https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


# Urban atmospheric $CO_2$ plumes from space - Part 2: Representation of fine-scale structures

Alohotsy Rafalimanana[1], Thomas Lauvaux[1,2], Charbel Abdallah[1], Mali Chariot[2,3], Michel Ramonet[2], Josselin Doc[2], Olivier Laurent[2], Morgan Lopez[2], Anja Raznjevic[4], Maarten Krol[4], Leena Järvi[5], Andreas Christen[6], Dana Looschelders[6], Leslie David[7], Olivier Sanchez[7], Laura Bignotti[8], Benjamin Loubet[8], Sue Grimmond[9], and William Morrison[10]

[1]Climate Impacts on Environment Laboratory (CIEL), Université de Reims-Champagne Ardenne, EMR CNRS 7007 AEROLAB, Reims, France
[2]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif sur Yvette Cedex, France
[3]Origins.earth, Reims, France
[4]Meteorology and Air Quality, Wageningen University, Wageningen, the Netherlands
[5]Institute for Atmospheric and Earth System Research (INAR) / Physics, Faculty of Science, University of Helsinki, Helsinki, Finland
[6]Chair of Environmental Meteorology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany
[7]Airparif, Paris, France
[8]ECOSYS, University Paris Saclay, INRAE, AgroParisTech, Palaiseau, France
[9]University of Reading, Reading, United Kingdom
[10]School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, United Kingdom

**Correspondence:** Alohotsy Rafalimanana: alohotsy.rafalimanana@univ-reims.fr

**Abstract.** Quantifying urban $CO_2$ emissions using spaceborn total column observations requires atmospheric transport models capable of resolving fine-scale spatial and temporal variability across metropolitan areas. Using the *Grand Paris* region as a testbed, we evaluated the sensitivity of simulated $CO_2$ fields to spatial resolution and model physics within the Weather Research and Forecasting model coupled to a high-resolution fossil fuel emissions inventory. Simulations were conducted at mesoscale (900 m) and Large-Eddy Simulation (LES) resolutions (300 m and 100 m), and evaluated against dense urban $CO_2$ observations from the Paris Picarro and the rooftop mid-cost sensor network. Horizontal resolution strongly affects plume morphology: high-resolution simulations produce sharper gradients, higher peak mixing ratios, and more realistic temporal variability. LES simulations capture localized enhancements and intra-urban variability more accurately than the 900 m mesoscale run, which systematically underestimates spatial contrasts by a factor of 2–3. Aggregation experiments indicate that averaging LES simulations to 900 m significantly modifies plume magnitudes and spatial gradients. Domain-mean relative errors for near-surface $CO_2$ reach 29–34% in winter and 59–61% in summer, while column-integrated $XCO_2$ shows 21–22% errors in winter and 58–59% in summer, with maxima exceeding 70%. While LES improves the representation of urban $CO_2$ variability, it also increases sensitivity to local wind and inventory uncertainties, meaning misplaced point sources can cause large local biases. These findings emphasize the need for inversion frameworks able to incorporate scale-dependent transport and inventory uncertainties to ensure unbiased urban top-down emission estimates and accurate spatial attribution within city domains.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


## 1 Introduction

Urban carbon dioxide ($CO_2$) emissions are heavily concentrated in cities, which contribute the majority of global fossil fuel related releases (Churkina, 2016; Crippa et al., 2021). With urban populations projected to comprise two-thirds of the world's total by mid-century, reducing city emissions is a central focus of climate mitigation efforts (Lwasa et al., 2022). Within metropolitan areas, complex surface structures, diverse emission sources and urban-influenced meteorological conditions produce highly variable atmospheric $CO_2$ patterns (Ye et al., 2020). Satellite based monitoring offers a way to quantify these emissions, whether through direct plume analyses or by integrating observations with atmospheric transport models. Such models must capture fine scale spatial gradients to accurately link observed $CO_2$ mixing ratios to underlying surface fluxes, enabling robust, multi-scale emission estimates within comprehensive monitoring frameworks (Che et al., 2024).

Recent advances in urban monitoring, including dense in situ sensor networks deployed in cities such as Paris (Lian et al., 2024) and San Francisco (Delaria et al., 2021), together with high resolution satellite imaging (*e.g.*, the OCO-3 Snapshot Area Mode; Taylor et al., 2020; Kiel et al., 2021), have revealed fine scale plume structures embedded within broader urban enhancements. These observations highlight a fundamental scale challenge: atmospheric transport models typically operate at kilometer scale resolution, whereas the processes governing urban dispersion, including building induced turbulence, plume meandering, and localized emission hotspots occur at much finer scales (Ye et al., 2025).

Mesoscale models parameterize sub-grid turbulence and urban canopy processes, potentially smoothing concentration gradients and underrepresenting peak enhancements (Gaudet et al., 2017). In contrast, large-eddy simulation (LES) explicitly resolves dominant turbulent eddies and provides a more detailed representation of dispersion processes. However, LES is computationally demanding, and its added value for urban $CO_2$ monitoring and inversion applications remains insufficiently quantified. Understanding how horizontal resolution influences plume structure, variability, and observational agreement is therefore essential for designing robust multi-scale modeling frameworks.

Using the Weather Research and Forecasting (WRF) model (Skamarock et al., 2019) with passive tracer capabilities (Lauvaux et al., 2012), this study investigates the sensitivity of simulated urban $CO_2$ plumes to horizontal resolution and model physics over the Paris metropolitan area. A physically consistent modeling configuration is adopted to ensure realistic representation of surface fluxes and boundary layer processes, allowing isolation of resolution effects.

This work forms part 2 of a two-part investigation of urban $CO_2$ modeling over Paris. In a companion study, we examine the sensitivity of simulated $CO_2$ fields to urban canopy and planetary boundary layer parameterizations, establishing a physically consistent modeling framework. The present study focuses exclusively on the impact of horizontal resolution and turbulence resolving approaches, using a fixed configuration to isolate scale dependent effects. The analysis presented here evaluates how grid refinement influences: plume morphology and spatial gradients, near-surface concentration variability, boundary layer development, agreement with ground-based and column observations, and representativeness errors arising from scale mismatch. We further compare native kilometer-scale simulations with spatially aggregated LES outputs to quantify potential biases introduced by spatial averaging.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


Paris provides an exceptional observational testbed due to its dense, multi-scale monitoring network, which includes high-precision surface $CO_2$ measurements, rooftop sensor networks, total column $CO_2$ observations ($XCO_2$), as well as Doppler wind LIDAR and planetary boundary layer height (PBLH) measurements. Combined with extensive meteorological observations across the Île-de-France region, this network enables a detailed evaluation of atmospheric transport and urban $CO_2$ dynamics at multiple scales.

By explicitly quantifying the trade-offs between mesoscale and turbulence resolving simulations, this study provides guidance for the design of multi-scale urban inversion systems (Staufer et al., 2016; Lauvaux et al., 2016; Lian et al., 2023) and the integration of high resolution transport modeling into operational $CO_2$ monitoring frameworks.

The paper is structured as follows. Section 2 describes the modeling framework and experimental design. Section 3 presents the resolution sensitivity analysis and aggregation experiments. Section 4 discusses implications for inversion systems and representativeness error. Section 5 summarizes the main findings.

## 2 Methodology, Models and Data

This section provides a concise summary of the detailed methodology presented in Part 1, along with complementary information specific to the resolution sensitivity experiments conducted in this Part 2.

The simulations were conducted for two intensive periods representing winter (10–31 January 2024) and summer (01–20 June 2024), selected to capture contrasting planetary boundary-layer dynamics, surface energy partitioning, and seasonal variability in $CO_2$ emissions and meteorological conditions. Each experiment included a spin-up time allowing the large-scale circulation and background atmospheric composition to adjust to the model configuration. The spin-up periods (01–09 January and 23–31 May) were excluded from the analysis, resulting in effective simulation lengths of 21 days for winter and 19 days for summer.

The study focuses on the Paris metropolitan area and the surrounding Île-de-France region, a densely urbanized environment with strong anthropogenic $CO_2$ emissions, complex land–atmosphere interactions, and pronounced urban heat island effects. The region encompasses a heterogeneous mix of urban, suburban, and peri-urban land-use types, providing a representative setting for investigating contrasts in boundary layer structure, and atmospheric transport. As one of Europe's largest metropolitan areas, Paris exhibits substantial seasonal variability in energy consumption, traffic activity, and biospheric fluxes, making it well suited for evaluating model performance under contrasting meteorological and emission conditions. The Île-de-France region also benefits from an extensive network of meteorological and atmospheric observations that supports robust model evaluation.

### 2.1 WRF–LES model configurations

The simulations were carried out with the Advanced Research WRF model (WRF–ARW, version 4.6.1) (Skamarock et al., 2019), coupled with the WRF–Chem module (Grell et al., 2005; Peckham et al., 2017). In WRF–Chem, $CO_2$ is treated as a passive tracer, with transport and mixing processes represented explicitly, while chemical transformations are neglected (Lauvaux et al., 2012).

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


The model configuration consists of a parent domain and four one-way nested domains centered over Paris (Figure 1). The horizontal resolution ranges from 8.1 km × 8.1 km in the outermost domain (D01) to 100 m × 100 m in the innermost domain (D05), allowing the investigation of model sensitivity across a wide range of spatial scales. The nesting strategy follows a one-way configuration, in which the outer domain provides lateral boundary conditions to the inner domains without feedback to coarser grids.

The WRF-Chem model uses 51 vertical levels extending from the surface up to 50 hPa. The vertical grid is stretched, with higher resolution near the surface and coarser spacing aloft; the lowest model layer has a thickness of approximately 10 m, and about 25 levels are located within the first 1000 m, capturing the bulk of the urban atmospheric boundary layer. This configuration ensures that turbulence, vertical mixing, and the evolution of $CO_2$ plumes are adequately represented. Table 1 summarizes the principal characteristics of the simulation domains.

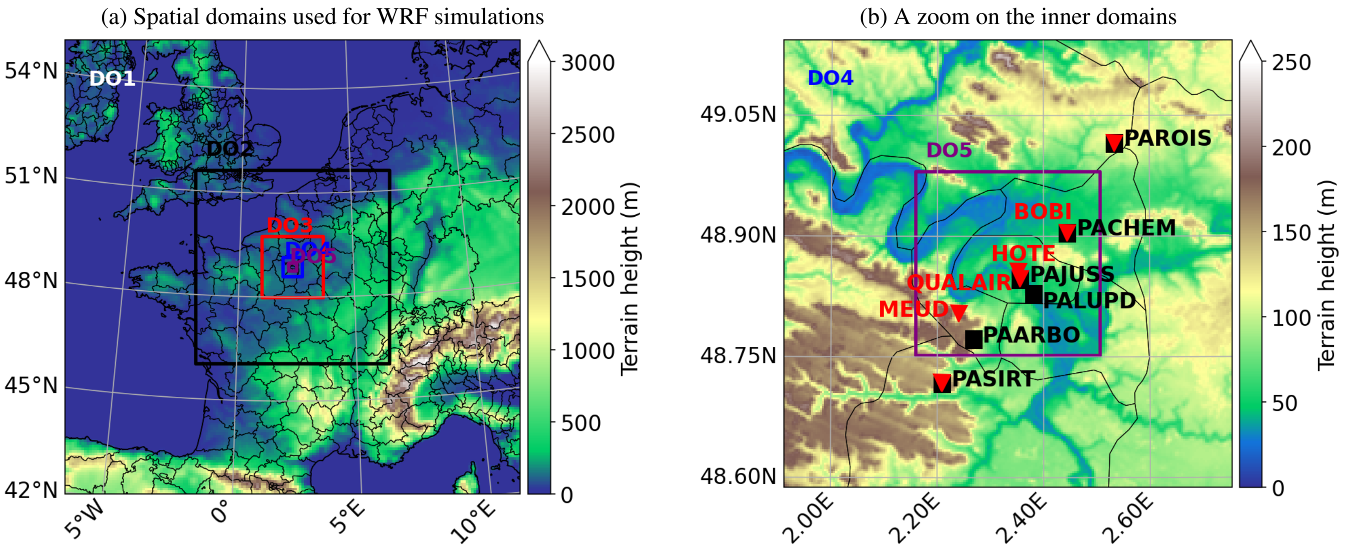


**Figure 1.** WRF domain configuration. (a) Spatial domains used for WRF simulations. The horizontal resolution varies from 8100 m to 100 m. (b) A zoom on the inner domains: D04 at 300m and D05 at 100m resolution. The colormap represents the terrain height in meter. Black symbols represent Doppler wind profile LIDAR positions and red symbols indicate the Automatic Lidar Ceilometers (ALC) locations used for planetary boundary layer height (PBLH) retrievals. Some sites host collocated instruments, notably PAROIS, PACHEM, and PASIRT.

The physical parameterizations are configured to allow the simultaneous representation of mesoscale dynamics in the outer domains and Large Eddy Simulation (LES) physics in the innermost domains. LES explicitly resolves the dominant turbulent eddies and parameterizes subgrid-scale turbulence, enabling a more detailed representation of urban dispersion processes. The microphysics scheme is set to the Thompson aerosol-aware scheme (Thompson et al., 2008) for the parent domain and to the WRF Single–Moment 5-class scheme (Hong et al., 2004) for the nested domains. Both longwave and shortwave radiation are computed using the Rapid Radiative Transfer Model for GCMs (RRTMG; Iacono et al., 2008) consistently across all domains.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


**Table 1.** Configuration of the nested simulation domains used in this study. For each domain, the table reports the horizontal grid resolution (m) in the $x$ and $y$ directions, the number of horizontal grid points, and the number of vertical levels used in the model

| Domains | Resolution (m) | Grid points number | Vertical level number |
|---|---|---|---|
| D01 | 8100×8100 | 182×182 | |
| D02 | 2700×2700 | 232×232 | |
| D03 | 900×900 | 220×220 | 51 |
| D04-LES | 300×300 | 208×208 | |
| D05-LES | 100×100 | 256×256 | |

The planetary boundary layer (PBL) is represented by the Mellor–Yamada–Janjic (MYJ) turbulence closure scheme (Janjic, 1990) in the mesoscale domains (D01–D03). In the two innermost domains (D04–D05), the PBL scheme is disabled and the model is operated in LES mode to explicitly resolve turbulent motions. The surface layer uses the Janjic Eta similarity scheme (Janjic, 1996), coupled with the Noah land surface model (Chen and Dudhia, 2001) across all domains. Urban surface processes are represented using the multi-layer Building Energy Model (BEM; Salamanca and Martilli, 2010), which provided the optimal configuration as described in Part 1. The outer domain (8.1 km) lies within the gray zone of convection, where it is only partially resolved, while convective processes become progressively better resolved toward the inner domains. No cumulus parameterization is applied in any domain. To improve model performance, the Four-Dimensional Data Assimilation (FDDA; Stauffer and Seaman, 1994) scheme is applied to the mesoscale domains (D01–D03) using meteorological measurements from the World Meteorological Organization (WMO). No FDDA is applied in the LES domains to prevent numerical instabilities associated with explicitly resolved turbulence.

### 2.2 Model Input Data

The model input data, including meteorological boundary conditions, surface characteristics, and anthropogenic and biogenic $CO_2$ fluxes, are the same as those used in Part 1.

Initial and boundary meteorological fields were provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), the fifth generation ECMWF atmospheric reanalysis (ERA5; Hersbach et al., 2020), offering hourly fields at $0.25° \times 0.25°$ horizontal resolution to drive WRF-Chem and ensure realistic synoptic conditions.

High-resolution surface datasets were used to represent topography and land cover consistently with the fine model grids. Topography for the innermost domain (D05, 100 m resolution) was obtained from the 1 arc-second Copernicus Global Digital Elevation Model (CGDEM) dataset (∼30 m; Agency, 2024), for D04 (300 m) and D03 (900 m) from 3 arc-seconds CGDEM (∼90 m), and for the outer domains (D01–D02) from Global Multi-resolution Terrain Elevation Data (GMTED) (30 arc-seconds, ∼1 km). Land use and land cover (LULC) were represented using the 100 m Copernicus Global Land Service Land Cover , Moderate Resolution Imaging Spectroradiometer, and Local Climate Zones (CGLC-MODIS-LCZ) hybrid dataset (Demuzere et al., 2023), which includes 61 land cover classes, of which 10 correspond to urban built types.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


Anthropogenic $CO_2$ emissions were based on a combination of the Netherlands Organisation for Applied Scientific Research (TNO; ) inventory (∼1 km; Super et al., 2020) and the high-resolution Airparif inventory over Île-de-France (500 m; Figure 2a), providing a sectorally resolved and spatially detailed representation of emissions. Because the Airparif inventory does not distinguish between point and area sources, point-source data from the TNO inventory were used to adjust the Airparif emissions, providing a consistent representation of large emitters. The annual $CO_2$ emissions from the inventory are first mapped onto the WRF-Chem grid and then converted to hourly emission rates using a scaling factor. The emissions are released at the first vertical model level, approximately 10 m above ground level, to represent near-surface sources throughout the domain, given the absence of source height data in the emission inventories. This simplification does not explicitly represent elevated point sources such as stacks or rooftop chimneys and may consequently lead to an underestimation of tracer concentrations at rooftop sensor elevations, where vertical mixing governs the transport from the surface layer. Biogenic $CO_2$ fluxes were simulated using the Vegetation Photosynthesis and Respiration Model (VPRM; Mahadevan et al., 2008), with satellite-derived inputs including Enhanced Vegetation Index (EVI), Land Surface Water Index (LSWI) and vegetation fraction generated using the `wrf_preprocessor` module of the PyVPRM package (Glauch et al., 2025). Vegetation inputs were derived from the Copernicus Global Land Cover 100 m dataset (Buchhorn et al., 2020) and MODIS/Terra Surface Reflectance 8-Day 500 m product (Eric, 2021).

Boundary and background $CO_2$ mixing ratios were provided by the Copernicus Atmosphere Monitoring Service (CAMS; Agustí-Panareda et al., 2023; Inness et al., 2019) global reanalysis (1° × 1°, 3-hourly), which assimilates both satellite and in situ observations to constrain large scale $CO_2$ variability. This setup ensures that WRF-Chem simulations of $CO_2$ tracers are consistent with observed atmospheric conditions while enabling the analysis of horizontal resolution sensitivity in urban environments.

### 2.3 Observational datasets

For model evaluation in this study, the same observational datasets as in Part 1 were used, focusing on variables relevant to the resolution sensitivity experiments.

Surface meteorology was validated using data from approximately 86 automatic weather stations distributed across Paris and the Île-de-France region, part of the French operational network coordinated by Météo-France and ACTRIS-FR, with data accessed through the AERIS infrastructure (ACTRIS-FR and Météo-France, 2024). Air temperature, relative humidity, wind speed, and wind direction were considered, with a 6-minute temporal resolution. Performance metrics were calculated using subsets of stations with valid observations for each parameter.

To assess the vertical structure of the urban boundary layer, observations from the urbisphere project (Fenner et al., 2024) were used. A multi-institution Doppler wind LIDAR dataset from the urbisphere Paris campaign (Morrison et al., 2025) was used at 10-minute temporal resolution. Planetary boundary layer height (PBLH) was derived from Automatic LIDAR Ceilometers (ALC) following the data processing chain described in Kotthaus et al. (2023), with ceilometer observations from the urbisphere campaign (Morrison et al., 2025; Looschelders et al., 2025) at 1 minute resolution, and additional data from PALAISEAU and

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


QUALAIR stations at 15 minute resolution accessed via the Institut Pierre-Simon Laplace (IPSL) Compute and Data Centre. Aerosol-based mixed-layer heights were determined using the STRATfinder algorithm (Kotthaus et al., 2020).

Two surface $CO_2$ networks were used to evaluate modeled concentrations (Figure 2b): a high-precision network of cavity ring-down spectroscopy sensors (Picarro; Doc et al., 2024) and a broader distributed rooftop network (mid-cost; Lian et al., 2024) using SenseAir HPP and K96 instruments. Both networks provide hourly $CO_2$ mixing ratios.

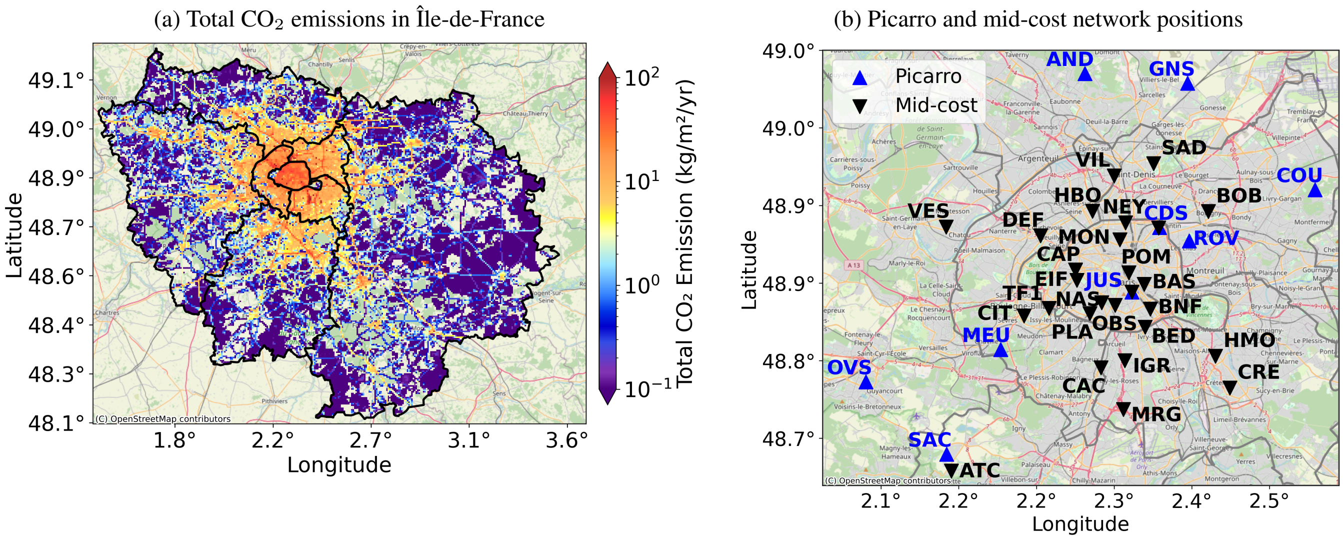


**Figure 2.** (a) Total $CO_2$ emissions in Île-de-France from Airparif inventories (2019). (b) Paris area showing Picarro and mid-cost (HPP and K96) network positions.

For validation of total column $CO_2$ ($XCO_2$), observations from Total Carbon Column Observing Network (TCCON; Té et al., 2022) instruments were used. TCCON data are retrieved using the GGG2020 algorithm (Washenfelder et al., 2006; Laughner et al., 2024), which defines the a-priori profiles and processing settings for consistent greenhouse gas column measurements. The temporal resolution of TCCON data is irregular, scan-based, with multiple measurements per day typically spanning daytime hours and gaps of minutes to hours between scans depending on weather and daylight. The Paris TCCON station is located at Jussieu (JUS), Sorbonne Université, Campus Pierre et Marie Curie, 4 Place Jussieu, Paris 05, France (48.85°N, 2.36°E). It is part of a global network of instruments providing validated total column $CO_2$ measurements. Unlike in-situ observations that reflect near surface concentrations strongly influenced by local boundary layer processes, total column measurements integrate $CO_2$ over the entire atmospheric column and therefore provide complementary information on large-scale transport and vertical distribution. These $XCO_2$ observations from TCCON (JUS, Paris) are used to evaluate WRF-Chem's performance in simulating the vertical distribution of $CO_2$ and to assess whether higher-resolution LES simulations (300 m and 100 m) capture column variability more accurately over urban areas. They also provide a benchmark for evaluating the temporal evolution of column

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


$CO_2$, allowing assessment of whether the model reproduces synoptic-scale variability and whether LES runs improve agreement compared to coarser simulations.

### 2.4 Evaluation strategy and statistical metrics

To examine the influence of horizontal grid resolution on WRF simulated near-surface meteorology and $CO_2$ mixing ratios, model results from three simulations, mesoscale WRF at 900 m, WRF-LES at 300 m, and WRF-LES at 100 m, are compared with observations over a 21 day winter period (10-31 January 2024) and a 19 day summer period (01–20 June 2024). Model output is stored at 10 minute intervals. Observations with finer temporal resolution (e.g., 1 minute) are averaged to match model output intervals, while hourly observations are compared against hourly averaged model output. To ensure temporal consistency between simulations and observations, model $CO_2$ mixing ratios were averaged over a ±30 minute window around each observation time. This approach reduces representativeness errors associated with comparing instantaneous model output with hourly observational data.

Near-surface meteorological variables (air temperature, relative humidity, wind speed, and wind direction) and $CO_2$ mixing ratios are extracted at the model level closest to the corresponding measurement height. Planetary boundary layer height (PBLH) is diagnosed using the bulk Richardson number method (Stull, 1988; Vogelezang and Holtslag, 1996), and model values are extracted from the grid cell corresponding to each observation site. Discrepancies between modeled and observed PBLH should not be interpreted at face value, given the fundamental methodological differences underlying each estimate: observational retrievals are based on aerosol backscatter profiling, whereas model diagnostics rely on critical Richardson number thresholds.

Model performance is assessed using standard statistical metrics: correlation coefficient (R), mean bias error (MBE), mean absolute error (MAE) and root mean square error (RMSE).

In addition to bulk metrics, scale-dependent variability of wind speed is examined using Power Spectral Density (PSD), allowing assessment of the distribution of variance across temporal frequencies and the representation of sub-hourly fluctuations. Vertical structure of the wind field is evaluated using profiles of MAE and RMSE against Doppler wind LIDAR observations, enabling diagnosis of resolution-dependent errors in wind speed and direction within the PBL.

For $CO_2$ and $XCO_2$, the evaluation considers multiple complementary aspects to assess the impact of horizontal resolution. Temporal variability is examined at selected stations, while probability distributions and Taylor diagram analysis are used to simultaneously assess correlation, variance, and overall error across the different simulations. Spatial aspects of $CO_2$ plume structure are evaluated through near-surface concentration gradients along transects, as well as the impact of spatial aggregation on both near-surface $CO_2$ (35 m AGL) and $XCO_2$. To specifically evaluate the impact of aggregation, high-resolution LES simulations at 100 m and 300 m were aggregated to the 900 m grid of the mesoscale simulation and compared with the native 900 m run. Both near-surface $CO_2$ and $XCO_2$ were considered to capture local emission and vertically integrated signals, respectively. Because no independent observations exist at this aggregated scale, the analysis focuses on internal consistency across resolutions. Time-averaged $CO_2$ and $XCO_2$ enhancements were computed by subtracting the background concentration to isolate plume structures from the dominant background field. Error maps were generated using a symmetric relative error

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


metric computed as:

$$E_{\text{sym}} = 100 \times \frac{|C_{\text{agg}} - C_{\text{nat}}|}{\max\left(|C_{\text{agg}}|, |C_{\text{nat}}|\right)}, \tag{1}$$

where $C_{\text{agg}}$ denotes the aggregated LES field (100–300 m remapped to 900 m), and $C_{\text{nat}}$ denotes the native 900 m simulation. This formulation follows the symmetric relative error family described in Törnqvist et al. (1985), which normalises by the larger of the two values to ensure the metric is bounded in [0, 100%], symmetric with respect to the choice of reference field, and robust to near-zero and mixed-sign enhancements. This avoids both the asymmetry of standard relative error and the unboundedness of log-ratio metrics (Tofallis, 2015), making it well suited for $CO_2$ enhancement fields where background subtraction can produce near-zero or weakly negative values. Analyses were stratified into daytime and nighttime periods to examine how aggregation affects plume morphology and amplitude under differing boundary layer and mixing conditions. This multi-faceted evaluation framework allows a systematic assessment of both temporal and spatial resolution effects on urban $CO_2$ and $XCO_2$ representation, linking fine-scale plume variability from high resolution LES to mesoscale simulations and inversion scale observables.

To ensure a consistent comparison with TCCON observations, the WRF $CO_2$ vertical profiles were converted into total column dry-air mole fractions ($XCO_2$) using the averaging kernels and prior profiles provided with the TCCON retrieval product. Following the standard approach for comparisons between models and column retrievals, the model profiles were sampled along the observation line of sight and smoothed with the averaging kernel to account for the vertical sensitivity of the measurements as well as the viewing geometry, which depends on the solar zenith angle (Rodgers and Connor, 2003; Wunch et al., 2011). The resulting smoothed profiles were then vertically integrated using pressure weighting to compute model equivalent $XCO_2$ values. This procedure ensures that the model derived columns are sampled with the same vertical sensitivity and observational geometry as the measurements.

## 3 Results

### 3.1 Assessment of the sensitivity of WRF model simulations to horizontal grid resolution using ground based meteorological observations

Although the primary focus of this study is $CO_2$ emissions and dispersion, validation of the meteorological fields is essential because simulated temperature, humidity, and wind directly control plume transport and concentration patterns. Table 2 summarizes the statistical performance of the WRF simulations at 900 m, 300 m, and 100 m resolution for winter and summer.

For temperature, all resolutions show high skill during winter (R≈0.95-0.96) with small biases and nearly identical MAE and RMSE values. The 100 m simulation exhibits a marginally lower RMSE (1.5 °C), indicating only slight improvement at the finest resolution. In summer, correlations decrease modestly (R≈0.88-0.89) and all simulations display a small negative bias, but differences among resolutions remain limited. This weak sensitivity of near surface temperature to horizontal grid spacing is consistent with previous mesoscale LES comparisons, which show that bulk thermodynamic variables are often well represented across scales, with limited added value from LES except under strongly stable or heterogeneous conditions (Talbot et al., 2012).

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


For relative humidity, correlations are moderate and nearly identical across resolutions (R≈0.75 in winter; 0.81-0.82 in summer). Error metrics vary only slightly, with the 300 m configuration showing marginally reduced winter bias. Overall, humidity displays limited systematic improvement with increasing resolution. Similar behavior has been reported in multi-resolution mesoscale model inter-comparisons, where modeled relative humidity remains relatively insensitive to grid refinement because moisture fields are largely governed by large scale forcing rather than small scale turbulent structures (Tao et al., 2020; Singh et al., 2021).

In contrast, wind speed shows a clear dependence on resolution. The 900 m simulation consistently outperforms the higher resolution runs, yielding the highest correlations (0.92 in winter; 0.81 in summer) and the lowest MAE and RMSE in both seasons. The 300 m simulation exhibits substantially degraded performance, particularly in winter (R = 0.79; RMSE = 2.9 m $s^{-1}$), along with a pronounced positive bias. The 100 m simulation improves upon the 300 m case but does not fully recover the performance of the 900 m run. A similar pattern is observed for wind direction, where the 900 m configuration produces the highest correlations and lowest directional errors in both seasons, while the 300 m and 100 m simulations show increased bias and RMSE, especially in winter. Previous study (Udina et al., 2020) reported comparable behavior in mesoscale versus LES configurations, showing that increasing horizontal resolution does not necessarily translate into improved deterministic skill for near-surface wind speed and direction. In particular, LES simulations, while providing a more detailed representation of turbulence and small-scale variability, may yield larger scatter and higher errors when evaluated against point observations. This apparent degradation can partly be attributed to the so-called "double-penalty effect", whereby high resolution models are penalized twice in traditional verification metrics: once for spatial displacement errors and once for amplitude differences. As a result, finer resolution simulations may appear statistically worse than coarser runs, even though they better resolve physically realistic flow structures and turbulence characteristics (Nurmi, 2003; Ebert, 2008).

This non monotonic behavior, in which intermediate resolutions perform worse than both coarser and finer grids, is also consistent with previous WRF and WRF-LES comparisons. Studies have shown that grid spacings in the so-called turbulence grey zone inadequately represent turbulent energy spectra: eddies are neither fully resolved nor properly parameterized (Wyngaard, 2004; Muñoz Esparza et al., 2017). Consequently, wind statistics can degrade at intermediate resolutions before improving again at true LES scales. Our results suggest that the 300 m configuration lies within this grey zone regime, whereas the 900 m run benefits from internally consistent PBL parameterization and the 100 m run begins to resolve finer scale turbulent motions more explicitly.

While the verification metrics in Table 2 focus on time-averaged performance, Figure 3 presents the Power Spectral Density (PSD) of wind speed at two sites, Eiffel Tower (EIF) and Villacoublay (VIL), for both seasons, in order to examine the scale dependent distribution of wind variability. Unlike bulk error metrics, PSD quantifies how variance is distributed across temporal frequencies and therefore allows assessment of short time scale fluctuations. Given the 10 minutes temporal resolution of the data, the analyzed frequency range primarily represents sub-hourly to multi-hour variability rather than microscale turbulence occurring at second level time scales; therefore, the spectra do not resolve the inertial sub-range or permit assessment of Kolmogorov scaling, but instead provide insight into model smoothing and the distribution of energy between synoptic and local variability scales.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


**Table 2.** Statistical evaluation of near-surface meteorological variables simulated by WRF at different resolutions: 900 m, 300 m and 100 m resolution for winter (10-31 January) and summer (01-20 June) periods. Statistics include the correlation coefficient (R), mean bias error (MBE), mean absolute error (MAE), and root mean square error (RMSE).

| | | **Winter (January)** | | | | **Summer (June)** | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Resolution** | R | MBE | MAE | RMSE | R | MBE | MAE | RMSE |
| Temperature | 900 m | 0.96 | 0.3 | 1.2 | 1.6 | 0.88 | -0.7 | 1.3 | 1.8 |
| | 300 m | 0.95 | -0.3 | 1.2 | 1.6 | 0.89 | -1.3 | 1.7 | 2.1 |
| | 100 m | 0.96 | -0.3 | 1.2 | 1.5 | 0.88 | -1.1 | 1.6 | 2.0 |
| Relative Humidity | 900 m | 0.75 | -3.5 | 6.9 | 8.7 | 0.81 | 1.0 | 7.0 | 9.4 |
| | 300 m | 0.75 | -1.5 | 6.4 | 7.9 | 0.82 | 2.5 | 7.5 | 9.7 |
| | 100 m | 0.75 | -2.1 | 6.6 | 8.1 | 0.81 | 1.5 | 7.4 | 9.7 |
| Wind Speed | 900 m | 0.92 | 0.0 | 1.0 | 1.5 | 0.81 | -0.3 | 1.1 | 1.6 |
| | 300 m | 0.79 | 1.7 | 2.2 | 2.9 | 0.67 | 0.9 | 1.8 | 2.3 |
| | 100 m | 0.79 | 0.9 | 1.8 | 2.4 | 0.67 | 0.6 | 1.6 | 2.2 |
| Wind Direction | 900 m | 0.86 | 5.1 | 18.0 | 29.6 | 0.83 | 2.9 | 26.7 | 43.3 |
| | 300 m | 0.78 | 14.8 | 26.2 | 38.8 | 0.80 | 7.0 | 31.2 | 47.0 |
| | 100 m | 0.77 | 10.1 | 26.7 | 40.5 | 0.78 | 5.9 | 33.1 | 49.5 |

Temperature is in °C, relative humidity in %, wind speed in m $s^{-1}$, and wind direction in degrees. Wind direction statistics are restricted to periods with wind speed exceeding 1 m $s^{-1}$ to avoid ill-defined directions under weak wind conditions.

The PSD analysis reveals distinct seasonal behavior. In winter, low frequency fluctuations (<$10^{-4}$ Hz) are best represented by the 900 m simulation, while the 100 m and 300 m configurations slightly overestimate variability. At higher frequencies ($10^{-4}$–$10^{-3}$ Hz), the 100 m simulation exhibits spectral energy closer to observations, indicating a better representation of short term turbulent variability. In contrast, during summer, the 100 m and 300 m simulations systematically overestimate PSD across both low and high frequencies, whereas the 900 m configuration remains closer to observations. This spectral over-amplification of variability in the LES configurations contributes to their larger deterministic errors.

Overall, the PSD analysis highlights the regime dependent impact of horizontal resolution on wind variability. The benefits of LES are confined to specific frequency ranges and stability conditions, and do not systematically translate into improved agreement with observations. In particular, under convective summer conditions, LES exhibits a systematic overestimation of spectral energy across frequencies, which contributes to their larger deterministic errors. Although LES explicitly resolves turbulent structures, their intermittent spatial and temporal characteristics may not coincide with point measurements, thereby increasing error metrics even when variability is physically realistic.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


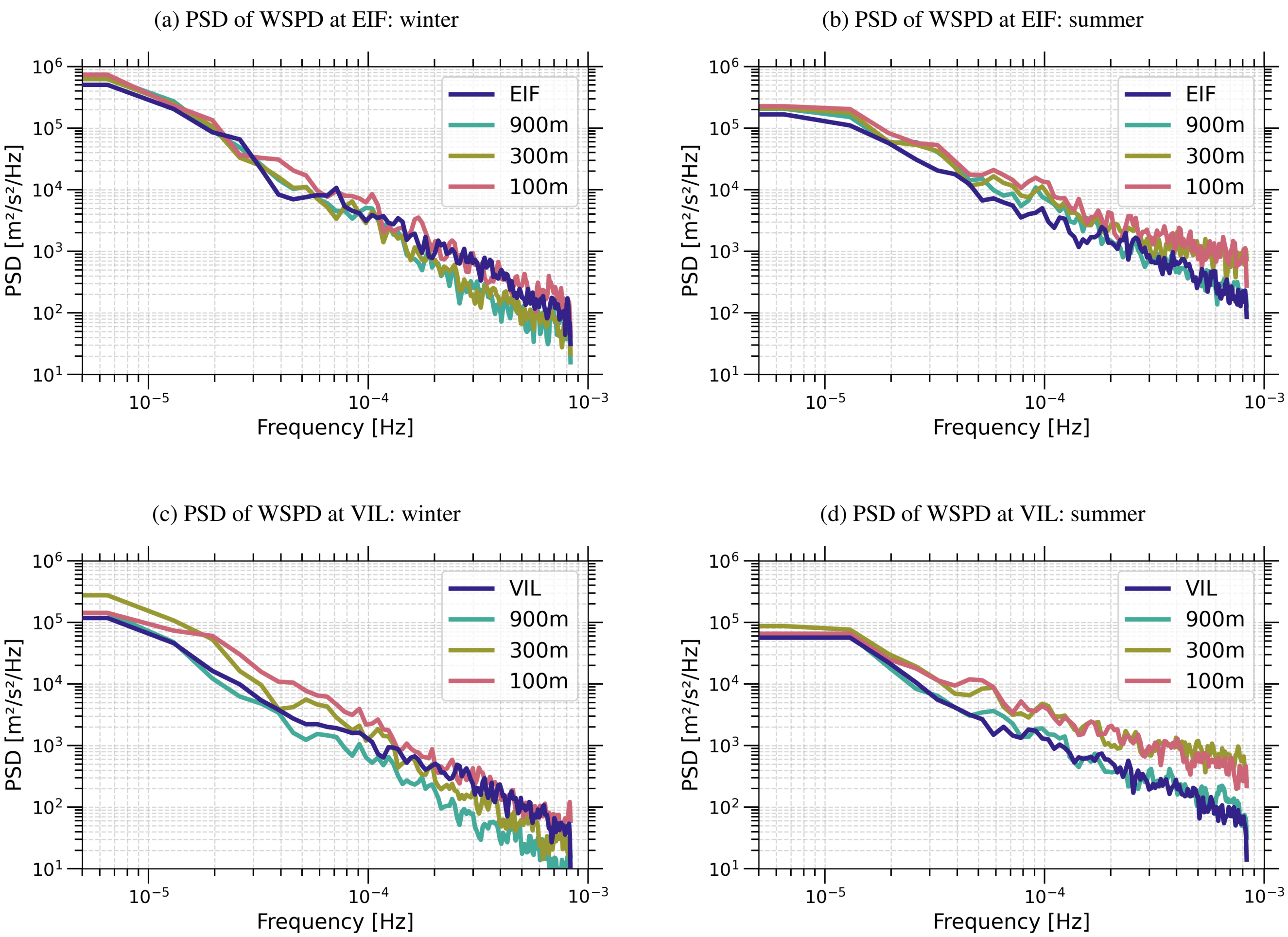


**Figure 3.** Power spectral density (PSD) of 10-min wind speed (WSPD) at Eiffel Tower (EIF) and Villacoublay (VIL) for winter (10-31 January) and summer (01-20 June) 2024. The spectra illustrate the distribution of variance across temporal frequencies for the different horizontal resolutions, highlighting seasonal and scale-dependent differences in wind variability representation.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


### 3.2 Impact of horizontal grid resolution on WRF simulated wind profiles: evaluation with Doppler wind LIDAR observations

To assess the sensitivity of model performance to horizontal resolution, Figure 4 presents the vertical profiles of MAE and RMSE between the WRF simulations and LIDAR observations for both wind speed and wind direction during winter (10-31 January 2024). Only the four LIDAR stations (PAARBO, PALUPD, PAJUSS, and PACHEM) that are located within all three model domains are included to ensure consistency in the inter-resolution comparison. The two stations situated outside the 100 m domain are excluded from this analysis.

Overall, the 900 m simulation provides the most consistent and accurate representation of both wind speed and wind direction profiles. In contrast, the finer resolution simulations (100 m and 300 m) exhibit larger errors, particularly within the lowest 400 m above ground level (AGL).

For wind speed, the 100 m simulation exhibits the largest MAE and RMSE throughout most of the vertical profile compared to the 900 m and 300 m configurations. In the lower atmosphere, the 900 m MAE increases from approximately 1.4 m $s^{-1}$ near the surface to around 2 m $s^{-1}$ at 400 m, while the 300 m simulation shows relatively little change, ranging from 1.7 to 1.9 m $s^{-1}$. The 100 m simulation increases from 1.9 to 2.2 m $s^{-1}$ over the same layer. Above 400 m AGL, the errors for all simulations slightly decrease, but the 300 m profile does not converge with the others and maintains the lowest errors at higher altitudes, whereas the 900 m and 100 m configurations converge around 500 m AGL. This pattern highlights that error differences are most pronounced in the lower atmosphere, with finer and coarser resolutions exhibiting convergence aloft, while intermediate resolutions may not follow the same vertical trend due to partially resolved turbulence in the gray-zone regime.

For wind direction, the MAE and RMSE in the lower atmosphere (below 400 m AGL) differ by approximately 2–6° between the 900 m simulation and the LES at 300 m and 100 m, with the 900 m configuration consistently exhibiting the lowest errors. The 300 m and 100 m simulations display similar error magnitudes and vertical structures. At the lowest levels (below 400 m AGL), the MAE decreases from 20–28° near the surface to approximately 16° at 400 m, while the RMSE decreases from 36–42° to about 28° over the same layer. Above 400 m, wind direction errors further decrease and the profiles converge, with all resolutions exhibiting comparable error levels (MAE $\approx$ 12–16° and RMSE $\approx$ 22–30°), and differences among resolutions become negligible.

The vertical wind profile errors mirror the surface findings, confirming the non-monotonic sensitivity to horizontal resolution. The 900 m simulation consistently provides the most accurate representation of both wind speed and direction, while the 300 m run shows degraded performance in the lower atmosphere due to partially resolved turbulence in the grey-zone regime, and the 100 m simulation only partially recovers skill. Differences between resolutions are most pronounced below 400–500 m AGL and converge with height.

These results indicate that increasing horizontal resolution does not necessarily improve deterministic skill in the near-surface or lower atmosphere, and that mesoscale configurations can outperform finer LES runs for wind fields. This is partly because three-dimensional turbulence is intrinsically stochastic and cannot be reproduced exactly in space and time, so coarser resolutions with consistent PBL parameterization can provide a better overall balance between turbulence representation and mean flow.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


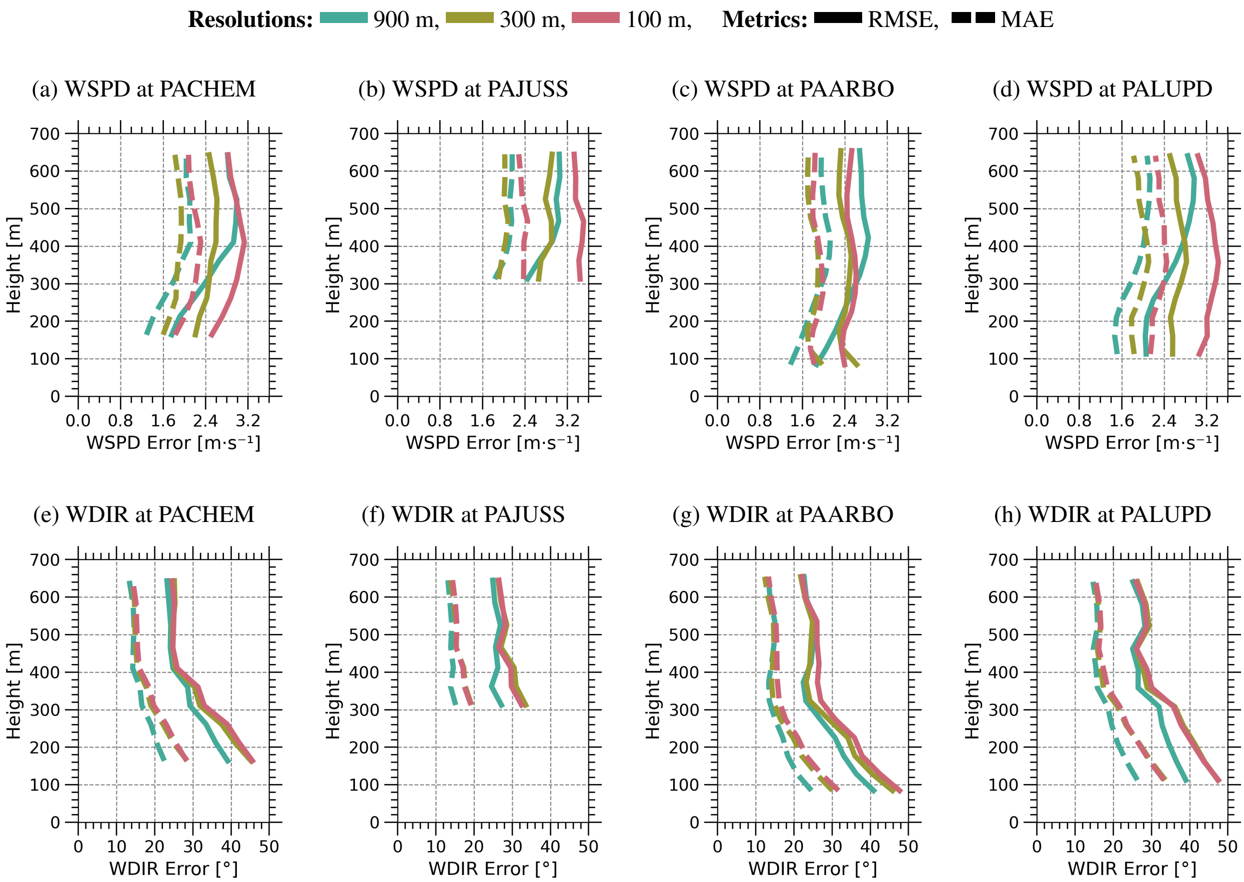


**Figure 4.** Vertical profiles of root mean square error (RMSE) and mean absolute error (MAE) between WRF simulations and Doppler wind lidar observations for wind speed (WSPD, top row) and wind direction (WDIR, bottom row). Results are shown for four sites (PACHEM, PAJUSS, PAARBO, and PALUPD) during wintertime conditions (10–31 January 2024). Simulations were performed at 900 m, 300 m and 100 m horizontal resolution.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


### 3.3 Effects of horizontal grid resolution on PBLH

PBLH for all simulations was diagnosed using the bulk Richardson number (Ri) method to ensure consistency across resolutions. Large-eddy simulations (LES) at 300 m and 100 m do not provide an internally diagnosed PBLH; using the Ri-based approach allows a direct comparison with the 900 m simulation and observations. Figures 5a and 5b show the diurnal cycles of PBLH at QUALAIR in winter and summer 2024, comparing model simulations at 900 m mesoscale resolution and LES at 300 m and 100 m with observations.

In winter, PBLH at 900 m overestimate observations by approximately 100–250 m. The LES at 100 m closely matches observations throughout the diurnal cycle, while LES at 300 m reproduces daytime PBLH well but underestimates nighttime PBLH by roughly 100 m. Overall, the 900 m simulations tend to overestimate PBLH, whereas LES better captures observed diurnal variability, particularly at 100 m resolution.

During summer, in the early morning (05–10 UTC) as PBLH begins to grow, all simulations match observations reasonably well. From 10 UTC onward, around 1000 m, the 900 m simulation follows the observed increase but the aggregated mean begins to decrease after 14 UTC, whereas observations maintain a level near 1500 m until 17 UTC. The LES runs capture the general growth trend but with smaller magnitudes, reaching a composite maximum around 1300 m, with the cycle-averaged PBLH showing lower values after 14 UTC, producing larger afternoon differences relative to observations. During nighttime, the 900 m simulation aligns well with observations, while LES underestimates PBLH.

In winter, the 900 m simulations tend to overestimate PBLH, whereas LES captures diurnal variability more accurately. In summer, morning growth is well represented by all simulations, but LES underestimates afternoon maxima, and the composite-mean 900 m PBLH shows a reduction slightly earlier than observed. These results highlight both the strengths and limitations of LES (100 m and 300 m) versus the 900 m simulation.

Figures 5c and 5d further examine the wintertime spatial structure of the PBLH along an urban–rural transect crossing the Paris metropolitan area. The left panel shows the geographical location of the transect and the intersected observational stations, while the right panel presents the mean PBLH averaged between 12:00 and 15:00 UTC during winter, corresponding to the period of maximum daytime boundary layer development. Observed PBLH at station locations is compared with WRF simulated PBLH sampled continuously along the transect for the different horizontal resolutions. Note that some stations lie outside the innermost 100 m and also the 300 m domains.

The transect analysis confirms the behavior identified in the diurnal cycle evaluation. The 900 m simulation systematically overestimates PBLH along most of the transect, consistent with its positive daytime bias in winter. In contrast, the LES configurations (300 m and 100 m) produce PBLH values closer to observations across the urban core, where all domains share common station coverage. Within this central urban sector, the LES runs show similar magnitudes and spatial gradients, suggesting a consistent representation of urban boundary layer development at finer resolutions.

Furthermore, urban plume effects driven by prevailing wind directions may contribute to observed PBLH enhancement, as discussed in detail in part 1 (Rafalimanana et al., 2026).

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


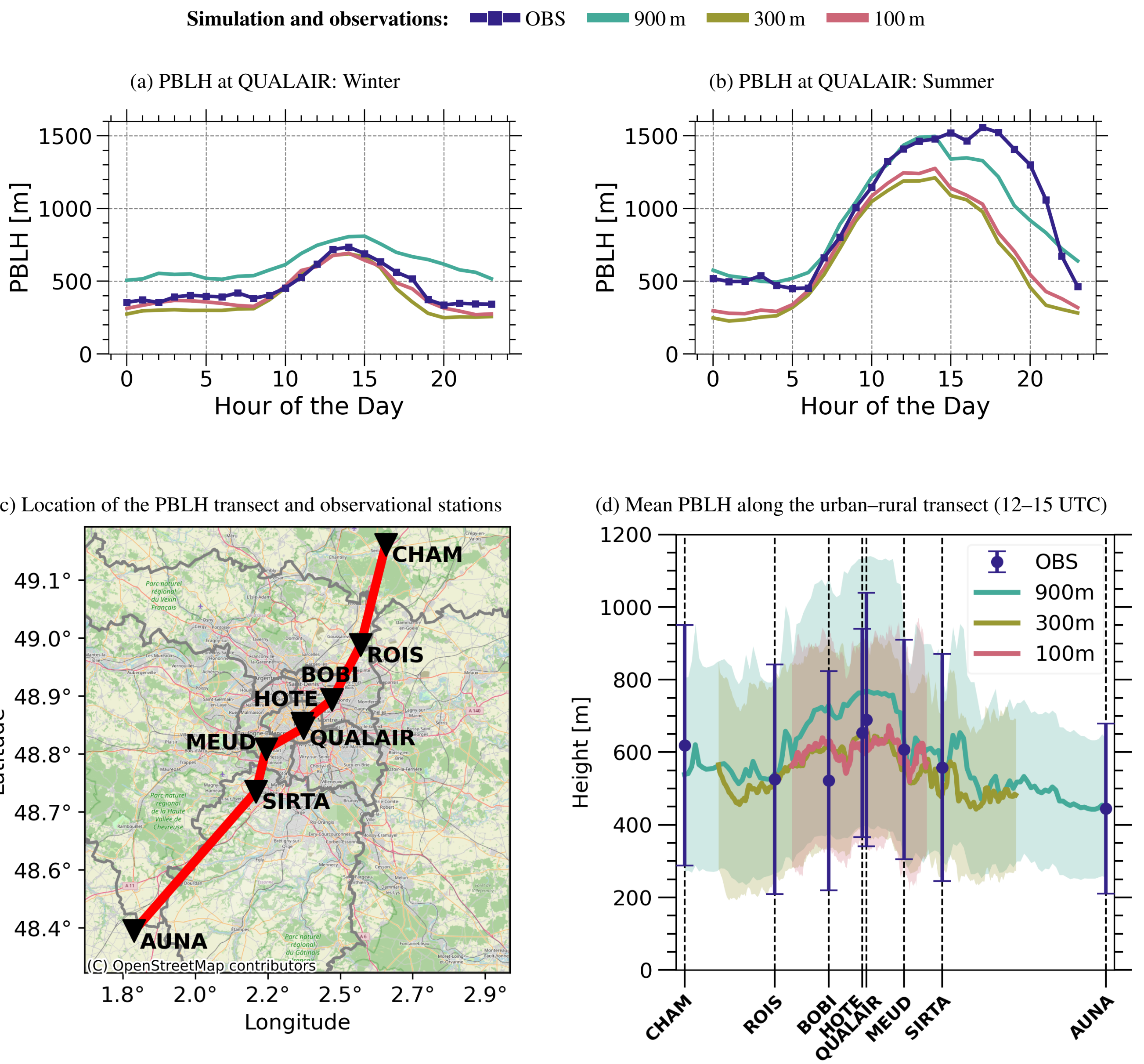


**Figure 5.** Temporal and spatial variability of planetary boundary layer height (PBLH) over the Paris metropolitan area in winter and summer 2024. Model simulations at 900 m and LES at 300 m and 100 m are evaluated against observations (Looschelders and et al., 2026, in prep; to be updated with the preprint citation). (a–b) Diurnal cycles of PBLH at QUALAIR during winter and summer. (c) Location of the urban–rural transect and observational stations across Paris. (d) Mean PBLH along the transect averaged between 12:00 and 15:00 UTC during winter, corresponding to the period of maximum boundary layer development. Model simulations at 900 m and LES at 300 m and 100 m are evaluated against observations. PBLH is diagnosed using the bulk Richardson number for all runs. In (d), shaded areas and bars indicate standard deviation.

Overall, the spatial structure reinforces the temporal analysis: during winter, the 900 m configuration tends to overestimate daytime PBLH, whereas the LES simulations provide a more realistic depiction of the urban boundary layer, particularly within the densely built environment.

### 3.4 Plume structure, dispersion, numerical diffusion

Here, we examine how model resolution influences the simulated structure and dispersion of a $CO_2$ plume. Inter-comparing LES and mesoscale simulations provides insight into the extent to which coarse resolutions smooth concentration fields and dilute localized signals (representation and aggregation errors), compared to finer grids that better preserve plume integrity and spatial variability. Figure 6 presents WRF-simulated $CO_2$ mixing ratios for a winter case (11 January 2024 at 13:00 UTC) and a summer case (10 June 2024 at 15:00 UTC). The upper panels show near-surface $CO_2$ at 35 m above ground level, while the lower panels display the corresponding vertical cross-sections along the source-centered transect indicated in the maps, with PBLH overlaid. For consistency of visual comparison, the spatial extent shown for the 900 m and 300 m simulations is cropped to match the innermost 100 m domain, although the coarser domains extend beyond this area within the nested configuration.

The transect was deliberately defined to pass directly through the dominant point emission sources within the Paris metropolitan area, oriented approximately along the prevailing wind direction at that time. These sources are dominated by public power generation (large combustion installations) and municipal waste incineration facilities. In the northern sector (Saint-Ouen-sur-Seine), the transect intersects a co-located district heating plant and waste-to-energy incinerator. In the south-western sector (Boulogne-Billancourt), it crosses a waste incineration facility with associated combustion units, while in the south-eastern sector (Ivry-sur-Seine), it intersects a high-capacity waste-to-energy complex coupled with urban heat recovery. This configuration allows a direct examination of the plume core and its vertical development, providing a consistent basis for assessing how horizontal resolution affects mixing ratio gradients and dispersion. By intersecting the emission center, the cross-section captures the region of strongest scalar gradients, where the effects of numerical diffusion are most pronounced and resolution dependent smoothing can be clearly diagnosed.

The comparison between LES and mesoscale simulations highlights significant differences in plume structure, dispersion, and the effects of numerical diffusion. In the low-resolution domains, the plume appears broader and less concentrated due to enhanced numerical diffusion, which artificially disperses $CO_2$ and smooths horizontal and vertical gradients, in addition to the coarse representation of the underlying point sources. Conversely, the high-resolution domains capture a more compact, concentrated plume with elevated concentrations near the emission sources, where finer grid spacing reduces numerical diffusion and better preserves plume structure.

Under winter conditions, characterized by a relatively shallow boundary layer (PBLH reaching approximately 500–600 m), $CO_2$ accumulates within the confined lower layer, leading to enhanced concentrations throughout the boundary layer depth. The limited vertical mixing strengthens horizontal gradients and accentuates differences between model resolutions. In contrast, during the summer case, the deeper and more convective boundary layer promotes stronger vertical mixing, resulting in a more vertically homogeneous distribution of $CO_2$ within the PBL. Consequently, elevated concentrations remain primarily confined near the emission source, while background concentrations appear more uniform away from the plume core.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


Differences across model physics and resolutions also reflect variations in the representation of local wind fields and boundary layer turbulence, which influence plume direction, spread, and vertical mixing. Topographic effects become more apparent at higher resolutions, where local channeling and uplift affect dispersion patterns, producing localized shifts in plume direction. In both winter and summer cases, the WRF-LES simulation at 100 m resolution exhibits substantially greater spatial variability in PBLH compared to the coarser 900 m configuration. This enhanced heterogeneity reflects the finer-scale representation of surface forcing and turbulence, which modulate boundary layer depth locally and, in turn, influence plume confinement and vertical dispersion. In contrast, the 900 m simulation produces a comparatively smoother and more horizontally uniform PBLH field, consistent with increased spatial averaging at coarse resolution.

Overall, our results show that spatial resolution and model physics strongly influence the simulated $CO_2$ plume structure and spatial variability across different boundary layer regimes, highlighting the sensitivity of urban-scale $CO_2$ fields to grid spacing.

### 3.5 Impact of horizontal resolution on transect-based $CO_2$ gradients and station agreement

To evaluate how the simulated plume gradients compare with ground-based observations, $CO_2$ mixing ratios are sampled along the same source-centered transects shown in Fig. 6. Spatial gradients between two endpoints are analyzed to assess the model's ability to resolve concentration variability along the prevailing wind direction. Two case studies are examined: 11 January 2024 at 13:00 UTC (winter) and 10 June 2024 at 15:00 UTC (summer). In both cases, transects are selected such that the mean wind direction falls within ±25° of the bearing between the two endpoints, ensuring alignment with plume transport.

For the winter case, the transect extends from the northeast of Paris (48.9766°N, 2.3700°E) to the southwest (48.7609°N, 2.2111°E), crossing the urban core. Stations located within 2.5 km of the transect are included to evaluate how modeled gradients compare with observations. The corresponding $CO_2$ mixing ratios are shown in Fig. 7a. The perpendicular distances of the stations from the transect are: VIL (0.23 km), NEY (2.36 km), HBO (0.57 km), EIF (0.79 km), TF1 (0.29 km), and CIT (1.92 km). As the transect intersects the plume core, clear resolution dependent differences emerge. The 900 m simulation produces a broader enhancement that begins earlier along the transect, reflecting the larger grid spacing and associated numerical diffusion. In contrast, the 300 m and 100 m simulations exhibit a sharper and more localized increase in $CO_2$, with the onset occurring closer to the grid cell containing the source. Because the transect does not necessarily pass exactly through the grid cell of maximum concentration at higher resolution, the 100 m simulation may delay the appearance of the peak until the transect intersects the corresponding grid cell. For example, between VIL and HBO, the 900 m simulation shows a gradual increase immediately after VIL and reaches its maximum earlier, whereas the 300 m and 100 m simulations display a more abrupt rise. Downwind of the peak, concentrations decrease more gradually in all simulations, consistent with plume dispersion. When the transect directly intersects the grid cell of maximum enhancement, the 100 m simulation produces the highest peak values, followed by 300 m, while the 900 m simulation yields the lowest maximum due to spatial smoothing. Away from the plume core, all three resolutions produce comparable magnitudes and generally agree with the observations. However, near the source region, discrepancies between modeled and observed values can be large when a station lies within a neighboring grid cell rather than the one intersected by the transect. This grid alignment sensitivity becomes more pronounced as resolution increases, reflecting the stronger spatial gradients resolved in the 300 m and 100 m simulations.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


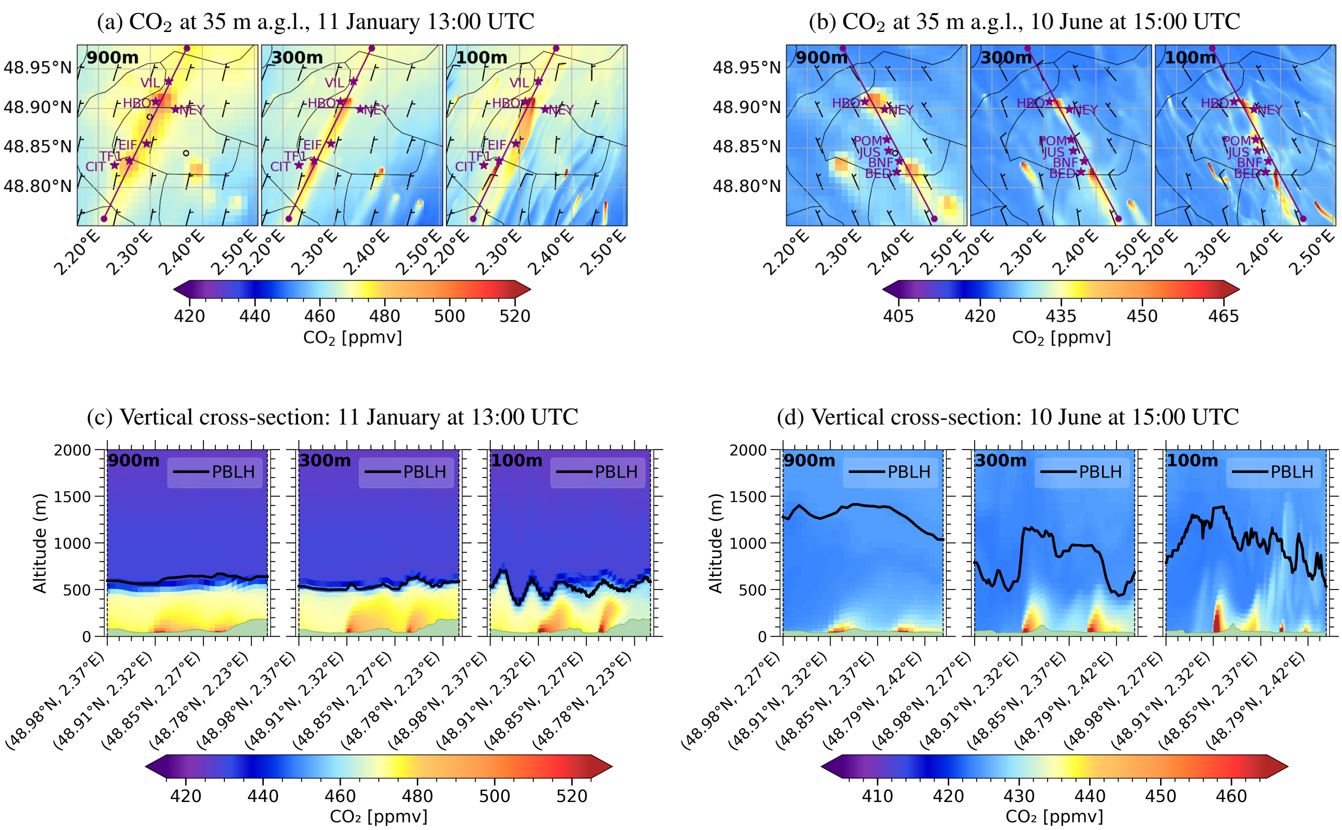


**Figure 6.** WRF-simulated $CO_2$ mixing ratios over Paris for a winter case (11 January 2024, 13:00 UTC) and a summer case (10 June 2024, 15:00 UTC). (a,b) Near-surface $CO_2$ at 35 m above ground level from the 900 m mesoscale simulation and the 300 m and 100 m LES configurations. The wind barbs depict the instantaneous wind speed and direction at the corresponding time. Ground based stations located within 2.5 km of the transect are indicated. (c,d) Vertical cross-sections along the source centered transect for the corresponding dates, with PBLH overlaid. PBLH is diagnosed using the bulk Richardson number for all runs. The spatial extent shown for the 900 m and 300 m simulations is cropped to match the innermost 100 m domain for consistent comparison.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


For the summer case, the transect extends from the northwest of Paris (48.9766°N, 2.2680°E) to the southeast (48.7609°N, 2.4430°E), again crossing the urban area. Stations within 2.5 km of the transect are included, and the resulting profiles are shown in Fig. 7b. Station offsets from the transect are: JUS (1.16 km), NEY (1.14 km), HBO (0.86 km), POM (0.60 km), BNF (0.45 km), and BED (1.55 km). The summer analysis exhibits behavior consistent with the winter case. The 900 m simulation produces smoother and broader enhancements, while the 300 m and 100 m simulations resolve sharper gradients and higher peak concentrations when the transect intersects the plume core. Differences between resolutions are most evident near the source region, whereas background areas show relatively small inter-resolution variability.

Overall, these results highlight the trade-off between numerical diffusion at coarse resolution and increased spatial variability, and corresponding grid alignment sensitivity, at higher resolution when evaluating concentration gradients against point observations. By explicitly sampling the model both along the transect and at the exact station grid cells, it becomes clear that part of the apparent discrepancy between transect values and station measurements arises from spatial sampling rather than intrinsic model error. In the present cases, all selected stations are located outside the grid cell containing the source. Consequently, the model values sampled at the station locations show relatively small differences across resolutions and generally agree well with the observations. In contrast, the transect sampling can intersect the grid cell containing the plume core, particularly at higher resolution, producing sharper and higher peak concentrations. This explains why strong resolution dependent differences are evident along the transect, while the station-based comparisons exhibit more modest inter resolution variability. Even a displacement of one to two grid cells from the source is sufficient to substantially reduce the simulated magnitude at high resolution due to the sharper gradients resolved in the 300 m and 100 m simulations. Thus, grid resolution influences not only the representation of plume structure, but also the likelihood that a point observation samples the plume core. At high resolution, a station located outside the source grid cell may miss the peak entirely, whereas at coarser resolution the larger grid cells are more likely to encompass both the source and the station, producing smoother but more spatially averaged enhancements. Conversely, when the source and station coincide within the same high resolution grid cell, the finer simulation can capture higher and more localized maxima than coarser grids, which dilute the peak through spatial averaging. These results demonstrate that differences between model resolutions in station based evaluation may reflect sampling geometry and representativeness effects in addition to intrinsic differences in plume dispersion.

### 3.6 Resolution dependent representation of $CO_2$ temporal variability at selected stations

While the transect analysis highlights the influence of grid resolution and sampling geometry on the spatial structure of $CO_2$ enhancements, it is equally important to assess how these resolution-dependent effects manifest in the temporal variability recorded at individual monitoring stations. In this subsection, modeled $CO_2$ time series for winter (10–31 January) and summer (01–20 June) 2024 are compared against observations from a representative urban monitoring site, Jussieu (JUS), part of the Picarro network (Figure 8). Figures 8a and 8b show short time periods, highlighting the detailed variability of $CO_2$ at different model resolutions, while Figures 8c and 8c present the full winter and summer periods. This station, selected from the broader Picarro measurement network across the Paris region (Figure 2b), provides a clear example of the model's ability to capture seasonal contrasts, short-term fluctuations, and localized emission signatures within diverse urban environments.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


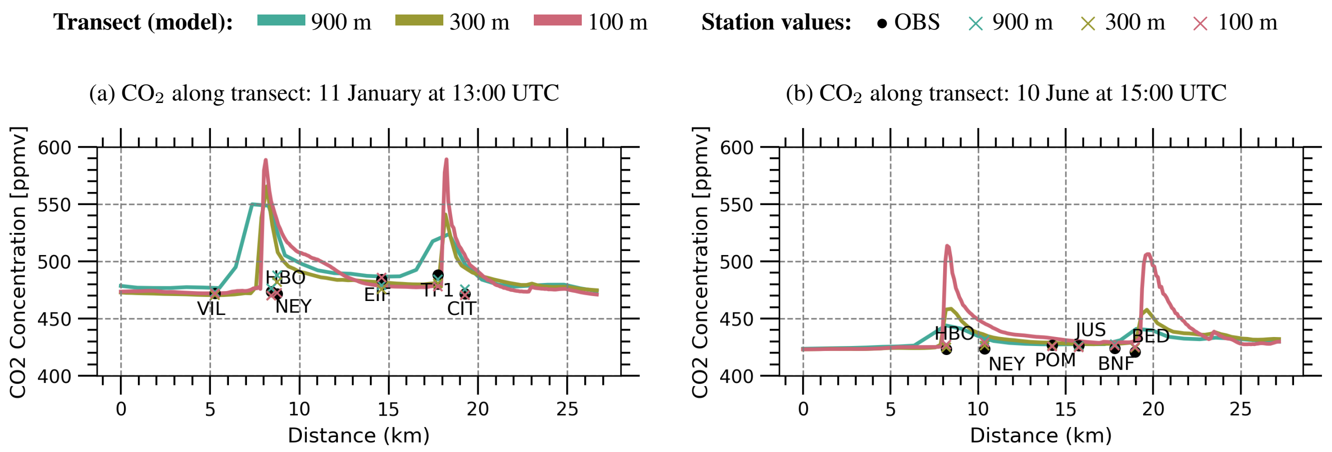


**Figure 7.** WRF-simulated $CO_2$ mixing ratios at 35 m above ground level along the source-centered transects for (a) the winter case (11 January 2024 at 13:00 UTC) and (b) the summer case (10 June 2024 at 15:00 UTC), shown for the mesoscale at 900 m and WRF-LES at 300 m, and 100 m resolutions. Markers denote ground-based stations located within 2.5 km of the transect. For the winter case (a), the perpendicular distances of the stations from the transect are: VIL (0.23 km), NEY (2.36 km), HBO (0.57 km), EIF (0.79 km), TF1 (0.29 km), and CIT (1.92 km). For the summer case (b), the corresponding station distances are: JUS (1.16 km), NEY (1.14 km), HBO (0.86 km), POM (0.60 km), BNF (0.45 km), and BED (1.55 km).

The time series analysis indicates that the 100 m simulation captures pronounced $CO_2$ peaks more clearly than the 300 m and 900 m configurations, not only in terms of peak magnitude but also in the timing and short-term fluctuations. Events observed on 16 and 27 January, as well as on 4 and 18 June 2024, are more distinctly represented at 100 m, showing both sharper spikes and higher temporal variability, whereas the 900 m simulation exhibits smoother and attenuated enhancements with much of the short-term structure lost. The 300 m simulation captures some of these features, but with reduced intensity and intermittency compared to 100 m. This behavior is consistent with the plume and transect-based $CO_2$ gradient analysis, where coarser resolutions dilute concentration maxima and short-term variability through enhanced numerical diffusion and spatial averaging

However, the high resolution simulation occasionally generates concentration peaks that are not evident in the observations and generally exhibits higher variability compared to the 900 m run. This behavior reflects both increased sensitivity to localized emission sources and small-scale turbulent transport processes, as well as representativeness effects: at high resolution, the sharper spatial gradients in the plume can cause station-based sampling to either over- or underestimate concentrations depending on whether the observation falls inside or outside the source grid cell. In addition, a bias in wind direction of about 20°–30° (Table 2, Figure 4) can lead to a lateral displacement of the plume relative to the observation site, further amplifying these discrepancies, particularly at higher resolution where gradients are stronger. More generally, discrepancies between simulated and observed $CO_2$ variability may arise from uncertainties in meteorological forcing, including wind speed and PBLH, which

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


strongly regulate urban-scale transport and mixing. Enhanced high frequency wind variability in the LES configurations, as demonstrated by the PSD analysis, further contributes to intermittent and amplified concentration fluctuations at fixed stations.

Overall, the time series comparison reinforces the spatial findings and highlights a clear trade-off: increasing horizontal resolution enhances the representation of localized plume structures, episodic $CO_2$ enhancements, and short term variability, but may also increase sampling sensitivity and introduce greater fluctuation amplitude and occasional overestimation under certain conditions.

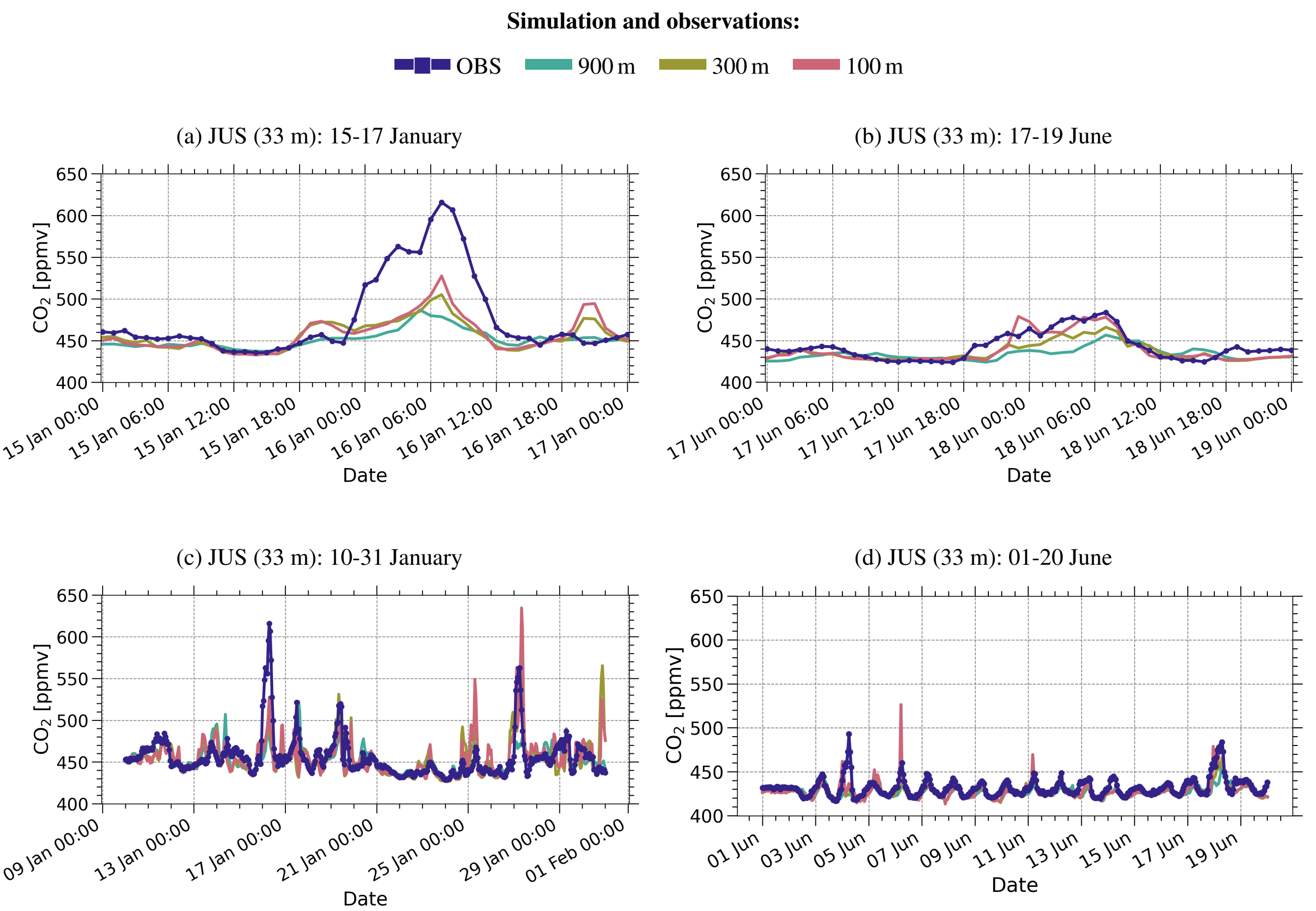


**Figure 8.** Time series of near surface $CO_2$ mixing ratios at Jussieu (JUS) site. (a, b) Observations and WRF-simulated $CO_2$ at 33 m above ground during short periods in winter (15–17 January 2024) and summer (17–19 June 2024). (c, d) Full winter (10–31 January 2024) and summer (01–20 June 2024) periods. Simulations include the 900 m mesoscale configuration and the 300 m and 100 m LES runs, with $CO_2$ sampled at the grid cell corresponding to the station. The observations and model output have an hourly resolution. These time series illustrate the ability of different resolutions to capture localized emission events, episodic peaks, and temporal variability, as well as the effects of grid resolution and LES on the magnitude and intermittency of $CO_2$ enhancements in the urban environment.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


### 3.7 Resolution dependent representation of $XCO_2$ temporal variability at selected stations

To complement the analyses of near-surface $CO_2$ mixing ratios and planetary boundary layer height (PBLH), we evaluated the model's ability to reproduce the total atmospheric column by comparing WRF-simulated total column $CO_2$ ($XCO_2$) with ground-based observations from the TCCON instrument at Jussieu (JUS). This comparison provides an independent constraint on the vertically integrated $CO_2$ field and allows us to assess whether the resolution-dependent differences identified near the surface also appear in the total column.

Figure 9 shows the $XCO_2$ time series at JUS for two winter days (Figures 9a and 9c, 11 and 16 January 2024) and two summer days (Figures 9b and 9d, 07 and 09 June 2024) when observations are available. Overall, the simulations reproduce a similar temporal variability across the different model resolutions. The 100 m WRF-LES simulation exhibits finer-scale temporal fluctuations, whereas the 900 m mesoscale simulation produces smoother variations.

On 11 January, the simulations alternate between over- and underestimation relative to the TCCON observations. After approximately 13 UTC, a clear divergence appears between the model resolutions: the 100 m simulation follows the observed $XCO_2$ more closely, while the 900 m and 300 m simulations tend to overestimate the column values. On 16 January, all model configurations systematically underestimate $XCO_2$ compared with the observations. This behaviour is consistent with the underestimation of near-surface $CO_2$ previously identified at JUS in Section 3.6, suggesting that differences in near-surface concentrations may also influence the integrated column.

During the summer cases (07 and 09 June), the differences between model resolutions are relatively small. All simulations slightly underestimate $XCO_2$ compared with TCCON, although the magnitude of the bias remains limited.

Overall, the results indicate that increasing spatial resolution introduces additional temporal variability in the simulated $XCO_2$ column, consistent with the finer-scale dynamics resolved by the WRF-LES simulations, while the general temporal evolution remains comparable across the different model configurations. Differences in $XCO_2$ magnitude between the analysed dates suggest the influence of varying local conditions and boundary layer dynamics. In particular, the stronger variability simulated by the 100 m run at JUS is consistent with its enhanced sensitivity to local emissions and turbulence, as previously observed in the near-surface $CO_2$ analyses. These results highlight the role of boundary layer processes and local representativeness in shaping column-averaged $CO_2$, reinforcing the link between PBLH representation, near-surface concentration variability, and total column behaviour.

### 3.8 Statistical evaluation of WRF-simulated $CO_2$ against observations

This section assesses the statistical consistency of WRF-simulated $CO_2$ with ground-based measurements from Picarro and mid-cost stations presented in Figure 2b. The objective of this evaluation is to quantify model skills across resolutions using standard statistical metrics, thereby complementing the qualitative insights gained from plume structure, spatial gradient and time series comparisons.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


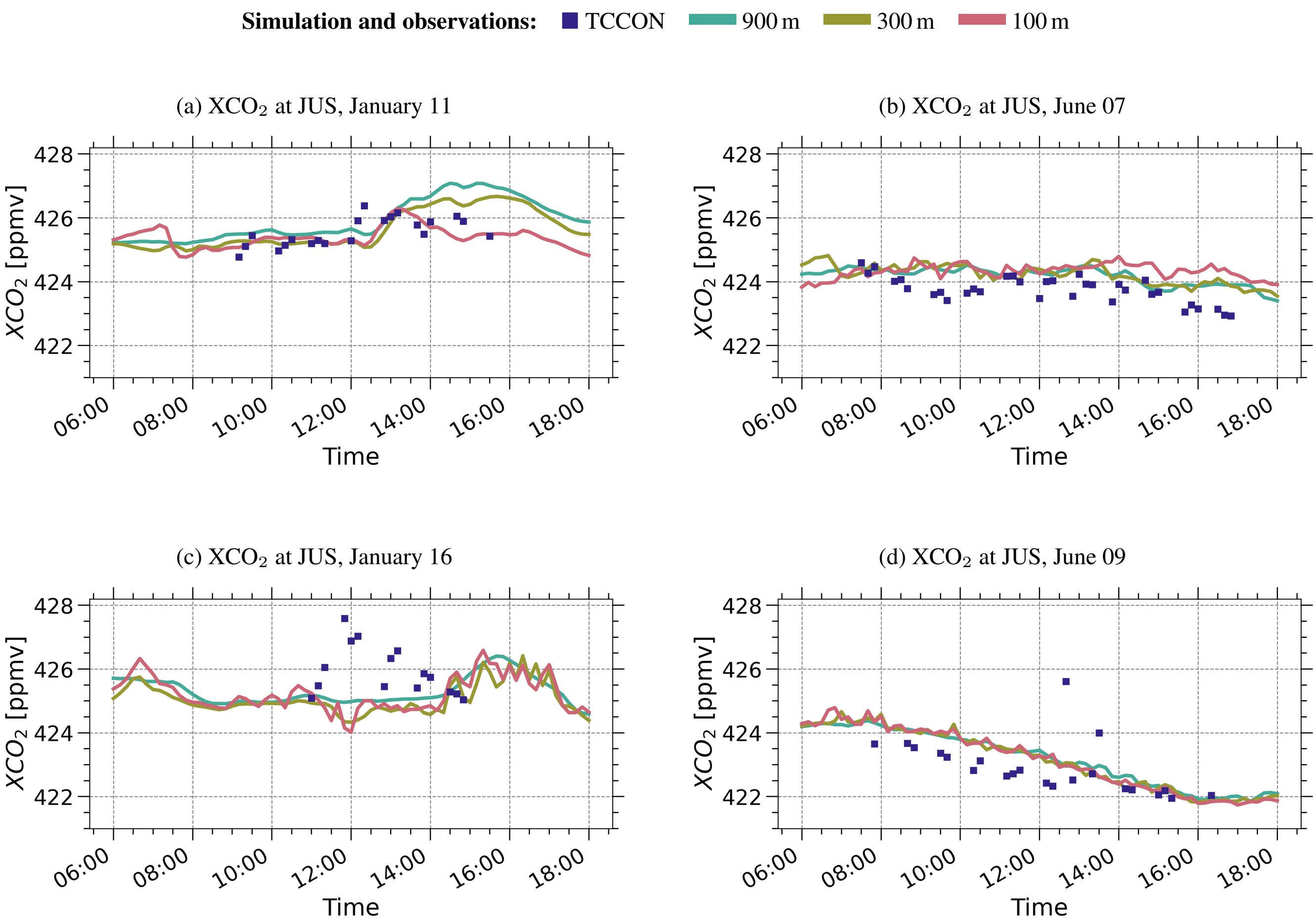


**Figure 9.** Time series of column-averaged $CO_2$ ($XCO_2$) at the Jussieu (JUS) site comparing TCCON observations with WRF simulations at multiple spatial resolutions. The mesoscale WRF simulation is performed at 900 m resolution, while WRF-LES simulations are conducted at 300 m and 100 m resolutions. Panels (a) and (c) correspond to winter cases (11 and 16 January 2024), and panels (b) and (d) correspond to summer cases (7 and 9 June 2024). TCCON measurements are irregularly sampled throughout the day, while model outputs are provided every 10 minutes.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


#### 3.8.1 Resolution dependent evaluation of $CO_2$ probability distributions against observations

The probability density functions (PDFs) of $CO_2$ mixing ratios simulated by WRF at horizontal resolutions of 900 m, 300 m, and 100 m are compared with observations in Figure 10.

The linear scale plots emphasize the central portion of the distributions and allow direct assessment of the dominant concentration ranges. In winter (Figures 10a, 10c), the 900 m simulation provides the best agreement in peak location with the observed distributions. The corresponding density is slightly overestimated, but still captures the most frequently occurring $CO_2$ levels realistically. In contrast, the 300 m and 100 m LES simulations exhibit higher peaks and narrower distributions at the Picarro stations, centered at lower concentration values (approximately 430 ppmv). For the mid-cost network, the distributions are also narrower and shifted toward lower concentrations, with peaks around 435 ppmv. Both differ from the 900 m configuration, which shows broader distributions centered at higher values (around 440–450 ppmv). The narrower distributions in the 100 m and 300 m simulations thus indicate that finer resolution suppresses the high-mixing ratio episodes that broaden the observed distribution.

In summer (Figures 10b, 10d), all three model resolutions produce distributions with higher peaks and slightly narrower shape compared to observations, although the modal concentration remains close to the observed value (around 425 ppmv). This indicates that while the general central tendency is captured across resolutions, the model tends to over-concentrate probability around the mean concentration during convective summer conditions.

For the logarithmic plots (Figures 10e, 10f, 10g, 10h), which emphasize the high-concentration tails, the differences between model resolutions become more pronounced. In winter, the 900 m simulation exhibits a sharp decline in probability density around ∼500 ppmv, while in summer the drop occurs near ∼460 ppmv, diverging from the observations and failing to reproduce the gradual tail of rare, high-concentration $CO_2$ events. In contrast, the 300 m and 100 m LES simulations better capture the observed tail behavior, with the 100 m configuration providing the closest match, although it occasionally produces minor spurious peaks in the tail not present in the measurements. These anomalies occur at very low probability densities (below $5 \times 10^{-4}$) and are consistent with the time series analysis in Section 3.6. They likely arise from the accumulation of local emissions combined with insufficient mixing in stable boundary layers, a well known challenge in winter, where fine-resolution simulations sometimes overestimate episodic peaks and may generate artificial rare event variability. While these artificial fluctuations are minor in terms of overall probability, they highlight the increased sensitivity of fine-resolution LES to localized emissions and turbulent transport, which enhances high-concentration events but can also amplify variability beyond observed levels.

These distributional analyses are consistent with the plume, transect, and time series results. In winter, the 300 m and 100 m LES simulations produce narrower central distributions, reflecting high resolution grids' ability to preserve localized emission maxima, whereas the 900 m simulation yields broader distributions due to spatial averaging. In summer, all resolutions show similar linear-scale distributions, indicating that the central tendency is captured regardless of grid spacing. Across both seasons, the logarithmic-scale PDFs reveal that finer resolutions better capture high concentration tails, although the 100 m simulation occasionally introduces minor spurious peaks at very low probabilities. Overall, the PDFs confirm that horizontal resolution

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


strongly influences the representation of extreme concentrations and localized plume structure, complementing the spatial and temporal analyses.

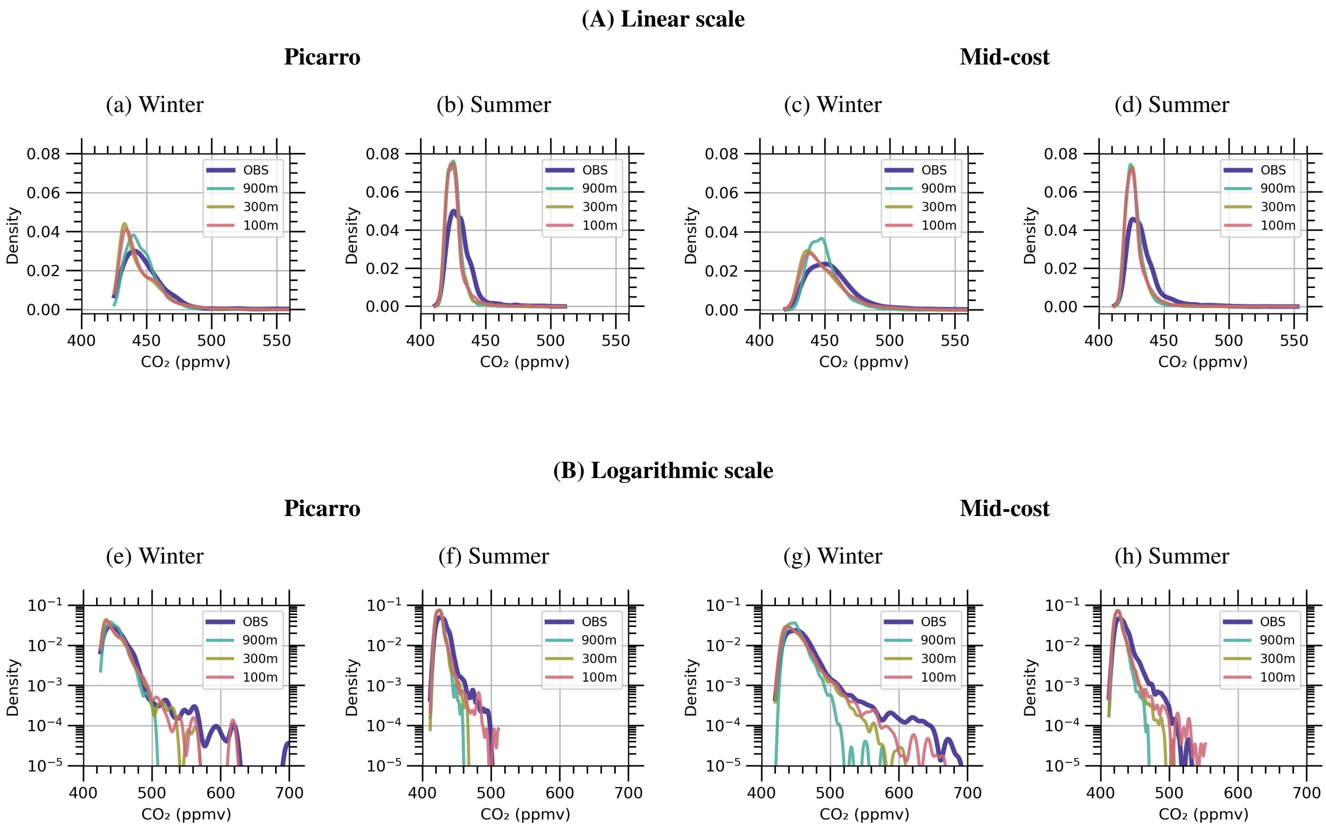


**Figure 10.** Probability density functions (PDFs) of $CO_2$ mixing ratios from WRF simulations and in-situ measurements aggregated across all stations. The left panels correspond to Picarro observations and the right panels to Mid-cost sensor measurements. Winter (January) and summer (June) distributions are shown for each instrument type. Panels (A) present PDFs on a linear scale, while panels (B) show the same distributions on a logarithmic scale to better highlight differences in the distribution tails and high concentration events.

#### 3.8.2 Taylor diagram analysis of WRF-simulated $CO_2$ mixing ratios compared to observations

Taylor diagrams (Figure 11) provide an integrated assessment of WRF-simulated $CO_2$ against observations from the Picarro (Figures 11a, 11b) and mid-cost (Figures 11c, 11d) networks for winter and summer 2024, allowing evaluation of how horizontal resolution affects the balance between capturing variability and minimizing error.

The results show that in both seasons the 100 m simulation exhibits the highest standard deviation, closely approaching the observed variability and indicating an improved representation of sub-kilometer urban $CO_2$ variability. In contrast, the 900 m mesoscale simulation substantially underestimates variability at both Picarro and mid-cost stations.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


For the Picarro stations, two distinct groups emerge in the 100 m and 300 m simulations. Stations located within Paris (JUS at 23 m and 33 m above ground, and CDS) exhibit standard deviations closely matching observations, demonstrating that high-resolution simulations effectively capture urban variability. Stations outside Paris, however, show variability similar to the 900 m mesoscale simulation, indicating a more limited impact of increased resolution at these sites. This pattern is consistent in both winter and summer. Correlation coefficients and RMSD remain relatively similar across resolutions in winter, whereas in summer, higher-resolution simulations improve correlations and reduce RMSD compared with the 900 m configuration. For urban stations, as illustrated by the time series analysis in section 3.6, higher-resolution simulations improve the representation of peak events and short-term variability, but may also introduce occasional overestimations, which can in turn influence the statistical metrics represented in the Taylor diagrams.

For the mid-cost network, a similar pattern is observed; however, since most stations are located within Paris, the variability exhibits a broader spread rather than forming distinct groups. In winter, the 900 m simulation produces slightly higher correlation coefficients and smaller RMSD compared with the 100 m and 300 m runs. In summer, by contrast, the higher-resolution simulations (300 m and 100 m) yield the highest correlations and lowest RMSD relative to the 900 m mesoscale configuration.

Overall, the 100 m simulation provides the closest representation of observed variability across both networks, whereas the 900 m run tends to smooth the temporal signal and underestimates the amplitude of short-term enhancements. The Taylor diagrams highlight the importance of horizontal resolution for representing urban $CO_2$ variability. Increasing resolution enhances the model's ability to reproduce observed temporal variability and, particularly in summer, leads to higher correlations and reduced RMSD relative to the 900 m configuration. In winter, correlations and RMSD remain relatively similar across resolutions, and higher-resolution simulations can occasionally exhibit slightly larger RMSD due to stronger representation of localized emission peaks and short-term fluctuations. Coarser resolutions, in contrast, smooth these fluctuations, leading to lower standard deviations and underestimation of extreme $CO_2$ enhancements while maintaining comparable correlation levels.

These findings reinforce earlier observations from the time series and PBLH analyses: higher resolution improves sensitivity to local variability and better captures standard deviations, whereas coarser resolutions underestimate variability. An important factor influencing these statistics is the interaction between grid resolution and station location, producing a "double penalty" effect. In high-resolution simulations (100 m), $CO_2$ peaks may be either missed or captured depending on whether the station falls within the grid cell containing the enhancement: a station outside the peak cell records lower concentrations, whereas a station intersecting the peak captures it accurately. In contrast, coarse-resolution simulations (900 m) smooth the plume over a larger area, reducing the likelihood of extreme over or underestimation. However, this smoothing masks fine-scale spatial variability, so apparent agreement with observations does not necessarily imply a more realistic representation of the underlying plume. This effect likely contributes to the observed differences in standard deviation, RMSD, and correlation across resolutions and between the Picarro and mid-cost sensor networks.

### 3.9 Impact of spatial aggregation on $CO_2$ and $XCO_2$ plume structure and implications for urban inversions

Having established the resolution dependent differences in plume structure, gradients, temporal variability, and statistical performance, we now examine whether spatial aggregation of high resolution simulations LES-run (100 m and 300 m) to 900 m

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


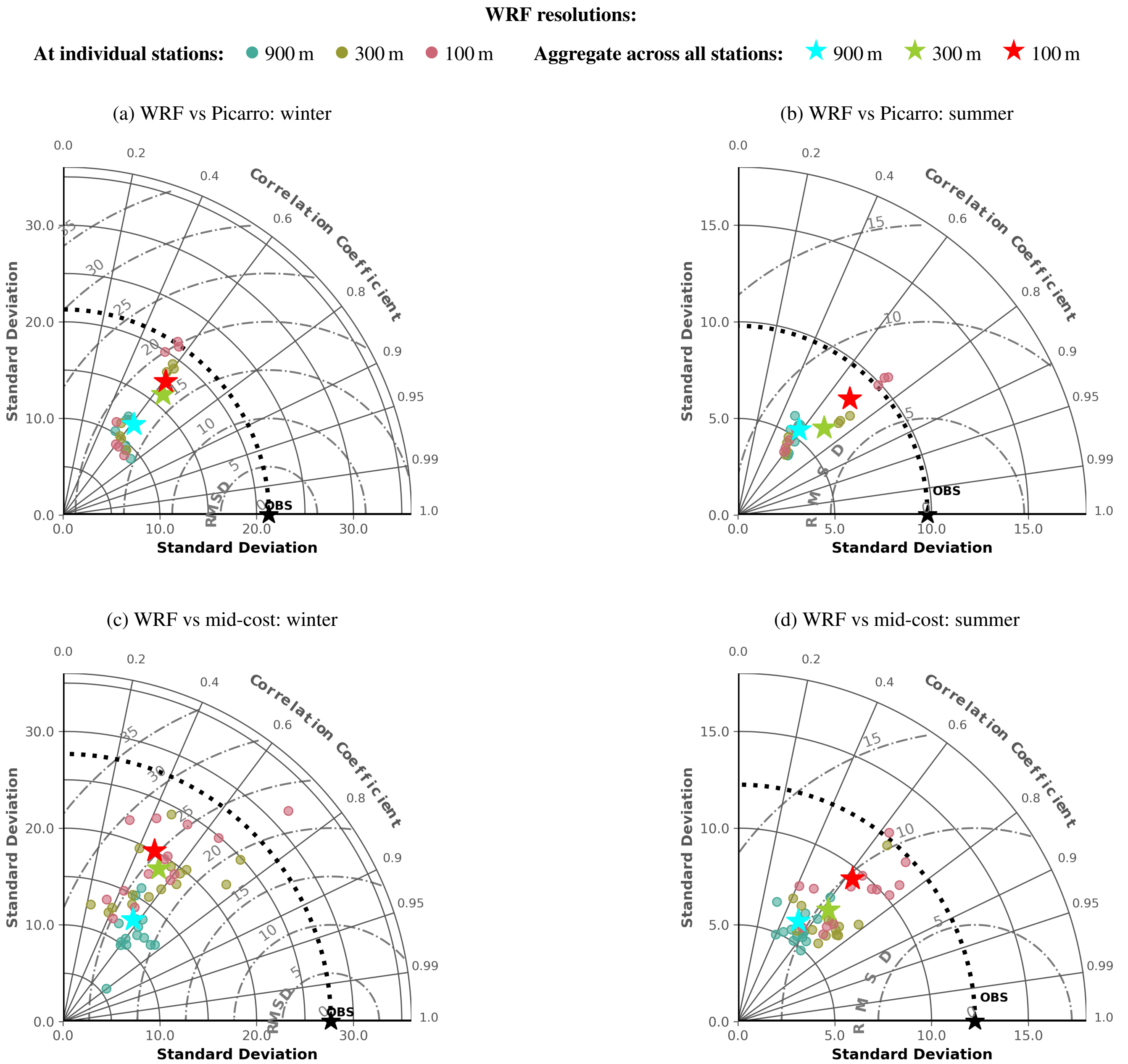


**Figure 11.** Taylor diagrams comparing WRF-simulated $CO_2$ mixing ratios with observations from the Picarro (a, b) and mid-cost (c, d) networks for winter (10–31 January) and summer (01–20 June) 2024. The diagrams summarize model performance in terms of correlation, standard deviation, and root-mean-square difference (RMSD) relative to observations. Colored symbols indicate the horizontal resolution of the WRF simulations: mesoscale at 900 m and LES at 300 m and 100 m. The corresponding colored stars show the aggregate statistics across all stations for each resolution, providing an overall measure of model skill.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


yields behavior comparable to the native mesoscale 900 m simulation. This analysis focuses on both near-surface $CO_2$ at 35 m above ground level and total column $XCO_2$ for winter (10–31 January 2024) and summer (01–20 June 2024). The 35 m level lies within the surface layer, above the first model level at 10 m, which helps avoid near-surface numerical noise while still capturing clearly expressed emission plumes. This framework allows us to assess how aggregation affects the representation of urban plume patterns relevant for inversion applications.

Table 3 provides an overview of domain-averaged consistency metrics for both tracers across seasons and periods, offering a quantitative basis for the spatial analysis presented in the following subsections.

**Table 3.** Domain averaged consistency metrics between aggregated high resolution LES run (100 m and 300 m) and native mesoscale 900 m WRF simulations for $CO_2$ mixing ratio at 35 m above ground level and $XCO_2$. Metrics are computed grid-cell by grid-cell over the time dimension and then averaged across the full domain, with minimum and maximum values reflecting the spatial variability across all grid cells. Results are reported for winter and summer, separated into daytime and nighttime periods. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are given in ppmv, symmetric relative error ($E_{\mathrm{sym}}$) in %, and $R$ denotes the Pearson correlation coefficient.

| Tracer | Period | Metric | **Winter** | | | | | | **Summer** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 300 m → 900 m | | | 100 m → 900 m | | | 300 m → 900 m | | | 100 m → 900 m | | |
| | | | Mean | Min | Max | Mean | Min | Max | Mean | Min | Max | Mean | Min | Max |
| **$CO_2$** | Daytime | RMSE | 11.64 | 4.38 | 30.27 | 10.99 | 4.65 | 36.05 | 2.45 | 0.75 | 13.01 | 2.84 | 0.65 | 24.13 |
| | | MAE | 7.32 | 2.84 | 22.89 | 6.73 | 2.97 | 25.67 | 1.12 | 0.19 | 9.18 | 1.29 | 0.16 | 15.40 |
| | | $E_{\mathrm{sym}}$ | 33.69 | 23.07 | 56.52 | 28.77 | 18.83 | 53.13 | 59.11 | 36.62 | 86.14 | 61.05 | 38.72 | 88.4 |
| | | $R$ | 0.67 | 0.27 | 0.92 | 0.70 | 0.27 | 0.92 | 0.74 | 0.48 | 0.93 | 0.68 | 0.36 | 0.90 |
| | Nighttime | RMSE | 15.50 | 7.60 | 44.45 | 16.58 | 7.68 | 52.77 | 6.75 | 2.03 | 18.50 | 8.21 | 2.13 | 23.96 |
| | | MAE | 9.21 | 4.62 | 24.40 | 9.56 | 4.81 | 28.24 | 3.70 | 1.15 | 10.94 | 4.18 | 1.11 | 14.27 |
| | | $E_{\mathrm{sym}}$ | 29.08 | 23.58 | 38.81 | 27.60 | 22.51 | 38.45 | 55.94 | 41.10 | 63.21 | 55.68 | 42.17 | 63.74 |
| | | $R$ | 0.63 | 0.38 | 0.80 | 0.63 | 0.47 | 0.78 | 0.57 | 0.26 | 0.80 | 0.54 | 0.11 | 0.80 |
| **$XCO_2$** | Daytime | RMSE | 0.47 | 0.19 | 1.10 | 0.49 | 0.22 | 1.15 | 0.15 | 0.06 | 0.38 | 0.17 | 0.06 | 0.50 |
| | | MAE | 0.26 | 0.11 | 0.63 | 0.28 | 0.12 | 0.69 | 0.05 | 0.01 | 0.19 | 0.05 | 0.01 | 0.23 |
| | | $E_{\mathrm{sym}}$ | 21.13 | 13.76 | 33.65 | 22.06 | 15.90 | 33.50 | 57.92 | 45.04 | 72.10 | 59.27 | 44.34 | 73.18 |
| | | $R$ | 0.79 | 0.57 | 0.94 | 0.78 | 0.54 | 0.94 | 0.75 | 0.41 | 0.92 | 0.71 | 0.36 | 0.92 |
| | Nighttime | RMSE | 0.29 | 0.19 | 0.52 | 0.34 | 0.21 | 0.63 | 0.04 | 0.02 | 0.15 | 0.06 | 0.02 | 0.20 |
| | | MAE | 0.16 | 0.10 | 0.33 | 0.20 | 0.12 | 0.42 | 0.02 | 0.01 | 0.07 | 0.02 | 0.01 | 0.09 |
| | | $E_{\mathrm{sym}}$ | 15.14 | 11.25 | 25.72 | 18.09 | 14.81 | 25.22 | 48.26 | 38.54 | 56.15 | 50.71 | 38.20 | 63.02 |
| | | $R$ | 0.85 | 0.71 | 0.91 | 0.80 | 0.64 | 0.89 | 0.90 | 0.51 | 0.97 | 0.85 | 0.46 | 0.97 |

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


#### 3.9.1 Effect of spatial aggregation on near surface $CO_2$ (35 m AGL)

As summarized in Table 3, aggregation reproduces the main spatial and temporal patterns of CO2, but noticeable differences remain. RMSE values are systematically larger in winter than in summer, reflecting enhanced variability under stable boundary-layer conditions. The symmetric relative error ($E_{\mathrm{sym}}$) values are higher in summer, consistent with a weaker atmospheric signal and reduced absolute concentration gradients. Correlation coefficients exhibit a clear dependence on diurnal stability, with generally higher agreement during daytime and reduced performance at night, highlighting the influence of boundary-layer structure on aggregation consistency. Overall, while aggregation captures the dominant behavior of the native 900 m simulation, it does not fully reproduce the variability and magnitude of near-surface $CO_2$, with discrepancies varying with season and time of day. These differences are further examined through the spatial patterns and their analysis below.

Figure 12 presents the mean $CO_2$ enhancement fields at 35 m above ground level, with winter daytime and nighttime cases given in Figures 12a and 12b. A clear contrast between daytime and nighttime conditions is evident. During winter daytime, mean concentrations are generally lower and exhibit similar large scale spatial patterns across all three configurations. The aggregated 100 m and 300 m simulations produce slightly higher and more spatially concentrated enhancements near emission hotspots, whereas the native 900 m run displays a broader and more spatially diffused plume structure. This behavior is consistent with a well-developed and efficiently mixed planetary boundary layer (PBL), as indicated by comparable daytime PBL heights across resolutions, with the 900 m configuration showing a slight overestimation. Under these convective conditions, enhanced vertical mixing reduces concentration gradients and limits resolution dependent differences.

In contrast, winter nighttime conditions (Figure 12b) reveal substantially larger discrepancies. The aggregated high-resolution simulations exhibit markedly higher $CO_2$ mixing ratios over the Paris area compared to the native 900 m run. This reflects the shallower nocturnal boundary layers simulated at 100 m and 300 m resolution, consistent with the earlier analysis of the diurnal PBLH cycle. Reduced vertical mixing at night amplifies near-surface accumulation, sharpens concentration gradients, and intensifies local plume maxima. The native 900 m simulation, by contrast, produces a smoother and weaker enhancement due to stronger effective mixing and spatial averaging. These pronounced differences arise because, while vertical resolution is consistent across simulations, the effective transport differs with horizontal resolution. Higher-resolution simulations better resolve local flow structures and turbulence intermittency, which limits vertical mixing in shallow nocturnal boundary layers and promotes $CO_2$ accumulation near emission sources. It is worth noting that, as described in the methodology (Section 2.2), emissions are released at the lowest model level (10 m); this may amplify near-surface accumulation, particularly under the stable nocturnal conditions better resolved at finer resolution. Additionally, the 35 m observation height corresponds to the second model level, making it sensitive to reduced mixing in the lowermost layer, whereas the coarser 900 m grid effectively reduces concentration gradients due to horizontal averaging across the larger grid cells.

Figures 12c and 12d present the corresponding summer daytime and nighttime averages. The overall behavior is qualitatively similar to winter: aggregated high resolution simulations maintain more concentrated plume structures and higher local amplitudes compared to the native 900 m configuration. However, absolute concentration enhancements are substantially lower in summer. The winter daytime mean near surface enhancements (averaged over 10–31 January 2024) locally exceed

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


45 ppmv, whereas summer enhancements are generally below 15 ppmv, reflecting stronger atmospheric mixing and reduced net accumulation during the warmer season.

These results demonstrate that spatial aggregation of high resolution simulations does not fully reproduce the plume structure and magnitude of a native 900 m simulation. Differences are modest during well mixed daytime conditions but become pronounced at night, when boundary layer dynamics strongly control plume confinement and amplification.

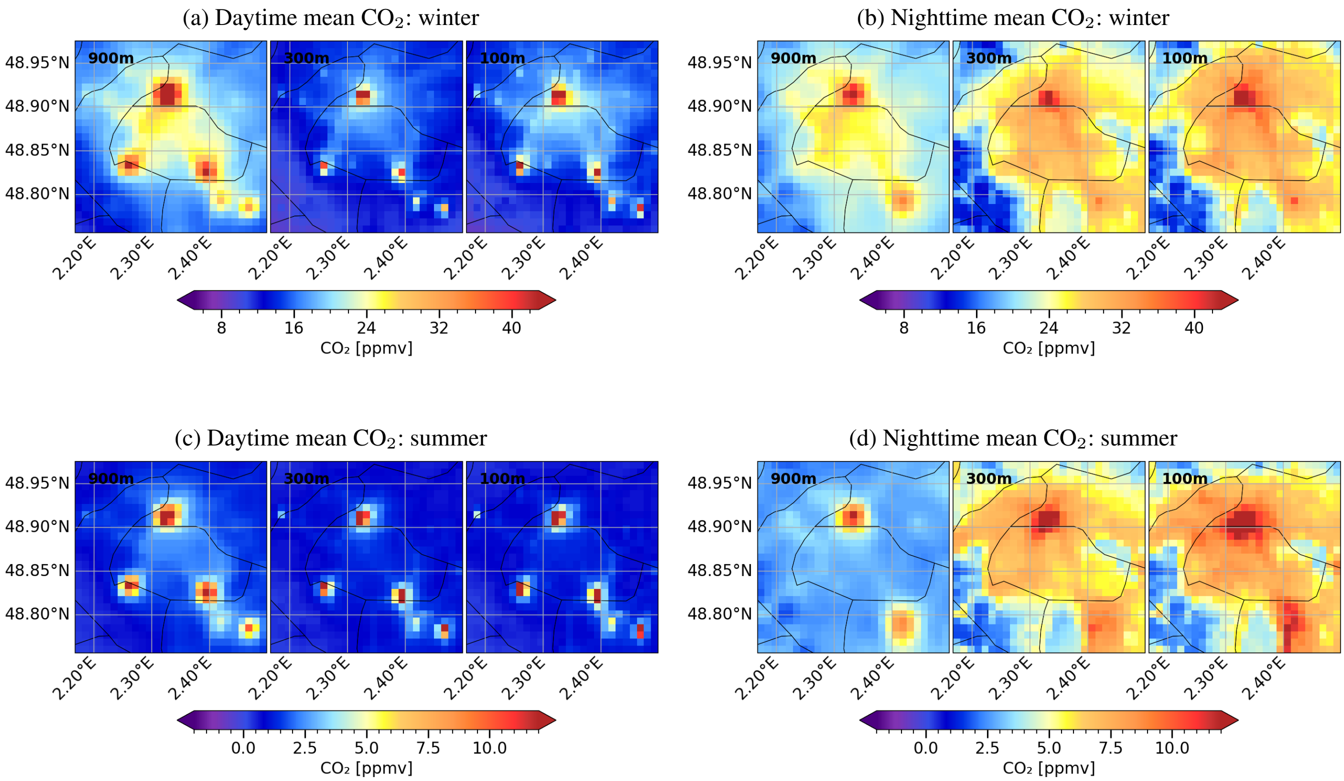


**Figure 12.** Mean $CO_2$ enhancement fields at 35 m above ground level for winter (10–31 January 2024) and summer (01–20 June 2024). Panels show daytime (left) and nighttime (right) averages for each season. Each map compares the native 900 m simulation with the 100 m and 300 m simulations aggregated to 900 m resolution. The fields represent $CO_2$ mixing ratio (background subtracted). The figure illustrates how spatial aggregation influences plume magnitude, spatial extent, and concentration gradients under contrasting boundary layer conditions.

Figures 15a and 15b present the spatial distribution of $E_{\mathrm{sym}}$ between the aggregated (100 m and 300 m mapped to 900 m) and the native 900 m simulations for winter daytime and nighttime conditions. These maps complement the mean fields shown in Figure 12 by quantifying where aggregation preserves plume structure and where systematic distortions occur.

During winter daytime, $E_{\mathrm{sym}}$ values are primarily concentrated within the urban core of Paris and in the vicinity of major point sources. These sources form an approximate triangular configuration, with prominent emitters located north, southwest, and southeast of the Paris metropolitan area, as well as additional sources farther southeast. In these regions, maximum $E_{\mathrm{sym}}$

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


locally exceed 40%. The spatial pattern closely mirrors the enhanced and more concentrated plume structures seen in the aggregated simulations in Figure 12. $E_{\mathrm{sym}}$ values are highest along sharp emission gradients, where sub-grid variability resolved at 100–300 m cannot be perfectly preserved through spatial averaging. In contrast, areas weakly influenced by plumes exhibit substantially lower $E_{\mathrm{sym}}$ values, often below 20%, indicating that background regions and smoothly varying fields are more robust to aggregation.

Winter nighttime conditions (Figure 15b) display a different structure. $E_{\mathrm{sym}}$ becomes more spatially homogeneous, with values typically ranging between 28–34% across much of the domain. Although elevated errors remain near major point sources, the contrast between plume affected and background areas is less pronounced than during daytime. This behavior is consistent with the shallower nocturnal boundary layer discussed previously: enhanced confinement increases overall concentration levels in the aggregated high resolution runs, producing broader domain wide differences relative to the native 900 m simulation rather than sharply localized discrepancies.

A similar but seasonally amplified behavior is observed in summer (Figures 15c and 15d). During daytime, $E_{\mathrm{sym}}$ values again peak near strong emission sources, but values can reach up to 70%, substantially higher than in winter. This increase reflects the smaller absolute plume amplitudes in summer (Figure 12); when background mixing is stronger and perturbations are weaker, even moderate absolute differences translate into larger $E_{\mathrm{sym}}$ values. Interestingly, during summer nighttime, the near point source grid cells sometimes exhibit lower $E_{\mathrm{sym}}$ values compared to surrounding areas. This suggests that under certain stable conditions, aggregation can partially smooth extreme local maxima in a way that reduces proportional differences relative to the native coarse resolution simulation.

An interesting feature is that $E_{\mathrm{sym}}$ at the exact grid cell containing a major point source is sometimes lower than in the surrounding grid cells. This behavior likely reflects differences in plume concentration and spatial spreading between resolutions. In the high resolution simulations (100–300 m), emissions remain strongly concentrated near the source before being aggregated to 900 m. When averaged to the coarser grid, the central enhancement may still align reasonably well with the native 900 m maximum, resulting in a comparatively smaller proportional difference at that specific grid cell. In contrast, the surrounding grid cells are more sensitive to differences in plume dispersion and wind direction. As discussed in Section 3.6, a bias in wind direction can lead to lateral plume displacement, and this effect is more pronounced at finer resolution where concentration gradients are stronger. Since time-averaged enhancements in neighboring cells depend on how frequently and consistently the plume is advected toward them, resolution-dependent differences in plume width and lateral dilution are amplified in the surrounding cells relative to the source cell itself. The native 900 m simulation tends to produce a broader and smoother plume due to numerical diffusion and coarser mixing, while the aggregated high-resolution fields retain sharper gradients and more localized structure, further contributing to larger $E_{\mathrm{sym}}$ values in neighboring cells. This pattern therefore reflects not only differences in plume magnitude, but also differences in spatial distribution and dispersion characteristics across resolutions.

Overall, the $E_{\mathrm{sym}}$ analysis confirms the conclusions drawn from the mean fields: spatial aggregation does not uniformly reproduce the native 900 m plume structure. $E_{\mathrm{sym}}$ values are strongly modulated by boundary layer dynamics and plume intensity. Daytime differences are primarily localized around sharp emission gradients, whereas nighttime differences reflect

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


broader structural changes linked to vertical confinement and mixing. Moreover, the larger $E_{sym}$ values in summer highlight the sensitivity of proportional metrics to seasonal variations in plume amplitude.

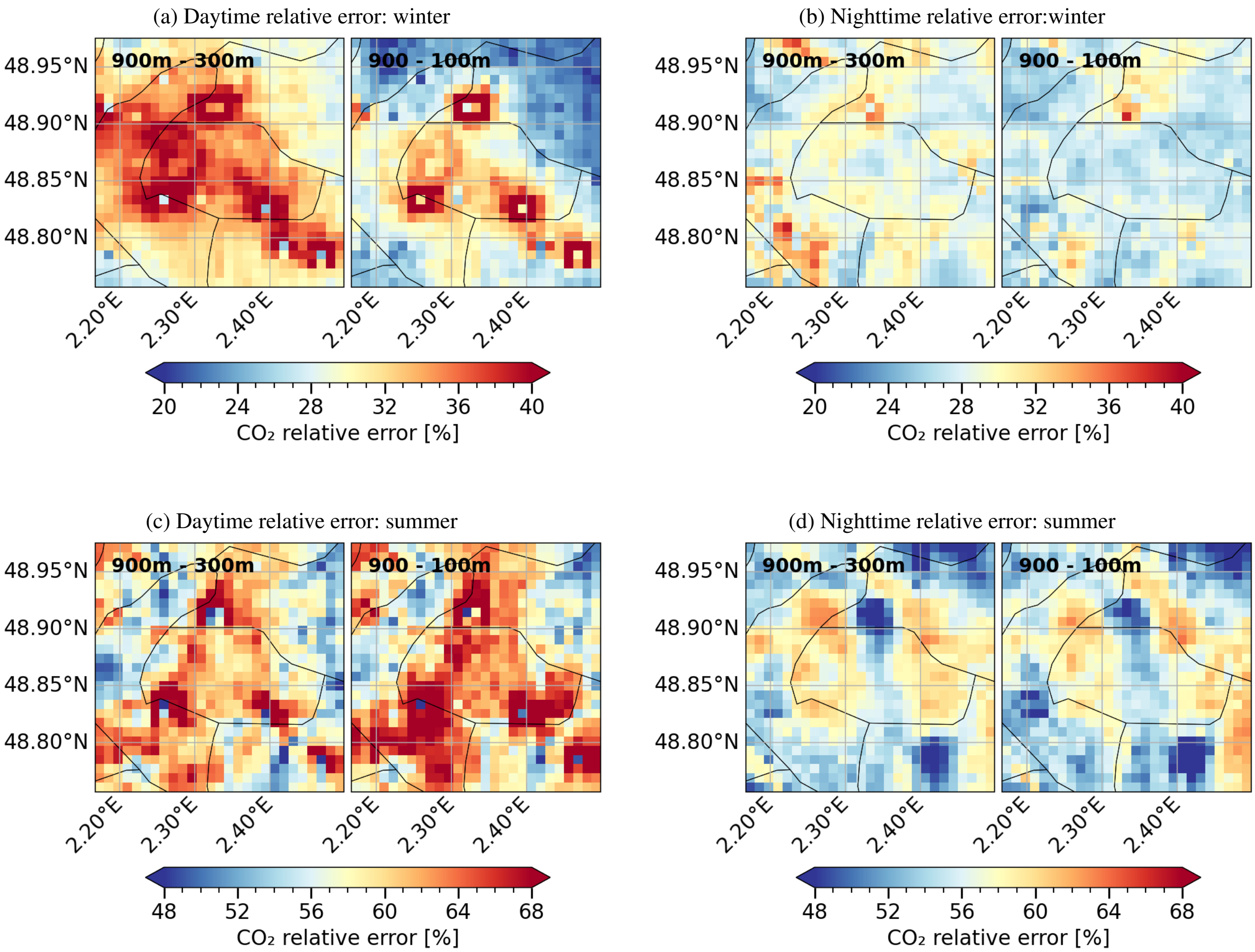


**Figure 13.** Spatial distribution of $E_{sym}$ between aggregated high resolution simulations (100 m and 300 m mapped to 900 m) and the native 900 m simulation for $CO_2$ at 35 m above ground level. Panels show daytime (left) and nighttime (right) conditions for winter (top row; 10–31 January 2024) and summer (bottom row; 01–20 June 2024). $E_{sym}$ is computed on $CO_2$ perturbations (background subtracted) to isolate plume differences. The figure highlights how spatial aggregation modifies plume magnitude and structure under contrasting seasonal and boundary layer regimes.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


#### 3.9.2 Effect of spatial aggregation on vertically integrated $XCO_2$

To evaluate consistency at the column scale relevant for satellite retrievals, the same analysis was applied to $XCO_2$. As indicated in Table 3, $XCO_2$ exhibits substantially lower RMSE and reduced variability compared to near-surface CO2, consistent with the smoothing effect of vertical column integration. Correlation coefficients indicate good overall agreement between aggregated and reference simulations, with generally higher values during daytime and reduced performance at night, reflecting a stronger sensitivity to boundary-layer stability even at the column scale. Summer conditions are further characterized by increased $E_{\mathrm{sym}}$, consistent with weaker atmospheric gradients and reduced signal strength. Overall, aggregation captures the dominant large-scale structure of $XCO_2$, although discrepancies are more pronounced in summer. The spatial characteristics of these differences are discussed in the following paragraphs.

Figures 14a and 14b show that in winter, daytime $XCO_2$ enhancements typically range between about 0.5 and 2 ppmv, while nighttime values are slightly lower (approximately 0.5–1.5 ppmv). The diurnal contrast is therefore opposite to that observed near the surface: daytime column enhancements are larger because the deeper and more efficiently mixed boundary layer allows surface emissions to influence a larger fraction of the atmospheric column. At night, the shallow stable boundary layer traps $CO_2$ close to the surface; since $XCO_2$ retrievals are less sensitive to near-surface layers due to the shape of the averaging kernel, the latter may amplify apparent differences by weighting plumes differently in the vertical, especially when high-resolution simulations place peaks slightly above or below the kernel's center of sensitivity.

Despite the reduced amplitude, resolution-dependent differences remain visible in the spatial structure of $XCO_2$. Aggregated high-resolution simulations (100 m and 300 m mapped to 900 m) display finer spatial gradients and localized plume signatures that are largely smoothed in the native 900 m simulation. As in the near-surface $CO_2$ case, this behavior reflects the fact that simple spatial aggregation does not fully reproduce the dynamical smoothing inherent in a coarser-resolution model run. Sub-grid heterogeneities in emissions and vertical mixing remain partially preserved after aggregation, producing residual small-scale variability that would otherwise be diffused in the native coarse-resolution simulation. Additional differences can also arise from small displacements of emission plumes due to local wind patterns or variations in PBLH, which can alter the period-averaged $XCO_2$ maps.

During summer (Figures 14c, 14d), the same structural behavior is observed but with substantially smaller amplitudes. Daytime $XCO_2$ enhancements generally range from about 0.02 to 0.35 ppmv, while nighttime values are typically very small (around 0.03–0.14 ppmv), approaching zero over most of the domain except in the immediate vicinity of strong emission sources. As in winter, differences between aggregated and native simulations remain discernible, indicating that the effects of spatial aggregation on plume structure persist even when the column signal itself is weak.

These results are consistent with the behavior observed for $CO_2$ at 35 m: aggregation of high-resolution simulations does not fully reproduce the plume magnitude and spatial structure of a native coarse resolution run. Although the discrepancies are smaller for $XCO_2$ due to vertical averaging, residual spatial variability and localized enhancements remain detectable, highlighting that aggregation and native coarse resolution simulations are not dynamically equivalent.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


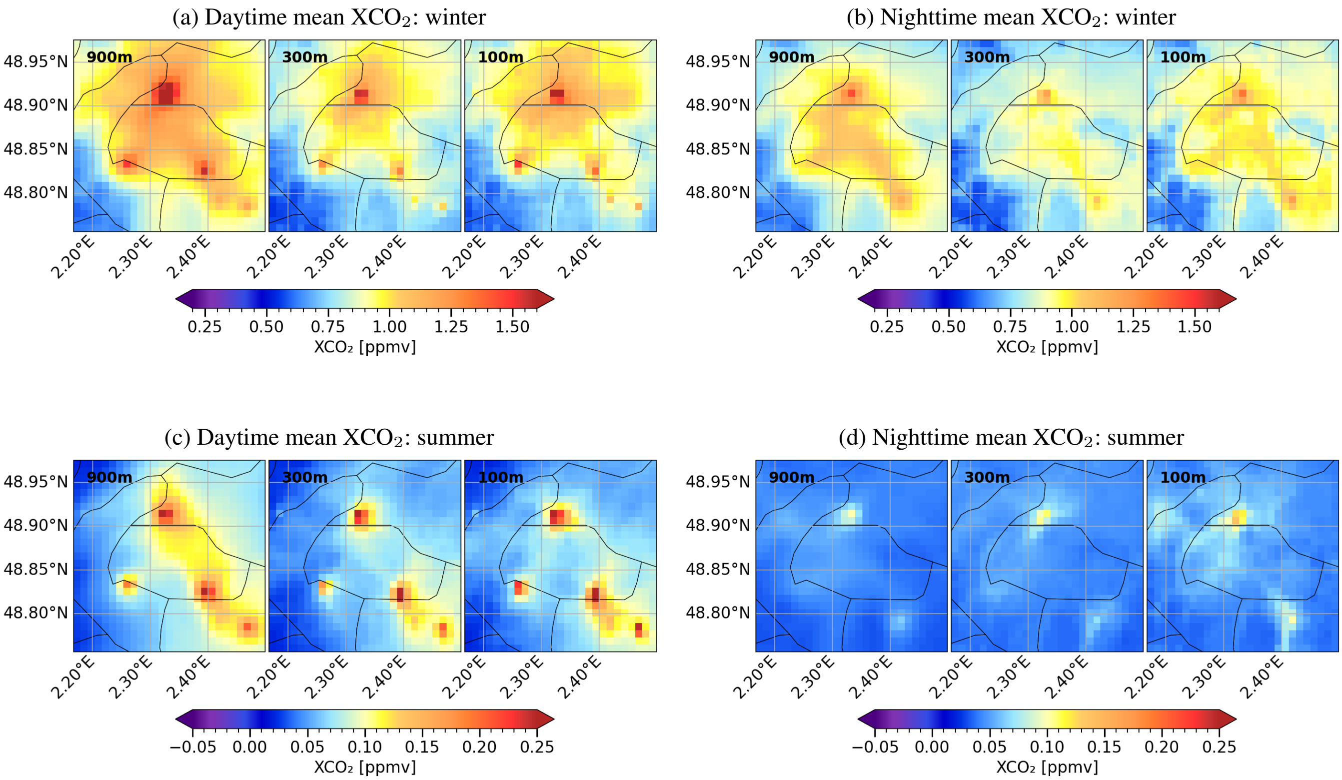


**Figure 14.** Mean $XCO_2$ enhancement fields for winter (10–31 January 2024) and summer (01–20 June 2024). Panels show daytime (left) and nighttime (right) averages for each season. Each map compares the native 900 m simulation with the 100 m and 300 m simulations aggregated to 900 m resolution. The fields represent $XCO_2$ (background subtracted). The figure illustrates how spatial aggregation influences plume magnitude, spatial extent, and gradients under contrasting boundary layer conditions.

Quantitatively, aggregation from 100 m to 900 m produces systematically larger $E_{\mathrm{sym}}$ values than aggregation from 300 m to 900 m, reflecting the larger scale mismatch and the stronger small-scale variability resolved at finer resolution (Figure 15).

For winter (10–31 January 2024; Figures 15a and 15b), daytime $E_{\mathrm{sym}}$ values typically range from about 13–34% for the 300 m to 900 m aggregation (mean ∼21%) and from roughly 16–34% for the 100 m to 900 m aggregation (mean ∼22%). Nighttime $E_{\mathrm{sym}}$ values are generally smaller, averaging about 15% (11–26%) for the 300 m aggregation and about 18% (15–25%) for the 100 m aggregation. The spatial distribution of these $E_{\mathrm{sym}}$ values closely reflects the structure of the mean $XCO_2$ fields discussed earlier (Figures 14a and 14b). In particular, the largest $E_{\mathrm{sym}}$ values occur in and around the Paris urban area, where strong local emission gradients and complex sub-grid variability are present. The spatial distribution of high-emission facilities around Paris, particularly those situated along the city's periphery, appears to shape the observed error patterns.

During daytime, the deeper and well-mixed boundary layer allows emissions from these sources to influence a larger fraction of the atmospheric column, leading to stronger horizontal gradients and larger discrepancies between aggregated and native

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


simulations. At night, the boundary layer becomes shallow and stable, trapping emissions near the surface and reducing their contribution to the column integrated signal over the urban core. Consequently, $E_{\mathrm{sym}}$ values over central Paris are generally lower at night than during the day, although elevated $E_{\mathrm{sym}}$ values persist near individual point sources where nocturnal plumes remain highly localized and therefore sensitive to model resolution.

For summer (01–20 June 2024; Figures 15c and 15d), $E_{\mathrm{sym}}$ values are substantially larger, with daytime mean $E_{\mathrm{sym}}$ of about 59% for the 100 m to 900 m aggregation (range 44–73%) and about 58% for the 300 m aggregation (45–72%). Nighttime $E_{\mathrm{sym}}$ values remain similarly high, averaging approximately 51% (38–63%) for the 100 m aggregation and about 48% (39–56%) for the 300 m aggregation. These larger $E_{\mathrm{sym}}$ values are primarily a consequence of the much weaker $XCO_2$ signal in summer, as shown in the mean fields (Figures 14c and 14d). Because column enhancements are typically very small during this period, even minor absolute differences between aggregated and native simulations translate into large $E_{\mathrm{sym}}$ values.

Overall, these results confirm the behavior already suggested by the mean $XCO_2$ fields: although aggregation preserves the general spatial structure of the column signal, it does not fully reproduce the magnitude and spatial variability simulated by the native coarse resolution model. This difference reflects the fact that simple spatial averaging cannot replicate the effective smoothing and transport representation inherent to the native 900 m simulation, where sub-grid emission heterogeneity, atmospheric mixing, and local transport processes are dynamically represented at the model scale. The discrepancy increases as the aggregation scale gap widens (100 m versus 300 m) and becomes particularly pronounced when the underlying atmospheric signal is weak, as in summer. Nevertheless, the preservation of localized spatial structures suggests that aggregated high resolution simulations may still provide valuable information for inversion frameworks seeking to retain emission driven spatial signals while operating at coarser computational grids.

## 4 Discussion

Because atmospheric transport and dispersion are primarily controlled by wind fields and boundary-layer dynamics, the ability of the model to reproduce near-surface meteorological conditions is a critical prerequisite for accurately simulating urban $CO_2$ distributions. The comparison with observations shows that the different model configurations reproduce the overall temporal variability of wind speed, temperature and relative humidity but systematic differences emerge as the horizontal resolution changes. The 900 m simulation generally provides smoother wind fields and smaller variability due to the spatial averaging inherent to coarser grid spacing. In contrast, the higher-resolution simulations (300 m and 100 m) resolve finer-scale flow structures and stronger local variability. While this increased variability improves the representation of localized flow patterns, it can also lead to larger point-wise errors when evaluated against individual meteorological stations. These results highlight the challenges associated with the so-called turbulence "gray zone", where model grid spacing becomes comparable to the dominant turbulent eddies but is still too coarse to fully resolve them. Similar sensitivity to meteorological forcing has been reported in the intercomparison of high-resolution transport models by Brunner et al. (2023) which showed that even turbulence resolving models remain strongly dependent on boundary layer structure and wind field accuracy.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


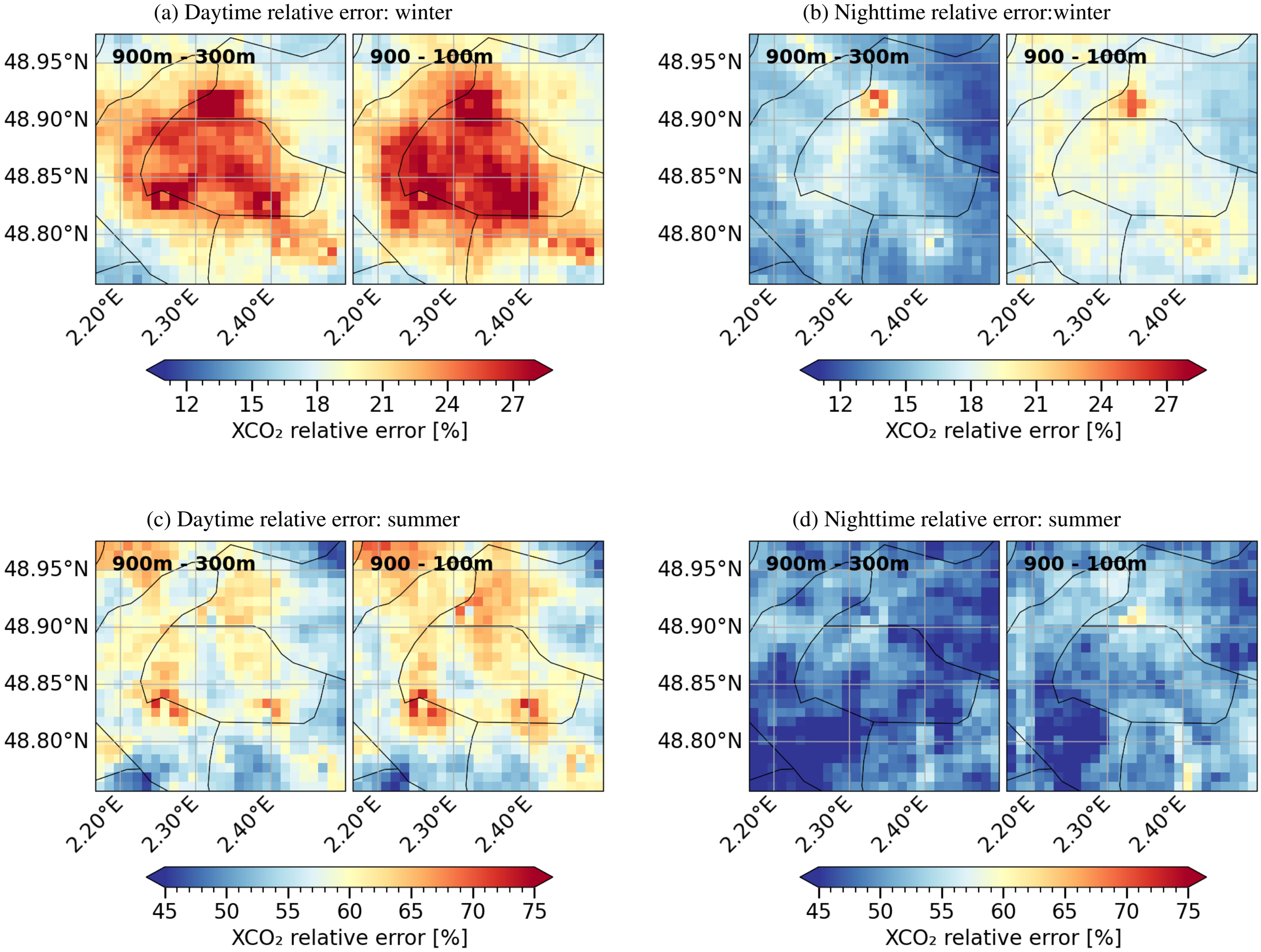


**Figure 15.** Spatial distribution of $E_{\mathrm{sym}}$ between aggregated high resolution LES simulations (100 m and 300 m mapped to 900 m) and the native 900 m simulation for $XCO_2$. Panels show daytime (left) and nighttime (right) conditions for winter (top row; 10–31 January 2024) and summer (bottom row; 01–20 June 2024). $E_{\mathrm{sym}}$ is computed on $XCO_2$ perturbations (background subtracted) to isolate plume differences. The figure highlights how spatial aggregation modifies plume magnitude and structure under contrasting seasonal and boundary layer regimes.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


The results presented in this study demonstrate that horizontal resolution strongly influences the simulation of urban $CO_2$ plumes and atmospheric transport over the Paris metropolitan region. Differences between mesoscale (900 m) and turbulence-resolving simulations (300 m and 100 m) affect plume morphology, concentration variability, and agreement with observations. These findings highlight both the advantages and limitations of high-resolution modeling for representing urban atmospheric transport and interpreting multi-scale observations.

The comparison between winter and summer simulations highlights the strong role of boundary-layer dynamics in controlling the sensitivity of $CO_2$ fields to model resolution. During winter, the planetary boundary layer is typically shallow and frequently stable, which limits vertical mixing and promotes the accumulation of $CO_2$ near the surface. Under these conditions, horizontal concentration gradients become stronger and differences between model resolutions are amplified. High-resolution simulations therefore produce larger local enhancements and sharper plume structures than the 900 m configuration. In contrast, summer conditions are characterized by deeper and more convective boundary layers, which enhance vertical mixing and reduce near-surface gradients. As a result, differences between model resolutions become less pronounced. This seasonal dependence is consistent with observational evidence reported in Mitchell et al. (2018), where boundary-layer dynamics were identified as a primary driver of temporal variability in urban $CO_2$. It should be noted that differences in PBLH definitions between model and observations may also contribute to discrepancies: in the simulations, PBLH is computed from model output using a bulk Richardson number criterion, whereas observations are derived from aerosol-layer detection using STRATfinder. Because this gradient-based approach may occasionally identify elevated residual layers rather than the layer closest to the surface (Kotthaus et al., 2020), observed nighttime PBLH may be subject to additional uncertainty. These results reinforce the conclusion that the benefits of high-resolution modeling are strongest under stable conditions when transport is less diffusive and plume structures are more distinct.

The analysis of plume morphology confirms that coarse resolution simulations tend to smooth concentration fields and dilute localized emission signals. In the 900 m simulation, emission plumes appear broader and less intense, reflecting numerical diffusion and spatial averaging across heterogeneous emission sources. In contrast, the LES simulations at 300 m and 100 m preserve sharper gradients and stronger concentration peaks, providing a more detailed representation of urban emission variability. However, the improved representation of plume structure at high resolution also increases sensitivity to errors in wind direction and emission location. Even slight shifts in plume position relative to observation sites can produce substantial differences in simulated concentrations. Consequently, discrepancies between modeled and observed concentrations at individual stations may partly reflect representativeness effects rather than purely model deficiencies.

Comparisons with the ground-based observation networks reveal a trade-off between simulated variability and conventional error metrics. The 100 m simulation reproduces observed temporal variability more realistically, reflecting its ability to resolve localized plumes and fine-scale transport structures. However, this increased variability can also produce larger apparent errors when a monitoring station lies outside the simulated plume grid cell. Conversely, the 900 m simulation averages concentrations over larger grid cells, smoothing plume structures and reducing variability. While this smoothing decreases the physical realism of the simulated plume field, it may incidentally improve statistical metrics such as root-mean-square deviation (RMSD), particularly during winter when concentration gradients are strongest. A similar trade-off was identified in Brunner et al. (2023),

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


where models with more realistic plume variability did not necessarily outperform simpler models in standard evaluation metrics. This apparent contradiction reflects representativeness errors, which are also emphasized in Turnbull et al. (2019) where mismatches between model resolution and observation scale were shown to significantly affect inferred emissions. Together, these studies highlight the need to interpret model–data mismatches in the context of spatial representativeness rather than purely model skill.

The total column $XCO_2$ comparisons against TCCON measurements demonstrate that resolution effects propagate vertically through the atmospheric column. LES simulations reproduce enhanced variability in total column $CO_2$, indicating that improved representation of near-surface plume structures can influence the vertically integrated signal. Nevertheless, occasional mismatches highlight the sensitivity of $XCO_2$ to boundary-layer dynamics and vertical mixing processes.

Aggregation experiments highlight that averaging high-resolution simulations to a coarser grid does not reproduce the same spatial and temporal structures as a native mesoscale simulation. Although aggregation preserves the general location of major plumes, it significantly modifies plume magnitude and spatial gradients because the underlying atmospheric transport and turbulent mixing processes differ across resolutions. In other words, the coarse-resolution simulation represents transport processes that are dynamically consistent with its grid spacing, whereas the aggregated LES fields simply smooth fine-scale variability without reproducing the corresponding mesoscale transport dynamics. This difference is reflected in the magnitude of the aggregation errors. For near-surface $CO_2$, domain-mean $E_{\mathrm{sym}}$ values reach approximately 30–35% during winter daytime conditions and exceed 50% during summer when plume amplitudes are weaker. For column-integrated $XCO_2$, vertical integration reduces absolute differences but substantial discrepancies remain, particularly near strong emission sources. The persistence of these differences indicates that sub-grid plume variability contributes significantly to the simulated concentration field and cannot be fully represented by simple spatial averaging. The larger $E_{\mathrm{sym}}$ values observed in summer primarily reflect the weaker absolute signal associated with enhanced convective mixing. When plume amplitudes are small, even modest absolute differences introduced by aggregation translate into larger $E_{\mathrm{sym}}$ values. In winter, by contrast, stronger emission accumulation within the shallow boundary layer produces larger plume amplitudes and more coherent spatial gradients, which reduces the relative impact of aggregation. These findings highlight an important limitation when comparing models and observations operating at different spatial scales. Satellite retrievals and some observational networks integrate signals over footprints that are substantially larger than the spatial scales resolved by LES simulations. If the model fields are aggregated to match these footprints without accounting for sub-grid plume variability, significant representation errors may be introduced. Such mismatches can propagate directly into atmospheric inversion frameworks, where the relationship between emissions and observed concentrations depends critically on the spatial consistency between model transport fields and observational footprints. Overall, the aggregation experiments demonstrate that high-resolution simulations contain physically meaningful variability that is not preserved when fields are simply averaged to coarser grids. This result emphasizes the need for careful treatment of scale mismatches in urban greenhouse gas monitoring systems and inversion frameworks.

Although the 100 m simulations provide the most detailed representation of plume structure and variability, this resolution does not fully resolve the entire spectrum of urban turbulence. Interactions between urban morphology and atmospheric turbulence often occur on spatial scales smaller than 100 m. Processes such as street-canyon circulations and localized shear

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


layers therefore remain partially unresolved and must still be parameterized, introducing uncertainties in near-surface transport and dispersion. Additional uncertainties arise from emission inventories and meteorological conditions. Small positional errors in emission sources or uncertainties in temporal emission can strongly influence simulated concentrations in high-resolution models. Similarly, errors in wind speed, wind direction, or planetary boundary layer height can substantially modify plume transport and dispersion. This trade-off has also been emphasized in prior urban $CO_2$ studies, where higher-resolution models improved plume realism but exposed the limitations of comparing grid-cell averages directly with point observations (Gaudet et al., 2017).

Despite these limitations, the results demonstrate the clear added value of high-resolution modeling for capturing fine-scale plume structures and intra-urban variability. Such capabilities are particularly important for atmospheric inversion systems and satellite-based monitoring frameworks that aim to attribute emissions to specific urban sectors or sources. However, the combined evidence from this study and prior work (Lauvaux et al., 2016; Gaudet et al., 2017; Turnbull et al., 2019; Brunner et al., 2023) indicates that careful consideration must be given to spatial resolution, observational representativeness, and scale mismatch. In particular, the interaction between model resolution and observational footprint remains a key challenge for future urban greenhouse gas monitoring systems.

## 5 Conclusion

This study evaluated the impact of the model resolution and physics on the simulation of $CO_2$ plumes over Paris using the WRF-GHG framework at resolutions of 900 m, 300 m and 100 m. Our results demonstrate that resolution strongly influences plume fidelity. High-resolution runs (100 m) better capture plume structures and gradients compared to mesoscale simulations, reducing numerical diffusion and capturing episodic emission events. However, these benefits come with trade-offs: finer resolutions increase computational cost, and may amplify biases related to PBL dynamics. Higher-resolution runs reproduce observed variability more realistically but are more sensitive to representativeness errors at observation sites, whereas coarser (900 m) simulations smooth the variability, producing reduced physical realism.

Importantly, our results indicate that the differences observed between LES and mesoscale simulations are not primarily due to model physics, but rather to aggregation errors associated with coarse spatial averaging of heterogeneous emission fields. While the 900 m configuration provides better meteorological performance, especially for wind speed and direction, it fails to represent fine-scale emission structures because local point sources are spatially diluted. Thus, the discrepancy between resolutions reflects the inability of coarse grids to aggregate high-resolution emission variability, rather than a deficiency of the transport model itself.

Boundary layer representation is a critical factor controlling model performance. In winter, mesoscale (900 m) simulations systematically overestimate daytime PBLH by 100–250 m, whereas LES (100 m and 300 m) better capture both the diurnal cycle and urban spatial gradients, particularly within the dense urban core. At night, LES tends to underestimate PBLH, contributing to excessive surface $CO_2$ accumulation. In summer, morning PBLH growth is well captured by all simulations, but LES underestimates afternoon maxima, and the 900 m simulation declines slightly earlier than observations.

Our results highlight the importance of improved turbulence schemes and assimilation of vertical profile data. Vertical and horizontal resolution effects are tightly coupled: errors in near-surface transport propagate to the total column $XCO_2$ and inter-station gradients, affecting both surface network interpretation and satellite-based inversion accuracy.

Finally, the analysis underscores the importance of observation representativeness in model evaluation. Model–observation agreement depends on how sensor locations align with grid cells, as small spatial offsets can lead to large apparent discrepancies at high resolution. Taken together, our results indicate that horizontal resolutions of approximately 100–300 m are a minimal requirement for accurately representing intra-urban $CO_2$ variability, whereas coarser ( 1 km) grids systematically underestimate spatial contrasts and peak intensities, not because of intrinsic model deficiencies, but due to scale-dependent aggregation of emissions and transport fields, an effect that must be explicitly addressed in future urban inversion systems.

Looking ahead, a hybrid inversion framework could mitigate the resolution-dependent challenges identified in this study. Running inversions at high spatial resolution (100–300 m) using detailed emission fields, while constraining transport with interpolated meteorological outputs from the more robust 900 m simulations, would preserve fine-scale emission realism without compromising wind accuracy. Such an approach would reduce aggregation errors while maintaining computational feasibility. Future work should focus on integrating these findings into urban inversion systems to enhance the accuracy and robustness of emission estimates. In particular, further research is needed to quantify how scale effects and aggregation errors propagate into flux inversions, and to determine whether sub-kilometer modeling provides measurable improvements in inferred emissions. The effective assimilation of high-resolution LIDAR observations will also be critical for constraining boundary layer dynamics and improving transport representation. Overall, hybrid approaches that couple mesoscale and LES configurations offer a practical compromise between computational cost, physical realism, and inversion accuracy. The Paris case study provides a transferable framework for future applications, demonstrating how multi-scale modeling and dense observations can be combined to design robust, next generation urban $CO_2$ monitoring and verification systems.

*Data availability.* The wind lidar data used in this study are available on Zenodo at https://zenodo.org/records/17710093. Surface meteorological observations were obtained from the ACTRIS-FR and Météo-France networks and are accessible through the AERIS data portal at https://www.aeris-data.fr. The Copernicus Global Digital Elevation Model is available at https://doi.org/10.5069/G9028PQB. The CAMS Greenhouse Gas Reanalysis dataset can be accessed through the Copernicus Atmosphere Data Store at https://doi.org/10.24380/8fck-9w87. The hybrid 100-m global land cover dataset with Local Climate Zones used for WRF model configuration is available on Zenodo at https://zenodo.org/records/7670653. Paris TCCON column $CO_2$ data are available from the CaltechDATA repository at https://data.caltech.edu/records/6cj5y-spd74. Near-surface $CO_2$ station data were obtained from the ICOS Carbon Portal and can be accessed at https://data.icos-cp.eu/portal/. The planetary boundary layer height (PBLH) data from the Qualair platform are accessible through the IPSL Compute and Data Centre server at https://thredds-x.ipsl.fr/thredds/catalog/catalog.html. The Airparif emission inventories and the PBLH data from the Urbisphere network are confidential and are not publicly available. Readers interested in accessing these datasets should contact the respective data providers directly.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


*Author contributions.* AR and TL developed the idea of the paper. TL provided overall supervision and project leadership. AR performed the experiments, analyzed the results and wrote the manuscript. CA contributed to the modeling framework and its implementation. MC, MR, JD, and OL provided $CO_2$ observational data. AR, MK, and LJ participated in regular project meetings and discussions. AC and DL provided PBLH, Doppler Wind LIDAR data and flux data and also participated in regular project meetings and discussions. LD and OS provided $CO_2$ Airparif inventory data. LB and BL provided additional flux data. All authors contributed through critical review of the manuscript.

*Competing interests.* The authors declare that they have no competing interests.

*Acknowledgements.* The CATRINE project (grant agreement 101135000) is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them. We acknowledge collaboration with the Atmosphere and Climate Competence Center (ACCC), funded by the Research Council of Finland (decision no. 359340). This work was supported by the Pilot Applications in Urban Landscapes (PAUL) - Towards integrated city observatories for greenhouse gases (ICOS Cities), funded by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 101037319. T. Lauvaux was supported by the French Ministry of Research and Education (MESRI) through the Chaire de Professeur Junior (CASAL project). The authors acknowledge the QUALAIR team and infrastructure for their scientific support and especially Cristelle Cailteau-fischbach and Camille Viatte for letting access to the JUS site. QUALAIR is an observation platform of the OSU Ecce Terra of Sorbonne University operated by LATMOS with the support of IPSL. We acknowledge Jeremie Depuydt, Pedro Coimbra, Alain Fortineau and Anaïs Feron for helping getting the data of the JUS site. We also acknowledge the ERC Urbisphere project (Grant Agreement No 855005) for providing access to data and fruitful collaboration.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


## References

ACTRIS-FR and Météo-France: Surface meteorological observations from the French automatic station network, 6-minute temporal resolution, https://www.aeris-data.fr, accessed via AERIS;, 2024.

Agency, E. S.: Copernicus Global Digital Elevation Model, https://doi.org/10.5069/G9028PQB, 2024.

Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., Bousserez, N., Boussetta, S., Buchwitz, M., Cantarello, L., Crevoisier, C., Engelen, R., Eskes, H., Flemming, J., Garrigues, S., Hasekamp, O., Huijnen, V., Jones, L., Kipling, Z., Langerock, B., McNorton, J., Meilhac, N., Noël, S., Parrington, M., Peuch, V.-H., Ramonet, M., Razinger, M., Reuter, M., Ribas, R., Suttie, M., Sweeney, C., Tarniewicz, J., and Wu, L.: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020, Atmospheric Chemistry and Physics, 23, 3829–3859, https://doi.org/10.5194/acp-23-3829-2023, 2023.

Brunner, D., Kuhlmann, G., Henne, S., Koene, E., Kern, B., Wolff, S., Voigt, C., Jöckel, P., Kiemle, C., Roiger, A., Fiehn, A., Krautwurst, S., Gerilowski, K., Bovensmann, H., Borchardt, J., Galkowski, M., Gerbig, C., Marshall, J., Klonecki, A., Prunet, P., Hanfland, R., Pattantyús-Ábrahám, M., Wyszogrodzki, A., and Fix, A.: Evaluation of simulated $CO_2$ power plant plumes from six high-resolution atmospheric transport models, Atmospheric Chemistry and Physics, 23, 2699–2728, https://doi.org/10.5194/acp-23-2699-2023, 2023.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Li, L., and Tarko, A.: Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual, https://doi.org/10.5281/zenodo.3938963, 2020.

Che, K., Lauvaux, T., Taquet, N., Stremme, W., Xu, Y., Alberti, C., Lopez, M., García-reynoso, A., Ciais, P., Liu, Y., Ramonet, M., and Grutter, M.: $CO_2$ Emissions Estimate From Mexico City Using Ground- and Space-Based Remote Sensing, Journal of Geophysical Research: Atmospheres, 129, https://doi.org/10.1029/2024jd041297, 2024.

Chen, F. and Dudhia, J.: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity, MWR, 129, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2, 2001.

Churkina, G.: The Role of Urbanization in the Global Carbon Cycle, Frontiers in Ecology and Evolution, Volume 3 - 2015, https://doi.org/10.3389/fevo.2015.00144, 2016.

Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hutfilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, Environmental Research Letters, 16, 074 033, https://doi.org/10.1088/1748-9326/ac00e2, 2021.

Delaria, E. R., Kim, J., Fitzmaurice, H. L., Newman, C., Wooldridge, P. J., Worthington, K., and Cohen, R. C.: The Berkeley Environmental Air-quality and $CO_2$ Network: field calibrations of sensor temperature dependence and assessment of network scale $CO_2$ accuracy, Atmospheric Measurement Techniques, 14, 5487–5500, https://doi.org/10.5194/amt-14-5487-2021, 2021.

Demuzere, M., He, C., Martilli, A., and Zonato, A.: A hybrid 100-m global land cover dataset with Local Climate Zones for WRF, https://doi.org/10.5281/zenodo.7670653, 2023.

Doc, J., Ramonet, M., Bréon, F.-M., Combaz, D., Chariot, M., Lopez, M., Delmotte, M., Cailteau-Fischbach, C., Nief, G., Laporte, N., Lauvaux, T., and Ciais, P.: The monitoring network of greenhouse gas ($CO_2$, $CH_4$) in the Paris region, EGUsphere, 2024, 1–32, https://doi.org/10.5194/egusphere-2024-2826, 2024.

Ebert, E. E.: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework, Meteorological Applications, 15, 51–64, https://doi.org/10.1002/met.25, 2008.

Eric, V.: MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V061, https://doi.org/10.5067/MODIS/MOD09A1.061, 2021.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026



Fenner, D., Christen, A., Grimmond, S., Meier, F., Morrison, W., Zeeman, M., Barlow, J., Birkmann, J., Blunn, L., Chrysoulakis, N., Clements, M., Glazer, R., Hertwig, D., Kotthaus, S., König, K., Looschelders, D., Mitraka, Z., Poursanidis, D., Tsirantonakis, D., Bechtel, B., Benjamin, K., Beyrich, F., Briegel, F., Feigel, G., Gertsen, C., Iqbal, N., Kittner, J., Lean, H., Liu, Y., Luo, Z., McGrory, M., Metzger, S., Paskin, M., Ravan, M., Ruhtz, T., Saunders, B., Scherer, D., Smith, S. T., Stretton, M., Trachte, K., and Hove, M. V.: urbisphere-Berlin Campaign: Investigating Multiscale Urban Impacts on the Atmospheric Boundary Layer, Bulletin of the American Meteorological Society, 105, E1929 – E1961, https://doi.org/10.1175/BAMS-D-23-0030.1, 2024.

Gaudet, B. J., Lauvaux, T., Deng, A., and Davis, K. J.: Exploration of the impact of nearby sources on urban atmospheric inversions using large eddy simulation, Elementa: Science of the Anthropocene, 5, 60, https://doi.org/10.1525/elementa.247, 2017.

Glauch, T., Marshall, J., Gerbig, C., Botía, S., Gałkowski, M., Vardag, S. N., and Butz, A.: *pyVPRM*: a next-generation vegetation photosynthesis and respiration model for the post-MODIS era, Geoscientific Model Development, 18, 4713–4742, https://doi.org/10.5194/gmd-18-4713-2025, 2025.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled online chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation, Monthly Weather Review, 132, 103–120, https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2, 2004.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, JGR: Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Janjic, Z. I.: The Step-Mountain Coordinate: Physical Package, MWR, 118, 1429–1443, https://doi.org/10.1175/1520-0493(1990)118<1429:TSMCPP>2.0.CO;2, 1990.

Janjic, Z. I.: The surface layer in the NCEP eta model in the eleventh conference on numerical weather prediction, Norfolk, VA, 19–23 August 1996, Amer. Meteor. Soc., Boston, MA, pp. 354–356, 1996.

Kiel, M., Eldering, A., Roten, D. D., Lin, J. C., Feng, S., Lei, R., Lauvaux, T., Oda, T., Roehl, C. M., Blavier, J.-F., and Iraci, L. T.: Urban-focused satellite $CO_2$ observations from the Orbiting Carbon Observatory-3: A first look at the Los Angeles megacity, Remote Sensing of Environment, 258, 112 314, https://doi.org/10.1016/j.rse.2021.112314, 2021.

Kotthaus, S., Haeffelin, M., Drouin, M.-A., Dupont, J.-C., Grimmond, S., Haefele, A., Hervo, M., Poltera, Y., and Wiegner, M.: Tailored Algorithms for the Detection of the Atmospheric Boundary Layer Height from Common Automatic Lidars and Ceilometers (ALC), Remote Sensing, 12, https://doi.org/10.3390/rs12193259, 2020.

Kotthaus, S., Bravo-Aranda, J. A., Collaud Coen, M., Guerrero-Rascado, J. L., Costa, M. J., Cimini, D., O'Connor, E. J., Hervo, M., Alados-Arboledas, L., Jiménez-Portaz, M., Mona, L., Ruffieux, D., Illingworth, A., and Haeffelin, M.: Atmospheric boundary layer

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


height from ground-based remote sensing: a review of capabilities and limitations, Atmospheric Measurement Techniques, 16, 433–479, https://doi.org/10.5194/amt-16-433-2023, 2023.

Laughner, J. L., Toon, G. C., Mendonca, J., Petri, C., Roche, S., Wunch, D., Blavier, J.-F., Griffith, D. W. T., Heikkinen, P., Keeling, R. F., Kiel, M., Kivi, R., Roehl, C. M., Stephens, B. B., Baier, B. C., Chen, H., Choi, Y., Deutscher, N. M., DiGangi, J. P., Gross, J., Herkommer, B., Jeseck, P., Laemmel, T., Lan, X., McGee, E., McKain, K., Miller, J., Morino, I., Notholt, J., Ohyama, H., Pollard, D. F., Rettinger, M., Riris, H., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Wofsy, S. C., Zhou, M., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2020 data version, Earth System Science Data, 16, 2197–2260, https://doi.org/10.5194/essd-16-2197-2024, 2024.

Lauvaux, T., Schuh, A., Uliasz, M., Richardson, S., Miles, N., Andrews, A., Sweeney, C., Diaz-Isaac, L., Martins, D., Shepson, P., and Davis, K.: Constraining the CO2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmospheric Chemistry and Physics, 12, 337–354, https://doi.org/10.5194/acp-12-337-2012, 2012.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., Song, Y., Karion, A., Oda, T., Patarasuk, R., Razlivanov, I., Sarmiento, D., Shepson, P., Sweeney, C., Turnbull, J., and Wu, K.: High-resolution atmospheric inversion of urban $CO_2$ emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Journal of Geophysical Research: Atmospheres, 121, 5213–5236, https://doi.org/10.1002/2015JD024473, 2016.

Lian, J., Lauvaux, T., Utard, H., Bréon, F.-M., Broquet, G., Ramonet, M., Laurent, O., Albarus, I., Chariot, M., Kotthaus, S., Haeffelin, M., Sanchez, O., Perrussel, O., Denier van der Gon, H. A., Dellaert, S. N. C., and Ciais, P.: Can we use atmospheric $CO_2$ measurements to verify emission trends reported by cities? Lessons from a 6-year atmospheric inversion over Paris, Atmospheric Chemistry and Physics, 23, 8823–8835, https://doi.org/10.5194/acp-23-8823-2023, 2023.

Lian, J., Laurent, O., Chariot, M., Lienhardt, L., Ramonet, M., Utard, H., Lauvaux, T., Bréon, F.-M., Broquet, G., Cucchi, K., Millair, L., and Ciais, P.: Development and deployment of a mid-cost $CO_2$ sensor monitoring network to support atmospheric inverse modeling for quantifying urban $CO_2$ emissions in Paris, EGUsphere, 2024, 1–26, https://doi.org/10.5194/egusphere-2024-125, 2024.

Looschelders, D. and et al.: [Title to be updated], in preparation, 2026.

Looschelders, D., Christen, A., Grimmond, S., Kotthaus, S., Fenner, D., Dupont, J.-C., Haeffelin, M., and Morrison, W.: Inter-Instrument Variability of Vaisala CL61 Lidar-Ceilometer's Attenuated Backscatter, Cloud Properties and Mixed-Layer Height, Meteorological Applications, 32, e70 088, https://doi.org/10.1002/met.70088, e70088 MET-25-0027.R2, 2025.

Lwasa, S., Seto, K., X., B., Blanco, H., Gurney, K., Kilkis, S., Lucon, O., Murakami, J., Pan, J., Sharifi, A., and Yamagata, Y.: 2022: Urban systems and other settlements, In IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, https://doi.org/10.1017/9781009157926.010, 2022.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochemical Cycles, 22, https://doi.org/10.1029/2006GB002735, 2008.

Mitchell, L. E., Lin, J. C., Bowling, D. R., Pataki, D. E., Strong, C., Schauer, A. J., Bares, R., Bush, S. E., Stephens, B. B., Mendoza, D., Mallia, D., Holland, L., Gurney, K. R., and Ehleringer, J. R.: Long-term urban carbon dioxide observations reveal spatial and temporal dynamics related to urban characteristics and growth, Proceedings of the National Academy of Sciences of the United States of America, 115, 2912–2917, https://doi.org/10.1073/pnas.1702393115, 2018.

Morrison, W., Looschelders, D., Céspedes, J., Claxton, B., Drouin, M.-A., Dupont, J.-C., Faucheux, A., Haeffelin, M., Holst, C. C., Kotthaus, S., Masson, V., McGregor, J., Price, J., Zeeman, M., Grimmond, S., and Christen, A.: Harmonised boundary layer wind profile dataset from six

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


ground-based Doppler wind lidars in a transect across Paris, France, Earth System Science Data, 17, 6507–6529, https://doi.org/10.5194/essd-17-6507-2025, 2025.

Muñoz Esparza, D., Lundquist, J. K., Sauer, J. A., Kosović, B., and Linn, R. R.: Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign: From weather to boundary-layer eddies, Journal of Advances in Modeling Earth Systems, 9, 1572–1594, https://doi.org/10.1002/2017MS000960, 2017.

Nurmi, P.: Recommendations on the Verification of Local Weather Forecasts, ECMWF Technical Memorandum 430, European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom, https://doi.org/10.21957/y1z1thg5l, 2003.

Peckham, S. E., Grell, G. A., McKeen, S. A., Ahmadov, R., Wong, K. Y., Barth, M., Pfister, G., Wiedinmyer, C., Fast, J. D., Gustafson, W. I., Ghan, S. J., Zaveri, R., Easter, R. C., Barnard, J., Chapman, E., Hewson, M., Schmitz, R., Salzmann, M., and Beck, V. Freitas, S. R.: WRF-Chem version 3.8.1 user's guide, Earth System Research Laboratory (U.S.), Global Systems Division, https://doi.org/10.7289/V5/TM-OAR-GSD-48, 2017.

Rafalimanana, A., Lauvaux, T., Abdallah, C., Chariot, M., Ramonet, M., Doc, J., Laurent, O., Lopez, M., Raznjevic, A., Krol, M., Järvi, L., Christen, A., Looschelders, D., David, L., Sanchez, O., Bignotti, L., Loubet, B., Grimmond, S., and Morrison, W.: Urban atmospheric CO2 plumes from space - Part 1: Atmospheric modeling of the urban boundary layer, manuscript submitted to Atmospheric Chemistry and Physics, 2026.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002299, 2003.

Salamanca, F. and Martilli, A.: A new Building Energy Model coupled with an Urban Canopy Parameterization for urban climate simulations—part II. Validation with one dimension off-line simulations, Theoretical and Applied Climatology, 99, 345–356, https://doi.org/10.1007/s00704-009-0143-8, 2010.

Singh, J., Singh, N., Ojha, N., Sharma, A., Pozzer, A., Kiran Kumar, N., Rajeev, K., Gunthe, S. S., and Kotamarthi, V. R.: Effects of spatial resolution on WRF v3.8.1 simulated meteorology over the central Himalaya, Geoscientific Model Development, 14, 1427–1443, https://doi.org/10.5194/gmd-14-1427-2021, 2021.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. ., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Version 4, NCAR Tech. Note NCAR/TN-556+STR, p. 145pp., https://doi.org/10.5065/1dfh-6p97, 2019.

Staufer, J., Broquet, G., Bréon, F.-M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel $CO_2$ emissions based on atmospheric inversion, Atmospheric Chemistry and Physics, 16, 14 703–14 726, https://doi.org/10.5194/acp-16-14703-2016, 2016.

Stauffer, D. R. and Seaman, N. L.: Multiscale Four-Dimensional Data Assimilation, Journal of Applied Meteorology and Climatology, 33, 416 – 434, https://doi.org/10.1175/1520-0450(1994)033<0416:MFDDA>2.0.CO;2, 1994.

Stull, B. R.: An Introduction to Boundary Layer Meteorology., Journal of Fluid Mechanics, 224, 660–661, https://doi.org/10.1017/S0022112091211921, 1988.

Super, I., Dellaert, S. N. C., Visschedijk, A. J. H., and Denier van der Gon, H. A. C.: Uncertainty analysis of a European high-resolution emission inventory of $CO_2$ and CO to support inverse modelling and network design, Atmospheric Chemistry and Physics, 20, 1795–1816, https://doi.org/10.5194/acp-20-1795-2020, 2020.

Talbot, C., Bou-Zeid, E., and Smith, J.: Nested Mesoscale Large-Eddy Simulations with WRF: Performance in Real Test Cases, Journal of Hydrometeorology, 13, 1421 – 1441, https://doi.org/10.1175/JHM-D-11-048.1, 2012.

Tao, Z., Chin, M., Gao, M., Kucsera, T., Kim, D., Bian, H., Kurokawa, J., Wang, Y., Liu, Z., Carmichael, G. R., Wang, Z., and Akimoto, H.: Evaluation of NU-WRF model performance on air quality simulation under various model resolutions – an investigation within the framework of MICS-Asia Phase III, Atmospheric Chemistry and Physics, 20, 2319–2339, https://doi.org/10.5194/acp-20-2319-2020, 2020.

Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg, R., Fisher, B., Crisp, D., Basilio, R., Bennett, M., Cervantes, D., Chang, A., Dang, L., Frankenberg, C., Haemmerle, V. R., Keller, G. R., Kurosu, T., Laughner, J. L., Lee, R., Marchetti, Y., Nelson, R. R., O'Dell, C. W., Osterman, G., Pavlick, R., Roehl, C., Schneider, R., Spiers, G., To, C., Wells, C., Wennberg, P. O., Yelamanchili, A., and Yu, S.: OCO-3 early mission operations and initial (vEarly) XCO2 and SIF retrievals, Remote Sensing of Environment, 251, 112 032, https://doi.org/10.1016/j.rse.2020.112032, 2020.

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, MWR, 136, 5095–5115, https://doi.org/10.1175/2008MWR2387.1, 2008.

Tofallis, C.: A better measure of relative prediction accuracy for model selection and model estimation, The Journal of the Operational Research Society, 66, 1352–1362, http://www.jstor.org/stable/24505756, 2015.

Törnqvist, L., Vartia, P., and Vartia, Y.: How Should Relative Changes be Measured, The American Statistician, 39, 43–46, https://api.semanticscholar.org/CorpusID:121490598, 1985.

Turnbull, J. C., Karion, A., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Sweeney, C., McKain, K., Lehman, S. J., Gurney, K. R., Patarasuk, R., Liang, J., Shepson, P. B., Heimburger, A., Harvey, R., and Whetstone, J.: Synthesis of Urban CO2 Emission Estimates from Multiple Methods from the Indianapolis Flux Project (INFLUX), Environmental Science & Technology, 53, 287–295, https://doi.org/10.1021/acs.est.8b05552, 2019.

Té, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2020.R0, https://doi.org/10.14291/tccon.ggg2020.paris01.R0, 2022.

Udina, M., Montornès, A., Casso, P., Kosović, B., and Bech, J.: WRF-LES Simulation of the Boundary Layer Turbulent Processes during the BLLAST Campaign, Atmosphere, 11, 1149, https://doi.org/10.3390/atmos11111149, 2020.

Vogelezang, D. H. P. and Holtslag, A. A. M.: Evaluation and model impacts of alternative boundary-layer height formulations, Boundary-Layer Meteorology, 81, 245–269, https://doi.org/10.1007/BF02430331, 1996.

Washenfelder, R. A., Toon, G. C., Blavier, J.-F., Yang, Z., Allen, N. T., Wennberg, P. O., Vay, S. A., Matross, D. M., and Daube, B. C.: Carbon dioxide column abundances at the Wisconsin Tall Tower site, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2006JD007154, 2006.

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of $CO_2$ total columns from space, Atmospheric Chemistry and Physics, 11, 12 317–12 337, https://doi.org/10.5194/acp-11-12317-2011, 2011.

Wyngaard, J. C.: Toward Numerical Modeling in the "Terra Incognita", Journal of the Atmospheric Sciences, 61, 1816 – 1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, 2004.

https://doi.org/10.5194/egusphere-2026-2581
Preprint. Discussion started: 18 August 2026


Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., Yang, E. G., and Wu, D.: Constraining Fossil Fuel CO2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO2, Journal of Geophysical Research: Atmospheres, 125, e2019JD030 528, https://doi.org/10.1029/2019JD030528, e2019JD030528 2019JD030528, 2020.

Ye, X., Li, W., Lauvaux, T., Lin, S., Zhang, Z., Lin, Y., Hua, J., You, Q., and Lin, J.: Constraining anthropogenic CO2 emissions using dense observations from OCO-3: Disentangling observed imprints over Xiamen-Zhangzhou-Quanzhou metropolitan area and challenges, Urban Climate, 63, 102 591, https://doi.org/10.1016/j.uclim.2025.102591, 2025.