



HAL
open science

Improvement of downscaled ozone concentrations from the transnational scale to the kilometric scale: Need, interest and new insights

Honorine Gauthier-Manuel, Frédéric Mauny, Mathieu Boilleaut, Marie Ristori, Sophie Pujol, Florent Vasbien, Anne-Laure Parmentier, Nadine Bernard

► To cite this version:

Honorine Gauthier-Manuel, Frédéric Mauny, Mathieu Boilleaut, Marie Ristori, Sophie Pujol, et al.. Improvement of downscaled ozone concentrations from the transnational scale to the kilometric scale: Need, interest and new insights. *Environmental Research*, 2022, 210, pp.112947. 10.1016/j.envres.2022.112947 . hal-03723752

HAL Id: hal-03723752

<https://hal.science/hal-03723752v1>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Improvement of downscaled ozone concentrations from the transnational scale to the kilometric scale: need, interest and new insights

Honorine GAUTHIER-MANUEL^{a,b,}, Frédéric MAUNY^{a,b}, Mathieu BOILLEAUT^c, Marie RISTORI^c, Sophie PUJOL^{a,b}, Florent VASBIEN^d, Anne-Laure PARMENTIER^{a,b}, Nadine BERNARD^{a,e}*

a: UMR 6249, Laboratoire Chrono-environnement, Université de Bourgogne Franche-Comté, CNRS, 25000 Besançon, France

b: Unité de méthodologie en recherche clinique, épidémiologie et santé publique (uMETH), Inserm CIC 1431, CHU, 25030 Besançon cedex, France

c: ATMO Bourgogne-Franche-Comté, 25000 Besançon, France

d: ATMO Grand Est, 67300 Schiltigheim, France

e: Centre National de la Recherche Scientifique, UMR 6049, Laboratoire ThéMA, Université de Bourgogne Franche-Comté, 25000 Besançon, France

*Corresponding author.

E-mail address: honorine.gauthier-manuel@univ-fcomte.fr (H. Gauthier-Manuel)

Abstract:

Background: Ground-level ozone is a major public health issue worldwide. An accurate assessment of ozone exposure is necessary. Modeling tools have been developed to tackle this issue in large areas. However, these models could present inaccuracies at the local scale.

Objectives: The objective of this study was i) to assess whether O₃ concentrations estimated by transnational modeling at the kilometric scale (9 km²) could be improved, ii) to propose a potential correction of these downscaled ozone concentrations and iii) to evaluate the efficiency and applicability of such a correction.

Method: The present work was carried out in three phases. First, the performance of a transnational modeling platform (PREV'EST) was assessed at 6 geographic points by comparison with data from 6 air quality monitoring stations. Performance indicators were used for this purpose (MBE (mean bias error), MAE (mean absolute error), RMSE (root mean square error), r (Pearson correlation coefficient), and target plots). Second, several corrections were developed using MARS (multivariate adaptive regression splines) and integrating different sets of variables (mean temperature, relative humidity, rainfall amount, wind speed, elevation, and date). Their performance was evaluated. Third, external validation of the corrections was conducted using the data from six additional air quality monitoring stations.

Results: The uncorrected PREV'EST model presented a lack of exactitude and precision. These concentrations did not reproduce the interday variability of the measurements, leading to a lack of temporal contrast in exposure data. For the best performance enhancement, the correction applied improved MBE, MAE, RMSE and r from 14.67, 19.23, 23.18 and 0.67 to 0.00, 8.00, 10.19 and 0.91, respectively. External validation confirmed the efficiency of the corrections at the regional scale.

Conclusions: We propose a validated and efficient methodology integrating local environmental variables. The methodology is adaptable according to the context, needs and data available.

Keywords: Ozone, modeling, improvement, performance, MARS model

Study Funding

This work was supported by the agglomeration community of Grand Besançon Métropole (PhD scholarship, 2019) and the “Pôle fédératif de recherche et de formation en santé publique Bourgogne-Franche-Comté” (2019 and 2020 research project calls).

1. Introduction:

Outdoor air pollution is a major global public health issue involving both primary and secondary pollutants (Lelieveld., 2015; Cohen et al., 2017). Ground-level ozone (O_3) is a secondary air pollutant which causes adverse human health effects (Brunekreef and Holgate, 2002; Zhang et al., 2019) that can range from eye and nose irritation to pulmonary and cardiovascular system disorders (Henrotin et al., 2010; Marx et al., 2019; Nuvolone et al., 2018; Zhang et al., 2019). Therefore, it is necessary to assess its spread in the environment and human exposure to this pollutant. Tropospheric ozone is particularly challenging to characterize in expology because it highly fluctuates in space and time. From 2000 to 2017, ozone rose in urbanized areas (with increasing annual ozone-related premature deaths) and decreased in rural areas in the EU-28 (Sicard et al., 2021). However, rural areas still remain more exposed. In response to climate change, ozone exposure is prone to be exacerbated in polluted regions (Lu et al., 2019). Indeed, this photooxidant is formed by photochemical reactions initiated in the presence of solar radiation and primary pollutant precursors (NO_x , volatile organic compounds (VOCs)) (Royal Society (Great Britain), 2008; Wang et al., 2017). Its distribution depends on numerous environmental parameters, such as solar radiation (and, indirectly, temperature), wind speed, relative humidity, and elevation (Ramos et al., 2018; Royal Society, 2008). It also depends on ozone formation regime, strongly influenced by the VOCs-to- NO_x ratio (Li et al., 2021). Ozone concentration evaluation is complex due to its extreme variability in space and time (diurnal cycle and seasonality) (Castell-Balaguer et al., 2012).

Networks of air quality monitoring stations provide an assessment of population exposure and monitor regulatory threshold exceedances. These data are nevertheless insufficient because they can offer limited spatial and temporal coverage (Chen et al., 2014; Garcia et al., 2010). Co-kriging approaches can be applied to interpolate background stations data (Sicard et al., 2013) and provide a spatially continuous estimation of ozone levels. Approaches combining local regression and residuals kriging can reach high spatial resolution (250 x 250m) (Sicard et al., 2016). However, these approaches are tributary to the number and spatial distribution of air quality monitoring stations (Wang et al., 2016).

Land-Use Regression (LUR) and chemical transport models (CTMs) are the most commonly used modeling approaches in the long-term assessment of population exposure (Joly et al., 2021; Wang et al., 2016). CTMs such as CHIMERE (Mailler et al., 2017; Menut et al., 2013), CMAQ or MOCAGE, provide grid meshes that commonly range from 4 to 12km

(U.S. EPA, 2018). However, the grid cell size is proportional to the spatial coverage of the model. CTMs covering Europe generally handle resolutions from 12 to 25 km, those covering a single country usually use 4 to 10 km resolutions and more rarely can reach ~1 km on some complex terrains regions (Schaap et al., 2015). CHIMERE is one of the most popular CTM (Thunis et al., 2016) and exploits the outputs of a meteorological model (Weather Research and Forecasting (WRF) model, for example) by modeling platforms delivering European, national or regional air quality forecasts (Honore et al., 2005; Riviere et al., 2019; Rouil et al., 2009). PREV'AIR is a modeling platform providing air quality forecasts at European and national scales (France) and integrates WRF and CHIMERE models. The data stemming from these forecasting platforms are used in exposure risk assessment (Riviere et al., 2019). However, regional air quality modeling tools can present biases and inaccuracies (Steven Porter et al., 2015). Improvements could be needed prior to using these data for: i) exposure assessment (Crooks and Özkaynak, 2014) and ii) better depict both spatial and temporal variability of ozone.

Modeling platforms, as PREV'AIR, require large computational capacity and extensive input data (Braun et al., 2017; Wang et al., 2016), but they constantly generate predictions to prevent and assess population exposure. We propose a correction of the outputs of a CTM predicted at the finest spatial scale, already available, to develop a tool which is more accessible, faster-processing, with lower computational cost than *ex nihilo* implementing LUR or CTM, and with better performance.

The objective of this study was i) to assess whether downscaled O₃ concentrations estimated by transnational modeling (initial resolution 45 km x 45 km, i.e., 2025 km²) to the kilometric scale (9 km²) should be improved, ii) to propose a potential correction of these downscaled ozone concentrations and iii) to evaluate the efficiency and applicability of such a correction.

2. Materials and methods:

The present work was conducted in three phases. First, the performance of the transnational model was assessed by comparison with air quality monitoring data at 6 geographic points. Second, several corrections were added to the model, and their performance were re-evaluated. Third, external validation of the corrections was conducted using the data from six additional air quality monitoring stations.

2.1 Assessment of the performance of the uncorrected PREV'EST model

2.1.1. Study site and data sources

The mesoscale meteorological model, WRF, is already used in numerous air quality modeling approaches (Thunis et al., 2016). The CHIMERE model, an Eulerian chemistry-transport model, integrated WRF outputs, emission inventories and chemical boundary conditions to forecast air quality (Mailler et al., 2017; Menut et al., 2013). The PREV'EST modeling platform used in this study is a regional/transnational platform developed and managed by ATMO Grand Est (regional air quality monitoring association). Its functioning is analogous to PREV'AIR, a modeling platform used at the national and European scales (Honore et al., 2005; Riviere et al., 2019; Rouil et al., 2009), and integrates the WRF and CHIMERE models. PREV'EST integrates Europe PREVAIR's forecasts into the forecasting mode as boundary conditions to improve the modeling by limiting the edge effect. The PREV'EST and PREV'AIR modeling platforms are both subjected to daily performance assessments, comparing modeled concentrations with measurements (bias, percentage of errors lower than a certain level and correlation) (Honore et al., 2005; Rouil et al., 2009). ATMO Grand Est also applies corrections on yearly aggregated data.

The PREV'EST modeling platform generated outputs on three grid meshes of different sizes: 45 km x 45 km (over Europe), 15 km x 15 km (over France and adjacent countries), and 3 km x 3 km (eastern France and the boundary area of France, Germany, Switzerland and Luxembourg). The finest grid mesh used in this study covered 274,428 km², or a 194 x 154 pixel grid of 9 km². The modeling and nesting processes were conducted by ATMO Grand Est, who delivered the outputs of the finest grid mesh (3 km x 3 km). We conducted the optimization processes (described below) on this finest grid mesh. CHIMERE's parametrization was adapted to this area thanks to the expertise and field knowledge of the team who developed and managed this platform (ATMO Grand Est). Each pixel represented a basic spatial unit associated with a daily average modeled ozone concentration (called [O₃] modelled).

[Figure 1]

Figure 1: PREV'EST modeling area (for the 3 km x 3 km grid mesh, on the top map) and location of the air quality monitoring stations (ATMO BFC) and weather stations (Météo-France).

Six air quality monitoring stations (three urban, two suburban and one rural station) measured daily average O₃ concentrations (called [O₃]_{measured}) in the Franche-Comté region (France) from 2008 to 2019 (Figure 1). Each monitoring station was related to the pixel of the PREV'EST model on which it is located. These daily measured concentrations were considered the gold standard for the O₃ concentrations at the local place of the monitoring station (Thunis et al., 2016). ATMO Bourgogne-Franche-Comté (ATMO BFC) provided the measured data.

For each of the six stations, the daily discrepancy between modeled and measured concentrations was quantified by the differences between daily average O₃ concentrations:

$$\text{Daily Discrepancy} = [\text{O}_3]_{\text{modeled}} - [\text{O}_3]_{\text{measured}} \quad (1)$$

2.1.2. Performance indicators

The following performance indicators were retained and are described in Appendix 1: mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE) and Pearson correlation coefficient (r). MBE quantifies the mean model-measure discrepancy and a potential overestimation (MBE>0) or underestimation (MBE<0) of the model (Kato, 2016; Martin, 2011). MAE is a precision indicator measuring the mean absolute model-measure discrepancy. MAE is higher when the model-measure discrepancies are large and scattered. RMSE is also a precision indicator (Abdulelah Al-Sudani et al., 2019; Lesmeister, 2019; Sekulic and Kowalski, 1992; Yilmaz et al., 2018) that gives more weight than MAE to high model-measure discrepancies (Bouthevillain and Mathis, 1995). RMSE is higher for large discrepancies (regardless of the sign, positive or negative) with high variability. The Pearson correlation coefficient (r) assesses the linear correlation between modeled and measured concentrations.

In addition to these average performance indicators, target plots were used to represent the performance of the model at the station scale. The localization of the station in the circle inside the target plot indicates respect to these quality criteria: the vertical axis displays the normalized bias (BIAS_{Norm}), while the horizontal axis displays the CRMSE_{Norm} (normalized centered RMSE) value, both normalized by the measurement uncertainty associated with the pollutant. An extended description of both the quality criterion and target plots is provided in Appendix 1. Target plots were built using R software ('dartle' package).

2.2. Development and performance assessment of the corrections

The statistical associations between model-measure discrepancies and environmental variables (meteorological variables and elevation) were analyzed using MARS (multivariate adaptive regression splines) models (Friedman, 1991; Friedman and Roosen, 1995).

2.2.1. *Statistical relationship between the discrepancy and environmental variables*

Daily meteorological data were collected from 2008 to 2019 and provided by the French meteorological institute (Météo-France). The meteorological parameters were mean temperature (Celsius degrees (°C)), rainfall amount (mm), mean relative humidity (expressed in percentage (%)), average wind speed at a height of 10 meters ($\text{m}\cdot\text{s}^{-1}$), live maximum wind speed ($\text{m}\cdot\text{s}^{-1}$), and maximum wind speed measured over 10 minutes ($\text{m}\cdot\text{s}^{-1}$). The data from the nearest weather station were assigned to each air quality monitoring station, minimizing the difference in elevation between the meteorological station and air quality monitoring stations. The elevation of each air quality monitoring station was recorded by ATMO-BFC.

The relationships between observed discrepancies and environmental variables were analyzed using MARS models (Eq. (2)). MARS models generate a piecewise linear model by finding knots (or “turning points”) in the relationship between Y (the variable we want to predict) and each of the explanatory variables (Figures A.1 and A.2).

These models were used as the prediction function of the daily discrepancy according to environmental variables (Eq. (2)). The predicted daily discrepancy was then used to correct the raw (called “uncorrected”) data. The performance of the corrected O_3 concentrations was assessed by the process described in Section 2.1.2.

$$\text{Daily observed discrepancy} = f(T^\circ, \text{RA}, \text{RH}, \text{WS}, \text{elevation}, \text{date}) \quad (2)$$

where T° is the mean temperature, RA is the rainfall amount, RH is the mean relative humidity, and WS is the wind speed variables.

Different sets of explanatory variables were iteratively used to assess the relevance of integrating some variables less frequently recorded (*e.g.*, RH and WS) (Table 1): four weather factors (4-WF correction, every meteorological variable and the date), two weather factors and elevation (2-WF elevation correction, mean temperature, rainfall amount, elevation and the date) and two weather factors (2-WF correction, mean temperature, rainfall amount and the date).

[Table 1]

The model-measure discrepancy predicted by MARS was then used to calculate corrected modeled O₃ concentrations as follows:

$$\text{Daily } [\text{O}_3]_{\text{corrected}} = \text{daily } [\text{O}_3]_{\text{modelled}} - \text{predicted daily discrepancy} \quad (3)$$

A positive discrepancy reflected an overestimation by the uncorrected PREV'EST model, and conversely, a negative discrepancy corresponded to an underestimation.

The efficiency of each correction was assessed using the performance indicators described in Section 2.1.2. (MBE, MAE, RMSE, r, and target plots) on the daily [O₃]_{corrected} values.

2.2.2. Modeling approaches

Four modeling approaches of increasing complexity were developed (Table 1). The global approach included the data of the six stations during a 12-year period. The annual approach included the 6 stations together but was separately conducted for each of the 12 years (one independent model per year, n = 12). The territorial approach included the data of the 12 years but was separately conducted for each of the six stations (one independent model per station, n = 6). In the territorial and annual approach, one independent model was performed per year and per station (n = 6 x 12) (Table 1). The predictions of all these models were recompiled to calculate the performance indicators and to keep these modeling approaches comparable (Figure A.3). Indeed, although we split the dataset (per year and/or per station), we kept the same number of statistical units (22,640 “discrepancies” used to train the models).

2.3.) External validation of the corrections

External validation was conducted to ensure that this methodology could be applied to a regional territory and not only to the six stations considered to train the model. Six additional air quality monitoring stations (location presented in Figure A.4) were used in this phase. These stations have temporarily measured O₃ concentrations. However, the additional stations were subject to the same quality approach as the six previous stations. The processes described in Sections 2.2.1 and 2.2.2 were applied for the periods when the additional stations were active. Finally, the efficiency of each correction (and for each approach) was assessed using the performance indicators described in Section 2.1.2. The quality of the correction was assessed by the difference between the corrected O₃ concentration and measured O₃ concentration (Eq. (4)).

$$([\text{O}_3]_{\text{modeled}} - \text{predicted daily discrepancy using the function in Eq. (2)}) - [\text{O}_3]_{\text{measured}} \quad (4)$$

The corrected O₃ concentration was defined by:

$$([\text{O}_3]_{\text{modeled}} - \text{predicted discrepancy using the function in Eq. (2)}) \quad (5)$$

All statistical treatments were performed with R software (version 4.0.3), and the MARS models were implemented with the 'earth' package.

3) Results and Discussion

3.1) Evaluation of uncorrected PREV'EST model performance

The discrepancy between the uncorrected kilometric scale modeling and the local measurements was substantial ($+14.67 \mu\text{g m}^{-3}$ (SD = 17.9)). This discrepancy was nonlinearly structured according to some environmental parameters, particularly temperature (Figure A.5). Table 2 presents the average performance indicators reflecting the general performance averaged among all stations. The average performance of the uncorrected PREV'EST model (Table 2) suggested a need for improvement of both exactitude (MBE = $+14.67 \mu\text{g m}^{-3}$) and precision (RMSE = $23.18 \mu\text{g m}^{-3}$, MAE = $19.23 \mu\text{g m}^{-3}$) to be used in exposure assessment at the local scale. According to the literature, the national model PREV'AIR had a very comparable performance to the study PREV'EST model (Meleux and Ung, 2020).

[Table 2]

The performance of the uncorrected PREV'EST model was also separately estimated for the six air quality monitoring stations (Figure 2). The stations are not centered on the horizontal axis (Figure 2). They are also positioned in the left part of the target plot, which indicates a lack of correlation between uncorrected PREV'EST data and measurements (Janssen et al., 2017). The PREV'EST model also overestimated the ozone concentrations, as seen previously (Table 2, MBE = $14.67 \mu\text{g m}^{-3}$). The stations are differently distributed on the vertical axis (bias axis) (Figure 2): the PREV'EST model overestimated the global data but more (Station C) or less (Stations D and E) according to the station.

[Figure 2]

Figure 2: Target plots of the uncorrected PREV'EST modeling.

Different letters represent the different air quality monitoring stations.

According to the current guidelines, the performance of the uncorrected PREV'EST model could be qualified as good. However, this model, already well calibrated, presented faults: a lack of precision and exactitude (a tendency to overestimate differentially according to the station).

3.2) Evaluation of the performance of the corrections

The average performance indicators presented in Table 2 were calculated by compiling the predictions of the models according to the different corrections proposed. The mean bias error (MBE) was drastically reduced regardless of the correction applied (initial value: 14.67; range of MBE after correction: 0.00 to 0.04 $\mu\text{g m}^{-3}$) (Table 2). This indicated an improvement in the exactitude of the modeled ozone concentrations. The MBE values varied very little between the corrections and were thus not discriminant to identifying the best correction.

According to the other indicators (MAE, RMSE, and r), the 4-WF correction could be considered the most efficient, followed by the 2-WF elevation correction and then the 2-WF correction, regardless of the approach considered. The refinement of the type of approach led to a gradual improvement in performance indicators (*i.e.*, decrease of MBE, MAE and RMSE and increase of r). The territorial and annual approach coupled to the 4-WF correction could be seen as the most efficient (MAE = 8.00 $\mu\text{g m}^{-3}$, RMSE = 10.19 $\mu\text{g m}^{-3}$, and $r = 0.91$). Taking into account the specificities of the site of each station provided a more precise correction.

Figure 3 allowed us to explore the potential difference in the performance of the corrections between the six stations. We observed an improvement in the performance of the modeled ozone concentrations (Figure 3) compared with uncorrected PREV'EST (Figure 2). All stations displayed a null bias and were closer to the target center after correction. A differential bias persisted with the 4-WF correction in the global and annual approaches (Figure 3). The differential bias could be damaging when the exposure assessment is simultaneously conducted in different areas. The 2-WF elevation correction allowed for the suppression of this differential bias. The integration of the elevation in the global model with the 2-WF elevation correction offers the opportunity to consider the local characteristics of the sites of each station. Moreover, elevation is well known in the literature to be a key factor in ozone distribution (Brönnimann et al., 2000; Chevalier et al., 2007; Ezcurra et al., 2013). A study conducted by Martin (2010) revealed some lack of precision in a comparable model (a regional application of the PREV' AIR modeling platform, AIRES). Martin (2010) suggested refining the spatial resolution of the model to better consider topography. This lack of local precision could be offset by the integration of local meteorological data and elevation. Moreover, the elevation is easily available over large territories thanks to digital elevation models. Other modeling methods, as LUR or (co)-kriging, integrate the digital elevation models, land use, population density, etc., and provide finer resolution. These methods

present some advantages: the ability to characterize small spatial scale variability, ease to use, and require far fewer computational resources than CTMs (Braun et al., 2017; Wang et al., 2016). However, they use internal data to validate the model through a cross-validation process (Ma et al., 2020) which can overestimate the predictive ability of the model (Wang et al., 2013). LUR focuses on spatial variation (Ren et al., 2020) and can fail to depict the temporal variation of ozone (Ma et al., 2020). CTMs rely on topography, emissions, meteorology, and physicochemical processes. They provide a well captured spatio-temporal ozone variability and a large spatial coverage (Wang et al., 2016). Nevertheless, they present coarser grid mesh and require large computational capacity and extensive input data (Braun et al., 2017), reducing their accessibility.

[FIGURE 3]

Figure 3: Target plots of the uncorrected PREV'EST modeling, global approach 4-WF and 2-WF elevation corrections, annual approach 4-WF and 2-WF elevation corrections, territorial approach 4-WF correction and territorial and annual approach 4-WF correction target plots.

Different letters represent the different air quality monitoring stations.

In terms of application, if we consider a correction dedicated to numerous stations or a regional projection, the 2-WF elevation correction in the annual approach should be favored. This correction suppressed the differential bias between stations and is easily applicable. Some meteorological variables (such as wind speed) were not recorded by most of the weather stations. Our 2WF-elevation correction used common and few variables.

The territorial and territorial and annual approaches had the best performance (Figure 3). From an operational point of view, these approaches could be used to locally correct the exposure data for a few stations and their surrounding areas. The implementation of the territorial and annual approach is very efficient but time-consuming. The territorial approach could be chosen instead of the territorial and annual approach to reach a very satisfactory performance faster. In any case, the territorial (and annual) approach should be applied with caution because the area around a station on which the model is reliable is unknown and hardly definable. External validation will allow the determination of whether these territorial corrections could be extended to other areas while remaining reliable or not.

Figure 4 shows the time series of the concentrations measured, uncorrected and corrected, using the annual 2-WF elevation correction. These corrected concentrations provided better

interday variability than uncorrected PREV'EST data. The latter did not reproduce the interday variability of the measurements and smoothed the curve of exposure assessment data. This led to a lack of contrast in exposure data and could make the identification of a harmful effect of acute O₃ exposure difficult. The day-to-day contrast is essential to identify ozone peaks (Ren et al., 2020) which appear as a key element in the occurrence of health effects (Di et al., 2017). Such a phenomenon has already been observed in air quality modeling (Hogrefe et al., 2004). Hogrefe and his collaborators (2004) suggested that a regional scale resolution was insufficient to well represent ozone variations in the short term (daily, interday, and intraday variations). The methodology proposed in this study offers a potential solution to this lack of precision and variability in the short-term assessment of ozone levels at the regional scale.

Considering the spatial variability of ozone, by definition, a grid cell displays a homogenous exposure level. The assessed level is the same for all individuals living in the same grid cell. This could lead to under or overestimation at the individual scale. However, we could apply the methodology presented in this article to models with finer spatial resolution, to better capture the spatial variability of ozone.

[Figure 4]

Figure 4: Daily mean ozone concentrations measured, modeled by uncorrected PREV'EST modeling and corrected by the annual 2-WF elevation correction, during 2018 at Station A.

3.3. External validation

In this validation, 14,846 statistical units (number of discrepancies modeled) were implied. The corrections of ozone concentrations led to an improvement of every average indicator (Table 3). The initial value of MBE, MAE, RMSE and r were 14.83, 17.63, 21.35 and 0.74, respectively; and their ranges after correction, according to the correction, were 0.04 to 4.87, 9.48 to 12.76, 12.11 to 16.11 and 0.75 to 0.85, respectively (Table 3). All average indicators converged to identify the annual approach with the 4-WF correction as the most efficient. However, the territorial and territorial and annual approaches were only applied to the models of two stations (A and D among the six available stations) because of the locations of these stations (Figure A.4).

[Table 3]

As already observed in the internal validation, the target plots presented in Figure 5 displayed a differential bias among stations in the 4-WF correction, which was reduced by the

2-WF elevation correction. This latter correction could be seen as the most efficient and flexible in this external validation.

This external validation confirmed the main points of the observed performance of the different corrections assessed. If the user desires to apply a correction with fewer spatial constraints, the annual approach will give the best performance. The correction to be favored could be the 4-WF correction because it presented the best average performance (Table 3). However, the 2-WF elevation correction presented a lower differential bias among stations (Figure 5) and very satisfactory average performance (Table 3). This correction was more flexible and required fewer variables that were commonly recorded by most of the weather stations. The 2-WF correction showed an equivalent average performance but lower performance at the station scale.

[FIGURE 5]

Figure 5: External validation: uncorrected PREV'EST modeling, annual approach 4-WF correction, annual approach 2-WF elevation correction, and annual and territorial approach 4-WF and 2-WF correction target plots.

Different letters represent the different air quality monitoring stations.

4 Strengths of the methodology and comparison with existing methodologies

The main strengths of this methodology are its efficiency in improving the precision and exactitude of exposure data, its ability to recover a good contrast in these data, and its transposability in other regions. Furthermore, this methodology is adaptable according to the context, needs and data available. The quantification of the benefit of the correction especially using external data represents a complete and powerful validation process. Indeed, kriging methods or LUR models commonly use cross-validation which can overestimate the predictive ability of the model (Wang et al., 2013). Our approach allows also to better estimate the ozone concentration peaks, which are of great importance in daily exposure assessment and could be underestimated by other modeling approaches.

MARS models have been used instead of other classical models (such as generalized additive models (GAMs) and artificial neural networks (ANNs)) because they present a better performance, the possibility to introduce interactions between variables and the ability to capture complex nonlinear relationships between variables (Bordignon et al., 2002; Leathwick et al., 2006; Martin, 2011). MARS models are simple to apply and interpret, and require a short calculation time.

MARS models have already been efficaciously used to locally model atmospheric pollutant concentrations (García Nieto et al., 2015; García Nieto and Álvarez Antón, 2014; Roy et al., 2018). To our knowledge, they have never been used to adjust a pre-existing and well-established air quality model. Guillas et al. (2008) used model diagnostic and correction (MDC) integrating variables associated with wind speed, precipitation amounts and the diurnal cycle to predict the discrepancy and correct ozone concentrations. However, this approach only used linear regressions. This was not suitable for our data because of the nonlinear relationship between our discrepancy and our environmental variables.

5. Conclusion:

We identified a lack of exactitude and precision of transnational modeling downscaled at the kilometric scale compared with local measurements. We proposed a validated and efficient methodology, integrating local environmental variables, to remedy this and demonstrated its broader applicability. This methodology can be transferable to other areas, applicable to variable ranges of concentration and periods, and requires only a few environmental variables. This procedure is also flexible according to the context, needs and data available. Our approach allows us to improve the ozone exposure assessment. This is a more accessible tool, faster-processing, lower-cost than computing LUR or CTM, externally validated, and which better captures ozone peaks that can be associated with health effects (Di et al., 2017). The quantification of the benefit of the correction, especially using external data, constitutes a complete and powerful validation process. In the context of climate change, such an improvement will be of paramount importance for impact assessment of ozone exposure, including on wildlife (Sanderfoot and Holloway, 2017) or vegetation (Proietti et al., 2021).

Acknowledgments

The authors wish to thank everyone who contributed to the development of the project and would like to acknowledge Météo-France, especially Mr. Vermot-Desroches, for providing the meteorological data.

References

Abdulelah Al-Sudani, Z., Salih, S.Q., sharafati, A., Yaseen, Z.M., 2019. Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *J. Hydrol.* 573, 1–12. <https://doi.org/10.1016/j.jhydrol.2019.03.004>

- Bordignon, S., Gaetan, C., Lisi, F., 2002. Nonlinear models for ground-level ozone forecasting. *Stat. Methods Appl.* 11, 227–245. <https://doi.org/10.1007/BF02511489>
- Bouthevillain, C., Mathis, A., 1995. Prévisions : mesures, erreurs et principaux résultats. *Econ. Stat.* 285, 89–100. <https://doi.org/10.3406/estat.1995.5982>
- Braun, S., Achermann, B., De Marco, A., Pleijel, H., Karlsson, P.E., Rihm, B., Schindler, C., Paoletti, E., 2017. Epidemiological analysis of ozone and nitrogen impacts on vegetation – Critical evaluation and recommendations. *Science of The Total Environment* 603–604, 785–792. <https://doi.org/10.1016/j.scitotenv.2017.02.225>
- Brönnimann, S., Schuepbach, E., Zanis, P., Buchmann, B., Wanner, H., 2000. A climatology of regional background ozone at different elevations in Switzerland (1992–1998). *Atmos. Environ., Sixth Scientific Conference of the International Global Atmospheric* 34, 5191–5198. [https://doi.org/10.1016/S1352-2310\(00\)00193-X](https://doi.org/10.1016/S1352-2310(00)00193-X)
- Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. *The Lancet* 360, 1233–1242. [https://doi.org/10.1016/S0140-6736\(02\)11274-8](https://doi.org/10.1016/S0140-6736(02)11274-8)
- Castell-Balaguer, N., Téllez, L., Mantilla, E., 2012. Daily, seasonal and monthly variations in ozone levels recorded at the Turia river basin in Valencia (Eastern Spain). *Environ. Sci. Pollut. Res.* 19, 3461–3480. <https://doi.org/10.1007/s11356-012-0881-5>
- Chen, G., Li, J., Ying, Q., Sherman, S., Perkins, N., Sundaram, R., Mendola, P., 2014. Evaluation of observation-fused regional air quality model results for population air pollution exposure estimation. *Sci. Total Environ.* 485–486, 563–574. <https://doi.org/10.1016/j.scitotenv.2014.03.107>
- Chevalier, A., Gheusi, F., Delmas, R., Ordóñez, C., Sarrat, C., Zbinden, R., Thouret, V., Athier, G., Cousin, J.-M., 2007. Influence of altitude on ozone levels and variability in the lower troposphere: a ground-based study for western Europe over the period 2001–2004. *Atmospheric Chem. Phys.* 7, 4326.
- Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 389, 1907–1918. [https://doi.org/10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6)
- Crooks, J.L., Özkaynak, H., 2014. Simultaneous statistical bias correction of multiple PM2.5 species from a regional photochemical grid model. *Atmos. Environ.* 95, 126–141. <https://doi.org/10.1016/j.atmosenv.2014.06.024>
- Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017. Association of Short-term Exposure to Air Pollution With Mortality in Older Adults. *JAMA* 318, 2446–2456. <https://doi.org/10.1001/jama.2017.17923>
- Ezcurra, A., Benech, B., Echelecou, A., Santamaría, J.M., Herrero, I., Zulueta, E., 2013. Influence of local air flow regimes on the ozone content of two Pyrenean valleys. *Atmos. Environ.* 74, 367–377. <https://doi.org/10.1016/j.atmosenv.2013.03.051>
- Friedman, J.H., 1991. Multivariate Adaptive Regression Splines. *Ann. Stat.* 19, 1–67. <https://doi.org/10.1214/aos/1176347963>
- Friedman, J.H., Roosen, C.B., 1995. An introduction to multivariate adaptive regression splines. *Stat. Methods Med. Res.* 4, 197–217. <https://doi.org/10.1177/096228029500400303>

- García Nieto, P.J., Álvarez Antón, J.C., 2014. Nonlinear air quality modeling using multivariate adaptive regression splines in Gijón urban area (Northern Spain) at local scale. *Appl. Math. Comput.* 235, 50–65. <https://doi.org/10.1016/j.amc.2014.02.096>
- García Nieto, P.J., Antón, J.C.Á., Vilán, J.A.V., García-Gonzalo, E., 2015. Air quality modeling in the Oviedo urban area (NW Spain) by using multivariate adaptive regression splines. *Environ. Sci. Pollut. Res.* 22, 6642–6659. <https://doi.org/10.1007/s11356-014-3800-0>
- Garcia, V.C., Foley, K.M., Gego, E., Holland, D.M., Rao, S.T., 2010. A Comparison of Statistical Techniques for Combining Modeled and Observed Concentrations to Create High-Resolution Ozone Air Quality Surfaces. *J. Air Waste Manag. Assoc.* 60, 586–595. <https://doi.org/10.3155/1047-3289.60.5.586>
- Guillas, S., Bao, J., Choi, Y., Wang, Y., 2008. Statistical correction and downscaling of chemical transport model ozone forecasts over Atlanta. *Atmos. Environ.* 42, 1338–1348. <https://doi.org/10.1016/j.atmosenv.2007.10.027>
- Henrotin, J.-B., Zeller, M., Lorgis, L., Cottin, Y., Giroud, M., Béjot, Y., 2010. Evidence of the role of short-term exposure to ozone on ischaemic cerebral and cardiac events: the Dijon Vascular Project (DIVA). *Heart Br. Card. Soc.* 96, 1990–1996. <https://doi.org/10.1136/hrt.2010.200337>
- Hogrefe, C., Biswas, J., Lynn, B., Civerolo, K., Ku, J.-Y., Rosenthal, J., Rosenzweig, C., Goldberg, R., Kinney, P.L., 2004. Simulating regional-scale ozone climatology over the eastern United States: model evaluation results. *Atmos. Environ.* 38, 2627–2638. <https://doi.org/10.1016/j.atmosenv.2004.02.033>
- Honore, C., Rouil, L., Meleux, F., Malherbe, L., Bessagnet, B., Vautard, R., Poisson, N., Peuch, V.-H., Dufour, A., 2005. The PREV’AIR system, an operational system for large scale air quality forecasts over Europe; applications at the local scale. Presented at the International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes (HARMO 10), Crete, Greece, pp. 123–127.
- Janssen, S., Guerreiro, C., Viaene, P., Georgieva, E., Thunis, P., 2017. Guidance Document on Modelling Quality Objectives and Benchmarking.
- Joly, D., Gilbert, D., Diaz-de-Quijano, M., Hilal, M., Joly, M., Bernard, N., 2021. Enhancing air quality forecasts by geomatic downscaling: an application to daily PM10 concentrations in France. *Theor. Appl. Climatol.* 143, 327–339. <https://doi.org/10.1007/s00704-020-03418-7>
- Kato, T., 2016. Chapter 4 - Prediction of photovoltaic power generation output and network operation, in: Funabashi, T. (Ed.), *Integration of Distributed Energy Resources in Power Systems*. Academic Press, pp. 77–108. <https://doi.org/10.1016/B978-0-12-803212-1.00004-0>
- Leathwick, J.R., Elith, J., Hastie, T., 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecol. Model., Predicting Species Distributions* 199, 188–196. <https://doi.org/10.1016/j.ecolmodel.2006.05.022>
- Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* 525, 367–371. <https://doi.org/10.1038/nature15371>
- Lesmeister, C., 2019. *Mastering Machine Learning with R* Ed. 3. Packt Publishing.
- Li, R., Xu, M., Li, M., Chen, Z., Zhao, N., Gao, B., Yao, Q., 2021. Identifying the spatiotemporal variations in ozone formation regimes across China from 2005 to 2019 based on polynomial simulation and causality analysis. *Atmospheric Chemistry and Physics* 21, 15631–15646. <https://doi.org/10.5194/acp-21-15631-2021>

- Lu, X., Zhang, L., Shen, L., 2019. Meteorology and Climate Influences on Tropospheric Ozone: a Review of Natural Sources, Chemistry, and Transport Patterns. *Curr. Pollut. Rep.* 5, 238–260. <https://doi.org/10.1007/s40726-019-00118-3>
- Ma, R., Ban, J., Wang, Q., Li, T., 2020. Statistical spatial-temporal modeling of ambient ozone exposure for environmental epidemiology studies: A review. *Science of The Total Environment* 701, 134463. <https://doi.org/10.1016/j.scitotenv.2019.134463>
- Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Létinois, L., Markakis, K., Meleux, F., 2017. CHIMERE-2017: from urban to hemispheric chemistry-transport modeling. *Geosci. Model Dev.* 10, 2397–2423. <https://doi.org/10.5194/gmd-10-2397-2017>
- Martin, E.P., 2011. Comparative Performance of Different Statistical Models for Predicting Ground-Level Ozone (O₃) and Fine Particulate Matter (PM_{2.5}) Concentrations in Montréal, Canada (masters). Concordia University.
- Martin, N., 2010. Performances des modélisations déterministes d’ozone à méso-échelle et à micro-échelle dans les Alpes-Maritimes. *Cybergeo Rev. Eur. Géographie Eur. J. Geogr.* <http://cybergeo.revues.org/index23183.html>.
- Marx, T., Bernard, N., Parmentier, A.-L., Puyraveau, M., Martin, B., Gantelet, M., Pretalli, J.-B., Dalphin, J.-C., Mauny, F., Desmettre, T., 2019. Does air pollution really impact the onset of spontaneous pneumothorax? A French case-crossover study. *Environ. Int.* 127, 317–323. <https://doi.org/10.1016/j.envint.2019.01.056>
- Meleux, F., Ung, A., 2020. Performances Prev’Air en 2017 et 2018. LCSQA.
- Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., Vivanco, M.G., 2013. CHIMERE 2013: a model for regional atmospheric composition modelling. *Geosci. Model Dev.* 6, 981–1028. <https://doi.org/10.5194/gmd-6-981-2013>
- Nuvolone, D., Petri, D., Voller, F., 2018. The effects of ozone on human health. *Environ. Sci. Pollut. Res.* 25, 8074–8088. <https://doi.org/10.1007/s11356-017-9239-3>
- Proietti, C., Fornasier, M.F., Sicard, P., Anav, A., Paoletti, E., De Marco, A., 2021. Trends in tropospheric ozone concentrations and forest impact metrics in Europe over the time period 2000–2014. *J. For. Res.* 32, 543–551. <https://doi.org/10.1007/s11676-020-01226-3>
- Ramos, Y., Requia, W.J., St-Onge, B., Blanchet, J.-P., Kestens, Y., Smargiassi, A., 2018. Spatial modeling of daily concentrations of ground-level ozone in Montreal, Canada: A comparison of geostatistical approaches. *Environ. Res.* 166, 487–496. <https://doi.org/10.1016/j.envres.2018.06.036>
- Ren, X., Mi, Z., Georgopoulos, P.G., 2020. Comparison of Machine Learning and Land Use Regression for fine scale spatiotemporal estimation of ambient air pollution: Modeling ozone concentrations across the contiguous United States. *Environment International* 142, 105827. <https://doi.org/10.1016/j.envint.2020.105827>
- Riviere, E., Bernard, J., Hulin, A., Virga, J., Dugay, F., Charles, M.-A., Cheminat, M., Cortinovis, J., Ducroz, F., Laborie, A., Malherbe, L., Piga, D., Real, E., Robic, P.-Y., Zaros, C., Seyve, E., Lepeule, J., 2019. Air pollution modeling and exposure assessment during pregnancy in the French Longitudinal Study of Children (ELFE). *Atmos. Environ.* 205, 103–114. <https://doi.org/10.1016/j.atmosenv.2019.02.032>
- Rouil, L., Honoré, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F., Dufour, A., Elichegaray, C., Flaud, J.-M., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., Poisson, N., 2009. Prev’air: An Operational Forecasting and Mapping System for

- Air Quality in Europe. *Bull. Am. Meteorol. Soc.* 90, 73–84. <https://doi.org/10.1175/2008BAMS2390.1>
- Roy, S.S., Pratyush, C., Barna, C., 2018. Predicting Ozone Layer Concentration Using Multivariate Adaptive Regression Splines, Random Forest and Classification and Regression Tree, in: Balas, V.E., Jain, L.C., Balas, M.M. (Eds.), *Soft Computing Applications, Advances in Intelligent Systems and Computing*. Springer International Publishing, Cham, pp. 140–152. https://doi.org/10.1007/978-3-319-62524-9_11
- Royal Society (Great Britain), 2008. *Ground-level ozone in the 21st century: future trends, impacts and policy implications*. The Royal Society, London.
- Sanderfoot, O.V., Holloway, T., 2017. Air pollution impacts on avian species via inhalation exposure and associated outcomes. *Environ. Res. Lett.* 12, 083002. <https://doi.org/10.1088/1748-9326/aa8051>
- Schaap, M., Cuvelier, C., Hendriks, C., Bessagnet, B., Baldasano, J.M., Colette, A., Thunis, P., Karam, D., Fagerli, H., Graff, A., Kranenburg, R., Nyiri, A., Pay, M.T., Rouil, L., Schulz, M., Simpson, D., Stern, R., Terrenoire, E., Wind, P., 2015. Performance of European chemistry transport models as function of horizontal resolution. *Atmospheric Environment* 112, 90–105. <https://doi.org/10.1016/j.atmosenv.2015.04.003>
- Sekulic, S., Kowalski, B.R., 1992. MARS: A tutorial. *J. Chemom.* 6, 199–216. <https://doi.org/10.1002/cem.1180060405>
- Sicard, P., Agathokleous, E., De Marco, A., Paoletti, E., Calatayud, V., 2021. Urban population exposure to air pollution in Europe over the last decades. *Environ Sci Eur* 33, 28. <https://doi.org/10.1186/s12302-020-00450-2>
- Sicard, P., De Marco, A., Troussier, F., Renou, C., Vas, N., Paoletti, E., 2013. Decrease in surface ozone concentrations at Mediterranean remote sites and increase in the cities. *Atmospheric Environment* 79, 705–715. <https://doi.org/10.1016/j.atmosenv.2013.07.042>
- Sicard, P., Serra, R., Rossello, P., 2016. Spatiotemporal trends in ground-level ozone concentrations and metrics in France over the time period 1999–2012. *Environmental Research* 149, 122–144. <https://doi.org/10.1016/j.envres.2016.05.014>
- Steven Porter, P., Rao, S.T., Hogrefe, C., Gego, E., Mathur, R., 2015. Methods for reducing biases and errors in regional photochemical model outputs for use in emission reduction and exposure assessments. *Atmos. Environ.* 112, 178–188. <https://doi.org/10.1016/j.atmosenv.2015.04.039>
- Thunis, P., Miranda, A., Baldasano, J.M., Blond, N., Douros, J., Graff, A., Janssen, S., Juda-Rezler, K., Karvosenoja, N., Maffei, G., Martilli, A., Rasoloharimahefa, M., Real, E., Viaene, P., Volta, M., White, L., 2016. Overview of current regional and local scale air quality modelling practices: Assessment and planning tools in the EU. *Environ. Sci. Policy, Multidisciplinary research findings in support to the EU air quality policy: experiences from the APPRAISAL, SEFIRA and ACCENT-Plus EU FP7 projects* 65, 13–21. <https://doi.org/10.1016/j.envsci.2016.03.013>
- U.S. EPA, 2018. *Modeling Guidance for Demonstrating Air Quality Goals for Ozone, PM2.5, and Regional Haze (No. EPA 454/R-18-009)*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Air Quality Assessment Division, Research Triangle Park, NC, USA.
- Wang, M., Beelen, R., Basagana, X., Becker, T., Cesaroni, G., de Hoogh, K., Dedele, A., Declercq, C., Dimakopoulou, K., Eeftens, M., Forastiere, F., Galassi, C., Gražulevičienė, R., Hoffmann, B., Heinrich, J., Iakovides, M., Künzli, N., Korek, M., Lindley, S., Mölter, A., Mosler, G., Madsen, C., Nieuwenhuijsen, M., Phuleria, H.,

- Pedeli, X., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Stempfelet, M., Tsai, M.-Y., Lanki, T., Udvardy, O., Varró, M.J., Wolf, K., Weinmayr, G., Yli-Tuomi, T., Hoek, G., Brunekreef, B., 2013. Evaluation of Land Use Regression Models for NO₂ and Particulate Matter in 20 European Study Areas: The ESCAPE Project. *Environ. Sci. Technol.* 47, 4357–4364. <https://doi.org/10.1021/es305129t>
- Wang, M., Sampson, P.D., Hu, J., Kleeman, M., Keller, J.P., Olives, C., Szpiro, A.A., Vedal, S., Kaufman, J.D., 2016. Combining Land-Use Regression and Chemical Transport Modeling in a Spatiotemporal Geostatistical Model for Ozone and PM_{2.5}. *Environ. Sci. Technol.* 50, 5111–5118. <https://doi.org/10.1021/acs.est.5b06001>
- Wang, T., Xue, L., Brimblecombe, P., Lam, Y.F., Li, L., Zhang, L., 2017. Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects. *Sci. Total Environ.* 575, 1582–1596. <https://doi.org/10.1016/j.scitotenv.2016.10.081>
- Yilmaz, B., Aras, E., Nacar, S., Kankal, M., 2018. Estimating suspended sediment load with multivariate adaptive regression spline, teaching-learning based optimization, and artificial bee colony models. *Sci. Total Environ.* 639, 826–840. <https://doi.org/10.1016/j.scitotenv.2018.05.153>
- Zhang, J., Wei, Y., Fang, Z., 2019. Ozone Pollution: A Major Health Hazard Worldwide. *Front. Immunol.* 10, 2518. <https://doi.org/10.3389/fimmu.2019.02518>

[Figure 1]

Figure 1: PREV'EST modeling area (for the 3 km x 3 km grid mesh, on the top map) and location of the air quality monitoring stations (ATMO BFC) and weather stations (Météo-France).

[Figure 2]

Figure 2: Target plots of the uncorrected PREV'EST modeling.

Different letters represent the different air quality monitoring stations.

[FIGURE 3]

Figure 3: Target plots of the uncorrected PREV'EST modeling, global approach 4-WF and 2-WF elevation corrections, annual approach 4-WF and 2-WF elevation corrections, territorial approach 4-WF correction and territorial and annual approach 4-WF correction target plots.

Different letters represent the different air quality monitoring stations.

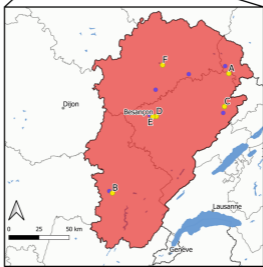
[FIGURE 4]

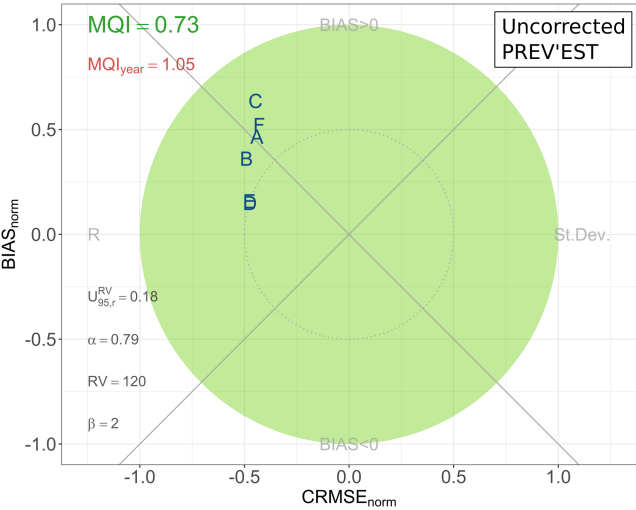
Figure 4: Daily mean ozone concentrations measured, modeled by uncorrected PREV'EST modeling and corrected by the annual 2-WF elevation correction, during 2018 at Station A.

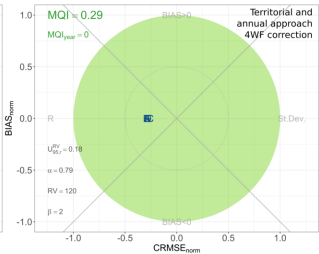
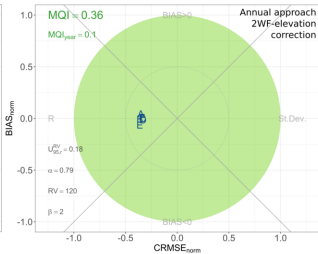
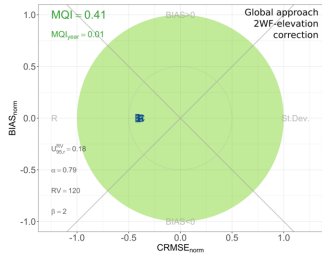
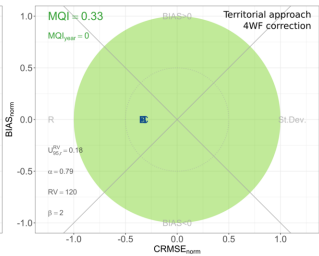
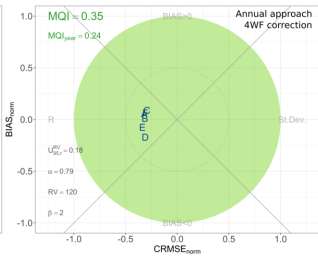
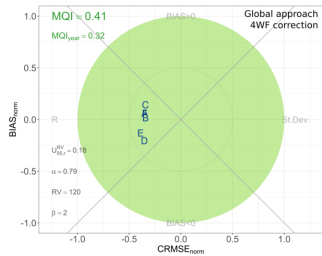
[FIGURE 5]

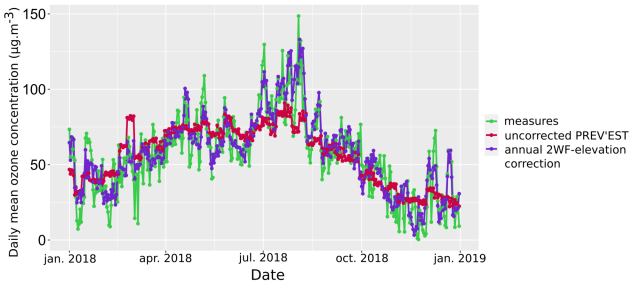
Figure 5: External validation: uncorrected PREV'EST modeling, annual approach 4-WF correction, annual approach 2-WF elevation correction, and annual and territorial approach 4-WF and 2-WF correction target plots.

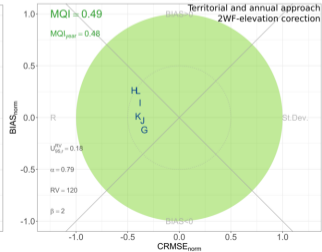
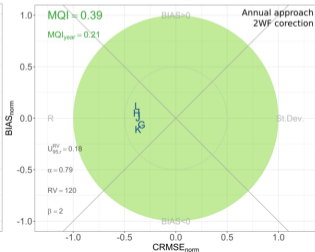
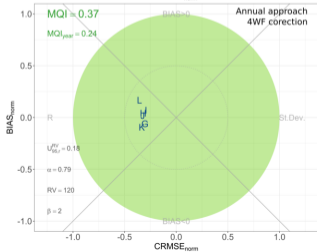
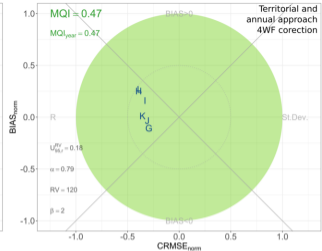
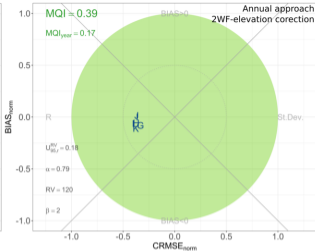
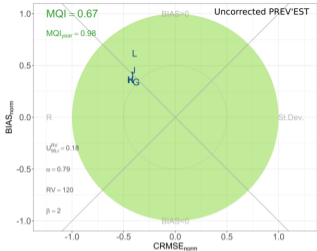
Different letters represent the different air quality monitoring stations.











Title: Improvement of downscaled ozone concentrations from the transnational scale to the
kilometric scale: need, interest and new insights

Table 1: MARS modeling approaches and corrections

Type of approach	Kind of correction applied ²	Number of models per approach	Number of models per kind of correction
Global	4WF	3	1
	2WF-elevation		1
	2WF		1
Annual	4WF	36	12
	2WF-elevation		12
	2WF		12
Territorial¹	4WF	12	6
	2WF		6
Territorial and annual¹	4WF	130 ³	65 ³
	2WF		65 ³
Total		179	

¹The territorial and, the territorial and annual approaches did not include a 2WF-elevation correction because the elevation was steady (only one station considered per model).

² 4WF: correction including 4 weather factors (Temperature (T°), Relative Humidity (RH), Rainfall Amounts (RA), Wind speed parameters (WS)) and the date; 2WF – elevation: correction including 2 weather factors (T°, RA), the date and the elevation; 2WF: correction including 2 weather factors (T°, RA), and the date.

³ Theoretically we should have 72 models but the station E opened only in 2013 and the station B opened only in 2010.

Title: Improvement of downscaled ozone concentrations from the transnational scale to the
kilometric scale: need, interest and new insights

Table 2: Average performance indicators of the different approaches and corrections

Kind of correction applied	MBE ^a	MAE ^a	RMSE ^a	r ^b
Uncorrected PREV'EST	14.67	19.23	23.18	0.67
Global 4WF correction	0.00	10.78	13.64	0.84
Global 2WF-elevation correction	0.00	11.37	14.32	0.81
Global 2WF correction	0.03	12.21	15.34	0.78
Annual 4WF correction	0.02	9.40	11.92	0.87
Annual 2WF-elevation correction	0.01	10.04	12.67	0.85
Annual 2WF correction	0.04	10.81	13.60	0.83
Territorial 4WF correction	0.00	9.41	11.96	0.88
Territorial 2WF correction	0.00	10.92	13.76	0.83
Territorial and annual 4WF correction	0.01	8.00	10.19	0.91
Territorial and annual 2WF correction	0.02	9.52	12.05	0.87

a Mean bias error (MBE), mean absolute error (MAE) and root mean square error (RMSE) should be minimized to reflect better a performance.

b r: Pearson correlation coefficient should be maximized to reach better a performance.

Title: Improvement of downscaled ozone concentrations from the transnational scale to the
kilometric scale: need, interest and new insights

Table 3: Average performance indicators calculated for the external validation.

	MBE	MAE	RMSE	r
Uncorrected PREV'EST	14.83	17.63	21.35	0.74
Global 4WF correction	0.54	10.24	13.00	0.85
Global 2WF-elevation correction	-0.96	11.64	14.70	0.77
Global 2WF correction	-0.23	11.55	14.56	0.78
Annual 4WF correction	0.51	9.48	12.11	0.85
Annual 2WF-elevation correction	-2.01	10.74	13.59	0.81
Annual 2WF correction	0.04	10.76	13.61	0.81
Territorial 4WF correction	4.87	12.26	15.61	0.79
Territorial 2WF correction	3.67	12.76	16.11	0.75
Territorial and annual 4WF correction	3.45	11.11	14.29	0.81
Territorial and annual 2WF correction	3.42	12.31	15.67	0.76

Number of statistical units (number of discrepancies predicted): 14 846

The prediction was performed on data of the air quality monitoring station which were not integrated into the MARS models for its training (additional stations, n = 6).