The Annals of Applied Statistics
2026, Vol. 20, No. 2, 1364–1387
https://doi.org/10.1214/26-AOAS2153
© Institute of Mathematical Statistics, 2026

TRANSPARENT SEQUENTIAL LEARNING AND MONITORING OF SPATIOTEMPORAL DISEASE INCIDENCE RATES

BY YUHANG ZHOU[a] AND PEIHUA QIU[b]

Department of Biostatistics, University of Florida, [a]*zhouyuhang@ufl.edu,* [b]*pqiu@ufl.edu*

Our society is under constant threat from outbreaks of various infectious diseases, such as COVID-19, Zika, and others. The recent COVID-19 pandemic has claimed millions of lives and caused devastating disruption to our daily routines. Prompt detection of outbreaks and effective disease surveillance are critical yet challenging, due to the complex spatiotemporal dynamics of infectious disease spread. Existing analytical tools often rely on restrictive assumptions, such as data independence and specific parametric distributions, that are rarely valid in practical scenarios. Additionally, effective disease surveillance demands sequential decision-making since decisions should be made or updated whenever new data become available. But many existing methods were designed for retrospective data observed in a prespecified time interval and would not be effective for disease surveillance. To address these limitations, we develop a cumulative sum (CUSUM) control chart for sequentially monitoring the evolution of spatiotemporal disease incidence rates. The new chart is constructed under the sequential learning framework that continuously incorporates new data in updating the estimated baseline model. Unlike traditional methods, the new method can capture complex data structure including spatiotemporal data variation and correlation. Numerical studies indicate that it achieves faster and more reliable detection of disease outbreaks compared to some traditional methods.

1. Introduction. Our society has been facing increasing threats of infectious diseases, including outbreaks of COVID-19, H5N1, Zika, and many others. In addition to natural outbreaks, there is the ongoing risk of bioterrorist attacks. The recent COVID-19 pandemic alone has claimed millions of lives and disrupted nearly every aspect of life worldwide. To combat these destructive diseases, it is crucial to promptly detect outbreaks and understand their spread dynamics (Murray and Cohen (2017)). This allows governments and the public to implement timely interventions to minimize their impact. This paper aims to develop an effective analytical tool to achieve such goals.

To collect infectious disease data, various national and regional disease reporting systems have been established worldwide. For instance, the Outpatient Influenza-like Illness (ILI) Surveillance Network (ILINet) is a national system that reports weekly ILI incidence in all 50 states of the U.S., where ILI is a respiratory infection caused by influenza viruses that has the major symptoms of fever, cough, and/or sore throat. Figure 1 shows the observed ILI incidence rates in the 48 contiguous U.S. states for Week 46 of 2018 (left) and Week 46 of 2019 (right), where darker colors indicate larger values.

Figure 1 illustrates that disease incidence rates at each time are typically collected across multiple locations, making the data *spatial* in nature. These spatial data are often gathered sequentially over time, resulting in a *spatiotemporal data stream* where new data continuously emerge. One major complexity in modeling and surveilling infectious diseases stems from the data structure itself. Observed disease incidence data often exhibit intricate spatiotemporal variation and correlation. For instance, data collected at different locations and times

Received November 2025; revised January 2026.
Key words and phrases. Control charts, correlation, disease incidence rates, process monitoring, sequential learning, spatiotemporal data.

FIG. 1. *Observed ILI incidence rates in the* 48 *contiguous U.S. states for Week* 46 *of* 2018 (*Nov* 12–18, 2018; *left*) *and Week* 46 *of* 2019 (*Nov* 11–17, 2019; *right*). *Darker colors indicate larger incidence rates*.

are usually correlated, with stronger correlations observed among those closer in space or time. This pattern reflects the influence of various confounding risk factors, such as weather, lifestyle, and other demographic, cultural, or environmental factors as well as the transmission nature of infectious diseases. Because these impacts are difficult to quantify, the spatiotemporal correlations in observed disease incidence data are also challenging to describe and analyze. Another major complexity in infectious disease surveillance is its nature as a *sequential decision-making problem*: a statistical decision regarding the presence of a disease outbreak must be made each time new data are collected. Sequential decision-making is inherently challenging because the temporal dynamics of the underlying process are difficult to study, and justifying the performance of numerous sequential decisions is equally complex.

Given the critical importance of disease surveillance, this issue has been extensively discussed in the literature. Early methods, such as the Knox, Mantel, and k-nearest neighbor methods (e.g., Jacquez (1996), Knox and Bartlett (1964), Mantel (1967)), have been used to identify irregular space-time patterns in observed disease incidence data. Other popular approaches rely on spatial or spatiotemporal scan statistics (e.g., Weinstock (1981), Kulldorff (1997, 2001), Takahashi et al. (2008)) to test whether the observed disease incidences in specific windows, often circular or other flexible shapes, are significantly higher than expected under the null hypothesis of no outbreaks. Additional methods for disease surveillance include those based on Bayesian spatial modeling (e.g., Besag, York and Mollié (1991), Lawson et al. (2000), Best, Richardson and Thomson (2005), Zhou and Lawson (2008)), generalized linear mixed-effects modeling (e.g., Kleinman, Lazarus and Platt (2004)), parametric and semiparametric regression modeling (e.g., Pelat et al. (2007), Stroup et al. (1999)), time series modeling (e.g., Heisterkamp, Dekkers and Heijne (2006), Reis and Mandl (2003), Zhao et al. (2011)), and point process modeling (e.g., Diggle, Rowlingson and Su (2005)), among others. For a recent comprehensive review of Bayesian spatiotemporal modeling approaches addressing a range of epidemiological problems, see Wang, Chen and Xue (2024). In addition, point-process techniques have facilitated disease monitoring through the estimation of disease velocity (Rodriguez Avellaneda, Mateu and Moraga (2024)). However, most of these methods are retrospective, meaning they analyze data over a set time interval, which makes them less effective for addressing the challenge of sequential disease surveillance, where new data continually emerge (Sonesson and Bock (2003), Lawson and Kleinman (2005), Marshall, Spitzner and Woodall (2007)). Additionally, the validity of their results is often difficult to justify, as the model assumptions required are hard to meet in practice (Zhang et al. (2015), Qiu (2024)).

Statistical process control (SPC) charts are widely used for handling sequential decision-making problems. However, most SPC methods are designed for monitoring univariate or multivariate data streams, assuming that sequential observations are independent and identically distributed (i.i.d.) when the underlying processes are performing satisfactorily or are

in-control (IC) (Qiu (2014)). While a few SPC charts have been developed to monitor spatial data streams, their applicability is often limited. For instance, Yan, Paynabar and Shi (2018) proposed a method that decomposes observed spatiotemporal data into functional mean, anomaly, and random error components, under the assumptions that errors are i.i.d. and normally distributed across space and time. Similarly, Yang and Qiu (2020) introduced a cumulative sum (CUSUM) chart based on a spatiotemporal nonparametric regression model estimated from a large IC dataset collected in advance, which may not be available in practice. In addition, recent SPC developments, including self-starting control charts designed for high-dimensional and short-run processes (Bodnar, Bodnar and Schmid (2023)) and rank-based EWMA charts that are robust to heteroscedasticity (Wang, Goedhart and Zwetsloot (2023)), have broadened the applicability of SPC to complex data structures. Another closely related line of research involves online change-point detection, which seeks to estimate the occurrence time of distributional shifts. While such methods are effective in identifying potential change points, many classical approaches (e.g., Hawkins, Qiu and Kang (2003), Gee et al. (2018), Yu et al. (2023)) depend on independence or parametric distributional assumptions that are often violated in spatiotemporal contexts. Despite these advances, most SPC methods either rely on model assumptions that are difficult to satisfy in practice or fail to fully utilize the available information during online monitoring, thereby limiting their adaptability and detection accuracy. Overcoming these limitations could substantially enhance the effectiveness of SPC charts in identifying anomalies within complex spatiotemporal data streams.

To address the limitations of existing methods, we propose an approach that integrates recent advancements in machine learning through a framework known as Transparent Sequential Learning (TSL) (Qiu and Xie (2022)). TSL has demonstrated its potential in SPC applications for monitoring serially correlated univariate or multivariate data streams by incorporating self-starting process monitoring techniques (Hawkins (1987), Xie and Qiu (2023)). In this paper we extend the TSL framework to spatiotemporal data streams, introducing a spatiotemporal TSL approach. In this method, observed data at the current time point are incorporated into the IC dataset, and estimates of key IC quantities are recursively updated during process monitoring, provided the underlying spatial process is deemed IC at the current time point. The statistical consistency of these estimated IC quantities is established under several regularity conditions, including a commonly used assumption in the literature that the temporal dependence structure exhibits an exponentially decaying α-mixing coefficient (e.g., Bradley (2005), Yang and Qiu (2019)). Details are provided in Theorems 2.1 and 2.2 in Section 2. The sequential learning strategy TSL empowers the online monitoring procedure to fully exploit historical information at each observation time. Through continuous data integration and adaptive updating of IC quantity estimates, the proposed method offers a robust framework for detecting disease outbreaks in complex spatiotemporal data streams, substantially improving adaptability and detection accuracy.

The remainder of the paper is organized as follows: Section 2 provides a detailed description of the proposed method. Section 3 evaluates its numerical performance through simulation studies. In Section 4 the method is illustrated using a real-data example about influenza-like illness surveillance. Finally, several concluding remarks are presented in Section 5.

2. Spatiotemporal disease surveillance by transparent sequential learning and monitoring.

The proposed method is demonstrated in Figure 2. It comprises the following major steps:

Initial baseline model estimation: A baseline model is estimated from an initial IC dataset to describe the regular spatiotemporal pattern of disease incidence rates under monitoring.

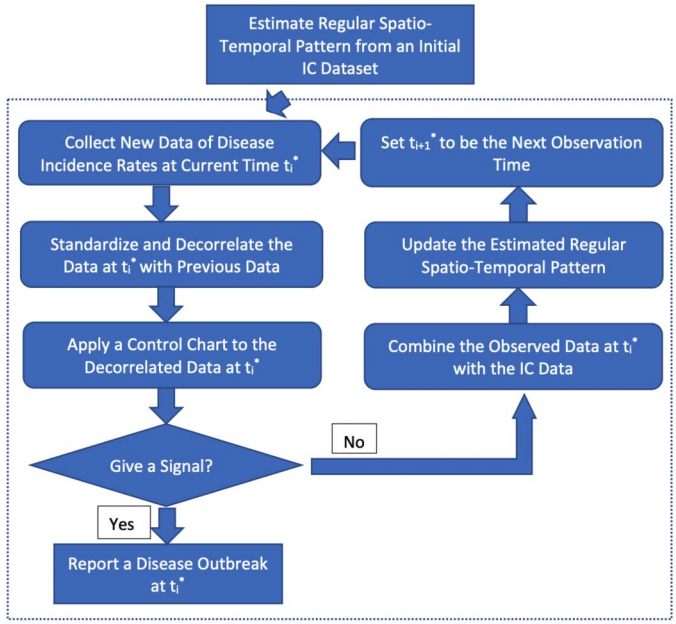

FIG. 2. *Demonstration of the proposed spatiotemporal disease surveillance method. The dotted rectangle highlights the Transparent Sequential Learning procedure.*

Data preprocessing: After a new batch of data is collected at the current observation time, it is standardized and decorrelated recursively with all previous data. This is achieved using the estimated regular spatiotemporal pattern and a dedicated data preprocessing procedure.

Sequential disease surveillance: A CUSUM chart is developed and applied to the preprocessed data at the current observation time for disease surveillance. If a disease outbreak is detected, the chart signals an alert. Otherwise, the newly observed data are merged with the IC dataset, and the estimated baseline model is updated recursively to support disease surveillance at the next time point.

This method is built on a *sequential learning* framework, where the regular spatiotemporal pattern is learned incrementally from the continuously expanding IC dataset. The learning process is *transparent* in the sense that the knowledge learned from the IC dataset is clearly defined and interpretable. Further details of these key steps are provided in the following subsections.

2.1. *Initial estimation of the regular spatiotemporal pattern.* Assume that an initial IC spatiotemporal dataset, denoted by $\{y(t_i, \boldsymbol{s}_{ij}) : j = 1, \ldots, m_i; i = 1, \ldots, n\}$, is available prior to online monitoring. This IC dataset follows the nonparametric spatiotemporal model:

$$y(t_i, \boldsymbol{s}_{ij}) = \mu(t_i, \boldsymbol{s}_{ij}) + \epsilon(t_i, \boldsymbol{s}_{ij}), \quad j = 1, \ldots, m_i; i = 1, \ldots, n, \tag{1}$$

where $y(t_i, \boldsymbol{s}_{ij})$ represents the observed disease incidence rate at time $t_i \in [0, T]$ and location $\boldsymbol{s}_{ij} \in \Omega$. The term $\mu(t_i, \boldsymbol{s}_{ij})$ denotes the mean, and $\epsilon(t_i, \boldsymbol{s}_{ij})$ is a zero-mean random error. Here m_i represents the number of observation locations at time t_i, and n is the total number of observation times in the initial IC dataset. If seasonality is relevant to the disease, then the time interval $[0, T]$ should span a full seasonal cycle.

The spatiotemporal correlation structure is captured by the covariance function

$$V(t, t'; \boldsymbol{s}, \boldsymbol{s}') = \mathrm{Cov}\big(y(t, \boldsymbol{s}), y(t', \boldsymbol{s}')\big), \quad \text{for } (t, \boldsymbol{s}), (t', \boldsymbol{s}') \in [0, T] \times \Omega.$$

The mean function $\mu(t,\boldsymbol{s})$ in model (1) can be estimated using the spatiotemporal local linear kernel smoothing procedure proposed by Yang and Qiu (2018),

$$\begin{aligned}\arg\min_{\boldsymbol{\theta}\in\mathbb{R}^4}&\sum_{i=1}^{n}\sum_{j=1}^{m_i}\big[y(t_i,\boldsymbol{s}_{ij})-\theta_\mu-\theta_t(t_i-t)-\theta_u(s_{x,ij}-s_x)-\theta_v(s_{y,ij}-s_y)\big]^2\\&\times K_t\bigg(\frac{t_i-t}{h_t}\bigg)K_s\bigg(\frac{d_E(\boldsymbol{s}_{ij},\boldsymbol{s})}{h_s}\bigg),\end{aligned}\tag{2}$$

where $\boldsymbol{\theta}=(\theta_\mu,\theta_t,\theta_u,\theta_v)^\top$, $\boldsymbol{s}=(s_x,s_y)^\top$, $\boldsymbol{s}_{ij}=(s_{x,ij},s_{y,ij})^\top$, $K_t(\cdot)$ and $K_s(\cdot)$ are kernel functions, $h_t>0$ and $h_s>0$ are bandwidths, and $d_E(\cdot,\cdot)$ denotes Euclidean distance.

The initial estimate of $\mu(t,\boldsymbol{s})$ is obtained as the solution to θ_μ in the minimization problem (2), which is expressed as

$$\widehat{\mu}^{(0)}(t,\boldsymbol{s})=\boldsymbol{\zeta}^\top(\mathbf{X}^\top\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{W}\mathbf{Y},\tag{3}$$

where $\boldsymbol{\zeta}=(1,0,0,0)^\top$, and:

- $\mathbf{X}$ is the design matrix with rows $(1,t_i-t,s_{x,ij}-s_x,s_{y,ij}-s_y)$.
- $\mathbf{W}$ is a diagonal weight matrix with elements $K_t(\frac{t_i-t}{h_t})K_s(\frac{d_E(\boldsymbol{s}_{ij},\boldsymbol{s})}{h_s})$.
- $\mathbf{Y}$ is the observation vector containing all $y(t_i,\boldsymbol{s}_{ij})$ values.

In the minimization problem (2), both $K_t(\cdot)$ and $K_s(\cdot)$ are selected as the Epanechnikov kernel function, defined as

$$K_t(u)=K_s(u)=0.75\big(1-u^2\big)\mathbb{I}\big(|u|\le 1\big).$$

The Epanechnikov kernel is employed here because it possesses certain optimality properties under specific regularity conditions (Epsnečnikov (1969), Qiu (2005)). The two bandwidths, h_t and h_s, are selected using the modified cross-validation (MCV) method proposed by Yang and Qiu (2018), which accounts for spatiotemporal correlation in the data. They can be computed using the function `mod_cv()` in the R package `SpTe2M`.

After obtaining the initial mean estimate $\widehat{\boldsymbol{\mu}}^{(0)}(t,\boldsymbol{s})$, the residuals can be defined as

$$\widehat{\boldsymbol{\epsilon}}(t_i,\boldsymbol{s}_{ij})=y(t_i,\boldsymbol{s}_{ij})-\widehat{\boldsymbol{\mu}}^{(0)}(t_i,\boldsymbol{s}_{ij}),\quad j=1,\dots,m_i,i=1,\dots,n.$$

When $(t,\boldsymbol{s})\neq(t',\boldsymbol{s}')$, the initial covariance function $V(t,t';\boldsymbol{s},\boldsymbol{s}')$ can be estimated using the following initial kernel estimate:

$$\widehat{V}^{(0)}\big(t,t';\boldsymbol{s},\boldsymbol{s}'\big)=\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=1}^{n}\sum_{l=1}^{m_k}\widehat{\boldsymbol{\epsilon}}(t_i,\boldsymbol{s}_{ij})\widehat{\boldsymbol{\epsilon}}(t_k,\boldsymbol{s}_{kl})w_v(i,j,k,l;t,t',\boldsymbol{s},\boldsymbol{s}')}{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=1}^{n}\sum_{l=1}^{m_k}w_v(i,j,k,l;t,t',\boldsymbol{s},\boldsymbol{s}')},\tag{4}$$

where

$$w_v\big(i,j,k,l;t,t',\boldsymbol{s},\boldsymbol{s}'\big)=K_t\bigg(\frac{t_i-t}{g_t}\bigg)K_s\bigg(\frac{d_E(\boldsymbol{s}_{ij},\boldsymbol{s})}{g_s}\bigg)K_t\bigg(\frac{t_k-t'}{g_t}\bigg)K_s\bigg(\frac{d_E(\boldsymbol{s}_{kl},\boldsymbol{s}')}{g_s}\bigg),$$

g_t and g_s are two bandwidths which may differ from h_t and h_s used in (2) and can be determined by minimizing the mean squared prediction error (MSPE), as discussed in Yang and Qiu (2019). They can be computed using the function `cv_mspe()` in the R package `SpTe2M`. When $(t,\boldsymbol{s})=(t',\boldsymbol{s}')$, $V(t,t;\boldsymbol{s},\boldsymbol{s})$ represents the variance of $y(t,\boldsymbol{s})$, denoted as $\sigma^2(t,\boldsymbol{s})$. This variance can be estimated using the following initial kernel estimate:

$$\widehat{\sigma^2}^{(0)}(t,\boldsymbol{s})=\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widehat{\boldsymbol{\epsilon}}^2(t_i,\boldsymbol{s}_{ij})w_\sigma(i,j;t,\boldsymbol{s})}{\sum_{i=1}^{n}\sum_{j=1}^{m_i}w_\sigma(i,j;t,\boldsymbol{s})},\tag{5}$$

where

$$w_\sigma(i,j;t,\boldsymbol{s}) = K_t\Big(\frac{t_i - t}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}_{ij},\boldsymbol{s})}{g_s}\Big).$$

The estimates $\widehat{\boldsymbol{\mu}}^{(0)}(t,\boldsymbol{s})$, $\widehat{V}^{(0)}(t,t';\boldsymbol{s},\boldsymbol{s}')$, and $\widehat{\sigma^2}^{(0)}(t,\boldsymbol{s})$, defined in (3)–(5), characterize the initially estimated regular spatiotemporal pattern of the observed disease incidence rates. These estimates are statistically consistent under certain regularity conditions (Yang and Qiu (2018, 2019)). It should be noted that the mean and covariance functions, $\boldsymbol{\mu}(t,\boldsymbol{s})$ and $V(t,t';\boldsymbol{s},\boldsymbol{s}')$, in Model (1) are not assumed to be space–time separable, even though separate kernel functions in time and space are employed in their local kernel estimation for simplicity. In expressions (2)–(5), the tensor product of temporal and spatial kernel functions, $K_t(\cdot)K_s(\cdot)$, can be readily replaced by a joint spatiotemporal kernel function, $K_{t,s}(\cdot,\cdot)$. However, such a replacement would substantially increase computational complexity, while our numerical experience indicates that it yields little improvement in estimation accuracy.

2.2. *Data preprocessing and online monitoring*. In this part we discuss online monitoring of the observed disease incidence rates in the time interval (T,∞). To this end, the estimated regular spatiotemporal pattern should first be extended periodically from $[0,T]$ to $[0,\infty)$ to account for seasonality. For simplicity, it is assumed that T is the length of one period in this paper. Assume that the current observation time is $t_i^* \in (T,\infty)$, for $i \geq 1$, and the observed spatiotemporal data at time t_i^* are $\{y(t_i^*,\boldsymbol{s}_{ij}^*),\ j=1,\ldots,m_i^*\}$, where $\{\boldsymbol{s}_{ij}^* \in \Omega,\ j=1,\ldots,m_i^*\}$ are m_i^* observation locations. Here "*" is used in some notations to distinguish them from those in the initial IC data.

Because the observations $\{y(t_i^*,\boldsymbol{s}_{ij}^*),\ j=1,\ldots,m_i^*,\ i\geq 1\}$ are spatiotemporally correlated and its IC distribution changes over time, the conventional SPC charts would not be reliable to monitor such data directly (Apley and Tsung (2002), Qiu, Li and Li (2020), Yang and Qiu (2020)). To address this issue, we can preprocess the observed data using the following data standardization and decorrelation procedure:

- At the first time point t_1^*, define the standardized and decorrelated data to be

$$\widehat{\boldsymbol{e}}(t_1^*) = \widehat{\boldsymbol{\Sigma}}_1^{-1/2}\widehat{\boldsymbol{\epsilon}}(t_1^*),$$

  where $\widehat{\boldsymbol{\epsilon}}(t_1^*) = \boldsymbol{y}(t_1^*) - \widehat{\boldsymbol{\mu}}^{(0)}(t_1^*)$, $\boldsymbol{y}(t_1^*) = (y(t_1^*,\boldsymbol{s}_{11}^*),\ldots,y(t_1^*,\boldsymbol{s}_{1m_1^*}^*))^{\mathsf{T}}$, $\widehat{\boldsymbol{\mu}}^{(0)}(t_1^*) = (\widehat{\mu}^{(0)}(t_1^*,\boldsymbol{s}_{11}^*),\ldots,\widehat{\mu}^{(0)}(t_1^*,\boldsymbol{s}_{1m_1^*}^*))^{\mathsf{T}}$, and $\widehat{\boldsymbol{\Sigma}}_1$ is the estimate of the covariance matrix $\boldsymbol{\Sigma}_1 = \mathrm{Var}(\boldsymbol{y}(t_1^*))$ that can be computed from $\widehat{V}^{(0)}(t,t';\boldsymbol{s},\boldsymbol{s}')$.
- At time t_i^*, for $i\geq 2$, let $\boldsymbol{y}(t_i^*) = (y(t_i^*,\boldsymbol{s}_{i1}^*),\ldots,y(t_i^*,\boldsymbol{s}_{im_i^*}^*))^{\mathsf{T}}$ be the vector of all observations at t_i^*, and $\mathbf{Y}_{i-1} = (\boldsymbol{y}^{\mathsf{T}}(t_1^*),\ldots,\boldsymbol{y}^{\mathsf{T}}(t_{i-1}^*))^{\mathsf{T}}$ be the vector of all previous observations. Then the vector of residuals at time t_i^* can be defined to be $\widehat{\boldsymbol{\epsilon}}(t_i^*) = \boldsymbol{y}(t_i^*) - \widehat{\boldsymbol{\mu}}^{(i-1)}(t_i^*)$, where $\widehat{\boldsymbol{\mu}}^{(i-1)}(\cdot)$ denotes the estimated mean vector using data up to the previous time t_{i-1}^*, which is defined in Section 2.3 and below. Similarly, $\widehat{\boldsymbol{\epsilon}}_{i-1}$ can be defined to be the long vector of all residuals at previous observation times. Then the standardized and decorrelated data at t_i^* are defined to be

$$\widehat{\boldsymbol{e}}(t_i^*) = \widehat{\boldsymbol{\Sigma}}_{ii\cdot i-1}^{-1/2}\big[\widehat{\boldsymbol{\epsilon}}(t_i^*) - \widehat{\boldsymbol{\Sigma}}_{i-1,i}^{\mathsf{T}}\widehat{\boldsymbol{\Sigma}}_{i-1,i-1}^{-1}\widehat{\boldsymbol{\epsilon}}_{i-1}\big],$$

  where $\widehat{\boldsymbol{\Sigma}}_{ii\cdot i-1} = \widehat{\boldsymbol{\Sigma}}_i - \widehat{\boldsymbol{\Sigma}}_{i-1,i}^{\mathsf{T}}\widehat{\boldsymbol{\Sigma}}_{i-1,i-1}^{-1}\widehat{\boldsymbol{\Sigma}}_{i-1,i}$. $\widehat{\boldsymbol{\Sigma}}_i$, $\widehat{\boldsymbol{\Sigma}}_{i-1,i}$, and $\widehat{\boldsymbol{\Sigma}}_{i-1,i-1}$ are estimates of $\mathrm{Var}(\boldsymbol{y}(t_i^*))$, $\mathrm{Cov}(\mathbf{Y}_{i-1},\boldsymbol{y}(t_i^*))$, $\mathrm{Cov}(\mathbf{Y}_{i-1},\mathbf{Y}_{i-1})$, respectively, that can be computed from $\widehat{V}^{(i-1)}(t,t';\boldsymbol{s},\boldsymbol{s}')$, which is the covariance function estimate obtained at time t_{i-1}^* that is defined in Section 2.3. To reduce the computational burden, the inverse matrix $\widehat{\boldsymbol{\Sigma}}_{i-1,i-1}^{-1}$

can be computed recursively across different values of i, as discussed in Remark 1 of Yang and Qiu (2020). Furthermore, to enhance computational efficiency, it is often reasonable in practice to assume that the serial correlation in the observed data is short-ranged, as described at the end of Section 2.3. The computational complexity of the proposed method is also discussed there.

After the data standardization and decorrelation, the transformed data become $\{\widehat{\boldsymbol{e}}(t_i^*) = (\widehat{e}(t_i^*, \boldsymbol{s}_{i1}^*), \ldots, \widehat{e}(t_i^*, \boldsymbol{s}_{im_i^*}^*))^{\mathsf{T}}, i \geq 1\}$. In cases when the spatial process under monitoring is IC, then they would be asymptotically uncorrelated with each other, and each of them would have the asymptotic mean of $\mathbf{0}$ and the asymptotic variance of the identity matrix. Then the following CUSUM charting statistic can be defined for detecting disease outbreaks:

$$C_i = \max\left\{0, C_{i-1} + \frac{\sum_{j=1}^{m_i^*} \widehat{e}(t_i^*, s_{ij}^*)^2 - m_i^*}{\sqrt{2m_i^*}} - k\right\}, \quad i \geq 1, \tag{6}$$

where $C_0 = 0$ and $k > 0$ is an allowance constant. The chart gives a signal at time t_i^* when $C_i > L$, where $L > 0$ is a control limit.

Performance of a control chart is typically evaluated using the IC average run length (ARL_0) and out-of-control (OC) average run length (ARL_1), where ARL_0 is defined to be the average number of observation times from the beginning of process monitoring to the signal time when the process under monitoring is IC, and ARL_1 is the average number of observation times from the occurrence of a process shift (i.e., an abrupt change in the process distribution) to the signal time after the process becomes OC. In applications, ARL_0 is usually prespecified, together with the allowance constant k. Then the control limit L can be determined so that the prespecified ARL_0 value is reached. For the CUSUM chart (6), its control limit L can be first determined using a block bootstrap procedure, described below from the initial IC data, and occasionally updated using the updated IC data, as described in Section 2.3:

(i) Standardize and decorrelate the initial IC data as follows. At any t_i for $i = 1, \ldots, n$, compute the residuals $\widehat{\boldsymbol{\epsilon}}(t_i) = \boldsymbol{y}(t_i) - \widehat{\boldsymbol{\mu}}^{(0)}(t_i)$, where $\boldsymbol{y}(t_i) = (y(t_i, \boldsymbol{s}_{i1}), \ldots, y(t_i, \boldsymbol{s}_{im_i}))^T$, and $\widehat{\boldsymbol{\mu}}^{(0)}(t_i) = (\widehat{\mu}^{(0)}(t_i, \boldsymbol{s}_{i1}), \ldots, \widehat{\mu}^{(0)}(t_1, \boldsymbol{s}_{im_i}))^T$. Then the data standardization and decorrelation procedure can be applied to the residuals, and the resulting transformed data are denoted as $\{\widehat{\boldsymbol{e}}(t_i), i = 1, \ldots, n\}$. These data are used to form $n - \tau + 1$ overlapping blocks $\{\mathcal{B}_j = \{\widehat{\boldsymbol{e}}(t_i), i = j, j+1, \ldots, j+\tau-1\}, j = 1, \ldots, n-\tau+1\}$, where τ is the block size.

(ii) Sample with replacement from the set of blocks $\{\mathcal{B}_j\}$. The selected blocks are combined sequentially to form the bth bootstrap sample of process observations for online monitoring, denoted as $\{\widehat{\boldsymbol{e}}_i^{(b)}, i \geq 1\}$. Repeat this block bootstrap sampling process for a total of B times.

(iii) For a given value of control limit L, apply the CUSUM chart (6) to each bootstrap sample of process observations for online process monitoring and record the run length (RL) value that is defined to be the number of observation times from the beginning of process monitoring to the first signal of the CUSUM chart. The average of the B values of RL can be used to estimate the ARL_0 value, denoted as $\text{ARL}_0(L)$. Then a bisection or other numerical algorithms can be used to find the value of L such that $\text{ARL}_0(L)$ reaches the prespecified ARL_0 value.

2.3. *Sequential update of IC parameter estimates.* Under the TSL framework, in the case when the CUSUM chart (6) does not signal at the current time t_i^*, the observed data at t_i^*, $\boldsymbol{y}(t_i^*)$, should be combined with the IC data, and the estimated regular spatiotemporal pattern described by the estimates of $\mu(t, \boldsymbol{s})$, $V(t, t'; \boldsymbol{s}, \boldsymbol{s}')$, and $\sigma^2(t, \boldsymbol{s})$ in the time interval $[0, T]$

should be updated using the combined IC data. To this end, because the regular spatiotemporal pattern is assumed to be periodic with the period time T (cf. discussions in the first paragraph of Section 2.2), each observation time t_i^* should be projected into the time interval $[0, T]$ using the periodic property of the regular spatiotemporal pattern as follows. Let $\widetilde{t}_i$ be the time in $[0, T]$ such that $t_i^* = \widetilde{t}_i + \ell_i T$ where $\ell_i \geq 0$ is an integer. Then the projection of t_i^* is defined to be $\widetilde{t}_i$, for each $i \geq 1$. Because the sequential update of the related estimates needs to be implemented at each observation time when the chart does not signal, computation could be heavy if all updated estimates are computed from scratch. To reduce computational burden, we derive recursive formulas for the updates in this subsection, which are described below.

Recursive update of the estimate of the IC mean function $\mu(t, \boldsymbol{s})$: Assume that the estimate of $\mu(t, \boldsymbol{s})$ at the previous time point t_{i-1}^* is $\widehat{\mu}^{(i-1)}(t, \boldsymbol{s})$. After the underlying spatial process is delared IC at the current time t_i^* and the observed data $\boldsymbol{y}(t_i^*)$ is combined with the IC dataset, we would like to compute the updated estimate of $\mu(t, \boldsymbol{s})$, denoted as $\widehat{\mu}^{(i)}(t, \boldsymbol{s})$, using the combined IC data $\{\boldsymbol{y}(t_l^*), -n + 1 \leq l \leq i\}$ where $\{\boldsymbol{y}(t_l^*), -n + 1 \leq l \leq 0\}$ is the initial IC dataset discussed in Section 2.1. For simplicity of notation, we denote by m_i the number of observation locations for all i in this section. Thus, m_i corresponds to m_i^* when $i > 0$.

From (3) we have

$$\widehat{\mu}^{(i)}(t, \boldsymbol{s}) = \zeta^\top \big[(\boldsymbol{X}^{(i)})^\top \boldsymbol{W}^{(i)} \boldsymbol{X}^{(i)}\big]^{-1} (\boldsymbol{X}^{(i)})^\top \boldsymbol{W}^{(i)} \boldsymbol{Y}^{(i)}, \tag{7}$$

where $\boldsymbol{X}^{(i)}$, $\boldsymbol{W}^{(i)}$ and $\boldsymbol{Y}^{(i)}$ are defined similarly to those in (3), using the combined IC data $\{\boldsymbol{y}(t_l^*), -n + 1 \leq l \leq i\}$. It is obvious that

$$\boldsymbol{X}^{(i)} = \begin{bmatrix} \boldsymbol{X}^{(i-1)} \\ \boldsymbol{X}(t_i^*) \end{bmatrix}, \qquad \boldsymbol{Y}^{(i)} = \begin{bmatrix} \boldsymbol{Y}^{(i-1)} \\ \boldsymbol{y}(t_i^*) \end{bmatrix}, \qquad \boldsymbol{W}^{(i)} = \begin{bmatrix} \boldsymbol{W}^{(i-1)} & 0 \\ 0 & \boldsymbol{W}(t_i^*) \end{bmatrix},$$

where $\boldsymbol{X}(t_i^*)$ is an $m_i \times 4$ matrix with the rows $\{(1, \widetilde{t}_i - t, s_{x,ij} - s_x, s_{y,ij} - s_y), j = 1, \ldots, m_i^*\}$ and $\boldsymbol{W}(t_i^*)$ is an $m_i \times m_i$ diagonal weight matrix with the elements $\{K_t((\widetilde{t}_i - t)/h_t) K_s(d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})/h_s), j = 1, \ldots, m_i^*\}$.

From (7) we have

$$\begin{aligned} \widehat{\mu}^{(i)}(t, \boldsymbol{s}) = \zeta^\top \big[&(\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{X}^{(i-1)} + (\boldsymbol{X}(t_i^*))^\top \boldsymbol{W}(t_i^*) \boldsymbol{X}(t_i^*)\big]^{-1} \\ &\times \big[(\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{Y}^{(i-1)} + (\boldsymbol{X}(t_i^*))^\top \boldsymbol{W}(t_i^*) \boldsymbol{y}(t_i^*)\big]. \end{aligned} \tag{8}$$

In Expression (9) the quantities $(\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{X}^{(i-1)}$ and $(\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{Y}^{(i-1)}$ have been computed at the previous time t_{i-1}^*. Thus, the updated estimate $\widehat{\mu}^{(i)}(t, \boldsymbol{s})$ can be computed by (7) and the following recursive formulas:

$$\begin{aligned} (\boldsymbol{X}^{(i)})^\top \boldsymbol{W}^{(i)} \boldsymbol{X}^{(i)} &= (\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{X}^{(i-1)} + (\boldsymbol{X}(t_i^*))^\top \boldsymbol{W}(t_i^*) \boldsymbol{X}(t_i^*), \\ (\boldsymbol{X}^{(i)})^\top \boldsymbol{W}^{(i)} \boldsymbol{Y}^{(i)} &= (\boldsymbol{X}^{(i-1)})^\top \boldsymbol{W}^{(i-1)} \boldsymbol{Y}^{(i-1)} + (\boldsymbol{X}(t_i^*))^\top \boldsymbol{W}(t_i^*) \boldsymbol{y}(t_i^*). \end{aligned}$$

Under some regularity conditions, the recursively updated estimate $\widehat{\mu}^{(i)}(t, \boldsymbol{s})$ has the statistical properties described in the following theorem.

THEOREM 2.1. *In Model (1), assume that the error term $\epsilon(t, \boldsymbol{s})$ has a bounded second moment at any $(t, \boldsymbol{s}) \in [0, T] \times \Omega$, and the mean function $\mu(t, \boldsymbol{s})$ is twice continuously differentiable with respect to both t and $\boldsymbol{s}$ within a neighborhood of $(t, \boldsymbol{s})$. For any open neighborhood U of $(t, \mathbf{s})$, the number of points $\{(\widetilde{t}_l, \mathbf{s}_{lj}^*) \in U, 1 \leq j \leq m_l^*, 1 \leq l \leq i\}$ is assumed to tend to infinity when i increases. At each time t_l^*, the spatial locations $\{\mathbf{s}_{lj}^*, j = 1, \ldots, m_l^*\}$ are assumed to be independent of the random errors $\{\epsilon(t_l^*, \mathbf{s}_{lj}^*), j = 1, \ldots, m_l^*\}$ and follow the*

density function $f(\mathbf{s})$ on Ω, which is bounded away from zero in any subregion of Ω where f is positive. The error terms $\{\epsilon(t_l^*, \mathbf{s}_{lj})\}$ are assumed to satisfy a strong mixing condition in the time domain, with the mixing coefficient

$$\alpha(q) = \sup_{i \geq 1, 1 \leq l \leq i-q} \sup_{A \in \mathcal{F}_1^l, B \in \mathcal{F}_{l+q}^i} |P(A \cap B) - P(A)P(B)|,$$

where $\mathcal{F}_{q_0}^{q_1}$ is the σ-algebra generated by $\{\epsilon(t_l^*, \mathbf{s}_{lj}), q_0 \leq l \leq q_1, j = 1, \ldots, m_l\}$. For this coefficient it is assumed that there exist constants $C_0, C_1 > 0$ such that $\alpha(q) \leq C_0 \exp(-C_1 q)$. Then, when $h_t = o(1)$, $h_s = o(1)$, and $1/(h_t h_s^2 N_i) = o(1)$, where $N_i = \sum_{l=-n+1}^{i} m_l^*$, we have

$$|\widehat{\mu}^{(i)}(t, \mathbf{s}) - \mu(t, \mathbf{s})| = O_p\left(h_t^2 + h_s^2 + \frac{1}{\sqrt{h_t h_s^2 N_i}}\right) \quad \text{for any } (t, \boldsymbol{s}) \in (0, T) \times \Omega.$$

Recursive update of the IC covariance function $V(t, t'; \boldsymbol{s}, \boldsymbol{s}')$: Let the estimated covariance function at the previous time t_{i-1}^* be denoted as $\widehat{V}^{(i-1)}(t, t'; \boldsymbol{s}, \boldsymbol{s}')$. After the mean estimate being updated to be $\widehat{\mu}^{(i)}(t, \boldsymbol{s})$, the residuals at t_i^* become $\widehat{\boldsymbol{\epsilon}}(t_i^*) = \boldsymbol{y}(t_i^*) - \widehat{\boldsymbol{\mu}}^{(i)}(t_i^*)$. Using the combined IC data $\{\boldsymbol{y}(t_l^*), -n+1 \leq l \leq i\}$ and Expression (4), the updated estimate of the covariance function at t_i^* has the expression

$$\widehat{V}^{(i)}(t, t'; \boldsymbol{s}, \boldsymbol{s}') = \frac{\sum_{l=-n+1}^{i} \sum_{j=1}^{m_l} \sum_{k=-n+1}^{i} \sum_{q=1}^{m_k} \widehat{\epsilon}(t_l^*, \boldsymbol{s}_{lj}^*) \widehat{\epsilon}(t_k^*, \boldsymbol{s}_{kq}^*) \widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}')}{\sum_{l=-n+1}^{i} \sum_{j=1}^{m_l} \sum_{k=-n+1}^{i} \sum_{q=1}^{m_k} \widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}')}, \tag{9}$$

where $\widetilde{w}_v(i, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}') =$
$K_t((\widetilde{t}_i - t)/g_t) K_s(d_E(\boldsymbol{s}_{ij}, \boldsymbol{s})/g_s) K_t((\widetilde{t}_k - t')/g_t) K_s(d_E(\boldsymbol{s}_{kq}, \boldsymbol{s}')/g_s)$.

From (9), $\widehat{V}^{(i)}(t, t'; \boldsymbol{s}, \boldsymbol{s}')$ can be computed recursively from $\widehat{V}^{(i-1)}(t, t'; s, s')$, as shown in the following formula:

$$\begin{aligned}
\widehat{V}^{(i)}(t, t'; \boldsymbol{s}, \boldsymbol{s}') = & \Bigg[\sum_{l=-n+1}^{i-1} \sum_{j=1}^{m_l} \sum_{k=-n+1}^{i-1} \sum_{q=1}^{m_k} \widehat{\epsilon}(t_l^*, \boldsymbol{s}_{lj}^*) \widehat{\epsilon}(t_k^*, \boldsymbol{s}_{kq}^*) \widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \\
& + 2 \sum_{l=-n+1}^{i-1} \sum_{j=1}^{m_l} \sum_{q=1}^{i} \widehat{\epsilon}(t_l^*, \boldsymbol{s}_{lj}^*) \widehat{\epsilon}(t_i^*, \boldsymbol{s}_{iq}^*) \widetilde{w}_v(l, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \\
& + \sum_{j=1}^{i} \sum_{q=1}^{i} \widehat{\epsilon}(t_i^*, \boldsymbol{s}_{ij}^*) \widehat{\epsilon}(t_i^*, \boldsymbol{s}_{iq}^*) \widetilde{w}_v(i, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \Bigg] \\
& \Big/ \Bigg[\sum_{l=-n+1}^{i-1} \sum_{j=1}^{m_l} \sum_{k=-n+1}^{i-1} \sum_{q=1}^{m_k} \widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \\
& + 2 \sum_{l=-n+1}^{i-1} \sum_{j=1}^{m_l} \sum_{q=1}^{i} \widetilde{w}_v(l, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \\
& + \sum_{j=1}^{i} \sum_{q=1}^{i} \widetilde{w}_v(i, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') \Bigg],
\end{aligned} \tag{10}$$

where

$$\widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}') = K_t\left(\frac{\widetilde{t}_l - t}{g_t}\right) K_s\left(\frac{d_E(\boldsymbol{s}_{lj}^*, \boldsymbol{s})}{g_s}\right)$$

$$\times K_t\Big(\frac{\widetilde{t}_k - t'}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}^*_{kq}, \boldsymbol{s}')}{g_s}\Big),$$

$$\widetilde{w}_v(l, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') = K_t\Big(\frac{\widetilde{t}_l - t}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}^*_{lj}, \boldsymbol{s})}{g_s}\Big)$$
$$\times K_t\Big(\frac{\widetilde{t}_i - t'}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}^*_{iq}, \boldsymbol{s}')}{g_s}\Big),$$

$$\widetilde{w}_v(i, j, i, q; t, t', \boldsymbol{s}, \boldsymbol{s}') = K_t\Big(\frac{\widetilde{t}_i - t}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}^*_{ij}, \boldsymbol{s})}{g_s}\Big)$$
$$\times K_t\Big(\frac{\widetilde{t}_i - t'}{g_t}\Big)K_s\Big(\frac{d_E(\boldsymbol{s}^*_{iq}, \boldsymbol{s}')}{g_s}\Big),$$

and the symbol "$\sim$" is used in the notation of these quantities because the projection $\widetilde{t}_i$ of t^*_i (see the related discussion in the first paragraph of Section 2.3) is used in their computation.

The quantities $\sum_{l=-n+1}^{i-1}\sum_{j=1}^{m_l}\sum_{k=-n+1}^{i-1}\sum_{q=1}^{m_k}\widehat{\epsilon}(t^*_l, \boldsymbol{s}^*_{lj})\widehat{\epsilon}(t^*_k, \boldsymbol{s}^*_{kq})\widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}')$ and $\sum_{l=-n+1}^{i}\sum_{j=1}^{m_i}\sum_{k=-n+1}^{i}\sum_{q=1}^{m_k}\widetilde{w}_v(l, j, k, q; t, t', \boldsymbol{s}, \boldsymbol{s}')$ in (10) have been computed at the previous time t^*_{i-1} when computing $\widehat{V}^{(i-1)}(t, t'; \boldsymbol{s}, \boldsymbol{s}')$.

Recursive update of the IC variance function $\sigma^2(t, \boldsymbol{s})$: The estimated variance function at the previous time t^*_{i-1} is denoted as $\widehat{\sigma^2}^{(i-1)}(t, \boldsymbol{s})$. After the mean function estimate is updated to $\widehat{\mu}^{(i)}(t, \boldsymbol{s})$, the estimate $\widehat{\sigma^2}^{(i)}(t, \boldsymbol{s})$ can be updated recursively from $\widehat{\sigma^2}^{(i-1)}(t, \boldsymbol{s})$ as shown in the following formula:

$$\widehat{\sigma^2}^{(i)}(t, \boldsymbol{s}) = \frac{\sum_{l=-n+1}^{i}\sum_{j=1}^{m_l}\widehat{\epsilon}^2(t^*_l, \boldsymbol{s}^*_{lj})\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})}{\sum_{l=-n+1}^{i}\sum_{j=1}^{m_l}\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})}$$
$$= \frac{(\sum_{l=-n+1}^{i-1}\sum_{j=1}^{m_l}\widehat{\epsilon}^2(t^*_l, \boldsymbol{s}^*_{lj})\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})) + (\sum_{j=1}^{m_i}\widehat{\epsilon}^2(t^*_i, \boldsymbol{s}^*_{ij})\widetilde{w}_\sigma(i, j; t, \boldsymbol{s}))}{(\sum_{l=-n+1}^{i-1}\sum_{j=1}^{m_l}\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})) + (\sum_{j=1}^{m_i}\widetilde{w}_\sigma(i, j; t, \boldsymbol{s}))}, \tag{11}$$

where $\widetilde{w}_\sigma(l, j; t, \boldsymbol{s}) = K_t((\widetilde{t}_l - t)/g_t)K_s(d_E(\boldsymbol{s}^*_{lj}, \boldsymbol{s})/g_s)$, and $\widetilde{w}_\sigma(i, j; t, \boldsymbol{s}) = K_t((\widetilde{t}_i - t)/g_t) \times K_s(d_E(\boldsymbol{s}^*_{ij}, \boldsymbol{s})/g_s)$.

In (12) the quantities $\sum_{l=-n+1}^{i-1}\sum_{j=1}^{m_l}\widehat{\epsilon}^2(t^*_l, \boldsymbol{s}^*_{lj})\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})$ and $\sum_{l=-n+1}^{i-1}\sum_{j=1}^{m_l}\widetilde{w}_\sigma(l, j; t, \boldsymbol{s})$ have been computed at time t^*_{i-1} when computing $\widehat{\sigma^2}^{(i-1)}(t, \boldsymbol{s})$.

THEOREM 2.2. *Besides the assumptions in Theorem 2.1, it is further assumed that* $\log(i)/(h_t h_s^2 N_i) = o(1)$, $h_t/h_s = o(1)$, $g_t = o(1)$, $g_s = o(1)$, *and* $\log(i)/(g_t g_s^2 N_i) = o(1)$. *In addition, the variance function $\sigma^2(t, \boldsymbol{s})$ and the covariance function $V(t, t'; \boldsymbol{s}, \boldsymbol{s}')$ are both assumed to be twice continuously differentiable. Then for any $(t, \boldsymbol{s}) \in (0, T) \times \Omega$, we have*

$$|\widehat{\sigma^2}^{(i)}(t, \boldsymbol{s}) - \sigma^2(t, \boldsymbol{s})| = O_p\Big(h_t^2 + h_s^2 + g_t^2 + g_s^2 + \frac{\log(i)}{\sqrt{h_t h_s^2 N_i}} + \frac{1}{\sqrt{g_t g_s^2 N_i}}\Big),$$

$$|\widehat{V}^{(i)}(t, t'; \boldsymbol{s}, \boldsymbol{s}') - V(t, t'; \boldsymbol{s}, \boldsymbol{s}')| = O_p\Big(h_t^2 + h_s^2 + g_t^2 + g_s^2 + \frac{\log(i)}{\sqrt{h_t h_s^2 N_i}} + \frac{1}{\sqrt{g_t g_s^2 N_i}}\Big).$$

The proofs of Theorems 2.1 and 2.2 are outlined in the Supplementary Material (Zhou and Qiu (2026)).

Update of the control limit L: In theory, after the estimates of the IC functions $\mu(t,\boldsymbol{s})$, $V(t,t';\boldsymbol{s},\boldsymbol{s}')$ and $\sigma^2(t,\boldsymbol{s})$ get updated at the current time t_i^*, the control limit L of the CUSUM chart (6) should also be calibrated to maintain the prespecified ARL_0 value. To this end, the block bootstrap procedure described in Section 2.2 can be applied to the updated IC data $\{\boldsymbol{y}(t_l^*), -n+1 \le l \le i\}$, after the initial estimates of $\mu(t,\boldsymbol{s})$, $V(t,t';\boldsymbol{s},\boldsymbol{s}')$ and $\sigma^2(t,\boldsymbol{s})$ are replaced by the updated estimates $\widehat{\mu}^{(i)}(t,\boldsymbol{s})$, $\widehat{V}^{(i)}(t,t';\boldsymbol{s},\boldsymbol{s}')$ and $\widehat{\sigma^2}^{(i)}(t,\boldsymbol{s})$. However, the computational burden to calibrate L at each time point is quite heavy. In Section 3.1 below, numerical studies have been performed to confirm that the value of L can be calibrated every d observation times without sacrificing much of the IC performance of the CUSUM chart, where d is an integer in the interval $[1, 25]$.

Short-range serial correlation: In practice, it is often reasonable to assume that the serial correlation over time is short-ranged in the sense that $V(t,t';\boldsymbol{s},\boldsymbol{s}') \approx 0$ when $|t-t'| \ge b_{\max}$, where $b_{\max} > 0$ is a parameter specifying the range of serial correlation (Qiu and Xie (2022)). In such cases, computation involved in many quantities can be simplified. For instance, the initial estimate of the IC covariance function, $\widehat{V}^{(0)}(t,t';\mathbf{s},\mathbf{s}')$, can be computed as follows:

$$\widehat{V}^{(0)}(t,t';\mathbf{s},\mathbf{s}') = \frac{\sum_{|t_i-t|\le b_{\max}}\sum_{j=1}^{m_i}\sum_{|t_k-t'|\le b_{\max}}\sum_{l=1}^{m_k}\widehat{\epsilon}(t_i,\mathbf{s}_{ij})\widehat{\epsilon}(t_k,\mathbf{s}_{kl})w_\sigma(i,j,k,l;t,t',\mathbf{s},\mathbf{s}')}{\sum_{|t_i-t|\le b_{\max}}\sum_{j=1}^{m_i}\sum_{|t_k-t'|\le b_{\max}}\sum_{l=1}^{m_k} w_\sigma(i,j,k,l;t,t',\mathbf{s},\mathbf{s}')}.$$

The data decorrelation procedure discussed earlier can also be simplified. Specifically, the observed data at the current time point t_i^* only need to be decorrelated from the data observed at the previous $b_{\max}$ time points, which can substantially reduce the computational burden.

Computational complexity: From Expressions (7)–(12), the computational complexity for updating the estimated regular spatiotemporal pattern at each observation time t_i^*, represented by $\widehat{\mu}^{(i)}(t,\mathbf{s})$, $\widehat{V}^{(i)}(t,t';\mathbf{s},\mathbf{s}')$, and $\widehat{\sigma^2}^{(i)}(t,\mathbf{s})$, is $O(m_i^2)$. In most disease surveillance applications, the number of observation locations is on the order of tens or hundreds. Therefore, the computation of the proposed ST-TSL method is quite efficient, as confirmed by the results in Table 4 of Section 4.

3. Simulation studies. In this section we evaluate the numerical performance of our proposed method, denoted as ST-TSL, by performing some simulation studies. In these studies the true IC mean function $\mu(t,\boldsymbol{s})$ in Model (1) is assumed to be

$$\mu(t,\boldsymbol{s}) = 0.01\cos(2\pi t) + 0.01e^{-[(s_x-0.5)^2+(s_y-0.5)^2]} + 0.01, \tag{12}$$

where $\boldsymbol{s} = (s_x, s_y)' \in \Omega = [0,1]\times[0,1]$ is the location, and $t \in [0,T] = [0,1]$ is the time. This function is shown in Figure 3 when $t = 0, 0.25, 0.75$, and 1.0. For simplicity, it is assumed that the observation times $\{t_i = i/n, i = 1,\ldots,n\}$ are equally spaced in $[0,1]$ and the observation locations $\{\boldsymbol{s}_{ij}, j = 1,\ldots,m\}$ remain the same over time and are also equally spaced in $[0,1]\times[0,1]$. In such cases, notations for the observation locations are simplified to $\{\boldsymbol{s}_j, j = 1,\ldots,m\}$.

In Model (1) the random errors $\{\epsilon(t_i,\boldsymbol{s}_j)\}$ are generated from a Gaussian random field using the R function `spatialnoise()` from the *neuRosim* package in which the parameter ρ controls both spatial and temporal data correlation and is set to be 0.1 (relatively weak correlation), 0.3 (medium correlation), or 0.5 (relatively strong correlation), and the variance parameter σ^2 controls the data variability and is set to be 0.01^2.

To compute the ARL_0 value of the chart ST-TSL, an initial IC dataset is first generated from Model (1), as described above. A spatiotemporal data stream is then generated from Model (1) as well for online process monitoring. The run length (RL) value of the chart (i.e.,

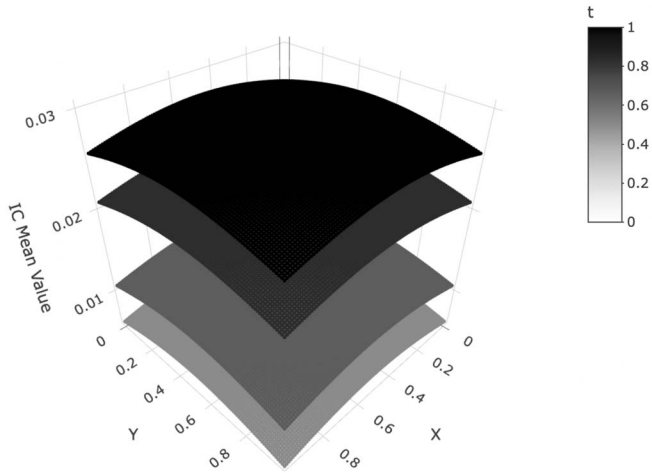

FIG. 3. *IC mean function* $\lambda(t, \boldsymbol{s})$ *when* $t = 0, 0.25, 0.75,$ *and* 1.0.

the number of observation times from the beginning of process monitoring to the signal time) is recorded for each simulation run of online process monitoring. Then the online process monitoring is repeated for 1000 times and the ARL_0 value is estimated by the sample mean of the 1000 RL values. Because this ARL_0 estimate depends on the initial IC data generated before process monitoring, the entire simulation process discussed above, from generation of the initial IC data to computing the ARL_0 estimate, is repeated for 100 times. The average of these 100 ARL_0 values is then used as the final estimate of the actual ARL_0 value of the chart ST-TSL. The computation of the ARL_1 value for the ST–TSL chart follows the same procedure as described above, with the distinction that the spatiotemporal data streams used for online monitoring are generated from a predefined OC model (see Section 3.2).

3.1. *IC performance*. We first study the IC performance of the proposed method ST-TSL in cases when the IC mean function $\mu(t, \boldsymbol{s})$ and the IC covariance function $V(t, \boldsymbol{s}; t', \boldsymbol{s}')$ are both unknown. For a given nominal ARL_0 value, the actual ARL_0 value of ST-TSL should ideally be close to the nominal value across the cases considered to ensure good IC performance. To verify this, we first estimate $\mu(t, \boldsymbol{s})$ and $V(t, \boldsymbol{s}; t', \boldsymbol{s}')$ using an IC dataset. Here the IC data are generated at n equally spaced time points in $[0, 1]$ and m equally spaced positions in the region $[0, 1] \times [0, 1]$. Then the initial control limit L of the CUSUM chart (6) is determined from the initial IC dataset using the block bootstrap procedure with $B = 1000$. For the block size τ, numerical studies have shown $\tau = 5$ to be suitable for both high and low spatiotemporal correlation (Yang and Qiu (2020), Qiu and Yang (2023)), which is adopted here.

It is important to highlight that the frequency of control limit recalibration might impact the IC performance of the control chart ST-TSL, as discussed at the end of Section 2. Recalibrating too often increases computational burden, while not recalibrating at all can cause the control chart's IC performance to drift from the prespecified ARL_0 level. Therefore, we first address this issue using a simulation example, where we set $n = 300$, $m = 49$, $k = 0.1$, $\text{ARL}_0 = 50$, $b_{\max} = 5$, and $\rho = 0.3$, and the control limit L is recalibrated every d observation times with d changing from 1, 5, 10, 25, to 50. The results are presented in Figure 4 from which it can be seen that the chart ST-TSL has a satisfactory IC performance when $d \in [1, 25]$. Therefore, we choose $d = 25$ in all subsequent simulations.

Next, we investigate the IC performance of the proposed chart ST-TSL when the allowance constant k is set to be 0.1, $m = 49$ or 100, $n = 100$, 200, 300, or 400, and the nominal ARL_0

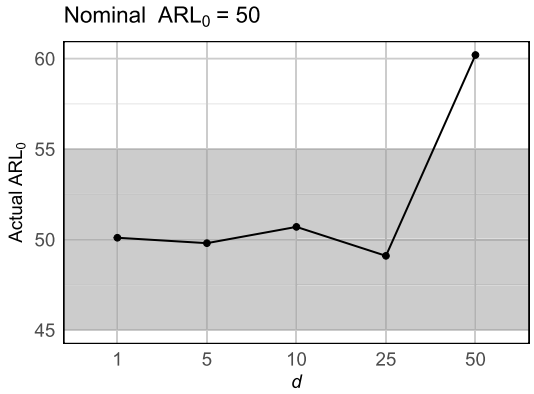

FIG. 4. *Actual* ARL_0 *values of the proposed chart ST-TSL when its control limit is recalibrated every d observation times.*

value of the chart is set to be 50 or 100. The calculated actual ARL_0 values of the chart are presented in Table 1 along with the corresponding standard errors. From the table it can be seen that: (i) as n increases, the actual ARL_0 value becomes closer to the nominal ARL_0 value, and the IC performance seems to stabilize when $n \geq 300$, since the difference bewteen the results for $n = 300$ and $n = 400$ is small, and (ii) as a comparison, the impact of m and ρ on the actual ARL_0 value seems to be quite small.

To check robustness of the proposed method to more complex IC mean structure, additional scenarios with spatially varying seasonal amplitude/phase and anisotropic spatial effects are considered; results are summarized in the Supplementary Material and are consistent with the main findings here.

3.2. *OC performance*. Next, we evaluate the OC performance of the proposed ST-TSL method. To this end, a control chart with a prespecified ARL_0 level is considered more effective in detecting a given process distributional shift if it yields a smaller ARL_1 value. In this simulation study, for simplicity, we consider a mean shift in the process distribution that occurs at the first observation time. Process observations are generated from an OC model identical to Model (1), except that the OC mean function is defined as $\mu(t, \boldsymbol{s}) + \delta(t, \boldsymbol{s})$, where $\mu(t, \boldsymbol{s})$ is the IC mean function and $\delta(t, \boldsymbol{s})$ represents the shift function. The following three shift functions are considered:

Type (i) $\delta_1(t, \boldsymbol{s}) = 0.0025\nu$, for $\nu = 1, 2, 3, 4$;

TABLE 1
Actual ARL_0 *values and their standard errors* (*in parentheses*) *of the proposed chart ST-TSL in different cases considered*

		$\rho = 0.1$		$\rho = 0.3$		$\rho = 0.5$	
n	m	$\mathrm{ARL}_0 = 50$	$\mathrm{ARL}_0 = 100$	$\mathrm{ARL}_0 = 50$	$\mathrm{ARL}_0 = 100$	$\mathrm{ARL}_0 = 50$	$\mathrm{ARL}_0 = 100$
100	49	41.21 (0.49)	80.44 (0.77)	41.48 (0.79)	79.81 (0.42)	41.61 (0.98)	79.28 (0.54)
	100	40.96 (0.86)	80.32 (0.65)	40.90 (0.91)	78.76 (0.92)	42.32 (0.88)	81.23 (0.77)
200	49	50.33 (1.15)	86.63 (1.45)	51.98 (1.13)	86.84 (1.60)	51.06 (1.40)	86.62 (1.95)
	100	49.43 (1.06)	84.67 (1.35)	51.05 (1.18)	87.34 (1.60)	53.03 (1.22)	89.88 (1.75)
300	49	49.61 (1.30)	97.47 (2.01)	49.31 (1.60)	99.67 (2.38)	50.69 (1.55)	101.09 (2.30)
	100	50.04 (1.29)	97.83 (2.09)	49.02 (1.37)	99.72 (2.30)	50.57 (1.40)	102.92 (2.05)
400	49	50.35 (1.43)	99.17 (2.69)	49.63 (1.47)	100.83 (2.63)	50.12 (1.44)	100.20 (2.46)
	100	50.10 (1.52)	100.43 (2.70)	49.89 (1.55)	100.24 (2.73)	50.24 (1.56)	100.40 (2.78)

TABLE 2
Calculated ARL_1 *values of ST-TSL, along with their standard errors* (*in parentheses*), *in various cases when* $k = 0.1$, $\text{ARL}_0 = 50$, $n = 300$, $m = 49$, $\rho = 0.1, 0.3,$ *or* 0.5, *and the shift function takes one of the three types*

Shift Type	ν	$\rho = 0.1$	$\rho = 0.3$	$\rho = 0.5$
(i)	1	10.35 (0.51)	12.76 (0.62)	13.46 (0.71)
	2	6.87 (0.33)	8.99 (0.54)	9.99 (0.69)
	3	5.16 (0.22)	6.60 (0.44)	7.59 (0.58)
	4	1.04 (0.01)	1.05 (0.04)	1.26 (0.09)
(ii)	1	10.29 (0.50)	12.68 (0.62)	13.42 (0.70)
	2	10.26 (0.50)	12.57 (0.62)	13.23 (0.71)
	3	10.25 (0.49)	12.49 (0.61)	13.11 (0.71)
	4	10.15 (0.49)	12.41 (0.61)	13.07 (0.71)
(iii)	1	5.73 (0.26)	7.36 (0.47)	8.72 (0.64)
	2	3.17 (0.16)	4.02 (0.29)	4.81 (0.36)
	3	1.62 (0.09)	2.36 (0.17)	2.90 (0.21)
	4	1.13 (0.04)	1.50 (0.10)	1.90 (0.14)

Type (ii) $\delta_2(t, \boldsymbol{s}) = 0.0025 + 0.01t\nu$, for $\nu = 1, 2, 3, 4$;
Type (iii) $\delta_3(t, \boldsymbol{s}) = 0.0025 + 0.01t + 0.001(s_x^2 + s_y^2)\nu$, for $\nu = 1, 2, 3, 4$.

It is obvious that $\delta_1(t, \boldsymbol{s})$ is a constant over both time and space, $\delta_2(t, \boldsymbol{s})$ changes over time but not space, and $\delta_3(t, \boldsymbol{s})$ changes over both time and space.

To compute the ARL_1 values of ST-TSL, the control limit L is determined in the same way as described in Section 3.1 to achieve a prespecified ARL_0 value. For the cases where $k = 0.1$, $\text{ARL}_0 = 50$, $n = 300$, $m = 49$, $\rho = 0.1$, 0.3, or 0.5 and where the shift function corresponds to one of the three functions discussed above, the computed ARL_1 values of ST-TSL, along with their standard errors, are presented in Table 2. From the table it can be seen that: (i) ARL_1 gets smaller when ν increases from 1 to 4 for all three shift functions, and (ii) ARL_1 gets larger when ρ is larger. These results are all intuitively reasonable.

In the simulation studies above, the bandwidths h_t, h_s, g_t, and g_s were selected using the MCV and MSPE procedures described in Section 2.1. In the Supplementary Material, we also report results obtained with alternative bandwidth choices. These results confirm that both the IC and OC performance of the proposed ST–TSL method improve when the bandwidths are selected using the MCV and MSPE procedures.

3.3. *Comparison with some alternative methods.* In this part we compare the proposed method ST-TSL with four representative existing methods described briefly below:

- Zhao et al. (2011) proposed a disease surveillance method, using kernel smoothing (KS) to estimate the IC mean function and an AR(2) model to describe temporal correlation. This method, denoted as KS, assumed that the observed data were normally distributed and independent across spatial locations.
- Yan, Paynabar and Shi (2018) suggested a real-time monitoring method for spatiotemporal processes based on the spatiotemporal smooth sparse decomposition (SSD). This method, denoted as SSD, can accommodate space/time-varying mean structure and has been shown effective for detecting sparse anomalies. But it assumes that the observed data are independent and normally distributed.
- Chen, Zi and Zou (2016) suggested a nonparametric approach for monitoring location parameters of multivariate processes with a small reference dataset. Its test statistic is constructed based on empirical distributions, and its control limit is determined based on the

observed data. However, as a rank-based method, much information in the observed data would be lost, potentially affecting its sensitivity in detecting anamolies. This method is denoted as RK.

- Yang and Qiu (2020) proposed a CUSUM chart for online monitoring of spatiotemporal data streams, denoted as ST-CUSUM. This method does not rely on any parametric assumptions regarding spatiotemporal data variation, correlation, or distribution, making it flexible for applications. This method estimates the regular spatiotemporal pattern of the observed data based solely on an IC data collected before online process monitoring, without using the observed data collected during online process monitoring.

Comparison of the IC performance. We begin by comparing the IC performance of the five methods described above. The threshold value in the KS method is determined by a parametric bootstrap procedure, as described in Zhao et al. (2011). The control limit of the SSD chart is determined through the simulation-based procedure discussed in Yan, Paynabar and Shi (2018). The weighting parameter in the RK method is set to be 0.2, as suggested in Chen, Zi and Zou (2016). The allowance constant in the ST-CUSUM chart is set to be 0.1, as discussed in Yang and Qiu (2020), and the allowance constant k in the proposed ST-TSL chart is also set to be 0.1. Then the calculated actual ARL_0 values of the five charts in cases when $\rho = 0.3$, $m = 50$ or 100, $n = 100, 200, 300$, or 400, and the nominal ARL_0 values of all charts are set to be 50 or 100 are shown in Figure 5. From the figure it can be seen that: (i) the proposed ST-TSL chart has the best IC performance in all cases considered since its actual ARL_0 values are closest to the nominal values among the five charts, (ii) the ST-CUSUM chart has a reliable IC performance only when n is quite large ($n \geq 300$ when the nominal ARL_0 value is 50, and $n \geq 400$ when the nominal ARL_0 value is 100), and (iii) the actual ARL_0 values of the other three charts are beyond 10% of the nominal ARL_0 values in all cases considered.

Comparison of the OC performance. We then study the OC performance of the five control charts for detecting the shifts considered in Section 3.2 in cases when $n = 300$, $m = 49$, the nominal ARL_0 values of the five charts are all set to be 50, and $\rho = 0.1, 0.3$, or 0.5. To make the comparison fair, the control limit (or the threshold value) of each chart has been adjusted so that its actual ARL_0 value equals the nominal value of 50. Additionally, procedure parameters in each chart (e.g., the allowance constant in ST-TSL and ST-CUSUM) are chosen such that its ARL_1 value reaches the minimum for detecting a given shift. Namely, the optimal ARL_1 values of the five charts are compared here, because it has been pointed out in the literature that the OC performance of different charts with prespecified parameter values may not be comparable otherwise (Qiu, Li and Li (2020)).

The optimal ARL_1 values of the five charts are presented in Figure 6. From the figure it can be seen that: (i) the proposed chart ST-TSL has the best performance in all cases considered, (ii) the chart ST-CUSUM performs well in cases with relatively large shifts (e.g., Shift Type (iii)), and (iii) the other three charts, especially KS, have relatively weak OC performance.

3.4. *A potential improvement to handle masking effect.* A common challenge in self-starting SPC procedures is the masking effect, which arises when OC observations inadvertently contaminate the IC dataset. This contamination can lead to biased estimation of IC parameters and reduce the effectiveness of shift detection (see (Qiu (2014)), Section 4.5). Our proposed method, ST-TSL, may be susceptible to this limitation, as it recursively updates IC parameter estimates during sequential monitoring, especially when the related shift is small (cf. Section 2.3).

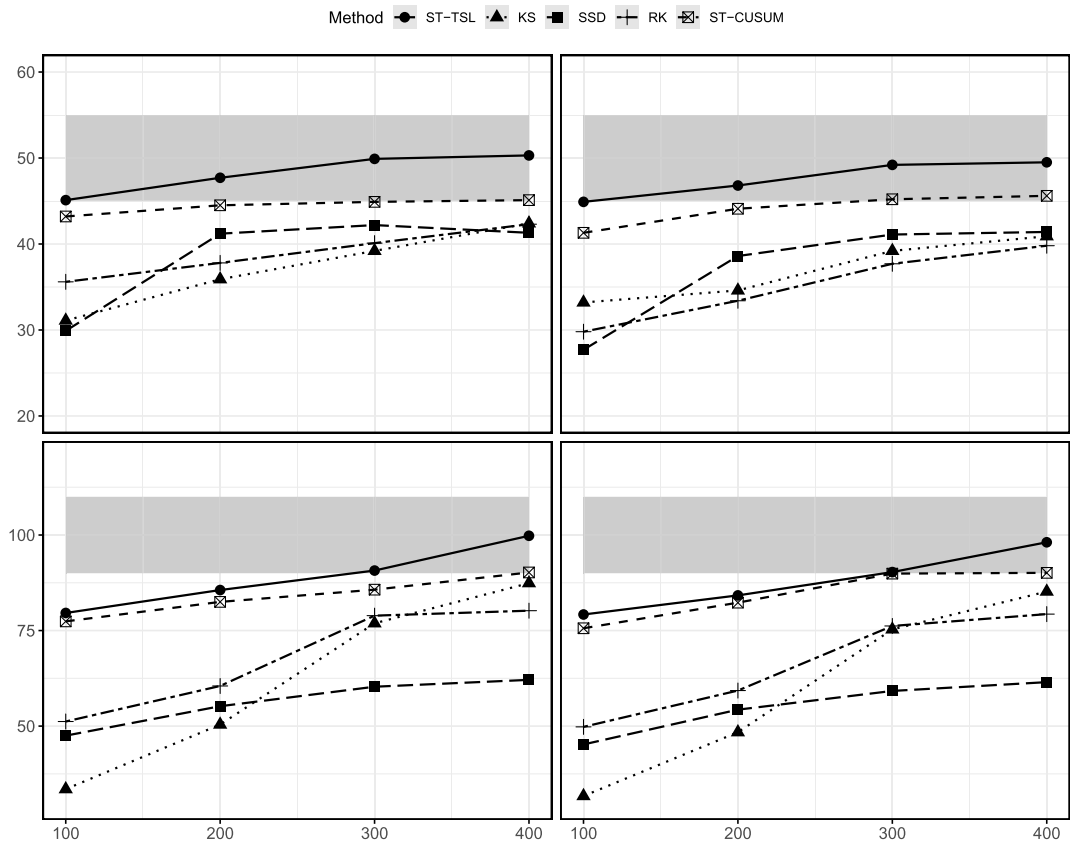

FIG. 5. *Calculated actual* ARL_0 *values of the five charts in cases when* $\rho = 0.3$, $m = 49$ (*first column*) *or* 100 (*second column*), $n = 100, 200, 300,$ *or* 400, *and the nominal* ARL_0 *values of all charts are set to be* 50 (*first row*) *or* 100 (*second row*). *Shaded region in each plot indicates the* ARL_0 *values within* 10% *of the nominal* ARL_0 *value*.

To mitigate this issue, we propose leveraging the natural restart mechanism of the CUSUM procedure (6). Specifically, the CUSUM charting statistic C_i resets to zero at time t_i^* when

$$C_{i-1} + \frac{\sum_{j=1}^{m_i^*} \widehat{e}(t_i^*, s_{ij}^*)^2 - m_i^*}{\sqrt{2m_i^*}} \leq k,$$

indicating that the evidence for a shift in the observed data up to time t_i^* is weak. To take advantage of this property, we propose a modification to the ST-TSL procedure: in the sequential learning algorithm described in Section 2.3, the estimates of the relevant IC quantities are updated at time t_i^* only when the CUSUM statistic restarts at t_i^*. This modification largely avoids the masking effect, as the likelihood that the data observed at t_i^* is OC is minimal when the chart restarts at that time.

To compare the modified ST-TSL procedure with the original version, we consider scenarios where the OC mean function is given by $\mu(t, \mathbf{s}) + \delta_4(t, \mathbf{s})$, where $\mu(t, \mathbf{s})$ is the IC mean function defined in (12), and $\delta_4(t, \mathbf{s})$ is a shift function defined as follows:

Type (iv) $\delta_4(t, s) = 0.005 \times 0.3^{\nu-1}$, for $\nu = 1, 2, 3, 4$.

As ν increases from 1 to 4, the shift size decreases exponentially, which amplifies the masking effect in the original ST-TSL procedure. Consequently, the advantage of the modified procedure is expected to become more pronounced with larger values of ν.

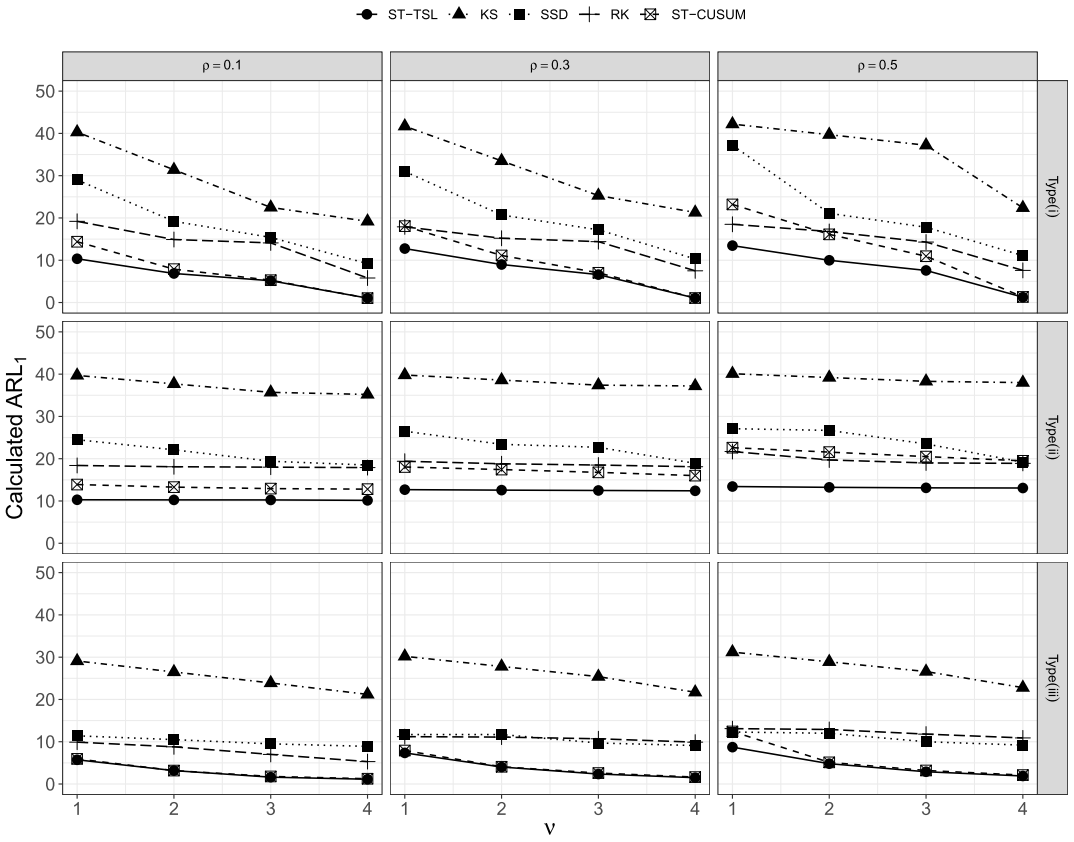

FIG. 6. *Optimal* ARL_1 *values of the five charts for detecting three types of shifts in cases when* $n = 300$, $m = 49$, *the nominal* ARL_0 *values of the five charts are all set to be* 50, *and* $\rho = 0.1, 0.3,$ *or* 0.5.

Under the settings $n = 300$, $\rho = 0.3$, $m = 49$, and nominal $\mathrm{ARL}_0 = 50$ (with all other configurations matching those in Section 3.3), the optimal ARL_1 values for both methods are presented in Figure 7. As shown, the modified ST–TSL consistently outperforms the original method, with the improvement becoming more substantial as ν increases, aligning with our expectations.

3.5. *Robustness to spatiotemporal nonseparability.* As discussed at the end of Section 2.1, the proposed ST-TSL method does not require the mean and covariance structures of the observed data to be space-time separable. To demonstrate this property, we consider two scenarios in which the covariance structure is space–time separable and nonseparable, respectively:

Separable case Let $\boldsymbol{\epsilon}(t_i) = (\epsilon(t_i, s_1), \ldots, \epsilon(t_i, s_m))^T$ for each i, where $\boldsymbol{\epsilon}(t_i)$ is generated from the following vector AR(1) model:

$$\boldsymbol{\epsilon}(t_i) = \rho_t \boldsymbol{\epsilon}(t_{i-1}) + \sqrt{1 - \rho_t^2}\boldsymbol{\eta}(t_i),$$

where $\boldsymbol{\eta}(t_i) = (\eta(t_i, s_1), \ldots, \eta(t_i, s_m))^T$ are independent over time t_i, and each follows a Gaussian spatial process with covariance given by

$$\mathrm{Cov}\big(\eta(t_i, s_j), \eta(t_i, s_l)\big) = \exp\left(-\frac{d_E(s_j, s_l)}{\rho_s}\right).$$

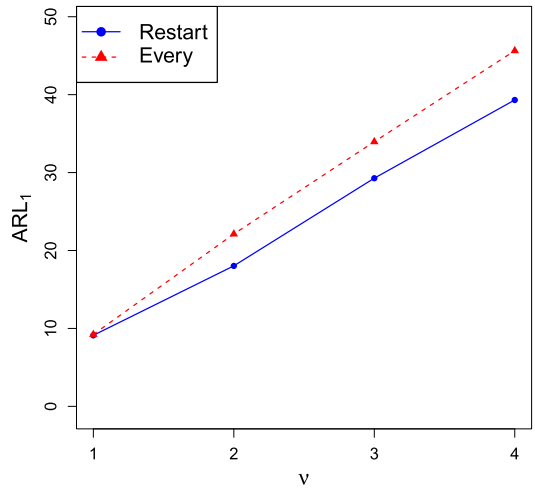

FIG. 7. *Comparison of the optimal* ARL_1 *values between the modified and original ST-TSL methods under the settings*: $n = 300$, $\rho = 0.3$, $m = 49$, *nominal* $\mathrm{ARL}_0 = 50$, *and shift type* (*iv*).

The resulting spatiotemporal covariance structure of Model (1) is given by

$$V(t_i, t_k; s_j, s_l) = \rho_t^{|k-i|} \exp\left(-\frac{d_E(s_j, s_l)}{\rho_s}\right),$$

where $\rho_t > 0$ and $\rho_s > 0$ are parameters that control the temporal and spatial correlations, respectively, with larger values indicating stronger correlations. Thus, in this case the spatiotemporal covariance structure is separable in space and time. In this example, we set $(\rho_t, \rho_s) = (0.3, 0.3)$, representing a moderate correlation scenario.

Inseparable case The spatiotemporal covariance structure follows the nonseparable model:

$$\mathrm{Cov}\big(\eta(t_i, s_j), \eta(t_k, s_l)\big) = \xi\big(d_E(s_j, s_l), |t_i - t_k|\big),$$

where

$$\xi(u, v) = \frac{1}{(a|v| + 1)^{0.5}} \exp\left(-\frac{cu}{(a|v| + 1)^{0.5}}\right),$$

and $(a, c) = (0.3, 0.3)$, representing a moderate correlation scenario.

In each case, we set $n = 300$, $m = 49$, and $\mathrm{ARL}_0 = 50$. All other setups are the same as those described in Section 3.1, where the process under monitoring is assumed to be IC. When the process becomes OC, the shift type (iv) with $\nu = 3$, as described in Section 3.4, is considered. Table 3 reports the actual ARL_0 and ARL_1 values of the proposed ST–TSL method. As shown in the table, its performance in the separable and nonseparable cases is very similar, indicating that ST–TSL is robust to spatiotemporal covariance nonseparability.

TABLE 3

Actual ARL_0 *and* ARL_1 *values and their standard errors* (*in parentheses*) *of the proposed ST–TSL method in the Separable and Inseparable Cases when* $n = 300$, $m = 49$, *and* $\mathrm{ARL}_0 = 50$

	ARL_0	ARL_1
Separable Case	49.00 (2.07)	24.21 (1.02)
Inseparable Case	50.62 (2.11)	23.98 (0.89)

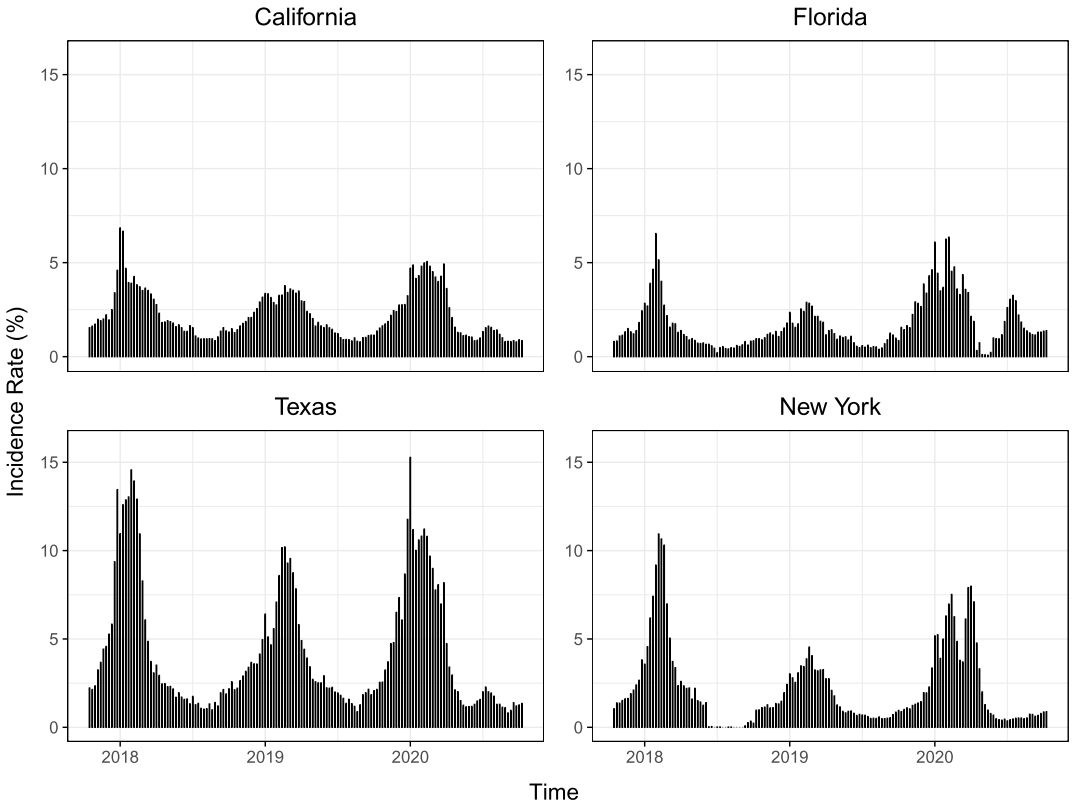

FIG. 8. *Weekly ILI incidence rates for California, Florida, Texas, and New York from Week* 41 *of* 2017 *to Week* 40 *of* 2020.

4. Monitoring the ILI data in the United States. In this section we apply the proposed ST–TSL method to monitor influenza-like illness (ILI) data in the United States, as introduced in Section 1. Monitoring ILI is crucial for the early detection of flu outbreaks, enabling timely public health interventions and efficient resource allocation. It also provides valuable insights into vaccine effectiveness and supports ongoing vaccine development efforts.

The ILI data analyzed in this study were collected weekly by the Outpatient Influenza-Like Illness Surveillance Network (ILINet), which consists of a nationwide network of healthcare providers in the United States. These data are publicly available from the CDC's website (https://www.cdc.gov/fluview/surveillance/usmap.html). Because our analysis accounts for the spatial characteristics of the data, we focus on the 48 contiguous U.S. states, excluding Hawaii and Alaska. The geometric centroids of the states are used to represent their locations, denoted as $\{\mathbf{s}_j, j = 1, \ldots, 48\}$, following standard practice in the epidemiological literature (Hanigan, Hall and Dear (2006)). The Euclidean distance between the geometric centroids of any two states is used to quantify their spatial separation. Figure 8 displays the weekly ILI incidence rates for California, Florida, Texas, and New York from Week 41 of 2017 to Week 40 of 2020. The plots clearly exhibit a seasonal pattern, with peaks consistently occurring during the winter months.

For disease surveillance, each flu season is defined as spanning from Week 41 of one year to Week 40 of the following year. Data from Week 41 of 2017 through Week 40 of 2019 are used as the initial IC data, while data from Week 41 of 2019 through Week 40 of 2020 are used for online process monitoring.

Before process monitoring, it is necessary to validate the presence of short-range serial correlation in the ILI data. Location-specific autocorrelation functions (ACFs), denoted as $\rho_{\mathbf{s}_j}(l)$

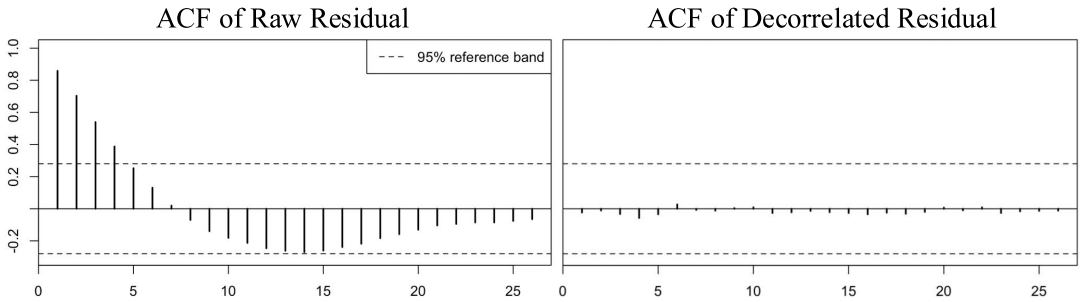

FIG. 9. *Pooled autocorrelation functions (ACFs) across all locations. The left panel displays the average ACF of the raw residuals, which shows strong temporal dependence lasting up to four weeks. The right panel presents the average ACF of the decorrelated residuals, where all autocorrelations lie within the* 95% *reference band (dashed lines), indicating that temporal dependence has been effectively removed.*

for $j = 1, \ldots, m$ and $l = 1, \ldots, \ell$, were computed using the IC residuals, where $m = 48$ is the number of locations and $\ell = 26$ is the maximum lag considered. At the 95% reference level, the ACFs at nearly all locations fall within lag 5, supporting the choice of $b_{\max} = 5$. ACFs were also computed for the decorrelated IC residuals, and no significant autocorrelation is found at the 95% reference level. This result confirms that temporal dependence was effectively removed and that the subsequent control-limit calibration is reliable. We then pool all location-specific ACFs using the Fisher's z-transformation as follows:

$$\bar{\rho}(l) = \tanh\left(\frac{1}{m}\sum_{j=1}^{m} \operatorname{atanh}\big(\rho_{\mathbf{s}_j}(l)\big)\right), \quad l = 1, \ldots, \ell.$$

The pooled ACFs computed from the original IC residuals and their decorrelated values are displayed in Figure 9. From the figure it can be observed that the original IC residuals exhibit significant autocorrelation up to lag 4, whereas no significant autocorrelation is present in the decorrelated IC residuals.

The proposed chart, ST–TSL, together with four competing charts KS, RK, SSD, and ST–CUSUM were constructed as described in Section 3, with nominal $\mathrm{ARL}_0 = 50$ for all methods. These five charts are displayed in Figure 10. As shown in the figure, the ST–CUSUM, KS, RK, SSD, and ST–TSL charts issue their first signals at Weeks 47 (2019), 4 (2020), 50 (2019), 8 (2020), and 46 (2019), respectively. Notably, the ST–TSL chart detects the signal earliest, whereas the RK chart lags behind by 14 weeks.

Figure 1 in Section 1 displays the observed ILI incidence rates for Week 46 of 2018 and Week 46 of 2019 (the first signal time of ST–TSL), revealing a pronounced spike in 2019 in Louisiana and several surrounding southern states. This shift in the data distribution confirms the outbreak detected by the ST–TSL chart. Table 4 presents the computation time of the proposed ST–TSL method, demonstrating its computational efficiency and suitability for real-world applications. In addition, an interactive Shiny application has been developed to visualize the observed ILI incidence rates across different times and locations, with the source code provided in the Supplementary Material.

5. Concluding remarks. In this article we introduced a novel disease surveillance method, ST–TSL, which incorporates transparent sequential learning (TSL) into spatiotemporal process monitoring. The method allows for recursive updates of IC parameter estimates, enhancing performance under both IC and OC conditions. ST–TSL accommodates flexible spatiotemporal variations, correlations, and distributions. Numerical results demonstrate that ST–TSL compares favorably with several representative existing methods, establishing it as a powerful tool for spatiotemporal disease surveillance.

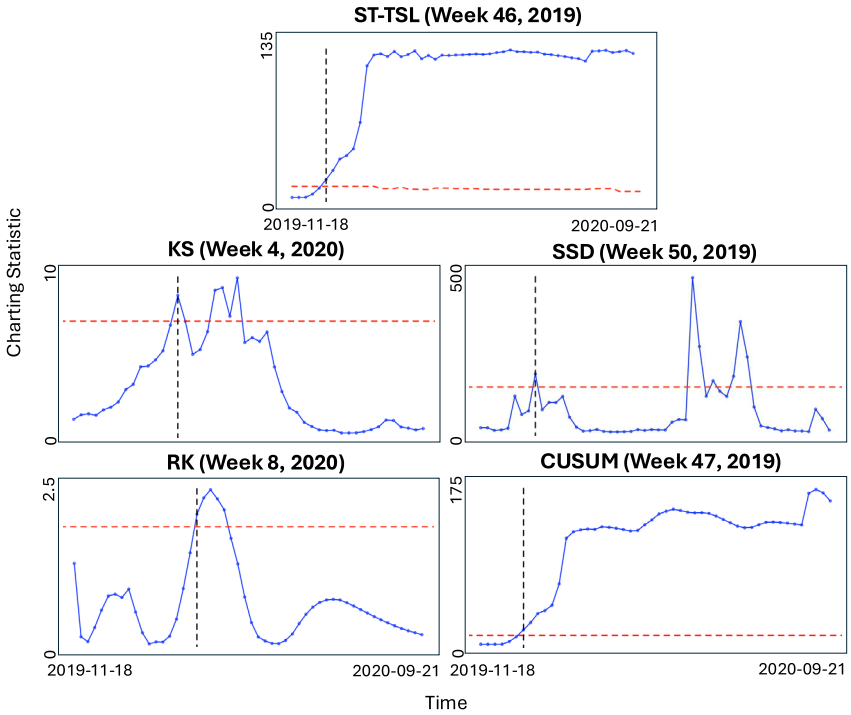

FIG. 10. *Control charts ST–CUSUM, KS, RK, SSD, and ST–TSL used to monitor ILI incidence rates across the* 48 *contiguous U.S. states from Week* 41 *of* 2019 *to Week* 40 *of* 2020. *In each plot, the dashed horizontal line denotes the control limit, while the dashed vertical line marks the first signal time, which is also reported in parentheses next to each chart label.*

Several challenges remain to be addressed in future research. While ST–TSL is effective in signaling disease outbreaks, it does not directly pinpoint the exact location or timing of the initial shift in disease incidence rates. Developing a robust post-signal diagnostic procedure to determine these specifics is essential for effective intervention planning. Although recursive updating of IC parameter estimates reduces the computational burden of ST–TSL, the block bootstrap procedure used to determine the control limit L of the CUSUM chart (6) remains time-intensive and more efficient strategies for calculating L are needed. Additionally, the current framework focuses on incidence rates expressed as continuous measures. Extending the ST–TSL approach to discrete outcomes, such as counts or bounded proportions—while appropriately accounting for their underlying distributional constraints—represents an important direction for future research, though it is beyond the scope of the present study. Finally, disease incidence rates are often influenced by covariates (e.g., weather conditions). Properly incorporating such covariates into the modeling framework constitutes another valuable avenue for future investigation.

TABLE 4

Wall-clock time (in seconds) for the real-data application. The column "Time for Computing L" presents the computing time for the block bootstrap procedure to determine the control limit L, while "Time for Monitoring" refers to the computing time of the subsequent ST–TSL monitoring process. All computations were performed on an Apple M1 Pro CPU using a standard gfortran build

Hardware (CPU)	Time for Computing L	Time for Monitoring
Apple M1 Pro	2.05	9.96

Significance statement. Public health agencies depend on the timely detection of abnormal disease activity to respond effectively to outbreaks. However, many existing statistical tools struggle with data whose distributions evolve over both space and time and exhibit complex correlations. We propose an adaptive monitoring framework that continuously learns from incoming observations, enabling earlier and more reliable detection of unusual patterns than traditional methods. The framework enhances robustness in real-world surveillance settings. Its sequential, automatic updating makes it well suited for modern streaming data and supports faster, data-driven decision-making in infectious disease control and other spatiotemporal monitoring applications.

Acknowledgments. The authors thank the Editor, Associate Editor, and four referees for their constructive comments and suggestions, which have greatly improved the quality of this paper.

Dr. Peihua Qiu is the corresponding author for this paper.

SUPPLEMENTARY MATERIAL

Supplementary.zip (DOI: 10.1214/26-AOAS2153SUPPA; .zip). This file contains the proofs of Theorems 2.1 and 2.2 and some extra numerical results.

Codes.zip (DOI: 10.1214/26-AOAS2153SUPPB; .zip). This ZIP file contains the Fortran code used to compute the ST-TSL charting statistics presented in Figure 10, as well as R code for a Shiny app that visualizes ILI incidence rates across states over time.

REFERENCES

APLEY, D. W. and TSUNG, F. (2002). The autoregressive T^2 chart for monitoring univariate autocorrelated processes. *J. Qual. Technol.* **34** 80–96.

BESAG, J., YORK, J. and MOLLIÉ, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Statist. Math.* **43** 1–59. MR1105822 https://doi.org/10.1007/BF00116466

BEST, N., RICHARDSON, S. and THOMSON, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Stat. Methods Med. Res.* **14** 35–59. MR2135924 https://doi.org/10.1191/0962280205sm388oa

BODNAR, R., BODNAR, T. and SCHMID, W. (2023). Sequential monitoring of high-dimensional time series. *Scand. J. Stat.* **50** 962–992. MR4630611 https://doi.org/10.1111/sjos.12607

BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** 107–144. MR2178042 https://doi.org/10.1214/154957805100000104

CHEN, N., ZI, X. and ZOU, C. (2016). A distribution-free multivariate control chart. *Technometrics* **58** 448–459. MR3556613 https://doi.org/10.1080/00401706.2015.1049750

DIGGLE, P., ROWLINGSON, B. and SU, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* **16** 423–434. MR2147534 https://doi.org/10.1002/env.712

EPSNEČNIKOV, V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Theory Probab. Appl.* **14** 153–158. MR0250422

GEE, A. H., CHANG, J., GHOSH, J. and PAYDARFAR, D. (2018). Bayesian online changepoint detection of physiological transitions. In *Proceedings of the* 40*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (*EMBC*) 45–48, Honolulu, HI, USA.

HANIGAN, I., HALL, G. and DEAR, K. B. (2006). A comparison of methods for calculating population exposure estimates of daily weather for health research. *Int. J. Health Geogr.* **5** 38.

HAWKINS, D. M. (1987). Self-starting cusums for location and scale. *Statistician* **36** 299–315.

HAWKINS, D. M., QIU, P. and KANG, C. W. (2003). The changepoint model for statistical process control. *J. Qual. Technol.* **35** 355–366.

HEISTERKAMP, S. H., DEKKERS, A. L. M. and HEIJNE, J. C. M. (2006). Automated detection of infectious disease outbreaks: Hierarchical time series models. *Stat. Med.* **25** 4179–4196. MR2307584 https://doi.org/10.1002/sim.2674

JACQUEZ, G. M. (1996). A k nearest neighbor test for spacetime interaction. *Stat. Med.* **15** 1935–1949.

KLEINMAN, K., LAZARUS, R. and PLATT, R. (2004). A generalized linear mixed models approach for detecting incident clusters of disease in small areas, with an application to biological terrorism. *Amer. J. Epidemiol.* **159** 217–224.

KNOX, E. and BARTLETT, M. (1964). The detection of space-time interactions. *J. R. Stat. Soc., Ser. C* **13** 25–30.
KULLDORFF, M. (1997). A spatial scan statistic. *Comm. Statist. Theory Methods* **26** 1481–1496. MR1456844 https://doi.org/10.1080/03610929708831995
KULLDORFF, M. (2001). Prospective time periodic geographical disease surveillance using a scan statistic. *J. Roy. Statist. Soc. Ser. A* **164** 61–72. MR1819022 https://doi.org/10.1111/1467-985X.00186
LAWSON, A. B., BIGGERI, A. B., BOEHNING, D., LESAFFRE, E., VIEL, J.-F., CLARK, A., SCHLATTMANN, P. and DIVINO, F. (2000). Disease mapping models: An empirical evaluation. *Stat. Med.* **19** 2217–2241.
LAWSON, A. B. and KLEINMAN, K. (2005). *Spatial and Syndromic Surveillance for Public Health*. Wiley, New York.
MANTEL, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27** 209–220.
MARSHALL, J. B., SPITZNER, D. J. and WOODALL, W. H. (2007). Use of the local Knox statistic for the prospective monitoring of disease occurrences in space and time. *Stat. Med.* **26** 1579–1593. MR2359160 https://doi.org/10.1002/sim.2603
MURRAY, J. and COHEN, A. L. (2017). Infectious disease surveillance. In *International Encyclopedia of Public Health* (S. R. Quah, ed.) 222–229.
PELAT, C., BOËLLE, P.-Y., COWLING, B. J., CARRAT, F., FLAHAULT, A. and ANSART, S. (2007). Online detection and quantification of epidemics. *BMC Med. Inform. Decis. Mak.* **7** article 29.
QIU, P. (2005). *Image Processing and Jump Regression Analysis*. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2111430 https://doi.org/10.1002/0471733156
QIU, P. (2014). *Introduction to Statistical Process Control*. CRC Press/CRC, Boca Raton, FL.
QIU, P. (2024). *Statistical Methods for Dynamic Disease Screening and Spatio-Temporal Disease Surveillance*. CRC Press/CRC, Boca Raton, FL.
QIU, P., LI, W. and LI, J. (2020). A new process control chart for monitoring short-range serially correlated data. *Technometrics* **62** 71–83. MR4058600 https://doi.org/10.1080/00401706.2018.1562988
QIU, P. and XIE, X. (2022). Transparent sequential learning for statistical process control of serially correlated data. *Technometrics* **64** 487–501. MR4506615 https://doi.org/10.1080/00401706.2021.1929493
QIU, P. and YANG, K. (2023). Spatiotemporal process monitoring using exponentially weighted spatial LASSO. *J. Qual. Technol.* **55** 163–180.
REIS, B. Y. and MANDL, K. D. (2003). Time series modeling for syndromic surveillance. *BMC Med. Inform. Decis. Mak.* **3**. article 2.
RODRIGUEZ AVELLANEDA, F., MATEU, J. and MORAGA, P. (2024). Estimating velocities of infectious disease spread through spatiotemporal log-Gaussian Cox point processes. arXiv preprint. Available at arXiv:2409.05036. https://doi.org/10.48550/arXiv.2409.05036
SONESSON, C. and BOCK, D. (2003). A review and discussion of prospective statistical surveillance in public health. *J. Roy. Statist. Soc. Ser. A* **166** 5–21. MR1973851 https://doi.org/10.1111/1467-985X.00256
STROUP, D., WHARTON, M., KAFADAR, K. and DEAN, A. (1999). Evaluation of a method for detecting aberrations in public health surveillance data. *Amer. J. Epidemiol.* **137** 373–380.
TAKAHASHI, K., KULLDORFF, M., TANGO, T. and YIH, K. (2008). A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *Int. J. Health Geogr.* **7** article 14.
WANG, Y., CHEN, X. and XUE, F. (2024). A review of Bayesian spatiotemporal models in spatial epidemiology. *ISPRS Int. J. Geo-Inf.* **13** 97. https://doi.org/10.3390/ijgi13030097
WANG, Z., GOEDHART, R. and ZWETSLOOT, I. M. (2023). Monitoring high-dimensional heteroscedastic processes using rank-based EWMA methods. *Comput. Ind. Eng.* **184** 109544. https://doi.org/10.1016/j.cie.2023.109544
WEINSTOCK, M. A. (1981). A generalized scan statistic test for the detection of clusters. *Int. J. Epidemiol.* **10** 289–293.
XIE, X. and QIU, P. (2023). Control charts for dynamic process monitoring with an application to air pollution surveillance. *Ann. Appl. Stat.* **17** 47–66. MR4539021 https://doi.org/10.1214/22-aoas1615
YAN, H., PAYNABAR, K. and SHI, J. (2018). Real-time monitoring of high-dimensional functional data streams via spatio-temporal smooth sparse decomposition. *Technometrics* **60** 181–197. MR3804247 https://doi.org/10.1080/00401706.2017.1346522
YANG, K. and QIU, P. (2018). Spatiotemporal incidence rate data analysis by nonparametric regression. *Stat. Med.* **37** 2094–2107. MR3802936 https://doi.org/10.1002/sim.7622
YANG, K. and QIU, P. (2019). Nonparametric estimation of the spatio-temporal covariance structure. *Stat. Med.* **38** 4555–4565. MR4010035 https://doi.org/10.1002/sim.8315
YANG, K. and QIU, P. (2020). Online sequential monitoring of spatiotemporal disease incidence rates. *IISE Trans.* **52** 1218–1233.
YU, Y., MADRID PADILLA, O. H., WANG, D. and RINALDO, A. (2023). A note on online change point detection. *Sequential Anal.* **42** 438–471. MR4678160 https://doi.org/10.1080/07474946.2023.2276170

ZHANG, J., KANG, Y., YANG, Y. and QIU, P. (2015). Statistical monitoring of the hand, foot and mouth disease in China. *Biometrics* **71** 841–850. MR3402620 https://doi.org/10.1111/biom.12301

ZHAO, Y., ZENG, D., HERRING, A. H., ISING, A., WALLER, A., RICHARDSON, D. and KOSOROK, M. R. (2011). Detecting disease outbreaks using local spatiotemporal methods. *Biometrics* **67** 1508–1517. MR2872402 https://doi.org/10.1111/j.1541-0420.2011.01585.x

ZHOU, H. and LAWSON, A. B. (2008). EWMA smoothing and Bayesian spatial modeling for health surveillance. *Stat. Med.* **27** 5907–5928. MR2597751 https://doi.org/10.1002/sim.3409

ZHOU, Y. and QIU, P. (2026). Supplement to "Transparent Sequential Learning and Monitoring of Spatiotemporal Disease Incidence Rates." https://doi.org/10.1214/26-AOAS2153SUPPA, https://doi.org/10.1214/26-AOAS2153SUPPB