Received: 5 October 2022 | Accepted: 15 March 2023

DOI: 10.1111/biom.13861

BIOMETRIC METHODOLOGY

Biometrics A JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY WILEY

# Detecting the spatial clustering of exposure–response relationships with estimation error: a novel spatial scan statistic

Wei Wang[1] | Sheng Li[1] | Tao Zhang[1] | Fei Yin[1] | Yue Ma[1,2]

[1]West China School of Public Health and West China Fourth hospital, Sichuan University, Chengdu, China

[2]Institute of Systems Epidemiology, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

**Correspondence**
Yue Ma, West China School of Public Health and West China Fourth hospital, Sichuan University, Chengdu, China.
Email: gordonrozen@scu.edu.cn

**Funding information**
National Natural Science Foundation of China, Grant/Award Number: 81803332

## ABSTRACT

Detecting the spatial clustering of the exposure–response relationship (ERR) between environmental risk factors and health-related outcomes plays important roles in disease control and prevention, such as identifying highly sensitive regions, exploring the causes of heterogeneous ERRs, and designing region-specific health intervention measures. However, few studies have focused on this issue. A possible reason is that the commonly used cluster-detecting tool, spatial scan statistics, cannot be used for multivariate spatial datasets with estimation error, such as the ERR, which is often defined by a vector with its covariance estimated by a regression model. Such spatial datasets have been produced in abundance in the last decade, which suggests the importance of developing a novel cluster-detecting tool applicable for multivariate datasets with estimation error. In this work, by extending the classic scan statistic, we developed a novel spatial scan statistic called the estimation-error-based scan statistic (EESS), which is applicable for both univariate and multivariate datasets with estimation error. Then, a two-stage analytic process was proposed to detect the spatial clustering of ERRs in practical studies. A published motivating example and a simulation study were used to validate the performance of EESS. The results show that the clusters detected by EESS can efficiently reflect the clustering heterogeneity and yield more accurate ERR estimates by adjusting for such heterogeneity.

KEYWORDS
cluster analysis, DLNM, exposure–response association, spatial heterogeneity

## 1 | INTRODUCTION

Studying the relationships between environmental risk factors and health-related outcomes, hereafter referred to as exposure–response relationships (ERRs), is a key component of environmental epidemiology. With improvements in the environmental monitoring system and disease surveillance system, many high spatiotemporal resolution environment and disease surveillance datasets have been obtained, from which region-specific time-series data with different administrative division levels have become available. Based on such data, time-series regressions have been one of the main tools used to characterize ERRs (Shah et al. 2013, Requia et al. 2018). As ERRs often exhibit regional heterogeneity, multi-region studies are commonly used to characterize such heterogeneous ERRs (Gasparrini et al. 2012, Meng et al. 2021, Samet et al. 2000). In these studies, a two-stage strategy is generally employed, particularly for complex nonlinear ERRs. Specifically, in the first stage, common region-stratified ERR estimation models are built to obtain the region-specific estimates of ERRs. When the ERR exhibits nonlinearity and lag effects, which are often found in ERRs of environmental factors, a generalized additive model (GAM), for example, distributed lag nonlinear model (DLNM) (Gasparrini 2014, Gasparrini et al. 2010), is often built for each region. In these models, the ERR is defined by multiple parameters, that is, a vector, via a group of splines. In the second stage, multivariate meta-regression, including the region-level predictors, is used to study the intensity and causes of heterogeneity, as well as to obtain a more stable estimate of average ERR across all regions by adjusting for the heterogeneity.

It has been acknowledged that various factors, such as meteorological factors, economic development levels, education levels, human behaviors, and sanitary conditions, may affect the ERRs between environmental factors and health outcomes (Hu et al. 2019, Shah et al. 2015). These factors are often spatially unevenly distributed and spatially aggregated, therefore potentially making the heterogeneity of ERRs spatially clustered. Studying the spatial clustering of ERRs will greatly help identifying highly sensitive regions, exploring the causes of clustering (or heterogeneity), and developing region-specific health intervention measures, which play an important role in disease control and prevention. For example, if we detect a spatial cluster in which the ERR between high temperature and mortality is significantly stronger than that in other regions, a temperature-related intervention measure should be taken in that region to decrease the mortality. We can also explore the reason for the clustering by looking further into differences in factors between clusters and non-cluster areas. However, the commonly used two-stage strategy does not consider the spatial distribution of regional heterogeneity in ERRs and therefore ignores such spatial clustering of ERRs. There are no other methods that focus on the spatial clustering of ERRs.

Kulldorff's scan statistic (Kulldorff et al. 1995) is a frequently used spatial clustering detection method, which can efficiently identify the location of potential clusters and simultaneously make a statistical inference to test for significance. It has been extended to work with various types of datasets (Abolhassani et al. 2021), such as Poisson-distributed (Kulldorff 1997), binomial-distributed (Kulldorff 1997), zero-inflated (Cançado et al. 2014, 2017), survival (Huang et al. 2007), normal-distributed (Kulldorff et al. 2009), multivariate Gaussian-distributed (Cucala et al. 2017), and graph-based (Cadena et al. 2017) data. However, these scan statistics were developed for observed raw data without estimation errors. For example, in detecting the spatial clustering of birth weights, all the individual's birth weight values must be observed rather than estimated. Although the distribution-free spatial scan statistic by Cucala (2014) and the weighted normal spatial scan statistic by Huang et al. (2009) can deal with such estimated spatial datasets by assuming unequal variances, they are only applicable for univariate data. For the ERRs of our interest, due to the epidemiological mechanism, they are usually nonlinear and are defined by vectors. Then, ERRs are often estimated multivariate values with nonignorable estimated covariances from GAMs. Therefore, the existed spatial scan statistics cannot be used due to the presence of estimation errors and multivariate. With a large number of spatial datasets for ERRs with estimation error having been and being generated, developing a method for detecting spatial clustering in ERRs from the standpoint of theory and application is necessary.

In this work, motivated by detecting the spatial clustering of ERRs, we develop a novel spatial scan statistic called the estimation-error-based scan statistic (EESS), which extends the classic scan statistic to apply to multivariate data with estimation error. When data are univariate, EESS is also applicable. Furthermore, based on EESS, a two-stage analytic strategy is proposed to detect the spatial clustering of ERRs. In Section 2, we introduce a published study as an example to illustrate EESS, mainly focusing on obtaining the region-specific ERRs in the first stage. In Section 3, based on the illustrative example, we present the detailed methodology of EESS. In Section 4, we employ EESS to detect the spatial clustering of ERRs in the second stage, providing an example. In Section 5, a simulation study is used to further evaluate the performance of the EESS. In Section 6, a general discussion is given.

## 2 | MOTIVATING EXAMPLE

In this section, we describe an illustrative example to provide a context for employing the proposed EESS to detect the spatial clustering of ERRs. This example derives from Xiao et al.'s work (2017) regarding the relationship between temperature and hand, foot, and mouth disease (HFMD) from 143 prefecture-level cities in China.

### 2.1 | Exposure–response relationship between temperature and hand, foot, and mouth disease

HFMD, caused by enterovirus, has become a predominant acute childhood infectious disease in the Asia-Pacific region during the last two decades (Zhuang et al. 2015). Especially in mainland China, HFMD has caused a heavy disease burden, with the highest disability-adjusted life-years in children and more than one million cases reported annually (Koh et al. 2018). It is well known that temperature is one of the most important environmental factors related to the disease, affecting the transmission of HFMD by impacting virus reproduction, survival, and children's behaviors (Belanger et al. 2009, Bertrand et al. 2012). Various studies (Cheng et al. 2018) have shown that the relationship between temperature and HFMD may differ across regions due to the heterogeneity of effect modifiers, such as the natural environment and economic development levels. Furthermore, the natural environment and economic development levels usually exhibit spatial clustering, which may lead to the spatial clustering of ERRs between temperature and HFMD. Conditioning on the clustering amounts to conditioning on the unobserved effect modifiers. Therefore, studying the spatial clustering of such ERRs will benefit to characterize a more accurate and comprehensive ERR, and also assist to identify temperature-sensitive regions and to explore the causes of clustering heterogeneity of ERRs, which play important roles in HFMD control and prevention.

## 2.2 | Data

In Xiao et al.'s well-presented work, the daily clinical cases of HFMD aged 0–12 years for each of 143 prefecture-level cities of mainland China between 2009 and 2014 were recorded. Data for a total of 3,060,450 cases were collected. The locations of the 143 cities were distributed as shown in Figure 1A. The daily relative mean temperature was used as the studied environmental factor. In addition, the daily relative humidity, air pressure, rainfall, and sunshine hours were also collected as potential confounders in the ERR between temperature and HFMD.

## 2.3 | Modeling city-specific exposure–response relationships

In the first stage, a batch of time-series regressions with the same model structure in terms of the variable of interest is independently constructed to estimate the city-specific ERRs. The general algebraic definition is given by the following:

$$Y_{it} \sim f(\mu_{it}) \text{ and } g(\mu_{it}) = \alpha + s(x_{it}; \theta_i) + \sum_{j=1}^{J} h_{ji}(z_{tji}, \gamma_{ji}), \tag{1}$$

where $f(\cdot)$ is a probability distribution with respect to the observed value $Y_{it}$ which indicates the number of disease cases in city $i$ at time $t$. The term $g(\cdot)$ is a monotonic link function of the expected case $\mu_{it} \equiv E(Y_{it})$, $\alpha$ is the intercept and $x_{it}$ is usually a vector in terms of the objective exposures within a specific range of lag, for example, $x_{it} = (x_{i,t}, x_{i,t-1}, x_{i,t-2}, x_{i,t-3})^T$ when the lag range is 0–3. The term $s(\cdot)$ is a function characterizing the objective ERR, which is often a spline function, cross-basis function, or linear function. The vector $\boldsymbol{\theta}_i$ is the ERR-related parameter to be estimated. The term $h_{ji}(\cdot)$ indicates the confounding effect, in which $z_{tji}$ is the confounding variable and $\gamma_{ji}$ is the confounding-related parameter to be estimated.

In Xiao et al.'s work, because temperature ranges vary much between cities, if all cities shared absolute temperature knots for $s(\cdot)$, some of them would get missing values of ERRs in the first stage (Gasparrini et al. 2012), which hinders the second-stage analysis. Therefore, temperature was scaled based on city-specific percentiles to unify the temperature ranges across cities. A quasi-Poisson distribution with overdispersion and a logarithmic link function were selected to construct a quasi-Poisson regression. The cross-basis function, using 5 degrees of freedom (df) natural cubic splines for the exposure–response relationship and 4-df natural cubic splines for the lag–response relationship, was used to characterize the complex nonlinear exposure–lag–response relationship between temperature and HFMD. A lag range of 4–14 days was selected. For the confounding terms, the daily relative humidity, air pressure, rainfall, and sunshine hours were adjusted using an exponentially weighted moving average. The effects of holidays, weekends, long-term trends, and seasonality were also adjusted as confounders.

By maximizing the likelihood, the city-specific $\hat{\boldsymbol{\theta}}_i$ and its covariance were obtained. Furthermore, by integrating the lag–response relationship over the lag range of $l_0 - l_1$, a five-dimensional vector $\hat{\boldsymbol{\beta}}_i$ along with its covariance $\mathbf{S}_i$ is obtained for each city to characterize the accumulated ERR between temperature and HFMD. The detailed method has been well presented in Gasparrini et al.'s works (2014). Such region-specific estimation values with estimation error, defining the ERR, have also been obtained in many other research fields in the last decade.

# 3 | METHODS: AN ESTIMATION-ERROR-BASED SCAN STATISTIC

In the second stage, based on the city-specific $\hat{\boldsymbol{\beta}}_i$ with estimation error, EESS is used to detect the spatial clustering of ERRs. The following will detail the methodology of EESS.

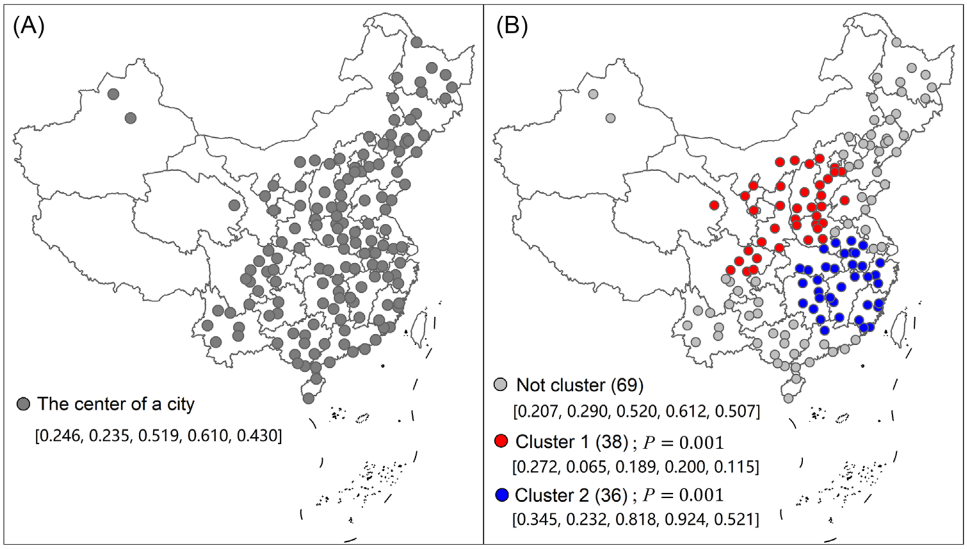

FIGURE 1 Distribution of cities and the detected clusters in the motivation example. (.) is the number of cities located in the corresponding cluster or non-cluster area. [.] is the weighted average multiple parameters characterizing the ERRs between temperature and HFMD. $P$ is calculated based on the Monte Carlo permutation with 999 replicas. When $P < 0.05$, the identified cluster is statistically significant. ERR, exposure–response relationship; HFMD, hand, foot, and mouth disease.

The illustration will use a general term to emphasize that EESS cannot only be used for detecting the spatial clustering of ERRs described in the motivating example, but can also be used for other derived (or second-hand) spatial data with estimation error.

## 3.1 | Observations and locations

The observations are a batch of scalars or vectors with estimation error, which are often estimated with statistical models or obtained from published data, such as the city-specific ERR $\hat{\boldsymbol{\beta}}_i$ with its covariance $\mathbf{S}_i$, $i = 1, 2, 3, , N$, in the motivating example. Each observation must be assigned to a location $k$, $k = 1, 2, 3, , K$, with a two-dimensional spatial coordinate. Each location permits more than one observation. As such, $N \geq K$.

## 3.2 | Scan window for clusters

Similar to the classic spatial scan statistic, through a large number of overlapping circular windows, a circular scan statistic is defined to detect clusters. Specifically, by taking each location in turn as the center with a radius varying continuously from zero to a preset upper limit, a large number of overlapping circles are defined. For each circle $z$, the logarithm of likelihood ratio LLR($z$) is calculated as in the next section, and the maximal LLR($z$) across all circles is used as the test statistic. A single cluster cannot be larger than the area outside the cluster, otherwise the outside area will be defined as a cluster; thus, the maximum size of circle $z$ is commonly set to be that containing 50% of all locations. For a specific study, a maximum size less than 50% could also be set according to the research requirements.

## 3.3 | Calculation of LLR($z$)

Under the null hypothesis that no cluster exists, that is, all the locations should have homogeneous true values, labeled $\boldsymbol{\eta}$, a probability model can be defined as

$$\hat{\boldsymbol{\beta}}_i = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_i, \ \boldsymbol{\varepsilon}_i \sim \mathrm{MN}\left(\mathbf{0}, \mathbf{S}_i\right) \tag{2}$$

where MN(.) is a multivariate normal distribution and $\mathbf{S}_i$ is the known estimated covariance matrix of $\hat{\boldsymbol{\beta}}_i$ from the GAM in the first stage. When the observation is a scalar, the univariate normal distribution is used. Therefore, the likelihood under the null hypothesis can be calculated as $L_0\left(\boldsymbol{\eta}\right) = \prod_i \left|2\pi\mathbf{S}_i\right|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta})^T \mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta})\}$ and then the log-likelihood is

$$LL_0\left(\boldsymbol{\eta}\right) = \ln\left(L_0\left(\boldsymbol{\eta}\right)\right) = -\frac{Nq}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{i=1}^{N}\ln\left(\left|\mathbf{S}_i\right|\right) - \frac{1}{2}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}\right)^T \mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}\right) \tag{3}$$

where $q$ is the dimension of $\hat{\boldsymbol{\beta}}_i$. The maximum likelihood estimation of $\boldsymbol{\eta}$ is

$$\hat{\boldsymbol{\eta}} = \left(\sum_{i=1}^{N} \mathbf{S}_i^{-1}\right)^{-1} \sum_{i\,=\,1}^{N} \mathbf{S}_i^{-1}\hat{\boldsymbol{\beta}}_i \text{ and cov } (\hat{\boldsymbol{\eta}}) = \left(\sum_{i=1}^{N} \mathbf{S}_i^{-1}\right)^{-1}. \tag{4}$$

As seen, $\hat{\boldsymbol{\eta}}$ is a weighted average of $\hat{\boldsymbol{\beta}}_i$ based on the inverse of covariance. Thus, the maximal log-likelihood under the null hypothesis is $\mathrm{MLL}_0 = \max_{\boldsymbol{\eta}} LL_0(\boldsymbol{\eta}) = \mathrm{LL}_0\ (\hat{\boldsymbol{\eta}}) = -\frac{Nq}{2}\ln(2\pi) - \frac{1}{2}\sum_{i\,=\,1}^{N}\ln(|\mathbf{S}_i|) - \frac{1}{2}\sum_{i\,=\,1}^{N}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}})^T\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}})$.

Under the alternative hypothesis that the circle $z$ is a cluster, the locations inside $z$ have different true values from those outside $z$. With the true values set as $\boldsymbol{\eta}_c$ and $\boldsymbol{\eta}_b$ for the locations inside and outside $z$, respectively, the probability model can be defined as

$$\hat{\boldsymbol{\beta}}_i = \begin{cases} \boldsymbol{\eta}_c + \boldsymbol{\varepsilon}_i, & i \in z \\ \boldsymbol{\eta}_b + \boldsymbol{\varepsilon}_i, & i \notin z \end{cases} \text{ where } \boldsymbol{\varepsilon}_i \sim \mathrm{MN}\left(\mathbf{0}, \mathbf{S}_i\right). \tag{5}$$

Therefore, the log-likelihood can be written as $\mathrm{LL}_z(\boldsymbol{\eta}_c, \boldsymbol{\eta}_b) = -\frac{Nq}{2}\ln(2\pi) - \frac{1}{2}\sum_{i\,=\,1}^{N}\ln(|\mathbf{S}_i|) - \frac{1}{2}\sum_{i(i\in z)} (\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}_c)^{\mathrm{T}}\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}_c) - \frac{1}{2}\sum_{i(i\notin z)}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}_b)^{\mathrm{T}}\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\eta}_b).\ \ .$ Thus, the maximum likelihood estimation of $\boldsymbol{\eta}_c$ and $\boldsymbol{\eta}_b$ can be calculated as $\hat{\boldsymbol{\eta}}_c = (\sum_{i(i\in z)}\mathbf{S}_i^{-1})^{-1}\sum_{i(i\in z)}\mathbf{S}_i^{-1}\hat{\boldsymbol{\beta}}_i$ and $\hat{\boldsymbol{\eta}}_b = (\sum_{i(i\notin z)}\mathbf{S}_i^{-1})^{-1}\sum_{i(i\notin z)}\mathbf{S}_i^{-1}\hat{\boldsymbol{\beta}}_i$, respectively. Then, the maximum log-likelihood under the alteative hypothesis is

$$\begin{aligned}\mathrm{MLL}_z &= \max_{\boldsymbol{\eta}_c,\boldsymbol{\eta}_b}\mathrm{LL}_z\left(\boldsymbol{\eta}_c, \boldsymbol{\eta}_b\right) = \mathrm{LL}_0\left(\hat{\boldsymbol{\eta}}_c, \hat{\boldsymbol{\eta}}_b\right) = -\frac{Nq}{2}\ln\left(2\pi\right) \\ &\quad - \frac{1}{2}\sum_{i\,=\,1}^{N}\ln\left(|\mathbf{S}_i|\right) - \frac{1}{2}\sum_{i(i\in z)}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c\right)^{\mathrm{T}}\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c\right) \\ &\quad - \frac{1}{2}\sum_{i(i\notin z)}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b\right)^{\mathrm{T}}\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b\right).\end{aligned} \tag{6}$$

Furthermore, for circle $z$, the LLR is

$$\begin{aligned}\mathrm{LLR}\ (z) = \mathrm{MLL}_z\ -\ \mathrm{MLL}_0 &= \frac{1}{2}\sum_{i\,=\,1}^{N}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}\right)^T\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}\right) \\ &\quad - \frac{1}{2}\sum_{i(i\in z)}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c\right)^{\mathrm{T}}\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c\right) \\ &\quad - \frac{1}{2}\sum_{i(i\notin z)}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b\right)^{\mathrm{T}}\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b\right).\end{aligned} \tag{7}$$

Finally, the maximal $\mathrm{LLR}(z)$ across all circles is used as the test statistic for the most likely cluster, which is $\max_z \mathrm{LLR}(z) = \max_z \Big\{\frac{1}{2}\sum_{i=1}^{N}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}})^T\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}) - \frac{1}{2}\sum_{i(i\in z)}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c)^T\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_c) - \frac{1}{2}\sum_{i(i\notin z)}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b)^T\mathbf{S}_i^{-1}(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\eta}}_b)\Big\}$ As the first term does not depend on $z$, the most likely cluster is one that maximizes the likelihood under the alternative hypothesis, which is intuitive and conforms to the principle of classic scan statistics. Although the calculation of $\mathrm{LLR}(z)$ above is based on vector observations, it is also applicable for scalar observations, in which $\hat{\boldsymbol{\beta}}_i$, $\hat{\boldsymbol{\eta}}$, $\hat{\boldsymbol{\eta}}_b$, $\hat{\boldsymbol{\eta}}_c$ will be scalars and $\mathbf{S}_i$ is a variance. Compared to the multivariate Gaussian scan statistics (MGSS) by Cucala et al. (2017), EESS allows unequal covariances among spatial locations by sufficiently utilizing the information of uncertainty from the estimated covariances in the first stage. When the estimated covariances are ignored and then all locations share a covariance estimated from the point values $\{\hat{\boldsymbol{\beta}}_i\}$, EESS amounts to MGSS.

### 3.4 | Statistical test for clusters

To control the false positive error rate for multiple testing, the Monte Carlo hypothesis test, based on spatial permutation, is used to test the significance of the detected clusters. Specifically, a large number of random datasets (e.g., $M = 999$) are generated by randomly permutating the pairs of observations and their corresponding covariances (or variances) across all the locations. That is, all the random datasets are derived from the hypothesis of no existing cluster. Note that the permutation requires that each location has at least one observed value to avoid missing data in some locations.

For each of $M$ random datasets, its maximal $\mathrm{LLR}(z)$ is calculated using the method above. Thus, $M$ maximal $\mathrm{LLR}(z)$ values from random datasets and a maximal $\mathrm{LLR}(z)$ value from the real dataset are obtained. If the latter value is among the $\alpha$, such as 5%, the highest of all the values of maximal $\mathrm{LLR}(z)$, that is, $M + 1$ values, the detected most likely cluster is statistically significant with a test level of $\alpha$. More specifically, let $R$ be the rank of the maximal $\mathrm{LLR}(z)$ from the real dataset in $M + 1$ values; then, the $P$ value of the most likely cluster is $R/(M + 1)$. As such, under the null hypothesis, the probability that the Monte Carlo hypothesis test observes a $p$-value less than or equal to $\alpha$ is exactly $\alpha$, that is, the false positive error rate is exactly $\alpha$. This is true regardless of the choice of $M$, but a larger $M$ will provide greater statistical power. To obtain $P$ value with a finite decimal and maintain sufficient statistical power, $M$ is usually set as 999, 1999, or 9999.

When the secondary clusters are also of interest, the $P$ values of the specific secondary clusters can be calculated

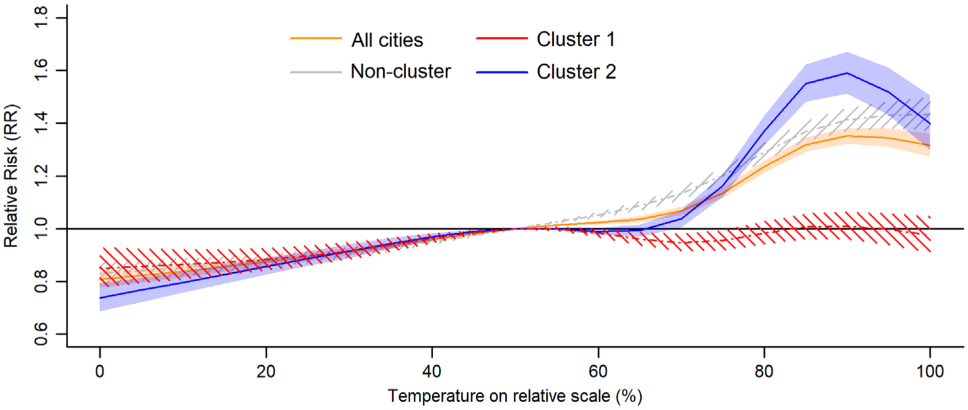

**FIGURE 2** The average ERRs with 95% confidence intervals across the cities in clusters and non-cluster areas. ERR, exposure–response relationship.

by replacing the maximal LLR($z$) from the real dataset with the corresponding secondary LLR($z$) from the real dataset. As the final result, the detected clusters with $P$ values smaller than 0.05, commonly also without geographic overlap, are reported in the decreasing order of LLR($z$).

## 4 | RESULTS: DETECTING THE SPATIAL CLUSTERING OF EXPOSURE–RESPONSE RELATIONSHIP IN THE MOTIVATING EXAMPLE

Based on the city-specific vector $\hat{\boldsymbol{\beta}}_i$ and its covariance $\mathbf{S}_i$ estimated in the first stage, defining the ERR between temperature and HFMD, we used the proposed EESS to detect the spatial clustering of the temperature–HFMD ERR in the motivating example in the second stage. In EESS, the maximal size of the scan window was set to contain 50% of the locations as a default, and $M$ was set as 999 to balance the computational burden with statistical power. The test level was set as 0.05. The statistically significant clusters without geographic overlap, including the most likely cluster and the secondary clusters, were reported. As seen in Figure 1B, a total of two significant clusters ($P = 0.001$ for both) were detected, containing 38 and 36 cities, respectively. For the two clusters and the non-cluster area, we independently calculated their average ERRs, along with the covariances, using a weighted average method as in Equations (4). Then, we compared the ERRs in the clusters with those in the non-cluster area, and the average ERR across all cities under the null hypothesis was also provided as a reference. As in Xiao et al.'s work (2017), taking the 50% percentile of temperature as a reference, we presented the relative risk curves to provide an intuitive comparison.

As shown in Figure 2, the ERRs in clusters 1 and 2 were significantly different from those in the non-cluster area with non-intersect 95% confidence intervals (CIs) at a considerable range of relative temperatures. In addition, even though the sample size in each subgroup was much smaller than the total sample size, the widths of the 95% CIs from the subgroups were inflated slightly compared with those from all cities. A previous study (Wang et al. 2022) showed that incorporating accurate clusters as fixed effects into a regression model will improve the model performance; therefore, we incorporated the two detected clusters as fixed effects in the commonly used multivariate meta regression model. The results showed that incorporating the detected clusters considerably improved the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values over the classic multivariate meta model (AIC: 443.35 vs. 527.93 and BIC: 580.52 vs. 619.38).

## 5 | SIMULATION STUDY

To further evaluate the efficiency of the proposed EESS, a simulation study with various simulation scenarios was carried out. Although MGSS is not developed for spatial datasets with estimation error, we also used MGSS as a comparison since it is applicable for the multivariate spatial datasets when ignoring estimation errors.

### 5.1 | Simulation data generation process

The size and number of clusters will affect the clustering strength (Kulldorff et al. 2003) and thus affect the performance of EESS. Based on the motivating example, we set a total of 12 simulation scenarios, including 11 scenarios with various sizes and numbers of clusters for evaluating the power and 1 without any cluster for evaluating the false positive rate. Specifically, four sizes were set, that is, a cluster consisting of 5, 10, 20, or 40 cities. The number of clusters varied from 1 to 3. For scenarios with clusters consisting of 40 cities, the number of clusters was set as only 1 or 2 due to the heavy geographic overlap for the 3-cluster setting. In addition, the different locations and distances among clusters, as well as heterogeneity between clusters, were also incorporated in these scenarios to reflect the relatively complex distribution of clusters in practical studies. The detailed spatial distributions of these artificial clusters are presented in Figure 3. To make the artificial ERRs more realistic, the true ERRs of clusters and non-cluster areas were set as the estimates in the motivating example based on the detected clusters. That is, referring to Figure 1B, the true multiple parameters characterizing the ERR in non-cluster cities were set as (0.207, 0.290, 0.520, 0.612, 0.507), and two types of true multiple parameters for clusters, that is, (0.272, 0.065, 0.189, 0.200, 0.115) and (0.345, 0.232, 0.818, 0.924, 0.521), were set. The city-specific covariances, $\mathbf{S}_i$, was also obtained from the motivating example. As such, 1000 random datasets were simulated for each scenario.

For simplification, a scenario was labeled with the size and number of clusters; for example, Scen_C2S10 refers to the scenario with two clusters consisting of 10 cities, and Scen_C0S0 refers to the scenario without any clusters.

## 5.2 | Simulation results

The parameter settings for EESS were the same as those in the motivating example. Power and false positive error were used to evaluate the ability of EESS to find clusters. The four classic accuracy performance measures, including sensitivity, specificity, positive predictive value (PPV), and misclassification rate, were also used to evaluate the accuracy of clusters detected by EESS, which were calculated as $a/(a+b)$, $d/(d+c)$, $a/(a+c)$, $(b+c)/(a+b+c+d)$, respectively, where $a$, $b$, $c$, and $d$ indicate the numbers of the four types of cities, that is, cities in detected clusters and true clusters, cities in true clusters and not in detected clusters, cities in detected clusters and not in true clusters, and cities in neither true clusters nor detected clusters, respectively. When denominators were zero, the values were set as not available (NA).

These performances were estimated by taking the average over 1,000 simulation datasets in each scenario. As seen in Table 1, for the proposed EESS, the probability that a cluster was detected in the scenario without any clusters, Scen_C0S0, was very close to the test level 0.05, suggesting that EESS is able to control the false positive error as expected. For scenarios with clusters, EESS achieved different levels of power in different scenarios. As expected, as the size and number of clusters increased, the power increased, and the simulated power even reached 1 in some scenarios with multiple clusters and large clusters. The sensitivity also increased, as expected, as the size increased, especially in the scenarios with a single large-size cluster, that is, Scen_C1S40, where the sensitivity was more than 0.99. Excluding the scenario Scen_C3S20, the specificity was always more than 0.95, and the misclassification and PPV also indicated the good performance. For MGSS, although the power and sensitivity are high, the other performances, that is, specificity, PPV, and misclassification are poor, suggesting that MGSS is not an efficient selection in detecting the spatial clustering of ERRs.

Getting insight into the reason why EESS exhibited unsatisfactory performance in Scen_C3S20, as seen in Figure 3, we found that the two type-2 clusters in Scen_C3S20 were close together, which may lead EESS to detect one large cluster containing the two true clusters as the most likely cluster, rather than either of the two true clusters. As presented in Figure 4, this large cluster included many cities between the two true clusters, leading to the poor specificity, PPV, and misclassification. Therefore, we used a smaller maximum scan size, containing 20% cities, to redetect the clusters in Scen_C3S20; as expected, EESS exhibited a much-improved performance in this analysis. That is, the power, sensitivity, specificity, PPV and misclassification were improved to 1, 0.9498, 0.9461, 0.9336, and 0.0523, respectively.

# 6 | DISCUSSION

Motivated by identifying the spatial clustering of ERRs, in this work, we proposed a novel spatial scan statistic called EESS, which extends Kulldorff's scan statistic to apply to data with estimation error. This type of data can be either scalar or vector, which usually come from derived (or second-hand) datasets estimated by a statistical model, such as the estimated regression coefficients with variances (or covariances for vectors) characterizing the region-specific ERRs. This type of data has been generated in a large number of multi-region ERR studies, which have become the mainstream approach for characterizing ERR. Furthermore, both a motivating example dataset and a batch of simulation datasets demonstrated that EESS performed well in detecting the spatial clustering of ERRs.

Specifically, in the motivating example of detecting the spatial clustering of ERRs between temperature and HFMD, two clusters were detected. One was located in central China, and the other was located in southeastern China. As seen in Figure 2, HFMD risk was more sensitive to temperature in southeastern China, especially when temperature rose, indicating that the government should be more alert to an outbreak of HFMD under those conditions. In central China, HFMD risk was not sensitive to high temperatures. Obtaining insight into the differences

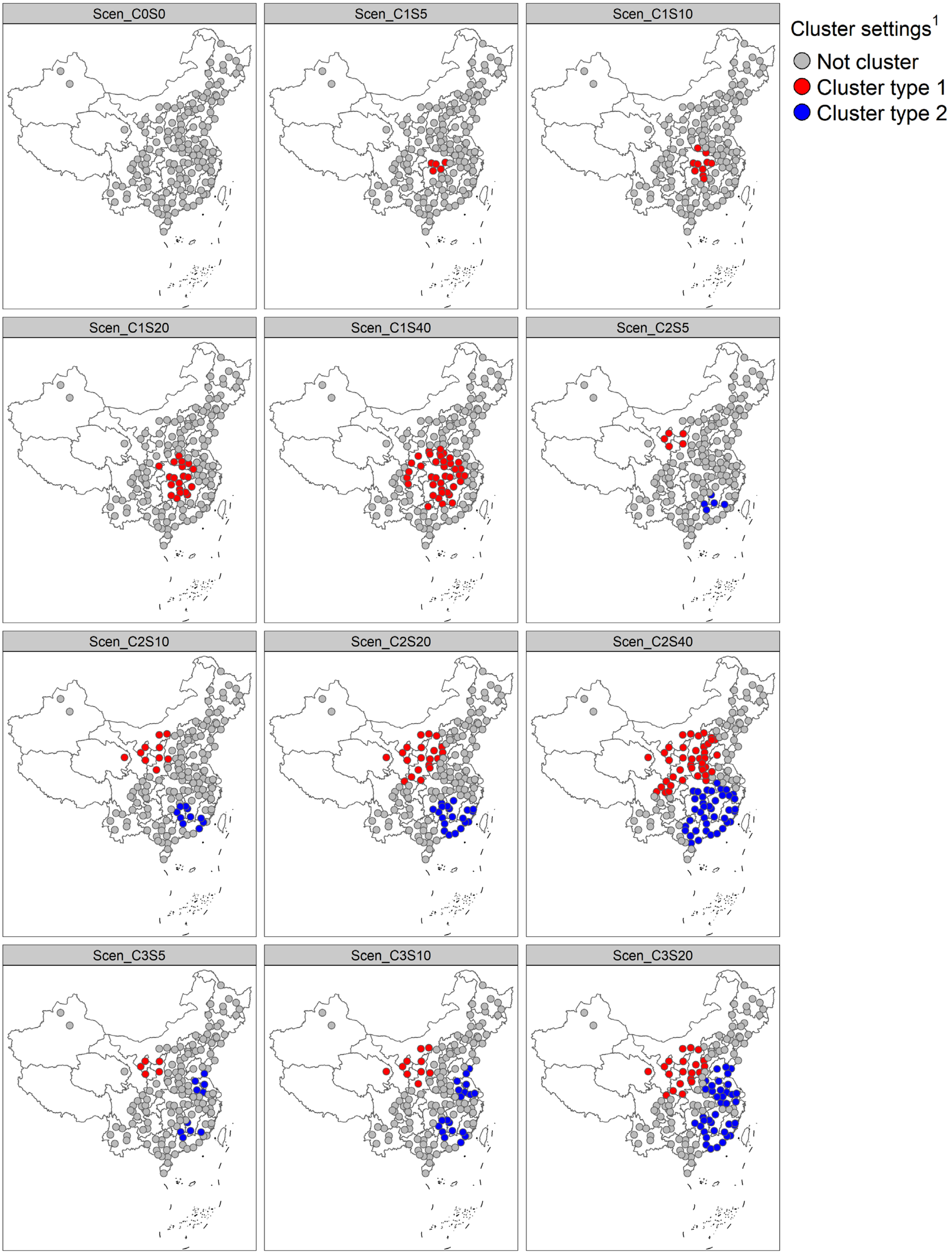

**FIGURE 3** The distribution of the artificial clusters for the 12 scenarios. [1] In scenarios with three clusters, two clusters belong to cluster type 2, and one belongs to cluster type 1. The true multiple parameters characterizing the ERRs were set as (0.207, 0.290, 0.520, 0.612, 0.507) (0.272, 0.065, 0.189, 0.200, 0.115), and (0.345, 0.232, 0.818, 0.924, 0.521) for cities in not cluster (gray), cluster type 1 (red), and cluster type 2 (blue), respectively. ERR, exposure–response relationship.

**TABLE 1** The average performance over 1000 simulation random datasets in each scenario for EESS and MGSS

| | EESS | | | | | MGSS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Scenario | Power | Sensitivity | Specificity | PPV | Misclassification | Power | Sensitivity | Specificity | PPV | Misclassification |
| Scen_C0S0 | 0.051 [1] | NA | 0.9923 | NA | 0.0077 | 1 | NA | 0.5303 | NA | 0.4697 |
| Scen_C1S5 | 0.565 | 0.5378 | 0.9863 | 0.8047 | 0.0294 | 1 | 0.8876 | 0.5762 | 0.075 | 0.4129 |
| Scen_C1S10 | 0.969 | 0.8997 | 0.9889 | 0.9144 | 0.0173 | 1 | 0.935 | 0.627 | 0.1688 | 0.3515 |
| Scen_C1S20 | 1 | 0.9618 | 0.9894 | 0.9531 | 0.0144 | 1 | 0.957 | 0.6786 | 0.3423 | 0.2825 |
| Scen_C1S40 | 1 | 0.9954 | 0.9921 | 0.9857 | 0.007 | 1 | 0.9912 | 0.654 | 0.5435 | 0.2517 |
| Scen_C2S5 | 0.844 | 0.5057 | 0.9865 | 0.9025 | 0.0471 | 1 | 0.9048 | 0.6188 | 0.1612 | 0.3612 |
| Scen_C2S10 | 1 | 0.8616 | 0.9905 | 0.9599 | 0.0276 | 1 | 0.95 | 0.6711 | 0.3377 | 0.2899 |
| Scen_C2S20 | 1 | 0.8932 | 0.9559 | 0.9359 | 0.0616 | 1 | 0.9773 | 0.7063 | 0.5871 | 0.2179 |
| Scen_C2S40 | 1 | 0.988 | 0.9751 | 0.983 | 0.0177 | 1 | 0.9878 | 0.8006 | 0.8736 | 0.0947 |
| Scen_C3S5 | 0.974 | 0.4235 | 0.9953 | 0.9707 | 0.0647 | 1 | 0.9137 | 0.6327 | 0.2374 | 0.3379 |
| Scen_C3S10 | 1 | 0.7461 | 0.965 | 0.9206 | 0.0809 | 1 | 0.9585 | 0.6415 | 0.4351 | 0.292 |
| Scen_C3S20 | 1 | 0.9429 | 0.7798 | 0.762 | 0.1518 | 1 | 0.9738 | 0.4466 | 0.5689 | 0.3322 |

*Note*: [1]The value indicates false positive error since Scen_C0S0 contains no cluster. EESS, estimation-error-based scan statistic; MGSS, multivariate Gaussian scan statistics; PPV, positive predictive value.

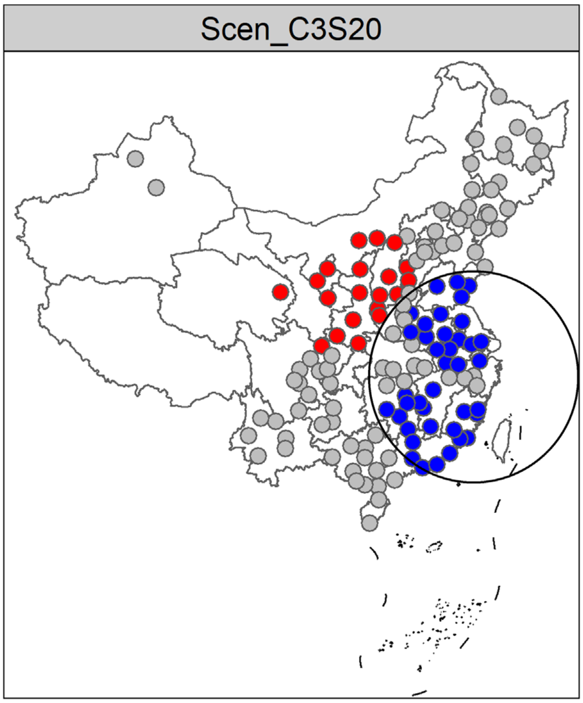

**FIGURE 4** Illustration of a large false cluster detected containing the two true clusters in Scen_C3S20.

in other factors, such as economic and natural conditions, among clusters and non-cluster areas, the detected clusters may assist in identifying the source of heterogeneity of ERRs. In this case, the low sensitivity of HFMD risk to high temperature in central China may partly derive from the low precipitation and relative humidity, which restrict the reproduction and transmission of enterovirus even at an optimal high temperature (Wang et al. 2011). In addition, we incorporated the two detected clusters as fixed effects into the commonly used multivariate meta regression model, and the performance of the model in terms of AIC and BIC was considerably improved (AIC: 443.35 vs. 527.93 and BIC: 580.52 vs. 619.38), which provided a support to the efficiency of the EESS identifying clusters or heterogeneity. On the other hand, the improved adjusted goodness-of-fit indices, that is, AIC and BIC, also suggested that incorporating the detected clusters into meta-analysis model benefits to achieve more accurate region-specific ERR estimates, which are the key to obtain accurate attributable burdens.

In the simulation study, EESS also exhibited good performance, as expected, in terms of false positive error, power of detecting clusters, sensitivity, specificity, PPV and misclassification in almost all scenarios. Generally, EESS exhibited larger power, sensitivity, and PPV when the size of the cluster increased because the larger size represented stronger clustering. A seemingly abnormal result was obtained; EESS exhibited poor performance in terms of specificity, PPV and misclassification in Scen_C3S20, while exhibiting good performance in all the other scenarios. Alternatively, we used a smaller maximum scan size, containing 20% cities, to redetect the clusters in Scen_C3S20; as expected, EESS exhibited a much-improved performance in this analysis. This result suggests that selecting an appropriate maximum scan size is also crucial to EESS, which is similar to the classic scan statistics (Wang et al. 2020). Fortunately, the Gini coefficient (Han et al. 2016), Maximum clustering set-proportion (MCS-P) (Ma et al. 2016), and Maximum clustering heterogeneous set-proportion (MCHS-P) (Wang et al. 2020) have been proposed to tune an appropriate maximum scan size or maximum reported size in practical studies to guarantee the accuracy. Therefore, a modified Gini coefficient, MCS-P, and MCHS-P may also be appropriate for EESS, which will be the focus of our future work.

Similar to the analytic process in the motivating example, based on the proposed EESS, a two-stage analytic process can be used to detect the spatial clustering of ERRs in practical studies. In the first stage, the common region-stratified method is used to estimate the region-specific regression coefficients characterizing the ERRs. When the ERR is assumed to be a simple linear relationship, the regression coefficient is a scalar with variance, commonly from a general linear model, logistical model, Poisson model, or survival model. When the ERR presents a complex nonlinear relationship even with lag effects, the regression coefficient is usually represented as a vector with covariance, commonly from a GAM, such as DLNM. In the second stage, based on the estimated region-specific observations with each respective variance (or covariance), EESS is used to detect the spatial clustering of ERRs. As such, EESS does not depend on the raw data and only depends on the estimation results in the first stage, which suggests that a large number of existing datasets from published studies can be used to carry out a further analysis in terms of identifying the clustering of ERRs. The detected clusters will help to achieve a more accurate ERR estimation and to further identify high-sensitivity or ERR-heterogeneous regions, which play important roles in designing reasonable region-specific intervention measures. To facilitate the further analysis, we strongly recommend that researchers provide the covariance (or variance) along with the region-specific ERR estimations from the first stage in a multi-region ERR study, at least as a supporting information, even though the region-specific ERRs are not their focus.

Although in this work EESS is used to detect the spatial clustering of ERRs, it can also be used for other objectives of spatial clustering detection based on datasets with estimation error. For example, we may aim to detect the spatial clustering of the blood pressure level, which can reflect the overall health status of an individual and is much associated with dietary pattern (Tseng et al. 2021). However, in many regions, the raw datasets, that is, the individual's blood pressure levels, are not accessible, and, thus, the classic scan statistic is unavailable. The sample mean of blood pressure level and its standard error for each region may be available and, in such cases, EESS makes the spatial clustering detection of blood pressure level possible.

At present, we must acknowledge that EESS is computationally intensive due to the test process based on standard Monte Carlo randomization. For the classic scan statistic, the sequential Monte Carlo (Besag et al. 1991) and Gumbel approximation (Abrams et al. 2010) have been proposed as alternatives to reduce the computational time requirements. However, whether such alternatives can also be applied for EESS requires further research. When calculating the $p$-value for secondary clusters, we did not adjust for the most likely cluster, which may lead to the loss of power in testing potential secondary clusters. Alternatively, a sequential version of EESS may help to improve this issue as in Zhang et al.'s work (2010). In addition, we only developed EESS based on circular windows, which may not obtain enough power for non-circular clusters, in such case, the EESS based on elliptical (Kulldorff et al. 2006) or irregular windows (Duczmal et al. 2007) deserves to be developed.

## ACKNOWLEDGMENTS

The authors would like to thank Xiong Xiao for providing the city-specific ERRs from the first stage in the motivating example.

## DATA AVAILABILITY STATEMENT

The data and codes that support the findings in this paper are available in the Supporting information of this paper.

## REFERENCES

Abolhassani, A. & Prates, M.O. (2021) An up-to-date review of scan statistics. *Statistics Surveys*, 15, 111–153.

Abrams, A.M., Kleinman, K. & Kulldorff, M. (2010) Gumbel based $p$-value approximations for spatial scan statistics. *International Journal of Health Geographics*, 9, 61. https://doi.org/10.1186/1476-072x-9-61

Belanger, M., Gray-Donald, K., O'Loughlin, J., Paradis, G. & Hanley, J. (2009) Influence of weather conditions and season on physical activity in adolescents. *Annals of Epidemiology*, 19(3), 180–186. https://doi.org/10.1016/j.annepidem.2008.12.008

Bertrand, I., Schijven, J.F., Sanchez, G., Wyn-Jones, P., Ottoson, J., Morin, T., et al. (2012) The impact of temperature on the inactivation of enteric viruses in food and water: A review. *Journal of Applied Microbiology*, 112(6), 1059–1074. https://doi.org/10.1111/j.1365-2672.2012.05267.x.

Besag, J. & Clifford, P. (1991) Sequential Monte Carlo $p$-values. *Biometrika*, 78(2), 301–304.

Cadena, J., Chen, F. & Vullikanti, A. (2017. Near-optimal and practical algorithms for graph scan statistics. Paper read at Proceedings of the 2017 SIAM International Conference on Data Mining, SIAM.

Cançado, A.L.F., da-Silva, C.Q. & Silva, M.F. (2014) A spatial scan statistic for zero-inflated Poisson process. *Environmental and Ecological Statistics*, 21(4), 627–650.

Cançado, A.L.F., Fernandes, L.B. & da-Silva, C.Q. (2017) A Bayesian spatial scan statistic for zero-inflated count data. *Spatial Statistics*, 20, 57–75.

Cheng, Q., Bai, L.J., Zhang, Y.W., Zhang, H., Wang, S.S., Xie, M.Y., et al. (2018) Ambient temperature, humidity and hand, foot, and mouth disease: a systematic review and meta-analysis. *Science of the Total Environment*, 625, 828–836. https://doi.org/10.1016/j.scitotenv.2018.01.006

Cucala, L. (2014) A distribution-free spatial scan statistic for marked point processes. *Spatial Statistics*, 10, 117–125. https://doi.org/10.1016/j.spasta.2014.03.004

Cucala, L., Genin, M., Lanier, C. & Occelli, F. (2017) A multivariate Gaussian scan statistic for spatial data. *Spatial Statistics*, 21, 66–74.

Duczmal, L., Cançado, A.L.F., Takahashi, R.H.C. & Bessegato, L.F. (2007) A genetic algorithm for irregularly shaped spatial scan statistics. *Computational Statistics & Data Analysis*, 52(1), 43–52.

Gasparrini, A. (2014) Modeling exposure–lag–response associations with distributed lag non-linear models. *Statistics in Medicine*, 33(5), 881–899. https://doi.org/10.1002/sim.5963

Gasparrini, A., Armstrong, B. & Kenward, M.G. (2010) Distributed lag non-linear models. *Statistics in Medicine*, 29(21), 2224–2234. https://doi.org/10.1002/sim.3940

Gasparrini, A., Armstrong, B. & Kenward, M.G. (2012) Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*, 31(29), 3821–3839. https://doi.org/10.1002/sim.5471

Han, J., Zhu, L., Kulldorff, M., Hostovich, S., Stinchcomb, D.G., Tatalovich, Z., et al. (2016) Using Gini coefficient to determining optimal cluster reporting sizes for spatial scan statistics. *International Journal of Health Geographics*, 15, 27. https://doi.org/10.1186/s12942-016-0056-6

Hu, K.J., Guo, Y.T., Hoehrainer-Stigler, S., Liu, W., See, L., Yang, X.C., et al. (2019) Evidence for urban-rural disparity in temperature-mortality relationships in Zhejiang province, China. *Environmental Health Perspectives*, 127(3), 37001. https://doi.org/10.1289/ehp3556

Huang, L., Kulldorff, M. & Gregorio, D. (2007) A spatial scan statistic for survival data. *Biometrics*, 63(1), 109–118.

Huang, L., Tiwari, RC., Zou, Z., Kulldorff, M. & Feuer, E.J. (2009) Weighted normal spatial scan statistic for heterogeneous population data. *Journal of the American Statistical Association*, 104(487), 886–898. https://doi.org/10.1198/jasa.2009.ap07613

Koh, W.M., Badaruddin, H., La, H., Chen, M.I.C. & Cook, A.R. (2018) Severity and burden of hand, foot and mouth disease in Asia: a modelling study. *Bmj Global Health*, 3(1), e000442. https://doi.org/10.1136/bmjgh-2017-000442

Kulldorff, M. (1997) A spatial scan statistic. *Communications in Statistics: Theory and Methods*, 26(6), 1481–1496. https://doi.org/10.1080/03610929708831995

Kulldorff, M., Tango, T. & Park, P.J. (2003) Power comparisons for disease clustering tests. *Computational Statistics & Data Analysis*, 42(4), 665–684. https://doi.org/10.1016/s0167-9473(02)00160-3

Kulldorff, M., Huang, L. & Konty, K. (2009) A scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics*, 8(1), 1–9.

Kulldorff, M., Huang, L., Pickle, L. & Duczmal, L. (2006) An elliptic spatial scan statistic. *Statistics in Medicine*, 25(22), 3929–3943.

Kulldorff, M. & Nagarwalla, N. (1995) Spatial disease clusters: detection and inference. *Statistics in Medicine*, 14(8), 799–810.

Ma, Y., Yin, F., Zhang, T., Zhou, X.A. & Li, X.S. (2016) Selection of the maximum spatial cluster size of the spatial scan statistic by using the maximum clustering set-proportion statistic. *Plos One*, 11(1), e0147918. https://doi.org/10.1371/journal.pone.0147918

Meng, X., Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A.M., Milojevic, A.i., et al. (2021) Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities. *Bmj British Medical Journal*, 372, n534. https://doi.org/10.1136/bmj.n534

Requia, W.J., Adams, M.D., Arain, A., Papatheodorou, S., Koutrakis, P. & Mahmoud, M. (2018) Global association of air pollution and cardiorespiratory diseases: a systematic review, meta-analysis, and investigation of modifier variables. *American Journal of Public Health*, 108, S123–S130. https://doi.org/10.2105/ajph.2017.303839

Samet, J.M., Zeger, S.L., Dominici, F., Curriero, F., Coursac, I., Dockery, D.W., et al. (2000) The national morbidity, mortality, and air pollution study. Part ii: Morbidity and mortality from air pollution in the United States. *Research Report (Health Effects Institute)*, 94(Pt 2), 5–70. discussion 71–79.

Shah, A.S.V., Langrish, J.P., Nair, H., McAllister, D.A., Hunter, A.L., Donaldson, K., et al. (2013) Global association of air pollution and heart failure: a systematic review and meta-analysis. *Lancet*, 382(9897), 1039–1048. https://doi.org/10.1016/s0140-6736(13)60898-3

Shah, A.S.V., Lee, K.K., McAllister, D.A., Hunter, A., Nair, H., Whiteley, W., et al. (2015) Short term exposure to air pollution and stroke: systematic review and meta-analysis. *British Medical Journal*, 350, h1295. https://doi.org/10.1136/bmj.h1295

Tseng, E., Appel, L.J., Yeh, H.-C., Pilla, S.J., Miller, E.R., Juraschek, S.P., et al. (2021) Effects of the dietary approaches to stop hypertension diet and sodium reduction on blood pressure in persons with diabetes. *Hypertension*, 77 (2), 265–274. https://doi.org/10.1161/hypertensionaha.120.14584

Wang, W., Xiao, X., Qian, J., Chen, S.Q., Liao, F., Yin, F., et al. (2022) Reclaiming independence in spatial-clustering datasets: a series of data-driven spatial weights matrices. *Statistics in Medicine*, 10, 41(15), 2939–2956. https://doi.org/10.1002/sim.9395

Wang, W., Zhang, T., Yin, F., Xiao, X., Chen, S.Q., Zhang, X.Y., et al. (2020) Using the maximum clustering heterogeneous set-proportion to select the maximum window size for the spatial scan statistic. *Scientific Reports*, 10(1), 4900. https://doi.org/10.1038/s41598-020-61829-y

Wang, Yu, Feng, Z., Yang, Y., Self, S., Gao, Y., Longini, I.M., et al. (2011) Hand, foot, and mouth disease in china patterns of spread and transmissibility. *Epidemiology*, 22(6), 781–792. https://doi.org/10.1097/EDE.0b013e318231d67a

Xiao, X., Gasparrini, A., Huang, J., Liao, Q.H., Liu, F.F., Yin, F., et al. (2017) The exposure–response relationship between temperature and childhood hand, foot and mouth disease: a multicity study from mainland China. *Environment International*, 100, 102–109. https://doi.org/10.1016/j.envint.2016.11.021

Zhang, Z., Assunção, R. & Kulldorff, M. (2010) Spatial scan statistics adjusted for multiple clusters. *Journal of Probability and Statistics*, 2010, 642379. https://doi.org/10.1155/2010/642379

Zhuang, Z.C., Kou, Z.Q., Bai, Y.J., Cong, X., Wang, L.H., Li, C., et al. (2015) Epidemiological research on hand, foot, and mouth disease in mainland China. *Viruses*, 7(12), 6400–6411. https://doi.org/10.3390/v7122947

## SUPPORTING INFORMATION

Supporting Information.

Web appendices for the data and R codes in Sections 2 and 3 are available with this paper at the Biometrics website on Wiley Online Library.

**How to cite this article:** Wang, W., Li, S., Zhang, T., Yin, F. & Ma, Y. (2023) Detecting the spatial clustering of exposure–response relationships with estimation error: a novel spatial scan statistic. *Biometrics*, 1–11. https://doi.org/10.1111/biom.13861