Environmetrics

WILEY

RESEARCH ARTICLE OPEN ACCESS

Generalized Additive Model With Dynamic Coefficients for Spatiotemporal Ozone Predictions

Abdollah Jalilian[1] | Claudia Cappello[2] | Monica Palma[2] | Sandra De Iaco[2,3]

[1]Lancaster Medical School, Lancaster University, Lancaster, UK | [2]Department of Economic Sciences, University of Salento, Lecce, Italy | [3]National Biodiversity Future Center - NBFC, Palermo, Italy

Correspondence: Sandra De Iaco (sandra.deiaco@unisalento.it)

Received: 1 June 2025 | **Revised:** 10 January 2026 | **Accepted:** 19 January 2026

Keywords: additive models | dynamic coefficients | ozone concentration

ABSTRACT

Accurate prediction of surface-level ozone concentrations is critical for air quality management and public health protection. This study develops a flexible spatiotemporal statistical modeling framework to predict daily mean O_3 concentrations across Italy by integrating satellite-derived ozone estimates with ground-based observations and high-resolution environmental predictors. The proposed model is based on a linear regression with dynamic intercept and slope that relate in situ O_3 measurements to satellite data, explicitly addressing additive (systematic shifts) and multiplicative (scaling) biases in satellite-derived ozone estimates. These spatiotemporally varying coefficients are modeled through a generalized additive model framework, allowing the capture of complex and potentially nonlinear relationships between ozone levels and environmental covariates. This unified and interpretable approach enables a detailed understanding of bias patterns in satellite data. Model diagnostics and crossvalidation demonstrate superior explanatory power and predictive performance compared to simpler models. The interpretability of the model is illustrated by revealing the influence of elevation, nitrogen dioxide concentrations, and seasonal variation on bias structures. Furthermore, the model's downscaling capability is demonstrated by producing fine-scale ozone concentration predictions over Italy and its surrounding regions. The proposed modeling framework offers an accurate, scalable, and interpretable tool for mapping surface-level ozone, supporting improved environmental monitoring and informing policy decisions.

1 | Introduction

Ambient ozone (O_3), a secondary pollutant formed through photochemical reactions involving precursors such as nitrogen oxides and volatile organic compounds, poses serious risks to human health globally (Nuvolone et al. 2018; Zhang et al. 2019). It also negatively affects crop quality and yield, thereby threatening food security (Ramya et al. 2023). Accurate estimation of O_3 exposure is therefore essential and requires the ability to monitor and predict surface-level concentrations at high spatial and temporal resolutions.

To this end, physics-based photochemical grid models and data-driven approaches, including statistical and machine learning models, have been developed over the past two decades to predict O_3 concentrations across space and time. Physics-based models, such as the Comprehensive Air Quality Model with extensions (CAMx), simulate atmospheric processes and use mechanistic equations that represent the physical and chemical interactions between O_3 and environmental variables to provide a process-oriented understanding and predictive capability (Emery et al. 2024). In contrast, machine learning models aim to uncover complex underlying relationships in available O_3 and environ-

mental data using flexible algorithms, such as random forests and neural networks, that do not rely on prior assumptions about the underlying relationships (Meng et al. 2022; Pan et al. 2023). Several machine learning methods were recently carried out for air pollution prediction, including random forests (Zhan et al. 2018; Guo et al. 2021), support vector regression (Castelli et al. 2020), artificial neural networks (Sipilä et al. 2025), and deep learning approaches (Niu et al. 2023). Some researchers also proposed a combination of chemical transport models, which simulate diffusion and transfer of atmospheric processes based on physical and chemical laws and numerical methods, with machine learning procedures to produce high spatial resolution estimations of the investigated pollutants (Silibello et al. 2021; Cedeno Jimenez et al. 2023). Machine learning algorithms are leveraged for their high prediction accuracy and ability to model complex, nonlinear relationships among pollutants, meteo-climatic variables, and geo-morphological conditions, anthroplogical factors.

Statistical models, on the other hand, provide a middle ground by relying on observed data and probabilistic assumptions to construct interpretable relationships between O_3 concentrations and relevant environmental variables, while also quantifying the uncertainty in predictions. This interpretability and ability to assess uncertainty in the results make statistical models particularly valuable in environmental analysis (Gelfand et al. 2019).

Among statistical approaches, time series models, such as the Seasonal AutoRegressive Integrated Moving Average (SARIMA) model, are commonly used for short-term forecasting and for understanding the temporal dynamics of air pollutants. These models capture key components such as trends, seasonality, and autocorrelation in pollutant concentrations over time within a specific area (Bhatti et al. 2021). Nonparametric regression models, such as bivariate splines, are used to model complex spatial patterns in surface-level O_3 (Ettinger et al. 2012). This improves prediction accuracy by enabling smooth variation in pollutant levels across geographic space. Other contributions coming from the geostatistical sphere are based on coregionalization modeling of spatial or spatio-temporal environmental scalar or vectorial variables (Wackernagel 2003; De Iaco et al. 2003, 2005; De Iaco 2017; Cappello, De Iaco, Maggio, and Posa 2020; De Iaco 2022; Muehlmann et al. 2022; De Iaco 2023a; 2023b). Hierarchical Bayesian models are flexible models suitable for complex spatiotemporal air quality data, as they can incorporate multiple levels of data structures. In McMillan et al. (2005), for example, the proposed hierarchical model incorporates linkages between O_3, and meteorology, including temperature, humidity, pressure, and wind speed and direction. For a more comprehensive modeling approach that includes the influence of environmental and geographic factors, land-use regression (LUR) models are commonly employed for pollutants like O_3 (Hoek et al. 2008). LUR models establish statistical relationships between pollutant concentrations measured at monitoring sites and various spatially resolved geographic and environmental predictors (van Nunen et al. 2017). These predictors typically include elevation, population density, proximity to roads, land-use and land-cover types, emission sources, and meteorological conditions. LUR models typically assume a multiple linear regression framework to relate O_3 concentrations to these predictors.

However, to better capture complex and possibly nonlinear spatial and temporal patterns, generalized additive models (GAMs) have been considered (Wang et al. 2020). Varying coefficient models were first introduced by Hastie and Tibshirani (1993), and GAMs incorporating varying coefficients have been widely applied in air pollution research. For instance, Hayn et al. (2009) used GAMs to investigate global spatiotemporal patterns of tropospheric NO_2 using satellite data from the Global Ozone Monitoring Experiment (GOME) covering the period 1996–2001. Similarly, Wang et al. (2020) applied GAMs to estimate daily mean ozone concentrations at 27 monitoring stations in Tianjin, northeast China, for the year 2016. In a broader context, Wang et al. (2022) used GAMs to model and predict daily mean concentrations of O_3, NO_2, $PM_{2.5}$, and PM_{10} across 48 to 85 monitoring stations throughout Great Britain between 2011 and 2015. To further account for unobserved heterogeneity or hierarchical structure in the data, random effects can be incorporated, resulting in a mixed-effects model formulation (Wang et al. 2022).

In this study, an additive regression model with spatiotemporal varying coefficients is proposed to relate the daily mean O_3 concentrations, measured at air quality monitoring stations across Italy, to satellite-derived O_3 estimates. Similar to LUR models, the model incorporates a set of environmental variables within an additive framework to capture variations in both the intercept and the slope coefficients across space and time. Notably, integrating an additive regression model with varying coefficients and random effects that describe the seasonal and cyclic patterns of O_3 within a unified framework represents a novel modeling approach for linking satellite observations to in situ measurements. Moreover, the proposed additive structure can contribute to increasing the model's flexibility and allows it to catch potentially nonlinear effects of covariates included in the linear relation. After evaluating the model's goodness of fit, its spatial and temporal predictive performance is assessed through an extensive cross-validation analysis. The best-fitted model is then applied for downscaling and forecasting daily mean O_3 concentrations on a high-resolution grid covering Italy and surrounding regions.

The remainder of the paper is structured as follows: Section 2 describes the data used in the analysis, Section 3 introduces the modeling framework, and Section 4 presents the evaluation of the model's adequacy and predictive performance. Finally, conclusions and remarks are provided in Section 5.

2 | Data Description

In this study, in situ O_3 concentrations obtained from air quality monitoring stations serve as the primary response variable. The covariates include satellite-based estimates of O_3 and a set of other environmental variables representing atmospheric composition, meteorological conditions, topography, land use, and human activity. The following subsections describe the structure, acquisition, and pre-processing procedures for the above data.

2.1 | Ozone Monitoring Data

The European Air Quality Portal, maintained by the European Environment Agency (EEA), provides access to air pollution

data collected from a network of participating and cooperating countries (European Environment Agency 2025). From this portal and the `R` (R Core Team 2025) package `EEAaq` (Maranzano et al. 2025), validated hourly O_3 concentration data were obtained for 402 air quality monitoring stations across Italy. These hourly records were aggregated to calculate the mean daily O_3 concentration at each station. Appendix A presents the `R` code used for data retrieval and pre-processing. The dataset spans an 11-year period from January 1, 2013, to December 30, 2023, comprising a total of 4015 days. Note that the station measurements can be reasonably treated as representing point locations, corresponding to the sites where the monitoring instruments are installed and reflecting the same effective area.

Data availability varied across stations; not all stations reported measurements for every day within the study period. The number of valid recording days per station ranged from 64 to 4006, with a mean of approximately 3006 days. Conversely, the number of stations reporting data on a given day ranged from 202 to 357, with an average of about 302 stations per day. Figure 1 displays the spatial distribution of the monitoring stations. The size of each circle reflects the total number of days with recorded O_3 concentrations, where larger circles indicate more complete coverage of the data. Moreover, the stations are distributed across the entire territory, and the inclusion of land-use as a covariate can help capture spatial variability and the effects of different processes occurring over the domain. The bottom panel of the figure shows the daily number of active stations throughout the study period, highlighting an overall increasing trend in station coverage, although with noticeable fluctuations. Sudden increases are evident around the beginning of 2014, 2016, and 2018. These variations in validated data availability are primarily attributed to administrative and technical factors, such as equipment maintenance, operational interruptions, or network-level changes in data reporting, and not to the underlying O_3 concentration levels themselves. Therefore, it is reasonable to assume that the missing data mechanism is missing at random. Under this assumption, the analysis proceeds using only the available data without requiring imputation or explicit modeling of the missingness process.

The mean daily O_3 concentration values exhibit substantial variation across time and space and are notably right-skewed: most days show moderate levels, while a smaller number record high concentrations. To reduce this skewness and approximate a more symmetric distribution, a square-root transformation was applied, which, as shown in the Supporting Information, reduced the right-skewness of the data. The square root transformation of O_3 data is commonly used (see, e.g., Dou et al. 2010; Lu et al. 2018) and was also supported, as demonstrated in the Supporting Information, by applying the Box–Cox family of transformations to the data (Box and Cox 1964; Eastoe 2009). Figure 2a shows boxplots of the square root of mean daily O_3 concentrations, grouped by month. These plots reveal a distinct seasonal cycle: concentrations rise in spring, peak in summer, and decline through autumn and winter. This seasonal pattern is well-documented and is primarily driven by meteorological factors such as sunlight and temperature. Nonetheless, the timing and intensity of O_3 peaks can vary depending on geographic and local environmental conditions (Yan et al. 2018). The boxplots also show an increase in extremely high values beginning in early 2022, likely due to the reduced and fluctuating number of monitoring stations contributing data during this period, as illustrated in the bottom panel of Figure 1.

2.2 | Satellite-Based Covariates

The Copernicus Atmosphere Monitoring Service (CAMS) provides hourly reanalysis data on air quality across Europe at a spatial resolution of 0.1 degrees, approximately 10 km, using the median of outputs from eleven different air quality data assimilation models (Copernicus Atmosphere Monitoring Service 2021). The dataset currently spans the period from 2013 to 2023. Although CAMS offers up to 19 air quality variables, only four are consistently available at surface level over the full period on a regular $0.1^\circ \times 0.1^\circ$ grid covering Italy and surrounding areas: O_3, nitrogen dioxide (NO_2), fine particulate matter ($PM_{2.5}$; particles smaller than 2.5 μm), and coarse particulate matter (PM_{10}; particles smaller than 10 μm). The hourly values of these pollutants were aggregated to obtain daily mean concentrations.

Figure 2b shows the square root of daily mean O_3 concentrations from CAMS, extracted from grid cells that contain monitoring stations over the 11-year study period. Differences between CAMS estimates and in situ measurements are expected due to several factors, including vertical gradients in O_3 concentrations, the relatively coarse spatial resolution of CAMS (approximately 10 km), and limitations inherent to satellite retrievals. Although satellite-based O_3 products have improved in accuracy and resolution, they still face challenges near the surface due to weak signals, complex land surface features, and interference from clouds and aerosols (Xu et al. 2024). These factors contribute to discrepancies between satellite-derived and ground-based measurements. Despite these limitations, the square-root transformed O_3 concentrations from CAMS and air quality stations show a strong overall linear relationship, with a Pearson correlation coefficient of 0.9139 (95% CI: 0.9135–0.9143). However, Figure 2c reveals temporal fluctuations in this relationship over the study period. In addition, the empirical variograms of the differences between square-root transformed CAMS and station values, computed per year and per day of the week and shown in Figure 2d,e, respectively, indicate no spatial structure in the discrepancies. This suggests that the differences between satellite-based and in situ measurements are not systematically influenced by geographic proximity.

To account for meteorological influences, climate variables were obtained from the European ReAnalysis Land (ERA5-Land) dataset, provided by the Copernicus Climate Data Store (Muñoz Sabater 2019). ERA5-Land offers consistent hourly reanalysis data of land surface variables over several decades, with the same spatial resolution and grid structure as CAMS. Using the same aggregation method, hourly values for 2 m air temperature, surface net solar radiation, total precipitation, wind speed and direction, relative humidity, boundary layer height, and leaf area index (for both high and low vegetation) were aggregated to daily means for the period 2013–2023. For the categorical variable wind direction, daily values were derived by computing the mode (most frequently occurring direction) from the hourly observations.

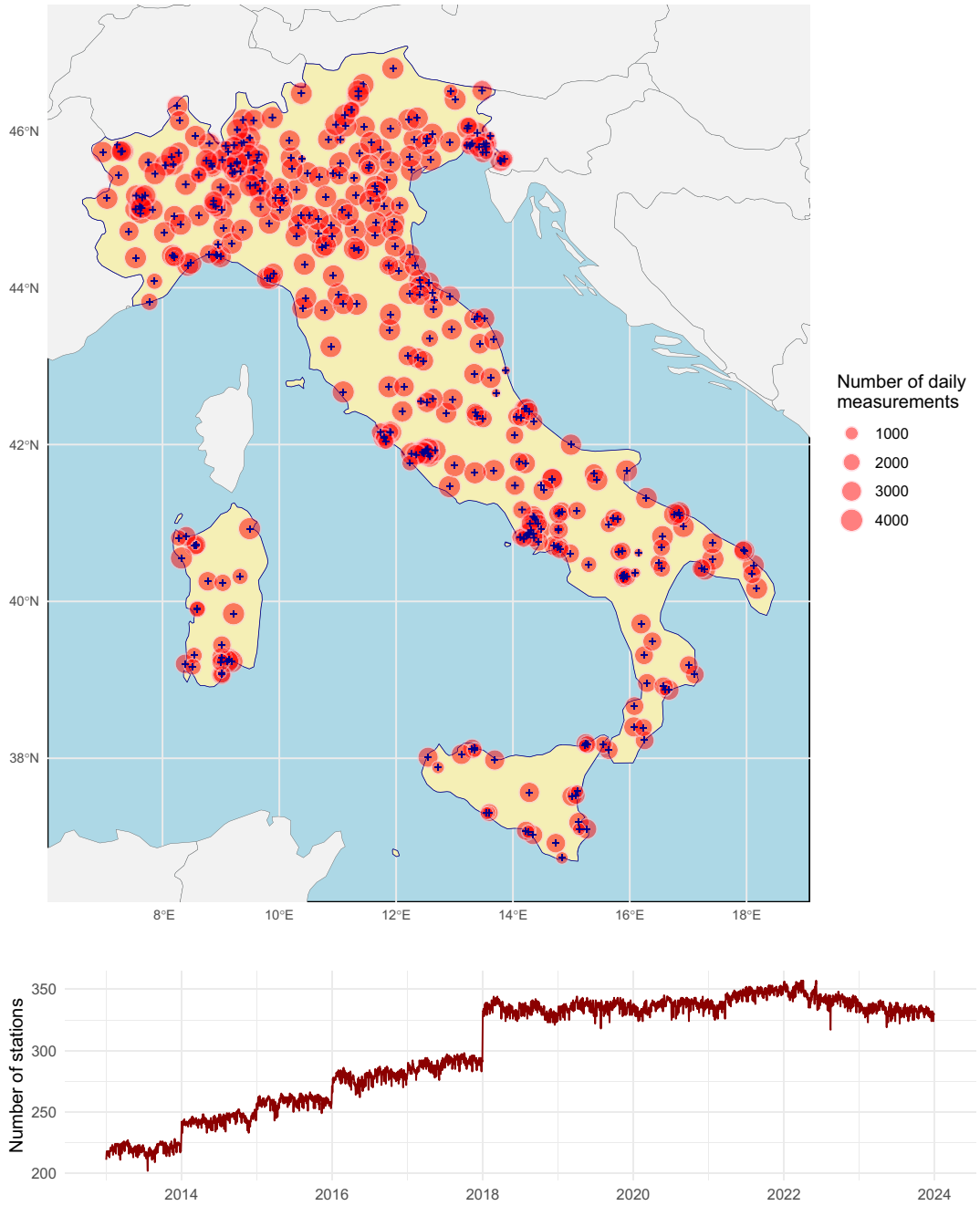

FIGURE 1 | Air quality monitoring stations in Italy from 2013 to 2023: spatial distribution (top) and daily number of stations recording ozone concentration (bottom). In the map (top), each station is marked with a cross, and the circle size indicates the number of available daily measurements at each location.

In addition to spatiotemporal atmospheric and climate variables, several spatial environmental and anthropogenic covariates were incorporated across the gridded domain covering Italy and its surrounding regions. Elevation data were sourced from the Copernicus Digital Elevation Model, Global 90-meter (GLO-90) resolution (European Space Agency 2025), which has a spatial resolution of 0.0008°. Elevation influences atmospheric processes that affect O_3 transport, dispersion, and accumulation, especially in complex terrains. Population density data were extracted from the 2015 version of NASA's Gridded Population of the World (GPWv4) dataset (Center For International Earth Science Information Network, Columbia University 2018), with a spatial resolution of 0.045°. This serves as a proxy for human activity and related anthropogenic emissions, which contribute to O_3 precursor concentrations. Road density data were obtained from the Global Roads Inventory Project (GRIP) (Meijer et al. 2018), providing global estimates of road density (including highways, primary, secondary, tertiary, and local roads) at a resolution of 0.083°. Road infrastructure is a key indicator of traffic-related emissions affecting O_3 formation. Land cover and land use data were derived from the Global Land Cover and Land Use Change dataset based on Landsat archive imagery from 2015

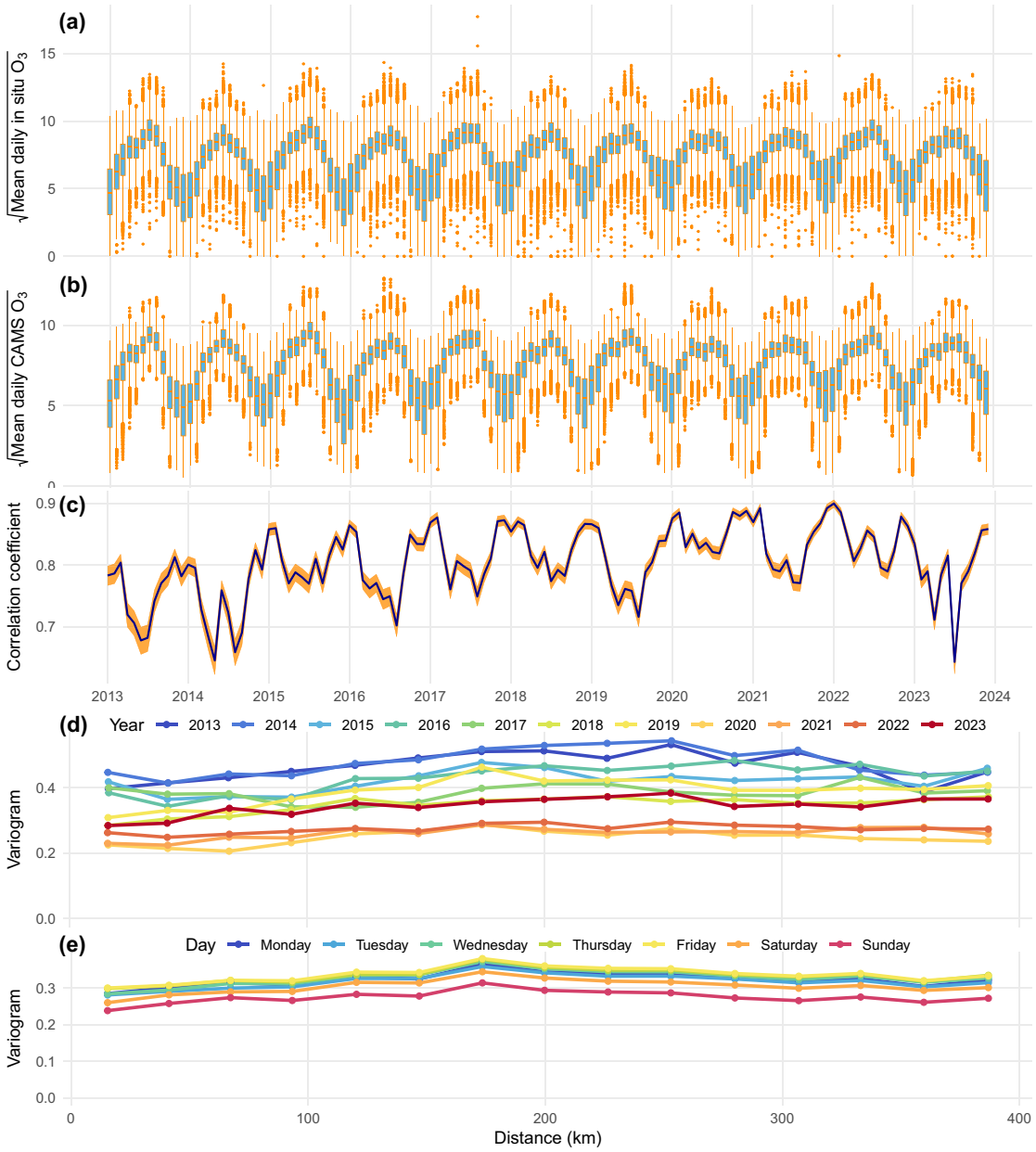

FIGURE 2 | Boxplots of the square root of mean daily ozone (O_3) concentrations grouped by month from January 2013 to December 2023, based on data from (a) monitoring stations and (b) the corresponding grid cells from the Copernicus Atmosphere Monitoring Service (CAMS). Panel (c) shows the monthly Pearson correlation coefficients between the two data sources, along with 95% confidence intervals. Panels (d) and (e) present the empirical variograms of the mean differences between station and CAMS values across all monitoring sites, per year and per day of the week, respectively.

(Potapov et al. 2022). The land surface was classified into seven major categories: semi-arid, dense short vegetation, tree cover, wetland, cropland, built-up area, and ocean. These categories influence biogenic emissions, surface albedo, and atmospheric chemistry, thereby affecting local and regional O_3 dynamics.

A summary of all variables included in the analysis, along with their hypothesized roles in O_3 formation and their Spearman correlation with observed in situ O_3 concentrations at monitoring stations, is provided in Table 1.

Finally, it is worth noting that several air quality datasets for the Italian territory have recently been collected and harmonized within the research project AgrImOnIA (Agricultural Impact On Italian Air), available at https://doi.org/10.5281/zenodo.6620529 (Fassó et al. 2023). In particular, extensive datasets on air quality, meteorology, emissions, livestock, and land use across Italy have been organized and made freely accessible to policymakers, researchers, students, and the general public. Furthermore, under the GRINS project (https://grins.it), a national partnership involving universities, research institutes, and private companies, large multisource datasets have been gathered and integrated within the open-data AMELIA platform. This platform supports the collection, integration, analysis, and dissemination of high-volume data related to the environment, population, and economics. All datasets are openly available on Zenodo (https://doi.org/10.5281/zenodo.15699805).

TABLE 1 | Environmental variables, their relevance to ground-level ozone (O_3), Spearman correlation (Corr) with observed in situ O_3 concentrations at monitoring stations and data sources. Note that Spearman correlation is not computed for the categorical variables: wind direction and land cover/land use.

Variable (Unit)	Relevance to O_3	Corr	Source
O_3 ($\mu g/m^3$)	Estimate of near-surface O_3	0.89	CAMS[a]
Nitrogen dioxide ($\mu g/m^3$)	Major O_3 precursor	−0.57	CAMS
$PM_{2.5}$ ($\mu g/m^3$)	Influences O_3 chemistry and light scattering	−0.24	CAMS
PM_{10} ($\mu g/m^3$)	Affects atmospheric chemistry and precursor dynamics	−0.21	CAMS
Solar radiation (W/m^2)	Provides energy for O_3 forming reactions	0.67	ERA5-Land[b]
Air temperature at 2m (°C)	Speeds up O_3 production	0.52	ERA5-Land
Boundary layer height (m)	Controls vertical mixing and O_3 dispersion	0.52	ERA5-Land
Relative humidity (%)	High levels may reduce O_3 via radical suppression	−0.45	ERA5-Land
Leaf area index: Low vegetation (m^2/m^2)	Minor effect on O_3 through surface exchange	0.31	ERA5-Land
Leaf area index: High vegetation (m^2/m^2)	Removes NOx, indirectly affecting O_3	0.29	ERA5-Land
Wind speed (m/s)	Facilitates horizontal O_3 transport	0.16	ERA5-Land
Wind direction (↑, →, ↓, ←)	Influences transport of O_3 precursors	—	ERA5-Land
Total precipitation (m)	Can reduce O_3 via scavenging of precursors	−0.06	ERA5-Land
Elevation (m)	Influences background O_3 and atmospheric mixing	0.17	CDEM[c]
Population density (people/km^2)	Proxy for anthropogenic emissions	−0.15	GPWv4[d]
Road density (m/km^2)	Indicator of traffic-related emissions	−0.16	GRIP[e]
Land cover and land use (10 categories)	Influences surface energy balance and emissions	—	GLCLUC[f]

[a]Copernicus Atmosphere Monitoring Service (CAMS).
[b]European ReAnalysis Land (ERA5-Land).
[c]Copernicus Digital Elevation Model (GLO-90).
[d]NASA's Gridded Population of the World (GPWv4, Revision 4.11).
[e]Global Roads Inventory Project (GRIP).
[f]Global Land Cover and Land Use Change (GLCLUC).

3 | Theoretical Framework

Let $Y(\mathbf{u}, t)$ denote the underlying stochastic process that governs the square root of the mean daily O_3 concentration measured by in situ air quality monitoring stations at location $\mathbf{u} \in \mathbb{R}^2$ on day $t \in \mathbb{Z}$. Let $z(c_0(\mathbf{u}), t)$ represent the mean daily surface-level O_3 concentration obtained from satellite-based CAMS data on day t, corresponding to the satellite grid cell $c_0(\mathbf{u})$ that contains the location $\mathbf{u}$. In addition, for $j = 1, \ldots, p$, let $x_j(c_j(\mathbf{u}), t)$ denote the value of the jth environmental covariate (as listed in Table 1) obtained from satellite data for day t, corresponding to the grid cell $c_j(\mathbf{u})$ that contains location $\mathbf{u}$.

3.1 | Modeling Framework

The spatiotemporal process $Y(\mathbf{u}, t)$ is modeled as

$$Y(\mathbf{u}, t) = \mu(\mathbf{u}, t) + \alpha(\mathbf{u}, t)\, z\big(c_0(\mathbf{u}), t\big) + \epsilon(\mathbf{u}, t), \tag{1}$$

where $\mu(\mathbf{u}, t)$ is a varying intercept term capturing systematic discrepancies between satellite and in situ O_3 measurements. The coefficient $\alpha(\mathbf{u}, t)$ is a dynamic slope that allows the strength of the association between satellite-based and in situ O_3 measurements to change across locations and over time. For easier interpretability, we refer to $\mu(\mathbf{u}, t)$ as a "varying intercept" and $\alpha(\mathbf{u}, t)$ as a "dynamic slope", while in the literature these terms are more commonly known as "varying-coefficients" (Hastie and Tibshirani 1993) or "smooth terms" (Wood 2017). The term $\epsilon(\mathbf{u}, t)$ represents an unstructured random error, accounting for measurement noise and other unexplained variability. Here, $\epsilon(\mathbf{u}, t)$ is assumed to be an independent and identically distributed (i.i.d.) process following a Student's t-distribution with $\nu > 2$ degrees of freedom. This choice enables the model to accommodate potential outliers and heavy-tailed behavior in the data, thus improving robustness compared to conventional models assuming Gaussian errors.

Note that this formulation allows the model (1) to flexibly capture complex spatiotemporal dynamics in the association especially between O_3 measurements at air quality monitoring stations, $Y(\mathbf{u}, t)$, and satellite-derived estimates of surface-level CAMS O_3 concentration, $z(c_0(\mathbf{u}), t)$, via the dynamic coefficient $\alpha(\mathbf{u}, t)$, while also adjusting for potential additive biases through $\mu(\mathbf{u}, t)$. Similar modeling strategies employing spatiotemporally varying coefficients were adopted in previous studies investigating the association between in situ and satellite-based O_3 measurements. For instance, Paci et al. (2013) and Berrocal et al. (2014) used models with spatially and temporally varying intercepts and slopes to analyze daily 8-hour mean O_3 concentrations across 717 monitoring stations in the eastern United States over a two-week period (August 1–14, 2011). Das and Ghosal (2017) extended this approach using quantile regression with varying intercepts and slopes to study the daily maximum 8-hour O_3 concentrations

from 2006 to 2015, based on data from 1629 monitoring stations across the United States.

In the following, however, both the varying intercept $\mu(\mathbf{u}, t)$ and the dynamic slope coefficient $\alpha(\mathbf{u}, t)$ are explicitly related to environmental covariates, enhancing interpretability and enabling a more structured understanding of the sources of variability.

According to the procedure proposed in Li et al. (2007), De Iaco et al. (2016), Cappello et al. (2018) and Cappello, De Iaco, and Posa (2020), the separability test has been performed on 10 spatial pairs of monitoring stations at a distance from 225 km to 450 km and on 2 temporal lags (1, 2). At a 5% significance level, the assumption of separability has been rejected (test statistic = 365.699, p value = 2.57e − 65, df = 20) as well as for the covariates, which help explain the presence of spatiotemporal interaction, in addition to the evident seasonal variation in the year and the weekly variation in both the intercept and slope.

3.2 | Structure of Dynamic Coefficients

The null model (denoted M_0), defined by $\mu(\mathbf{u}, t) \equiv \beta_0$ and $\alpha(\mathbf{u}, t) \equiv \alpha_0$ corresponds to the assumption that satellite-based O_3 concentrations are simply noisy versions of the in situ measurements, after adjusting for a constant additive bias β_0 and a constant multiplicative bias α_0. A more flexible and widely used alternative is the land-use regression (LUR) model, which introduces a linear structure to model both the systematic discrepancy and the scaling of satellite estimates (Hoek et al. 2008). More specifically, the model (denoted by M_1) assumes

$$\mu(\mathbf{u}, t) = \beta_0 + \sum_{j=1}^{p} \beta_j \, x_j(c_j(\mathbf{u}), t), \quad \alpha(\mathbf{u}, t) \equiv \alpha_0, \tag{2}$$

where the coefficients $\beta_0, \ldots, \beta_p$ and α_0 are unknown parameters to be estimated. This formulation allows the model to adjust for systematic discrepancy (intercept) between satellite and in situ measurements using a linear combination of environmental variables, while maintaining a constant strength of association between the two measurement sources. To account for weekly variability in O_3 concentration, indicators for the days of the week are included among the covariates x_j's. Their corresponding coefficients β_j's capture potential weekly cycles such as the "weekend ozone effect", where O_3 levels may be higher on weekends despite lower precursor emissions (Schipa et al. 2009; Ma et al. 2020).

To allow for nonlinear effects of covariates on the intercept and slope, an alternative approach is to adopt additive models

$$\mu(\mathbf{u}, t) = \beta_0 + \sum_{j=1}^{p} f_j\big(x_j(c_j(\mathbf{u}), t)\big) + s(d(t)) + \tau_{w(t)} \tag{3}$$

$$\alpha(\mathbf{u}, t) = \alpha_0 + \sum_{j=1}^{p} \tilde{f}_j\big(x_j(c_j(\mathbf{u}), t)\big) + \tilde{s}(d(t)) + \tilde{\tau}_{w(t)}, \tag{4}$$

for the systematic bias component $\mu(\mathbf{u}, t)$ and the strength of association $\alpha(\mathbf{u}, t)$, respectively leading to the more flexible models M_2 and M_3. Here, $f_j, \tilde{f}_j : \mathbb{R} \to \mathbb{R}$ are smooth functions, typically represented using thin plate regression splines (see Wood 2017, section 4.1.5). This additive structure increases the model's flexibility and allows it to capture potentially nonlinear effects of covariates on the linear relation (1). Moreover, $d(t) \in \{1, \ldots, 366\}$ represents the day of the year, and $s, \tilde{s} : [1, 366] \to \mathbb{R}$ are penalized cyclic cubic regression splines (see Wood 2017, section 4.1.3), used to capture seasonal variation in the intercept $\mu(\mathbf{u}, t)$ and slope $\alpha(\mathbf{u}, t)$ of the linear relationship between in situ and satellite-derived O_3 concentrations. In addition to seasonal effects, Gaussian i.i.d. random effects $\tau_{w(t)}$ and $\tilde{\tau}_{w(t)}$, with mean zero and variances σ^2_τ and $\sigma^2_{\tilde{\tau}}$, are included, where the index $w(t)$ denotes the year−month combination (e.g., "2013−01") associated with calendar date t. These terms account for any remaining temporal variability at the monthly scale that is not explained by the other components of the model.

Although it is possible to include spatial random processes, for example, $\xi(\mathbf{u})$ and $\tilde{\xi}(\mathbf{u})$ with Matérn covariance structures, in the models for the varying intercept (3) and dynamic slope (4) to account for additional spatial correlation, our goodness-of-fit and cross-validation analyses (detailed in the following subsections) indicate that such terms are not necessary. Specifically, including these spatial random effects did not substantially improve model fit and, in some cases, reduced predictive performance (see cross-validation results in the Supporting Information file). This may be attributed to the fact that the major sources of spatial variation are already well captured by the model structure and the suite of environmental covariates, leaving little residual spatial dependence to warrant the added model complexity.

The specifications given in Equations (3) and (4) make the proposed model in Equation (1) a special case of generalized additive models (GAMs), which are well-suited for modeling complex nonlinear relationships and can be efficiently applied to large datasets (Wood et al. 2017; Li and Wood 2020). The model is implemented using the `mgcv` package in `R`, which employs a restricted maximum likelihood approach for parameter estimation (Marra and Wood 2011). In this framework, penalty terms are applied to the smoothing functions $f_j(\cdot)$ and $\tilde{f}_j(\cdot)$ to control model complexity (Wood 2017). The smoothing parameters, along with other model coefficients, are estimated by maximizing the restricted likelihood, ensuring a balance between goodness-of-fit and smoothness (Kim and Gu 2004).

4 | Model Evaluation

In this section, the proposed model (1) with the spatially and temporally varying intercept (3) and dynamic slope (4), is evaluated with regard to its adequacy and predictive performance. Specifically, the model is assessed to see how well it reproduces the observed data and how effectively it generalizes to new data not used in model fitting. To examine the necessity of allowing non-constant and nonlinear forms for the intercept (3) and slope (4), the full model M_3,

$$\begin{aligned} Y(\mathbf{u}, t) = {} & [\beta_0 + \sum_{j=1}^{p} f_j\big(x_j(c_j(\mathbf{u}), t)\big) + s(d(t)) + \tau_{w(t)}] \\ & + [\alpha_0 + \sum_{j=1}^{p} \tilde{f}_j\big(x_j(c_j(\mathbf{u}), t)\big) + \sim s(d(t)) + \tilde{\tau}_{w(t)}] \\ & \cdot z\big(c_0(\mathbf{u}), t\big) + \epsilon(\mathbf{u}, t), \end{aligned}$$

is compared against three simpler alternatives:

- M_0: a baseline model with a constant intercept $\mu(\mathbf{u}, t) = \beta_0$ and constant slope $\alpha(\mathbf{u}, t) \equiv \alpha_0$, that is,

$$Y(\mathbf{u}, t) = \beta_0 + \alpha_0 \; z\big(c_0(\mathbf{u}), t\big) + \epsilon(\mathbf{u}, t),$$

- M_1: a model with a linearly varying intercept based on environmental covariates, and a constant slope, as defined in Equation (2), that is,

$$Y(\mathbf{u}, t) = \beta_0 + \sum_{j=1}^{p} \beta_j \; x_j(c_j(\mathbf{u}), t) + \alpha_0 \; z\big(c_0(\mathbf{u}), t\big) + \epsilon(\mathbf{u}, t),$$

- M_2: a model with an additive nonlinear intercept $\mu(\mathbf{u}, t)$ that includes seasonal and weekly components, as in Equation (3) and a constant slope $\alpha(\mathbf{u}, t) \equiv \alpha_0$, that is,

$$Y(\mathbf{u}, t) = \left[\beta_0 + \sum_{j=1}^{p} f_j\big(x_j(c_j(\mathbf{u}), t)\big) + s(d(t)) + \tau_{w(t)}\right] + \alpha_0 \; z\big(c_0(\mathbf{u}), t\big) + \epsilon(\mathbf{u}, t).$$

This comparative analysis is intended to evaluate whether the increased flexibility in M_3 improves both the model's ability to fit the data and its predictive performance, particularly in terms of downscaling at unmonitored locations and forecasting ozone levels at future time points.

4.1 | Goodness of Fit

Table 2 presents adjusted coefficient of determination (R^2_{adj}), the percentage of null deviance explained, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the full model (M_3) and its simpler nested alternatives (M_0, M_1, and M_2). From the baseline model M_0 to the full model M_3, a consistent improvement can be seen in all these four goodness of fit measures: both R^2_{adj} and the percentage of explained deviation increase, while AIC and BIC decrease, indicating a better model fit. To assess whether these improvements are statistically significant, for $k = 1, 2, 3$, likelihood ratio tests are performed to compare each model M_k with its nested simpler submodel M_{k-1}. The resulting χ^2 statistics and associated p values are also reported in Table 2. In all cases, the p values are small, indicating that the additional complexity of M_k relative to M_{k-1} accounts for a significant amount of variation in the response variable. These results provide strong evidence that M_3 offers a significantly better fit than simpler nested models M_0–M_2.

The full model M_3 explains 72.2% of the null deviance, with an adjusted R^2 of 0.865. This indicates that approximately 13% of the variation in the square-root-transformed in situ O_3 measurements remains unexplained by the model. Figure 3 presents diagnostic plots, including fitted versus observed values, residuals versus fitted values, and a histogram of the residuals. While there is some scope for improvement, the histogram of residuals closely matches the probability density function of a Student's t-distribution with estimated degrees of freedom $\hat{\nu} = 3.57$, suggesting an adequate fit to the assumed error distribution. Although a few outliers are present, the fitted values generally align well with the observed values. Moreover, the residuals show no clear structure or trends when plotted against the fitted values, supporting the assumption of homoscedasticity and indicating that the model successfully captures the main patterns in the data.

The proposed structure in Equation (1) provides a clear and interpretable modeling framework for assessing how each environmental variable contributes to the additive bias, represented by the varying intercept term $\mu(\mathbf{u}, t)$, and the multiplicative bias, captured by the dynamic slope term $\alpha(\mathbf{u}, t)$, in surface-level satellite-derived O_3 measurements relative to in situ observations. To illustrate this, Figure 4 displays the estimated effects of nitrogen dioxide concentration, wind speed, elevation, and seasonal variation associated with the day of the year on both $\mu(\mathbf{u}, t)$ and $\alpha(\mathbf{u}, t)$. These are partial effects, meaning that all other covariates are held constant at their respective mean values for continuous variables, and at their reference category for categorical variables, including land use and land cover, wind direction, and day of the week.

TABLE 2 | Adjusted coefficient of determination (R^2_{adj}), percentage of the null deviance explained by the model (%Dev), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) for comparing the goodness of four models. Model M_0 is the baseline, with a constant intercept and slope. Models M_1, M_2, and M_3 incorporate varying intercepts and slopes informed by environmental variables, as well as seasonal and weekly patterns: linear varying intercept and constant slope (M_1), additive nonlinear varying intercept with seasonal and weekly components and a constant slope (M_2) and both intercept and slope are modeled as additive nonlinear functions with seasonal and weekly components (M_3). The test statistic (χ^2) and p value of likelihood ratio tests, used to compare each model M_k, $k = 1, 2, 3$, against its nested simpler counterpart M_{k-1} are also provided. The last column reports the computational time (in seconds), measured on a personal computer equipped with an Intel$^{\text{R}}$ Core$^{\text{TM}}$ Ultra 5 135U processor (1600 MHz, 12 cores, 14 threads) and 16GB of RAM.

	$\mu(\mathbf{u}, t)$	$\alpha(\mathbf{u}, t)$	R^2_{adj}	%Dev	AIC	BIC	χ^2	p value	Time
M_0	β_0	α_0	0.830	68.6	2,918,694	2,918,754	—	—	38.0
M_1	$\beta_0 + \sum_{j=1}^{p} \beta_j x_j(\mathbf{u}, t)$	α_0	0.851	70.6	2,754,652	2,755,048	164,098	< 0.001	129.5
M_2	$\beta_0 + \sum_{j=1}^{p} f_j\big(x_j(\mathbf{u}, t)\big) + s(d(t)) + \tau_{w(t)}$	α_0	0.860	71.6	2,684,639	2,687,116	70,360	< 0.001	74.7
M_3	$\beta_0 + \sum_{j=1}^{p} f_j\big(x_j(\mathbf{u}, t)\big) + s(d(t)) + \tau_{w(t)}$	$\tilde{\beta}_0 + \sum_{j=1}^{p} \tilde{f}_j\big(x_j(\mathbf{u}, t)\big) + \tilde{s}(d(t)) + \tilde{\tau}_{w(t)}$	**0.865**	**72.2**	**2,639,100**	**2,642,248**	45,650	< 0.001	184.7

Note: Bold Indicates the best model.

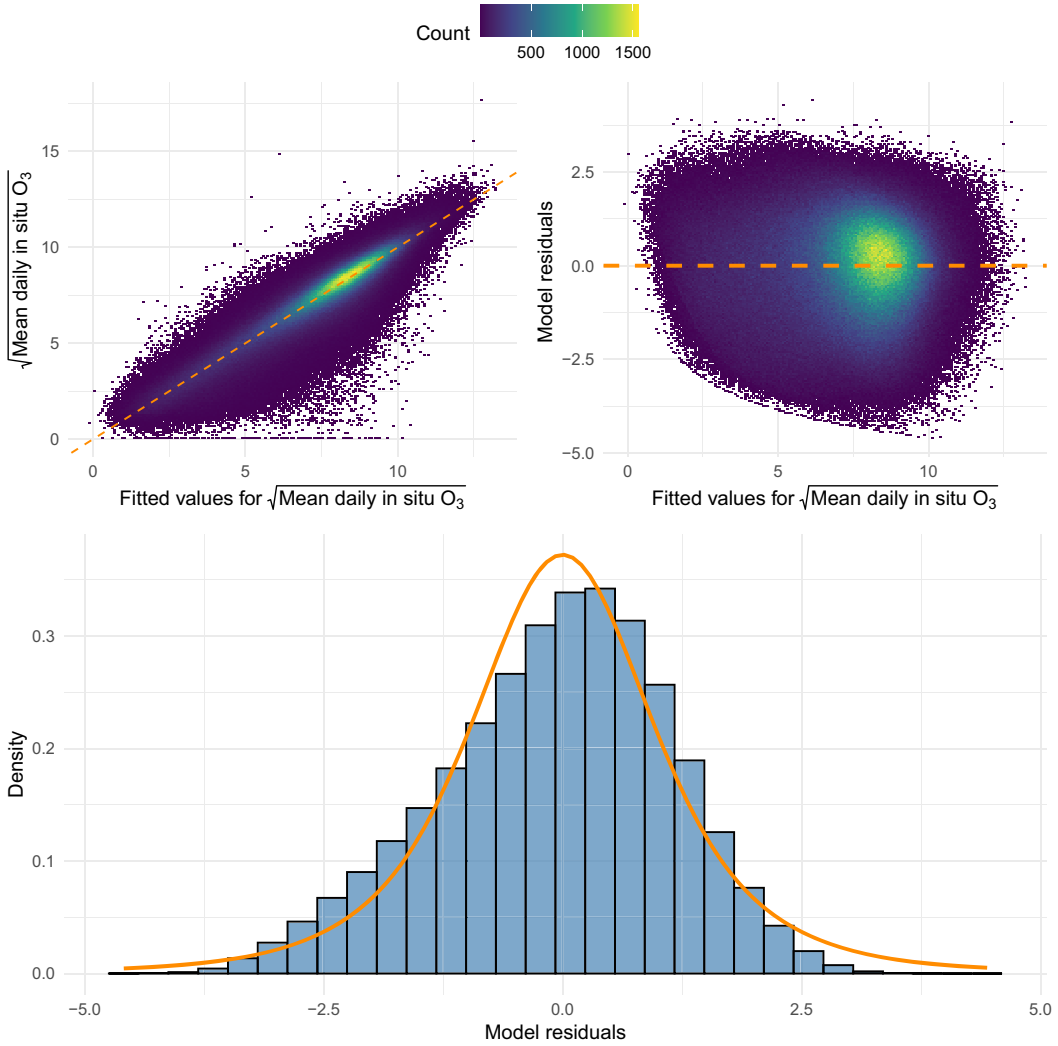

FIGURE 3 | Diagnostic plots for the fitted model. Top left: observed square root of ozone concentration versus fitted values. Top right: model residuals plotted against fitted values. Bottom: histogram of the residuals overlaid with the probability density function of a Student's t-distribution using the estimated degrees of freedom $\hat{\nu}$.

It can be seen from Figure 4 that both nitrogen dioxide and elevation have a substantial positive effect on the additive bias $\mu(\mathbf{u}, t)$, while windspeed, when sufficiently strong, exerts a significant negative effect. In contrast, their influence on the multiplicative bias $\alpha(\mathbf{u}, t)$ is less pronounced, slightly in the opposite direction, and due to its multiplicative nature show high variability, as reflected by the wide confidence bands. Seasonal variations associated with the day of the year affect both $\mu(\mathbf{u}, t)$ and $\alpha(\mathbf{u}, t)$, with patterns that appear to be reciprocal. These differences reflect the distinct roles of the two bias components: $\mu(\mathbf{u}, t)$ represents systematic shifts in baseline O_3 concentrations, whereas $\alpha(\mathbf{u}, t)$ adjusts the scaling of satellite measurements relative to in situ observations. For example, increases in variables such as nitrogen dioxide or elevation tend to raise baseline O_3 levels (reflected in μ) while slightly reducing the scaling factor α, indicating that satellite measurements are proportionally smaller compared to in situ observations. Elevated NO_2 is frequently co-located with increased aerosol loads and complex surface albedo (e.g., urban surfaces or water bodies), which can alter top-of-atmosphere radiance and introduce systematic offsets and reduced contrast in the retrieved O_3 field, providing a plausible explanation for the observed additive and multiplicative biases (Geddes et al. 2012). Similarly, higher wind speeds are associated with lower baseline O_3 concentrations and may also modify the scaling between satellite and in situ measurements. Overall, the positive effects of nitrogen dioxide and elevation, the negative influence of wind speed and the cyclic impact of seasonal variations on baseline ozone (μ) are consistent with previous findings (e.g., Zhu et al. 2023), whereas their effects on the scaling term (α) provide additional insight into divergences between satellite and in situ measurements.

4.2 | Cross-Validation

To evaluate the predictive performance of the proposed model in general and across both spatial and temporal dimensions, the following cross-validation framework is adopted that systematically partitions the full dataset into k mutually exclusive subsets. As recommended in the literature, for each spatial and temporal

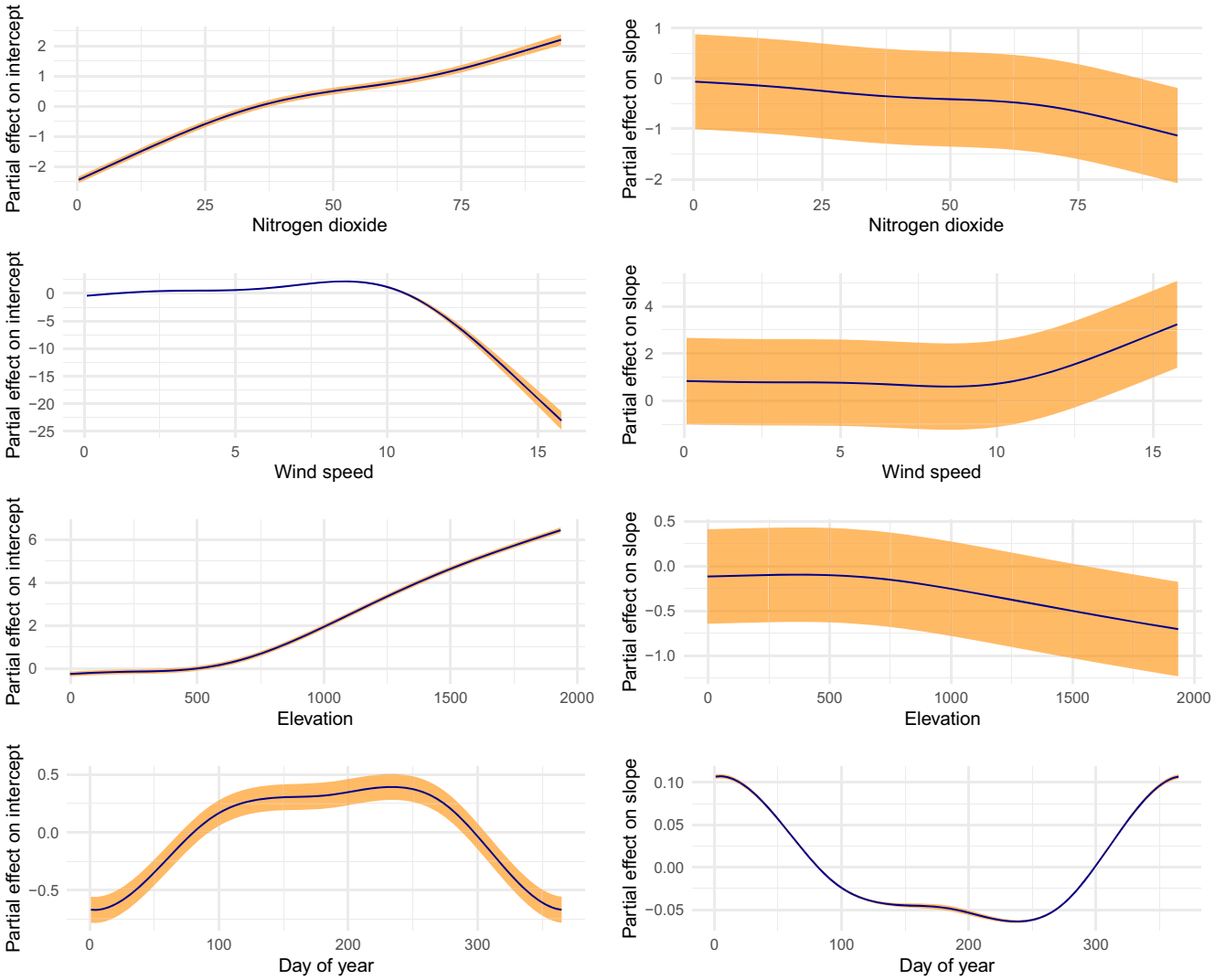

FIGURE 4 | Partial effects of elevation, nitrogen dioxide, and day of the year on the varying intercept and dynamic slope in the linear model describing the relationship between in situ O_3 measurements and corresponding satellite-derived surface O_3 estimates (note that each plot is displayed with a distinct vertical axis scale for easier interpretability).

cluster a buffer zone has been defined in order to ensure that the training and testing data are approximately independent, minimizing the influence of spatial and temporal autocorrelation on model evaluation (Wenger and Olden 2012; Roberts et al. 2017; Meyer et al. 2018; Otto et al. 2024).

1. The 402 air quality monitoring stations (Figure 1) are clustered into $k_1 = 10$ spatial groups, $\mathcal{C}_1, \ldots, \mathcal{C}_{k_1}$, using the k_1-medoid algorithm (see Kaufman and Rousseeuw 1990, Chapter 2) based on their spatial proximity. A buffer zone around each cluster $\mathcal{C}_j$ is defined as the union of 50-km radius circular buffers centered on every station belonging to $\mathcal{C}_j$. These clusters and their corresponding buffer zones are shown in the top panel of Figure 5. Then, for $j = 1, \ldots, k = k_1$, all observations from stations in cluster $\mathcal{C}_j$ over the entire 4016-day study period are set aside as the testing set $\mathcal{P}_j$, while all observations from stations outside the buffer zone of cluster $\mathcal{C}_j$ over the same period are used to estimate the model parameters. The fitted model is then used to predict O_3 concentrations for the held-out subset $\mathcal{P}_j$. This spatial cross-validation evaluates the model's ability to predict O_3 concentrations at locations not used during model training and assesses the model's spatial generalization ability.

2. The full study period $T = [t_{\min}, t_{\max}]$, spanning 4016 days, is divided into $k_2 = 8$ overlapping equal-length time intervals $\mathcal{I}_1, \ldots, \mathcal{I}_{k_2}$, each lasting four years, as shown in the bottom panel of Figure 5. Each interval $\mathcal{I}_j$ is further partitioned into a three-year training period (to capture at least three seasonal cycles), followed by a six-month buffer zone and then a six-month testing period. Then, for $j = 1, \ldots, k = k_2$, observations from all 402 stations with dates within the six-month testing period of $\mathcal{I}_j$ are set aside as the testing set $\mathcal{P}_j$, while observations from all 402 stations during the corresponding 3-year training period of $\mathcal{I}_j$ are used to estimate the model parameters. The fitted model is then used to predict O_3 concentrations for the held-out subset $\mathcal{P}_j$. This temporal cross-validation mimics realistic forecasting by ensuring that only past data are used in model training,

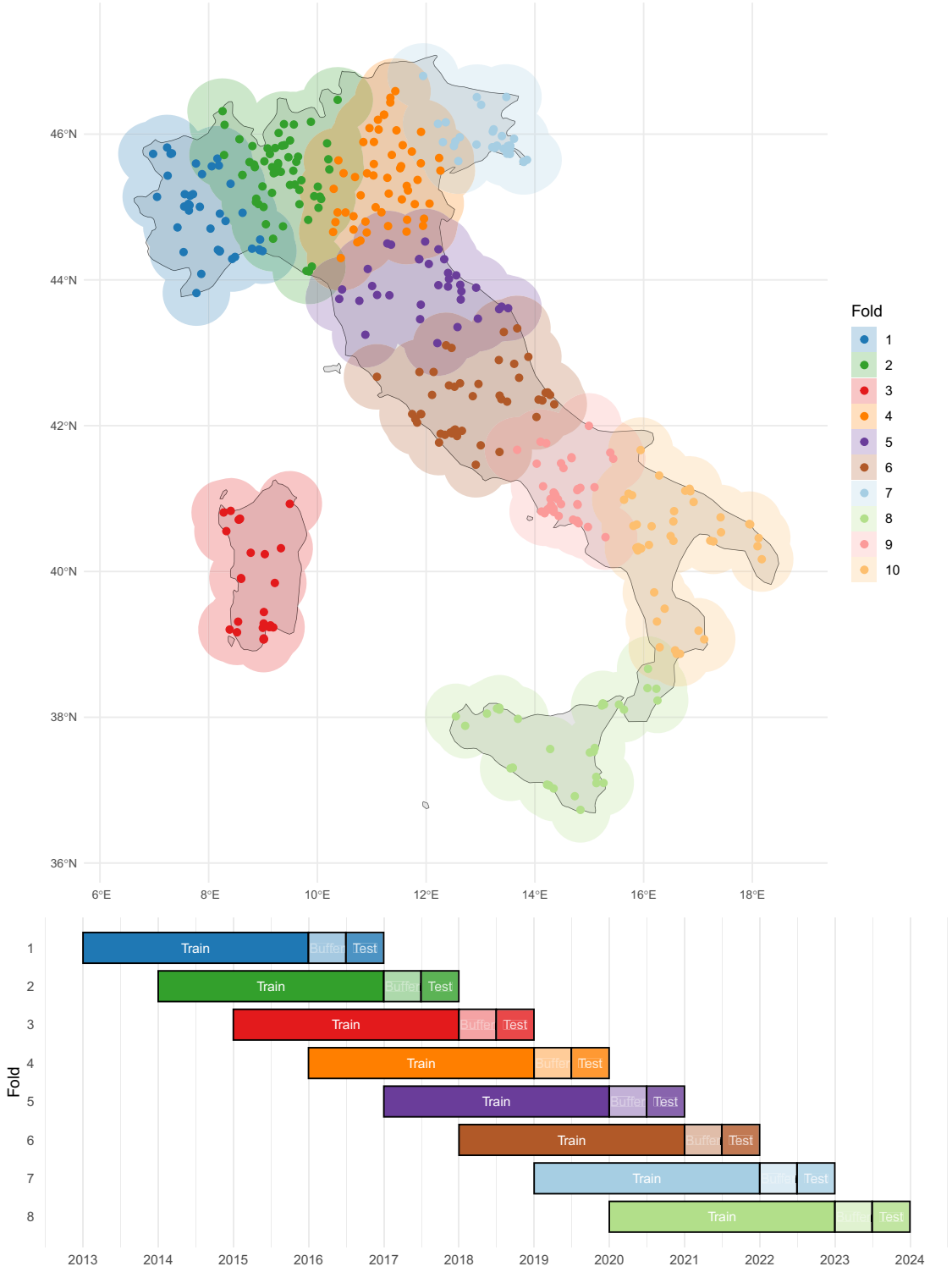

FIGURE 5 | Spatial and temporal partitioning for cross-validation. The top panel shows the grouping of 402 air quality monitoring stations into $k_1 = 10$ spatial clusters using the k-medoid clustering algorithm. The bottom panel illustrates the division of the full time period (2013-01-01 to 2023-12-30) into $k_2 = 8$ temporal folds, each spanning four years. For spatial cross-validation, each cluster is treated as a test set and predicted using data from the remaining clusters, after excluding stations located within buffer zones. For temporal cross-validation, each interval is partitioned into a 3-year training period, followed by a 6-month buffer, and then a 6-month test period.

thereby avoiding temporal information leakage and allowing a robust evaluation of temporal prediction accuracy.

3. Each fold $j = 1, \ldots, k = k_1 \times k_2$ corresponds to a unique pair of spatial cluster $\mathcal{C}_{j_1}, j_1 = 1, \ldots, k_1$, and temporal interval $\mathcal{I}_{j_2}, j_2 = 1, \ldots, k_2$, where the test set $\mathcal{P}_j$ consists of all observations from stations in $\mathcal{C}_{j_1}$ during the 6-month test period of $\mathcal{I}_{j_2}$. The model parameters are estimated using only observations from stations lying outside the buffer zone of $\mathcal{C}_{j_1}$ and recorded during the corresponding 3-year training period of $\mathcal{I}_{j_2}$ and then fitted model is used to predict O_3 concentrations for the held-out subset $\mathcal{P}_j$. This spatiotemporal cross-validation evaluates the model's ability to predict O_3 concentrations at both locations and times not used during model training and assesses the model's spatial and temporal generalization ability.

Once the parameters of model in Equation (1) have been estimated, the square root of the O_3 concentration at location $\mathbf{u}$ and time t is predicted with

$$\widehat{Y}(\mathbf{u}, t) = \widehat{\mu}(\mathbf{u}, t) + \widehat{\alpha}(\mathbf{u}, t) z(\mathbf{u}, t),$$

where $\widehat{\mu}(\mathbf{u}, t)$ and $\widehat{\alpha}(\mathbf{u}, t)$ are the estimated dynamic coefficients. To obtain a prediction for the O_3 concentration (i.e., transforming the response to the original scale), the identity

$$\mathbb{E}\left[(Y(\mathbf{u}, t))^2\right] = (\mu(\mathbf{u}, t) + \alpha(\mathbf{u}, t) z(\mathbf{u}, t))^2 + \mathbb{V}\mathrm{ar}[Y(\mathbf{u}, t)]$$

is used, which leads to the prediction

$$\widehat{Y^2}(\mathbf{u}, t) = (\widehat{\mu}(\mathbf{u}, t) + \widehat{\alpha}(\mathbf{u}, t) z(\mathbf{u}, t))^2 + \frac{\widehat{\sigma}^2 \widehat{\nu}}{\widehat{\nu} - 2},$$

where $\widehat{\sigma}$ and $\widehat{\nu}$ are the estimated scale and degrees of freedom parameters from the assumed Student t-distribution of the error term.

Then, for each of the above cross-validation strategies, the model's predictive performance is measured using standard prediction accuracy metrics (Wang et al. 2022): the Root Mean Square Prediction Error (RMSPE), the Mean Absolute Prediction Error (MAPE), and cross-validated coefficient of determination (CVR^2), defined as

$$\mathrm{RMSPE} = \frac{1}{k} \sum_{j=1}^{k} \sqrt{\frac{1}{\#(\mathcal{P}_j)} \sum_{(\mathbf{u},t) \in \mathcal{P}_j} \left(Y^2(\mathbf{u}, t) - \hat{Y}^2(\mathbf{u}, t) \right)^2},$$

$$\mathrm{MAPE} = \frac{1}{k} \sum_{j=1}^{k} \left(\frac{1}{\#(\mathcal{P}_j)} \sum_{(\mathbf{u},t) \in \mathcal{P}_j} \left| Y^2(\mathbf{u}, t) - \hat{Y}^2(\mathbf{u}, t) \right| \right),$$

$$\mathrm{CVR}^2 = \frac{1}{k} \sum_{j=1}^{k} \left(1 - \frac{\sum_{(\mathbf{u},t) \in \mathcal{P}_j} (Y(\mathbf{u}, t) - \hat{Y}(\mathbf{u}, t))^2}{\sum_{(\mathbf{u},t) \in \mathcal{P}_j} \left(Y(\mathbf{u}, t) - \frac{1}{\#(\mathcal{P}_j)} \sum_{(\mathbf{u},t) \in \mathcal{P}_j} Y(\mathbf{u}, t) \right)^2} \right).$$

Here, $Y^2(\mathbf{u}, t)$ denotes the observed O_3 concentration (i.e., the square of the response) at location $\mathbf{u}$ and on day t, and $\widehat{Y^2}(\mathbf{u}, t)$ is the corresponding model prediction. The term $\#(\mathcal{P}_j)$ represents the number of observations in subset $\mathcal{P}_j$. The RMSPE emphasizes larger errors due to the squaring, while MAPE provides a more interpretable measure of average absolute prediction error. Together, these metrics, which are widely recognized indices of model performance (Hodson 2022), offer a comprehensive evaluation of the model's prediction accuracy across all folds.

To evaluate the contribution of the varying intercept (3) and the dynamic slope (4) to predictive performance, the full model M_3 is compared with the three simpler nested models M_0–M_2 in predicting O_3 concentration $Y^2(\mathbf{u}, t)$. Table 3 summarizes the predictive performance of the models using RMSPE, MAPE, and CVR^2 for spatial, temporal, and spatiotemporal cross-validation frameworks described above. The full model, M_3, consistently outperforms the null model M_0 and the intermediate models M_1 and M_2 in terms of RMSPE and CVR^2 across all three cross-validation frameworks: spatial, temporal, and spatiotemporal. In both the spatial and temporal frameworks, the progressive reduction in predictive error from M_0 to M_1, to M_2, and finally to M_3 confirms that each added level of complexity contributes meaningful predictive power. Only in the more challenging spatiotemporal framework with a relatively small training dataset, the linear model M_1 exhibits a slightly lower MAPE than M_3 and a smaller RMSPE than M_2. Nevertheless, the full model M_3 achieves the best overall predictive performance with an RMSPE of 10.9 $\mu g/m^3$, a MAPE of 5.98 $\mu g/m^2$, and a CVR^2 exceeding 0.8, which demonstrates robust generalization to both unseen locations and future time periods.

After confirming that the full model M_3 adequately explains the variability in the data—based on goodness-of-fit metrics—and does not exhibit signs of overfitting, as evidenced by its cross-validation performance, it can be reliably used to predict in situ O_3 concentrations at any location and time where the environmental variables listed in Table 1 are available. Note also that after inspecting the residual correlation, there is no spatial pattern left in the data. As an illustrative example, Figure 6 presents predicted in situ O_3 concentrations at a spatial resolution of $0.1° \times 0.1°$ across Italy and surrounding regions on twelve selected dates during the study period. The displayed maps well reproduce the spatiotemporal variability of the investigated pollutant over Italian territories, with ozone concentrations often higher in the pre-Alpine (in the Northern part of Italy) and Mediterranean areas, with respect to the other parts of the Italian peninsula. In particular, in Northern Italy, the elevation factor and the consequent high values of UV radiation play a crucial role. Indeed, very large O_3 concentrations are recorded/predicted over the pre-Alpine area. The seasonal pattern for O_3 predicted values (higher levels during warmer periods than those recorded during colder periods) is also evident from the different maps, where the highest levels have been predicted in May, June, July, and August, due to the increasing air temperature and solar radiation.

5 | Conclusions

This study proposed and evaluated a flexible modeling framework for predicting surface-level O_3 concentrations by integrating satellite-based data with spatially and temporally varying intercepts and dynamic slope coefficients. The model extends the generalized additive model (GAM) framework by incorporating a structure that allows both the intercept and slope

TABLE 3 | Comparison of spatial, temporal and spatiotemporal predictive performance for four models predicting O_3 concentrations: baseline model M_0 with constant intercept and slope, model M_1 with linear models for intercept and constant slope, model M_2 with additive nonlinear model for intercept and constant slope and the full model M_3 with additive nonlinear models for both intercept and slope. Performance metrics include root mean squared prediction error (RMSPE), mean absolute percentage error (MAPE), and cross-validated R^2 (CVR^2).

Model	Spatial CV			Temporal CV			Spatiotemporal CV		
	RMSPE	MAPE	CVR^2	RMSPE	MAPE	CVR^2	RMSPE	MAPE	CVR^2
M_0	11.784	6.260	0.775	11.710	5.986	0.832	11.558	6.172	0.785
M_1	11.101	6.019	0.797	10.910	5.646	0.854	10.941	**5.931**	0.806
M_2	11.002	6.057	0.802	10.721	5.610	0.859	10.953	6.023	0.808
M_3	**10.932**	**5.985**	**0.802**	**10.604**	**5.511**	**0.862**	**10.897**	5.983	**0.807**

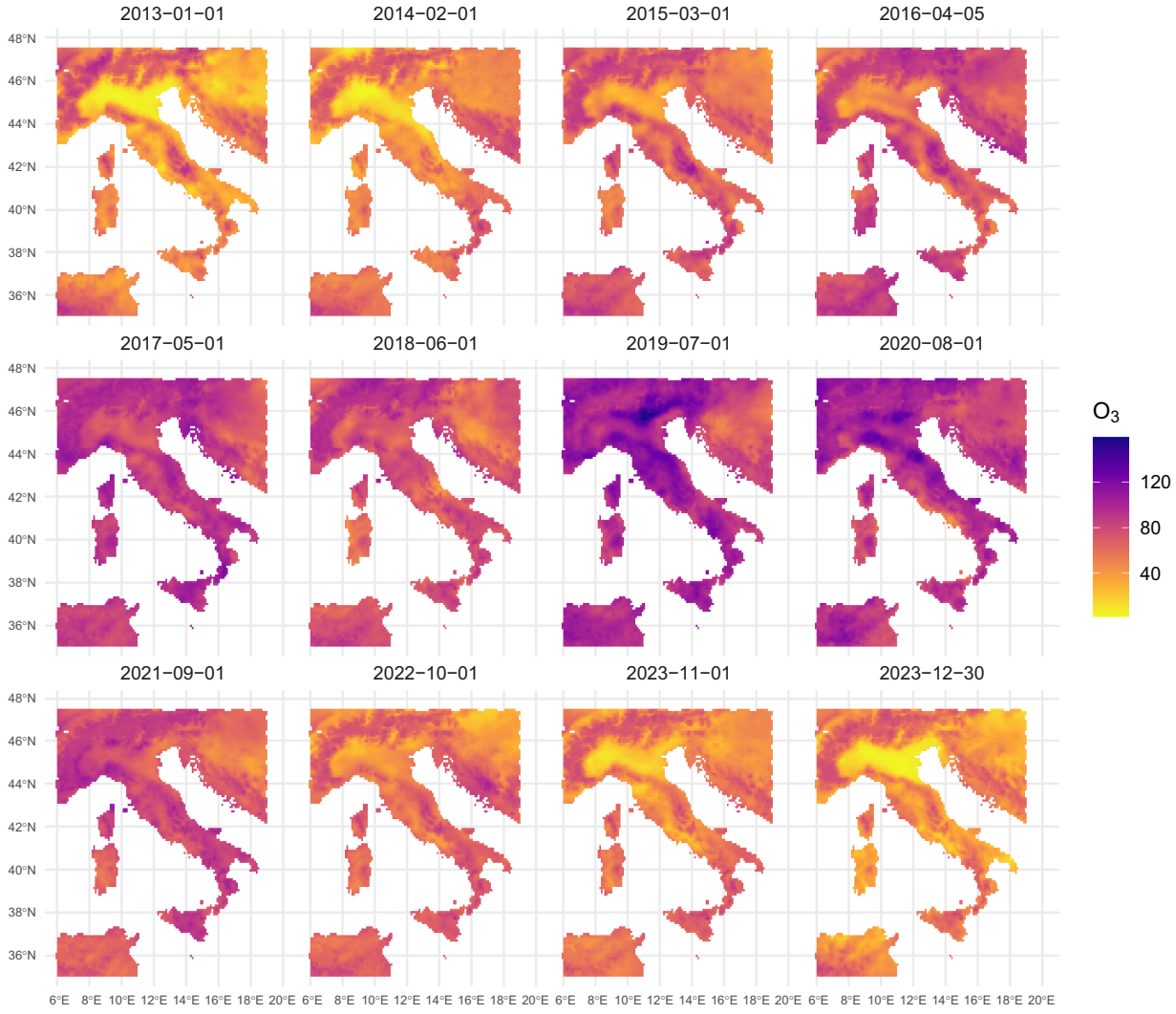

FIGURE 6 | Predicted in situ O_3 concentrations at a spatial resolution of 0.1° × 0.1° across Italy and its surrounding regions on twelve selected dates during the study period.

to vary across space and time, thereby capturing local deviations between satellite-derived estimates and in situ observations. While GAMs and varying coefficient models have each been employed independently in previous studies to model pollutant concentrations and to relate satellite observations to in situ measurements, the novelty of this work lies in combining these approaches within a unified and interpretable framework. This enables a better understanding of the sources of additive biases (i.e., systematic shifts) and multiplicative biases (i.e., scaling differences) in satellite-based measurements of surface-level O_3 relative to in situ observations.

The proposed model is a special case of a GAM and, consequently, parameter estimation is already efficiently implemented via the

mgcv package in R. Cross-validation and goodness-of-fit diagnostics demonstrated that incorporating varying intercepts and dynamic slopes substantially improves model fit and predictive performance compared to simpler models. In particular, the full model with additive specifications and seasonal components for these coefficients yielded the lowest prediction errors and highest explained variance. Residual diagnostics further supported the model's adequacy, indicating reasonable agreement between observed and fitted values and no significant departures from distributional assumptions.

While seasonal components in both the varying intercept and dynamic slope models contributed to improved model performance, incorporating spatial random processes did not lead to substantial gains in explanatory power or predictive accuracy for the available data. This likely reflects that the included environmental covariates and the model's existing structure already capture the primary sources of spatial variation. To demonstrate model interpretability, the framework provided a clear explanation of the effects of elevation, nitrogen dioxide, and the seasonal variation associated with the day of the year on both additive and multiplicative biases in surface-level satellite measurements of O_3 compared with in situ observations. The model was applied to provide downscaling by predicting O_3 concentrations on a fine spatial grid covering Italy and its surrounding areas for a selected set of dates.

Overall, the findings underscore the value of integrating dynamic model components in air pollution modeling, especially for aligning satellite-based proxies with ground-level observations, enabling better interpretability and explanation of the relationship between satellite measurements and in situ observations. The proposed approach is not only accurate but also scalable, making it suitable for large datasets and operational forecasting. Future work could explore multivariate extensions, incorporate additional covariates, assess performance across other pollutants and regions, and develop joint models for multiple pollutants.

Acknowledgments

The authors thank the Editors and the referees for their constructive suggestions during the reviewing process. Open access publishing facilitated by Universita del Salento, as part of the Wiley - CRUI-CARE agreement.

Funding

This study was supported by the European Union–NextGenerationEU with the Cascade Open Calls published by ALMA MATER STUDIORUM – University of Bologna, inside the Project GRINS funded by PNRR – Mission 4, Component 2, Investment 1.3 "Partnership extended to Universities, Research Centers, Firms and research projects funding", D.D. 341 of 15/03/2022, "ECoST-DATA, Exploring Spatio-Temporal Environmental Conditions: Harmonized Databases and Analytical Techniques", CUP: J33C22002910001.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Berrocal, V., A. Gelfand, and D. Holland. 2014. "Assessing Exceedance of Ozone Standards: A Space-Time Downscaler for Fourth Highest Ozone Concentrations." *Environmetrics* 25, no. 4: 279–291.

Bhatti, U. A., Y. Yan, M. Zhou, et al. 2021. "Time Series Analysis and Forecasting of Air Pollution Particulate Matter (Pm 2.5): An SARIMA and Factor Analysis Approach." *IEEE Access* 9: 41019–41031.

Box, G. E., and D. R. Cox. 1964. "An Analysis of Transformations." *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 26, no. 2: 211–243.

Cappello, C., S. De Iaco, S. Maggio, and D. Posa. 2020. "Modeling Ocean Currents Through Complex Random Fields Indexed in Time." *Mathematical Geosciences* 53, no. 5: 999–1025. https://doi.org/10.1007/s11004-020-09880-3.

Cappello, C., S. De Iaco, and D. Posa. 2018. "Testing the Type of Non-Separability and Some Classes of Space-Time Covariance Function Models." *Stochastic Environmental Research and Risk Assessment* 32, no. 1: 17–35.

Cappello, C., S. De Iaco, and D. Posa. 2020. "Covatest: An R Package for Selecting a Class of Space-Time Covariance Functions." *Journal of Statistical Software* 94: 1–42.

Castelli, M., F. M. Clemente, A. Popovic, S. Silva, and L. Vanneschi. 2020. "A Machine Learning Approach to Predict Air Quality in California." *Complexity* 2020: 1–23.

Cedeno Jimenez, J. R., A. J. Pugliese Viloria, and M. A. Brovelli. 2023. "Estimating Daily NO_2 Ground Level Concentrations Using Sentinel-5P and Ground Sensor Meteorological Measurements." *ISPRS International Journal of Geo-Information* 12, no. 3: 107.

Center For International Earth Science Information Network, Columbia University. 2018. *Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11*. Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H49C6VHW.

Copernicus Atmosphere Monitoring Service. 2021. *CAMS European air Quality Reanalyses*. Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store. https://doi.org/10.24381/7cc0465a.

Das, P., and S. Ghosal. 2017. "Analyzing Ozone Concentration by Bayesian Spatio-Temporal Quantile Regression." *Environmetrics* 28, no. 4: e2443.

De Iaco, S. 2017. "The Cgeostat Software for Analyzing Complex-Valued Random Fields." *Journal of Statistical Software* 79, no. 5. https://doi.org/10.18637/jss.v079.i05.

De Iaco, S. 2022. "New Spatio-Temporal Complex Covariance Functions for Vectorial Data Through Positive Mixtures." *Stochastic Environmental Research and Risk Assessment* 36, no. 9: 2769–2787. https://doi.org/10.1007/s00477-022-02171-9.

De Iaco, S. 2023a. "Spatio-Temporal Generalized Complex Covariance Models Based on Convolution." *Computational Statistics & Data Analysis* 183: 107709. https://doi.org/10.1016/j.csda.2023.107709.

De Iaco, S. 2023b. "Families of Complex-Valued Covariance Models Through Integration." *Environmetrics* 34, no. 3. https://doi.org/10.1002/env.2779 Portico.

De Iaco, S., D. E. Myers, and D. Posa. 2003. "The Linear Coregionalization Model and the Product–Sum Space–Time Variogram." *Mathematical Geology* 35, no. 1: 25–38. https://doi.org/10.1023/a:1022425111459.

De Iaco, S., M. Palma, and D. Posa. 2005. "Modeling and Prediction of Multivariate Space–Time Random Fields." *Computational Statistics & Data Analysis* 48, no. 3: 525–547. https://doi.org/10.1016/j.csda.2004.02.011.

De Iaco, S., M. Palma, and D. Posa. 2016. "A General Procedure for Selecting a Class of Fully Symmetric Space-Time Covariance Functions." *Environmetrics* 27, no. 4: 212–224.

Dou, Y., N. D. Le, and J. V. Zidek. 2010. "Modeling Hourly Ozone Concentration Fields." *Annals of Applied Statistics* 4, no. 3: 1183–1213.

Eastoe, E. F. 2009. "A Hierarchical Model for Non-Stationary Multivariate Extremes: A Case Study of Surface-Level Ozone and NOx Data in the UK." *Environmetrics: The Official Journal of the International Environmetrics Society* 20, no. 4: 428–444.

Emery, C. A., K. R. Baker, G. M. Wilson, and G. Yarwood. 2024. "Comprehensive Air Quality Model With Extensions, v7. 20: Formulation and Evaluation for Ozone and Particulate Matter Over the US." *Geoscientific Model Development Discussions* 2024: 1–48.

Ettinger, B., S. Guillas, and M.-J. Lai. 2012. "Bivariate Splines for Ozone Concentration Forecasting." *Environmetrics* 23, no. 4: 317–328.

European Environment Agency. 2025. "European air Quality Portal." https://www.eea.europa.eu/themes/air/air-quality-index.

European Space Agency. 2025. "Copernicus Digital Elevation Model." https://doi.org/10.5270/ESA-c5d3d65.

Fassó, A., J. Rodeschini, A. Fusta Moro, et al. 2023. "Agrimonia: A Dataset on Livestock, Meteorology and Air Quality in the Lombardy Region, Italy." *Scientific Data* 10: 143.

Geddes, J. A., J. G. Murphy, J. M. O'Brien, and E. A. Celarier. 2012. "Biases in Long-Term NO_2 Averages Inferred From Satellite Observations due to Cloud Selection Criteria." *Remote Sensing of Environment* 124: 210–216.

Gelfand, A. E., M. Fuentes, J. A. Hoeting, and R. L. Smith. 2019. *Handbook of Environmental and Ecological Statistics*. CRC Press.

Guo, B., D. Zhang, L. Pei, et al. 2021. "Estimating $PM_{2.5}$ Concentrations via Random Forest Method Using Satellite, Auxiliary, and Ground-Level Station Dataset at Multiple Temporal Scales Across China in 2017." *Science of the Total Environment* 778: 146288.

Hastie, T., and R. Tibshirani. 1993. "Varying-Coefficient Models." *Journal of the Royal Statistical Society: Series B: Methodological* 55, no. 4: 757–796.

Hayn, M., S. Beirle, F. A. Hamprecht, U. Platt, B. H. Menze, and T. Wagner. 2009. "Analysing Spatio-Temporal Patterns of the Global NO2-Distribution Retrieved From GOME Satellite Observations Using a Generalized Additive Model." *Atmospheric Chemistry and Physics* 9, no. 17: 6459–6477.

Hodson, T. O. 2022. "Root-Mean-Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not." *Geoscientific Model Development* 15, no. 14: 5481–5487.

Hoek, G., R. Beelen, K. De Hoogh, et al. 2008. "A Review of Land-Use Regression Models to Assess Spatial Variation of Outdoor Air Pollution." *Atmospheric Environment* 42, no. 33: 7561–7578.

Kaufman, L., and P. J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons.

Kim, Y.-J., and C. Gu. 2004. "Smoothing Spline Gaussian Regression: More Scalable Computation via Efficient Approximation." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 66, no. 2: 337–356.

Li, B., M. G. Genton, and M. Sherman. 2007. "A Nonparametric Assessment of Properties of Space–Time Covariance Functions." *Journal of the American Statistical Association* 102, no. 478: 736–744.

Li, Z., and S. N. Wood. 2020. "Faster Model Matrix Crossproducts for Large Generalized Linear Models With Discretized Covariates." *Statistics and Computing* 30, no. 1: 19–25.

Lu, X., A. E. Gelfand, and D. M. Holland. 2018. "Local Real-Time Forecasting of Ozone Exposure Using Temperature Data." *Environmetrics* 29, no. 7: e2509.

Ma, Y., B. Ma, H. Jiao, Y. Zhang, J. Xin, and Z. Yu. 2020. "An Analysis of the Effects of Weather and Air Pollution on Tropospheric Ozone Using a Generalized Additive Model in Western China: Lanzhou, Gansu." *Atmospheric Environment* 224: 117342.

Maranzano, P., R. Borgoni, S. Doghmi, and A. Tassan Mazzocco. 2025. "EEAaq: Handle Air Quality Data From the European Environment Agency Data Portal. R Package Version 1.0.1."

Marra, G., and S. N. Wood. 2011. "Practical Variable Selection for Generalized Additive Models." *Computational Statistics & Data Analysis* 55, no. 7: 2372–2387.

McMillan, N., S. M. Bortnick, M. E. Irwin, and L. M. Berliner. 2005. "A Hierarchical Bayesian Model to Estimate and Forecast Ozone Through Space and Time." *Atmospheric Environment* 39, no. 8: 1373–1382.

Meijer, J. R., M. A. Huijbregts, K. C. Schotten, and A. M. Schipper. 2018. "Global Patterns of Current and Future Road Infrastructure." *Environmental Research Letters* 13, no. 6: 064006.

Meng, X., W. Wang, S. Shi, et al. 2022. "Evaluating the Spatiotemporal Ozone Characteristics With High-Resolution Predictions in Mainland China, 2013–2019." *Environmental Pollution* 299: 118865.

Meyer, H., C. Reudenbach, T. Hengl, M. Katurji, and T. Nauss. 2018. "Improving Performance of Spatio-Temporal Machine Learning Models Using Forward Feature Selection and Target-Oriented Validation." *Environmental Modelling & Software* 101: 1–9.

Muehlmann, C., S. De Iaco, and K. Nordhausen. 2022. "Blind Recovery of Sources for Multivariate Space-Time Random Fields." *Stochastic Environmental Research and Risk Assessment* 37, no. 4: 1593–1613. https://doi.org/10.1007/s00477-022-02348-2.

Muñoz Sabater, J. 2019. *ERA5-Land Hourly Data From 1950 to Present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac.

Niu, M., Y. Zhang, and Z. Ren. 2023. "Deep Learning-Based pm2.5 Long Time-Series Prediction by Fusing Multisource Data—A Case Study of Beijing." *Atmosphere* 14, no. 2: 340.

Nuvolone, D., D. Petri, and F. Voller. 2018. "The Effects of Ozone on Human Health." *Environmental Science and Pollution Research* 25: 8074–8088.

Otto, P., A. Fassò, and P. Maranzano. 2024. "A Review of Regularised Estimation Methods and Cross-Validation in Spatiotemporal Statistics." *Statistics Surveys* 18: 299–340.

Paci, L., A. E. Gelfand, and D. M. Holland. 2013. "Spatio-Temporal Modeling for Real-Time Ozone Forecasting." *Spatial Statistics* 4: 79–93.

Pan, Q., F. Harrou, and Y. Sun. 2023. "A Comparison of Machine Learning Methods for Ozone Pollution Prediction." *Journal of Big Data* 10, no. 1: 63.

Potapov, P., M. C. Hansen, A. Pickens, et al. 2022. "The Global 2000-2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results." *Frontiers in Remote Sensing* 3: 856903.

R Core Team. 2025. *R: A Language and Environment for Statistical Computing [Computer Software Manual]*. R Core Team. https://www.R-project.org/.

Ramya, A., P. Dhevagi, R. Poornima, S. Avudainayagam, M. Watanabe, and E. Agathokleous. 2023. "Effect of Ozone Stress on Crop Productivity: A Threat to Food Security." *Environmental Research* 236: 116816.

Roberts, D. R., V. Bahn, S. Ciuti, et al. 2017. "Cross-Validation Strategies for Data With Temporal, Spatial, Hierarchical, or Phylogenetic Structure." *Ecography* 40, no. 8: 913–929.

Schipa, I., A. Tanzarella, and C. Mangia. 2009. "Differences Between Weekend and Weekday Ozone Levels Over Rural and Urban Sites in Southern Italy." *Environmental Monitoring and Assessment* 156: 509–523.

Silibello, C., G. Carlino, M. Stafoggia, et al. 2021. "Spatial-Temporal Prediction of Ambient Nitrogen Dioxide and Ozone Levels Over Italy Using a

Random Forest Model for Population Exposure Assessment." *Air Quality, Atmosphere & Health* 14, no. 6: 817–829.

Sipilä, M., C. Cappello, S. De Iaco, K. Nordhausen, and S. Taskinen. 2025. "Modelling Multivariate Spatio-Temporal Data With Identifiable Variational Autoencoders." *Neural Networks* 181: 106774. https://doi.org/10.1016/j.neunet.2024.106774.

van Nunen, E., R. Vermeulen, M.-Y. Tsai, et al. 2017. "Land Use Regression Models for Ultrafine Particles in Six European Areas." *Environmental Science & Technology* 51, no. 6: 3336–3345.

Wackernagel, H. 2003. *Multivariate Geostatistics*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-05294-5.

Wang, J., D. S. Cohan, and H. Xu. 2020. "Spatiotemporal Ozone Pollution Lur Models: Suitable Statistical Algorithms and Time Scales for a Megacity Scale." *Atmospheric Environment* 237: 117671.

Wang, W., D. Fecht, S. Beevers, and J. Gulliver. 2022. "Predicting Daily Concentrations of Nitrogen Dioxide, Particulate Matter and Ozone at Fine Spatial Scale in Great Britain." *Atmospheric Pollution Research* 13, no. 8: 101506.

Wenger, S. J., and J. D. Olden. 2012. "Assessing Transferability of Ecological Models: An Underappreciated Aspect of Statistical Validation." *Methods in Ecology and Evolution* 3, no. 2: 260–267.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction With R*. Chapman and Hall/CRC.

Wood, S. N., Z. Li, G. Shaddick, and N. H. Augustin. 2017. "Generalized Additive Models for Gigadata: Modeling the U.K. Black Smoke Network Daily Data." *Journal of the American Statistical Association* 112, no. 519: 1199–1210.

Xu, J., Z. Zhang, L. Rao, et al. 2024. "Remote Sensing of Tropospheric Ozone From Space: Progress and Challenges." *Journal of Remote Sensing* 4: 0178.

Yan, Y., A. Pozzer, N. Ojha, J. Lin, and J. Lelieveld. 2018. "Analysis of European Ozone Trends in the Period 1995–2014." *Atmospheric Chemistry and Physics* 18, no. 8: 5589–5605.

Zhan, Y., Y. Luo, X. Deng, M. L. Grieneisen, M. Zhang, and B. Di. 2018. "Spatiotemporal Prediction of Daily Ambient Ozone Levels Across China Using Random Forest for Human Exposure Assessment." *Environmental Pollution* 233: 464–473.

Zhang, J., Y. Wei, and Z. Fang. 2019. "Ozone Pollution: A Major Health Hazard Worldwide." *Frontiers in Immunology* 10: 2518.

Zhu, S., J. Xu, J. Zeng, et al. 2023. "Leso: A Ten-Year Ensemble of Satellite-Derived Intercontinental Hourly Surface Ozone Concentrations." *Scientific Data* 10, no. 1: 741.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1.** Supporting Information.

Appendix A

Retrieving and Pre-Processing Ozone Data From Monitoring Stations

Information on all air quality monitoring stations available through the European Environment Agency (EEA) database was retrieved using the `R` package `EEAaq`. The dataset was filtered to include only stations located in Italy, and the corresponding unique station codes and geographical coordinates were extracted.

```
library("EEAaq")                        # load package for accessing EEA air quality data
library("tidyverse")                    # load tidyverse for data manipulation and cleaning

stations <- EEAaq_get_stations() |>     # retrieve metadata for all monitoring stations
  filter(Country == "Italy") |>         # keep only Italian stations

years <- 2013:2023                      # define range of years to download
EEA_data_list <- vector("list", length=length(years))  #  empty list to store yearly data
options(timeout=1200)                   # extend download timeout to prevent interruptions
for (i in seq_along(years))             # loop through each year
{
  EEA_data_list[[i]] <- tryCatch({      # handle download errors gracefully
    EEAaq_get_data(                     # download hourly ozone data for the year at  stations
      IDstations=stations$AirQualityStationEoICode, pollutants="O3",
      from=paste0(years[i], "-01-01"), to=paste0(years[i], "-12-31"),
      verbose=TRUE) |>
      filter(is.finite(O3), !is.na(AirQualityStationEoICode)) |> #remove invalid/missing data
      mutate(date=as.Date(DatetimeBegin)) |>                     # extract date from timestamp
      group_by(AirQualityStationEoICode, date) |>                # group by station and date
      summarize(                        # compute daily summary statistics
        mean_O3=mean(O3, na.rm=TRUE), min_O3=min(O3, na.rm=TRUE),
        max_O3=max(O3, na.rm=TRUE), .groups="drop") |>
      left_join(stations, by=join_by(AirQualityStationEoICode)) |>  # add station coordinates
      relocate(Longitude, Latitude, .after=AirQualityStationEoICode) # reorder columns
  }, error=function(e) NULL)            # return null if download fails
}
which(unlist(lapply(EEA_data_list, is.null))) # check if there is no missing year
EEA_data <- bind_rows(EEA_data_list)  # combine all yearly data into one data frame
```