Emission inventories and modeling requirements for the development of air quality plans. Application to Madrid (Spain)”. *Science of The Total Environment* 466-467, pp. 809–819. DOI: https://doi.org/10.1016/j.scitotenv.2013.07.093. 2014.

Brousse, O., Martilli, A., Foley, M., Mills, G., and Bechtel, B. “WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid”. *Urban Climate* 17, pp. 116–134. DOI: https://doi.org/10.1016/j.uclim.2016.04.001. 2016.

Brown, P., Wakeling, D., Pang, Y., and Murrells, T. *Methodology for the UK ' s Road Transport Emissions Inventory*. Tech. rep. ED59803130. 1. United Kingdom Government. Department for Business, Energy & Industrial Strategy, 2018.

Byun, D.W. and Schere, K.L. “Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system”. *Applied Mechanics Reviews* 59, pp. 51–77. DOI: https://doi.org/10.1115/1.2128636. 2006.

Carslaw, D.C. and Beevers, S. “Investigating the potential importance of primary NO2emissions in a street canyon”. *Atmospheric Environment* 38.22, pp. 3585–3594. DOI: https://doi.org/10.1016/j.atmosenv.2004.03.041. 2004.

Carslaw, D.C., Murrells, Tim P, and Keenan, Matthew. “Have vehicle emissions of primary NO2 peaked ?” *Faraday Discussions* 189, pp. 439–454. DOI: https://doi.org/10.1039/c5fd00162e. 2016.

Chang, J.C. and Hanna, S.R. “Air quality model performance evaluation”. *Meteorology and Atmospheric Physics* 87.1-3, pp. 167–196. DOI: https://doi.org/10.1007/s00703-003-0070-7. URL: http://link.springer.com/10.1007/s00703-003-0070-7, 2004.

Clapp, L.J. and Jenkin, M.E. “Analysis of the relationship between ambient levels of O3, NO2 and NO as a function of NOx in the UK”. *Atmospheric Environment* 35.36, pp. 6391–6405. DOI: https://doi.org/10.1016/S1352-2310(01)00378-8. 2001.

Degraeuwe, B., Thunis, P., Clappier, A., Weiss, M., Lefebvre, W., Janssen, Stijn, and Vranckx, Stijn. “Impact of passenger car NOx emissions on urban NO2 pollution - Scenario analysis for 8 European cities”. *Atmospheric Environment* 171, pp. 21330–337. DOI: https://doi.org/10.1016/j.atmosenv.2017.10.040. 2017.

Duyzer, J., Hout, D. van den, Zandveld, P., and Ratingen, S. van. “Representativeness of air quality monitoring networks”. *Atmospheric Environment* 104, pp. 88–101. DOI: https://doi.org/10.1016/j.atmosenv.2014.12.067. 2015.

EEA. *Air quality in Europe — 2018 report (NO 12 / 2018)*. 12. 2018. DOI: https://doi.org/10.2800/777411.

Ferreira, J., Guevara, M., Baldasano, J. M., Tchepel, O., Schaap, M., Miranda, A. I., and Borrego, C. “A comparative analysis of two highly spatially resolved European atmospheric emission inventories”. *Atmospheric Environment* 75, pp. 43–57. DOI: https://doi.org/10.1016/j.atmosenv.2013.03.052. 2013.

Fisher, B., Kukkonen, J., Piringer, M., Rotach, M.W., and Schatzmann, M. “Meteorology Applied to Urban Air Pollution Problems - Final Report COST Action 715”. 2005.

Grimmond, C.S.B. and Oke, T.R. “Aerodynamic Properties of Urban Areas Derived from Analysis of Surface Form”. *Journal of Applied Meteorology* 38, pp. 1262–1292. DOI: https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2. 1999.

Guevara, M., Lopez-Aparicio, S., Cuvelier, C., Tarrason, L., Clappier, A., and Thunis, P. “A benchmarking tool to screen and compare bottom-up and top-down atmospheric emission inventories”. *Air Quality, Atmosphere and Health* 10.5, pp. 627–642. DOI: https://doi.org/10.1007/s11869-016-0456-6. 2017.

Guevara, M., Martınez, F., Arévalo, G., Gassó, S., and Baldasano, J.M. “An improved system for modelling Spanish emissions: HERMESv2.0”. *Atmospheric Environment* 81, pp. 209–221. DOI: https://doi.org/10.1016/j.atmosenv.2013.08.053. 2013.

Hood, C., Mackenzie, I., Stocker, J., Johnson, K., Carruthers, D., Vieno, M., and Doherty, R. “Air quality simulations for London using a coupled regional-to-local modelling system”. *Atmospheric Chemistry and Physics Discussions* February, pp. 1–44. DOI: https://doi.org/10.5194/acp-18-11221-2018. 2018.

Isakov, V. et al. “Air Quality Modeling in Support of the Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS)”. *International Journal of Environmental Research and Public Health* 11.9, pp. 8777–8793. DOI: https://doi.org/10.3390/ijerph110908777. 2014.

Janssen, S, Guerreiro, C., Viaene, P., Georgieva, E., and Thunis, P. *Guidance Document on Modelling Quality Objectives Benchmarking, Version 2.1, Forum for air quality modelling in Europe*. Tech. rep. February. 2017, pp. 1–56. URL: http://fairmode.jrc.ec.europa.eu/document/fairmode/WG1/Guidance%7B%5C_%7DMQO%7B%5C_%7DBench%7B%5C_%7Dvs2.2.pdf.

Jensen, S. S., Ketzel, M., Becker, T., Christensen, J., Brandt, J., Plejdrup, M., Winther, M., Nielsen, O.K., Hertel, O., and Ellermann, T. “High resolution multi-scale air quality modelling for all streets in Denmark”. *Transportation Research Part D: Transport and Environment* 52, pp. 322–339. DOI: https://doi.org/10.1016/j.trd.2017.02.019. 2017.

Jorba, O. et al. “The DAURE field campaign: meteorological overview”. *Atmospheric Chemistry and Physics Discussions* 11, pp. 4953–5001. DOI: https://doi.org/10.5194/acpd-11-4953-2011. 2011.

Kanda, M., Inagaki, A., Miyamoto, T., Gryschka, M., and Raasch, S. “A New Aerodynamic Parametrization for Real Urban Surfaces”. *Boundary-Layer Meteorology* 148.2, pp. 357–377. DOI: https://doi.org/10.1007/s10546-013-9818-x. 2013.

Kastner-Klein, P, Fedorovich, E, and Rotach, M W. “A wind tunnel study of organised and turbulent air motions in urban street canyons”. *Journal of Wind Engineering and Industrial Aerodynamics* 89, pp. 849–861. DOI: https://doi.org/10.1016/S0167-6105(01)00074-5. 2001.

Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., and Brode, R.W. “AERMOD : A Dispersion Model for Industrial Source Applications . Part I : General Model Formulation and Boundary Layer Characterization”. *Journal of Applied Meteorology* 44, pp. 682–693. DOI: https://doi.org/10.1175/JAM2227.1. 2005.
Clapp, L.J. and Jenkin, M.E. “Analysis of the relationship between ambient levels of O3, NO2 and NO as a function of NOx in the UK”. *Atmospheric Environment* 35.36, pp. 6391–6405. DOI: https://doi.org/10.1016/S1352-2310(01)00378-8. 2001.

Degraeuwe, B., Thunis, P., Clappier, A., Weiss, M., Lefebvre, W., Janssen, Stijn, and Vranckx, Stijn. “Impact of passenger car NOx emissions on urban NO2 pollution - Scenario analysis for 8 European cities”. *Atmospheric Environment* 171, pp. 21330–337. DOI: https://doi.org/10.1016/j.atmosenv.2017.10.040. 2017.

Duyzer, J., Hout, D. van den, Zandveld, P., and Ratingen, S. van. “Representativeness of air quality monitoring networks”. *Atmospheric Environment* 104, pp. 88–101. DOI: https://doi.org/10.1016/j.atmosenv.2014.12.067. 2015.

EEA. *Air quality in Europe — 2018 report (NO 12 / 2018)*. 12. 2018. DOI: https://doi.org/10.2800/777411.
Fagerli, H., Denby, B., and Wind, P. “Assessment of LRT contribution to cities in Europe using uEMEP?” 2019.
Ferreira, J., Guevara, M., Baldasano, J. M., Tchepel, O., Schaap, M., Miranda, A. I., and Borrego, C. “A comparative analysis of two highly spatially resolved European atmospheric emission inventories”. *Atmospheric Environment* 75, pp. 43–57. DOI: https://doi.org/10.1016/j.atmosenv.2013.03.052. 2013.

Fisher, B., Kukkonen, J., Piringer, M., Rotach, M.W., and Schatzmann, M. “Meteorology Applied to Urban Air Pollution Problems - Final Report COST Action 715”. 2005.

Grimmond, C.S.B. and Oke, T.R. “Aerodynamic Properties of Urban Areas Derived from Analysis of Surface Form”. *Journal of Applied Meteorology* 38, pp. 1262–1292. DOI: https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2. 1999.

Guevara, M., Lopez-Aparicio, S., Cuvelier, C., Tarrason, L., Clappier, A., and Thunis, P. “A benchmarking tool to screen and compare bottom-up and top-down atmospheric emission inventories”. *Air Quality, Atmosphere and Health* 10.5, pp. 627–642. DOI: https://doi.org/10.1007/s11869-016-0456-6. 2017.

Guevara, M., Martınez, F., Arévalo, G., Gassó, S., and Baldasano, J.M. “An improved system for modelling Spanish emissions: HERMESv2.0”. *Atmospheric Environment* 81, pp. 209–221. DOI: https://doi.org/10.1016/j.atmosenv.2013.08.053. 2013.

Hood, C., Mackenzie, I., Stocker, J., Johnson, K., Carruthers, D., Vieno, M., and Doherty, R. “Air quality simulations for London using a coupled regional-to-local modelling system”. *Atmospheric Chemistry and Physics Discussions* February, pp. 1–44. DOI: https://doi.org/10.5194/acp-18-11221-2018. 2018.

Isakov, V. et al. “Air Quality Modeling in Support of the Near-Road Exposures and Effects of Urban Air Pollutants Study (NEXUS)”. *International Journal of Environmental Research and Public Health* 11.9, pp. 8777–8793. DOI: https://doi.org/10.3390/ijerph110908777. 2014.

Janssen, S, Guerreiro, C., Viaene, P., Georgieva, E., and Thunis, P. *Guidance Document on Modelling Quality Objectives Benchmarking, Version 2.1, Forum for air quality modelling in Europe*. Tech. rep. February. 2017, pp. 1–56. URL: http://fairmode.jrc.ec.europa.eu/document/fairmode/WG1/Guidance%7B%5C_%7DMQO%7B%5C_%7DBench%7B%5C_%7Dvs2.2.pdf.

Jensen, S. S., Ketzel, M., Becker, T., Christensen, J., Brandt, J., Plejdrup, M., Winther, M., Nielsen, O.K., Hertel, O., and Ellermann, T. “High resolution multi-scale air quality modelling for all streets in Denmark”. *Transportation Research Part D: Transport and Environment* 52, pp. 322–339. DOI: https://doi.org/10.1016/j.trd.2017.02.019. 2017.

Jorba, O. et al. “The DAURE field campaign: meteorological overview”. *Atmospheric Chemistry and Physics Discussions* 11, pp. 4953–5001. DOI: https://doi.org/10.5194/acpd-11-4953-2011. 2011.

33

Kent, C.W., Grimmond, S., Barlow, J., Gatey, D., Kotthaus, S., Lindberg, F., and Halios, C.H. “Evaluation of Urban Local-Scale Aerodynamic Parameters: Implications for the Vertical Profile of Wind Speed and for Source Areas”. *Boundary-Layer Meteorol* 164.2, pp. 183–213. DOI: https://doi.org/10.1007/s10546-017-0248-z. URL: https://link.springer.com/content/pdf/10.1007%7B%5C%%7D2Fs10546-017-0248-z.pdf, 2017.

Kochanski, A.K., Pardyjak, E.R., Stoll, R., Gowardhan, A., Brown, M.J., and Steenburgh, W.J. “One-way coupling of the WRF-QUIC Urban dispersion modeling system”. *Journal of Applied Meteorology and Climatology* 54.10, pp. 2119–2139. DOI: https://doi.org/10.1175/JAMC-D-15-0020.1. 2015.

Lefebvre, W. et al. “Validation of the MIMOSA-AURORA-IFDM model chain for policy support: Modeling concentrations of elemental carbon in Flanders”. *Atmospheric Environment*. DOI: https://doi.org/10.1016/j.atmosenv.2011.08.033. 2011.

Macdonald, R.W., Griffiths, R.F., and Hall, D.J. “An improved method for the estimation of surface roughness of obstacle arrays”. *Atmospheric Environment* 32.11, pp. 1857–1864. DOI: https://10.1016/S1352-2310(97)00403-2. 1998.

Maiheu, B., Lefebvre, W., Walton, H., Dajnak, D., Janssen, S., Williams, M., Blyth, L., and Beevers, S. *Improved Methodologies for NO2 Exposure Assessment in the EU*. Tech. rep. 2. VITO, 2017. URL: http://ec.europa.eu/environment/air/publications/models.htm.

Martilli, A., Clappier, A., and Rotach, M.W. “An urban surface exchange parameterisation for mesoscale models”. *Boundary-Layer Meteorology*, pp. 261–304. DOI: https://doi.org/10.1023/A:1016099921195. 2002.

Monin, A.S. and Obukhov, A.M. “Osnovnye zakonomernosti turbulentnogo peremeshivanija v prizemnom sloe atmosfery (Basic Laws of Turbulent Mixing in the Atmosphere Near the Ground)”. *Trudy geofiz. inst. AN SSSR* 24.151, pp. 163–187. 1954.

Moreno-Garcia, M.C. “Intensity and form of the urban heat island in barcelona”. *International Journal of Climatology* 14.6, pp. 705–710. ISSN: 10970088. DOI: https://10.1002/joc.3370140609. 1994.

Oke, T.R. “Street design and urban canopy layer climate”. *Energy and Buildings* 11.1-3, pp. 103–113. DOI: https://10.1016/0378-7788(88)90026-6. 1988.

Pay, M.T., Martınez, F., Guevara, M., and Baldasano, J.M. “Air quality forecasts on a kilometer-scale grid over complex Spanish terrains”. *Geoscientific Model Development* 7.5, pp. 1979–1999. DOI: https://doi.org/10.5194/gmd-7-1979-2014. 2014.

QGIS Development Team. *QGIS Geographic Information System*. Tech. rep. Open Source Geospatial Foundation, 2017. URL: http://qgis.osgeo.org.

Skamarock, W.C. and Klemp, J.B. “A time-split nonhydrostatic atmospheric model for weather research and forecasting applications”. *Journal of Computational Physics* 227.7, pp. 3465–3485. DOI: https://doi.org/10.1016/j.jcp.2007.01.037. 2008.

Snyder, M.G. and Heist, D.K. *User’s Guide for R-LINE Model Version 1.2 A Research LINE source model for near-surface releases*. Tech. rep. USEPA, 2013. URL: https://www.cmascenter.org/r-line/documentation/1.2/RLINE%7B%5C_%7DUserGuide%7B%5C_%7D11-13-2013.pdf.

Snyder, M.G., Venkatram, A., Heist, D.K., Perry, S.G., Petersen, W.B., and Isakov, V. “RLINE: a line source dispersion model for near-surface releases”. *Atmospheric Environment* 77, pp. 748–756. DOI: https://doi.org/10.1016/j.atmosenv.2013.05.074. 2013.

Soret, A., Guevara, M., and Baldasano, J.M. “The potential impacts of electric vehicles on air quality in the urban areas of Barcelona and Madrid (Spain)”. *Atmospheric Environment* 99, pp. 51–63. DOI: https://doi.org/10.1016/j.atmosenv.2014.09.048. 2014.

Soulhac, L., Perkins, R.J., and Salizzoni, P. “Flow in a street canyon for any external wind direction”. *Boundary-Layer Meteorology* 126.3, pp. 365–388. DOI: https://10.1007/s10546-007-9238-x. 2008.

Soulhac, L., Salizzoni, P., Cierco, F. X., and Perkins, R. “The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model”. *Atmospheric Environment* 45.39, pp. 7379–7395. DOI: https://10.1016/j.atmosenv.2011.07.008. 2011.

Kanda, M., Inagaki, A., Miyamoto, T., Gryschka, M., and Raasch, S. “A New Aerodynamic Parametrization for Real Urban Surfaces”. *Boundary-Layer Meteorology* 148.2, pp. 357–377. DOI: https://doi.org/10.1007/s10546-013-9818-x. 2013.

Kastner-Klein, P, Fedorovich, E, and Rotach, M W. “A wind tunnel study of organised and turbulent air motions in urban street canyons”. *Journal of Wind Engineering and Industrial Aerodynamics* 89, pp. 849–861. DOI: https://doi.org/10.1016/S0167-6105(01)00074-5. 2001.

Kent, C.W., Grimmond, S., Barlow, J., Gatey, D., Kotthaus, S., Lindberg, F., and Halios, C.H. “Evaluation of Urban Local-Scale Aerodynamic Parameters: Implications for the Vertical Profile of Wind Speed and for Source Areas”. *Boundary-Layer Meteorol* 164.2, pp. 183–213. DOI: https://doi.org/10.1007/s10546-017-0248-z. URL: https://link.springer.com/content/pdf/10.1007%7B%5C%%7D2Fs10546-017-0248-z.pdf, 2017.

Kim, Y., Wu, Y., Seigneur, C., and Roustan, Y. “Multi-scale modeling of urban air pollution : development and application of a Street-in-Grid model by coupling MUNICH and”. *Geoscientific Model Development Discussions* September, pp. 1–24. DOI: https://doi.org/10.5194/gmd-11-611-2018. 2018.

Kochanski, A.K., Pardyjak, E.R., Stoll, R., Gowardhan, A., Brown, M.J., and Steenburgh, W.J. “One-way coupling of the WRF-QUIC Urban dispersion modeling system”. *Journal of Applied Meteorology and Climatology* 54.10, pp. 2119–2139. DOI: https://doi.org/10.1175/JAMC-D-15-0020.1. 2015.

Lefebvre, W. et al. “Validation of the MIMOSA-AURORA-IFDM model chain for policy support: Modeling concentrations of elemental carbon in Flanders”. *Atmospheric Environment*. DOI: https://doi.org/10.1016/j.atmosenv.2011.08.033. 2011.

Macdonald, R.W., Griffiths, R.F., and Hall, D.J. “An improved method for the estimation of surface roughness of obstacle arrays”. *Atmospheric Environment* 32.11, pp. 1857–1864. DOI: https://10.1016/S1352-2310(97)00403-2. 1998.

Maiheu, B., Lefebvre, W., Walton, H., Dajnak, D., Janssen, S., Williams, M., Blyth, L., and Beevers, S. *Improved Methodologies for NO2 Exposure Assessment in the EU*. Tech. rep. 2. VITO, 2017. URL: http://ec.europa.eu/environment/air/publications/models.htm.

Martilli, A., Clappier, A., and Rotach, M.W. “An urban surface exchange parameterisation for mesoscale models”. *Boundary-Layer Meteorology*, pp. 261–304. DOI: https://doi.org/10.1023/A:1016099921195. 2002.

Monin, A.S. and Obukhov, A.M. “Osnovnye zakonomernosti turbulentnogo peremeshivanija v prizemnom sloe atmosfery (Basic Laws of Turbulent Mixing in the Atmosphere Near the Ground)”. *Trudy geofiz. inst. AN SSSR* 24.151, pp. 163–187. 1954.

Moreno-Garcia, M.C. “Intensity and form of the urban heat island in barcelona”. *International Journal of Climatology* 14.6, pp. 705–710. DOI: https://10.1002/joc.3370140609. 1994.

Moussafir, J, Olry, C, Nibart, M, Albergel, A, Armand, P, Duchenne, C, and Thobois, L. “Aircity: a very high resolution atmospheric dispersion modeling system for Paris”. *American Society of Mechanical Engineers, Fluids Engineering Division (Publication) FEDSM* 1. DOI: https://doi.org/10.1115/FEDSM2014-21820. 2014.

Oke, T.R. “Street design and urban canopy layer climate”. *Energy and Buildings* 11.1-3, pp. 103–113. DOI: https://10.1016/0378-7788(88)90026-6. 1988.

Pay, M.T., Martınez, F., Guevara, M., and Baldasano, J.M. “Air quality forecasts on a kilometer-scale grid over complex Spanish terrains”. *Geoscientific Model Development* 7.5, pp. 1979–1999. DOI: https://doi.org/10.5194/gmd-7-1979-2014. 2014.

QGIS Development Team. *QGIS Geographic Information System*. Tech. rep. Open Source Geospatial Foundation, 2017. URL: http://qgis.osgeo.org.

Rotach, M. W. “Profiles of turbulence statistics in and above an urban street canyon”. *Atmospheric Environment* 29.13, pp. 1473–1486. DOI: 10.1016/1352-2310(95)00084-C. 1995.

Stocker, J., Hood, C., Carruthers, D., Seaton, M., and Johnson, K. “The development and evaluation of an automated system for nesting ADMS-URBAN in regional photochemical models”. In: Chapel Hill, NC: 13th Annual CMAS Conference, 2014, pp. 1–6.

Sunyer, J. et al. “Association between Traffic-Related Air Pollution in Schools and Cognitive Development in Primary School Children: A Prospective Cohort Study”. *PLoS medicine* 12.3, e1001792. DOI: https://10.1371/journal.pmed.1001792. 2015.

Thunis, P. and Cuvelier, C. *Concepts/User's, DELTA Version 5.4 Guide/Diagrams*. Tech. rep. 2016. URL: http://fairmode.jrc.ec.europa.eu/%20Document/fairmode/WG1/DELTA%7B%5C_%7DUserGuide%7B%5C_%7DV5%7B%5C_%7D4.pdf.

Valencia, A., Venkatram, A., Heist, D., Carruthers, D., and Arunachalam, S. “Development and evaluation of the R-LINE model algorithms to account for chemical transformation in the near-road environment”. *Transportation Research Part D: Transport and Environment* 59.2, pp. 464–477. DOI: https://doi.org/10.1016/j.trd.2018.01.028. 2018.

Vardoulakis, S., Fisher, B., Pericleous, K., and Gonzalez-Flesca, N. “Modelling air quality in street canyons: A review”. *Atmospheric Environment* 37.2, pp. 155–182. DOI: https://doi.org/10.1016/S1352-2310(02)00857-9. 2003.

Venkatram, A., Snyder, M.G., Heist, D.K., Perry, S.G., Petersen, W.B., and Isakov, Vlad. “Re-formulation of plume spread for near-surface dispersion”. *Atmospheric Environment* 77, pp. 846–855. DOI: https://doi.org/10.1016/j.atmosenv.2013.05.073. 2013.

Wild, R.J., Dubé, W.P., Aikin, K.C., Eilerman, S.J., Neuman, J.A., Peischl, Jeff, Ryerson, Thomas B., and Brown, Steven S. “On-road measurements of vehicle NO2/NOxemission ratios in Denver, Colorado, USA”. *Atmospheric Environment* 148.2, pp. 182–189. ISSN: 18732844. DOI: https://doi.org/10.1016/j.atmosenv.2016.10.039. 2017.

Salizzoni, P., Soulhac, L., and Mejean, P. “Street canyon ventilation and atmospheric turbulence”. *Atmospheric Environment* 43.32, pp. 5056–5067. DOI: 10.1016/j.atmosenv.2009.06.045. URL: http://dx.doi.org/10.1016/j.atmosenv.2009.06.045, 2009.

Skamarock, W.C. and Klemp, J.B. “A time-split nonhydrostatic atmospheric model for weather research and forecasting applications”. *Journal of Computational Physics* 227.7, pp. 3465–3485. DOI: https://doi.org/10.1016/j.jcp.2007.01.037. 2008.

Snyder, M.G. and Heist, D.K. *User's Guide for R-LINE Model Version 1.2 A Research LINE source model for near-surface releases*. Tech. rep. USEPA, 2013. URL: https://www.cmascenter.org/r-line/documentation/1.2/RLINE%7B%5C_%7DUserGuide%7B%5C_%7D11-13-2013.pdf.

Snyder, M.G., Venkatram, A., Heist, D.K., Perry, S.G., Petersen, W.B., and Isakov, V. “RLINE: a line source dispersion model for near-surface releases”. *Atmospheric Environment* 77, pp. 748–756. DOI: https://doi.org/10.1016/j.atmosenv.2013.05.074. 2013.

Soret, A., Guevara, M., and Baldasano, J.M. “The potential impacts of electric vehicles on air quality in the urban areas of Barcelona and Madrid (Spain)”. *Atmospheric Environment* 99, pp. 51–63. DOI: https://doi.org/10.1016/j.atmosenv.2014.09.048. 2014.

Soulhac, L., Perkins, R.J., and Salizzoni, P. “Flow in a street canyon for any external wind direction”. *Boundary-Layer Meteorology* 126.3, pp. 365–388. DOI: https://10.1007/s10546-007-9238-x. 2008.

Soulhac, L., Salizzoni, P., Cierco, F. X., and Perkins, R. “The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model”. *Atmospheric Environment* 45.39, pp. 7379–7395. DOI: https://10.1016/j.atmosenv.2011.07.008. 2011.

Stocker, J., Hood, C., Carruthers, D., Seaton, M., and Johnson, K. “The development and evaluation of an automated system for nesting ADMS-URBAN in regional photochemical models”. In: Chapel Hill, NC: 13th Annual CMAS Conference, 2014, pp. 1–6.

Sunyer, J. et al. “Association between Traffic-Related Air Pollution in Schools and Cognitive Development in Primary School Children: A Prospective Cohort Study”. *PLoS medicine* 12.3, e1001792. DOI: https://10.1371/journal.pmed.1001792. 2015.

Thunis, P. and Cuvelier, C. *Concepts/User's, DELTA Version 5.4 Guide/Diagrams*. Tech. rep. 2016. URL: http://fairmode.jrc.ec.europa.eu/%20Document/fairmode/WG1/DELTA%7B%5C_%7DUserGuide%7B%5C_%7DV5%7B%5C_%7D4.pdf.

Valencia, A., Venkatram, A., Heist, D., Carruthers, D., and Arunachalam, S. “Development and evaluation of the R-LINE model algorithms to account for chemical transformation in the near-road environment”. *Transportation Research Part D: Transport and Environment* 59.2, pp. 464–477. DOI: https://doi.org/10.1016/j.trd.2018.01.028. 2018.

Vardoulakis, S., Fisher, B., Pericleous, K., and Gonzalez-Flesca, N. “Modelling air quality in street canyons: A review”. *Atmospheric Environment* 37.2, pp. 155–182. DOI: https://doi.org/10.1016/S1352-2310(02)00857-9. 2003.

Venkatram, A., Snyder, M.G., Heist, D.K., Perry, S.G., Petersen, W.B., and Isakov, Vlad. “Re-formulation of plume spread for near-surface dispersion”. *Atmospheric Environment* 77, pp. 846–855. DOI: https://doi.org/10.1016/j.atmosenv.2013.05.073. 2013.

Wild, R.J., Dubé, W.P., Aikin, K.C., Eilerman, S.J., Neuman, J.A., Peischl, Jeff, Ryerson, Thomas B., and Brown, Steven S. “On-road measurements of vehicle NO2/NOxemission ratios in Denver, Colorado, USA”. *Atmospheric Environment* 148.2, pp. 182–189. DOI: https://doi.org/10.1016/j.atmosenv.2016.10.039. 2017.