A Constrained Spatial Autoregressive Model for Interval-valued data

Tingting Huang

School of Statistics, Capital University of Economics and Business, Beijing, 100070, China.

Contributing authors: tingth@cueb.edu.cn;

Abstract

Interval-valued data receives much attention due to its wide applications in the fields of finance, econometrics, meteorology and medicine. However, most regression models developed for interval-valued data assume observations are mutually independent, not adapted to the scenario that individuals are spatially correlated. We propose a new linear model to accommodate to areal-type spatial dependency existed in interval-valued data. Specifically, spatial correlation among centers of responses are considered. To improve the new model's prediction accuracy, we add three inequality constrains. Parameters are obtained by an algorithm combining grid search technique and the constrained least squares method. Numerical experiments are designed to examine prediction performances of the proposed model. We also employ a weather dataset to demonstrate usefulness of our model.

Keywords: interval-valued data, SAR model, areal data, linear regression, prediction

1 Introduction

Rapid development of information technology makes collect and store data at a low cost, resulting in huge, heterogeneous, imprecise, and structured data sets to be analyzed. Particularly, we may deal with a type of data, whose units of
information are intervals consisted of lower and upper bounds. The new-type data arise in various contexts.

For example, if our question of interest involves community features, then there is a need to aggregate classical numerical data table of individuals into a smaller data sheet of groups \cite{Billard2003, Wei2017}. The requirement of data confidentiality or lower computation complexity can produce such intervals as well. We summarize these sources of intervals as information gathering. What’s more, some real data endowed with variability are naturally represented by an interval, for instance, daily temperature with the maximum and the minimum values, stock prices at the opening and the closing \cite{Lim2016, Sun2022}. Imprecise observations and uncertainty are the third way causing intervals, which takes advantages of intervals to represent error measurements or error estimations of possible values.

We call this complex data interval-valued data. Modeling real data by intervals, we can better take fluctuations, oscillations of variables into consideration, and understand things from a more comprehensive view \cite{Wang2012}. Since Diday \cite{Diday1988} proposed the concept of symbolic data analysis, interval-valued data, as the most common type of symbolic data, has been extensively studied.

Regression analysis as one of the most popular approaches in statistics, exploring how a dependent variable changes as one or more covariates are changed, has been extended to investigate interval-valued data as well. Currently, discussing the relation between interval-valued responses and interval-valued predictors receives the most attention. As the data units are comprised of two bounds, not just single points, applying classical linear models to interval-valued data should be careful. Two paths of modelling interval-valued data in regressions have been developed \cite{Sun2016}. One presumes interval-valued data sit in a linear space, endowed with certain addition and scalar multiplication operations. Thus the aim is find a function of the interval-valued explanatory variables so that it is ”closest” to the outcomes. The other treats intervals as a two-dimensional vector, then establishing two separate regressions for lower and upper bounds, or centers and ranges. Although the second method may lose some geometric features of interval-valued data, it brings simpler mathematical operations and accordingly more flexibility.

We mainly focus on literature into the second method. Billard and Diday \cite{Billard2002} constructs two linear models for the lower and upper bounds of the response interval (the MinMax method). Lima Neto and de Carvalho \cite{Lima2008}
on the other way, suggests to fit mid-point and semi-length of the interval-valued dependent variable instead of the bounds (the center and range method, CRM). Lima Neto and de Carvalho (2010) then further incorporates a nonnegative constraint to the coefficient of range model so as to ensure the predicted lower boundaries smaller than the upper ones (the constrained center and range method, CCRM). To improve model’s prediction accuracy, Hao and Guo (2017) adds two other inequality restrictions to the constrained center and range model so that predicted intervals and true values have overlapping areas. Besides, variants of interval-valued linear model have also been extensively studied. Fagundes et al. (2014) shows an interval kernel regression. Lima Neto and de Carvalho (2017) introduces nonlinear interval-valued regression. Lim (2016) proposes nonparametric additive interval-valued regression. Lima Neto and de Carvalho (2018) presents a robust model for interval-valued data by penalizing outliers through exponential kernel functions. Yang et al. (2019) introduces artificial neural network techniques to interval-valued data prediction. More recently, de Carvalho et al. (2021) combines k-means algorithm with linear and nonlinear regressions and raises clusterwise nonlinear regression for interval-valued data. Xu and Qin (2022) extends the center and range method to the Bayesian framework. Notably, all these investigations are conducted with the assumption that observations are mutually independent.

However, we may encounter real data that are spatially correlated in practice. Like in weather data, seasonal precipitations of a city are often high if it’s neighbor cities have abundant rainfall. And in unemployment data, unemployment rates of different districts in the same city are spatially agglomerate. Housing price data and stock price exhibit similar patterns (Topa (2001); Anselin (1998); Lesage and Pace (2009)). An vital and common characteristic of these correlated data is that spatial dependence between individuals $i$ and $i'$ is determined by distance $d_{ii'}$ between them. The closer two spatial units are (the smaller $d_{ii'}$), the greater the spatial dependency. Causes of such correlation are mainly from human spatial behavior, spatial interaction and spatial heterogeneity. In applications, $d_{ii'}$ can be a more general distance, such as social distance, policy distance and economic distance (Isard and et al (1970)).

Areal data (lattice data) is one kind of this spatially correlated data. It arises when a fixed domain is carved up by a finite number of subareas. Resulting in each observation of areal data is aggregated outcomes in the same subregion. Lattice data receive much attention in the fields of econometrics and spatial statistics. We focus on modelling interval-valued data with areal-type spatial dependence in this research.
To better illustrate our motivation, we present a practical example. To investigate how temperature affects precipitation in major cities of China, monthly weather data from China Meteorological Yearbook of 2019 were collected. As temperature data are naturally interval-valued data, interval-valued linear regressions are employed to analyze the problem. In data preprocessing, the single-valued monthly rainfall and monthly temperature for each city were aggregate into intervals. Then to see whether the interval-valued outcomes are spatially dependent, the Moran’I statistic was utilized. We tested spatial correlation among centers and radii of the interval-valued precipitations, separately. The values of the statistics are 0.675 for centers with p-value equaling 0.00001, and −0.081 for ranges with p-value equaling 0.6149. Obviously, midpoints of the interval-valued response are correlated, and semi-lengths are not. If apply exiting interval-valued linear regressions to model the problem, the results obtained would be misleading as these models mostly assume observations are mutually independent. Therefore, there are reality needs to model interval-valued data with spatial correlation. Figure 1 (left and middle) shows Moran’I scatter plots of the centers and the ranges. Figure 1 (right) displays Moran’s I scatter plot of the interval-valued response, where the spatially lagged intervals on the vertical axis are aggregated from single-valued lagged precipitations. The spatial weight matrix $W$ used here will be illustrated at length in Section 5.

![Fig. 1](image) The Moran’I Scatter plot of the centers (left) and the ranges (middle), the predicted precipitation versus the true values (right).

When only the ordinary single-valued data are involved, spatial linear models, including spatial autoregressive model (SAR), spatial error model (SEM) and spatial Durbin model (SDM) are most commonly employed to analyze areal data. Through the utility of spatial weight matrix $W$ and spatial lag parameter $\rho$, spatial models incorporates spatial effects into classic linear regressions. Estimation methods for these regressions cover quasi-maximum likelihood estimation method (Lee (2004)), generalized moment estimators (Lee (2007)) and the Markov Chain Monte Carlo (MCMC) method (Lesage
and Pace (2009)). Their applications range from stock markets to social networks (Ma et al. (2020); Zhang et al. (2019)). As the SAR model is the most popular spatial model, we capitalize on it to struct our new model.

In this article, we propose a constrained spatial autoregressive model for interval-valued data (ICSM), which aims to predict interval-valued outcomes in the presence of areal-type spatial dependency. In particular, the CRM (center and range method) is utilized to handle inter-valued data. And we considered spatial correlation among centers of responses, employing the SAR model to form regression. We also take prediction accuracy into account by adding three inequality constrains in the model, motivated by Hao and Guo (2017). The parameters are obtained by the combination a grid search technique and the constrained least squares method. We also conduct numerical experiments to examine prediction performances of the proposed model. We find when spatial correlation is great and signal to noise is high, our model behaves better than models without spatial term and constrains. The weather data mentioned earlier is thoroughly analyzed by the new model.

We note that literatures related to spatial linear models of interval-valued data are relatively scanty. Han et al. (2016) discusses vector autoregressive moving average model for interval-valued data. Sun et al. (2018) proposes a new class of threshold autoregressive interval models. Although these investigations have considered correlation between intervals, they are mainly designed for interval-valued time series data, not interval-valued data with lattice-type spatial dependency.

The paper is organized as follows. Section 2 introduces the spatial autoregressive model and our proposed ICSM. Section 3 describes the algorithm to obtain the parameters. Numerical experiments are conducted in Section 4 to examine performances of the new model. Real data analysis are arranged in Section 5. We end up the article with a conclusion.

2 Model Specification

2.1 Spatial autoregressive (SAR) model

Ord (1975) proposed the spatial autoregressive (SAR) model, modeling spatial effects in the areal data. Suppose there are \( n \) spatial units on a lattice, and each individual occupies a position. We observe \( \{(x_i, y_i)\}_{i=1}^n \) from these entities. Denote \( \mathbf{x} = (x_1, x_2, \ldots, x_n)^\top \) and \( \mathbf{y} = (y_1, y_2, \ldots, y_n)^\top \), the SAR model is formulated as

\[
\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{x} \beta + \mathbf{\epsilon}, \quad \mathbf{\epsilon}_i \sim N(0, \sigma^2)
\]
where \( W = (w_{ii'})_{n \times n} \) is a spatial weight matrix, with each entry \( w_{ii'} \) representing the weight between unit \( i \) and unit \( i' \), \( \rho \in (-1, 1) \) is the spatial lag parameter to be estimated, \( \epsilon = (\epsilon_1, \cdots, \epsilon_n) \) is the noise term independent of \( x \) and i.i.d. following a normal distribution. \( \beta \) is the scalar coefficient. Here, \( W \) quantifies spatial structure of the lattice, and the spatial lag term \( \rho W y \) models spatial effects. The greater value \( \rho \) takes, the stronger spatial impacts are imposed on \( y \).

It should be emphasized that generally \( W \) is exogenous and pre-specified, where \( w_{ii'} \) is determined by the contiguity or distance between \( i \) and \( i' \) in different spatial scenarios. Rook matrix is among the most common spatial weight matrices, where

\[
w_{ii'} = \begin{cases} 1, & i \text{ and } i' \text{ share an edge} \\ 0, & \text{otherwise}. \end{cases}
\]

Bishop matrix is built in similar way by setting \( w_{ii'} = 1 \) when \( i \) and \( i' \) have the same vertex. The key point to construct such matrix is knowing the map indicating spatial arrangement of the individuals. In case that geographical locations of spatial units are given, \( w_{ii'} \) can be functions of distance \( d_{ii'} \) between the two units, for example,

\[
w_{ii'} = \begin{cases} \frac{1}{d_{ii'}}, & d_{ii'} \leq d_0 \\ 0, & d_{ii'} > d_0 \end{cases}
\]

where \( d_0 \) is a threshold distance. More general spatial weights can be derived from networks. For instance under a social network setting, \( w_{ii'} = 1 \) if person \( i \) and \( i' \) are friends. For more information of \( W \), refer to [Isard and et al. (1970)] and [Anselin (1998)]. In general, \( W \) is row-normalized so that summation of the row elements is unity. Entries on the diagonal of \( W \) are zeros.

### 2.2 Constrained spatial autoregressive model for interval-valued data (ICSM)

Following [Jenish and Prucha (2009)](#), we presume the spatial process is located on an unevenly spaced lattice \( L \in R^n \ (n \geq 1) \), and all units on \( L \) are endowed with positions (\( L \) is crucial for establishing the weight matrix \( W \)).

Denote an interval-valued variable as \( w^s = [w^l, w^u] \), where \( w^s \in S = \{ [a, b] \parallel a, b \in \mathbb{R}, a \leq b \} \). Here, we observe \( \{ y^s_i, x^s_i \}_{i=1}^n \) from lattice \( L \), and
\( \mathbf{y}_i^s = [y_i^l, y_i^u], \mathbf{x}_i^s = [x_i^l, x_i^u]. \) To deal with the interval-valued data, we employ the CRM method (Lima Neto and de Carvalho (2008)). Specifically, represent midpoint and radius of \( \mathbf{w}^s \) by \( \mathbf{w}_c \) and \( \mathbf{w}_r \), and write \( \mathbf{y}_c = (y_1^c, \ldots, y_n^c)^\top, \mathbf{y}_r = (y_1^r, \ldots, y_n^r)^\top, \mathbf{x} = (x^c, x^r) = ((x_1^c, \ldots, x_n^c)^\top, (x_1^r, \ldots, x_n^r)^\top), \) then the constrained spatial autoregressive interval-valued model (ICSM) is formulated as

\[
\begin{align*}
\mathbf{y}_c &= \rho \mathbf{W} \mathbf{y}_c + \mathbf{x} \beta_c + \mathbf{e}_c \\
\mathbf{y}_r &= \mathbf{x} \beta_r + \mathbf{e}_r \\
\mathbf{y}_c - \mathbf{y}_r &\leq (\mathbf{I}_n - \rho \mathbf{W})^{-1} \cdot \mathbf{x} \beta_c + \mathbf{x} \beta_r \\
\mathbf{y}_c + \mathbf{y}_r &\geq (\mathbf{I}_n - \rho \mathbf{W})^{-1} \cdot \mathbf{x} \beta_c - \mathbf{x} \beta_r \\
\mathbf{x} \beta_r &\geq 0
\end{align*}
\]  

where \( \mathbf{W} \) is the spatial weight matrix, \( \rho \in (-1, 1) \) is the spatial lag parameter, \( \beta_c = (\beta_1^c, \beta_2^c) \) and \( \beta_r = (\beta_1^r, \beta_2^r) \) are the coefficients, and \( \mathbf{e}_c = (e_1^c, \ldots, e_n^c)^\top, \mathbf{e}_r = (e_1^r, \ldots, e_n^r)^\top \) are the error terms with \( E(\mathbf{e}_c) = \mathbf{0}, Var(\mathbf{e}_c) = \sigma^2 \mathbf{I}_n, E(\mathbf{e}_r) = \mathbf{0}, Var(\mathbf{e}_r) = \sigma^2 \mathbf{I}_n. \)

Note that compared to precisely forecast semi-length \( \mathbf{y}_r \), accurately predicting midpoint \( \mathbf{y}_c \) would contribute more to obtain a good predictor of \( \mathbf{y}^s \). Therefore, considering simplicity and efficiency of the new model, we give priority to modeling spatial correlation of \( \mathbf{y}_c \) by the SAR regression, fitting \( \mathbf{y}_r \) by ordinary linear regression. We also considered effects from both centers and ranges of \( \mathbf{x}^s \).

Inspired by Hao and Guo (2017), to improve prediction accuracy of the model, we add three inequality constraints to the interval-valued spatial regression, where the first two ensure predicted intervals have overlapping areas with the true values, and the last guarantees predicted value of \( \mathbf{y}_r \) is positive. Concretely, in the first inequality, the forecasted upper bound of \( \mathbf{y}_s^c \) should be greater than the observed lower bound of \( \mathbf{y}_s^c \); and in the second, the forecasted lower bound of \( \mathbf{y}_s^c \) should be smaller than the true upper bound of \( \mathbf{y}_s^c \).

In model (5), \( \rho \) quantifies average spatial effects among \( \mathbf{y}_s^c \)’s central positions. In particular, when \( \rho = 0 \) the ICSM degenerates into the ICM (Hao and Guo (2017)). Thus (5) is a more general model.

### 3 Parameter estimation

Estimating \( \rho, \beta_c \) and \( \beta_r \) of (5) is not straightforward due to the presence of spatial lag term \( \rho \mathbf{W} \mathbf{y}_c \) and the constrains. Regarding \( \rho \), commonly used methods to obtain it’s estimator include maximum likelihood estimation method
(Lee (2004)), generalized moment estimators (Lee (2007)) and Markov Chain Monte Carlo method (Lesage and Pace (2009)). However, these approaches cannot adapt to the situation that model parameters are restricted by inequalities. We put forward an algorithm that combines a grid search method with the constrained least squares to obtain estimators.

Firstly, as $\rho$ is a scalar number taking values from $-1$ to $1$, we can give a sufficiently fine grid of $m$ values for $\rho$. For example, take step size by 0.01, then $m = 201$ and $(\rho_1, \rho_2, \rho_3, \cdots, \rho_{109}, \rho_{200}, \rho_{201}) = (-1, -0.99, -0.98, \cdots, 0.98, 0.99, 1)$.

Secondly, for each value of $\rho$, we solve a constrained least squares optimization problem. Provided value of $\rho_j$ ($j = 1, \cdots, m$), we can calculate $A_j = (I_n - \rho_j W)$.

Then model (5) turns into

\[
\begin{cases}
    y^c = A_j^{-1} \cdot (x\beta^c + \epsilon^c) \\
    y^r = x\beta^r + \epsilon^r
\end{cases}
\]

s.t.

\[
\begin{cases}
    y^c - y^r \leq A_j^{-1} \cdot x\beta^c + x\beta^r \\
    y^c + y^r \geq A_j^{-1} \cdot x\beta^c - x\beta^r \\
    x\beta^r \geq 0
\end{cases}
\]

Now that $A_j$ is known, we can employ least squares method to obtain $\hat{\beta}^c$ and $\hat{\beta}^r$. The sum of the squares of deviations is given by

\[
\min_{\beta^c, \beta^r} (\epsilon_j^c)^\top \epsilon_j^c + (\epsilon_j^r)^\top \epsilon_j^r 
\]

s.t.

\[
\begin{cases}
    -A_j^{-1} \cdot x\beta^c - x\beta^r \leq -y^c + y^r \\
    A_j^{-1} \cdot x\beta^c - x\beta^r \leq y^c + y^r \\
    -x\beta^r \leq 0
\end{cases}
\]

where $\epsilon_j^c = A_j \cdot y^c - x\beta^c$ and $\epsilon_j^r = y^r - x\beta^r$.

Once the solution for (2) is found for all candidates $\rho_j$, we select the $\rho_j$ with the least sum of square residuals as the $\hat{\rho}$. Accordingly, if the $j$th $\rho$ is the optimum $\rho$, then the $j$th estimated $\hat{\beta}^c$ and $\hat{\beta}^r$ are estimators of $\beta^c$ and $\beta^r$.

We summarize the whole process in Algorithm 1.
Algorithm 1 Main steps of the estimation procedure

Require: the dataset $y^c, y^r, x^c, x^r$ and the spatial matrix $W$.

1: Generate a sequence of $\rho$ from $-1$ to 1 with step size 0.01, that is $(\rho_1, \cdots, \rho_{200}, \rho_{201}) = (-1, \cdots, 0.99, 1)$.

2: for $k = 1$ to 201 do

3: compute $A = I_n - \rho_k W$

4: optimize

$$\min \|Y - Z\beta\|^2 \text{ s.t. } G\beta \leq h$$

where $G = \begin{pmatrix} -A^{-1}X & -X \\ A^{-1}X & -X \\ 0_{n \times 3} & -X \end{pmatrix}$, $h = \begin{pmatrix} y^r - y^c \\ y^c + y^r \\ 0_{n \times 1} \end{pmatrix}$, $Z = \begin{pmatrix} X \\ 0_{n \times 3} \\ X \end{pmatrix}$, $Y = \begin{pmatrix} Ay^c \\ y^r \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, $X = \begin{pmatrix} I_n \ x^c \ x^r \end{pmatrix}$, $\beta_1 = \begin{pmatrix} \beta^c_0 \\ \beta^c \end{pmatrix}$, $\beta_2 = \begin{pmatrix} \beta^r_0 \\ \beta^r \end{pmatrix}$.

5: end for

6: choose $\rho_j$ with the least $\|Y - Z\beta\|^2$, corresponding $\beta$ is the estimated coefficient.

Ensure: $\rho_j$ and the $j$th $\beta$

4 simulation study

In this section, numerical experiments are conducted to show the superiority of our new method. Firstly, we introduce the data generating process. Secondly, we briefly explain how unobserved individuals are predicted under a spatial scenario. Two competing regression models are also considered. Lastly, we present several criterions to measure model performances and display the experimental results.

4.1 data generating process

For the spatial scenarios, we take two spatial weight matrices $W$ into account, one is the commonly used rook matrix, and the other is the block matrix. These two matrices imply two cases of spatial structures, where the first corresponds to sparse networks, and the second corresponds to dense ones. More specifically, the spatial units under a rook scenario have a small number of neighbors, while the individuals under a block setting have a relatively greater number of neighbors. Here are details of generating the two matrices.

- Rook matrix assigns $w_{ii'} = 1$ if units $i$ and $i'$ share an edge, and $w_{ii'} = 0$ otherwise. We assume $n$ agents are randomly located on a regular square lattice with $R$ rows and $T$ columns, where each agent occupies a cell. Then
we have sample size \( n = R \times T \). In this context, the agents have 4 neighbors at most, and 2 neighbors at least, given that they may be in the inner field or in the corner of the square grid. We set \( n = \{10 \times 12, 12 \times 20, 20 \times 25\} = \{120, 240, 500\} \).

- Block matrix was first discussed by [Case (1991)](Case1991), where there are \( D \) number of districts, \( M \) members in each district. Specifically, each unit \( i \) in the block matrix has \( m - 1 \) neighbors with equal weights, i.e., we set \( w_{ii'} = \frac{1}{m-1} \) if \( i \) and \( i' \) are in the same district, \( w_{ii'} = 0 \) otherwise. We take \( n = D \times M = \{20 \times 6, 20 \times 12, 25 \times 20\} = \{120, 240, 500\} \).

The numerical data are generated using the following model

\[
\begin{align*}
y^c &= (I_n - \rho W)^{-1} \cdot (x^c \beta_1^c + x^r \beta_2^c + \epsilon^c) \\
y^r &= x^c \beta_1^r + x^r \beta_2^r + \epsilon^r
\end{align*}
\]

1) Regarding \( \rho \), we consider three cases.

i. \( \rho = 0 \), where spatial dependency does not present. The ICSM degenerates into the CCRJM of [Hao and Guo (2017)](HaoGuo2017).

ii. \( \rho = 0.4 \), where network effects are moderate.

iii. \( \rho = 0.8 \), where network effects are strong.

2) Centers and ranges of the interval covariates are generated by uniform distribution, i.e.

\[
x^c \sim U(0, 150), \quad x^r \sim U(5, 8)
\]

3) Coefficients of the covariates \( \beta_1^c, \beta_2^c \) are generated by uniform distribution, as well

\[
\beta_1^c \sim U(-2.5, -2), \quad \beta_2^c \sim U(2.5, 5), \quad \beta_1^r = 1, \quad \beta_2^r = 0.1
\]

4) To see how the prediction performance is influenced by different levels of noises of the center, we let

i. \( \epsilon_i^c \sim U(0, 11), \quad \epsilon_r \sim U(0, 5) \).

ii. \( \epsilon_i^c \sim U(0, 18), \quad \epsilon_r \sim U(0, 5) \).

4.2 Predicting formula and comparative models

Note that in a spatial context, predicting \( Y \) of an out-of-sample should be careful, as adding new observations will change the existing spatial structure, i.e., \( W \) is different. [Goulard et al. (2017)](Goulard2017) studied the problem. They discussed
out-of-sample prediction for the SAR model and concluded the “BP” predictor behaves well among other predictors. Denote in-sample observations by $(X_s, Y_s)$ and out-of-sample covariate by $X_o$, which are all known. Then the unknown out-of-sample response $Y_o$ can be calculated by “BP” predictor

$$
\hat{Y}_o^{BP} = \hat{Y}_o^{TC} - Q_o^{-1}Q_{os}(Y_s - \hat{Y}_s^{TC})
$$

$$
\hat{Y}_o^{TC} = (I_n - \hat{\rho}W)^{-1}X\hat{\beta} = \left(\hat{Y}_s^{TC} - \hat{Y}_o^{TC}\right)
$$

$$
Q = \frac{1}{\hat{\sigma}^2}(I_n - \hat{\rho}(W^T + W) + \hat{\rho}^2W^TW) = \begin{pmatrix} Q_s & Q_{so} \\ Q_{os} & Q_o \end{pmatrix}
$$

where $\hat{\sigma}^2$ is mean square error of the fitting model. In our simulated data set, we randomly select $\frac{9}{10}n$ individuals as the learning set, and leave the remaining $\frac{1}{10}n$ units as the test set.

We compare the ICSM’s predictive power with other two models: the constrained interval-valued model without spatial dependency (ICM), and the spatial interval-valued model without constrains (ISM), which can be formulated as

- The ICM (equivalent to the CCRJM by Hao and Guo (2017))

$$
\begin{cases}
\begin{align*}
y^c &= x^c\beta_1^c + x^r\beta_2^c + \epsilon^c \\
y^r &= x^c\beta_1^r + x^r\beta_2^r + \epsilon^r
\end{align*}
\end{cases}
\begin{align*}
-(x^c\beta_1^c + x^r\beta_2^c) - (x^c\beta_1^r + x^r\beta_2^r) &\leq -y^c + y^r \\
(x^c\beta_1^r + x^r\beta_2^r) - (x^c\beta_1^c + x^r\beta_2^c) &\leq y^c + y^r \\
-(x^c\beta_1^c + x^r\beta_2^r) &\leq 0.
\end{align*}
\end{cases}
$$

- The ISM

$$
\begin{cases}
\begin{align*}
y^c &= \rho Wy^c + x^c\beta_1^c + x^r\beta_2^c + \epsilon^c \\
y^r &= x^c\beta_1^r + x^r\beta_2^r + \epsilon^r
\end{align*}
\end{cases}
$$

### 4.3 measurements and comparison results

A number of criterion classes have been used to evaluate the performance of the interval linear regressions. Lim (2016) calculates root mean square error (RMSE) of the lower and upper bound of the predictive interval, i.e., $\text{RMSE}_l$ and $\text{RMSE}_u$. Hojati et al. (2005) employs the accuracy rate (AR) to access similarity of two intervals based on the percentage of overlapping areas. We
use these three criteria in our simulation, which are defined as

\[
RMSE_l = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i^l - \hat{y}_i^l)^2}
\]

\[
RMSE_u = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i^u - \hat{y}_i^u)^2},
\]

\[
AR = \frac{1}{n} \sum_{i=1}^{n} \frac{\mu(y_i \cap \hat{y}_i)}{\mu(y_i \cup \hat{y}_i)},
\]

where \(\mu(y_i \cap \hat{y}_i)\) and \(\mu(y_i \cup \hat{y}_i)\) are volumes of intersection and union of \(y_i\) and \(\hat{y}_i\), respectively.

Note that \(AR\) is an average measurement of the overlapping areas. It may happen that the value of \(AR\) is great, but some \(\hat{y}_i\) are mispredicted (the corresponding \(\mu(y_i \cap \hat{y}_i)\) equals 0). To know how many individuals are wrongly forecasted in a data set, we add a new indicator \(N_d\), which is the number of disjoint intervals, i.e.,

\[
N_d = \# \{ i \mid \mu(y_i \cap \hat{y}_i) = 0, \ i = 1, \ldots, n \}.
\]

The whole process was repeated 75 times in the experiment. We computed the average \(MSE_l, MSE_u, AR, N_d\) and their standard deviations, summarized in Table 1 (\(\rho = 0\)), Table 2 (\(\rho = 0.4\)), and Table 3 (\(\rho = 0.8\)).

5 Application

In this section, we give a detailed analysis of the weather data mentioned in Section 4 using the proposed new model. In particular, to examine the prediction power of the ICSM, we collected monthly temperature and monthly rainfall data in the year of 2020 from China Meteorological Yearbook. That is, data in 2019 acts as training set, and those in 2020 acts as test set. In the preprocessing of the data, we first take logarithm of the precipitation, then group all the single values into intervals. The ICSM is built as

\[
\begin{align*}
\left\{ 
\begin{array}{ll}
y_i^c &= \rho \sum_{i' \neq i} w_{ii'} y_{i'} + x_i^{c} \beta_1^c + x_i^{r} \beta_2^c + \epsilon_i^c \\
y_i^u &= x_i^{c} \beta_1^u + x_i^{r} \beta_2^u + \epsilon_i^u 
\end{array}
\right.
\end{align*}
\]
Table 1 Prediction performances of ICSM, ICM and ISM when $\rho = 0$.  

| $n$ | $e_i^*$ | $W$ | $MSE_l$ | $MSE_u$ | $MSE_l$ | $MSE_u$ |
|-----|---------|-----|---------|---------|---------|---------|
|     |         | ICSM | ICM     | ISM     | ICSM    | ICM     |
| 120 | N(0,11) | rook | 11.1014 | 11.0722 | 11.1052 | 11.2156 | 11.1783 | 11.2237 |
|     |         | block| 11.2169 | 11.1144 | 11.2134 | 11.2167 | 11.0920 | 11.2161 |
|     | N(0,18) | rook | 18.4828 | 18.3768 | 18.4603 | 18.8141 | 18.7145 | 18.5812 |
|     |         | block| 18.4331 | 18.4424 | 18.4669 | 18.3173 | 18.3298 | 18.3754 |
|     | 240 N(0,11) | rook | 11.0143 | 11.0286 | 11.0579 | 11.1054 | 11.1150 | 11.1523 |
|     |         | block| 11.2274 | 11.2125 | 11.2541 | 11.1916 | 11.1743 | 11.2184 |
|     | N(0,18) | rook | 18.1153 | 18.1168 | 17.9986 | 17.9500 | 17.9420 | 17.9134 |
|     |         | block| 17.9657 | 17.9348 | 17.7945 | 17.9281 | 17.9095 | 17.8560 |
|     | 500 N(0,11) | rook | 10.9784 | 10.9739 | 10.9848 | 11.0429 | 11.0348 | 11.0498 |
|     |         | block| 10.9696 | 10.9653 | 10.9791 | 11.0228 | 11.0189 | 11.0335 |
|     | N(0,18) | rook | 18.3842 | 18.3787 | 18.1039 | 18.1269 | 18.0955 | 18.0033 |
|     |         | block| 18.5208 | 18.5262 | 18.2646 | 18.4309 | 18.4353 | 18.3191 |

| $n$ | $e_i^*$ | $W$ | $AR$ | $Nd$ | $AR$ | $Nd$ |
|-----|---------|-----|------|------|------|------|
|     |         | ICSM | ICM | ISM | ICSM | ICM | ISM |
| 120 | N(0,11) | rook | 0.7749 | 0.7748 | 0.7747 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.7725 | 0.7743 | 0.7727 | 0.0000 | 0.0000 | 0.0000 |
|     | N(0,18) | rook | 0.6453 | 0.6467 | 0.6446 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.6709 | 0.6704 | 0.6706 | 0.0000 | 0.0000 | 0.0000 |
|     | 240 N(0,11) | rook | 0.7771 | 0.7768 | 0.7765 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.7939 | 0.7695 | 0.7687 | 0.0000 | 0.0000 | 0.0000 |
|     | N(0,18) | rook | 0.6672 | 0.6671 | 0.6666 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.6696 | 0.6703 | 0.6703 | 0.0000 | 0.0000 | 0.0000 |
|     | 500 N(0,11) | rook | 0.7755 | 0.7756 | 0.7755 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.7792 | 0.7793 | 0.7791 | 0.0000 | 0.0000 | 0.0000 |
|     | N(0,18) | rook | 0.6711 | 0.6712 | 0.6719 | 0.0000 | 0.0000 | 0.0000 |
|     | block | 0.6668 | 0.6667 | 0.6666 | 0.0000 | 0.0000 | 0.0000 |
Table 2  Prediction performances of ICSM, ICM and ISM when $\rho = 0.4$. 

| $n$  | $\varepsilon_i^c$ | $W$     | $MSE_i$     | $MSE_u$     | $MSE_i$ | $MSE_u$ |
|-----|----------------|---------|--------------|--------------|---------|---------|
|     |                |         | ICSM | ICM  | ISM  | ICSM | ICM  | ISM  |
| 120 | N(0,11)        | rook   | 11.4445 | 26.8926 | 11.4144 | 11.4818 | 27.2150 | 11.4899 |
|     |                | block  | 10.8221 | 29.6401 | 10.8411 | 10.8115 | 29.3916 | 10.8306 |
| 240 | N(0,11)        | rook   | 18.1000 | 32.3855 | 18.1169 | 18.3430 | 31.0777 | 18.1409 |
|     |                | block  | 17.4380 | 34.4353 | 17.4085 | 17.4482 | 33.6160 | 17.3586 |
| 500 | N(0,11)        | rook   | 11.2280 | 26.5330 | 11.2390 | 11.2742 | 26.6516 | 11.2438 |
|     |                | block  | 11.0554 | 22.5821 | 11.0636 | 10.9353 | 22.0217 | 10.9421 |
|     | N(0,18)        | rook   | 17.9222 | 31.9797 | 17.6607 | 18.0094 | 31.8204 | 17.6132 |
|     |                | block  | 17.6641 | 27.5415 | 17.5935 | 17.6786 | 26.2599 | 17.5969 |
| 240 | N(0,11)        | rook   | 11.4305 | 28.5840 | 11.2410 | 11.5519 | 28.0278 | 11.2111 |
|     |                | block  | 11.1313 | 18.3621 | 11.1339 | 11.0830 | 18.3696 | 11.0864 |
|     | N(0,18)        | rook   | 18.7708 | 33.8118 | 18.0857 | 18.3586 | 34.6654 | 18.0755 |
|     |                | block  | 18.2823 | 24.8806 | 18.0818 | 18.1978 | 24.6113 | 18.0764 |

| $n$  | $\varepsilon_i^c$ | $W$     | $AR$  | $Nd$  | $AR$  | $Nd$  |
|-----|----------------|---------|-------|-------|-------|-------|
|     |                |         | ICSM  | ICM  | ISM  | ICSM  | ICM  | ISM  |
| 120 | N(0,11)        | rook   | 0.7697 | 0.5660 | 0.7695 | 0.0000 | 0.2133 | 0.0000 |
|     |                | block  | 0.7795 | 0.5363 | 0.7791 | 0.0000 | 0.1333 | 0.0000 |
| 240 | N(0,11)        | rook   | 0.6658 | 0.5214 | 0.6640 | 0.0533 | 0.2133 | 0.9067 |
|     |                | block  | 0.6761 | 0.5075 | 0.6769 | 0.0000 | 0.2533 | 0.0000 |
| 500 | N(0,11)        | rook   | 0.7738 | 0.5728 | 0.7737 | 0.0000 | 0.1200 | 0.0000 |
|     |                | block  | 0.7739 | 0.6112 | 0.7738 | 0.0000 | 0.1067 | 0.0000 |
|     | N(0,18)        | rook   | 0.6755 | 0.5365 | 0.6780 | 0.0533 | 0.3600 | 0.9733 |
|     |                | block  | 0.6766 | 0.5674 | 0.6761 | 0.0133 | 0.2667 | 0.8400 |
| 500 | N(0,11)        | rook   | 0.7699 | 0.5633 | 0.7736 | 0.0000 | 0.3200 | 0.0000 |
|     |                | block  | 0.7732 | 0.6658 | 0.7732 | 0.0000 | 0.0667 | 0.0000 |
|     | N(0,18)        | rook   | 0.6700 | 0.5252 | 0.6707 | 0.0933 | 0.4133 | 1.0000 |
|     |                | block  | 0.6670 | 0.5922 | 0.6673 | 0.0800 | 0.2000 | 1.2267 |
| n  | $c_i^2$ | $W$     | $MSE_i$ | $MSE_u$ |
|----|---------|---------|---------|---------|
|    |         | ICSM    | ICM     | ISM     | ICSM    | ICM     | ISM     |
| 120| N(0,11) | rook    | 20.9170 | 136.9900 | 11.8538 | 22.1827 | 143.8223 | 11.9363 |
|    |         | block   | 10.0668 | 268.4357 | 10.8871 | 10.1485 | 276.4214 | 10.1281 |
|    | N(0,18) | rook    | 26.8112 | 138.2213 | 18.2134 | 25.2355 | 135.1789 | 18.2767 |
|    |         | block   | 17.7788 | 276.9172 | 17.5684 | 18.1281 | 286.9323 | 17.6841 |
| 240| N(0,11) | rook    | 22.5784 | 145.8142 | 11.5306 | 22.3133 | 155.3351 | 11.4477 |
|    |         | block   | 10.9630 | 193.0556 | 10.9685 | 10.7834 | 190.3062 | 10.7810 |
|    | N(0,18) | rook    | 29.6531 | 167.7000 | 17.4381 | 32.8813 | 165.4525 | 17.4703 |
|    |         | block   | 18.1834 | 195.6656 | 17.8889 | 18.1563 | 202.0845 | 17.8429 |
| 500| N(0,11) | rook    | 27.1713 | 173.1929 | 11.1199 | 25.5140 | 167.6693 | 11.0789 |
|    |         | block   | 11.0527 | 147.2848 | 11.0548 | 11.1206 | 142.7372 | 11.1228 |
|    | N(0,18) | rook    | 31.5665 | 180.5881 | 17.4332 | 32.3119 | 179.2781 | 17.4337 |
|    |         | block   | 18.2320 | 163.7546 | 17.7877 | 17.9188 | 161.2984 | 17.8166 |

| n  | $c_i^2$ | $W$     | $AR$     | $N_d$     |
|----|---------|---------|---------|---------|
|    |         | ICSM    | ICM     | ISM     | ICSM    | ICM     | ISM     |
| 120| N(0,11) | rook    | 0.6513  | 0.2094  | 0.7540  | 0.0000  | 0.5867  | 0.0000  |
|    |         | block   | 0.7910  | 0.1222  | 0.7911  | 0.0000  | 0.6800  | 0.0000  |
|    | N(0,18) | rook    | 0.6168  | 0.2234  | 0.6769  | 0.0267  | 0.7333  | 0.2800  |
|    |         | block   | 0.6611  | 0.1168  | 0.6590  | 0.0133  | 0.5200  | 0.4800  |
| 240| N(0,11) | rook    | 0.6641  | 0.2143  | 0.7792  | 0.0000  | 0.5467  | 0.0000  |
|    |         | block   | 0.7811  | 0.1660  | 0.7811  | 0.0000  | 0.5733  | 0.0000  |
|    | N(0,18) | rook    | 0.5594  | 0.1860  | 0.6680  | 0.0000  | 0.4933  | 0.7333  |
|    |         | block   | 0.6730  | 0.1602  | 0.6748  | 0.0267  | 0.6800  | 0.5200  |
| 500| N(0,11) | rook    | 0.6192  | 0.1867  | 0.7738  | 0.0000  | 0.4667  | 0.0000  |
|    |         | block   | 0.7718  | 0.2038  | 0.7717  | 0.0000  | 0.5733  | 0.0000  |
|    | N(0,18) | rook    | 0.5617  | 0.1746  | 0.6768  | 0.0133  | 0.5067  | 0.9867  |
|    |         | block   | 0.6665  | 0.1881  | 0.6670  | 0.1200  | 0.6267  | 1.1867  |
the spatial effects. Thus we set \( w_d k \) number of neighbors \( \rho \) latitudes. Figure determine a proper \( W \) that \( \rho, \beta_1, \beta_2, \beta_1^r, \beta_2^r \) are parameters to be estimated.

In model 5 we utilize the inverse distance to obtain \( w_{ii'} \), i.e., \( w_{ii'} = \frac{1}{d_{ii'}} \), where \( d_{ii'} \) is the distance between cities \( i \) and \( i' \) computed by longitudes and latitudes. Figure displays locations of these cities on map of China. To determine a proper \( W \) for the ICSM, we consider two factors. The first is the number of neighbors \( k \), which determines the sparsity of \( W \). The second is the distance threshold \( d_0 \). It is known that the greater the \( d_{ii'} \), the weaker the spatial effects. Thus we set \( w_{ii'} = 0 \) if \( d_{ii'} > d_0 \), and \( \frac{1}{d_{ii'}} \) otherwise. To choose an optimal pair of \( (k, d_0) \) for \( W \), we firstly set \( k = \{1, 2, 3, \ldots, 9\} \). Then regarding each \( k \), search for a \( d_0 \), which makes the Moran’s I statistic of \( y^c \) take the greatest value. Table 4 shows results for \( k \) and \( d_0 \). It can be seen that \( k = 2 \) with \( d_0 = 492.37 \) km is the best choice.

| \( k \) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Moran’s I | 0.678 | 0.687 | 0.636 | 0.641 | 0.643 | 0.648 | 0.646 | 0.646 | 0.646 |
| \( d_0 \) (km) | 446.57 | 492.37 | 492.37 | 492.37 | 492.37 | 492.37 | 492.37 | 492.37 | 492.37 |

Table 5 compares the ICSM and the ICM in terms of both fitting and prediction performances. It is obvious that the ICSM behaves better than the ICM. We can see spatial dependencies are eliminated in the residuals of the ICSM (with the Moran’s I statistic being \(-0.24\)), while still present in those of the ICM (with the Moran’s I statistic being \(0.46\)). From Figure 2 (left and middle) we can come to similar conclusion, which displays Moran’s I scatter plots of the residuals. Besides, MSE of the residuals of the ICSM is
0.580, smaller than that of the ICM (0.674). Regarding prediction performance, the ICSM has wider overlapping areas (44.23%) than the ICM (43.96%) by average, and smaller forecast errors of the lower and the upper bounds of the precipitation intervals. Figure 2 (right) exhibits predicted intervals versus the real values.

| Models | Moran’s I statistic (residuals of $y_i^c$) | MSE (fitted error) | MSE (prediction error) | AR MSE | $MSE_l$ MSE_u |
|--------|------------------------------------------|--------------------|------------------------|--------|---------------|
| ICSM   | -0.24 (0.901)                            | 0.580              | 0.668                  | 44.23% | 0.790 0.519   |
| ICM    | 0.46 (0.002)                             | 0.674              | 0.712                  | 43.96% | 0.822 0.583   |

Table 5 summarizes estimation results of the ICSM and the ICM. The spatial lag parameter $\hat{\rho}$ is 0.56, meaning spatial effects in the precipitation are medium. $\hat{\beta}_1$ and $\hat{\beta}_2$ are equivalent to 0.261 and 0.195, respectively, thus $y_i^c$s are positively related to both centers and ranges of the temperatures. However, $\hat{\beta}_1$ equals 0.108 and $\hat{\beta}_2$ is $-0.003$, indicating $y_i^r$s are mainly influenced by mid-points of the temperatures. And semi-lengths of the temperatures play a relatively small role.

| models | $\hat{\rho}$ | $\hat{\beta}_0$ | $\hat{\beta}_1^c$ | $\hat{\beta}_2^c$ | $\hat{\beta}_0^r$ | $\hat{\beta}_1^r$ | $\hat{\beta}_2^r$ |
|--------|--------------|-----------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| ICSM   | 0.56         | -0.167          | 0.261             | 0.195             | 0.710             | 0.108             | -0.003            |
| ICM    | —            | -0.250          | 0.510             | 0.284             | 0.710             | 0.108             | -0.003            |

Fig. 2 The Moran’I Scatter plot of the residuals of $y_i^c$s in the ICM (left) and the ICSM (middle), the predicted 2020 precipitation intervals versus the true values using the ICSM (right).
6 Conclusion

Interval-valued data receives much attention due to its wide applications in the area of meteorology, medicine, finance, etc. Regression models concentrated on this kind of data have also been extended to nonparametric paradigms and the Bayesian framework. Nevertheless, relative few literature considered spatial dependence among intervals. In the article, we put forward a constrained spatial autoregressive model (ICSM) to model interval-valued data with lattice-type spatial correlation. In particular, for the simplicity and efficiency of the new model, we only take spatial effects among centers of responses into account. And to improve prediction accuracy of the model, three inequality constrains for the parameters are added. We obtain the parameters by a grid search technique combined with the constrained least-squares method. Simulation studies show that our new model behaves better than the model without spatial lag term (ICM) and the model without constrains (ISM), when spatial dependencies are present and disturbances are great. And in cases that there are no spatial correlation or disturbances are small, our model performs similarly with the other two models. We also employ a real dataset of temperatures and precipitation to illustrate the utility of the ICSM.

Although we only considered spatial effects among mid-points of outcomes, spatial correlation among the semi-lengths can also be taken into consideration by adding a spatial lag term in the linear regression. Under this new model with two different spatial lag parameters, estimator can be gotten by the addition of a “for” loop, which searches for an adequate $\rho$ for $y^*$. Note that the additional iteration would greatly increase the computation time. In future, we will discuss more flexible models, such as the incorporation of two different spatial weight matrices, and more general assumptions of the noise terms.

References

Anselin, L. 1998. Spatial Econometrics: Methods and Models. Springer Science & Business Media.

Billard, L. and E. Diday. 2002. Symbolic Regression Analysis, In Classification, Clustering, and Data Analysis, eds. Bock, H.H., W. Gaul, M. Schader, K. Jajuga, A. Sokolowski, and H.H. Bock, 281–288. Berlin, Heidelberg: Springer Berlin Heidelberg. Series Title: Studies in Classification, Data Analysis, and Knowledge Organization. https://doi.org/10.1007/978-3-642-56181-8_31
Billard, L. and E. Diday. 2003, June. From the Statistics of Data to the Statistics of Knowledge: Symbolic Data Analysis. *Journal of the American Statistical Association* 98(462): 470–487. [https://doi.org/10.1198/016214503000242](https://doi.org/10.1198/016214503000242).

Case, A.C. 1991. Spatial Patterns in Household Demand. *Econometrica* 59(4): 953–965. [https://doi.org/10.2307/2938168](https://doi.org/10.2307/2938168).

de Carvalho, F.d.A., E.d.A. Lima Neto, and K.C. da Silva. 2021, May. A clusterwise nonlinear regression algorithm for interval-valued data. *Information Sciences* 555: 357–385. [https://doi.org/10.1016/j.ins.2020.10.054](https://doi.org/10.1016/j.ins.2020.10.054).

Diday, E. 1988. The symbolic approach in clustering and related methods of data analysis: The basic choices. In *Proceedings of IFCS’87 on Classification and Related Methods of Data Analysis*, pp. 673–684.

Fagundes, R.A., R.M. de Souza, and F.J.A. Cysneiros. 2014, March. Interval kernel regression. *Neurocomputing* 128: 371–388. [https://doi.org/10.1016/j.neucom.2013.08.029](https://doi.org/10.1016/j.neucom.2013.08.029).

Goulard, M., T. Laurent, and C. Thomas-Agnan. 2017. About predictions in spatial autoregressive models: optimal and almost optimal strategies. *Spatial Economic Analysis* 12(2-3): 304–325. [https://doi.org/10.1080/17421772.2017.1300679](https://doi.org/10.1080/17421772.2017.1300679) [https://arxiv.org/abs/https://doi.org/10.1080/17421772.2017.1300679](https://arxiv.org/abs/https://doi.org/10.1080/17421772.2017.1300679).

Han, A., Y. Hong, S. Wang, and X. Yun. 2016, June. A Vector Autoregressive Moving Average Model for Interval-Valued Time Series Data, In *Advances in Econometrics*, eds. González-Rivera, G., R.C. Hill, and T.H. Lee, Volume 36, 417–460. Emerald Group Publishing Limited. [https://doi.org/10.1108/S0731-905320160000036021](https://doi.org/10.1108/S0731-905320160000036021).

Hao, P. and J. Guo. 2017, December. Constrained center and range joint model for interval-valued symbolic data regression. *Computational Statistics & Data Analysis* 116: 106–138. [https://doi.org/10.1016/j.csda.2017.06.005](https://doi.org/10.1016/j.csda.2017.06.005).

Hojati, M., C. Bector, and K. Smimou. 2005. A simple method for computation of fuzzy linear regression. *European Journal of Operational Research* 166(1): 172–184. [https://doi.org/10.1016/j.ejor.2004.01.039](https://doi.org/10.1016/j.ejor.2004.01.039).
Isard, W. and et al. 1970. *General theory: social, political, economic and regional*. Cambridge, Mass./London: Mass. Inst. Technol.

Jenish, N. and I.R. Prucha. 2009. Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics* 150(1): 86 – 98. [https://doi.org/10.1016/j.jeconom.2009.02.009](https://doi.org/10.1016/j.jeconom.2009.02.009).

Lee, L.f. 2004. Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica* 72(6): 1899–1925. [https://doi.org/10.1111/j.1468-0262.2004.00558.x](https://doi.org/10.1111/j.1468-0262.2004.00558.x).

Lee, L.f. 2007. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics* 137(2): 489–514. [https://doi.org/10.1016/j.jeconom.2005.10.004](https://doi.org/10.1016/j.jeconom.2005.10.004).

Lesage, J. and R.K. Pace. 2009. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.

Lim, C. 2016, September. Interval-valued data regression using nonparametric additive models. *Journal of the Korean Statistical Society* 45(3): 358–370. [https://doi.org/10.1016/j.jkss.2015.12.003](https://doi.org/10.1016/j.jkss.2015.12.003).

Lima Neto, E.d.A. and F.d.A. de Carvalho. 2008, January. Centre and Range method for fitting a linear regression model to symbolic interval data. *Computational Statistics & Data Analysis* 52(3): 1500–1515. [https://doi.org/10.1016/j.csda.2007.04.014](https://doi.org/10.1016/j.csda.2007.04.014).

Lima Neto, E.d.A. and F.d.A. de Carvalho. 2010, February. Constrained linear regression models for symbolic interval-valued variables. *Computational Statistics & Data Analysis* 54(2): 333–347. [https://doi.org/10.1016/j.csda.2009.08.010](https://doi.org/10.1016/j.csda.2009.08.010).

Lima Neto, E.d.A. and F.d.A. de Carvalho. 2018, July. An exponential-type kernel robust regression model for interval-valued variables. *Information Sciences* 454-455: 419–442. [https://doi.org/10.1016/j.ins.2018.05.008](https://doi.org/10.1016/j.ins.2018.05.008).

Lima Neto, E.d.A. and F.d.A.T. de Carvalho. 2017, August. Nonlinear regression applied to interval-valued data. *Pattern Analysis and Applications* 20(3): 809–824. [https://doi.org/10.1007/s10044-016-0538-y](https://doi.org/10.1007/s10044-016-0538-y).
Ma, Y., R. Pan, T. Zou, and H. Wang. 2020. A Naive Least Squares Method for Spatial Autoregression with Covariates. *Statistica Sinica*. https://doi.org/10.5705/ss.202017.0135.

Ord, K. 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70(349): 120–126. https://doi.org/10.1080/01621459.1975.10480272.

Sun, L., L. Zhu, W. Li, C. Zhang, and T. Balezentis. 2022, August. Interval-valued functional clustering based on the Wasserstein distance with application to stock data. *Information Sciences* 606: 910–926. https://doi.org/10.1016/j.ins.2022.05.112.

Sun, Y. 2016, January. Linear regression with interval-valued data: Linear regression with interval-valued data. *Wiley Interdisciplinary Reviews: Computational Statistics* 8(1): 54–60. https://doi.org/10.1002/wics.1373.

Sun, Y., A. Han, Y. Hong, and S. Wang. 2018, October. Threshold autoregressive models for interval-valued time series data. *Journal of Econometrics* 206(2): 414–446. https://doi.org/10.1016/j.jeconom.2018.06.009.

Topa, G. 2001. Social Interactions, Local Spillovers and Unemployment. *The Review of Economic Studies* 68(2): 261–295. https://doi.org/10.1111/1467-937X.00169.

Wang, H., R. Guan, and J. Wu. 2012, June. CIPCA: Complete-Information-based Principal Component Analysis for interval-valued data. *Neurocomputing* 86: 158–169. https://doi.org/10.1016/j.neucom.2012.01.018.

Wei, Y., S. Wang, and H. Wang. 2017, August. Interval-valued data regression using partial linear model. *Journal of Statistical Computation and Simulation*: 1–20. https://doi.org/10.1080/00949655.2017.1360298.

Xu, M. and Z. Qin. 2022, January. A bivariate Bayesian method for interval-valued regression models. *Knowledge-Based Systems* 235: 107396. https://doi.org/10.1016/j.knosys.2021.107396.

Yang, Z., D.K. Lin, and A. Zhang. 2019, February. Interval-valued data prediction via regularized artificial neural network. *Neurocomputing* 331: 336–345. https://doi.org/10.1016/j.neucom.2018.11.063.
Zhang, W., X. Zhuang, and Y. Li. 2019, December. Spatial spillover around G20 stock markets and impact on the return: a spatial econometrics approach. *Applied Economics Letters* 26(21): 1811–1817. https://doi.org/10.1080/13504851.2019.1602703.