



Spatiotemporal Event Studies for Environmental Data Under Cross-Sectional Dependence: An Application to Air Quality Assessment in Lombardy

Paolo MARANZANO  and Matteo PELAGATTI

We propose a twofold adjustment for Event Studies considering spatiotemporal data in a multivariate time series framework where the data are characterized by spatial and temporal dependence. The first adjustment consists of modeling the spatiotemporal dynamics of the data by implementing several geostatistical models capable of handling both spatial and temporal components, as well as estimating the relationship between the response variable and a set of exogenous factors. With the second adjustment, we propose to use cross-sectional-adjusted test statistics directly accounting for spatial cross-correlation. The proposed methods are applied to the case of NO₂ concentrations observed in Northern Italy during the first wave of the COVID-19 pandemic. The key findings are as follows. First, all the considered geostatistical models estimate larger reductions in the major metropolitan and congested areas, while smaller reductions are estimated in rural plains and in the mountains. Second, the models are nearly equivalent in terms of fitting and are capable of identifying the true event window. Third, by using spatiotemporal models we ensure the residuals are uncorrelated across space and time, thus allowing Event Studies test statistics to provide reliable and realistic estimates. Fourth, as expected, all test statistics show significant reductions in NO₂ concentrations starting from the first few days of lockdown.

Supplementary materials accompanying this paper appear online.

Key Words: Multivariate time series; Spatial cross-sectional dependence; Event Studies; Air quality; Abnormal concentrations.

Mathematics Subject Classification C12 · C23 · C32 · G14 · Q53 · Q58

P. Maranzano (✉) · M. Pelagatti, Department of Economics, Management and Statistics (DEMS), University of Milano-Bicocca, Piazza dell'Ateneo Nuovo, 1, 20126 Milano, Italy (E-mail: paolo.maranzano@unimib.it); (E-mail: matteo.pelagatti@unimib.it).

P. Maranzano, Fondazione Eni Enrico Mattei (FEEM), Corso Magenta, 63, 20123 Milan, Milano, Italy.

© 2023 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics

<https://doi.org/10.1007/s13253-023-00564-z>

1. INTRODUCTION

1.1. EVENT STUDIES: FUNDAMENTALS AND MODELING STRATEGY

Event Studies, hereafter ES, are statistical tools used to assess whether a particular event of interest has induced relevant changes in the evolution of one or more time series. The changes may concern either the mean level, i.e., testing for a level shift, (Campbell et al. 1998) or (less frequently) the variability of the phenomenon, i.e., testing for a variance shift (Giaccotto and Sfridis 1996).

ES are grounded on two main pillars. The first pillar is the *interrupted time series* paradigm (McDowall et al. 2019), in which we assume that at a certain known date, an event (e.g., treatment or intervention) occurred, dividing the time series into two parts: before and after the event occurrence. The occurrence date is labeled as the *event date* and is assumed to be known. The final goal of an ES is to state if the event of interest generated a statistically significant impact on the time series under consideration. The event can be permanent or temporary, and it can be gradual or abrupt. Usually, ES focus on abrupt changes. The second pillar is the *offline hypothesis testing* (Basseville and Nikiforov 1993), in which a *no-change* scenario is compared to a *with-change* scenario. The two scenarios are compared through a statistical hypothesis test. Considering ES for a level shift, under the null hypothesis, we state that there are no abnormal variations at the time of the event, whereas the alternative hypothesis assumes the existence of a level shift in correspondence with the event of interest.

ES combine a regression-based approach for parameter estimation complemented with a validation strategy based on ad hoc tests. The standard procedure for ES consists of segmenting the timeline into two consecutive subsamples: The first part of the time series, i.e., the *estimation window*, is used to estimate a regression model, while the second part, called the *event window*, is used to test the statistical significance of the event. The regression model takes as the response variable the time series on which to calculate the impact of the event and typically employs a set of predictors able to explain the movements of the response. Thus, the regression step is used to control for confounding variables whose effects should be filtered out before testing for the presence of a shift. The estimated regression model is then used to make predictions about future values of the time series of interest. Eventually, the ES test statistics are computed using the prediction errors within the event window, i.e., using the observations not used in the model's estimation at the event date. In this regard, the ES approach is similar to modern statistical learning techniques for temporal data, in which a portion of the time series is used for in-sample model training and parameter estimation, whereas performance is evaluated over successive out-of-sample periods in cross-validation (Bergmeir et al. 2018). The reader can refer to the figure on page 332 of Benninga (2014) for an explicative graphical representation of the typical outline of an event study.

When studying the presence of an event-induced level shift, if the event truly generated a significant effect, the observations following the event should diverge from the model-based predictions, leading to prediction errors showing significant level changes. Conversely, if the event did not induce any significant level shift within the event window, the predicted and actual values should overlap, leading to forecast errors averaging zero.

ES can focus both on univariate time series and multiple time series. In the latter, the observations will be denoted by a pair of symbols, namely the index s for the cross-sectional units and the index t for time. An ES with multiple time series is used when it is intended to take advantage of the correlation existing between cross-sectional units, allowing the construction of aggregate statistics for the entire system. In addition, ESs can involve single events in time (typically single-day occurrences) or event windows composed of multiple instants of time, called cumulated event windows.

1.2. OUR CONTRIBUTION

We contribute to the empirical literature on ES by proposing a twofold adjustment for Event Studies considering spatiotemporal data. In particular, we are interested in ES applied to multivariate time series characterized by the presence of spatial (i.e., cross-sectional) and temporal dependence. Also, we focus on ES for multi-observation event windows by means of cumulated ES test statistics. Since spatial and temporal correlation is typically found in environmental data (recall the three Laws of Geography by [Tobler 1970](#); [Zhu and Turner 2022](#)), we are implicitly suggesting a research framework for addressing environmental Event Studies. Also, extending ES to the spatiotemporal context is an improvement that can provide great benefits to projects whose goal is to assess the impact of policies at the territorial level ([Fassó et al. 2023](#)).

The first adjustment concerns the modeling step. Usually, ES literature assumes a linear relationship between the response variable and the covariates ([Borghesi et al. 2022](#)). Moreover, the relationship is often estimated in a univariate framework ([Neill and Chen 2022](#)); that is, for each cross-sectional unit s the observations are modeled as linear functions of the set of predictors. This is generally suboptimal when concentrations observed at the stations are not mutually independent. Also, in linear regression contexts, estimates of regression coefficients could be biased due to not explicitly modeled spatial ([Paciorek 2010](#)) and temporal ([Lee and Lund 2008](#)) dependence into the residuals. Since ES statistics are calculated from the predicted residuals, the presence of confounding spatiotemporal correlation can adversely affect the values of the statistics and their uncertainty. We aim at relaxing the independence assumption by explicitly modeling the spatiotemporal dynamics of the data by implementing several geostatistical models capable of handling both spatial and temporal components, as well as estimating the relationship between the response variable and a set of exogenous factors. In particular, we will consider models belonging to the class of linear mixed models (LMMs) and generalized additive mixed models (GAMMs).

The second adjustment refers to the hypothesis testing step. Specifically, we aim at addressing the problem of adjusting ES against cross-correlation when dealing with spatially distributed observations. When the observations are affected by (even by small amounts of) positive cross-sectional dependence (CD), classical ES parametric test statistics reject the null hypothesis too often ([Pelagatti and Maranzano 2021b](#)). The same considerations hold when observations are affected by temporal autocorrelation ([Lee and Lund 2004, 2008](#)), or when considering correlated paired samples ([Dutilleul et al. 1993](#); [Zimmerman 2012](#)). Size-distortion effects still hold when negligible levels of cross-correlation are observed ([Pelagatti and Maranzano 2021b](#)). We adopt a strategy involving the use of CD-adjusted test statistics

to directly account for spatial cross-sectional dependence by means of a cross-correlation measure for multivariate time series.

In the following, we will treat spatial dependence and cross-sectional dependence as interchangeable terms. Aware that these are two different statistical concepts, we can still point out a close connection between the two. Indeed, the Global Moran's I (Moran 1950) can be interpreted as a generalization of Pearson's correlation coefficient with geographical weights (Chen 2013). Many of the ES test statistics used in the next sections are adjusted with respect to cross-sectional dependence by means of Pearson's linear correlation coefficient. While this is an approximate *aspatial* measure of autocorrelation, it can still provide a straightforward indication of the direction of the relationship and its strength. In fact, the bivariate-spatial correlation can be expressed as a fraction of Pearson's linear correlation coefficient, which acts as an upper-bound (Lee 2001).

Eventually, we provide an empirical application on the airborne pollutant concentrations observed on geo-referenced monitoring networks located in Northern Italy. As with many other environmental phenomena, air quality data are affected by positive spatial dependence (Dale and Fortin 2009). Air quality data are a natural example of the three Laws of Geography; that is, *near things are more related than distant things* (Tobler 1970), and *the more similar the geographic configurations of two points, the more similar the values (processes) of the target variable at these two points* (Zhu and Turner 2022). Indeed, by analyzing the measurements recorded in monitoring networks belonging to specific regions and assuming very similar environmental conditions, it is realistic to state that the concentrations of airborne pollutants observed in monitoring stations located at close distances will be very similar (Montero et al. 2021). Also, pollutant concentrations are strongly seasonal and persistent phenomena, leading to strong temporal autocorrelation among observations. In this context, the proposed adjustments seem ineluctable for obtaining credible estimates and thus reasonable policy guidance.

The remainder of the paper is organized as follows. In Sect. 2, we provide a short literature review on the use of ES in environmental and energy fields. In Sect. 3, we propose an Event Studies taxonomy tailored to the case of air quality data and briefly introduce the CD-adjusted test statistics implemented in the application section. In Sect. 4, we present the HDGM and discuss its interpretation and the major benefits of its implementation. In Sect. 5, we present an ES concerning the effect of the COVID-19-related restrictions on air quality in the Lombardy region (Italy) in 2020. Finally, Sect. 6 concludes the paper and proposes future developments of ES in spatiotemporal frameworks.

2. EVENT STUDIES FOR ENVIRONMENT AND ENERGY: STATE OF THE ART

Event Studies are only recently receiving attention in the environmental and energy fields. Among others, oil and fuels commodity markets provide some examples. Demirel and Kutan (2010) use ES methodology to examine the behavior of crude oil spot and futures markets around the OPEC conference, as well as the US strategic petroleum reserves announcements between 1983 and 2008. Zhang et al. (2009) use ES to test the impact of extreme events,

such as the Gulf War in 1991 and the Iraq War in 2003, on crude oil price volatility. Further, in [Zha et al. \(2018\)](#) the authors aim at assessing the impact of refined oil price adjustments to control air pollution in China between 2014 and 2015. In addition, ES methods have recently received great attention in climate policy analysis. Looking at the macroeconomic perspective, one can refer to the paper by [Barnett \(2019\)](#), in which the impacts of climate policy risk exposure on observable market outcomes such as oil production, stock returns, and oil prices are analyzed. In [Diaz-Rainey et al. \(2021\)](#), the authors examined the effect of policy interventions associated with the Paris Agreement (agreement and ratification) and the election of Donald Trump (election and withdrawal from the agreement) on stock returns of oil and gas companies. Other researchers focused on the effect of climate policies on stock returns and investment portfolios. We recall, for example, [Borghesi et al. \(2022\)](#) examining the behavior of green and brown portfolios around green policy-related announcements launched by European governments in 2020 to alleviate the adverse effect of climate change; [Birindelli and Chiappini \(2021\)](#), which examined investor reaction toward eight EU policy announcements over the years 2013–2018 on a large sample of EU firms; and [Huynh and Xia \(2020\)](#), which used ES to analyze the effect of climate change news on individual corporate bond returns.

Regarding air quality, most of the contributions using ES focus on city-level data with daily or hourly frequencies. However, to the best of our knowledge, none of them correct the statistics for CD or employ spatiotemporal models to filter the spatial and temporal dependence. Also, most of them focus on Asian-located case studies, whereas no examples are available for Europe or other regions. For instance, [Li et al. \(2019\)](#) investigate the effect of mega events on local air quality (daily $PM_{2.5}$) using a comparative ES involving CD-independent time series from a treatment location and a placebo location. Similarly to our goals, [Xiao et al. \(2022\)](#) investigate the benefits deriving from lockdown measures on air quality ($PM_{2.5}$) in 31 cities across China. The authors implement a univariate time series model using a 40-day-long event window but do not control for correlation among locations. Other papers use ES as a robustness check for causal inference tools, such as the difference-in-differences (see, for instance, [Li et al. 2022](#); [Djoundourian et al. 2022](#); [Weng et al. 2022](#)), or provide a combination of the two (see, for example, [Naqvi 2021](#); [Lin and Zhu 2019](#); [Xu et al. 2022](#)).

3. EVENT STUDIES FOR AIR QUALITY ASSESSMENT: TAXONOMY AND STATISTICS

Let $s = 1, \dots, S$ identify the cross-sectional units (spatial locations), and let t be the time index $t = 1, \dots, T$. ES rely on a validation strategy performed by splitting the whole temporal sequence into two disjoint subsamples, namely the *estimation window* and the *event window*. The observations in the estimation window are used to estimate the model parameters, while those in the event window are used to assess the event's effect.

Formally, the estimation window is the set of time indexes Ω_0 containing the first T_0 time points $1, \dots, T_0$, while the event window is the set of time points Ω_1 containing

$T_0 + 1, \dots, T$ and whose cardinality is $T_1 = T - T_0$. For completeness, we define also the full set $\Omega = \Omega_0 \cup \Omega_1$ with cardinality $T = T_0 + T_1$.

Let C_{st} be the observed airborne pollutant concentrations observed at time t and monitoring station s . Moreover, let X_{st} be a vector of conditioning covariates collected at the same time t and station s . Notice that the set X_{st} can include station-specific information (e.g., local weather measurements, traffic data, land cover) and information common to all the sensors (e.g., calendar effects) or further variables measured at an aggregate geographical level. Assuming the existence of a statistical relationship between the concentrations and conditioning information, for each spatial sampling point s and time t , the *normal concentration* (NC_{st}) can be defined as the conditional expectation of C_{st} given X_{st} , i.e.,

$$NC_{st} = \mathbb{E}[C_{st}|X_{st}].$$

It follows that the *abnormal concentration* (AC_{st} or ε_{st}) is defined as the difference between the observed concentrations at time t and expected concentration at time t for station s :

$$AC_{st} = \varepsilon_{st} = C_{st} - \mathbb{E}[C_{st}|X_{st}] = C_{st} - NC_{st}.$$

The abnormal concentrations in the event window Ω_1 can be interpreted as the abnormal values of C not explained by the conditioning information X_{st} , and potentially generated by the event of interest.

Finally, by the term *cumulated abnormal concentration* we mean the cumulative sum of abnormal concentrations in a given time window. We are particularly interested in the cumulated abnormal concentration in the event window, hereafter $CAC_{s\Omega_1}$. The cumulated abnormal concentration for station s over the event window Ω_1 is defined by

$$CAC_{s\Omega_1} = \sum_{t \in \Omega_1} AC_{st}. \quad (1)$$

Note that, to connect this notation to the previously existing literature, in Table S1 of Supplementary materials, we provide a synthetic *conversion* table mapping the air quality assessment taxonomy here proposed to the main statistical and financial notation.

In the following application, we apply and compare a subset of ES test statistics presented and discussed in Pelagatti and Maranzano (2021b). In particular, we will use the statistics directly accounting for cross-sectional dependence, thus being CD-adjusted. The proposed statistics test the null hypothesis of the absence of a level shift in the cumulated abnormal concentrations (CAC). We list the implemented test statistics in Table 1. The list includes both parametric and nonparametric specifications, in particular those belonging to the family of rank-based statistics (Kolari and Pynnönen 2011; Luoma 2011; Hagnäs and Pynnönen 2014). The main difference among the statistics lies in how they account for cross-sectional dependence. For instance, while the Patell and the BMP statistics make use of the linear correlation on the abnormal concentrations, the P_1 and P_2 statistics compute the correlation on the ranks of the abnormal concentrations. An extended discussion on the statistical properties, as well as simulated and empirical results about the performance of each statistic, is available in the same article.

Table 1. Test statistics for $H_0: \mathbb{E}[CAC_{\Omega_1}] = 0$

Specification	Description	Statistic	Reference
Param.	Patell's Z	$Z_{\text{patell_adj}}$	Kolari and Pynnönen (2011)
Param.	Boehmer–Musumeci–Poulsen	$Z_{\text{BMP_adj}}$	Kolari and Pynnönen (2011)
Nonpar.	Generalized rank T	T_{grank}	Kolari and Pynnönen (2011)
Nonpar.	Generalized rank Z adjusted	$Z_{\text{grank_adj}}$	Kolari and Pynnönen (2011)
Nonpar.	Campbell–Wasley	CumRank	Corrado (1989)
Nonpar.	Cumulated rank modified	CumRank_mod	Corrado and Zivney (1992)
Nonpar.	Cumulated rank T	CumRank_T	Corrado and Zivney (1992)
Nonpar.	Cumulated rank Z adjusted	CumRank_Z_adj	Hagnäs and Pynnönen (2014)
Nonpar.	Normalized rank	P_1	Pelagatti and Maranzano (2021b)
Nonpar.	van der Waerden rank	P_2	Pelagatti and Maranzano (2021b)
Nonpar.	Corrado–Tukey adjusted	CT_adj	Pelagatti and Maranzano (2021b)

4. GEOSTATISTICAL MODELS FOR AIR QUALITY

We propose four alternative spatiotemporal models to model air quality as a function of several predictors and to obtain abnormal concentrations to be employed by the Event Studies test statistics. Airborne pollutant concentrations are usually characterized by strong right skewness due to unexpectedly high concentrations ([Mudelsee and Alkio 2007](#)). To address this issue, we considered a log transformation for the original data which led to a Gaussian-like distribution of observations ([Maranzano et al. 2020](#)). Notice that since we considered rank-based nonparametric test statistics, their results are not affected by monotonic transformations, such as the logarithm. Thus, all the ES test statistics presented below are computed on the log-scaled abnormal concentrations.

All the considered models assume that log concentrations of NO_2 are generated by a spatiotemporal process $\{Y_{st} \in \mathbb{R} : s \in D, t = 1, \dots, T\}$, where D is the spatial domain composed of S locations and t represents a discrete point of time. Also, we assume that the concentrations are influenced by a set of p site-specific exogenous covariates, including weather conditions, land use, and calendar effects.

We propose four geostatistical models taking into account the spatial and temporal dependence of the data: (1) hidden dynamics geostatistical model (HDGM); (2) a generalized additive model (GAM) with a nonlinear smooth trend; (3) a generalized additive mixed model (GAMM) with a nonlinear smooth trend and site-specific random effects; and (4) a generalized additive mixed model with a nonlinear smooth trend, site-specific random effects, and temporal AR(1) structure for the site-specific errors (GAMM-ar1).

There are several differences between HDGM and the three models from the GAMM family. On the one hand, HDGM allows only purely linear relationships between regressors and the response variable, whereas GAMMs allow nonlinear smooth functions via spline basis expansions. On the other hand, while HDGM is a mixed model with a small-scale random spatiotemporal component based on autoregressive temporal processes and spatial Gaussian processes, GAMMs can only include one between temporal and spatial dependence in the small-scale component and allow flexible relationships in the fixed-effects component.

Numerous examples of environmental applications involving both classes of models can be found in the literature. For instance, the HDGM has been extensively used in air quality policy assessment (Fassó et al. 2021; Maranzano et al. 2023), bike-sharing system comprehension (Piter et al. 2022), off-shore coastal profile measurements for beach monitoring (Otto et al. 2021), and spatiotemporal interpolation of missing observations in land-use regression (Taghavi-Shahri et al. 2019). On the other side, GAMMs have been widely used in the epidemiological (Cabrera and Taylor 2019; Feng 2022), environmental (Padilla et al. 2014), and socioeconomic (Hu et al. 2022) fields due to their impressive flexibility.

4.1. HDGM

The HDGM (Calculli et al. 2015) belongs to the class of linear mixed models (LMMs) and entails a random-effect term w_{st} (i.e., the small-scale component) modeling the spatial and temporal dependence, a fixed-effect term v_{st} (i.e., the large-scale component) accounting for all exogenous regressive effects and a measurement error term ε_{st} . The model can be specified by the following system of equations:

$$Y_{st} = v_{st} + w_{st} + \varepsilon_{st} \quad (2)$$

with $\varepsilon_{st} \sim N(0, \sigma_\varepsilon^2)$ being the measurement error vector that is assumed to be independent and identically distributed (i.i.d.) across space and time. Note that, recalling the taxonomy of ES introduced in Sect. 3, the response variable Y_{st} is the equivalent of the observed concentrations C_{st} , while the measurement error term ε_{st} is equivalent to the abnormal concentrations AC_{st} .

The fixed-effects mean term can be specified as follows:

$$v_{st} = \mathbf{x}_{st}^\top \boldsymbol{\beta}, \quad (3)$$

where \mathbf{x}_{st} is a vector with p covariates observed at location s and time t , and $\boldsymbol{\beta}$ is a vector of p coefficients. The random effects term w_{st} assumes a separable space-time covariance for the random process Y_{st} , as the spatiotemporal dynamics is described by Markovian autoregressive temporal processes plus spatially correlated random effects ω_{st}

$$w_{st} = \phi_{HDGM} w_{st-1} + \omega_{st}, \quad (4)$$

where $|\phi_{HDGM}| < 1$ represents the common first-order temporal autocorrelation parameter. The innovation ω_{st} is assumed to be a realization of a Gaussian process independent in time and with the spatial exponential covariance function with the range parameter $\theta_{HDGM} > 0$.

HDGM can be represented through a state-space representation, and the maximum likelihood estimates of the parameters involved are computed using a spatiotemporal Kalman filter (Ferreira et al. 2022; Jurek and Katzfuss 2022, 2023) and the EM algorithm. Parameter estimation, as well as computation of the variance–covariance matrix, is implemented in the D-STEM package (Wang et al. 2021) for MATLAB.

4.2. GENERALIZED ADDITIVE MIXED MODELS

Generalized additive models (GAMs) (Pinheiro and Bates 2006) are linear additive models allowing for either linear or nonlinear relationships between the predictors and the response variable. Nonlinear relationships are included by means of smooth functions, typically represented by basis function expansions. A straightforward extension of GAMs is given by generalized additive mixed model (GAMMs) (Wood 2017), which keeps the same semi-parametric structure of GAMs but allows including an additive random effect. Such a random effects can be used to describe the spatiotemporal dependence of the data.

While in GAMMs the evolution of NO_2 concentrations can be expressed as in Equation (2), the measurement equation for GAMs is simply given by the sum of the large-scale component v_{st} and the measurement error ε_{st} , i.e.,

$$Y_{st} = v_{st} + \varepsilon_{st}. \quad (5)$$

In both cases, the large-scale component v_{st} can include a set of p linear and m nonlinear terms, that is:

$$v_{st} = \mathbf{x}_{L,st}^\top \boldsymbol{\beta}_L + \sum_{j=1}^m \alpha_{(j)} \mathbf{x}_{NL,j}, \quad (6)$$

where $\mathbf{x}_{L,st}$ is the vector of purely linear covariates observed at location s and time t , and $\alpha_{(j)} \mathbf{x}_{NL,j}$ is an additive term (basis expansion) with nonlinear influence function $\alpha_{(j)}$ of the j -th covariate for the m nonlinear covariates. To achieve the highest comparability possible among different models, we considered the same set of p purely linear covariates across the specifications. The spatiotemporal dependence is modeled both in the large- and small-scale components. Indeed, the nonlinear term of (6) is used to model the spatial dependence among observations through a single smooth function, i.e., $m = 1$. In particular, we considered a smooth surface of the sites' coordinates (i.e., longitude and latitude) following a Gaussian process with an exponential covariance function with range parameter θ . In the following, we will refer to the GAM model following Equations (5) and (6) as the *GAM* with range parameter θ_{GAM} .

We then considered two alternative specifications for the small-scale component of GAMMs. The first one includes the spatial dependence as a sequence of site-specific time-independent Gaussian-distributed random effects, i.e.,

$$w_{st} = w_s \sim N(\mathbf{0}, \sigma_{GAMM}), \quad (7)$$

with σ_{GAMM} being the variance of the random effects. In the following, we will refer to this model as *GAMM* being characterized by a range parameter θ_{GAMM} .

The second specification of the small-scale component includes either time dependence or spatial dependence effects. Indeed, the spatial dependence is embedded (as in the previous specification) by a sequence of site-specific time-independent Gaussian-distributed random effects. In contrast, the time dependence is modeled via a site-specific first-order autoregressive model on the residuals calculated at each site (within-group residuals), that is,

$$w_{st} = \phi(Y_{st-1} - v_{st-1} - w_s), \quad (8)$$

with ϕ being the autoregressive parameter representing the temporal dependence, v_{st-1} following Equation (6), and w_s following Equation (7). In the following, we will refer to this model as *GAMMar1* and is characterized by a range parameter $\theta_{GAMMar1}$.

The model estimation is performed via restricted maximum likelihood, which is implemented within the package `mgcv` available in R. The range parameters are estimated as proposed by [Kammann and Wand \(2003\)](#), thus by taking the maximum Euclidean distance computed on the sites' coordinates. Notice that, using the same monitoring network for the three specifications, this returns the estimated value for θ .

5. ASSESSING THE IMPACT OF COVID-19 LOCKDOWN MEASURES ON AIR QUALITY IN LOMBARDY

To fight the spreading of COVID-19 across the country, the Italian government imposed a total lockdown ([Presidenza del Consiglio dei Ministri Italia 2020](#)) from March 9 to May 18, 2020, for a total of 71 days. This period, also denoted as *first-wave COVID-19 lockdown*, was characterized by the closure of all non-essential activities and enterprises, and by the minimization of individual mobility and social distancing ([Pelagatti and Maranzano 2021a](#)). As a direct consequence of the limitations, a generalized reduction of car traffic and personal travel took place in the entire country ([Finazzi and Fassò 2020](#)).

Numerous studies have shown how general lockdowns imposed by governments have generated strong reductions in pollutant concentrations worldwide ([Higham et al. \(2020\)](#); [Zangari et al. \(2020\)](#); [Nakada and Urban \(2020\)](#); [Xin et al. \(2021\)](#)), particularly in large urban centers ([Baldasano 2020](#); [Rossi et al. 2020](#)). The Lombardy (Northern Italy) case study received remarkable scientific interest. In particular, the studies by [Collivignarelli et al. \(2020\)](#); [Cameletti \(2020\)](#); [Fassó et al. \(2021\)](#); [Maranzano and Fassó \(2021\)](#); [Granella et al. \(2021\)](#) showed that, due to the restrictions on mobility, oxide concentrations registered statistically significant reductions (up to 50%) throughout the region. On the contrary, the particulate matter remained stable or slightly reduced over the entire period. This indicates that the major emission sources of particulate matter in the region are other than vehicular traffic and industrial production. Consider, for example, the role of agriculture and livestock farming, which, through the production of ammonia, generates significant amounts of secondary particulate matter ([Lovarelli et al. 2020, 2021](#)).

5.1. EVENT STUDY STRATEGY: RECURSIVE WINDOW

We are interested in analyzing the effect of the lockdown restrictions on NO_2 concentrations registered in Lombardy. The null hypothesis we are testing is that restrictions did not have any effect on the cumulative abnormal NO_2 concentrations during the lockdown period (i.e., $CAC_{s\Omega_1}$ have null mean value). The alternative hypothesis is that the cumulated abnormal concentrations registered a significant reduction during the event window (i.e., $CAC_{i\Omega_1}$ have negative mean value). Therefore, the hypothesis test is one-sided on the

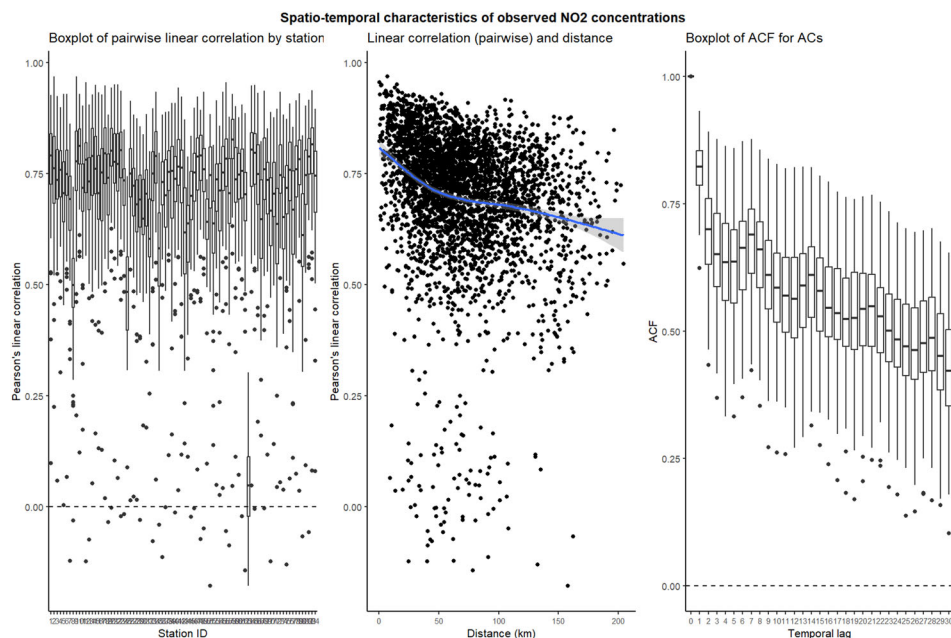


Figure 1. Spatiotemporal characteristics of observed NO_2 concentrations in Lombardy between January 1, 2018, and January 31, 2020. Left panel: boxplot of pairwise Pearson's linear correlation index by the station. Central panel: Linear correlation against geodetic distance among monitoring sites. Right panel: boxplot of ACF up to lag 30 days .

left tail. In other words, we are testing the presence of a negative level shift in the average NO_2 concentrations due to the lockdown. Previously existing literature confirmed significant reductions in NO_2 levels in Lombardy. Thus, a statistically significant negative sign of the statistics is expected.

We consider the average daily concentrations of NO_2 collected in $S = 84$ ground stations belonging to the ARPA Lombardia monitoring network. (The considered stations are represented in Figure S1 in Supplementary Materials, while for an extended description of the ARPA network we refer to the reader to [Maranzano 2022](#).) Air quality measurements are collected using the *ARPALData* package (release 1.3.1 available on CRAN) of the statistical software *R* ([R Core Team 2020](#)). For each monitoring site, we collected daily observations of NO_2 concentrations from January 1, 2018, to May 31, 2020, totaling $T = 881$ days. Figure 1 shows the main spatiotemporal features of NO_2 concentrations from 2018 to right before the pandemic started. The stations are generally characterized by high temporal persistence both at short and long lags and by strong positive linear correlation (with a median value around 75%). The latter is able to heavily bias classical ES statistics that do not consider cross-sectional dependence adjustments. The correlation tends to decrease as the distance among monitoring sites increases, while remaining sustained even for high distances, proving that linear correlation can be used as a proxy of spatial correlation.

We know that the official start date of the lockdown period is March 8, 2020. The classical approach to ES sets the event window in advance, and usually, its length does not exceed 30 days (citations). If we used the entire 71-day-long lockdown window, the expectation would

naturally be for a robust significant rejection of the null hypothesis and thus for a reduction in the average level of the concentrations. This expectation is confirmed by Figure S3 in Supplementary Material, which shows that during the lockdown period in 2020, all stations have concentrations well below those observed during the same period in previous years. Notice that given the historical magnitude of the event, although the event window is very long, we could reasonably assume that there were no other overlapping events capable of masking the impact of the restrictive measures during the period.

Therefore, in our exercise, it is more interesting to identify the minimum time window that each model needs to detect a significant effect on the average level of concentrations. We then propose a recursive testing approach that makes use of an increasing event window to estimate ES test statistics. Conversely, the estimation window is constant for all models and all sequential tests. The estimation window is composed of all the measurements lying between January 1, 2018, and January 31, 2020, i.e., $T_0 = 761$ days. The recursive algorithm starts on the February 1, 2020, and ends on the May 31, 2020, i.e., the maximum event window length is $T_1 = 120$ days. The recursive algorithm estimates the ES test statistics adding one day per iteration up to the last time stamp in the event window. That is, at the generic iteration $\tau = 1, \dots, 120$, the event window set $\Omega_{1\tau}$ includes all the estimated abnormal concentrations ranging from 1 to τ , and the corresponding ES test statistics are computed using $\Omega_{1\tau}$.

5.2. COVARIATES AND CONTROLS

The geostatistical models presented in Sect. 4 allow including a large set of time- and site-specific predictors to model the airborne pollutant concentrations. As stated above, our aim is to achieve the maximum comparability possible between models; thus, both HDGM and GAMMs will include the same linear predictors, whereas the spatiotemporal dynamics are left model-specific. In particular, our model will consider (1) local weather variables, (2) calendar effects, and (3) land-cover variables. Weather and land-cover covariates included in the large-scale component of the models have been chosen among those available from the Copernicus ERA-5 reanalysis database (Sabater 2019). ERA-5 provides observations with a $0.1^\circ \times 0.1^\circ$ grid spatial resolution. For each air quality station, we associated the meteorological measurement observed in the cell where the station is located. To explain the airborne pollutant concentrations, we considered a set of nine meteorological and land-cover variables: average daily temperature ($^\circ C$), daily cumulative precipitation (mm), relative humidity (%), atmospheric pressure (Pa), daily average eastward and the northward component of the wind (m/s), daily maximum eastward and northward wind speed (m/s), geopotential height (m^2/s^2) as a proxy of altitude, and high and low vegetation covering (measured as one-half of the total green leaf area per unit horizontal ground surface area, cf., Sabater 2019). While the geopotential height and land cover are time-invariant, the weather covariates are all time-varying.

Pollutant concentrations are strongly seasonal phenomena. In particular, they consistently follow the cyclical pattern of climatic seasons. Statistically speaking, they are characterized by annual seasonality and intra-weekly seasonality. Usually, the temperature is used as a proxy for the climatic season and thus serves as a driver of the infra-annual cycle of

pollutants. However, linearity might not be enough. Thus, we adopt two further corrections. First, we allow for a flexible relationship between daily temperature and NO_2 concentrations including the covariate as a cubic B-spline expansion with $k_{Temp} = 3$ basis. Second, we add as a proxy of the annual cycle a periodic Fourier spline with two harmonics, thus having $k_{Fourier} = 4$ basis (Ramsay and Silverman 2005). Similarly, we allow a flexible relationship among average wind speed and concentrations by including both eastward and northward components as cubic regression splines with $k_{Wind} = 3$ basis. This choice allows the model to detect possible anomalies (mainly storms or high wind days) in the wind movements, reducing the presence of outlier concentrations in the residuals.

Due to the strong seasonality, airborne pollutant concentrations are also characterized by strong autocorrelation, even for high temporal distances. In the case of NO_2 , the use of temperature might be not sufficient to suitably capture the temporal correlation. To resolve this issue, we included short- and long-term lags of the remaining time-varying covariates. In detail, we included the 1-day, 2-day, and 365-day lagged values of rainfall, pressure, daily maximum wind speed, and relative humidity among the regressors. Altogether, excluding the intercept, the total number of linear covariates is 44.

In addition to weather parameters, we included two dummy variables controlling for calendar effects. In particular, we included a dummy for the weekend effect, which allows us to control for typical reductions observed during the weekend, and a dummy accounting for the main Italian holidays across the year. Also, as suggested in Fassó et al. (2021), we included two sets of dummy variables controlling for local conditions near the monitoring site (i.e., station type) and for large-scale geographical conditions surrounding the air quality station (i.e., type of surrounding area). The station-type variables classified the monitoring sites as traffic, rural, industrial, and background (reference category), whereas the surrounding conditions classify the stations as metropolitan areas, mountains, urbanized plain, and rural plain (reference category).

5.3. GEOSTATISTICAL MODELING AND DIAGNOSTICS

In this section, we comment on the main results referring to the performance and diagnostics of the models in the estimation window. Extended results are available in Section S3 of Supplementary Materials. In particular, we provide estimates of the spatiotemporal parameters and further performance metrics.

In Figs. 2 and 3, we show the abnormal concentrations estimated using the four geostatistical models around the event window. Specifically, Fig. 2 shows the ACs by station type (i.e., local-scale conditions around the site), whereas Fig. 3 shows the ACs by surrounding area (i.e., large-scale conditions). We plot the ACs in the last part of the estimation window (blue-shaded areas) and during the lockdown period (red-shaded areas). Abnormal concentrations are computed by back-transforming the log-scaled concentrations to $\mu\text{g}/\text{m}^3$.

The insights provided by the charts are manifold. First, both classifications are mutually consistent and show that urbanized areas (traffic sites in Fig. 2 and metropolitan areas in Fig. 3) experienced the greatest reductions in NO_2 , reaching average values of $-20\mu\text{g}/\text{m}^3$ at the height of the lockdown (April 2020). This finding is consistent with the restrictions of the movement of transport vehicles, the primary source of nitrogen dioxide. In contrast, sites

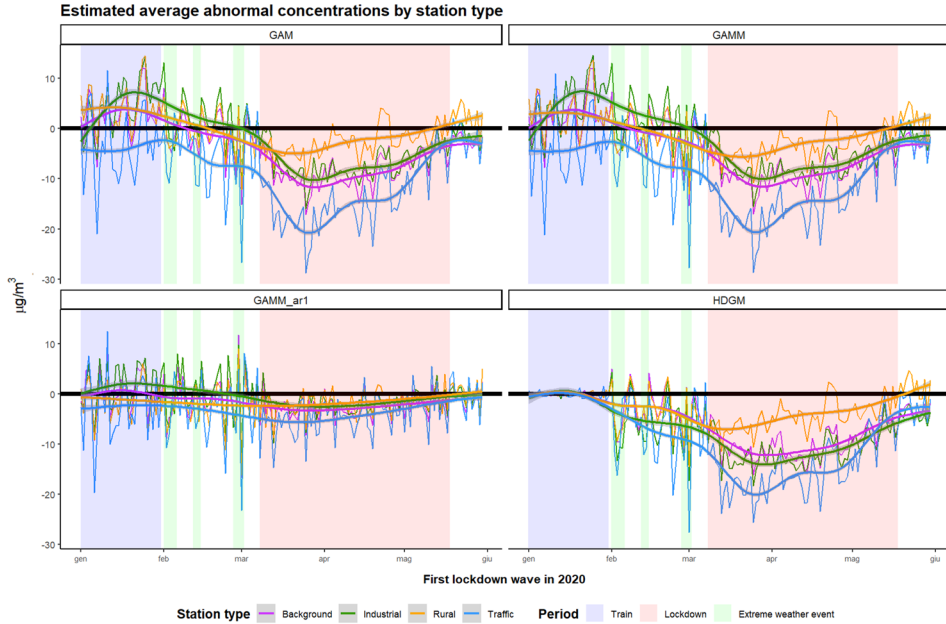


Figure 2. Estimated abnormal concentrations (back-transformed to $\mu\text{g}/\text{m}^3$) aggregated (average) by station type and model: GAM (upper left panel), GMM (upper right panel), GAMM with AR(1) dynamics (lower left panel), and HDGM (lower right panel). Solid-colored curves are the smooth estimates of the trends (splines) by station type. The blue-shaded area identifies the last part of the training window; the red-shaded area identifies the lockdown period; and the green-shaded areas identify three extreme weather events occurring within the event window .

with less human presence (rural, mountain, and rural lowlands) registered more moderate falls ($-10\mu\text{g}/\text{m}^3$ to $-12\mu\text{g}/\text{m}^3$); moreover, average levels returned in line with predictions (null ACs) before the end of the lockdown (mid-May 2020). Overall, the estimated reductions are consistent with those provided in other studies for Lombardy, such as [Lonati and Riva \(2021\)](#) and [Bontempi et al. \(2022\)](#). Second, the values estimated by the respective models do not show substantial differences in their trend or shape. In particular, ACs from GAM, GMM, and HDGM show strong overlaps. The GAMM-ar1 model, on the other hand, estimates reductions consistent with the classifications, but significantly smaller. One potential explanation lies in the fact that GAMM-ar1 includes the autoregressive term in the site-specific residuals; thus, it is capable of rapidly adapting to level shifts while absorbing the event-generated effect. Third, all models show some notable reductions prior to the establishment of the lockdown. These reductions, highlighted in the green areas, coincide with three extreme weather events recorded in Lombardy in February 2020. These events relate to sudden and large increases in atmospheric boundary layer height ([Fassó et al. 2023](#)), which increased local wind speeds and increased air recycling removing concentrations. In Figure S2 in Supplementary Materials, we show estimates of BHL at some sites monitored by the ARPA Lombardy Agency showing the peaks at the NO_2 fall. Finally, we note that the variability of the HDGM estimation-window residuals is very low when compared to that of the event-window residuals. This fact suggests that the HDGM may suffer from overfitting. However, the out-of-sample behavior of its residuals is similar to those of the GAM and

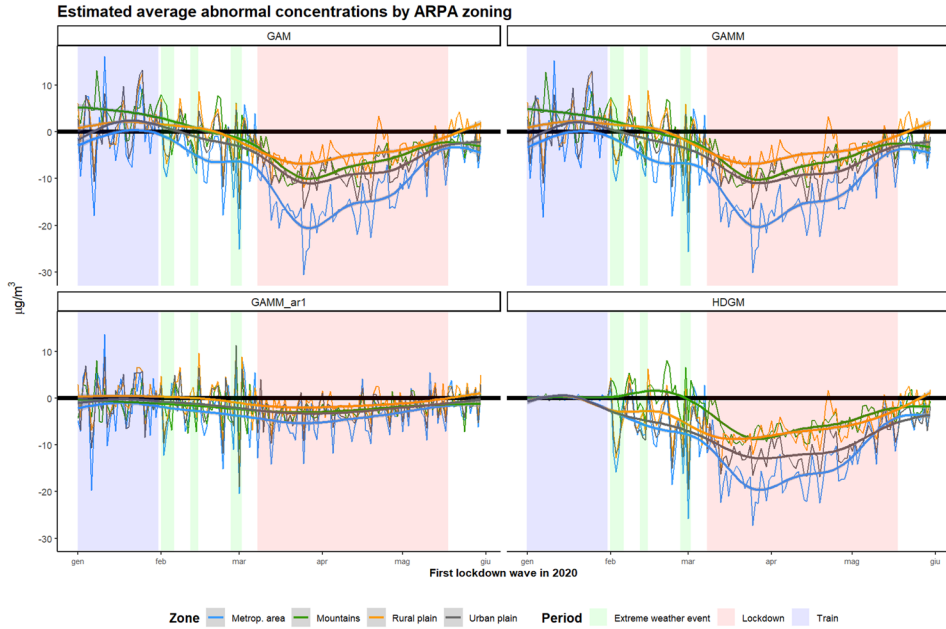


Figure 3. Estimated abnormal concentrations (back-transformed to $\mu\text{g}/\text{m}^3$) aggregated (average) by surrounding conditions and model: GAM (upper left panel), GMM (upper right panel), GAMM with AR(1) dynamics (lower left panel), and HDGM (lower right panel). Solid-colored curves are the smooth estimates of the trends (splines) by the surrounding area. The blue-shaded area identifies the last part of the training window; red-shaded area identifies the lockdown period; and the green-shaded areas identify three extreme weather events occurring within the event window.

GAMM models and the ES tests come to the same conclusions. Future developments should take into account this issue by implementing regularizations in the estimation process of the HDGM, as proposed in [Maranzano et al. \(2023\)](#).

In Fig. 4, we report spatiotemporal diagnostics useful to understand how to propose geostatistical models fit the observed concentrations and in particular, if they are able to handle the spatial and temporal dependences. On the left, we show the boxplots of ACF patterns from lag 1 day to 30 days, while on the right we depict the site-specific distribution of the pairwise linear correlation. Regarding the temporal dimension, the plots highlight that, with the only exception of the GAM, the models adequately capture the temporal correlation as it is on average very close to zero with little dispersion. However, the ACs from GAM are still strongly characterized by temporal persistence, especially in the short run. Regarding spatial dependence, the only model which is able to decrease it toward zero is the HDGM. In fact, all the stations have a distribution surrounding the null value, while the GAMM results in strongly linearly correlated ACs.

5.4. RECURSIVE TEST STATISTICS FOR NO_2 CONCENTRATIONS

Here, we present and discuss the results obtained from the recursive ES experiment in which the event window is iteratively expanded. The findings are summarized in Fig. 5.

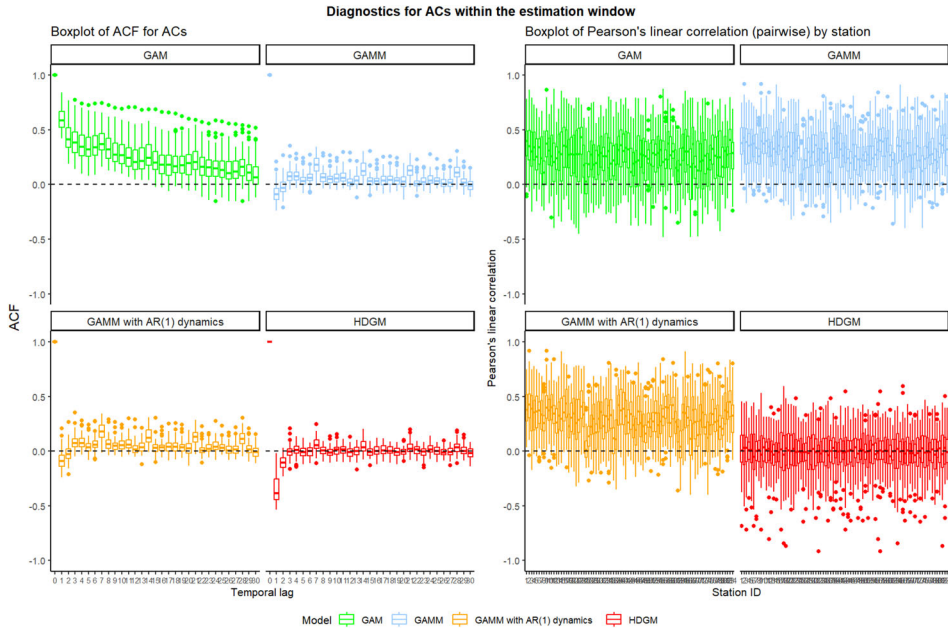


Figure 4. Autocorrelation and cross-sectional dependence diagnostics for abnormal concentrations (ACs) within the estimation window. Left panel: boxplots (computed across the 84 stations) of the ACF up to lag 30 days. Right panel: boxplot (computed across the 84 stations) of the pairwise Pearson's linear correlation index by station ID.

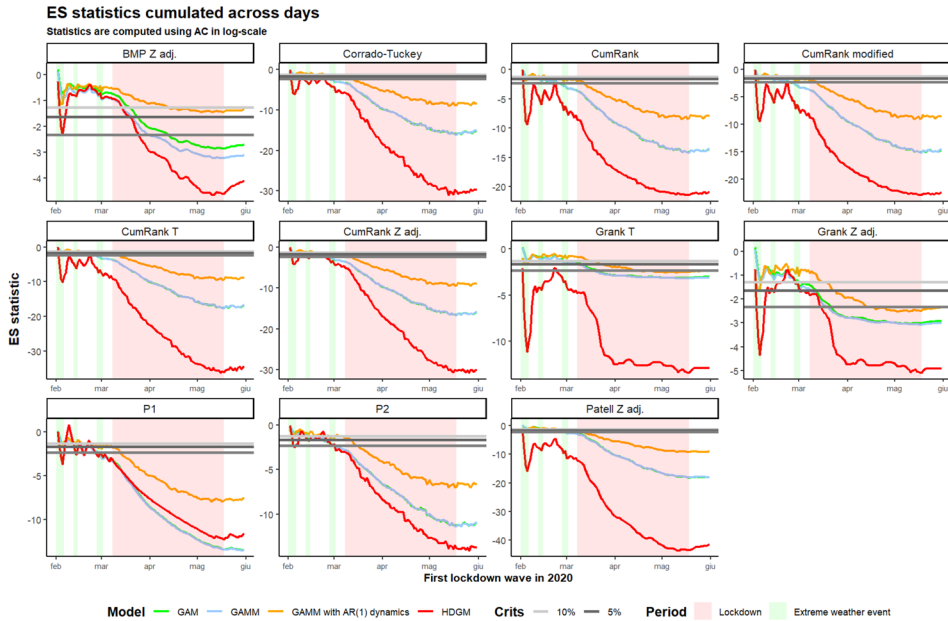


Figure 5. ES statistics using a recursive window from February 1 to May 31, 2020, by model. Red-shaded area identifies the lockdown period; green-shaded areas identify three extreme weather events occurring within the event window.

In February, many of the test statistics tend to fluctuate between significant and non-significant values. Indeed, at the real beginning of the recursive window, the recursive sample is still small and sensitive to single extreme cases. All test statistics are able to identify BLH-related extreme events as abnormal changes, which are immediately absorbed by the rebounds of the following days.

The permanent change takes place only at the beginning of March. In fact, one can notice that all the ES statistics start drifting between March 6 and 10, 2020, in correspondence with the actual enforcement of the lockdown restriction. Also, as the recursive window becomes large, the estimated values of the test statistics tend to stabilize. In this specific case, this occurs toward the beginning of April 2020, a time when the lockdown has already been active for weeks and the concentrations settle down at values strongly below the predictions. From the ES standpoint, this means that in the mid-stage of the lockdown, the change becomes persistent, and all test statistics confidently identify the presence of a level shift and not a sudden outlier event. Finally, as the reopening phase begins in mid-May 2020, they begin a slight upward phase that in the long run will lead to the absorption of the shock.

Finally, we notice that all four statistical models are fully equivalent in determining a statistically significant negative shock. Nevertheless, some of the statistics (e.g., Patell-Z, CumRank, and CumRank modified) computed on the HDGM prediction errors identify the onset of the shock well in advance. Also, we observe how the P1, P2, and Corrado–Tukey statistics are very robust tools for both the identification of isolated structural breaks (e.g., weather events) and structural changes in the level, while others (see, for instance, the BMP-adjusted or GRank-Z statistics) struggle in identifying abnormal events.

6. CONCLUSIONS AND FUTURE DEVELOPMENTS

In this paper, we contributed to the empirical literature on ES by proposing a twofold adjustment for Event Studies considering spatiotemporal data. In particular, we analyzed the case of ES applied to multivariate time series characterized by the presence of spatial (i.e., cross-sectional) and temporal dependence.

The first adjustment concerns the modeling step. Previously existing literature showed that the presence of confounding spatiotemporal correlation in the regression residuals can adversely affect the values of the statistics and their uncertainty. Thus, we proposed to model explicitly the spatiotemporal dynamics of the data by implementing several geostatistical models capable of handling both spatial and temporal components, as well as estimating the relationship between the response variable and a set of exogenous factors. In particular, we considered LMMs and GAMMs with spatial and temporal components. The second adjustment refers to the Event Studies hypothesis testing step. From the literature, we know that when the observations are affected by positive spatial dependence (even by small amounts), classical ES parametric test statistics are unreliable. We then proposed to use cross-sectional-adjusted test statistics directly accounting for spatial cross-sectional dependence by means of a cross-correlation measure for multivariate time series.

The proposed adjustments were applied to the case study of NO₂ concentrations in Lombardy, Northern Italy. In particular, we considered as the event of interest the lockdown

restrictions imposed on citizenship during the first wave of the COVID-19 pandemic. The main interest was to state if the lockdown generated significant reductions in the average concentrations of NO₂, i.e., we tested for a level shift after the event date.

The key findings can be summarized as follows. First, the reductions in the level of NO₂ concentrations provided by the geostatistical models are consistent with the characteristics of the Lombardy region. In particular, the largest reductions are estimated in the major metropolitan and congested areas, while smaller reductions are estimated in rural plains and in the mountains. Second, the proposed models are nearly equivalent both in terms of fitting and identifying the true event window (recursive experiment). Third, the adoption of models with spatial and temporal components ensures residuals that are cleaned from spatiotemporal correlation, thus allowing ES test statistics to provide reliable and realistic estimates. Fourth, as expected, all test statistics show significant reductions in NO₂ concentrations starting from the first few days of lockdown.

Overall, the very positive performance of the geostatistical models and the consistency of the test statistics demonstrate the adequacy of the proposed tools and point out the need to adopt corrections for spatial and temporal dependence in an Event Studies framework with spatiotemporal data.

We focused on modeling NO₂ concentrations using univariate spatiotemporal models. However, multivariate models could be implemented to take advantage of the cross-correlation among mutually correlated response variables to further improve predictions in the event window (Fassó et al. 2021; Ferreira et al. 2022). Furthermore, the ES test statistics could be explicitly adjusted for spatial cross-correlation (Chen 2015) and spatiotemporal cross-correlation (Ma et al. 2006; Gao et al. 2019) measures. Eventually, future works in ES, being strictly related to forecasting, should address the possible issues of models overfitting across time and space.

ACKNOWLEDGEMENTS

The authors would like to thank Andrea Algieri, PhD, from the Lombardy's Regional Agency for Environmental Protection (ARPA Lombardia) for providing data on HBL and for helping in building the empirical strategy and interpreting the results.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. This research was funded by Fondazione Cariplo under the grant 2020-4066 'AgriImOnIA: the impact of agriculture on air quality and the COVID-19 pandemic' from the 'Data Science for Science and Society' program.

Authors' contributions P.M. contributed to Section 1 'Introduction' (design, original writing, review), Section 2 'Event Studies for environment and energy: state of the art' (design, original writing, review), Section 3 'Event Studies for air quality assessment: taxonomy and statistics' (design, development of the definitions/taxonomy, original writing, review), Section 4 'Geostatistical models for air quality' (design, development of the empirical strategy, original writing, review), Section 5 'Assessing the impact of COVID-19 lockdown measures on air quality in Lombardy' (design, data collection, computation/code with R and MATLAB, original writing, review), and Section 6 'Conclusions and future developments' (design, original writing, review). M.P. was involved in Section 1 'Introduction' (review, supervision), Section 2 'Event Studies for environment and energy: state of the art' (review, supervision), Section 3 'Event Studies for air quality assessment: taxonomy and statistics' (design, review, supervision), Section 4 'Geostatistical models for air quality' (review), Section 5 'Assessing the impact of COVID-19 lockdown measures on air quality in Lombardy' (design, review), and Section 6 'Conclusions and future developments' (design, review).

Data Availability All the results presented in this paper can be reproduced using R and MATLAB software. Data and code have been uploaded to the journal's database. Data and code are available at the following GitHub web page: https://github.com/PaoloMaranzano/PM_MP_ESforAQ_JABES.git.

Declarations

Conflict of interest the authors declare that they have no competing interests

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

[Received October 2022. Revised May 2023. Accepted July 2023.]

REFERENCES

- Baldasano JM (2020) Covid-19 lockdown effects on air quality by no₂ in the cities of Barcelona and Madrid (Spain). *Sci Total Environ* 741(140):353. <https://doi.org/10.1016/j.scitotenv.2020.140353>
- Barnett MD (2019) A run on oil: climate policy, stranded assets, and asset prices. Thesis
- Basseville M, Nikiforov I (1993) Detection of abrupt change theory and application, vol 15. PTR Prentice-Hall
- Benninga S (2014) Financial modeling. MIT press
- Bergmeir C, Hyndman RJ, Koo B (2018) A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Comput Stat Data Anal* 120:70–83. <https://doi.org/10.1016/j.csda.2017.11.003>
- Birindelli G, Chiappini H (2021) Climate change policies: Good news or bad news for firms in the European union? *Corp Soc Responsib Environ Manag* 28(2):831–848. <https://doi.org/10.1002/csr.2093>
- Bontempi E, Carnevale C, Cornelio A et al (2022) Analysis of the lockdown effects due to the Covid-19 on air pollution in Brescia (Lombardy). *Environ Res* 212(113):193. <https://doi.org/10.1016/j.envres.2022.113193>
- Borghesi S, Castellini M, Comincioli N et al (2022) European green policy announcements and sectoral stock returns. *Energy Policy* 166(113):004. <https://doi.org/10.1016/j.enpol.2022.113004>
- Cabrera M, Taylor G (2019) Modelling spatio-temporal data of dengue fever using generalized additive mixed models. *Spatial Spatio-temporal Epidemiol* 28:1–13. <https://doi.org/10.1016/j.sste.2018.11.006>
- Calculli C, Fassó A, Finazzi F et al (2015) Maximum likelihood estimation of the multivariate hidden dynamic geostatistical model with application to air quality in apulia, italy. *Environmetrics* 26(6):406–417
- Cameletti M (2020) The effect of corona virus lockdown on air pollution: Evidence from the city of Brescia in Lombardia region (Italy). *Atmos Environ* 239(117):794. <https://doi.org/10.1016/j.atmosenv.2020.117794>
- Campbell JY, Lo AW, MacKinlay AC et al (1998) The econometrics of financial markets. *Macroecon Dyn* 2(4):559–562
- Chen Y (2013) New approaches for calculating Moran's index of spatial autocorrelation. *PLOS ONE* 8(7):e68336. <https://doi.org/10.1371/journal.pone.0068336>
- Chen Y (2015) A new methodology of spatial cross-correlation analysis. *PLOS ONE* 10(5):e0126158. <https://doi.org/10.1371/journal.pone.0126158>
- Collivignarelli MC, Abbà A, Bertanza G et al (2020) Lockdown for covid-2019 in Milan: What are the effects on air quality? *Sci Total Environ* 732(139):280. <https://doi.org/10.1016/j.scitotenv.2020.139280>
- Corrado CJ (1989) A nonparametric test for abnormal security-price performance in event studies. *J Financ Econ* 23:385–395

- Corrado CJ, Zivney TL (1992) The specification and power of the sign test in event study hypothesis tests using daily stock returns. *J Financ Quant Anal* 27(3):465–478
- Dale MRT, Fortin MJ (2009) Spatial autocorrelation and statistical tests: some solutions. *J Agric Biol Environ Stat* 14(2):188–206. <https://doi.org/10.1198/jabes.2009.0012>
- Demirer R, Kutun AM (2010) The behavior of crude oil spot and futures prices around opec and spr announcements: an event study perspective. *Energy Econ* 32(6):1467–1476. <https://doi.org/10.1016/j.eneco.2010.06.006>
- Diaz-Rainey I, Gehricke SA, Roberts H et al (2021) Trump vs. paris: The impact of climate policy on u.s. listed oil and gas firm returns and volatility. *Int Rev Financ Anal* 76:101746. <https://doi.org/10.1016/j.irfa.2021.101746>
- Djoundourian S, Marrouch W, Sayour N (2022) Adaptation funding and greenhouse gas emissions: Halo effect or complacency? *Energy J* 43(4):215–230. <https://doi.org/10.5547/01956574.43.4.sdjo>
- Dutilleul P, Clifford P, Richardson S et al (1993) Modifying the t test for assessing the correlation between two spatial processes. *Biometrics* 49(1):305–314. <https://doi.org/10.2307/2532625>
- Fassò A, Maranzano P, Otto P (2021) Spatiotemporal variable selection and air quality impact assessment of Covid-19 lockdown. *Spatial Stat.* <https://doi.org/10.1016/j.spasta.2021.100549>
- Fassò A, Rodeschini J, Fusta Moro A et al (2023) Agrimonia: a dataset on livestock, meteorology and air quality in the Lombardy region, Italy. *Sci Data* 10(1):143. <https://doi.org/10.1038/s41597-023-02034-0>
- Feng C (2022) Spatial-temporal generalized additive model for modeling Covid-19 mortality risk in Toronto, Canada. *Spatial Stat* 49(100):526. <https://doi.org/10.1016/j.spasta.2021.100526>
- Ferreira G, Mateu J, Porcu E (2022) Multivariate Kalman filtering for spatio-temporal processes. *Stoch Environ Res Risk Assess.* <https://doi.org/10.1007/s00477-022-02266-3>
- Finazzi F, Fassò A (2020) The impact of the Covid-19 pandemic on Italian mobility. *Significance* (Oxford, England) 17(3):17
- Gao Y, Cheng J, Meng H et al (2019) Measuring spatio-temporal autocorrelation in time series data of collective human mobility. *Geo-spatial Inf Sci* 22(3):166–173. <https://doi.org/10.1080/10095020.2019.1643609>
- Giacotto C, Sfiridis JM (1996) Hypothesis testing in event studies: the case of variance changes. *J Econ Bus* 48(4):349–370. [https://doi.org/10.1016/0148-6195\(96\)00019-7](https://doi.org/10.1016/0148-6195(96)00019-7)
- Granello F, Reis LA, Bosetti V et al (2021) Covid-19 lockdown only partially alleviates health impacts of air pollution in northern Italy. *Environ Res Lett* 16(3):035012
- Hagnäs T, Pynnönen S (2014) Testing for cumulative abnormal returns in event studies with the rank test. Available at SSRN 2479228
- Higham J, Ramírez CA, Green M, et al (2020) UK covid-19 lockdown: 100 days of air pollution reduction? Air quality, atmosphere and health pp 1–8
- Hu S, Xiong C, Younes H et al (2022) Examining spatiotemporal evolution of racial/ethnic disparities in human mobility and Covid-19 health outcomes: Evidence from the contiguous united states. *Sustain Cities Soc* 76(103):506. <https://doi.org/10.1016/j.scs.2021.103506>
- Huynh TD, Xia Y (2020) Climate change news risk and corporate bond returns. *J Financ Quant Anal* 56(6):1985–2009. <https://doi.org/10.1017/S0022109020000757>
- Jurek M, Katzfuss M (2022) Hierarchical sparse cholesky decomposition with applications to high-dimensional spatio-temporal filtering. *Stat Comput* 32(1):15. <https://doi.org/10.1007/s11222-021-10077-9>
- Jurek M, Katzfuss M (2023) Scalable spatio-temporal smoothing via hierarchical sparse Cholesky decomposition. *Environmetrics* 34(1):e2757. <https://doi.org/10.1002/env.2757>
- Kammann EE, Wand MP (2003) Geoadditive models. *J R Stat Soc Series C (Appl Stat)* 52(1):1–18. <https://doi.org/10.1111/1467-9876.00385>
- Kolari JW, Pynnönen S (2011) Nonparametric rank tests for event studies. *J Empir Financ* 18(5):953–971. <https://doi.org/10.1016/j.jempfin.2011.08.003>
- Lee SI (2001) Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's i. *J Geogr Syst* 3(4):369–385
- Lee J, Lund R (2004) Revisiting simple linear regression with autocorrelated errors. *Biometrika* 91(1):240–245. <https://doi.org/10.1093/biomet/91.1.240>

- Lee J, Lund R (2008) Equivalent sample sizes in time series regressions. *J Stat Comput Simul* 78(4):285–297. <https://doi.org/10.1080/10629360600758484>
- Li B, Wang F, Yin H et al (2019) Mega events and urban air quality improvement: A temporary show? *J Clean Prod* 217:116–126. <https://doi.org/10.1016/j.jclepro.2019.01.116>
- Li H, Zhang L, Chen T et al (2022) Environmental and health impacts of heating fuel transition: evidence from northern China. *Energy Build* 276(112):483. <https://doi.org/10.1016/j.enbuild.2022.112483>
- Lin B, Zhu J (2019) Is the implementation of energy saving and emission reduction policy really effective in Chinese cities? A policy evaluation perspective. *J Clean Prod* 220:1111–1120. <https://doi.org/10.1016/j.jclepro.2019.02.209>
- Lonati G, Riva F (2021) Regional scale impact of the Covid-19 lockdown on air quality: gaseous pollutants in the PO valley, northern Italy. *Atmosphere* 12(2):264
- Lovarelli D, Conti C, Finzi A et al (2020) Describing the trend of ammonia, particulate matter and nitrogen oxides: the role of livestock activities in northern Italy during covid-19 quarantine. *Environ Res* 191(110):048. <https://doi.org/10.1016/j.envres.2020.110048>
- Lovarelli D, Fugazza D, Costantini M et al (2021) Comparison of ammonia air concentration before and during the spread of Covid-19 in Lombardy (Italy) using ground-based and satellite data. *Atmos Environ* 259(118):534. <https://doi.org/10.1016/j.atmosenv.2021.118534>
- Luoma T (2011) Nonparametric event study tests for testing cumulative abnormal returns. *Acta Wasaensia* 254
- Maranzano P (2022) Air quality in lombardy, Italy: An overview of the environmental monitoring system of arpa lombardia. *Earth* 3(1):172–203
- Maranzano P, Fassó A (2022) The impact of the lockdown restrictions on air quality during COVID-19 pandemic in Lombardy, Italy. In: Steland A, Tsui K-L (eds) *Artificial intelligence, big data and data science in statistics: challenges and solutions in environmetrics, the natural sciences and technology*. Springer International Publishing, Cham, pp 343–374
- Maranzano P, Fassó A, Pelagatti M et al (2020) Statistical modeling of the early-stage impact of a new traffic policy in Milan, Italy. *Int J Environ Res Public Health* 17(3):1088
- Maranzano P, Otto P, Fassó A (2023) Adaptive lasso estimation for functional hidden dynamic geostatistical model. *Stoch Environ Res Risk Assess*. <https://doi.org/10.1007/s00477-023-02466-5>
- Ma J, Zeng D, Chen H (2006) Spatial-temporal cross-correlation analysis: A new measure and a case study in infectious disease informatics. In: Mehrotra S, Zeng DD, Chen H et al (eds) *Intelligence and Security Informatics*. Springer, Berlin Heidelberg, pp 542–547
- McDowall D, McCleary R, Bartos BJ (2019) *Interrupted time series analysis*. Oxford University Press
- Montero JM, Fernández-Avilés G, Laureti T (2021) A local spatial Stirpat model for outdoor NOX concentrations in the community of Madrid, Spain. *Mathematics* 9(6):677
- Moran PAP (1950) Notes on continuous stochastic phenomena. *Biometrika* 37(1/2):17–23. <https://doi.org/10.2307/2332142>
- Mudelsee M, Alkio M (2007) Quantifying effects in two-sample environmental experiments using bootstrap confidence intervals. *Environ Modell Softw* 22(1):84–96. <https://doi.org/10.1016/j.envsoft.2005.12.001>
- Nakada LYK, Urban RC (2020) Covid-19 pandemic: impacts on the air quality during the partial lockdown in São Paulo state, Brazil. *Sci Total Environ* 730(139):087. <https://doi.org/10.1016/j.scitotenv.2020.139087>
- Naqvi A (2021) Decoupling trends of emissions across EU regions and the role of environmental policies. *J Clean Prod* 323(129):130. <https://doi.org/10.1016/j.jclepro.2021.129130>
- Neill CL, Chen SE (2022) Food safety events versus media: nonlinear effects of egg recalls on US egg prices. *J Agric Res Econ* 47(1):23–37
- Otto P, Piter A, Gijsman R (2021) Statistical analysis of beach profiles: a spatiotemporal functional approach. *Coast Eng* 170(103):999. <https://doi.org/10.1016/j.coastaleng.2021.103999>
- Paciorek CJ (2010) The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Stat Sci Rev J Inst Math Stat* 25(1):107

- Padilla CM, Kihal-Talantikite W, Vieira VM et al (2014) Air quality and social deprivation in four French metropolitan areas—a localized spatio-temporal environmental inequality analysis. *Environ Res* 134:315–324. <https://doi.org/10.1016/j.envres.2014.07.017>
- Pelagatti M, Maranzano P (2021) Assessing the effectiveness of the Italian risk-zones policy during the second wave of Covid-19. *Health Policy* 125(9):1188–1199. <https://doi.org/10.1016/j.healthpol.2021.07.011>
- Pelagatti M, Maranzano P (2021) Nonparametric tests for event studies under cross-sectional dependence. *Q J Finance Account* 59:29
- Pinheiro J, Bates D (2006) *Mixed-effects models in S and S-PLUS*. Springer science & business media
- Piter A, Otto P, Alkhatib H (2022) The Helsinki bike-sharing system—insights gained from a spatiotemporal functional model. *J R Stat Soc Ser A* 185(3):1294–1318
- Presidenza del Consiglio dei Ministri Italia (2020) Decreto del presidente del consiglio dei ministri 8 marzo 2020. Report, *Gazzetta Ufficiale della Repubblica Italiana*, <https://www.gazzettaufficiale.it/eli/id/2020/03/08/20A01522/sg>
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, <https://www.R-project.org/>
- Ramsay J, Silverman B (2005) *Functional data analysis*. Springer Series in Statistics, Springer New York, NY, <https://doi.org/10.1007/b98888>
- Rossi R, Ceccato R, Gastaldi M (2020) Effect of road traffic on air pollution. Experimental evidence from Covid-19 lockdown. *Sustainability* 12(21):8984
- Sabater M (2019) Era5-land hourly data from 1981 to present. Copernicus climate change service (c3s) climate data store (cds). Accessed on 31 Jan 2022 <https://doi.org/10.24381/cds.e2161bac>. Report
- Taghavi-Shahri SM, Fassó A, Mahaki B, et al (2019) Concurrent spatiotemporal daily land use regression modeling and missing data imputation of fine particulate matter using distributed space-time expectation maximization. *Atmospheric Environment* p 117202
- Tobler WR (1970) A computer movie simulating urban growth in the detroit region. *Econ Geogr* 46(sup1):234–240. <https://doi.org/10.2307/143141>
- Wang Y, Finazzi F, Fassó A (2021) D-stem v2: a software for modeling functional spatio-temporal data. *J Stat Softw* 99(10):1–29. <https://doi.org/10.18637/jss.v099.i10>
- Weng Z, Wang Y, Yang X et al (2022) Effect of cleaner residential heating policy on air pollution: a case study in Shandong province, China. *J Environ Manage* 311(114):847. <https://doi.org/10.1016/j.jenvman.2022.114847>
- Wood SN (2017) *Generalized additive models: an introduction with R*, 2nd edn. Chapman and Hall/CRC
- Xiao B, Yin W, Zhu Z (2022) Does the air quality benefit from lockdown policy? evidence from major cities in china. In: *Advances in transdisciplinary engineering*, pp 683–693, <https://doi.org/10.3233/ATDE220341>
- Xin Y, Shao S, Wang Z et al (2021) Covid-2019 lockdown in Beijing: a rare opportunity to analyze the contribution rate of road traffic to air pollutants. *Sustain Cities Soc* 75(102):989. <https://doi.org/10.1016/j.scs.2021.102989>
- Xu H, Liang W, Xiang K (2022) The environmental consequences of place-based policies in china: an empirical study based on so₂ emission data. *China World Econ* 30(4):201–229. <https://doi.org/10.1111/cwe.12433>
- Zangari S, Hill DT, Charette AT et al (2020) Air quality changes in New York city during the covid-19 pandemic. *Sci Total Environ* 742(140):496. <https://doi.org/10.1016/j.scitotenv.2020.140496>
- Zha D, Zhao T, Kavuri AS et al (2018) An event study analysis of price adjustment of refined oil and air quality in China. *Environ Sci Pollut Res* 25(34):34236–34246. <https://doi.org/10.1007/s11356-018-3374-3>
- Zhang X, Yu L, Wang S et al (2009) Estimating the impact of extreme events on crude oil price: an emd-based event analysis method. *Energy Econ* 31(5):768–778. <https://doi.org/10.1016/j.eneco.2009.04.003>
- Zhu AX, Turner M (2022) How is the third law of geography different? *Ann GIS* 28(1):57–67. <https://doi.org/10.1080/19475683.2022.2026467>
- Zimmerman DW (2012) Correcting two-sample z and t tests for correlation: an alternative to one-sample tests on difference scores. *Psicol Int J Methodol Exp Psychol* 33(2):391–418