

Ancillary Heteroskedastic Structural Model

Ivair Silva

ivair@ufop.edu.br

Universidade Federal de Ouro Preto

Yan Zhuang

Connecticut College

Debanjan Bhattacharjee

Utah Valley University

Research Article

Keywords: Climate change, Linear structural model, Ancillary series, Forecasting

Posted Date: June 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10033791/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Ancillary Heteroskedastic Structural Model

Ivair Silva^{1*}, Yan Zhuang^{2†} and Debanjan Bhattacharjee^{3†}

^{1*}Computer Department, Federal University of Ouro Preto, MG, Brazil.

²Department of Mathematics and Statistics, Connecticut College, CT,
USA.

³Department of Mathematics, Utah Valley University, UT, USA.

*Corresponding author(s). E-mail(s): ivair@ufop.edu.br;

Contributing authors: yzhuang@conncoll.edu; debanjanB@uvu.edu;

[†]These authors contributed equally to this work.

Abstract

We consider modeling the relationships between a response time series of each given area in a map and its lagged variations in the relevant neighboring regions. This new modeling approach does not assume cross-sectional homoscedasticity, that is, the variability of each area is estimated as well as the area-specific impacts of time-variable covariates. The construction is based on embedding a regression structure in the level term of the linear structural model, which enables the derivation of ancillary time series, solving the point and interval estimation by means of the conventional likelihood principle. The new model is illustrated with an application to study the rise in global temperature and the use of fossil fuels, along with carbon emissions, of seven South American countries.

Keywords: Climate change, Linear structural model, Ancillary series, Forecasting

1 Introduction

The modeling of spatial-time data finds meaningful applications in almost every research field, including economics, finance, social studies, health studies, etc. In the fields of medicine and public health, figuring out geographic areas with high incidences or mortality of a specific disease would be very important. Disease mapping helps researchers to have a better understanding of the regional patterns of disease. [1] discussed the spatial data analysis used for disease mapping. In environmental studies, [2] developed a spatial analysis method for predicting landslide hazards. In

economics, there is a vast literature on street-level consumer behaviors, retailer spatial strategies, etc. [3] used spatial approach to study New Orleans business recovery after a disaster, hurricane Katrina.

The approaches for spatial data modeling vary according to many aspects. [4] studied the estimation and predictive performance of several estimators for the time-space dynamic panel data model with a spatial moving average random effects structure of the disturbances. [5] compared different methods for estimating spatial autoregressive panel models with randomly missing data in the dependent variable. They started with a random effects model and then generalized the model by introducing the spatial Mundlak approach. The model was estimated through a nonlinear least squares method and a generalized method of moments. [6] considered a class of spatio-temporal models by allowing the scalar coefficients for each location (or panel) to be different from each other. They proposed a generalized Yule–Walker estimation method by applying the least squares estimation to a Yule–Walker equation. [7] provided a model for daily temperature time series. Their model separated fixed effects from random effects and also separated space, years, and days for each type of effect. [8] considered identification, estimation, and inference for spatial panel data models with heterogeneous spatial lag coefficients, with and without (weakly) exogenous regressors, and subject to heteroskedastic errors. They provided a quasi-maximum likelihood estimation procedure for their models. [9] proposed a fixed effects partially linear additive spatial autoregressive model with space-time correlated disturbances. Their model captured spatial dependence of response, linearity and nonlinearity between response and regressors, spatial and serial correlations of disturbances. More generally, for instance, to model the spatial self-correlation of a time series and, simultaneously, the fixed and random effects from covariates, one can refer to [10], [11], [12], and [13], and [14]. For taking into account the auto-correlation structures, one is referred to [15], [16], [17], and [18].

In this paper, we introduce a model that is meant to take into account:

- the influence of the lagged time series variations of neighboring areas.
- the cross-sectional heteroscedasticity among the areas of the map.
- the area-specific set of covariates with random-effects.
- the covariates with fixed-effects.
- the area-specific autocorrelation structure.
- the area-specific trend and seasonal terms.

The main contribution of our work lies in the development of a comprehensive spatio-temporal panel data model that jointly captures spatial dependence, temporal dynamics, and heterogeneity across areas. Our approach of modeling extends existing dynamic spatial panel models by introducing area-level heterogeneity in both the mean and variance processes, and by explicitly modeling temporal decomposition into trend, seasonal, and stochastic components. The proposed model is useful, for instance, to comprehend how the income average of a given area in a geographical region may depend upon the economic dynamics among its neighboring areas. The correlation among the incomes of the regions in the map can also be explained by common sources of variation, such as countercyclical policies. Still, intrinsic stochastic

components, such as trend, seasonality, and autocorrelation, are usually embedded in the unobservable mechanism that generates the incomes. In addition, the statistically relevant covariates associated with the income variation of an area A may not be statistically relevant to area B. Estimating the parameters of such a tri-dimensional structure (space, time, and area-specific sets of covariates) is a key step for studying the *spin-off* effect. In econometrics terms, the *spin-off* effect denotes a spatial or sectoral spillover whereby improvements or innovations in one area induce measurable positive externalities in adjacent areas through channels such as knowledge transfer, capital flows, or production linkages.

The rest of this paper is organized as follows. Section 2 proposed the spatial-time heteroskedastic linear structural model, as well as the differencing strategy to generate a spatial-time variable, which makes the state parameters in the model ancillary. In Section 3, focusing on the model estimation, we provided the maximum likelihood estimation for all the parameters in the model, including their confidence interval estimations and hypothesis tests, as well as estimating the states. In Section 4, we discussed the steps for conducting forecasting, following the estimations we provided in Section 3. Section 5 is a simulation study. We started with a simulation set-up in Section 5.1, and then summarized the estimates in Sections 5.2 and 5.3 and provided the forecasting results in Section 5.4. We gave a real data application on exploring the association between the rise in global temperature and the use of fossil fuels, along with carbon emission, using data from seven South American countries. We concluded our work with some concluding remarks in Section 7.

2 Spatial-time Heteroskedastic Linear Structural Model

We assume that the response time series at time t , with $t = 1, 2, \dots, n$, for the i -th area in the map, $i = 1, \dots, M$, denoted by $Y_{i,t}$, is explained by deterministic and stochastic effects. The deterministic part is defined by the particular characteristics of each area, and the stochastic effect is the sum of two components: the common source of variations that all the areas perceive in t , and their reactions to the variations of neighboring areas in the recent past. Such a dynamic spatial-time behavior is modeled with the following linear structure:

$$\begin{aligned}
Y_{i,t} &= \mu_{i,t} + S_{i,t} + \epsilon_{i,t}, \\
\mu_{i,t} &= \mu_{i,t-1} + \beta_{i,t} + \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} (Y_{r,t-j} - Y_{r,t-j-1}) + \sum_{c \in C_i} \varphi_{i,c} x_{c,t} + \sum_{q=1}^Q \psi_q \chi_{q,t} + \eta_{i,t} \\
\beta_{i,t} &= \beta_{i,t-1} + \xi_{i,t}, \\
S_{i,t} &= S_{i,t-1} + \nu_{i,t},
\end{aligned} \tag{1}$$

where:

- $\mu_{i,t}$ is the average of $Y_{i,t}$ in t ,
- $\beta_{i,t}$ and $S_{i,t}$ are the trend and seasonal components of $Y_{i,t}$, respectively,

- $\theta_{i,r,j}$ is the reaction of $Y_{i,t}$ to the lagged variations $(Y_{r,t-j} - Y_{r,t-j-1})$ of its neighbor r , where $p_{i,r}$ is the number of such significant lags. Here R_i denotes the number of neighbors of area i .
- $\varphi_{i,c}$ is the reaction of $Y_{i,t}$ to the covariate $x_{c,t}$, with c taken in the set, C_i , of covariates with non-fixed effects measured for the i -th area in time t .
- ψ_q is the reaction of $Y_{i,t}$ to the fixed effect covariate $\chi_{q,t}$ measured in time t , where Q is the number of covariates with fixed effects,
- $\epsilon_{i,t}$, $\eta_{i,t}$, $\xi_{i,t}$ and $\nu_{i,t}$ are independent white noises.

2.1 Ancillary Statistics for $\mu_{i,t}$, $\beta_{i,t}$, and $S_{i,t}$

It is not necessary to estimate the states $\mu_{i,t}$, $\beta_{i,t}$, and $S_{i,t}$ to estimate $\theta_{i,r,j}$, $\varphi_{i,c}$, and ψ_q . This is possible by constructing ancillary $W_{i,j}$ statistics with respect to $\mu_{i,t}$, $\beta_{i,t}$, and $S_{i,t}$.

The conventional approach of applying simple consecutive differences of $Y_{i,t}$ to remove level, trend and seasonality would produce serial correlation, leading to intricate inference. Instead, for the construction of ancillary terms with respect to $\mu_{i,t}$, $\beta_{i,t}$, and $S_{i,t}$, we introduce a particular type of differencing process. It consists of taking disjoint pair of observations in a way that the resulting noise terms are independent. For this, define:

$$V_{i,l} = \begin{cases} Y_{i,2l} - Y_{i,2l-1}, & \text{if } n/2 \text{ is integer,} \\ Y_{i,2l+1} - Y_{i,2l}, & \text{otherwise,} \end{cases} \quad (2)$$

with $l = 1, \dots, N$. If n is an even number, then $N = n/2$, otherwise, $N = (n-1)/2$.

Now, define:

$$W_{i,u} = \begin{cases} V_{i,2u} - V_{i,2u-1}, & \text{if } N/2 \text{ is integer,} \\ V_{i,2u+1} - V_{i,2u}, & \text{otherwise,} \end{cases} \quad (3)$$

with $u = 1, \dots, k$. If N is an even number, then $k = N/2$, otherwise, $k = (N-1)/2$.

The distribution of $W_{i,u}$ does not depend on the states $\mu_{i,n}$, $\mu_{i,n-1}, \dots$, $\beta_{i,n}$, $\beta_{i,n-1}, \dots$, and $S_{i,n}$, $S_{i,n-1}, \dots$. To see this, without loss of generality, consider that n and N are even numbers. Thus,

$$\begin{aligned} V_{i,l} &= Y_{i,2l} - Y_{i,2l-1} \\ &= \mu_{i,2l} + S_{i,2l} + \epsilon_{i,2l} - \mu_{i,2l-1} - S_{i,2l-1} - \epsilon_{i,2l-1} \\ &= \mu_{i,2l-1} + \beta_{i,2l} + \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} (Y_{r,2l-j} - Y_{r,2l-j-1}) + \sum_{c \in C_i} \varphi_{i,c} x_{c,2l} + \sum_{q=1}^Q \psi_q \chi_{q,2l} + \\ &\quad + \eta_{i,2l} + S_{i,2l-1} + \nu_{i,2l} + \epsilon_{i,2l} - \mu_{i,2l-1} - S_{i,2l-1} - \epsilon_{i,2l-1} \\ &= \beta_{i,2l} + \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} (Y_{r,2l-j} - Y_{r,2l-j-1}) + \sum_{c \in C_i} \varphi_{i,c} x_{c,2l} + \sum_{q=1}^Q \psi_q \chi_{q,2l} + \\ &\quad + \epsilon_{i,2l} - \epsilon_{i,2l-1} + \eta_{i,2l} + \nu_{i,2l}. \end{aligned} \quad (4)$$

Therefore, for $2u = j^*$, :

$$\begin{aligned}
W_{i,u} &= V_{i,2u} - V_{i,2u-1} \\
&= \beta_{i,2j^*-2} + \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} V_{r,2u-j/2} + \sum_{c \in C_i} \varphi_{i,c} x_{c,2j^*} + \sum_{q=1}^Q \psi_q \chi_{q,2j^*} \\
&\quad + \xi_{i,2j^*-1} + \xi_{i,2j^*} + \epsilon_{i,2j^*} - \epsilon_{i,2j^*-1} + \eta_{i,2j^*} + \nu_{i,2j^*} \\
&\quad - \beta_{i,2j^*-2} - \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} V_{r,2u-1-j/2} - \sum_{c \in C_i} \varphi_{i,c} x_{c,2j^*-2} - \sum_{q=1}^Q \psi_q \chi_{q,2j^*-2} \\
&\quad - \epsilon_{i,2j^*-2} + \epsilon_{i,2j^*-3} - \eta_{i,2j^*-2} - \nu_{i,2j^*-2}.
\end{aligned} \tag{5}$$

With $\beta_{i,2j^*} = \beta_{i,2j^*-1} + \xi_{i,2j^*} = \beta_{i,2j^*-2} + \xi_{i,2j^*-1} + \xi_{i,2j^*}$, we obtain:

$$\begin{aligned}
W_{i,u} &= \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} (V_{r,2(u-j/4)} - V_{r,2(u-j/4)-1}) + \sum_{c \in C_i} \varphi_{i,c} (x_{c,4u} - x_{c,4u-2}) \\
&\quad + \sum_{q=1}^Q \psi_q (\chi_{q,4u} - \chi_{q,4u-2}) + \xi_{i,4u} + \xi_{i,4u-1} + \epsilon_{i,4u} - \epsilon_{i,4u-1} - \epsilon_{i,4u-2} + \epsilon_{i,4u-3} \\
&\quad + \eta_{i,4u} - \eta_{i,4u-2} + \nu_{i,4u} - \nu_{i,4u-2} \\
&= \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} W_{r,u-j/4} + \sum_{c \in C_i} \varphi_{i,c} (x_{c,4u} - x_{c,4u-2}) + \sum_{q=1}^Q \psi_q (\chi_{q,4u} - \chi_{q,4u-2}) + a_{i,u},
\end{aligned} \tag{6}$$

where $a_{i,u} = \xi_{i,4u} + \xi_{i,4u-1} + \epsilon_{i,4u} - \epsilon_{i,4u-1} - \epsilon_{i,4u-2} + \epsilon_{i,4u-3} + \eta_{i,4u} - \eta_{i,4u-2} + \nu_{i,4u} - \nu_{i,4u-2}$.

Important: $a_{i,u}$ and $a_{i,l}$ are formed by independent terms for $u \neq l$. This is so because they are formed by the sum of disjoint sets of noise terms. To see this, consider to examine the first two terms. For $u = 1$:

$$\begin{aligned}
a_{i,1} &= \xi_{i,4} + \xi_{i,3} + \\
&\quad + \epsilon_{i,4} - \epsilon_{i,3} - \epsilon_{i,2} + \epsilon_{i,1} + \\
&\quad + \eta_{i,4} - \eta_{i,2} + \\
&\quad + \nu_{i,4} - \nu_{i,2}.
\end{aligned}$$

For $u = 2$:

$$\begin{aligned}
a_{i,2} &= \xi_{i,8} + \xi_{i,7} + \\
&\quad + \epsilon_{i,8} - \epsilon_{i,7} - \epsilon_{i,6} + \epsilon_{i,5} + \\
&\quad + \eta_{i,8} - \eta_{i,6} + \\
&\quad + \nu_{i,8} - \nu_{i,6}.
\end{aligned}$$

Note that $a_{i,1}$ and $a_{i,2}$ have no noise terms in common. Hence, they are independent. The next terms, $a_{i,3}, a_{i,4}, a_{i,5}, \dots$, will have disjoint error terms as well.

From the pairwise independence of terms $a_{i,1}, a_{i,2}, a_{i,3}, \dots$, for fixed observed $V_{r,4u-j} - V_{r,4u-j-1} = v_{r,4u-j} - v_{r,4u-j-1}$, with $r \in R_i$ and $j = 1, \dots, p_{i,r}$, the variables $W_{i,1}, \dots, W_{i,k}$ are pairwise independent.

Furthermore, from (6), the expected value of $W_{i,u}$, say λ_i , based on given values of its previous $p_{i,r}$ terms, is given by:

$$\lambda_i \equiv \mathbb{E}(W_{i,u}) = \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} w_{r,u-j/4} + \sum_{c \in C_i} \varphi_{i,c} (x_{c,4u} - x_{c,4u-2}) + \sum_{q=1}^Q \psi_q (\chi_{q,4u} - \chi_{q,4u-2}), \quad (7)$$

and its variance is $\text{Var}(a_{i,u}) = \sigma_i^2$. Therefore, the distribution of $W_{i,u}$ does not depend of $\mu_{l,t}, \beta_{l,t}$, and $S_{l,t}$ for every $l \in \{1, \dots, k\}$. And we can break it down into the following three parts: 1) the spatial interaction between the i -th area and its neighbors involving θ parameters and the W values from the previous terms; 2) covariates involving φ parameters and x ; and 3) fixed effects involving ψ parameters and χ .

3 Inference

Because $W_{i,u}$ is ancillary for $\mu_{i,t}, \beta_{i,t}$, and $S_{i,t}$, the point estimation process for $\theta_{i,r,j}, \varphi_{i,c}$ and ψ_q is conducted in a first stage without the need to estimate the state parameters $\mu_{i,t}, \beta_{i,t}$, and $S_{i,t}$.

3.1 Maximum Likelihood Estimators for $\theta_{i,r,j}, \varphi_{i,c}, \psi_q$, and σ_i^2

Let θ_i denote the parameter vector with the $\theta_{i,r,j}$ coefficients for the outcomes of the relevant areas for area i , and φ_i denote the parameter vector with the $\varphi_{i,c}$ coefficients for the set of covariates relevant for area i .

Let $\vec{\phi} = (\theta_i, \varphi_i, \psi_q, \sigma_i^2)$ denote the vector of parameters to estimate. We assume that $a_{i,u}$ is a Gaussian noise with variance σ_i^2 for $i = 1, \dots, k$. Therefore, from (6) and (7), $W_{i,u} \sim N(\lambda_i, \sigma_i^2)$.

Let $\vec{W}_1 = (W_{1,1}, W_{2,1}, \dots, W_{M,1})$ denote the ancillary vector related to areas 1 to M for the first observable time, $\vec{W}_2 = (W_{1,2}, W_{2,2}, \dots, W_{M,2})$ for the second observable time, and so on. Thus, we have the sequence of vectors $\vec{W}_1, \vec{W}_2, \dots, \vec{W}_k$ for the observable times 1 to k .

Define $p = \max p_{i,r}$. For fixed observations $\vec{W}_1 = \vec{w}_1, \dots, \vec{W}_p = \vec{w}_p$, the likelihood based on the observations $\vec{W}_{p+1} = \vec{w}_{p+1}, \dots, \vec{W}_k = \vec{w}_k$ is given by:

$$\begin{aligned} L(\vec{\phi} | \vec{w}_{p+1}, \dots, \vec{w}_k) &= Pr[\vec{W}_{p+1} = \vec{w}_{p+1}, \dots, \vec{W}_k = \vec{w}_k | \vec{W}_1 = \vec{w}_1, \dots, \vec{W}_p = \vec{w}_p] \\ &= Pr[\vec{W}_{p+1} = \vec{w}_{p+1} | \vec{W}_1 = \vec{w}_1, \dots, \vec{W}_p = \vec{w}_p] \times \dots \end{aligned}$$

$$\begin{aligned}
& \times Pr[\vec{W}_{p+2} = \vec{w}_{p+2} | \vec{W}_2 = \vec{w}_2, \dots, \vec{W}_{p+1} = \vec{w}_{p+1}] \times \dots \\
& \dots \\
& \times Pr[\vec{W}_k = \vec{w}_k | \vec{W}_{k-p} = \vec{w}_{k-p}, \dots, \vec{W}_{k-1} = \vec{w}_{k-1}] \\
& = \prod_{u=p+1}^k \prod_{i=1}^M \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(w_{i,u}-\lambda_i)^2/(2\sigma_i^2)}. \tag{8}
\end{aligned}$$

To obtain the maximum likelihood estimates for the parameters $\theta_{i,r,j}$, $\varphi_{i,c}$ and ψ_q , one can use the Newton-Raphson method [19] as the likelihood is twice differentiable with respect to these parameters. For this, let $l(\vec{\phi})$ denote the log-likelihood, then, from (8):

$$l(\vec{\phi}) \propto \sum_{u=p+1}^k \sum_{i=1}^M -\frac{(w_{i,u} - \lambda_i)^2}{2\sigma_i^2} - \frac{k-p}{2} \sum_{i=1}^M \ln \sigma_i^2. \tag{9}$$

The first and the second partial derivatives of $l(\vec{\phi})$ with respect to each of the parameters are:

$$\frac{\partial}{\partial \theta_{i,r,j}} l(\vec{\phi}) = \frac{1}{2\sigma_i^2} \sum_{u=p+1}^k (w_{i,u} - \lambda_i) w_{r,4u-j}. \tag{10}$$

$$\frac{\partial^2}{\partial \theta_{i,r,j} \partial \theta_{i',r',j'}} l(\vec{\phi}) = -\frac{1}{2\sigma_i^2} \sum_{u=p+1}^k w_{r,4u-j} w_{r',4u-j'}, \tag{11}$$

$$\text{and } \frac{\partial^2}{\partial \theta_{i,r,j} \partial \theta_{i',r',j'}} l(\vec{\phi}) = 0 \text{ for } i \neq i'. \tag{12}$$

$$\frac{\partial^2}{\partial \theta_{i,r,j} \partial \sigma_i^2} l(\vec{\phi}) = -\frac{1}{2(\sigma_i^2)^2} \sum_{u=p+1}^k (w_{i,u} - \lambda_i) w_{r,4u-j}. \tag{13}$$

$$\frac{\partial^2}{\partial \theta_{i,r,j} \partial \varphi_{i,c}} l(\vec{\phi}) = -\frac{1}{2\sigma_i^2} \sum_{u=p+1}^k w_{r,4u-j} (x_{c,4u} - x_{c,4u-1}). \tag{14}$$

$$\frac{\partial^2}{\partial \theta_{i,r,j} \partial \psi_q} l(\vec{\phi}) = -\frac{1}{2\sigma_i^2} \sum_{u=p+1}^k w_{r,4u-j} (\chi_{q,4u} - \chi_{q,4u-1}). \tag{15}$$

$$\frac{\partial}{\partial \varphi_{i,c}} l(\vec{\phi}) = \frac{1}{2\sigma_i^2} \sum_{u=p+1}^k (w_{i,u} - \lambda_i)(x_{c,4u} - x_{c,4u-1}). \quad (16)$$

$$\frac{\partial^2}{\partial \varphi_{i,c} \partial \varphi_{i,c'}} l(\vec{\phi}) = -\frac{1}{2\sigma_i^2} \sum_{u=p+1}^k (x_{c,4u} - x_{c,4u-1})(x_{c',4u} - x_{c',4u-1}), \quad (17)$$

$$\text{and } \frac{\partial^2}{\partial \varphi_{i,c} \partial \varphi_{i',c'}} l(\vec{\phi}) = 0 \text{ for } i \neq i'. \quad (18)$$

$$\frac{\partial^2}{\partial \varphi_{i,c} \partial \sigma_i^2} l(\vec{\phi}) = -\frac{1}{2(\sigma_i^2)^2} \sum_{u=p+1}^k (w_{i,u} - \lambda_i)(x_{c,4u} - x_{c,4u-1}). \quad (19)$$

$$\frac{\partial^2}{\partial \varphi_{i,c} \partial \psi_q} l(\vec{\phi}) = -\frac{1}{2\sigma_i^2} \sum_{u=p+1}^k (x_{c,4u} - x_{c,4u-1})(\chi_{q,4u} - \chi_{q,4u-1}). \quad (20)$$

$$\frac{\partial}{\partial \psi_q} l(\vec{\phi}) = \sum_{u=p+1}^k \sum_{i=1}^M \frac{1}{2\sigma_i^2} (w_{i,u} - \lambda_i)(\chi_{q,4u} - \chi_{q,4u-1}). \quad (21)$$

$$\frac{\partial^2}{\partial \psi_q \partial \psi_{q'}} l(\vec{\phi}) = - \left[\sum_{u=p+1}^k (\chi_{q,4u} - \chi_{q,4u-1})(\chi_{q',4u} - \chi_{q',4u-1}) \right] \sum_{i=1}^M \frac{1}{2\sigma_i^2}. \quad (22)$$

$$\frac{\partial^2}{\partial \psi_q \partial \sigma_i^2} l(\vec{\phi}) = - \sum_{u=p+1}^k \sum_{i=1}^M \frac{1}{2(\sigma_i^2)^2} (w_{i,u} - \lambda_i)(\chi_{q,4u} - \chi_{q,4u-1}). \quad (23)$$

$$\frac{\partial}{\partial \sigma_i^2} l(\vec{\phi}) = \sum_{u=p+1}^k \frac{(w_{i,u} - \lambda_i)^2}{2(\sigma_i^2)^2} - \frac{k-p}{2\sigma_i^2}. \quad (24)$$

$$\frac{\partial^2}{\partial (\sigma_i^2)^2} l(\vec{\phi}) = - \sum_{u=p+1}^k \frac{(w_{i,u} - \lambda_i)^2}{(\sigma_i^2)^3} + \frac{k-p}{2(\sigma_i^2)^2}. \quad (25)$$

3.2 Hypothesis Test and Confidence Intervals for $\theta_{i,r,j}$, $\varphi_{i,c}$, ψ_q , and σ_i^2

Once the maximum likelihood estimator $\hat{\phi}$ of $\vec{\phi}$ is available, the confidence interval of each component in it, say ϕ_l , $l = 1, \dots, L$, where L is the dimension of $\vec{\phi}$, can be calculated through the classical asymptotic theory. The method is based on inverting the Wald's test, then the approximate $100 \times (1 - 2\alpha)\%$ confidence interval for ϕ_l is given by:

$$\left[\hat{\phi}_l - I_{ll}^{-1/2}(\hat{\phi}) z_\alpha, \hat{\phi}_l + I_{ll}^{-1/2}(\hat{\phi}) z_\alpha \right], \quad (26)$$

where $I_{ll}^{-1/2}(\hat{\phi})$ is the inverse of the squared root of the l^{th} , $l = 0, \dots, L$, diagonal element of the empirical Fisher Information matrix.

Now, consider to construct a decision rule for the following pair of hypotheses:

$$H_0 : \phi_l = \phi_{l,0} \text{ versus } H_1 : \phi_l \neq \phi_{l,0}. \quad (27)$$

The duality between hypothesis tests and confidence intervals leads to reject H_0 if $\phi_{l,0}$ is not in the confidence interval estimate, and the null hypothesis is not rejected otherwise.

3.3 Estimating the States $\mu_{i,t} + S_{i,t}$ and $\beta_{i,t}$

Besides the auto-correlation, trend, and noise terms, the level component $\mu_{i,t}$ in (1) contains the influence of neighboring areas and covariates through the regression structure:

$$\rho_{i,t} = \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \theta_{i,r,j} (Y_{r,t-j} - Y_{r,t-j-1}) + \sum_{c \in C_i} \varphi_{i,c} x_{c,t} + \sum_{q=1}^Q \psi_q \chi_{q,t}. \quad (28)$$

The original three-state model is:

$$\begin{aligned} Y_{i,t} &= \mu_{i,t} + S_{i,t} + \epsilon_{i,t}, \\ \mu_{i,t} &= \mu_{i,t-1} + \beta_{i,t} + \rho_{i,t} + \eta_{i,t} \\ \beta_{i,t} &= \beta_{i,t-1} + \xi_{i,t} \\ S_{i,t} &= S_{i,t-1} + \nu_{i,t}, \end{aligned} \quad (29)$$

In order to estimate states and conduct forecast on Y , we make two changes: The first one is to define:

$$Y_{i,t}^* = Y_{i,t} - \rho_{i,t}. \quad (30)$$

The second one is to combine two states $\mu_{i,t}^* + S_{i,t}$ as a new state $L_{i,t}^*$. Here, we also denote $\mu_{i,t}^* = \mu_{i,t} - \rho_{i,t}$.

Then, $Y_{i,t}^*$ follows a new regular linear structural model with two-states as:

$$\begin{aligned} Y_{i,t}^* &= L_{i,t}^* + \epsilon_{i,t}, \\ L_{i,t}^* &= L_{i,t-1}^* + \beta_{i,t} + \rho_{i,t-1} + \omega_{i,t} \end{aligned}$$

$$\beta_{i,t} = \beta_{i,t-1} + \xi_{i,t}, \quad (31)$$

where $\omega_{i,t} = \eta_{i,t} + \nu_{i,t}$.

Since the variances of the independent white noises are unknown, this approach demands to apply the maximum likelihood method once again to estimate the variances of $\epsilon, \eta + \nu, \xi$, which is already embedded in the usual Kalman Filter algorithm. After replacing $\rho_{i,t}$ by its maximum likelihood estimator, $\hat{\rho}_{i,t}$, we apply the conventional Kalman Filter [20] for eliciting the estimates $\hat{L}_{i,t}, \hat{\beta}_{i,t}$ for the states $\mu_{i,t}^* + S_{i,t}$ and $\beta_{i,t}$ respectively. Then $\hat{Y}_{i,t}$ is finally estimated through the relationship:

$$\hat{Y}_{i,t} = \hat{L}_{i,t} + \hat{\rho}_{i,t}. \quad (32)$$

It is worth mentioning that in this Section, specifically, as per (31), we are estimating two states, $L_{i,t}$ and $\beta_{i,t}$ instead of three states $\mu_{i,t}, S_{i,t}$, and $\beta_{i,t}$. The reason for doing this is that $\mu_{i,t}$ and $S_{i,t}$ enter the observation in the same way, and they are not individually identified unless we give the seasonal part some extra structure. Since our next goal is to forecast, namely, predicting h time points of $Y_{i,t+1}, \dots, Y_{i,t+h}$, changing the estimation from three states to two would not only be simpler but statistically optimal, since this set-up will pin down the combined signal perfectly.

4 Forecasting

Consider predicting $Y_{i,t+h}$, where h is a positive integer representing the forecasting horizon of interest. For this, we assume that predictions, or scenarios of interest, for the regression covariates, say $x_{c,t+1}, \dots, x_{c,t+h}$ and $\chi_{q,t+1}, \dots, \chi_{q,t+h}$, are available.

Let $\hat{Y}_{i,t}(h)$ and $\hat{\rho}_{i,t}(h)$ denote the predictions for $Y_{i,t+h}$, and $\rho_{i,t+h}$, respectively. From (28), we have:

$$\hat{\rho}_{i,t}(h) = \sum_{r \in R_i} \sum_{j=1}^{p_{i,r}} \hat{\theta}_{i,r,j} (\hat{Y}_{r,t-j}(h) - \hat{Y}_{r,t-j-1}(h)) + \sum_{c \in C_i} \hat{\varphi}_{i,c} x_{c,t} + \sum_{q=1}^Q \hat{\psi}_q \chi_{q,t}, \quad (33)$$

where $\hat{Y}_{r,t-j}(h) = Y_{r,t-j+h}$ for $t-j+h \leq n$, and for $t-j+h > n$:

$$\hat{Y}_{i,t}(h) = \hat{L}_{i,t}(h) + \hat{\rho}_{i,t}(h) \quad (34)$$

where $\hat{L}_{i,t}(h)$ is obtained from (32). The terms are obtained through the classical forecasting theory related to the Kalman Filter procedure to solve the system in (31).

To reach up to the prediction for time $t+h$, it is necessary to iterate through one time point at a time. Following the following steps:

- Step 1: Let $\hat{\beta}_{i,T+1} = \hat{\beta}_{i,T}$. This is reasonable because $\beta_{i,t}$ is a pure random walk, as assumed in (1). The slope equation has no drift term, so the best one-step forecast for β is to use the last value.
- Step 2: Obtain the estimation for the state $L_{i,T+1}$ through $L_{i,T+1} = L_{i,T} + \hat{\beta}_{i,T+1}$.

- Step 3: estimate one future time point of Y : $\hat{Y}_{i,T+1} = \hat{L}_{i,T+1} + \hat{\rho}_{i,T+1}$.
- Step 4: repeat this process, one time point at a time, until the h future time points of Y are estimated, that is, $\hat{Y}_{i,T+1}, \dots, \hat{Y}_{i,T+h}$.

5 Simulation Studies

In order to validate the estimation methods that we have discussed in Section 3 and Forecasting in Section 4, we run some simulation studies and briefly summarize them here. In Section 3, we took a two-stage estimation strategy, the first stage is to estimate the parameters, mainly focusing on θ , ϕ , and ψ . In the second stage, we utilized the estimated parameter values and moved on to estimate the states. With both stages completed, we would then be able to forecast on Y . The simulation is exactly set up in this strategy.

5.1 Simulation Set-Up

We assume that we have four areas, which means we will have data $Y_{i,t}$ with $i = 1, 2, 3, 4$. We also set $Tmax = 1,000$, which means that the total number of time points is up to 1,000, with time index $t = 1, \dots, 1,000$. We then define the spatial adjacency structure. We let site 1 be adjacent only to site 2. Site 2 is adjacent to sites 1 and 3. Site 3 is adjacent to sites 2 and 4. And finally, site 4 is adjacent only to site 3.

The number of lagged effects from areas to their neighbors, namely $p_{i,r}$, are in the following order:

$$P = (p_{i,r})_{i,r=1}^4 = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 3 & 0 & 5 & 0 \\ 0 & 3 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

We consider two dynamic covariates and two fixed effects covariates, both are simulated from standard normal distributions. For brevity, we only consider covariates vary by time (common for all sites).

To further simulate data, we assumed the following information. Please note that some values are randomly generated: θ are generated from $N(0, 0.05^2)$ and ϕ are generated from $N(0.2, 0.1^2)$.

- (1) Spatial-lag coefficients $\theta_{i,r,j}$
- (2) Dynamic covariate loadings. $\varphi_{i,c}$ denotes the dynamic effect from the c th covariate from the area i .
- (3) Fixed-effect loadings and noise standard deviations(s.d.).

We initialized $\mu_{i,t}$, $\beta_{i,t}$, and $S_{i,t}$ for each spatial unit $i = 1, \dots, M$ and the first two time steps $t = 1, 2$ using independent Gaussian draws:

$$\mu_{i,t} \sim \mathcal{N}(0, 0.1^2), \quad \beta_{i,t} \sim \mathcal{N}(0, 0.1^2), \quad S_{i,t} \sim \mathcal{N}(0, 0.1^2), \quad \text{for } t = 1, 2.$$

The observed response variable $Y_{i,t}$ was generated as:

Table 1 True spatial-lag coefficients $\theta_{i,r,j}$, sorted by area, neighbor, lag.

Area i	Neighbor r	Lag j	Value $\theta_{i,r,j}$
1	2	1	0.00983
1	2	2	0.03251
2	1	1	0.03355
2	1	2	-0.06421
2	1	3	-0.10131
2	3	1	0.11027
2	3	2	0.01157
2	3	3	0.01878
2	3	4	-0.05965
2	3	5	0.05663
3	2	1	0.09197
3	2	2	0.07639
3	2	3	-0.10118
3	4	1	-0.05203
3	4	2	0.00489
4	3	1	0.03484
4	3	2	-0.02719

$$\varphi_{i,c} = \begin{pmatrix} 0.11415 & 0.01745 \\ 0.22298 & 0.25088 \\ 0.10574 & 0.00037 \\ 0.40435 & 0.15057 \end{pmatrix}$$

$$Y_{i,t} = \mu_{i,t} + S_{i,t} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim \mathcal{N}(0, \sigma_y^2), \quad t = 1, 2.$$

To ensure reproducibility, we used the R function `set.seed(123)` prior to all random number generation. Up until this point, we have the data for the first two time stamps. The observations from future data stamps are then computed using the model we set up as per (1) as well as the variance set-up as in Table 3.

5.2 Maximum Likelihood Estimations for Parameters

As of now, we have the full simulated data, with 1,000 time stamps, (1) the response time series, Y_i ; (2) the fixed covariates, $\chi_{1,t}$ and $\chi_{2,t}$; (3) the random covariates, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$. Based on the derivations discussed in Section 3, we conduct the maximum likelihood estimation for the parameters, $\theta_{i,r,j}$, ϕ , and ψ , and summarize the results in Table 4. One may note that this is a simulation study; the true parameter values are used to generate the time series data, but in the estimation process, we pretend they

Table 3 True fixed-effect coefficients ψ_q and noise standard deviations(s.d.).

Parameter	Meaning	Value
ψ_1	first fixed covariate	0.50
ψ_2	second fixed covariate	-0.30
σ_μ	s.d. of η	0.05
σ_β	s.d. of ξ	0.05
σ_S	s.d. of ν	0.05
σ_y	s.d. of ϵ	0.10

are unknown. As can be seen from the first three columns of Table 4, the estimated values of $\theta_{i,r,j}$, ϕ , and ψ are very close to their true values across the board.

Now, with the maximum likelihood estimates and the empirical Fisher Information matrix, we construct 95% confidence intervals for these parameters. The results are summarized in the last two columns of Table 4. The true values of each of $\theta_{i,r,j}$, ϕ , and ψ are within their confidence interval ranges.

Table 4 Maximum likelihood estimates and Wald-type 95% CIs for simulated data ($T = 1000$).

Parameter	True value	Estimate	Std. Error	t -value	p-value	Lower	Upper
ψ_1	0.50000	0.49739	0.00341	145.71000	< 0.001	0.49070	0.50408
ψ_2	-0.30000	-0.29823	0.00316	-94.39000	< 0.001	-0.30442	-0.29203
$\varphi_{1,1}$	0.11415	0.12630	0.00612	20.63000	< 0.001	0.11431	0.13830
$\varphi_{2,1}$	0.22298	0.22069	0.00608	36.30000	< 0.001	0.20877	0.23260
$\varphi_{3,1}$	0.10574	0.11488	0.00655	17.54000	< 0.001	0.10205	0.12772
$\varphi_{4,1}$	0.40435	0.40069	0.00603	66.41000	< 0.001	0.38886	0.41251
$\varphi_{1,2}$	0.01745	0.03120	0.00587	5.31000	< 0.001	0.01969	0.04270
$\varphi_{2,2}$	0.25088	0.24746	0.00574	43.10000	< 0.001	0.23621	0.25871
$\varphi_{3,2}$	0.00037	-0.00197	0.00601	-0.33000	0.74300	-0.01375	0.00981
$\varphi_{4,2}$	0.15057	0.13354	0.00571	23.37000	< 0.001	0.12234	0.14473
$\theta_{1,2,1}$	0.00983	0.02082	0.01057	1.97000	0.04890	0.00011	0.04153
$\theta_{3,2,1}$	0.09197	0.07954	0.02319	3.43000	< 0.001	0.03409	0.12498
$\theta_{1,2,2}$	0.03251	0.04123	0.00987	4.18000	< 0.001	0.02190	0.06057
$\theta_{3,2,2}$	0.07639	0.06746	0.02160	3.12000	0.00180	0.02513	0.10980
$\theta_{2,1,1}$	0.03355	0.01362	0.03202	0.43000	0.67060	-0.04913	0.07637
$\theta_{2,1,2}$	-0.06421	-0.03946	0.03985	-0.99000	0.32220	-0.11756	0.03865
$\theta_{2,1,3}$	0.11027	-0.09757	0.03136	-3.11000	0.00190	-0.15904	-0.03609
$\theta_{2,3,1}$	0.01157	0.11855	0.03196	3.71000	< 0.001	0.05590	0.18120
$\theta_{4,3,1}$	0.03484	0.04010	0.01212	3.31000	< 0.001	0.01634	0.06386
$\theta_{2,3,2}$	0.01157	-0.00444	0.03958	-0.11000	0.91070	-0.08201	0.07313
$\theta_{4,3,2}$	-0.02719	-0.00340	0.01160	-0.29000	0.76960	-0.02614	0.01934
$\theta_{2,3,3}$	0.01878	0.02187	0.03128	0.70000	0.48440	-0.03944	0.08318
$\theta_{2,3,4}$	-0.05965	-0.05471	0.01569	-3.49000	< 0.001	-0.08546	-0.02397
$\theta_{2,3,5}$	0.05663	0.08573	0.01277	6.71000	< 0.001	0.06070	0.11076
$\theta_{3,2,3}$	-0.10010	-0.10010	0.01047	-9.56000	< 0.001	-0.12062	-0.07958
$\theta_{3,4,1}$	0.05257	-0.04415	0.02254	-1.96000	0.05020	-0.08834	0.00003
$\theta_{3,4,2}$	0.00489	0.01991	0.02125	0.94000	0.34880	-0.02174	0.06157

5.3 Estimating the States

We summarize the estimation of $\mu_{i,t} + S_{i,t}$ and $\beta_{i,t}$ through simulation, following the methods introduced in Section 3.3. We used a two-stage state-space approach. In the first stage, we construct $\hat{\rho}_{i,t}$, which is made possible through the estimations of the parameters, $\theta_{i,r,j}$, ϕ , and ψ . $\hat{\rho}_{i,t}$ captures three components: a spatial distributed-lag of neighbor differences, dynamic covariate effects, and fixed covariate effects. From here, the de-regressed series $Y_{i,t}^*$ is obtained through offsetting $\hat{\rho}_{i,t}$ from $Y_{i,t}$. In the second stage, we fit a 2-state local linear trend dynamic linear model (DLM). The variance components in the model are estimated by Gaussian maximum likelihood (ML) through the Kalman filter, while the state estimates are obtained through Kalman smoothing. We then evaluate state-estimation accuracy using root mean squared error (RMSE) and the Pearson correlation between the true states and their estimates.

One may note that in our simulation, we figured that the scales of $L_{i,t}$ and $\beta_{i,t}$ are very different, which can destabilize maximum-likelihood estimation. To address this problem, we re-parameterized the DLM with scaling the states: for $L_{i,t}$, it is scaled using $s_L = sd(Y_i^*)$; while for $\beta_{i,t}$, it is scaled using $s_\beta = sd(\Delta Y_i^*)$. Afterwards, we mapped the smoothed states back to the original units. Table 5 summarizes the accuracy measurements for states and their estimates. From the table, point estimates are very accurate for both $L_{i,t}$ and $\beta_{i,t}$: the estimates have near-perfect correlation with their true values and small RMSEs; the trend is also well recovered. Figures 1 and 3 show that the estimated values of $L_{i,t}$ and $\beta_{i,t}$ are very close to their true values, which are consistent with the summarized results in Table 5. We use black thick lines for the true values, while red thin lines for the estimates. The curves for both of the graphs largely overlap due to the high correlation and small errors. Thus, we added Figures 2 and 4 to clearly show the differences for the estimates and their true values. Again, those errors are small.

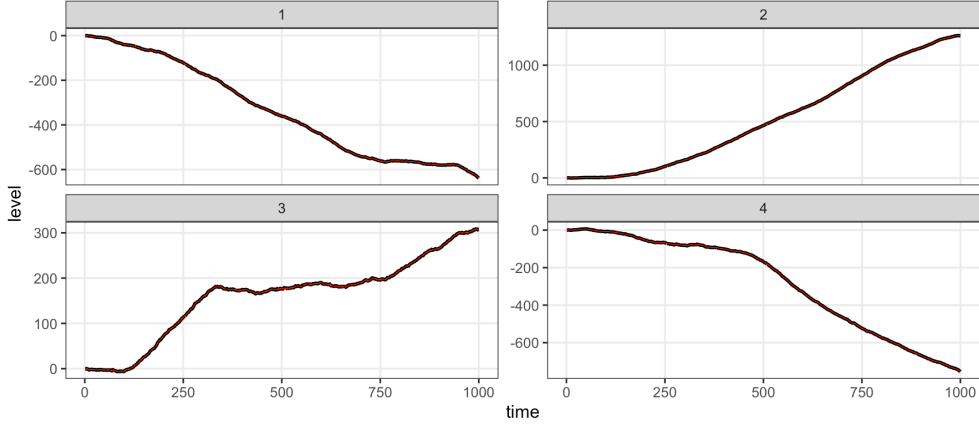


Fig. 1 Level $L_{i,t} = \mu_{i,t} + S_{i,t}$ by area: smoothed estimate (red) and truth (black).

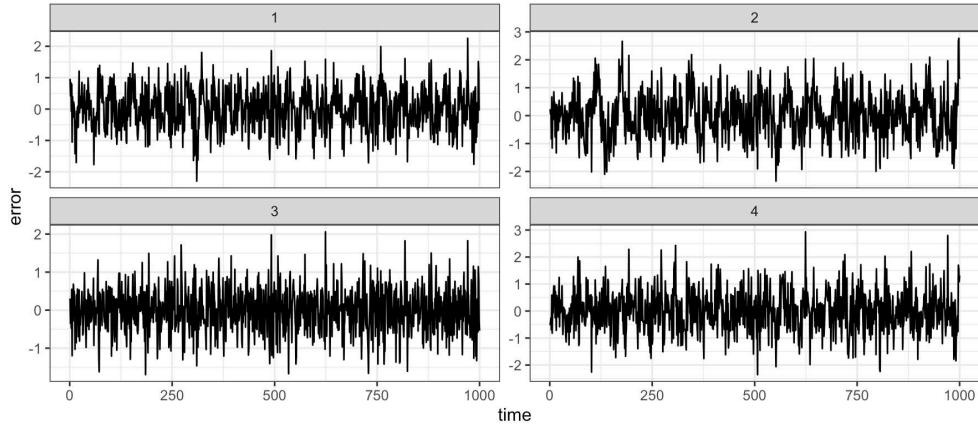


Fig. 2 The errors of estimating Level $L_{i,t}$ by area.

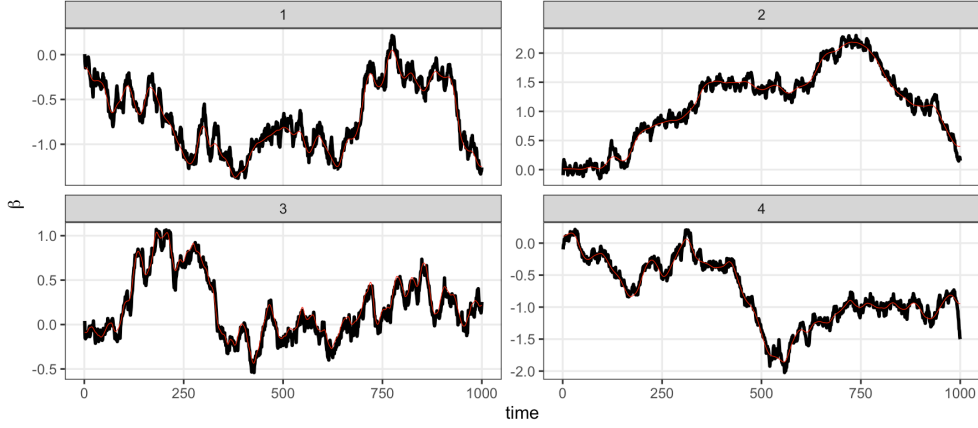


Fig. 3 Trend $\beta_{i,t}$ by area: smoothed estimate (red) and truth (black).

5.4 Forecasting

With parameters and states estimated, we are finally in a position to conduct forecasting on $Y_{i,t+h}$. Let's assume we want to forecast 10 time stamps forward; that means, continuing in our simulation, we will predict the values of $Y_{i,1001}, Y_{i,1002}, \dots, Y_{i,1010}$. We follow closely with the four steps as discussed in Section 4, and will illustrate our forecast in 5. For brevity, we simulated 10 more values for each dynamic and fixed

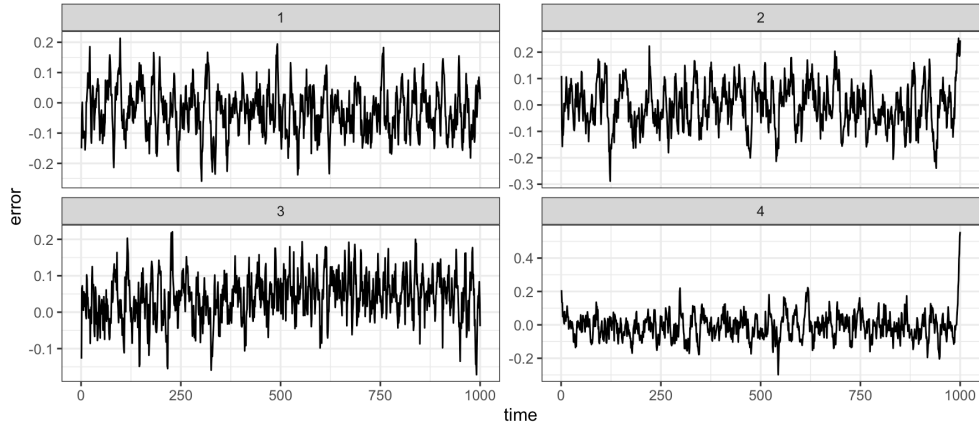


Fig. 4 The errors of estimating Level $\beta_{i,t}$ by area.

Table 5 Accuracy estimation on four representative areas. Coverage (Cov_{95}) is the empirical fraction of times the true state lies within the nominal 95% interval.

Area	RMSE_L	Corr_L	RMSE_β	Corr_β
1	0.6587477	0.9999951	0.08165167	0.9792557
2	0.8200148	0.9999981	0.08358957	0.9915324
3	0.5898973	0.9999771	0.07459951	0.9847251
4	0.7644290	0.9999954	0.07836182	0.9879145

covariates, and pretended they were observed. In this case, $\hat{\rho}_{i,T+1}$ can be obtained and thus a forecast can be made on Y .

6 Data Analysis Example

In this section, we perform some real data analyses to illustrate the methodologies discussed in the preceding sections. It is an well established fact that there may be some association between the rise in global temperature and the use of fossil fuels along with carbon emission. At the outset, we would like to make it clear that there is no intention of establishing of any causation through this analysis.

Data are available for several countries for several years. However, for brevity, we restricted our discussion to thirty years of data from 1994 to 2023 for seven countries of South America. These countries are, Argentina(1), Brazil(2), Chile(3), Colombia(4), Ecuador(5), Peru(6), and Venezuela(7). The response variable of interest was *Temperature-anomaly*. In Figures 6 and 7, we visually present the temperature anomalies for the years 1994 and 2023 for the whole world. We took *Carbon emissions* and *Fossil fuel use* as primary source of energy as two fixed effects. The source that we

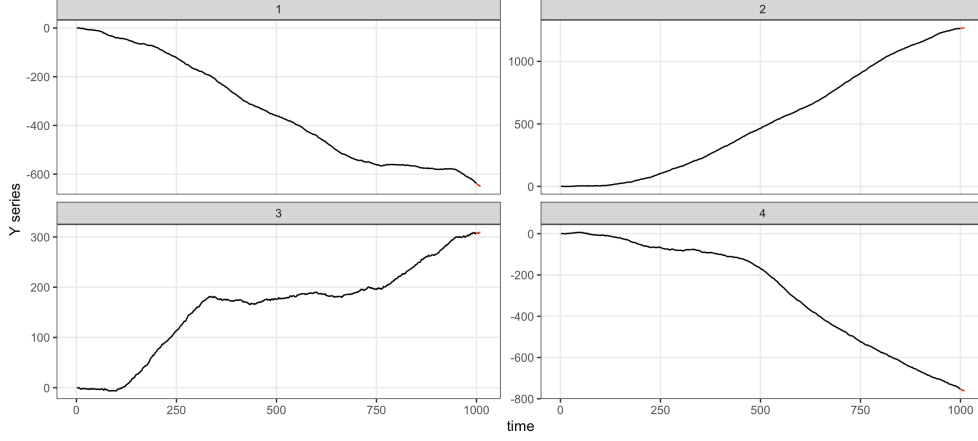


Fig. 5 The forecast of $Y_{i,t+h}$ by area, with $i = 1, 2, 3, 4$ and $h = 1, 2, \dots, 10$.

used is available publicly on the website: <https://ourworldindata.org/>. The values of the three variables of interest can be accessed from the following links:

- <https://ourworldindata.org/grapher/annual-temperature-anomalies>. Data sources: Contains modified Copernicus Climate Change Service information (2025) – with major processing by Our World In Data.
- <https://ourworldindata.org/grapher/annual-co2-emissions-per-country>. Data sources: Global Carbon Budget (2024) Population based on various sources (2024) – with major processing by Our World in Data.
- <https://ourworldindata.org/grapher/fossil-fuel-primary-energy>. Data sources: Energy Institute - Statistical Review of World Energy (2024) – with major processing by Our World in Data.

The index for each country mentioned above is given in the parenthesis. With respect to the neighbors of each country, the number of significant lags for each neighbor is taken to be one. Thus, following our original model (1), we have, $t = 30$, $M = 7$, $R_1 = R_3 = R_5 = R_7 = 2$, $R_2 = R_4 = R_6 = 4$, $Q = 2$, and $p_{i,r} = 1$ for all $r = 1, \dots, R_i$ and $i = 1, \dots, 6$. As an example, for Argentina, $i = 1$ and the model will be,

$$\begin{aligned}
 Y_{1,t} &= \mu_{1,t} + S_{1,t} + \epsilon_{1,t}, \\
 \mu_{1,t} &= \mu_{1,t-1} + \beta_{1,t} + \sum_{r \in R_1} \theta_{1,r,1} (Y_{r,t-1} - Y_{r,t-2}) + \sum_{q=1}^2 \psi_q \chi_{q,t} + \eta_{1,t} \\
 \beta_{1,t} &= \beta_{1,t-1} + \xi_{1,t}, \\
 S_{1,t} &= S_{1,t-1} + \nu_{1,t},
 \end{aligned} \tag{35}$$

for $t = 1, \dots, 30$ and $R_1 = \{2, 3\}$. The results of the analysis is presented in Table 6. Considering 5% to be the nominal significance level, the significant coefficients are distinguished using bold font.

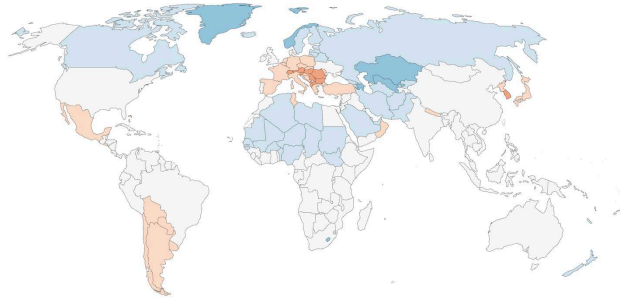


Fig. 6 Temperature anomalies of the world for 1994

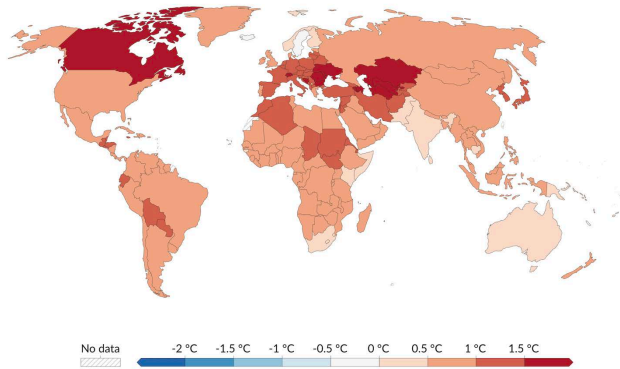


Fig. 7 Temperature anomalies of the world for 2023

The P-values for the coefficients associated with the fixed effects turns out to be high, apparently indicating no association. We may note carefully that the data analysis was performed by taking the first difference of the fixed effects. The correlation between the first difference of the fixed effects and the temperature anomaly may not be high. In fact we observed strong positive correlation between the original values of fixed covariates and temperature anomaly for all the seven countries which may not be surprising.

The model is particularly helpful in determining the spatial effect or association between neighbors. We find that lagged variation of temperature anomaly in Chile is significant for Argentina and vice versa. The lagged variation of temperature anomaly in Argentina is slightly significant for Brazil. The same from Colombia and Ecuador are significant for Peru. For brevity we have not kept any country specific random effects in our illustration.

Table 6 Data analysis for seven countries of South America

Coefficients	Estimate	Std.Error	t-value	P-value
ψ_1	-1.258×10^{-3}	1.694×10^{-3}	-0.743	0.458602
ψ_2	5.885×10^{-9}	6.494×10^{-9}	0.906	0.366120
$\theta_{1,2,1}$	2.014×10^{-1}	2.073×10^{-1}	0.972	0.332652
$\theta_{1,3,1}$	-6.194×10^{-1}	1.787×10^{-1}	-3.465	0.000673
$\theta_{2,1,1}$	3.561×10^{-1}	1.757×10^{-1}	2.027	0.044285
$\theta_{2,4,1}$	1.368	9.706×10^{-1}	1.409	0.160667
$\theta_{2,6,1}$	-1.274	7.453×10^{-1}	-1.709	0.089240
$\theta_{2,7,1}$	7.379×10^{-1}	3.612×10^{-1}	0.020	0.983724
$\theta_{3,1,1}$	-3.445×10^{-1}	1.470×10^{-1}	-2.344	0.020271
$\theta_{3,6,1}$	2.588×10^{-1}	1.935×10^{-1}	1.338	0.182824
$\theta_{4,2,1}$	-1.362×10^{-1}	3.191×10^{-1}	-0.427	0.669955
$\theta_{4,5,1}$	8.747×10^{-1}	4.552×10^{-1}	1.922	0.056352
$\theta_{4,6,1}$	-1.098	5.532×10^{-1}	-1.984	0.048840
$\theta_{4,7,1}$	1.822×10^{-1}	3.315×10^{-1}	0.550	0.583204
$\theta_{5,4,1}$	2.958×10^{-1}	7.482×10^{-1}	0.395	0.693124
$\theta_{5,6,1}$	-6.637×10^{-1}	6.307×10^{-1}	-1.052	0.294137
$\theta_{6,2,1}$	-1.130×10^{-1}	3.188×10^{-1}	-0.354	0.723470
$\theta_{6,3,1}$	-2.609×10^{-1}	1.970×10^{-1}	-1.325	0.187110
$\theta_{6,4,1}$	-1.173	5.619×10^{-1}	-2.088	0.038282
$\theta_{6,5,1}$	9.354×10^{-1}	4.578×10^{-1}	2.043	0.042586
$\theta_{7,2,1}$	-1.952×10^{-1}	3.594×10^{-1}	-0.543	0.587719
$\theta_{7,4,1}$	-1.186×10^{-1}	3.655×10^{-1}	-0.324	0.745972

7 Concluding Remarks

The proposed heteroskedastic structural model is flexible and robust. Still, there is place for potential useful extensions. For example, in Section 3.2 we used asymptotic theory for hypothesis testing and interval estimation. However, the asymptotic distributions may not hold in the presence of correlated covariates. Variable selection techniques guided by the basic structure in (1) is another meaningful topic for future investigations. A multivariate version of the basic structure is also a natural extension to be explored. Finally, as with any statistical model, noncompliance of the basic structure in (1) and/or of the conditions stated in Section 3 may lead to biased inferences. In particular, the trend and seasonal terms of some time series may not behavior as a pure random walk. New transformations for reaching the ancillary series could be tried in order to consider general AR(1) processes for $\beta_{i,t}$ and $S_{i,t}$.

Acknowledgements. The first author received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, Brazil, grant #302882/2022-7, and from Fundação de Amparo à Pesquisa do Estado de Minas Gerais - Fapemig, Brazil, grants PQ-00787-21 and RED-00133-21.

Financial Disclosure

None reported.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability

The data used in this manuscript was purely illustrative and does not serve to support the results. The data is publicly available in the links offered in the second paragraph of Section 6.

References

- [1] Banerjee, S.: Spatial data analysis. *Annual Review of Public Health* **37**(1), 47–60 (2016)
- [2] Gorsevski, P.V., Gessler, P.E., Foltz, R.B., Elliot, W.J.: Spatial prediction of landslide hazard using logistic regression and roc analysis. *Transactions in GIS* **10**(3), 395–415 (2006)
- [3] LeSage, J.P., Pace, R.K., Lam, N., Campanella, R., Liu, X.: New orleans business recovery in the aftermath of hurricane katrina. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **174**(4), 1007–1027 (2011)
- [4] Baltagi, B.H., Fingleton, B., Pirotte, A.: A time-space dynamic panel data model with spatial moving average errors. *Regional Science and Urban Economics* **72**, 45–54 (2018)
- [5] Wang, W., Lee, L.-f.: Estimation of spatial panel data models with randomly missing data in the dependent variable. *Regional Science and Urban Economics* **43**(3), 521–538 (2013)
- [6] Dou, B., Parrella, M.L., Yao, Q.: Generalized yule–walker estimation for spatio-temporal models with unknown diagonal coefficients. *Journal of Econometrics* **194**(2), 369–382 (2016)
- [7] Castillo-Mateo, J., Lafuente, M., Asín, J., Cebrián, A.C., Gelfand, A.E., Abaurrea, J.: Spatial modeling of day-within-year temperature time series: An examination of daily maximum temperatures in aragón, spain. *Journal of Agricultural, Biological and Environmental Statistics* **27**, 487–505 (2022)
- [8] Aquaro, M., Bailey, N., Pesaran, M.H.: Estimation and inference for spatial models with heterogeneous coefficients: An application to u.s. house prices. *Journal of Applied Econometrics* **36**(1), 18–44 (2021)
- [9] Li, B., Chen, J.: Gmm estimation of fixed effects partially linear additive sar model with space-time correlated disturbances. *Computational Statistics Data Analysis* **195**, 107930 (2025)

- [10] Breusch, T.S.: Maximum likelihood estimation of random effects models. *J Econometrics* **36**(3), 383–389 (1987)
- [11] Besag, J., Newell, J.: The detection of clusters in rare diseases. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **154**(1), 143–155 (1991)
- [12] Jones, K., Bullen, N.: Contextual models of urban house prices: a comparison of fixed and random-coefficients models developed by expansion. *Economic Geography* **70**(3), 252–272 (1994)
- [13] H, G.: *Multilevel statistical models*. Edward Arnoldg, London (2001)
- [14] Banerjee S, G.A. Carlin BP: *Hierarchical modeling and analysis for spatial data*. CRC Press (Taylor and Francis Group), Boca Raton [FL], London and New York (2004)
- [15] JP, E.: Serial and spatial autocorrelation. *Econ Letters* **100**(3), 422–424 (2008)
- [16] Baltagi, B.H., Li, D.: *Advances in spatial econometrics: methodology, tools, and applications*, pp. 283–295. Springer, Berlin, Heidelberg and New York (2004)
- [17] Baltagi, B.H.: Random effects and spatial autocorrelation with equal weights. *Econ Theory* **22**(5), 973–984 (2006)
- [18] Baltagi, B.H., Song, S.H., Jung, B.C., Koh, W.: Testing for serial correlation, spatial autocorrelation, and random effects using panel data. *Journal of Econometrics* **140**(1), 5–51 (2007)
- [19] Burden, R.L., Faires, J.D.: *Numerical Analysis*. PWS Publishers, ??? (1985)
- [20] Hamilton, J.D.: *Time Series Analysis*. Princeton University Press, ??? (1994)