Econometrics and Statistics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

A Dynamic Spatiotemporal Stochastic Volatility Model with an Application to Environmental Risks

Philipp Otto [a,*], Osman Doğan [b], Süleyman Taşpınar [c]

[a] School of Mathematics and Statistics, University of Glasgow, United Kingdom
[b] Economics Department, Istanbul Technical University, Istanbul
[c] Department of Economics, Queens College, The City University of New York

ARTICLE INFO

Article history:
Received 3 August 2023
Revised 3 November 2023
Accepted 4 November 2023
Available online xxx

Keywords:
Environmental risk
MCMC
Spatial dependence
Spatiotemporal stochastic volatility
Air quality

ABSTRACT

A dynamic spatiotemporal stochastic volatility (SV) model is introduced, incorporating explicit terms accounting for spatial, temporal, and spatiotemporal spillover effects. Alongside these features, the model encompasses time-invariant site-specific factors, allowing for differentiation in volatility levels across locations. The statistical properties of an outcome variable within this model framework are examined, revealing the induction of spatial dependence in the outcome variable. Additionally, a Bayesian estimation procedure employing the Markov Chain Monte Carlo (MCMC) approach, complemented by a suitable data transformation, is presented. Simulation experiments are conducted to assess the performance of the proposed Bayesian estimator. Subsequently, the model is applied in the domain of environmental risk modeling, addressing the scarcity of empirical studies in this field. The significance of climate variation studies is emphasized, illustrated by an analysis of local air quality in Northern Italy during 2021, which underscores pronounced spatial and temporal clusters and increased uncertainties/risks during the winter season compared to the summer season.

1. Introduction

When analyzing geo-referenced data, statistical models have to account for instantaneous spatial correlations due to the geographical proximity between the observations. This is commonly known as Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). This observation was already noted by Ronald A. Fisher in 1935 as follows, "the widely verified fact that patches in close proximity are commonly more alike, as judged by the yield of crops, than those which are further apart" (Fisher, 1935). Even though the similarity is typically considered to be in the (conditional) mean level at each location, there might also be spatial correlations in the (conditional) variance or variation of the random process. In particular, for small-scale spatial units, the variance of the process is increased (known as Arbia's law of geography, Arbia and Espa, 1996). In addition to the instantaneous spatial correlations, we also have to account for the natural temporal correlations, which usually occur if we repeatedly observe a

* Corresponding author.
E-mail address: philipp.otto@glasgow.ac.uk (P. Otto).

https://doi.org/10.1016/j.ecosta.2023.11.002

random process over time. The closer two observations are to each other in time, the more strongly they can correlate in general. This paper introduces a new model for spatial, temporal, and spatiotemporal correlations in the log-volatilities, allowing for additional random errors in the mean and volatility equation. Furthermore, we apply this model to environmental data and show for the first time how it can be used to analyze environmental risk factors such as air pollution.

There are generally two ways to account for spatial/cross-sectional correlations in spatial statistics. Firstly, it can be modeled in the covariance matrix of the process, where each entry is supposed to follow a certain (non-)parametric covariance function depending on the distance between their locations. This idea is typically known as the geostatistical approach (Cressie and Wikle, 2015; Zimmerman, 2019). The selection of a suitable parametric covariance function with certain properties such as stationarity, separability and full symmetry is one of the main modeling issues of this approach (e.g., Porcu et al., 2016; Huang et al., 2011), but also the scalability of these approaches to handle massive geospatial data, e.g., remotely-sensed data (e.g., Katzfuss and Cressie, 2011; Banerjee, 2020; Jurek and Katzfuss, 2021). See Gneiting et al. (2007) for a review on the spatiotemporal covariance functions suggested in the literature. Secondly, the observations on an outcome variable can be explicitly correlated with the adjacent observations, where the adjacency is defined fairly generally by a spatial weights matrix. This second approach is referred to as spatial autoregression in spatial econometrics, where the spatial lags of variables are used to model spatial correlations. See LeSage and Pace (2009); Anselin (1988); Elhorst (2014); Lee (2004); Kelejian and Prucha (2010) on the specification and estimation issues in spatial econometrics. Both approaches can be equivalent under certain conditions (see, e.g., Ver Hoef et al., 2018a, for simultaneous and conditionally autoregressive models).

In this paper, we consider the second approach. Our model consists of an outcome and log-volatility equation with separate independent error terms, whereby the log-volatility process introduces spatial dependence in the outcome variable. Specifically, the log-volatility equation allows for spatial, temporal, and spatiotemporal correlations, as well as time-invariant site-specific effects (unobserved heterogeneity). Also, assuming that the error terms in both equations have normal distributions, it is possible to show that the outcome variable has a leptokurtic symmetric distribution under our suggested model. To introduce a Bayesian estimation approach, we use a transformation approach such that the outcome equation becomes linear in the log-volatility terms. We use a Gaussian mixture distribution to approximate the distribution of the transformed error terms in the outcome equation. This approximation turns our model into a linear state-space model, where the log-volatility equation becomes the state equation. Following recent developments in the precision-based algorithms (Chan and Jeliazkov, 2009; Chan, 2017), we suggest a Gibbs sampler that consists of five steps for the estimation. We provide simulation evidence showing that the suggested sampler can perform satisfactorily.

Theoretically, our paper is related to the spatial econometric literature that addresses the presence of cross-sectional correlations in higher moments of spatial data. This strand of the literature considers the spatial extensions of generalized autoregressive conditional heteroskedasticity (GARCH) and stochastic volatility (SV) models to account for volatility clustering patterns observed over space (Otto et al., 2018; Hølleland and Karlsen, 2020b; Sato and Matsuda, 2021; Taşpınar et al., 2021; Robinson, 2009; Yan, 2007). Our model can be considered as the longitudinal data (panel data) extension of the cross-sectional spatial SV models suggested by Yan (2007), Robinson (2009), and Taşpınar et al. (2021). While these studies allow for the presence of spatial dependence in the log-volatility equations, they do not include temporal, spatiotemporal, and unobserved heterogeneity terms in the log-volatility equations. Our suggested process is also related to the separable and non-separable space-time filters considered in the spatial panel data models for modeling spatiotemporal interactions. For example, Parent and LeSage (2012, 2011) consider a separable space-time filter for the outcome variable, while Lee and Yu (2015) and Wang and Lee (2018) consider non-separable space-time filters for the disturbance terms of spatial panel data models. In contrast to these studies, we consider a general space-time filter that also allows for unobserved heterogeneity, i.e., the site-specific effects, for the log-volatility of an outcome variable.

In an empirical application, we use our suggested model to assess environmental risk stemming from the variations in the log-volatility of air quality predictions. Spatial and spatiotemporal interactions in the local climate and environmental risks were addressed in comparably few studies in previous empirical research, even though it has been shown that an increased variation in environmental processes can be harmful (Tewksbury et al., 2008; Paaijmans et al., 2013; Vasseur et al., 2014; Iaco et al., 2012). Most previous studies focused on correlations in the (conditional) mean levels of ecological processes (Ver Hoef et al., 2018b; Wilby et al., 2009, e.g.). In our case, we model the log-volatility of fine dust concentrations of particles having a diameter less than 10μ m in Lombardy, Northern Italy. Following the literature on the ecological processes (Ver Hoef et al., 2018b; Wilby et al., 2009, e.g.), we first model the variation in the conditional mean of our outcome variable through a conventional spatial panel data model that allows for the unobserved site and time heterogeneity. Our results from this initial model indicate that there is a strong and moderate spatial correlation in the outcome variable in the model with only site-fixed effects and the model with both site and time-fixed effects, respectively. In the next step, we use the errors from this initial model as an outcome variable in our suggested specification and aim to model the variations in its log-volatility terms. The estimation results from our suggested Bayesian approach indicate that the spatial and temporal effects in the log-volatilities are moderate, while the spatiotemporal effects appear to be of minor importance. We were also able to detect a noticeable variation in air pollution risk across the year and identify measurement stations that are associated with higher risks. These are mostly located in valleys in the Alpine regions.

The rest of this paper proceeds in the following way. In Section 2, we introduce our suggested model specification, including conditions ensuring the stability of the model and prior specifications. In Section 3, we investigate the statistical properties of the suggested model. In Section 4, we provide the details on the posterior analysis of our model and state

an algorithm for the estimation. Section 5 provides a simulation study on the performance of the suggested Gibbs sampler. Then, in Section 6, the stochastic volatility model is applied to environmental risks for the first time. More precisely, our focus is on local air quality modeling. Finally, in Section 7, we provide our concluding remarks.

2. Model Specification

Suppose that we observe the spatiotemporal process across a constant set of n locations in a geographical domain at T equidistant time points. These locations can be measurement stations (i.e., marked point data), atmospheric and remotely-sensed data, or sets of municipalities, counties, and states (i.e., areal data). Moreover, these locations do not have to be seen in a strictly geographical sense but can also be vertices in a network. Let $\mathbf{Y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ be the $n \times 1$ vector of the outcome variable at time t for $t = 1, 2, \dots, T$. We assume the following data-generating process (DGP) for $\mathbf{Y}_t$:

$$\mathbf{Y}_t = \mathbf{H}_t^{1/2}\mathbf{V}_t, \tag{1}$$

for $t = 1, 2, \dots T$, where $\mathbf{H}_t^{1/2} = \text{Diag}(e^{\frac{1}{2}h_{1t}}, \dots, e^{\frac{1}{2}h_{nt}})$ is the $n \times n$ diagonal matrix containing the log-volatility terms h_{it}'s, which are specified subsequently, and $\mathbf{V}_t = (v_{1t}, \dots, v_{nt})'$ is the $n \times 1$ vector of disturbance terms. We assume that v_{it}'s are i.i.d standard normal random variables. Let $\mathbf{h}_t = (h_{1t}, \dots, h_{nt})'$ be the $n \times 1$ vector of stochastic volatility at time t. We assume the following process for $\mathbf{h}_t$:

$$\mathbf{h}_t - \boldsymbol{\mu} = \rho_1\mathbf{W}(\mathbf{h}_t - \boldsymbol{\mu}) + \rho_2(\mathbf{h}_{t-1} - \boldsymbol{\mu}) + \rho_3\mathbf{W}(\mathbf{h}_{t-1} - \boldsymbol{\mu}) + \mathbf{U}_t, \tag{2}$$

for $t = 1, 2, \dots T$, where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$ is the $n \times 1$ vector of constant means, i.e., the time-invariant site-specific effects, and $\mathbf{U}_t = (u_{1t}, \dots, u_{nt})'$ is the $n \times 1$ vector of i.i.d. disturbance terms such that $u_{it} \sim N(0, \sigma^2)$ for all i and t. In (2), $\mathbf{W}$ is the $n \times n$ spatial weights matrix with zero diagonal elements. This matrix specifies how volatility terms are related over space. This matrix is equivalently specified as an adjacency matrix in a network setting. In practice, a suitable weights matrix can be chosen from a set of candidates based on the model's goodness-of-fit, e.g., assessed by the root-mean-square error (RMSE) or information criteria. Standard definitions can be included in the set of the candidate schemes, such as contiguity matrices, k-nearest-neighbors matrices, inverse-distance weights, or weights based on secondary covariates. Moreover, if the underlying physical processes causing the spatial dependence are known, they could be included in the definition of the weights matrix. The scalar parameter ρ_1 captures contemporaneous spatial correlation, ρ_2 measures the temporal effect, i.e., the time dynamic effect, and ρ_3 represents the spatiotemporal effect, i.e., the spatial diffusion effect.

Let $\mathbf{I}_n$ be the $n \times n$ identity matrix and $\mathbf{L}$ be the matrix of the time-lag operator such that $\mathbf{L}\mathbf{h}_t = \mathbf{h}_{t-1}$. Then, (2) can be written as

$$\big((\mathbf{I}_n - \rho_1\mathbf{W}) - (\rho_2\mathbf{L} + \rho_3\mathbf{W}\mathbf{L})\big)(\mathbf{h}_t - \boldsymbol{\mu}) = \mathbf{U}_t, \tag{3}$$

where $\big((\mathbf{I}_n - \rho_1\mathbf{W}) - (\rho_2\mathbf{L} + \rho_3\mathbf{W}\mathbf{L})\big)$ is called the general space-time filter (Parent and LeSage, 2011; 2012; Lee and Yu, 2015). Under the assumption that $\rho_3 = -\rho_1\rho_2$, this general filter is separable and decomposes into a product of the space filter $(\mathbf{I}_n - \rho_1\mathbf{W})$ and the time filter $(\mathbf{I}_n - \rho_2\mathbf{L})$. In our analysis, we do not impose this restrictive assumption.

Let $\mathbf{S}(\rho_1) = (\mathbf{I}_n - \rho_1\mathbf{W})$. Then, under the assumption that $\mathbf{S}(\rho_1)$ is invertible, the reduced form of volatility equation is

$$\mathbf{h}_t - \boldsymbol{\mu} = \mathbf{S}^{-1}(\rho_1)\mathbf{A}(\rho_2, \rho_3)(\mathbf{h}_{t-1} - \boldsymbol{\mu}) + \mathbf{S}^{-1}(\rho_1)\mathbf{U}_t, \tag{4}$$

where $\mathbf{A}(\rho_2, \rho_3) = (\rho_2\mathbf{I}_n + \rho_3\mathbf{W})$. When the cross-sectional dimension is fixed, the process for the log-volatility is stable if all eigenvalues of $\mathbf{S}^{-1}(\rho_1)\mathbf{A}(\rho_2, \rho_3)$ lie inside the unit ball (Hamilton, 1994, Proposition 10.1). Let $\vartheta_i(\mathbf{W})$ be the ith eigenvalue of $\mathbf{W}$ for $i = 1, \dots, n$. In the spatial econometric literature, there are alternative ways to specify the parameter spaces for spatial autoregressive parameters (see Anselin, 1988; LeSage and Pace, 2009; Lee, 2004; Kelejian and Prucha, 2010; Elhorst, 2014, among others). We assume that the parameter space of ρ_1, ρ_2 and ρ_3 are chosen such that the following conditions hold:

$$(i)\ \max_{1\le i\le n}\big|\vartheta_i(\rho_1\mathbf{W})\big| < 1, \quad \text{and } (ii)\ \max_{1\le i\le n}\big|\vartheta_i\big(\mathbf{S}^{-1}(\rho_1)\mathbf{A}(\rho_2, \rho_3)\big)\big| < 1. \tag{5}$$

The first condition is the sufficient condition for the invertibility of $\mathbf{S}(\rho_1)$ (Kelejian and Prucha, 2010, Lemma 1). By the spectral radius theorem (Horn and Johnson, 2012, Theorem 5.6.9), we can use any matrix norm to define relatively restrictive conditions that ensure the conditions in (5). Let $\|\cdot\|$ be any matrix norm. Then, the sufficient conditions for (5) are (i) $\|\rho_1\mathbf{W}\| < 1$ and (ii) $\|\mathbf{S}^{-1}(\rho_1)\mathbf{A}(\rho_2, \rho_3)\| < 1$. Note that

$$\begin{aligned}
\big\|\mathbf{S}^{-1}(\rho_1)\mathbf{A}(\rho_2, \rho_3)\big\| &\le \big\|\mathbf{S}^{-1}(\rho_1)\big\| \times \|\rho_2\mathbf{I}_n + \rho_3\mathbf{W}\| \\
&= \|\mathbf{I}_n + \rho_1\mathbf{W} + \rho_1^2\mathbf{W}^2 + \rho_1^3\mathbf{W}^3 + \dots\| \times \|\rho_2\mathbf{I}_n + \rho_3\mathbf{W}\| \\
&\le \big(\|\mathbf{I}_n\| + \|\rho_1\mathbf{W}\| + \|\rho_1\mathbf{W}\|^2 + \dots\big) \times \big(|\rho_2| + |\rho_3| \cdot \|\mathbf{W}\|\big) \\
&= \frac{1}{1 - \|\rho_1\mathbf{W}\|} \times \big(|\rho_2| + |\rho_3| \cdot \|\mathbf{W}\|\big),
\end{aligned} \tag{6}$$

where the last equality follows since we assume that $\|\rho_1\mathbf{W}\| < 1$. If we choose the matrix row sum norm $\|\cdot\|_\infty$ and assume that $\mathbf{W}$ is row normalized, then (6) reduces to $(|\rho_2| + |\rho_3|)/(1 - |\rho_1|)$. Thus, a further restrictive sufficient condition for the

stability of (4) is $|\rho_1|+|\rho_2|+|\rho_3|<1$. We will impose these restrictions during the sampling steps for ρ_1, ρ_2, and ρ_3 in our suggested Gibbs sampler.

Finally, to complete the model in (1), we assume the following prior distributions for the posterior analysis:

$$\begin{aligned}&\rho_1 \sim \text{Uniform}(-1,1), \quad \rho_2 \sim \text{Uniform}(-1,1),\ \rho_3 \sim \text{Uniform}(-1,1),\\ &\boldsymbol{\mu}|\mathbf{b}_\mu, \mathbf{B}_\mu \sim N(\mathbf{b}_\mu, \mathbf{B}_\mu),\ \sigma^2|a,b \sim \text{IG}(a,b),\end{aligned} \tag{7}$$

where Uniform(c_1, c_2) denotes the uniform distribution over the interval (c_1, c_2) and IG(a, b) denotes the inverse gamma distribution with the shape parameter a and the scale parameter b. The priors for ρ_1, ρ_2 and ρ_3 are subject to the stability conditions stated in (5). Our model specification can be extended by considering a mean process for the outcome variable. Such extensions pose no difficulty to the estimation process and can be easily integrated into our estimation algorithm provided in Section 4 by following Taşpınar et al. (2021, Section 5).

3. Statistical Properties

The outcome equation of our model can be written as

$$y_{it} = e^{\frac{1}{2}h_{it}}v_{it} \quad \text{for all } i=1,\ldots,n, \text{ and } t=1,\ldots,T. \tag{8}$$

Thus, the conditional variance of y_{it} given h_{it} is $\text{Var}(y_{it}|h_{it}) = e^{h_{it}}$, indicating that the conditional variance is both time and space varying. Following the time series literature, we refer to h_{it} as the log-volatility since $h_{it} = \log\big(\text{Var}(y_{it}|h_{it})\big)$. In order to determine the unconditional moments of y_{it}, we need to determine the distribution of $\mathbf{h} = (\mathbf{h}_1', \ldots, \mathbf{h}_T')'$. Let $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3)'$, and define the $nT \times nT$ matrix $\mathbf{J}(\boldsymbol{\rho})$ as

$$\mathbf{J}(\boldsymbol{\rho}) = \begin{pmatrix} \mathbf{S}(\rho_1) & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ -\mathbf{A}(\rho_2,\rho_3) & \mathbf{S}(\rho_1) & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & -\mathbf{A}(\rho_2,\rho_3) & \mathbf{S}(\rho_1) \end{pmatrix}. \tag{9}$$

Then, we can express the log-volatility equation as

$$\mathbf{J}(\boldsymbol{\rho})(\mathbf{h} - \boldsymbol{l}_T \otimes \boldsymbol{\mu}) = \begin{pmatrix} \mathbf{S}(\rho_1)(\mathbf{h}_1 - \boldsymbol{\mu}) \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_T \end{pmatrix}, \tag{10}$$

where $\boldsymbol{l}_T$ is the $T \times 1$ vector of ones. Using (4), we can express $\mathbf{S}(\rho_1)(\mathbf{h}_1 - \boldsymbol{\mu})$ as $\mathbf{S}(\rho_1)\big(\mathbf{h}_1 - \boldsymbol{\mu}\big) = \mathbf{A}(\rho_2,\rho_3)(\mathbf{h}_0 - \boldsymbol{\mu}) + \mathbf{U}_1$. We substitute $\mathbf{S}^{-1}(\rho_1)\big(\mathbf{h}_t - \boldsymbol{\mu}\big) = \mathbf{A}(\rho_2,\rho_3)(\mathbf{h}_{t-1} - \boldsymbol{\mu}) + \mathbf{U}_t$ for $t = 0, -1, -2, \ldots$, recursively into $\mathbf{S}(\rho_1)\big(\mathbf{h}_1 - \boldsymbol{\mu}\big) = \mathbf{A}(\rho_2,\rho_3)(\mathbf{h}_0 - \boldsymbol{\mu}) + \mathbf{U}_1$ to get the following expression:

$$\mathbf{S}(\rho_1)(\mathbf{h}_1 - \boldsymbol{\mu}) = \sum_{j=0}^{\infty} \big(\mathbf{A}(\rho_2,\rho_3)\mathbf{S}^{-1}(\rho_1)\big)^j \mathbf{U}_{1-j}, \tag{11}$$

This expression is based on the assumption that the spatial dynamic process for $\mathbf{h}_t$ has been operating for a long time under the stability conditions stated in (5). This result suggests that

$$\text{Var}(\mathbf{S}(\rho_1)(\mathbf{h}_1 - \boldsymbol{\mu})) = \sigma^2 \sum_{j=0}^{\infty} \big(\mathbf{A}(\rho_2,\rho_3)\mathbf{S}^{-1}(\rho_1)\big)^j \big(\mathbf{A}(\rho_2,\rho_3)\mathbf{S}^{-1}(\rho_1)\big)^{'j} = \sigma^2 \mathbf{K}(\boldsymbol{\rho}),$$

where $\mathbf{K}(\boldsymbol{\rho}) = \sum_{j=0}^{\infty} \big(\mathbf{A}(\rho_2,\rho_3)\mathbf{S}^{-1}(\rho_1)\big)^j \big(\mathbf{A}(\rho_2,\rho_3)\mathbf{S}^{-1}(\rho_1)\big)^{'j}$. Then, from (10), we obtain

$$\text{Var}\big((\mathbf{h} - \boldsymbol{l}_T \otimes \boldsymbol{\mu})\big) = \sigma^2 \mathbf{J}^{-1}(\boldsymbol{\rho}) \mathbf{P}(\boldsymbol{\rho}) \mathbf{J}^{'-1}(\boldsymbol{\rho}), \tag{12}$$

where

$$\mathbf{P}(\boldsymbol{\rho}) = \begin{pmatrix} \mathbf{K}(\boldsymbol{\rho}) & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{I}_n \end{pmatrix}. \tag{13}$$

Let $\boldsymbol{\Omega} = \sigma^2 \mathbf{J}^{-1}(\boldsymbol{\rho}) \mathbf{P}(\boldsymbol{\rho}) \mathbf{J}^{'-1}(\boldsymbol{\rho})$. Then, the distribution of $\mathbf{h}$ is

$$\mathbf{h}|\boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2 \sim N\big(\boldsymbol{l}_T \otimes \boldsymbol{\mu},\ \boldsymbol{\Omega}\big). \tag{14}$$

Note that when $\boldsymbol{\rho} = \mathbf{0}$, $\boldsymbol{\Omega}$ reduces to $\sigma^2\mathbf{I}_{nT}$, and thus the result in (14) becomes $\mathbf{h}|\boldsymbol{\mu}, \sigma^2 \sim N(\boldsymbol{\iota}_T \otimes \boldsymbol{\mu},\ \sigma^2\mathbf{I}_{nT})$. Consider the following partition of $\boldsymbol{\Omega}$:

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} & \dots & \boldsymbol{\Omega}_{1,T-1} & \boldsymbol{\Omega}_{1T} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} & \dots & \boldsymbol{\Omega}_{2,T-1} & \boldsymbol{\Omega}_{2T} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\Omega}_{T1} & \boldsymbol{\Omega}_{T2} & \dots & \boldsymbol{\Omega}_{T,T-1} & \boldsymbol{\Omega}_{TT} \end{pmatrix}, \tag{15}$$

where each $\boldsymbol{\Omega}_{st}$ for $s, t = 1, 2 \dots, T$ is an $n \times n$ sub-matrix of $\boldsymbol{\Omega}$. Let $\Omega_{ij,st}$ be the (i, j)th element of $\boldsymbol{\Omega}_{st}$ for $i, j = 1, 2, \dots, n$. Let $r \in \mathbb{N}$ be a natural even number. Then, the even moments of y_{it} can be expressed as

$$\mathbb{E}\left(y_{it}^r\right) = \mathbb{E}\left(e^{\frac{r}{2}h_{it}}\right)\mathbb{E}(v_{it}^r) = \exp\left(\frac{\mu_i r}{2} + \frac{r^2}{8}\Omega_{ii,tt}\right)\gamma(r) \tag{16}$$

where $\gamma(r) = \frac{r!}{2^{r/2}(r/2)!}$. Then, it follows that $\mathbb{E}(y_{it}^4)/\left(\mathbb{E}(y_{it}^2)\right)^2 - 3 = 3\left(\exp(\Omega_{ii,tt}) - 1\right) > 0$. Thus, our specification suggests that y_{it} has a leptokurtic symmetric distribution. Next, we consider the covariance between y_{it}^r and y_{js}^r:

$$\begin{aligned} \mathrm{Cov}\left(y_{it}^r, y_{js}^r\right) &= \mathbb{E}\left(e^{\frac{r}{2}(h_{it}+h_{js})}v_{it}^r v_{js}^r\right) - \mathbb{E}\left(e^{\frac{r}{2}h_{it}}v_{it}^r\right)\mathbb{E}\left(e^{\frac{r}{2}h_{js}}v_{js}^r\right) \\ &= \gamma^2(r)\exp\left(\frac{r(\mu_i+\mu_j)}{2} + \frac{r^2}{8}\left(\Omega_{ii,tt} + \Omega_{jj,ss} + 2\Omega_{ij,ts}\right)\right) \\ &\quad - \gamma^2(r)\exp\left(\frac{\mu_i r}{2} + \frac{r^2}{8}\Omega_{ii,tt}\right)\exp\left(\frac{\mu_j r}{2} + \frac{r^2}{8}\Omega_{jj,ss}\right) \\ &= \gamma^2(r)\exp\left(\frac{r(\mu_i+\mu_j)}{2} + \frac{r^2}{8}\left(\Omega_{ii,tt} + \Omega_{jj,ss}\right)\right)\left(\exp\left(\frac{r^2}{4}\Omega_{ij,ts}\right) - 1\right). \end{aligned} \tag{17}$$

This result indicates that our specification introduces spatial dependence in the outcome variable since $\mathrm{Cov}(y_{it}^r, y_{js}^r) \neq 0$ in general. Note that $\mathrm{Cov}(y_{it}^r, y_{js}^r) = 0$ when $\boldsymbol{\rho} = \mathbf{0}$ holds, because $(\exp(\frac{r^2}{4}\Omega_{ij,ts}) - 1) = 0$.

4. Posterior Analysis

To introduce a Bayesian MCMC estimation approach, we first transform our model such that the resulting outcome equation is linear in $\mathbf{h}_t$. We then determine the conditional likelihood function of the transformed model by approximating the distribution of the transformed disturbance term with a Gaussian mixture distribution (Kim et al., 1998; Chib et al., 2002; Omori et al., 2007). The conditional likelihood function of the transformed model facilitates the sampling steps for $\mathbf{h}_t$ and the auxiliary mixture component indicator defined subsequently. We also provide the conditional likelihood function of the original model, which we use to determine the sampling steps of other parameters in our model.

We square both sides of (8) and then take the logarithm to obtain

$$y_{it}^* = h_{it} + v_{it}^*, \tag{18}$$

where $y_{it}^* = \log y_{it}^2$ and $v_{it}^* = \log v_{it}^2$. Let $p(\cdot)$ be a density function and $p(\cdot|\cdot)$ be a conditional density function. Then, the density of v_{it}^* can be derived as

$$p(v_{it}^*) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(e^{v_{it}^*} - v_{it}^*)\right), \quad -\infty < v_{it}^* < \infty, \quad i = 1, 2, \dots, n,\ t = 1, \dots, T. \tag{19}$$

This density is highly left skewed with $\mathbb{E}(v_{it}^*) \approx -1.2704$ and $\mathrm{Var}(v_{it}^*) = \pi^2/2 \approx 4.9348$. Define $\mathbf{Y}_t^* = (y_{1t}^*, y_{2t}^*, \dots, y_{nt}^*)'$ and $\mathbf{V}_t^* = (v_{1t}^*, v_{2t}^*, \dots, v_{nt}^*)'$. Then, in vector form, we have

$$\mathbf{Y}_t^* = \mathbf{h}_t + \mathbf{V}_t^*. \tag{20}$$

In order to convert (20) into a linear Gaussian state-space model, we approximate $p(v_{it}^*)$ with an m-component Gaussian mixture distribution:

$$p(v_{it}^*) \approx \sum_{j=1}^{m} p_j \times \phi\left(v_{it}^*|\mu_j,\ \sigma_j^2\right), \tag{21}$$

where $\phi(v_{it}^*|\mu_j,\ \sigma_j^2)$ denotes the Gaussian density function with mean μ_j and variance σ_j^2, p_j is the probability of jth mixture component and m is the number of components. In particular, we use the ten-component Gaussian mixture distribution suggested by Omori et al. (2007) to approximate $p(v_{it}^*)$. We provide the parameter values of the ten-component Gaussian

Table 1
The ten-component Gaussian mixture for $p(v_{it}^*)$.

Components	p_j	μ_j	σ_j^2
1	0.00609	1.92677	0.11265
2	0.04775	1.34744	0.17788
3	0.13057	0.73504	0.26768
4	0.20674	0.02266	0.40611
5	0.22715	-0.85173	0.62699
6	0.18842	-1.97278	0.98583
7	0.12047	-3.46788	1.57469
8	0.05591	-5.55246	2.54498
9	0.01575	-8.68384	4.16591
10	0.00115	-14.65000	7.33342

mixture distribution in Table 1. The parameters in this table are chosen by matching the first four moments of the ten-component Gaussian mixture distribution with that of $p(v_{it}^*)$. This approach has two advantages. First, the Gaussian mixture distribution with the pre-determined parameter values in Table 1 provides a good enough approximation to $p(v_{it}^*)$ (Omori et al., 2007). Second, this approach does not pose any estimation difficulties since the mixture parameters in Table 1 are pre-determined.

We can equivalently write (21) in terms of an auxiliary discrete random variable $z_{it} \in \{1, 2, \ldots, m\}$ that serves as the mixture component indicator:

$$v_{it}^* | (z_{it} = j) \sim N\left(\mu_j,\, \sigma_j^2\right), \quad \text{and} \quad \mathbb{P}(z_{it} = j) = p_j, \quad j = 1, 2, \ldots, m, \tag{22}$$

where $\mathbb{P}(z_{it} = j) = p_j$ is the probability that z_{it} takes the jth value. Let $\mathbf{Z}_t = (z_{1t}, \ldots, z_{nt})'$, $\mathbf{d}_t = (\mu_{z_{1t}}, \ldots, \mu_{z_{nt}})'$ and $\mathbf{\Sigma}_t = \text{Diag}(\sigma_{z_{1t}}^2, \ldots, \sigma_{z_{nt}}^2)$. Then, from (22), we have $\mathbf{V}_t^* | \mathbf{Z}_t \sim N(\mathbf{d}_t,\, \mathbf{\Sigma}_t)$, which indicates that our model in (20) is now conditionally linear Gaussian given the component indicator variable. Thus, from (20), we have

$$\mathbf{Y}_t^* | \mathbf{Z}_t,\, \mathbf{h}_t \sim N\left(\mathbf{h}_t + \mathbf{d}_t,\, \mathbf{\Sigma}_t\right), \tag{23}$$

which facilitates the sampling steps for $\mathbf{h}_t$ and $\mathbf{Z}_t$ in our suggested Gibbs sampler given in Algorithm 1. The sampling steps for the remaining parameters require the following distribution:

$$\mathbf{h} | \boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2 \sim N\left(\boldsymbol{I}_T \otimes \boldsymbol{\mu},\, \mathbf{\Omega}\right). \tag{24}$$

The conditional distribution of $\mathbf{Y}_t$ given $\mathbf{h}_t$ is

$$\mathbf{Y}_t | \mathbf{h}_t \sim N(\mathbf{0},\, \mathbf{H}_t), \tag{25}$$

where $\mathbf{H}_t = \text{Diag}(e^{h_{1t}}, \ldots, e^{h_{nt}})$ is the $n \times n$ diagonal matrix with the (i, i)th diagonal element $e^{h_{it}}$.

We are now in a position to design a Gibbs sampler by using our results on (i) the mixture component indicators in (22), (ii) the conditional likelihood function of transformed model in (23), (iii) the conditional likelihood function of $\mathbf{Y}_t$ in (25), and (iv) the distribution of $\mathbf{h}$ in (14). Let $\mathbf{Y} = (\mathbf{Y}_1', \ldots, \mathbf{Y}_T')'$ and $\mathbf{Z} = (\mathbf{Z}_1', \ldots, \mathbf{Z}_T')'$. The joint posterior distribution $p(\mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2 | \mathbf{Y})$ can be expressed as

$$p(\mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2 | \mathbf{Y}) \propto p(\mathbf{Y}^* | \mathbf{Z}, \mathbf{h}) \times p(\mathbf{h} | \boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2) \times p(\mathbf{Z}) \times p(\boldsymbol{\mu}) \times p(\boldsymbol{\rho}) \times p(\sigma^2), \tag{26}$$

where $p(\mathbf{a})$ is the prior density function of $\mathbf{a}$ for $\mathbf{a} \in \{\mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2\}$. Then, our suggested Gibbs sampler for generating draws from $p(\mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2 | \mathbf{Y})$ consists of the steps given in Algorithm 1.

Algorithm 1 (Estimation Algorithm)**.**

1. Sampling step for $\mathbf{Z}$: Note that $\mathbf{Z}_t$ is a discrete random variable, and its conditional posterior probability mass function is

$$p\left(\mathbf{Z}_t|\mathbf{Y}_t, \mathbf{h}_t, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2\right) \propto p(\mathbf{Z}_t)p\left(\mathbf{Y}_t^*|\mathbf{Z}_t, \mathbf{h}_t\right) = \prod_{i=1}^{n} p\left(y_{it}^*|z_{it}, h_{it}\right)p(z_{it}), \tag{27}$$

for $t = 1, \ldots, T$. Thus,

$$\mathbb{P}\left(z_{it} = j|y_{it}^*\right) = \frac{p_j\phi\left(y_{it}^*|h_{it} + \mu_j,\ \sigma_j^2\right)}{\sum_{k=1}^{10} p_k\phi\left(y_{it}^*|h_{it} + \mu_k,\ \sigma_k^2\right)},\ j = 1, \ldots, 10,\ i = 1, \ldots, n, \tag{28}$$

for $t = 1, \ldots, T$, where the denominator is the normalization constant.

2. Sampling step for $\mathbf{h}$: Let $\boldsymbol{\Sigma} = \text{Diag}\left(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T\right)$ and $\mathbf{d} = (\mathbf{d}_1', \ldots, \mathbf{d}_T')'$. Using standard regression results on $p(\mathbf{h}|\mathbf{Y}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2) \propto \prod_{t=1}^{T} p(\mathbf{Y}_t^*|\mathbf{Z}, \mathbf{h}_t)p(\mathbf{h}|\boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2)$, we obtain

$$\mathbf{h}|\mathbf{Y}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2 \sim N\left(\hat{\mathbf{b}}_h, \hat{\mathbf{B}}_h\right), \tag{29}$$

where

$$\hat{\mathbf{B}}_h = \left(\boldsymbol{\Omega}^{-1} + \boldsymbol{\Sigma}^{-1}\right)^{-1}, \quad \hat{\boldsymbol{b}}_h = \hat{\mathbf{B}}_h\left(\boldsymbol{\Omega}^{-1}(\boldsymbol{l}_T \otimes \boldsymbol{\mu}) + \boldsymbol{\Sigma}^{-1}(\mathbf{Y}^* - \mathbf{d})\right).$$

3. Sampling step for $\boldsymbol{\mu}$: Using (9) and (13), we can find that

$$\boldsymbol{\Omega}^{-1} = \begin{pmatrix} \boldsymbol{\Omega}_{11}^* & \boldsymbol{\Omega}_{12}^* & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\Omega}_{21}^* & \boldsymbol{\Omega}_{22}^* & \boldsymbol{\Omega}_{23}^* & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\Omega}_{T-1,T-1}^* & \boldsymbol{\Omega}_{T-1,T}^* \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\Omega}_{T,T-1}^* & \boldsymbol{\Omega}_{TT}^* \end{pmatrix},$$

where

$$\boldsymbol{\Omega}_{11}^* = \sigma^{-2}\mathbf{S}'(\rho_1)\mathbf{K}^{-1}(\boldsymbol{\rho})\mathbf{S}(\rho_1) + \sigma^{-2}\mathbf{A}'(\rho_2, \rho_3)\mathbf{A}(\rho_2, \rho_3), \quad \boldsymbol{\Omega}_{TT}^* = \sigma^{-2}\mathbf{S}'(\rho_1)\mathbf{S}(\rho_1),$$
$$\boldsymbol{\Omega}_{ii}^* = \sigma^{-2}\mathbf{S}'(\rho_1)\mathbf{S}(\rho_1) + \sigma^{-2}\mathbf{A}'(\rho_2, \rho_3)\mathbf{A}(\rho_2, \rho_3), \quad i = 2, \ldots, T-1, \text{ and}$$
$$\boldsymbol{\Omega}_{i,i+1}^* = \boldsymbol{\Omega}_{i+1,i}^{*'} = -\sigma^{-2}\mathbf{A}'(\rho_2, \rho_3)\mathbf{S}(\rho_1), \quad i = 1, \ldots, T-1.$$

Then, from $p(\boldsymbol{\mu}|\mathbf{Y}, \mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2) \propto p(\mathbf{h}|\boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2)p(\boldsymbol{\mu})$, we obtain

$$\boldsymbol{\mu}|\mathbf{Y}, \mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho}, \sigma^2 \sim N\left(\hat{\mathbf{b}}_\mu, \hat{\mathbf{B}}_\mu\right), \tag{30}$$

where

$$\hat{\mathbf{B}}_\mu = \left(\mathbf{B}_\mu^{-1} + \sum_{j=1}^{T}\sum_{i=1}^{T}\boldsymbol{\Omega}_{ij}^*\right)^{-1}, \quad \hat{\mathbf{b}}_\mu = \hat{\mathbf{B}}_\mu\left(\mathbf{B}_\mu^{-1}\boldsymbol{b}_\mu + \sum_{j=1}^{T}\sum_{i=1}^{T}\boldsymbol{\Omega}_{ij}^*\mathbf{h}_i\right).$$

4. Sampling step for σ^2: Using $p(\sigma^2|\mathbf{Y}, \mathbf{h}, \boldsymbol{\mu}, \boldsymbol{\rho}) \propto p(\mathbf{h}|\boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2)p(\sigma^2)$, where $p(\sigma^2)$ is the prior density function of σ^2 given in (7), we obtain

$$\sigma^2|\mathbf{Y}, \mathbf{h}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\rho} \sim \text{IG}(\hat{a}, \hat{b}),$$

where

$$\hat{a} = a + nT/2, \quad \hat{b} = b + \frac{1}{2}\left(\mathbf{h} - (\boldsymbol{l}_T \otimes \boldsymbol{\mu})\right)'\left(\mathbf{J}'(\boldsymbol{\rho})\mathbf{P}^{-1}(\boldsymbol{\rho}, \boldsymbol{\lambda})\mathbf{J}(\boldsymbol{\rho})\right)\left(\mathbf{h} - (\boldsymbol{l}_T \otimes \boldsymbol{\mu})\right).$$

5. Sampling step for $\boldsymbol{\rho}$: The conditional posterior density of $\boldsymbol{\rho}$ does not take any known form in our model. We use the adaptive Metropolis (AM) algorithm suggested in Haario et al. (2001) and Roberts and Rosenthal (2009) to generate draws from $p(\boldsymbol{\rho}|\mathbf{Y}, \mathbf{h}, \mathbf{Z}, \sigma^2)$. Let $\mathbf{a}^{(g)}$ be the gth draw generated at the gth iteration for $\mathbf{a} \in \{\boldsymbol{\rho}, \boldsymbol{\mu}, \sigma^2\}$. Then, at the iteration g, we use the following proposal distribution to generate the candidate value $\tilde{\boldsymbol{\rho}}$:

$$f_g\left(\boldsymbol{\rho}|\boldsymbol{\rho}^{(0)}, \ldots, \boldsymbol{\rho}^{(g-1)}\right) = \begin{cases} N\left(\boldsymbol{\rho}^{(g-1)},\ \frac{(0.1)^2}{3} \times \mathbf{I}_3\right), & \text{for} \quad g \le g_0, \\ 0.95 \times N\left(\boldsymbol{\rho}^{(g-1)},\ \frac{c(2.38)^2}{3} \times \text{Cov}\left(\boldsymbol{\rho}^{(0)}, \ldots, \boldsymbol{\rho}^{(g-1)}\right)\right) \\ \quad +0.05 \times N\left(\boldsymbol{\rho}^{(g-1)},\ \frac{(0.1)^2}{3} \times \mathbf{I}_3\right), & \text{for} \quad g > g_0, \end{cases}$$

where g_0 is the length of the initial sampling period, $\text{Cov}\left(\boldsymbol{\rho}^{(0)},\ldots,\boldsymbol{\rho}^{(g-1)}\right)$ is the empirical covariance matrix of historical draws given by $\text{Cov}\left(\boldsymbol{\rho}^{(0)},\ldots,\boldsymbol{\rho}^{(g-1)}\right)=\frac{1}{g}\sum_{j=0}^{g-1}\boldsymbol{\rho}^{(j)}\boldsymbol{\rho}^{(j)'}-\bar{\boldsymbol{\rho}}^{(g-1)}\bar{\boldsymbol{\rho}}^{(g-1)'}$ with $\bar{\boldsymbol{\rho}}^{(g-1)}=\frac{1}{g}\sum_{j=0}^{g-1}\boldsymbol{\rho}^{(j)}$, and c is a scalar tuning parameter used to achieve a reasonable acceptance rate. For our application, we set $g_0=6$. We then check whether $\tilde{\boldsymbol{\rho}}$ satisfies the stability conditions in Section 2. If not, we regenerate $\tilde{\boldsymbol{\rho}}$ until it meets the stability conditions. We compute the following acceptance probability:

$$\mathbb{P}(\boldsymbol{\rho}^{(g-1)},\tilde{\boldsymbol{\rho}})=\min\left(\frac{p\left(\mathbf{h}|\tilde{\boldsymbol{\rho}},\boldsymbol{\mu}^{(g)},\sigma^{2(g)}\right)}{p\left(\mathbf{h}|\boldsymbol{\rho}^{(g-1)},\boldsymbol{\mu}^{(g)},\sigma^{2(g)}\right)},1\right),$$

where $p(\mathbf{h}|\boldsymbol{\rho},\boldsymbol{\mu},\sigma^2)$ is given in (14). Finally, we return $\tilde{\boldsymbol{\rho}}$ with probability $\mathbb{P}(\boldsymbol{\rho}^{(g-1)},\tilde{\boldsymbol{\rho}})$; otherwise return $\boldsymbol{\rho}^{(g-1)}$.

Remark 1. In Step 1, the result in (28) indicates that the mixture components are conditionally independent given y_{it}^*. Thus, each component is a discrete random variable taking integer values in the interval [1,10] with the conditional posterior probability $\mathbb{P}(z_{it}=j|y_{it}^*)$. The conditional posterior results in Steps 2, 3, and 4 are obtained from a standard Bayesian analysis as in a linear regression model. In the AM algorithm described in Step 5, the proposal distribution $f_g(\boldsymbol{\rho}|\boldsymbol{\rho}^{(0)},\ldots,\boldsymbol{\rho}^{(g-1)})$ has two parts. The first part is $N(\boldsymbol{\rho}^{(g-1)},\frac{(0.1)^2}{3}\times\mathbf{I}_3)$, and is used when the number of iterations is less than or equal to g_0. The second part consists of two normal distributions. The first component is specified as $N(\boldsymbol{\rho}^{(g-1)},\frac{c(2.38)^2}{3}\times\text{Cov}(\boldsymbol{\rho}^{(0)},\ldots,\boldsymbol{\rho}^{(g-1)}))$, where the covariance matrix is determined from the historical MCMC draws of $\boldsymbol{\rho}$. The second component is $N(\boldsymbol{\rho}^{(g-1)},\frac{(0.1)^2}{3}\times\mathbf{I}_3)$. The candidate values generated from $f_g(\boldsymbol{\rho}|\boldsymbol{\rho}^{(0)},\ldots,\boldsymbol{\rho}^{(g-1)})$ are subject to the stability conditions given in (5). Finally, we adjust the tuning parameter c during the estimation to achieve an acceptance rate that falls between 40 percent and 60 percent. Han and Lee (2016) and Yang et al. (2023) use this algorithm to generate draws for spatial parameters in spatial panel data models. Their results show that this algorithm can perform satisfactorily.

Remark 2. In the sampling step for $\boldsymbol{\rho}$, $\mathbb{P}(\boldsymbol{\rho}^{(g-1)},\tilde{\boldsymbol{\rho}})$ is calculated at each pass of the sampler, and therefore, $p(\mathbf{h}|\boldsymbol{\rho},\boldsymbol{\mu},\sigma^2)$ is evaluated twice at each pass of the sampler. In other words, $p(\mathbf{h}|\boldsymbol{\rho},\boldsymbol{\mu},\sigma^2)=(2\pi)^{-nT/2}|\boldsymbol{\Omega}|^{-1/2}\exp\left(-\frac{1}{2}(\mathbf{h}-\boldsymbol{l}_T\otimes\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\mathbf{h}-\boldsymbol{l}_T\otimes\boldsymbol{\mu})\right)$ must be calculated twice. Since $\boldsymbol{\Omega}=\sigma^2\mathbf{J}^{-1}(\boldsymbol{\rho})\mathbf{P}(\boldsymbol{\rho})\mathbf{J}'^{-1}(\boldsymbol{\rho})$, we have $|\boldsymbol{\Omega}|^{-1/2}=(\sigma^2)^{-nT/2}|\mathbf{J}(\boldsymbol{\rho})||\mathbf{P}(\boldsymbol{\rho})|^{-1/2}$. From (9), since $\mathbf{J}(\boldsymbol{\rho})$ is a triangular matrix, we have $|\mathbf{J}(\boldsymbol{\rho})|=\prod_{t=1}^{T}|\mathbf{S}(\rho_1)|=|\mathbf{S}(\rho_1)|^T$. Also, from (13), since $|\mathbf{P}(\boldsymbol{\rho})|$ is a block-diagonal matrix, we have $|\mathbf{P}(\boldsymbol{\rho})|=|\mathbf{K}(\boldsymbol{\rho})|$. These results can be used to improve the computational efficiency of the sampler.

5. Simulations

In this section, we provide simulation evidence to assess the sampling properties of the suggested Bayesian algorithm. The data generating process follows (1) and (2). More specifically, the elements of $\mathbf{V}_t$ and $\mathbf{U}_t$ are drawn independently from the standard normal distribution for $t=1,2,\ldots,T$. Therefore, the value of σ^2 is set to 0.25 in all experiments. The elements of $\boldsymbol{\mu}$ are drawn independently from the normal distribution with mean 3.3 and standard deviation 0.35. To initialize the process, we use (14), and the series expression for $\mathbf{K}(\boldsymbol{\rho})$ is truncated at 15. We consider two sets of values for $\boldsymbol{\rho}$, $\{(0.6,0.35,-0.025),(0.3,0.65,-0.025)\}$. These parameter values are chosen to ensure that the data-generating process mimics the findings from our empirical application in the next section. The number of spatial units n is set to 98, and the number of time periods T is fixed at 50.

For the spatial weights matrix $\mathbf{W}$, we consider a row-standardized queen contiguity weights matrix. To this end, we first generate a vector containing a random permutation of the integers from 1 to n without repeating elements. Then, we reshape this vector into a $k\times c$ rectangular lattice, where $c=n/k$. We use this lattice to determine the spatial locations and adjacency relations of the n locations. We set $w_{ij}=1$ if the jth observation is adjacent to or shares a border with the ith observation; otherwise, we set $w_{ij}=0$. We set $k=7$, and row-normalize the resulting weights matrices. For the prior distributions, we consider two cases: (i) $\sigma^2\sim\text{IG}(3,2)$ and $\boldsymbol{\mu}\sim N(\mathbf{0},10\mathbf{I}_n)$, and (ii) $\sigma^2\sim\text{IG}(0.01,0.01)$ and $\boldsymbol{\mu}\sim N(\mathbf{0},100\mathbf{I}_n)$. The length of the Markov chain is 22000 draws, and the first 2000 draws are discarded to dissipate the effects of the initial values.

Figures 1–4 show the simulation results when $\sigma^2\sim\text{IG}(3,2)$ and $\boldsymbol{\mu}\sim N(\mathbf{0},10\mathbf{I}_n)$. To determine the adequacy of the length of the chains and their mixing properties, we provide the trace plots in Figures 1 and 3. Figure 1 displays the trace plots when $\boldsymbol{\rho}=(0.6,0.35,-0.025)$, while Figure 3 presents the plots when $\boldsymbol{\rho}=(0.30,0.65,-0.025)$. In these plots, the red solid lines correspond to the estimated posterior means. We observe that our suggested sampler performs satisfactorily and provides chains that mix well in all cases. Also note that for both $\boldsymbol{\rho}$ and σ^2, the 95% credible intervals contain the true values chosen in the experiments.

For the $n\times 1$ vector $\boldsymbol{\mu}$ and the $nT\times 1$ vector $\mathbf{h}$, we provide evidence on the performance of our Bayesian estimator in Figures 2 and 4. In these plots, the true values are represented by solid lines, and the estimates are presented by dashed lines. In the first panel, we observe that the estimated posterior means for the components of $\boldsymbol{\mu}$ are generally close to the true values. Here, the shaded region refers to the 95% credible interval. For $\mathbf{h}$, we calculate the average of true values over i and over t, respectively, and plot them against the average posterior means over space and time, respectively. The second panel presents the case where the average is taken over the spatial locations i, and the last panel is the case where

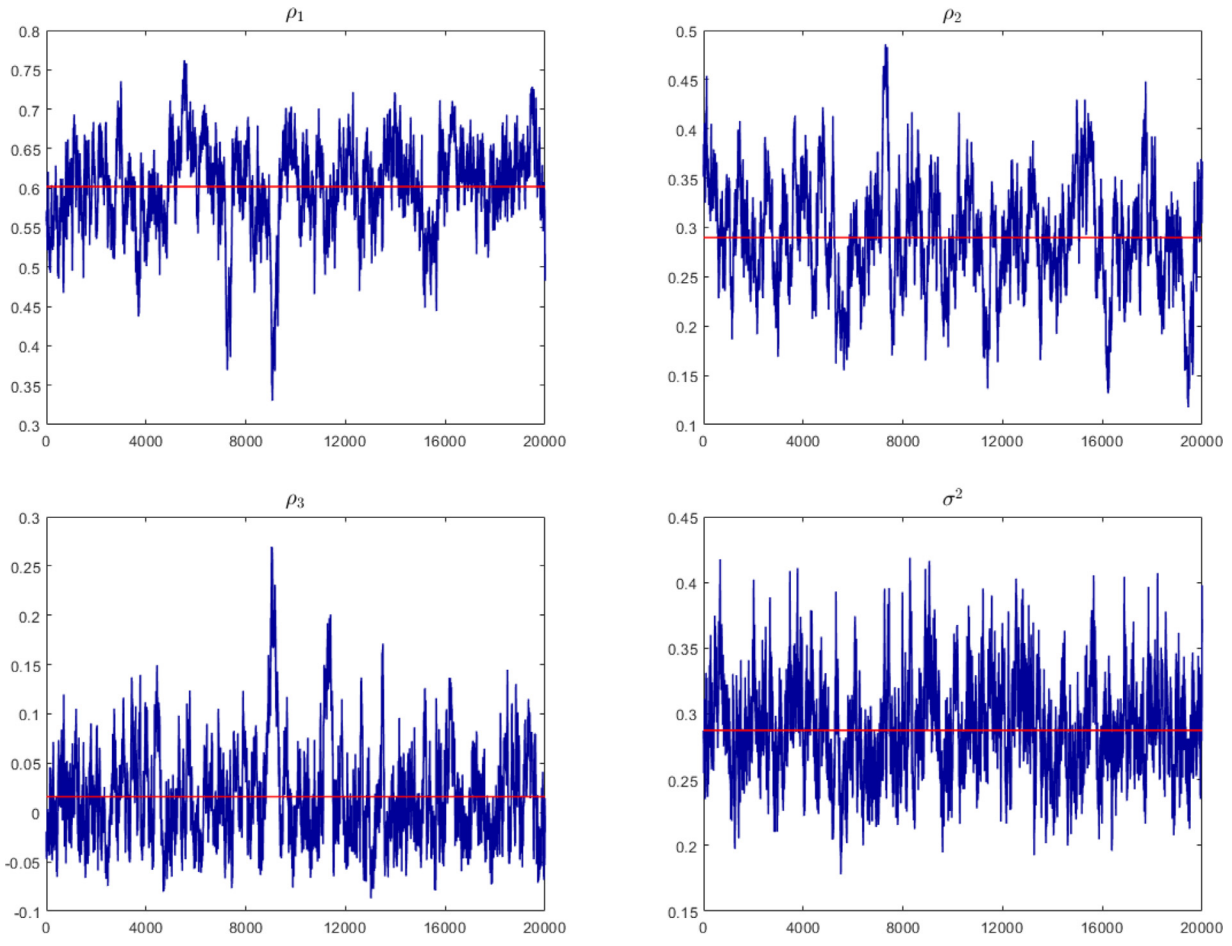

Fig. 1. Trace plots for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

the average is taken over t. We observe that the Bayesian estimator performs satisfactorily in terms of capturing the log volatility over cross-sections as well as over time.

Next, we consider relatively more uninformative priors for σ^2 and $\boldsymbol{\mu}$ by setting $\sigma^2 \sim \text{IG}(0.01, 0.01)$ and $\boldsymbol{\mu} \sim N(\mathbf{0}, 100\mathbf{I}_n)$. The simulation results for this case are presented in Appendix. The trace plots in Figures 11 and 13 display a pattern resembling a "fat hairy caterpillar," indicating no convergence issues in our sampler. Figures 12 and 14 show the estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ along with their true values. These figures suggest that the estimated posterior mean values closely align with the corresponding true values and track them. Overall, these results indicate that the proposed sampler is not sensitive to the values of the hyperparameters.

Finally, to display the spatiotemporal dynamics, we present volatility estimates for selected locations and time periods under the following prior distributions: $\sigma^2 \sim \text{IG}(0.01, 0.01)$ and $\boldsymbol{\mu} \sim N(\mathbf{0}, 100\mathbf{I}_n)$. These results are presented in the Appendix. Figure 15 shows the line plots of $\{\hat{h}_{13,t}\}_{t=1}^{50}$, $\{\hat{h}_{34,t}\}_{t=1}^{50}$, $\{\hat{h}_{73,t}\}_{t=1}^{50}$, $\{\hat{h}_{91,t}\}_{t=1}^{50}$ against their true corresponding values, while Figure 16 presents the line plots of $\{\hat{h}_{i,11}\}_{i=1}^{98}$, $\{\hat{h}_{i,23}\}_{i=1}^{98}$, $\{\hat{h}_{i,35}\}_{i=1}^{98}$, $\{\hat{h}_{i,43}\}_{i=1}^{98}$ against their true corresponding values. Overall, we observe that our sampler produces estimates that align with their corresponding true values.

6. Air Quality Modeling and Environmental Risks

In this section, we demonstrate the usage of the dynamic stochastic volatility model using an empirical example from environmental science. As for applications in financial economics, the volatility of the process can be interpreted as risks (i.e., environmental risks in our case). Previous ecological studies mostly focused on changes in the mean behavior (cf. Wilby et al., 2009), but also an increased variation might be harmful, e.g., Vasseur et al. (2014) showed that an increased temperature variation (i.e., changes in the variance) poses a greater risk than global warming (i.e., changes in the mean). Similar results on this topic were found by Paaijmans et al. (2013); Screen (2014) and Tewksbury et al. (2008). These varying risk levels are mostly observed locally. Huntingford et al. (2013) found no evidence for time-varying standard deviations on a global scale. There is also a connection between climate changes and the financial market, such that climate variations

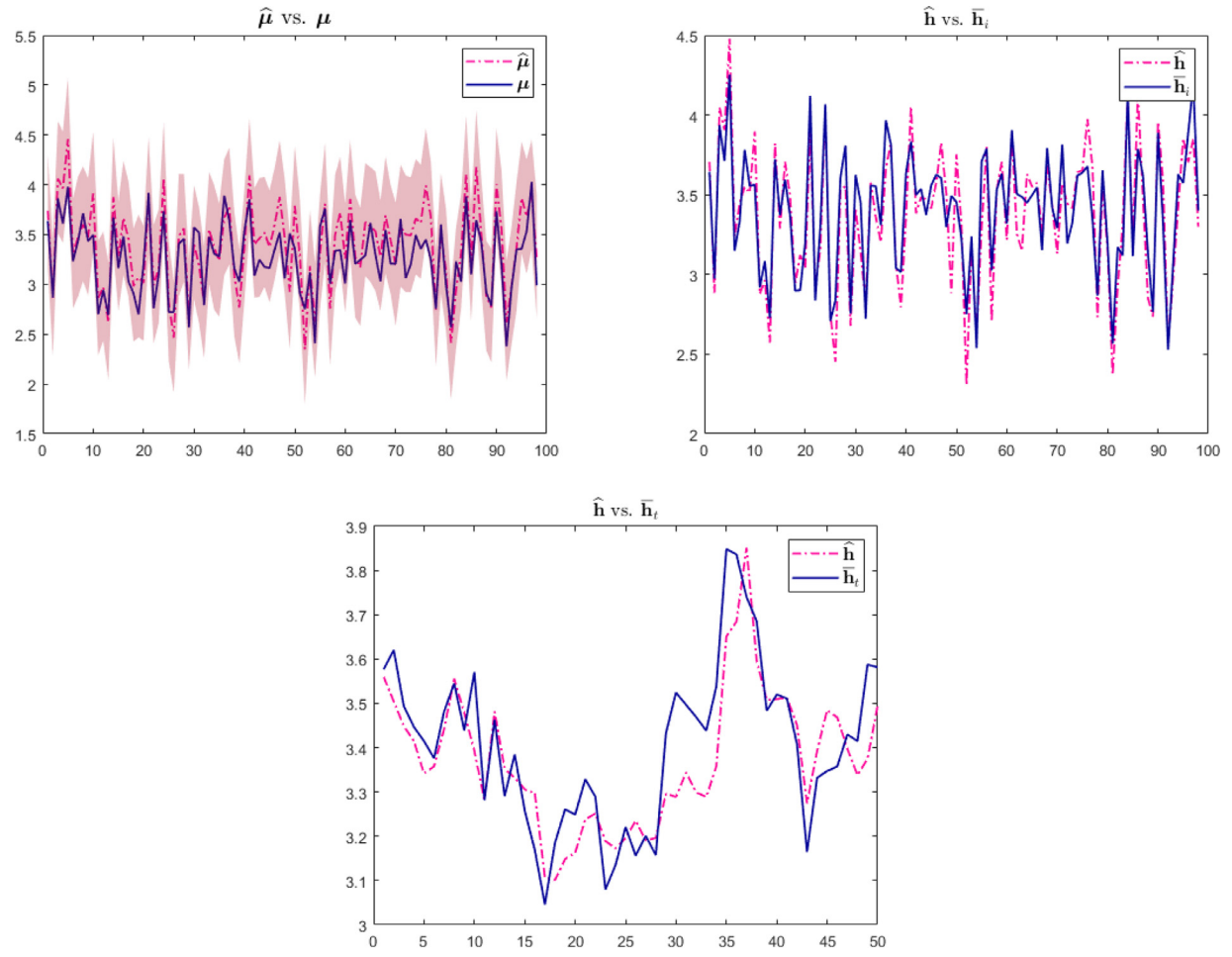

Fig. 2. Estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ (dashed lines) and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

might have an impact on the risk of financial markets (see also Giglio et al. 2021; Hong et al. 2020). In addition, temporal volatility models have been applied to describe varying model uncertainties in temperature series. More specifically, Hølleland and Karlsen (2020a) used generalized autoregressive conditional heteroscedasticity (GARCH) models for the residuals of an autoregressive model with exogenous regressors for the temperature over 44 years. They showed time-varying model uncertainties throughout the year, which are the highest during the winter seasons.

In this article, we analyze the variation of fine dust concentrations of particles having a diameter less than 10 μm, PM_{10}, in Lombardy, Northern Italy. Bounded by the Alps to the west, the region experiences reduced wind circulation, contributing to Lombardy being among the European regions with the lowest air quality, as discussed in (see Fassò et al., 2022). In the following empirical analysis, we use the daily PM_{10} concentrations from 1.1.2021 to 31.12.2021 from the official monitoring stations of the regional environmental authority, ARPA Lombardia (Maranzano, 2022). The data are open-source provided by the Agrimonia project (Fassò et al. 2023a). We used version 3.0 of the dataset for this analysis. We refer the interested reader to Fassò et al. (2023b) for a detailed description of the dataset. In total, there are $n = 103$ measurement stations and $T = 365$ daily observations. Missing values were imputed by the overall mean. To provide a first overview of the dataset, we depicted the median concentrations in $\mu g/m^3$ across space and time as a time-series plot (Figure 5, left) and displayed on the map (Figure 5, right), respectively. The code to reproduce the results can be found on GitHub philot789.github.io/SVM_EnvRisk/.

For our analysis, we first estimated a spatial panel model to describe the mean variations. The spatial correlations have been modeled in an autoregressive manner. To be precise, the mean model is given by

$$\mathbf{C}_t = \psi \mathbf{W}\mathbf{C}_t + \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{X}_t\boldsymbol{\gamma} + \mathbf{s} + a_t\mathbf{1}_n + \mathbf{E}_t\,. \tag{31}$$

The outcome variable $\mathbf{C}_t$ is the n-dimensional vector of PM_{10} concentrations at time point t, $\mathbf{W}$ is spatial weights matrix which we specify below in more detail, $\mathbf{X}_t$ is an $n \times p$ matrix of exogenous regressors, and $\mathbf{E}_t$ denotes the model error terms. Spatial interactions are included via the spatial autoregressive term $\psi\mathbf{W}\mathbf{C}_t$ with an unknown autoregressive parameter ψ. Moreover, spatial fixed effects $\mathbf{s}$ are present for each station (even if non-significant) because there are different types of

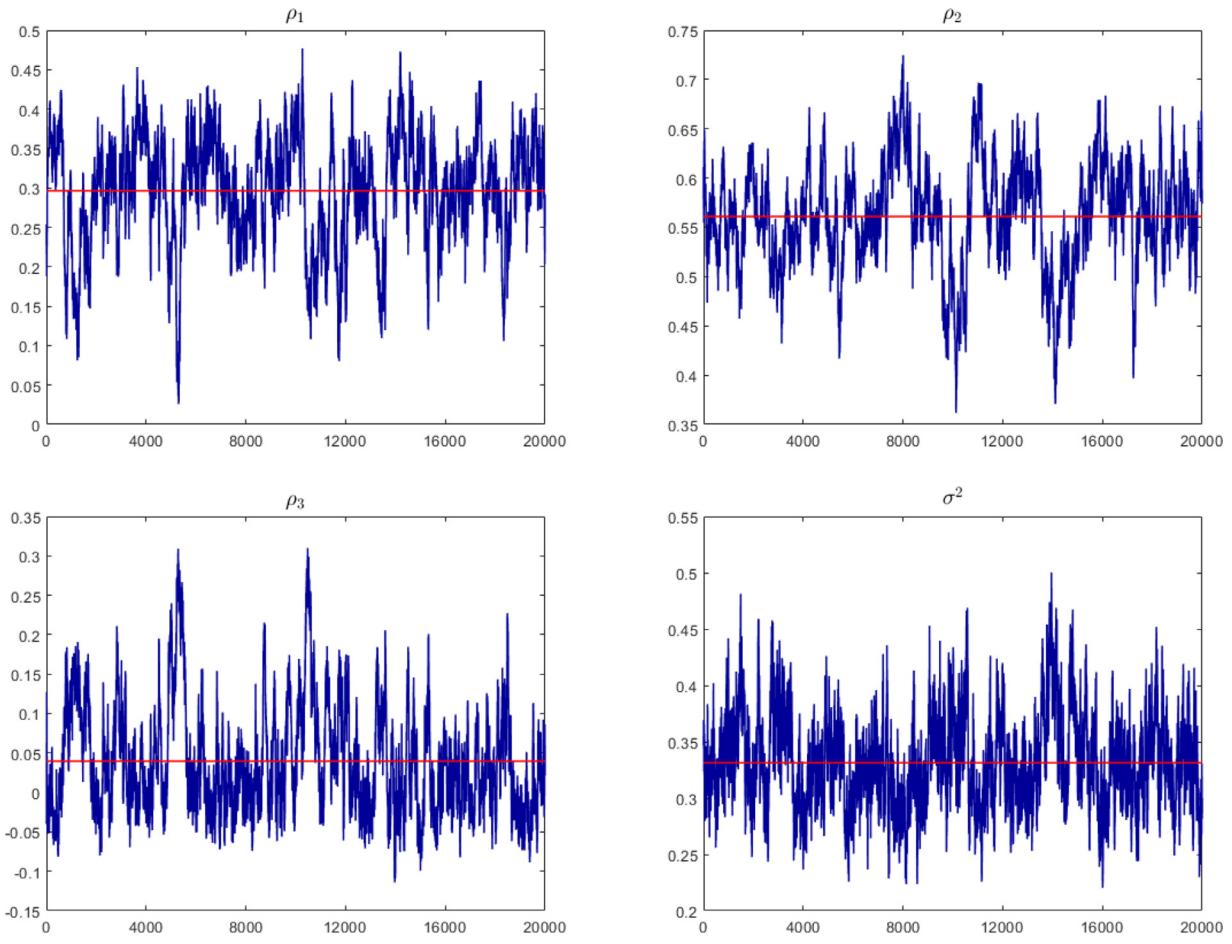

Fig. 3. Trace plots for $\boldsymbol{\rho} = (0.30, 0.65, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

stations included in the data set, e.g., urban traffic stations located at major roads in the cities or rural background stations located in the Alps. The spatial fixed effects describe the station-specific PM_{10} concentrations and thereby serve as the model intercept. The temporal fixed effects $a_t\mathbf{1}_n$ will remove any additional (station invariant) seasonal variation across time, which is not explained by the exogenous regressors. More precisely, we included $p = 4$ covariates that are known to influence PM concentrations, namely the maximum height of the planetary boundary layer (PBL), the relative humidity, the air temperature, and the pressure level. All covariates are available as daily observations for each measurement station. Additionally, they were standardized to compare the size of the effects.

The spatial weights matrix has been chosen as a row-standardized inverse-distance matrix with a cut-off distance of 32 miles. This weighting scheme shows the best goodness-of-fit across several alternative weight matrices based on binary and inverse-distance weights (see Figure 6). On average, each station has 21.05 neighbors within the cut-off distance, leading to a sparsity level of $\mathbf{W}$ of 79.56% (i.e., percentage of zero values in $\mathbf{W}$). Our estimation algorithm in Section 4 can be extended to include sampling steps for the parameters $\{\mathbf{s}, \mathbf{a}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \psi, \sigma_E^2\}$, where $\mathbf{a} = (a_1, \ldots, a_T)'$. By assigning suitable prior distributions to these parameters, we can use (23) to determine the conditional posterior distributions of these parameters, which will be conditional on the high-dimensional parameter $\mathbf{h}$. However, since our main focus is not the conditional mean equation of the outcome variable, we estimated all parameters $\{\mathbf{s}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \psi, \sigma_E^2\}$ using the maximum likelihood approach readily implemented in the spatial econometrics MATLAB toolbox (cf. Bivand and Piras 2015; LeSage 1999). The resulting estimates, including their asymptotic 95% confidence intervals and the direct/indirect/total effects, are summarized in Table 2. In addition, we report the estimated fixed effects in Figure 7.

The estimated parameters align with our expectations, and we see a good general fit of the model with $R^2 = 0.8161$. The pivotal variable with the most substantial total effect on PM concentrations is air pressure. It is noteworthy that air pressure exerts a positive influence on PM concentrations, implying that as air pressure increases, PM concentrations also rise. In addition, both temperature and relative humidity have a statistically significant and positive impact on PM concentrations. On the contrary, the maximum height of the planetary boundary layer has a negative effect, although this effect is not statistically significant. These findings align with a recent study by Otto et al. (2023) that compared various mean models

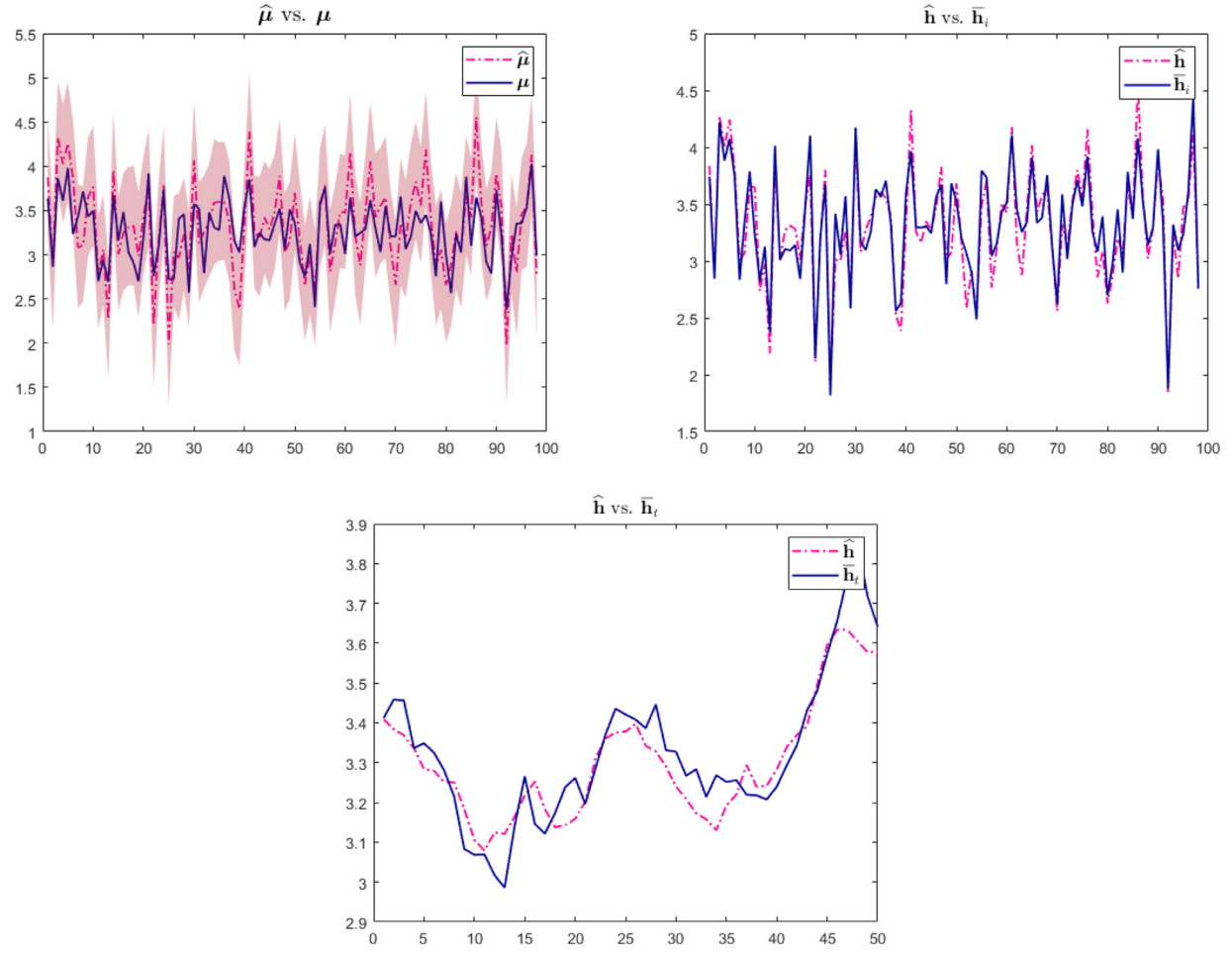

Fig. 4. Estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ (dashed lines) and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.30, 0.65, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

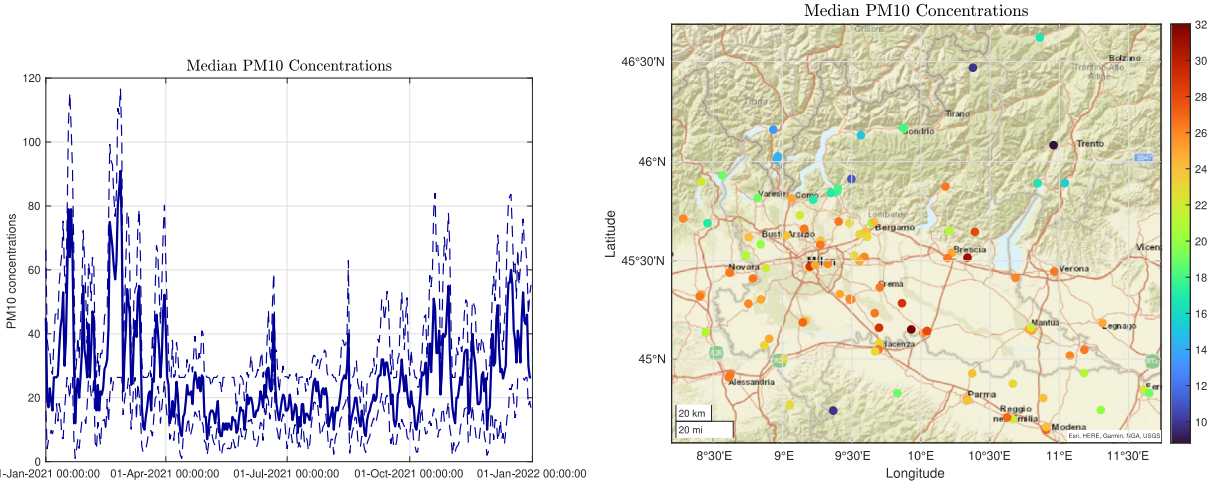

Fig. 5. Median PM_{10} concentrations ($\mu g/m^3$) and the 5% and 95% quantiles across space displayed as time series (left) and median concentrations across time shown on a map (right).

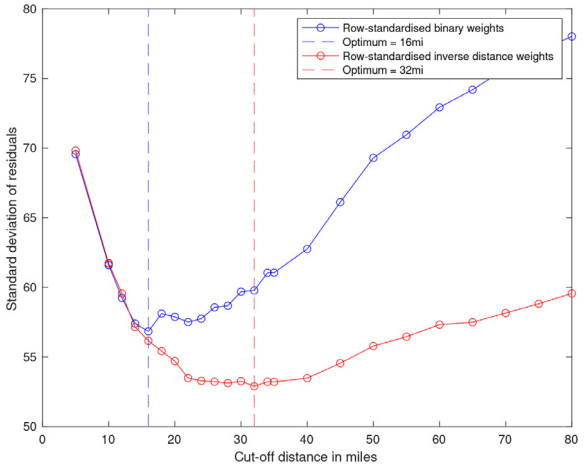

Fig. 6. Selection of the optimal cut-off distance for the mean model.

Table 2
Estimated parameters of the mean model.

	Estimate	Standard error	95% confidence interval (asymptotic)	Direct effect	Indirect effect	Total effect
Regressors						
β_1 (max PBL height)	-0.2101	0.3530	(-0.9021, 0.4818)	-0.2819	0.2203	-0.0616
β_2 (temperature)	-6.5128	1.2081	(-8.8807, -4.1449)	-6.2988	6.6933	0.3944
β_3 (relative humidity)	1.9376	0.2863	(1.3764, 2.4987)	1.7419	-1.2028	0.5391
β_4 (air pressure)	35.2495	4.1764	(27.0640, 43.4350)	34.0280	-20.1166	13.9114
Spatially lagged regressors						
γ_1 (max PBL height)	0.1736	0.4283	(-0.6658, 1.0131)			
γ_2 (temperature)	6.6011	1.4465	(3.7660, 9.4363)			
γ_3 (relative humidity)	-1.7913	0.3471	(-2.4715, -1.1110)			
γ_4 (air pressure)	-30.1394	5.0513	(-40.0397, -20.2392)			
Spatial dependence						
ψ	0.6470	0.0049	(0.6374, 0.6565)			
Error Variance						
σ_E^2	53.2667					
Coefficient of determination R^2	0.8161					

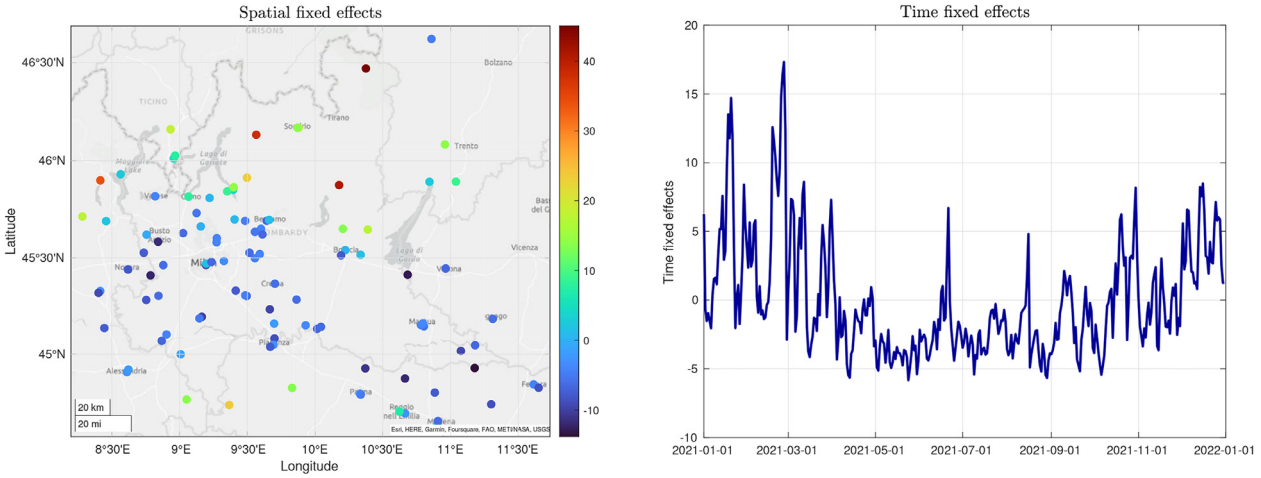

Fig. 7. Spatial (left) and temporal (right) fixed effects.

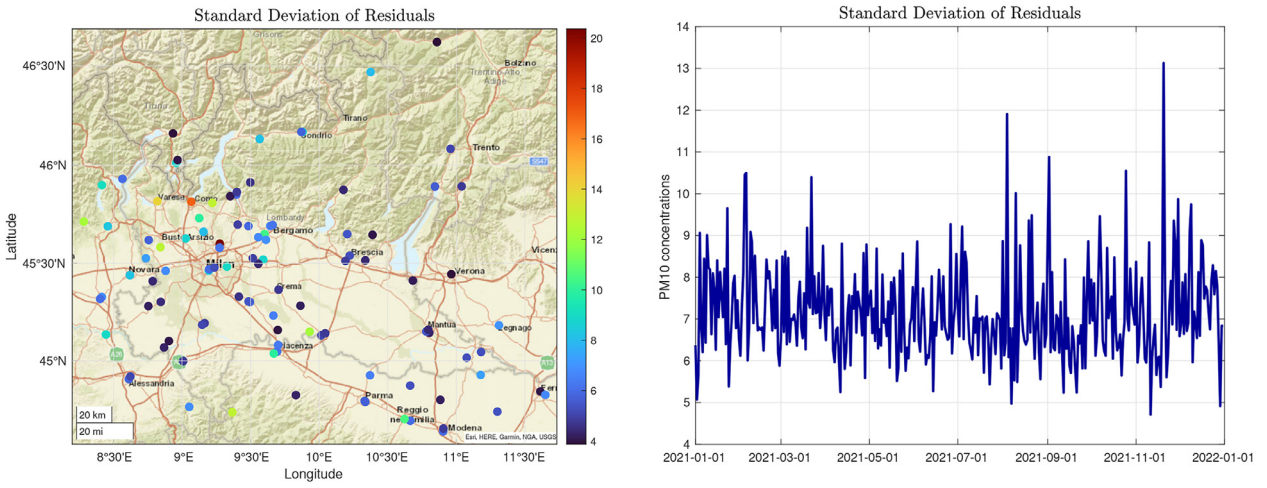

Fig. 8. Standard deviation of residuals $\mathbf{E}_t$ across space (left) and across time (right).

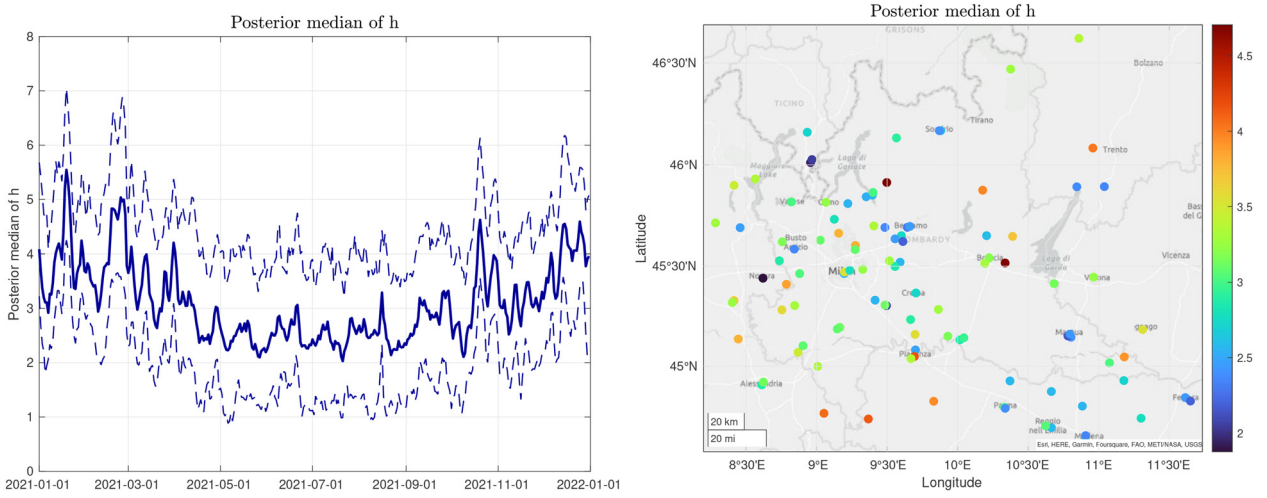

Fig. 9. Median posterior draws of $\mathbf{h}_t$ and the 5% and 95% quantiles across space displayed as time series (left) and median posterior draws of $\mathbf{h}_t$ across time shown on a map (right).

Table 3
Estimated parameters of the dynamic stochastic volatility model, medians and 95% posterior credible intervals.

	Median	95% credible interval
Intercept $\boldsymbol{\mu}$		
Average constant term	3.1679	(2.8751, 3.4505)
Space-time interactions		
ρ_1 (spatial)	0.6274	(0.6169, 0.6302)
ρ_2 (temporal)	0.3576	(0.3546, 0.3619)
ρ_3 (spatiotemporal)	-0.0113	(-0.0148, -0.0043)
Stochastic volatilities $\mathbf{h}_t$		
Average log-volatility	3.1475	(3.1307, 3.1647)
Error variance σ^2		
σ^2 (variance of $\mathbf{U}_t$)	0.3398	(0.3213, 0.3594)

on the same dataset. However, for this paper, our focus is not on the mean variations but on the model errors $\mathbf{E}_t$, i.e., the model uncertainties. Hence, we do not go into further detail on interpreting the mean models. The errors of the mean model represent the unexplained variations and, therefore, the environmental risks. When considering the standard deviation of $\mathbf{E}_t$ at each time point and/or for each measurement station, we see that the error variance varies across space and time (see Figure 8). There are time periods of increased variations, so-called volatility clusters, e.g., at the end of the year. Moreover, we observe a similar clustering across space. Measurement stations with higher volatility are located in close proximity, e.g.,

northeast of Milan or around Brescia. This provides motivation for estimating a dynamic stochastic volatility model for $\mathbf{E}_t$ in the second step, that is,

$$\mathbf{E}_t = \mathbf{H}_t^{1/2}\mathbf{V}_t \tag{32}$$

as defined in (1) and (2).

Table 3 reports the median of the posterior draws for each parameter, including the corresponding 95% credible intervals (4,000 posterior draws and 2,000 burn-in draws). The space-time interaction parameters ρ_1, ρ_2, and ρ_3 represent the degree of uncertainty/risk clusters and spillovers. In general, we observe moderate instantaneous spatial interactions with $\hat{\rho}_1 = 0.6274$ and temporal autoregressive interactions $\hat{\rho}_2 = 0.3576$, while spatiotemporal effects appear to be of minor importance (i.e., $\hat{\rho}_3 = -0.0113$). The posterior mean estimates for both the spatial and temporal autoregressive parameters are positive. That is, if the environmental risk (log-volatility h_{it}) is high, it is likely to influence neighboring regions and future time points. For this reason, spatial and temporal volatility clusters are formed. The negative signs of the posterior mean estimate for the spatiotemporal lag (i.e., the spillover or diffusion effect) oppose this behavior, but it is close to zero. Moreover, we observe that the spatial volatility clusters dominate the temporal ones. All space-time interaction parameters are significantly different from zero, indicating a space-time dependence in the local uncertainties of the environmental process (here: the particulate matter concentrations).

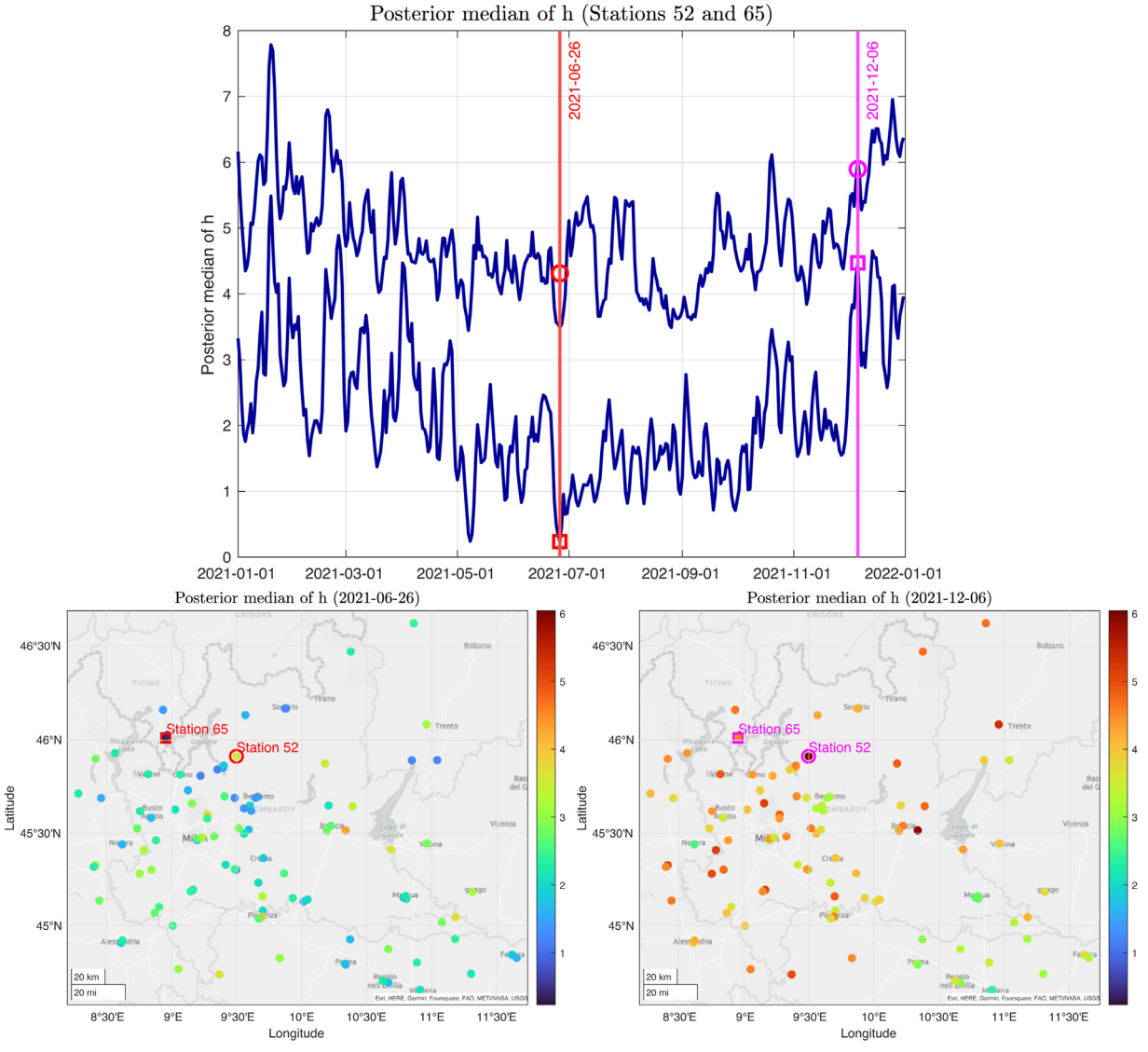

Fig. 10. Median posterior draws of $\mathbf{h}_t$ for two selected stations, Lugano in Switzerland (station 65) and Moggio (station 52), across time (top) and two selected time points 26 June 2021 (bottom left) and 6 December 2021 (bottom right). The log-volatility values highlighted with the red/magenta empty circle or square correspond to each other.

The overall volatility level represented by the average $\mathbf{h}_t$ of 3.1475 provides insights about the estimated overall uncertainty level across all stations and time points. If the individual log-volatilities h_{it} exceed the average volatility level, this specific location i has an increased environmental risk at time point t. Thus, we additionally analyzed the conditional volatilities for each station across the time horizon. The results are displayed in Figure 9 as a time-series plot (left) and on a map (right). More precisely, we computed the posterior medians of h_{it} for all locations and time points, leading to an estimate of the $n \times T$ matrix of the logarithmic conditional volatilities. Then, for visualization, we computed the median and the 5% and 95% quantiles of these estimates across all spatial locations (see time series plot) and across all time points (see map). From the time series plot, we see that there is a clear annual pattern – the model uncertainties are lower in summer and highest in winter. These varying conditional volatility levels, which are the highest in the winter season, can be interpreted as environmental risks. Similarly, measurement stations with greater uncertainty/risk can be identified from the map. They are mostly located in the south and the valleys in the mountain areas. These uncertainty/risk profiles can be obtained for each station, as we visualize for two selected locations in Figure 10. Station 52 near Moggio, Lombardy Italy, was one of the stations with the highest log volatilities, while Station 65, Lugano Switzerland, showed significantly lower volatilities. Moggio station is located in the mountains with unique climatic conditions that differ markedly from all other stations. This explains the larger log-volatilities for this station. In addition to the temporal variation (top plot in Figure 10), we also displayed the estimated risks for all other locations at two selected time points in the summer and winter (bottom plots in Figure 10).

7. Conclusion

We have introduced a novel spatiotemporal statistical model for stochastic volatilities, which are spatially, temporally, and spatiotemporally correlated and have different levels for each location, reflecting the heterogeneity of the process across space. The model includes two different error terms for the mean and the log-volatility equations. To estimate the parameters, we suggested a Bayesian MCMC approach. To this end, we applied a log-square transformation to transform our nonlinear state-space model into a linear one. Further, we used a Gaussian mixture distribution to approximate the distribution of the transformed error terms in the space equation, transforming our model into a linear Gaussian state-space model. We analyzed the estimation performance for different parameter settings and spatial interactions in a simulation study and showed that the suggested Bayesian sampler performs satisfactorily.

Moreover, we applied the dynamic spatiotemporal stochastic volatility model in a completely new empirical framework, namely, in the field of modeling environmental and climate risks. First, statistical modeling of the volatility process of climate variables has not been done extensively yet, even though there is a large scientific consensus that an increased variability of environmental processes, e.g., temperature variability, is harmful to the environment. Second, stochastic volatility models were predominantly applied to financial data because of the straightforward interpretation of the log conditional volatilities as the (return) risk of financial assets. We have transferred this idea to environmental risks and showed how the volatility of PM_{10} predictions are correlated across space and time. Accurate estimation of local model uncertainties is particularly relevant when making predictions of the fine dust concentration or assessing policy interventions, such as traffic or agricultural restrictions, in scenario analyses. In particular, measurement stations with an increased model uncertainty could be identified in this way. In addition, we showed in our application that there are significant temporal and spatial volatility clusters in these environmental risks. Also, the temporally lagged spatial spillovers, i.e., the spatiotemporal correlations, appear to play a minor role.

Both from a theoretical and applied perspective, spatiotemporal stochastic volatility models and environmental risk modeling are important topics for future research. The current model does not allow for temporally varying constant terms in the conditional volatilities, which are constant across space. Regarding the latter case, other environmental processes, like temperature variations, soil droughts, or atmospheric ozone concentration and optical depths, are important processes where a deep understanding of the spatiotemporal interactions in the variabilities is essential, both for obtaining accurate prediction intervals and the planning of interventions (and their impact on the environment).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

In this section, we provide additional simulation results under the following prior distributions: $\sigma^2 \sim \text{IG}(0.01, 0.01)$ and $\boldsymbol{\mu} \sim N(\mathbf{0}, 100\mathbf{I}_n)$. Figures 11 and 13 show the trace plots for $\boldsymbol{\rho}$ and σ^2, while Figures 12 and 14 show the estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ along with their true values. Figure 15 shows estimated volatility terms at certain locations. It shows the line plots of $\{\hat{h}_{13,t}\}_{t=1}^{50}$, $\{\hat{h}_{34,t}\}_{t=1}^{50}$, $\{\hat{h}_{73,t}\}_{t=1}^{50}$, $\{\hat{h}_{91,t}\}_{t=1}^{50}$ against their true corresponding values. Finally, Figure 15 shows the line plots of $\{\hat{h}_{i,11}\}_{i=1}^{98}$, $\{\hat{h}_{i,23}\}_{i=1}^{98}$, $\{\hat{h}_{i,35}\}_{i=1}^{98}$, $\{\hat{h}_{i,43}\}_{i=1}^{98}$ against their true corresponding values.

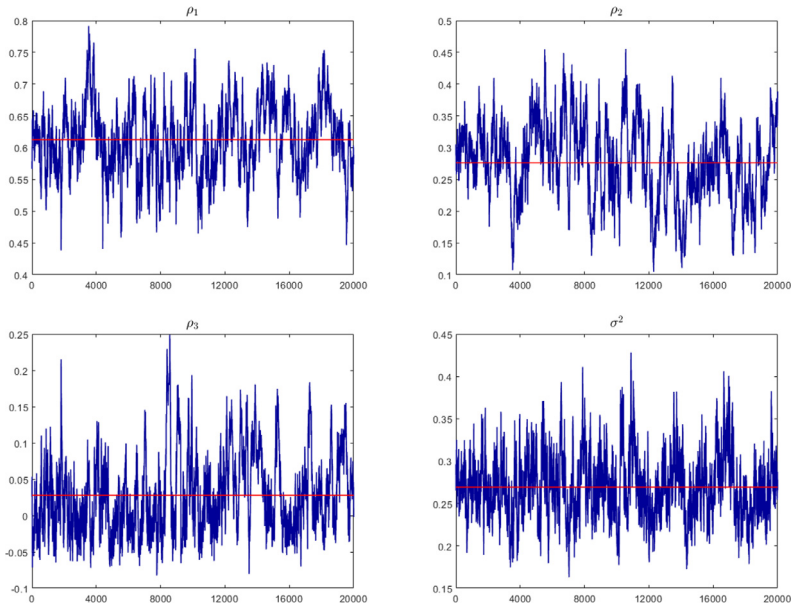

Fig. 11. Trace plots for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

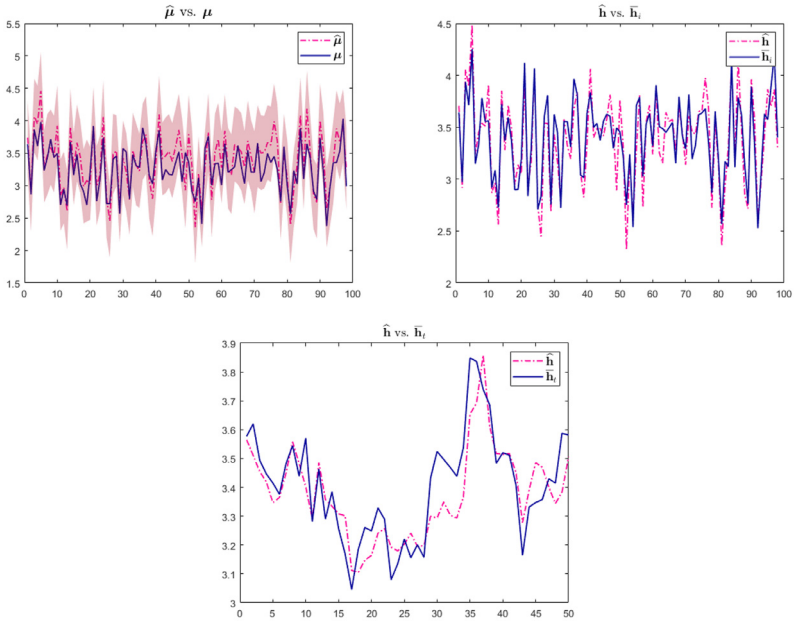

Fig. 12. Estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ (dashed lines) and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

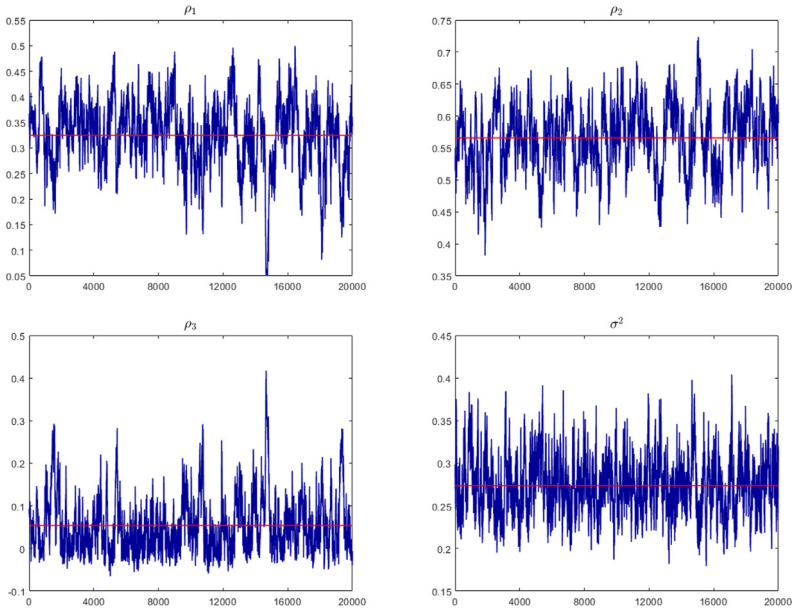

Fig. 13. Trace plots for $\boldsymbol{\rho} = (0.30, 0.65, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

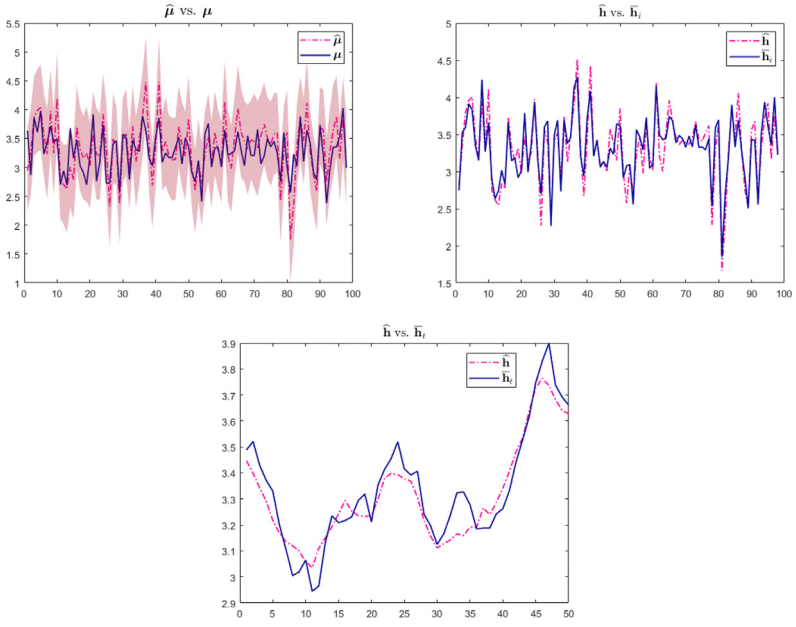

Fig. 14. Estimates of $\boldsymbol{\mu}$ and $\mathbf{h}$ (dashed lines) and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.30, 0.65, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

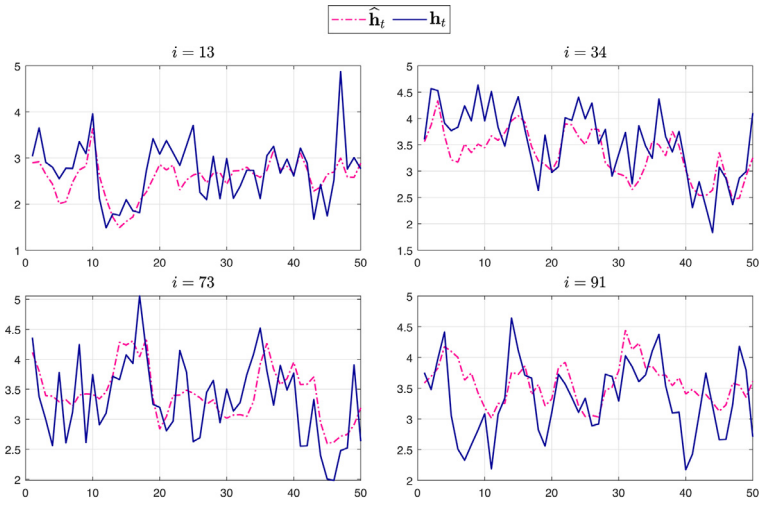

Fig. 15. Estimates of $\mathbf{h}_t$ (dashed lines) for four arbitrarily chosen entities and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

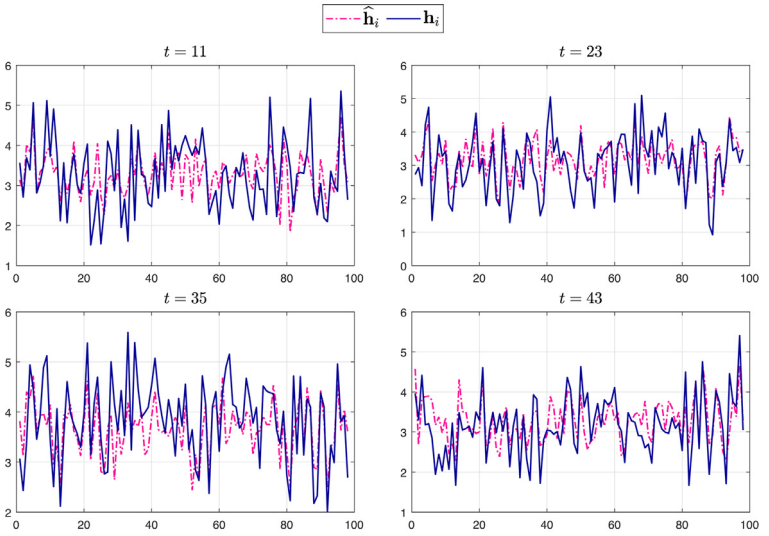

Fig. 16. Estimates of $\mathbf{h}_i$ (dashed lines) for four arbitrarily chosen periods and their corresponding true data-generating values (solid lines) for $\boldsymbol{\rho} = (0.6, 0.35, -0.025)$ and $\mathbf{W}$ is a queen contiguity matrix.

References

Anselin, L., 1988. Spatial econometrics: Methods and Models. Springer, New York.
Arbia, G., Espa, G., 1996. Effects of the maup on image classification. Journal of Geographical Systems.
Banerjee, S., 2020. Modeling massive spatial datasets using a conjugate Bayesian linear modeling framework. Spatial statistics 37, 100417.
Bivand, R., Piras, G., 2015. Comparing implementations of estimation methods for spatial econometrics. Journal of Statistical Software 63, 1–36.
Chan, J.C., Jeliazkov, I., 2009. Efficient simulation and integrated likelihood estimation in state space models. International Journal of Mathematical Modelling and Numerical Optimisation 1 (1-2), 101–120.
Chan, J.C.C., 2017. The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. Journal of Business & Economic Statistics 35 (1), 17–28.
Chib, S., Nardari, F., Shephard, N., 2002. Markov chain monte carlo methods for stochastic volatility models. Journal of Econometrics 108 (2), 281–316.
Cressie, N., Wikle, C.K., 2015. Statistics for spatio-temporal data. John Wiley & Sons.
Elhorst, J.P., 2014. Spatial Econometrics: From Cross-Sectional Data to Spatial Panels. Springer Berlin Heidelberg, New York.
Fassò, A., Maranzano, P., Otto, P., 2022. Spatiotemporal variable selection and air quality impact assessment of Covid-19 lockdown. Spatial Statistics 49.
Fassò, A., Rodeschini, J., Fusta Moro, A., Shaboviq, Q., Maranzano, P., Cameletti, M., Finazzi, F., Golini, N., Ignaccolo, R., Otto, P., 2023a. Agrimonia: Open access dataset correlating livestock and air quality in the Lombardy region, Italy. 10.5281/zenodo.7956006
Fassò, A., Rodeschini, J., Moro, A.F., Shaboviq, Q., Maranzano, P., Cameletti, M., Finazzi, F., Golini, N., Ignaccolo, R., Otto, P., 2023. Agrimonia: a dataset on livestock, meteorology and air quality in the Lombardy region, Italy. Scientific Data 10, 143.
Fisher, R.A., 1935. The design of experiments. Oliver and Boyd, Edinburgh.
Giglio, S., Kelly, B., Stroebel, J., 2021. Climate finance. Annual Review of Financial Economics 13, 15–36.

Gneiting, T., Genton, M.G., Guttorp, P., 2007. Geostatistical space-time models, stationarity, separability and full symmetry. In: Finkenstadt, B., Held, L., Isham, V. (Eds.), Statistical Methods for Spatio-Temporal Systems. Chapman and Hall/CRC, pp. 151–175.
Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive metropolis algorithm. Bernoulli 7 (2), 223–242.
Hamilton, J., 1994. Time Series Analysis. Princeton University Press.
Han, X., Lee, L.-F., 2016. Bayesian analysis of spatial panel autoregressive models with time-varying endogenous spatial weight matrices, common factors, and random coefficients. Journal of Business & Economic Statistics 34 (4), 642–660.
Hølleland, S., Karlsen, H.A., 2020. Decline in temperature variability on Svalbard. Journal of Climate 33 (19), 8475–8486.
Hølleland, S., Karlsen, H.A., 2020. A stationary spatio-temporal GARCH model. Journal of Time Series Analysis 41 (2), 177–209.
Hong, H., Karolyi, G.A., Scheinkman, J.A., 2020. Climate finance. The Review of Financial Studies 33 (3), 1011–1023.
Horn, R., Johnson, C., 2012. Matrix Analysis. Cambridge University Press.
Huang, W., Wang, K., Jay Breidt, F., Davis, R.A., 2011. A class of stochastic volatility models for environmental applications. Journal of Time Series Analysis 32 (4), 364–377.
Huntingford, C., Jones, P.D., Livina, V.N., Lenton, T.M., Cox, P.M., 2013. No increase in global temperature variability despite changing regional patterns. Nature 500 (7462), 327–330.
Iaco, S.D., Maggio, S., Palma, M., Posa, D., 2012. Advances in spatio-temporal modeling and prediction for environmental risk assessment. In: Haryanto, B. (Ed.), Air Pollution. IntechOpen, Rijeka doi:10.5772/51227.
Jurek, M., Katzfuss, M., 2021. Multi-resolution filters for massive spatio-temporal data. Journal of Computational and Graphical Statistics 30 (4), 1095–1110.
Katzfuss, M., Cressie, N., 2011. Spatio-temporal smoothing and EM estimation for massive remote-sensing data sets. Journal of Time Series Analysis 32 (4), 430–446.
Kelejian, H.H., Prucha, I.R., 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics 157, 53–67.
Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: Likelihood inference and comparison with arch models. The Review of Economic Studies 65 (3), 361–393.
Lee, L.-F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72 (6), 1899–1925.
Lee, L.-F., Yu, J., 2015. Estimation of fixed effects panel regression models with separable and nonseparable space-time filters. Journal of Econometrics 184 (1), 174–192.
LeSage, J., Pace, R.K., 2009. Introduction to Spatial Econometrics (Statistics: A Series of Textbooks and Monographs. Chapman and Hall/CRC, London.
LeSage, J. P., 1999. Spatial econometrics using MATLAB. The Web: www.econ.utoledo.edu.
Maranzano, P., 2022. Air quality in Lombardy, Italy: An overview of the environmental monitoring system of ARPA Lombardia. Earth 3 (1), 172–203.
Omori, Y., Chib, S., Shephard, N., Nakajima, J., 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. Journal of Econometrics 140 (2), 425–449.
Otto, P., Fusta Moro, A., Rodeschini, J., Shaboviq, Q., Ignaccolo, R., Golini, N., Cameletti, M., Maranzano, P., Finazzi, F., Fassò, A., 2023. Spatiotemporal modelling of $PM_{2.5}$ concentrations in Lombardy (Italy)–a comparative study. arXiv preprint arXiv:2309.07285.
Otto, P., Schmid, W., Garthoff, R., 2018. Generalised spatial and spatiotemporal autoregressive conditional heteroscedasticity. Spatial Statistics 26, 125–145.
Paaijmans, K.P., Heinig, R.L., Seliga, R.A., Blanford, J.I., Blanford, S., Murdock, C.C., Thomas, M.B., 2013. Temperature variation makes ectotherms more sensitive to climate change. Global change biology 19 (8), 2373–2380.
Parent, O., LeSage, J.P., 2011. A space-time filter for panel data models containing random effects. Computational Statistics & Data Analysis 55 (1), 475–490.
Parent, O., LeSage, J.P., 2012. Spatial dynamic panel data models with random effects. Regional Science and Urban Economics 42 (4), 727–738.
Porcu, E., Bevilacqua, M., Genton, M.G., 2016. Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. Journal of the American Statistical Association 111 (514), 888–898.
Roberts, G.O., Rosenthal, J.S., 2009. Examples of adaptive mcmc. Journal of Computational and Graphical Statistics 18 (2), 349–367.
Robinson, P.M., 2009. Large-sample inference on spatial dependence. Econometrics Journal 12.
Sato, T., Matsuda, Y., 2021. Spatial extension of generalized autoregressive conditional heteroskedasticity models. Spatial Economic Analysis 16 (2), 148–160.
Screen, J.A., 2014. Arctic amplification decreases temperature variance in northern mid-to high-latitudes. Nature Climate Change 4 (7), 577–582.
Taşpınar, S., Doğan, O., Chae, J., Bera, A.K., 2021. Bayesian inference in spatial stochastic volatility models: An application to house price returns in Chicago. Oxford Bulletin of Economics and Statistics 83 (5), 1243–1272.
Tewksbury, J.J., Huey, R.B., Deutsch, C.A., 2008. Putting the heat on tropical animals. Science 320 (5881), 1296–1297.
Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Economic geography 46 (sup1), 234–240.
Vasseur, D.A., DeLong, J.P., Gilbert, B., Greig, H.S., Harley, C.D., McCann, K.S., Savage, V., Tunney, T.D., O'Connor, M.I., 2014. Increased temperature variation poses a greater risk to species than climate warming. Proceedings of the Royal Society B: Biological Sciences 281 (1779), 20132612.
Ver Hoef, J.M., Hanks, E.M., Hooten, M.B., 2018. On the relationship between conditional (CAR) and simultaneous (SAR) autoregressive models. Spatial statistics 25, 68–85.
Ver Hoef, J.M., Peterson, E.E., Hooten, M.B., Hanks, E.M., Fortin, M.-J., 2018. Spatial autoregressive models for statistical inference from ecological data. Ecological Monographs 88 (1), 36–59.
Wang, W., Lee, L.-F., 2018. GMM estimation of spatial panel data models with common factors and a general space-time filter. Spatial Economic Analysis 13 (2), 247–269.
Wilby, R.L., Troni, J., Biot, Y., Tedd, L., Hewitson, B.C., Smith, D.M., Sutton, R.T., 2009. A review of climate risk information for adaptation and development planning. International Journal of Climatology: A Journal of the Royal Meteorological Society 29 (9), 1193–1215.
Yan, J., 2007. Spatial stochastic volatility for lattice data. Journal of Agricultural, Biological, and Environmental Statistics 12 (1), 25–40.
Yang, Y., Doğan, O., Taşpınar, S., 2023. Observed-data DIC for spatial panel data models. Empirical Economics 64 (3), 1281–1314.
Zimmerman, D.L., 2019. Geostatistics. Wiley StatsRef: Statistics Reference Online 1–10.