Bayesian pollution source identification via an inverse physics model

Youngdeok Hwang [a], Hang J. Kim [b,*], Won Chang [b], Kyongmin Yeo [c], Yongku Kim [d]

[a] Department of Statistics, Sungkyunkwan University, Republic of Korea
[b] Division of Statistics and Data Science, University of Cincinnati, USA
[c] IBM T.J. Watson Research Center, USA
[d] Department of Statistics, Kyungpook National University, Republic of Korea

ABSTRACT

The behavior of air pollution is governed by complex dynamics in which the air quality of a site is affected by the pollutants transported from neighboring locations via physical processes. To estimate the sources of observed pollution, it is crucial to take the atmospheric conditions into account. Traditional approaches to building empirical models use observations, but do not extensively incorporate physical knowledge. Failure to exploit such knowledge can be critically limiting, particularly in situations where near-real-time estimation of a pollution source is necessary. A Bayesian method is proposed to estimate the locations and relative contributions of pollution sources by incorporating both the physical knowledge of fluid dynamics and observed data. The proposed method uses a flexible approach to statistically utilize large-scale data from a numerical weather prediction model while integrating the dynamics of the physical processes into the model. This method is illustrated with a real wind data set.

1. Introduction

Locating the sources of air pollution provides the policy makers with the critical information necessary to curb the pollution, and may also serve to resolve international disputes between neighboring countries. In some instances, the pollution sources are clearly visible in the satellite images. For example, during the 2013 Southeast Asian Haze event, smoke from the massive wildfires in Indonesia was visible from space (Reid et al., 2013). In most cases, however, pollution is not readily observable and associated data are scarce, whereas pollution can be generated from various sources — traffic, fossil fuel combustion, or the burning of natural biomass (World Health Organization, 2005). Another important challenge is pinpointing and better tracking atmospheric releases of hazardous materials from industrial accidents or terrorism. When such an event occurs, it is crucial to find the location of the event to contain the releases and protect the public. Studies of computational fluid dynamics models that simulate airflow and dispersion in urban areas have been sponsored by the U.S. Department of Energy and Department of Homeland Security (e.g., Chan and Leach, 2007). Due to the importance of these environmental problems, an extensive effort to improve the speed and fidelity of pollution modeling has been undertaken by the applied mathematics and physics communities. Examples include the Weather Research and Forecasting coupled with Chemistry (WRF-Chem, Fast et al., 2006) and the Community Multiscale Air Quality Model (Byun and Schere, 2006).
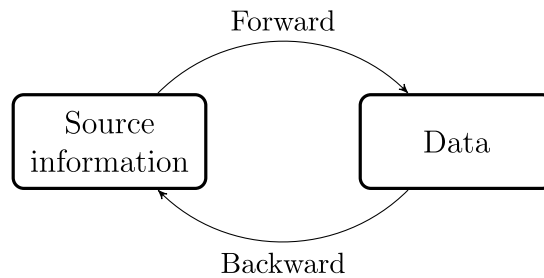

Fig. 1. Description of the forward and inverse problems.

Although there are differences among the models, their common goal is to provide a computational tool that simulates the behavior of a pollutant with the *given* source information.

Computational models are designed to produce the data from the known physical parameters and conditions; in the environmental context, the parameters are the location of pollution sources and the quantity of emissions from each source. This is called the *forward* problem because it focuses on producing predictive outputs from the known or assumed physical parameters, i.e., it looks forward in time. Another important problem involves locating the sources of pollution from the observed data, which is often called the *inverse* or *backward* problem (Biegler et al., 2011). The notion of forward and backward problems is depicted in Fig. 1. Traditional approaches to building statistical models for source estimation include chemical mass balance models such as receptor models (Park et al., 2001; Christensen and Gunst, 2004). The chemical mass balance approach uses a finite number of chemical fingerprints of sources to find a matrix factorizer. These models are useful for estimating the overall attribution of each source from the historical data, but are not suited to rapid identification of emission locations.

In statistics, the backward problem is called model calibration (Kennedy and O'Hagan, 2001; Higdon et al., 2008; Chang and Joseph, 2014; Chang et al., 2015). There have been studies in several different settings dealing with model calibration, especially in the context of pollution source estimation. For example, Section 2.2.1 of Kennedy and O'Hagan (2001) used a simple plume model to illustrate the Bayesian calibration problem, in which executing the physics model simply amounts to evaluating a closed-form mathematical function. Williams et al. (2011) have proposed a Bayesian method using a physics model called AERMOD (Cimorelli et al., 2005). The AERMOD data generation process is directly contained in the likelihood, so the physics model is run at each iteration of the Markov Chain Monte Carlo (MCMC) process to estimate the input parameter, since this model is computationally affordable. In these examples, the computational physics models are simple enough for many simulations to be run in a short period of time. The simplicity of this model, however, limits its accuracy and its capability to represent complex physical processes, which adversely affects the fineness of the information obtained using this methodology. For example, Cimorelli et al. (2005) only identifies the direction of an emission source from the observation location.

With recent advances in computational physics models, the data with a high level of granularity are available from numerical weather prediction (NWP) models. The physical knowledge is naturally woven into the outputs of an NWP model, but current statistical frameworks using such data have not caught up enough to exploit such data. In this context, our goal is to propose a flexible Bayesian physics–statistics hybrid approach to estimate the locations of the main pollution sources and their corresponding contribution. There were three main considerations in developing our methodology. First, the availability of data is asymmetrical. Real observations are scarce, and only available at limited locations, while the NWP output is available on a highly dense computational grid, which allows us to leverage physical knowledge to overcome the lack of the data. Second, physical interpretability is critical, and the results of the methodology should be communicated with the experts in physics. Third, we pursue computational affordability along with flexibility. Because the main target is to identify the source of emissions for an event producing abrupt spikes in pollution concentration levels, the method should be run quickly while providing useful information for the decision-making.

To develop a statistical framework that achieves these goals, we propose a flexible methodology using a Bayesian hierarchical model, coupled with a dispersion process. The dispersion process introduces the physical knowledge into the Bayesian analysis model while connecting the real observations and the NWP output data. This model incorporates the inherent spatial nature and high-dimensionality of the model parameters, as well as the physical constraints of the parameters.

The remainder of the paper is organized as follows. Section 2 describes the physical processes used as building blocks for our methodology. Section 3 presents the proposed statistical model. Section 4 presents simulations with real wind data sets for studying the proposed method. Section 5 provides some concluding remarks.

2. Background

In this section, we introduce the underlying physical processes and connect them to the statistical problem. For simplicity, the description is given as a two-dimensional problem, although our approach can be applied in a three-dimensional setting.

2.1. Forward dispersion model

The fundamental building block used to incorporate the physics into our approach is the *advection–diffusion* equation. We are interested in the pollution concentration in the spatial and temporal domains, $\mathbb{S}$ and $\mathbb{T}$, respectively. For $s \in \mathbb{S}$ and $t \in \mathbb{T}$, $z = z(s, t)$ denotes the pollution concentration, $u = (u_1(s, t), u_2(s, t))$ the wind velocity in the north and east directions, and $\theta = \theta(s, t)$ the pollution emission rate, respectively. We use the advection–diffusion equation to model the change of the pollution concentration over $\mathbb{S} \times \mathbb{T}$, given by

$$\frac{\partial z}{\partial t} = \theta - \nabla \cdot (uz) + \nabla \cdot (K_0 \cdot \nabla z) - \eta_0 z, \tag{1}$$

where $\nabla \equiv (\partial/\partial s_1, \partial/\partial s_2)$ is the gradient operator for the spatial domain, $K_0 = \mathrm{diag}(\kappa_0, \kappa_0)$ is the eddy diffusivity used to model the turbulent scalar fluxes, and η_0 is the uniform decay parameter.

Despite its complexity, the physical principle embedded in (1) is straightforward. The change in the pollution concentration at the spatial domain of interest, $\partial z / \partial t$, is decomposed into three elements. The first element is the newly added pollution at each location, θ, which is our main interest. Because it is defined at every location in the spatial domain, its dimension is equal to $|\mathbb{S}|$. The second element is the advection and the diffusion to and from the nearby locations. The net amount of advection, $-\nabla \cdot (uz)$, is determined by the speed and direction of the wind and the amount of pollutant concentration. The diffusion, expressed as $\nabla(K_0 \cdot \nabla z)$, is the movement of the pollutants from a region of high concentration to a region of low concentration, and is determined by the difference of pollutant concentration ∇z over $\mathbb{S}$ and the diffusion mixing rate κ_0. The last element is the removal process, in which we assume the rate of decay to be uniform with η_0.

To connect the physical processes with the observations, we define a detector function by

$$I(s, t; S, T) = \begin{cases} \frac{1}{|S||T|} & \text{if } s \in S \text{ and } t \in T \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The average concentration over region S during time interval T under the advection–diffusion in (1) is given by

$$\omega^F(S, T) \equiv \int_{\mathbb{S}} \int_{\mathbb{T}} z(s, t) I(s, t; S, T) \, dt \, ds. \tag{3}$$

Intuitively, Eq. (3) defines the average pollution concentration of S as the average of the pollution particles that have existed during T. This definition also directly takes into account the practical consideration that the most commonly available form of pollution observation data is the average amount of the pollutant in a unit volume at a specific location during a certain time interval (e.g., the one-hour average concentration at monitoring location i within region S_i, assuming the constant level of pollution concentration over region S_i). Eq. (1) views the process forward in time, i.e., the concentration field of the next period evolves from the current status, which is indicated by the superscript F in (3).

Hwang et al. (2018) considered an approach to estimate spatial pollution contribution using the forward model in (1). Their approach simplifies the spatial structure of θ by using a basis representation. It focuses on an estimation of the average long-term contribution and therefore requires a data set associated with a timescale of several weeks. In contrast, our work focuses on the abrupt pollution spike events with a limited data availability, which exploits the backward approach and will be described in the next section.

2.2. Backward model

We now describe the backward model developed from (1)–(3) beginning with the definition of a *conjugate* field. From the advection–diffusion equation in (1), the conjugate field $\chi = \chi(s, t; S, T)$ corresponding to the detector $l = I(s, t; S, T)$ is defined by an adjoint equation of the advection–diffusion as

$$\frac{\partial \chi}{\partial t} = -l - u \cdot (\nabla \chi) - \nabla \cdot (K_0 \cdot \nabla \chi) + \eta_0 \chi. \tag{4}$$

Contrary to (1), the adjoint equation of (4) is solved backward in time by reversing the direction of wind field, which gives the name "backward" model. The forward model computes pollution dispersion from a given source location, while the backward model gives the candidate locations of pollution sources for the pollutants detected by $I(s, t; S, T)$. Similar to $\omega^F(S, T)$ in (3), the average concentration associated with l from the backward model can be written as

$$\omega^B(S, T) \equiv \int_{\mathbb{S}} \int_{\mathbb{T}} \theta(s, t) \chi(s, t; S, T) \, dt \, ds, \tag{5}$$

now with the superscript B indicating the backward. From duality, it can be shown that $\omega^F(S, T) = \omega^B(S, T)$. This equivalence establishes the direct relationship between the backward and forward models. In plain words, the backward model implies that the pollutants observed at the monitoring stations can be traced back to their origin following the wind backward, whereas the forward model suggests that the pollutants emitted at each source location are transported and observed at the monitoring stations, following the wind forward, under the physical systems governed by (1) and (4).

2.3. Numerical discretization

The partial differential equations (PDE) in (1) and (4) are solved numerically by projecting the continuous functions onto a discrete space. We use the *finite difference method*, which is one of the most widely used methods to solve PDEs numerically. Here we give a brief sketch of the scheme.

First, we define a grid for the time domain by $\{m\delta_t : m = 1, \ldots, M\}$, where δ_t denotes the size of the time grid. The study period is partitioned into M intervals, and each interval is represented by a time point. Similarly, the spatial grid is defined to be $\{s = (s_1, s_2) = (a\delta_s, b\delta_s) : a = 1, \ldots, A, b = 1, \ldots, B\}$, and δ_s is the size of the spatial grids. The spatial domain is divided into $A \times B$ cells of size $\delta_s \times \delta_s$, each of which is represented by a single spatial point among the grid. We use the equidistant grid size for the east and north directions for the sake of brevity, but they can be chosen flexibly. Let $z_{(a,b)}^{(m)} = z(a\delta_s, b\delta_s, m\delta_t)$. In the finite difference method, the projection uses the Taylor expansion. Specifically, the first-order spatial derivative of the pollution concentration z at $s = (a\delta_s, b\delta_s)$ with respect to s_1 is

$$\frac{\partial z}{\partial s_1}\bigg|_{s=(a\delta_s, b\delta_s)} = \frac{z((a+1)\delta_s, b\delta_s, t) - z((a-1)\delta_s, b\delta_s, t)}{2\delta_s} + O(\delta_s).$$

Similarly, the second-order spatial derivative can be computed by

$$\frac{\partial^2 z}{\partial s_1^2}\bigg|_{s=(a\delta_s, b\delta_s)} = \frac{z((a+1)\delta_s, b\delta_s, t) - 2z(a\delta_s, b\delta_s, t) + z((a-1)\delta_s, b\delta_s, t)}{\delta_s^2} + O(\delta_s).$$

The derivatives can be approximated by the function values defined at the grid points. In the limit of $\delta_s \to 0$, the finite difference approximation on the right-hand side converges to the left-hand side of the equation at a rate of δ_s. The derivatives with respect to s_2 can be defined similarly.

Regarding t, we use *explicit Euler time advancement*. For a time-marching problem (i.e., a process that evolves over time), we define $f(s, t) = \partial z(s, t)/\partial t$. The explicit Euler method yields

$$f(s, t) = \frac{z(s, t + \delta_t) - z(s, t)}{\delta_t} + O(\delta_t). \tag{6}$$

From (6), one can see

$$z(s, t + \delta_t) = z(s, t) + f(s, t)\delta_t + O(\delta_t^2),$$

which indicates that the function value at the future step $t + \delta_t$ is computed from the present values. It is apparent that the discrete formula converges to the analytical expression in the limit of $\delta_t \to 0$.

As such, Eq. (1) for $z(a\delta_s, b\delta_s, m\delta_t)$ is approximated by the following finite difference equation:

$$\frac{z_{(a,b)}^{(m+1)} - z_{(a,b)}^{(m)}}{\delta_t} = \theta_{(a,b)}^{(m)} - \frac{u_{1(a+1,b)}^{(m)} z_{(a+1,b)}^{(m)} - u_{1(a-1,b)}^{(m)} z_{(a-1,b)}^{(m)}}{2\delta_s} - \frac{u_{2(a,b+1)}^{(m)} z_{(a,b+1)}^{(m)} - u_{2(a,b-1)}^{(m)} z_{(a,b-1)}^{(m)}}{2\delta_s}$$

$$+ \kappa_0 \frac{z_{(a+1,b)}^{(m)} - 2z_{(a,b)}^{(m)} + z_{(a-1,b)}^{(m)}}{\delta_s^2} + \kappa_0 \frac{z_{(a,b+1)}^{(m)} - 2z_{(a,b)}^{(m)} + z_{(a,b-1)}^{(m)}}{\delta_s^2} - \eta_0 z_{(a,b)}^{(m)}. \tag{7}$$

The approximation error of (7) with respect to the original PDE is $O(\delta_t) + O(\delta_s)$. The solution of the finite difference equation approaches closer to the original PDE solution as the grid becomes finer, but with a larger computational burden. There are other numerical considerations in employing (7). We refer readers to Strikwerda (2004) and Moin (2010) for a discussion of other issues, such as selection of the grid size.

Let region S_i be the $\delta s \times \delta s$ rectangular area enclosing the station i. In the discretized domain, the region S_i and observation time interval T are represented by $s_i = (a_i\delta_s, b_i\delta_s)$ and $\{m\delta_t : m = m_{\text{start}}, m_{\text{start}} + 1, \ldots, m_{\text{end}}\}$, respectively. Then, Eq. (4) is approximated by

$$\frac{\chi_{i(a,b)}^{(m)} - \chi_{i(a,b)}^{(m+1)}}{\delta_t} = -I[(a, b), m; (a_i, b_i), \{m_{\text{start}}, \ldots, m_{\text{end}}\}]$$

$$- u_{1(a,b)}^{(m)} \frac{\chi_{i(a+1,b)}^{(m)} - \chi_{i(a-1,b)}^{(m)}}{2\delta_s} - u_{2(a,b)}^{(m)} \frac{\chi_{i(a,b+1)}^{(m)} - \chi_{i(a,b-1)}^{(m)}}{2\delta_s}$$

$$- \kappa_0 \frac{\chi_{i(a+1,b)}^{(m)} - 2\chi_{i(a,b)}^{(m)} + \chi_{i(a-1,b)}^{(m)}}{\delta_s^2} - \kappa_0 \frac{\chi_{i(a,b+1)}^{(m)} - 2\chi_{i(a,b)}^{(m)} + \chi_{i(a,b-1)}^{(m)}}{\delta_s^2} + \eta_0 \chi_{i(a,b)}^{(m)}$$

where $\chi_{i(a,b)}^{(m)}$ denotes the conjugate field associated with station i whose detector function is given by

$$I[(a, b), m; (a_i, b_i), \{m_{\text{start}}, \ldots, m_{\text{end}}\}] = \frac{1}{(m_{\text{end}} - m_{\text{start}} + 1)\delta_t \delta_s^2}$$

when $(a, b) = (a_i, b_i)$ and $m_{\text{start}} \le m \le m_{\text{end}}$; and zero otherwise. Note that the conjugate field is the function of wind velocity, eddy diffusivity, and the decay parameter, but not of pollution emissions $\{\theta_{(a,b)}^{(m)}\}$ and that the conjugate field needs not be re-calculated for the estimation, as presented in the following section.

3. Statistical model and inference

In this section, we present a statistical framework to take advantage of the backward model. In the following subsections, we introduce a model derived from (5) and then present a statistical methodology to overcome challenges associated with estimation.

3.1. Statistical model with uncertainty

In practice, the forward and backward model outputs in (3) and (5) are approximated by numerical integration. Suppose that station i belongs to region S_i and measures the concentration during T that are represented by $s_i = (a_i \delta_s, b_i \delta_s)$ and $\{m_{start} \delta_t, \ldots, m_{end} \delta_t\}$ in the discretized domain, respectively. Then, Eq. (3) can be approximated by

$$\omega^F(S_i, T) \approx \tilde{\omega}_i^F \equiv \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{m=1}^{M} z_{(a,b)}^{(m)} I[(a, b), m; (a_i, b_i), \{m_{start}, \ldots, m_{end}\}] \delta_t \delta_s^2. \tag{8}$$

The discretization scheme implicitly assumes that all parameters are constant over region S_i at the discretized time point $m\delta_t$, so the measurement $\tilde{\omega}_i^F$ will be the same in S_i regardless of the precise location of station i within S_i. For the backward model in (5), the finite difference approximation gives

$$\omega^B(S_i, T) \approx \tilde{\omega}_i^B \equiv \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{m=1}^{M} \theta_{(a,b)}^{(m)} \chi_{i(a,b)}^{(m)} \delta_t \delta_s^2. \tag{9}$$

We assume that the emission rate is constant over the time of the study, i.e., $\theta_{(a,b)}^{(m)} = \theta_{(a,b)}$.

Several sources contribute to the discrepancy between the real measurements and the physics model output: (a) measurement error, (b) model inadequacy of the physics model in representing the real process, and (c) discretization error from using the finite difference approximation. To incorporate the uncertainties, we consider the model given by

$$y_i = \tilde{\omega}_i^F + \varepsilon_i^F = \tilde{\omega}_i^B + \varepsilon_i^B \tag{10}$$

where ε_i^F (or ε_i^B) is the discrepancy between the real measurements and the forward (or backward) physics model calculation consisting of (a)–(c), which is assumed to follow N$(0, \sigma^2)$. The random errors ε_i^F and ε_i^B for y_i may be similar to each other, but they are not identical because of the slight discrepancy between $\tilde{\omega}_i^F$ and $\tilde{\omega}_i^B$ with the discretization error. Note that, in our pollution emission context, the unit of $\theta_{(a,b)}$ is $(\mu g/m^3)/h$ as it denotes the emission rate. Meanwhile, $z_{(a,b)}^{(m)}$ in (8) is the pollution level at $m\delta_t$ and $\tilde{\omega}_i^B$ in (9) is the concentration amount averaged for $(m_{end} - m_{start} + 1)\delta_t$ hours, both with the units of $\mu g/m^3$.

We now define $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ where $p = A \times B$ is the total size of the spatial grid. The vector of the unknown intensity parameter $\boldsymbol{\beta}$ over the spatial domain is obtained by vectorizing the matrix $\theta_{(a,b)}$. The backward finite difference approximation in (10) is then written as a linear model, $y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i^B$ where $\boldsymbol{x}_i = \delta_t \delta_s^2 (\sum_{m=1}^{M} \chi_{i1}^{(m)}, \ldots, \sum_{m=1}^{M} \chi_{ip}^{(m)})^\top$. This linear approximation provides the compatibility with a statistical methodology.

3.2. Bayesian hierarchical linear model with regularization

In inverse problems, the number of spatial grid points p is usually much larger than the number of observations n. This high dimensionality is particularly troublesome for our problem because we are interested in detecting abrupt spike events, which limits the number of available observations. To address this challenge, we propose a Bayesian hierarchical linear model given by

$$\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim N(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}), \tag{11}$$

$$f(\boldsymbol{\beta}|\sigma^2, \lambda_1, \lambda_2) \propto N(\boldsymbol{\beta}; \boldsymbol{0}, \sigma^2(\lambda_1 \boldsymbol{I} + \lambda_2 \boldsymbol{W})^{-1}) \prod_{j=1}^{p} \mathbb{I}(\beta_j \geq 0) \tag{12}$$

where $\boldsymbol{X}$ is a $n \times p$ matrix whose ith row is $\boldsymbol{x}_i^\top$ and $\boldsymbol{I}$ is the $p \times p$ identity matrix. For notational convenience, we interchangeably use N to express the normal distribution, e.g., N$(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and its probability density function, e.g., N$(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. The indicator function $\mathbb{I}(\cdot)$ restricts the support of β_j to be *non-negative* as each pollution source only can add emissions to the ambient air. We assumed that many values of β_j's are close to zero for many grid points, i.e., they do not add meaningful pollution to the air. This *sparsity* is adjusted by the model parameter λ_1. Finally, the emission intensity is assumed to be spatially dependent because nearby sources tend to behave similarly. The model parameter λ_2 characterizes the strength of the *spatial dependency* between neighbors. Let $\{j_{x-}, j_{x+}\}$ and $\{j_{y-}, j_{y+}\}$ denote the neighboring cells of j along with the x-axis and y-axis, respectively. We define Ω_{x+} to be the set of j which has the neighbor j_{x+} in the study region and define Ω_{x-}, Ω_{y+}, and Ω_{y-} in a similar manner. Then, we define the adjacency matrix $\boldsymbol{W} = \{W_{jl}, j, l = 1, \ldots, p\}$ as

$$W_{jj} = \mathbb{I}(j \in \Omega_{x+}) + \mathbb{I}(j \in \Omega_{x-}) + \mathbb{I}(j \in \Omega_{y+}) + \mathbb{I}(j \in \Omega_{y-})$$

and

$$W_{jl} = \begin{cases} -1 & \text{if } l = j_{x+}, j_{x-}, j_{y+}, \text{ or } j_{y-} \\ 0 & \text{otherwise.} \end{cases}$$

The diagonal elements take into account the different number of neighborhood grids depending on where a pixel is located; W_{jj} is 4 for an interior, 3 for a border, and 2 for a corner grid of the study region. The off-diagonal elements incorporate the connected neighborhood structure.

For Bayesian inference, we consider the following prior distributions. For the scale parameter σ^2, the improper prior $f(\sigma^2) = 1/\sigma^2$ was often used (for example, Park and Casella, 2008; Kyung et al., 2010), but we find that this prior causes a too large value of σ^2 when p is large, which results in a poor fit of $\hat{y}_i$ to y_i. Instead, we use the proper prior $f(\sigma^2) \sim \text{Inv-Gamma}(a_\sigma, b_\sigma)$ by setting $a_\sigma = p + 1$ and $b_\sigma = p\mu_\sigma$, where μ_σ is the prior mean of σ^2, provided by expert knowledge. For λ_1 and λ_2, we set $f(\lambda_1) \sim \text{Gamma}(c_1, d_1)$ with a mean of c_1/d_1 and $f(\lambda_2) \sim \text{Gamma}(c_2, d_2)$ with mean c_2/d_2 for computational convenience.

3.3. Computation

For posterior inference of the model, we implemented the MCMC, using the Metropolis-within-Gibbs algorithm:

Step 1. Metropolis updating for β_j.

    (a) For each $j = 1, \ldots, p$, propose b_j from $\log\text{N}(\log \beta_j, s_\sigma^2)$ with a step-size of s_σ. Then, we update $\beta_j = b_j$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \lambda_1 \boldsymbol{b}^\top \boldsymbol{b} + \lambda_2 \boldsymbol{b}^\top \boldsymbol{W}\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right\}\right]} \frac{\log\text{N}(\beta_j; \log b_j, s_\sigma^2)}{\log\text{N}(b_j; \log \beta_j, s_\sigma^2)}.$$

    and $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, b_j, \beta_{j+1}, \ldots, \beta_p)^\top$.

    (b) For each neighbor (j, j_{x+}) or (j, j_{y-}), propose $\boldsymbol{b}$ where β_j and $\beta_{j_{x+}}$ are swapped, i.e., $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j_{x+}}, \beta_{j+1}, \ldots, \beta_{j_{x+}-1}, \beta_j, \beta_{j_{x+}+1}, \ldots, \beta_p)^\top$. Then, update $\boldsymbol{\beta} = \boldsymbol{b}$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \lambda_1 \boldsymbol{b}^\top \boldsymbol{b} + \lambda_2 \boldsymbol{b}^\top \boldsymbol{W}\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right\}\right]}.$$

Step 2. Gibbs sampling for σ^2.

    Update σ^2 from its closed form of the full conditional distribution,

$$\sigma^2 | \cdots \sim \text{Inv-Gamma}\left(a_\sigma + \frac{n+p}{2}, \ b_\sigma + \frac{1}{2}\left[(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right]\right).$$

Step 3. Metropolis updating for λ_1.

    Propose λ_1 from $\log\text{N}(\log \lambda_1, s_1^2)$ with a step-size of s_1. Then, update $\lambda_1 = l_1$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{|l_1 \boldsymbol{I} + \lambda_2 \boldsymbol{W}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(l_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right)\right\} \text{Gamma}(l_1; c_1, d_1)}{|\lambda_1 \boldsymbol{I} + \lambda_2 \boldsymbol{W}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right)\right\} \text{Gamma}(\lambda_1; c_1, d_1)} \frac{\log\text{N}(\lambda_1; \log l_1, s_1^2)}{\log\text{N}(l_1; \log \lambda_1, s_1^2)}.$$

Step 4. Metropolis updating for λ_2.

    Propose λ_2 from $\log\text{N}(\log \lambda_2, s_2^2)$ with a step-size of s_2. Then, update $\lambda_2 = l_2$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{|\lambda_1 \boldsymbol{I} + l_2 \boldsymbol{W}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + l_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right)\right\} \text{Gamma}(l_2; c_2, d_2)}{|\lambda_1 \boldsymbol{I} + \lambda_2 \boldsymbol{W}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\lambda_1 \boldsymbol{\beta}^\top \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^\top \boldsymbol{W}\boldsymbol{\beta}\right)\right\} \text{Gamma}(\lambda_2; c_2, d_2)} \frac{\log\text{N}(\lambda_2; \log l_2, s_2^2)}{\log\text{N}(l_2; \log \lambda_2, s_2^2)}.$$

In Step 1, we adopt Metropolis–Hastings updating instead of Gibbs sampling for three reasons. First, our Metropolis updating avoids the matrix inversion required for Gibbs sampling. With a large p, the matrix inversion is to be avoided because of the computational burden and numerical instability. Second, it is difficult to sample $\boldsymbol{\beta}$ directly from the multivariate truncated normal distribution when p is large. One may consider drawing β_j, given $\beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_p$, from the truncated normal distribution, but it is difficult to compute $\log N(\beta_j; \mu_\beta, \sigma^2) \boldsymbol{I}(\beta_j \geq 0)$ when μ_β has a very small negative value compared with σ. Third, we include the swapping operation in Step 1 (b). Due to the strong dependency between neighbors and the consequential impact on the acceptance probability, individually proposed values in Step 1 (a) are rejected too frequently, resulting in a slow mixing of the chain. In contrast, the swapping operation in Step (b) simultaneously updates the neighboring values, e.g., β_j and $\beta_{j_{x+}}$, with the swapped proposed values retaining their spatial dependency.

In Steps 3 and 4, we again use Metropolis updating instead of Gibbs sampling. A closed form of the full conditional distribution of λ_1 and λ_2 is needed to use Gibbs sampling, but the conditionals of λ_1 and λ_2 are complex.

3.4. Connection with Bayesian regularized regression

In this subsection, we discuss various Bayesian regularized regression approaches and compare them with our proposed method. Starting from the pioneering work of ridge (Hoerl and Kennard, 1970) and lasso regression (Tibshirani, 1996), there has been extensive research in lasso-type estimation (Tibshirani et al., 2005; Tibshirani and Taylor, 2011). The generalized lasso-type model can be written as

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \left\{ (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 h_1(\boldsymbol{\beta}) + \lambda_2 h_2(\boldsymbol{\beta}) \right\}, \tag{13}$$

where the second term with h_1 is the penalty function to regularize the sparsity of β_j, and the third term with h_2 is used to impose spatial dependency between β_j and $\beta_{k \sim j}$.

From Bayesian perspectives, Tibshirani (1996) showed that the Bayes posterior mode of $\boldsymbol{\beta}$ under the Laplace prior distribution is the solution of (13) with $h_1(\boldsymbol{\beta}) = \sum_{j=1}^{p} |\beta_j|$ and $\lambda_2 = 0$. Park and Casella (2008) proposed the Bayesian lasso, which introduces the closed form of full conditional for Gibbs sampling, estimates the value of λ_1 without cross validation, and incorporates the uncertainty in estimating λ_1 as well as $\boldsymbol{\beta}$. Extending these studies, Kyung et al. (2010) suggested Bayesian hierarchical models and their Gibbs samplers for generalized lasso estimation, consisting of the group lasso, fused lasso, and elastic net.

In our pollution source identification problems, we consider Bayesian versions of regularized regression approaches: *Bayesian lasso* of Park and Casella (2008) whose posterior mode is identical to the solution of (13) where h_1 is an L_1 norm, *Bayesian ridge* where h_1 is a L_2 norm, and *Bayesian fused lasso* of Kyung et al. (2010) in which both h_1 and h_2 are L_1 norms. Under the Bayesian regularized regression framework, our hierarchical model introduced in Section 3.2 can be viewed as a Bayesian regularized regression. The conditional posterior distribution of $\boldsymbol{\beta}$ defined in (11) and (12) is written as

$$f(\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2, \lambda_1, \lambda_2) \propto \prod_{j=1}^{p} \mathbb{I}(\beta_j \geq 0)$$

$$\cdot \exp\left[-\frac{1}{2\sigma^2} \left\{ (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 \sum_{j=1}^{p} \beta_j^2 + \lambda_2 \left(\sum_{j \in \Omega_{x+}} (\beta_{j_{x+}} - \beta_j)^2 + \sum_{j \in \Omega_{y-}} (\beta_{j_{y-}} - \beta_j)^2 \right) \right\} \right]. \tag{14}$$

The posterior mode of $\boldsymbol{\beta}$ from (14) is the solution to the generalized ridge formulation

$$\min_{\{\beta_j \geq 0, \, \forall j\}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1 \sum_{j=1}^{p} \beta_j^2 + \lambda_2 \left(\sum_{j \in \Omega_{x+}} (\beta_{j_{x+}} - \beta_j)^2 + \sum_{j \in \Omega_{y-}} (\beta_{j_{y-}} - \beta_j)^2 \right), \tag{15}$$

which is equivalent to (13) with h_1 and h_2 being L_2 norms, after including the non-negativity constraint on the β_j values. Because of its spatial nature and association with the ridge regression, we call our proposed model the *Bayesian spatial ridge* model.

In our simulation study, each method and its sampling scheme is modified to incorporate the non-negativity constraint of β_j imposed for the emission identification problem. Details of MCMC implementation for each method are provided in Appendix A.

3.5. A dissection of the spatial structure h_2

The L_2 norm of the difference $\sum_j (\beta_{j+1} - \beta_j)^2$ in (15) can be interpreted as a Gibbs distribution prior (see, for example, Banerjee et al., 2014, Chapter 4.2) which arises from commonly used Gaussian Markov random field (GMRF) models. To see this, the adjacency term of the exponent in (14) can be viewed as a prior density for $\boldsymbol{\beta}$, given by

$$f(\boldsymbol{\beta}) \propto \exp\left\{ -\frac{\lambda_2}{2\sigma^2} \sum_{j<k} (\beta_j - \beta_k)^2 \, \mathbb{I}(j \sim k) \right\}, \tag{16}$$

where $j \sim k$ denotes the adjacency of j and k. By Brook's lemma (Brook, 1964), it can be shown that this is equivalent to a GMRF model defined as

$$\beta_k | \beta_j, j \neq k \sim N \left(\sum_{j \sim k} \frac{\beta_j}{n(k)}, \frac{\sigma^2}{n(k)\lambda_2} \right), \tag{17}$$

where $n(k)$ is the number of neighbors of k. Hence one can view our Bayesian spatial ridge model given in (11)–(12) as a GMRF model in (17) equipped with the additional shrinkage prior enforcing *spatial sparsity* governed by λ_1 and the *non-negative* constraint imposed by the physics principle.

One may be interested in comparing our proposed model with the Bayesian fused lasso with the L_1 norm for the function h_2. In the simulation study, we tried fitting the fused lasso model using the sampling method suggested by Kyung et al. (2010), but could not obtain a result that is meaningfully interpretable. As a Bayesian version of the fused lasso, Kyung et al. (2010) suggests using the hierarchical model with the likelihood in (11) and

$$\boldsymbol{\beta} | \sigma^2, \boldsymbol{T}, \boldsymbol{W} \sim N(\boldsymbol{0}, \sigma^2 (\boldsymbol{T} + \boldsymbol{W})^{-1}), \tag{18}$$

$$f(\tau_j^2 | \lambda_1) = \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2 \tau_j^2}{2} \right), j = 1, \ldots, p, \tag{19}$$

$$f(w_j^2 | \lambda_2) = \frac{\lambda_2^2}{2} \exp\left(-\frac{\lambda_2^2 w_j^2}{2} \right), j = 1, \ldots, p-1, \tag{20}$$

where $\boldsymbol{T}$ is a diagonal matrix whose jth diagonal element is τ_j^{-2} and $\boldsymbol{W} = \{W_{j,l}, j, l = 1, \ldots, p\}$ is a matrix whose elements are given by

$$W_{j,j} = \begin{cases} w_j^{-2} & \text{if } j = 1, \\ w_{j-1}^{-2} + w_j^{-2} & \text{if } 2 \leq j \leq p-1, \\ w_{j-1}^{-2} & \text{if } j = p, \end{cases}$$

and $W_{j,j+1} = W_{j+1,j} = -w_j^{-2}$ for $j = 1, \ldots, p-1$. They also introduce the Gibbs sampler (full conditionals) for the Bayesian fused lasso, which updates each w_j^{-2} via the inverse Gaussian distribution. We found that the determinant $|\boldsymbol{T} + \boldsymbol{W}|$ in (18) is not factorized to the products of τ_j^2 and w_j^2 unlike the Bayesian lasso with $|\boldsymbol{T}| = \prod_j \tau_j^{-2}$. Thus, Kyung et al. (2010)'s sampling method does not lead to a posterior distribution for the model of (18)–(20). For a similar reason, it is unclear that the posterior mode of their model is the solution of the original fused lasso of Tibshirani et al. (2005) because τ_j^2 and w_j^2 cannot be integrated out from the joint distribution. We refer readers to Appendix A.3.1 for more discussion on this topic.

4. Numerical study

We conduct simulation studies to demonstrate the performance of our approach in estimating pollution emission sources. Simulation is particularly useful in our application, as the estimated location and contribution of pollution sources from our methodology cannot be validated using measurement data sets. Such data will be extremely costly to obtain, if not impossible. Thus, we create a pseudo-real observational data set by exploiting the real wind data set available from NWP and hypothetical source information, and compare how our methodology performs under such settings.

4.1. Simulation setting

We use the time grid $\delta_t = 1/60$ h and the spatial grid $\delta_s = 3.0$ km for numerical discretization. Wind data used in this study are imported from the IBM *Physical Analytics Integrated Data Repository* (PAIRS, Klein et al., 2015). The wind data covers 72 h starting from the midnight of August 1, 2016 ($m = 1, \ldots, 4320$) for 150×150 km^2 areas in the Midwest region ($a, b = 1, \ldots, 50$ with an equally-spaced grid) of the United States, including parts of Ohio, Indiana, Kentucky, and West Virginia. The original unit (m/s) of wind data is converted to km/h. We assume that there are 144 monitoring stations available, $i = 1, \ldots, 144$.

We assume an emission field $\{\theta_{(a,b)}\}$ with unit $\mu g/(m^3 h)$, which is equivalent to $\boldsymbol{\beta}$ of size $p = 2500$ after vectorization, as a smooth surface with two modes shown in Fig. 2(a). Given the emission field and wind data, the forward dispersion model in (7) with the diffusion parameter $\kappa_0 = 10$ and the decay parameter $\eta_0 = 0.1$ generates the simulated concentration field $\{z_{(a,b)}^{(m)}\}$ using the finite difference method. Fig. 2(b)–(d) display snapshots of the simulated concentration field every 24 h after the initial emission, i.e., $\{z_{(a,b)}^{(m)}\}$ with $m = 1440, 2880,$ and 4320, respectively.

We assume that ith monitoring station is located within $\boldsymbol{S}_i$, equivalently at discretized grid $\boldsymbol{s}_i = (a_i \delta_s, b_i \delta_s)$, and measures the average pollution during $T = (71, 72]$ in unit $\mu g/m^3$. The average concentration $\omega^F(\boldsymbol{S}_i, T)$ in (1)–(3) during T is approximated by $\tilde{\omega}_i^F$ in (8) with $A = B = 50$, $M = 4320$, $m_{start} = 4261$ and $m_{end} = 4320$. Some noise is inevitable when obtaining measurements as discussed in Section 3.1. We generate noisy simulation measurements by

$$y_i = \tilde{\omega}_i^F + \varepsilon_i^F \tag{21}$$

for $i = 1, \ldots, 144$.
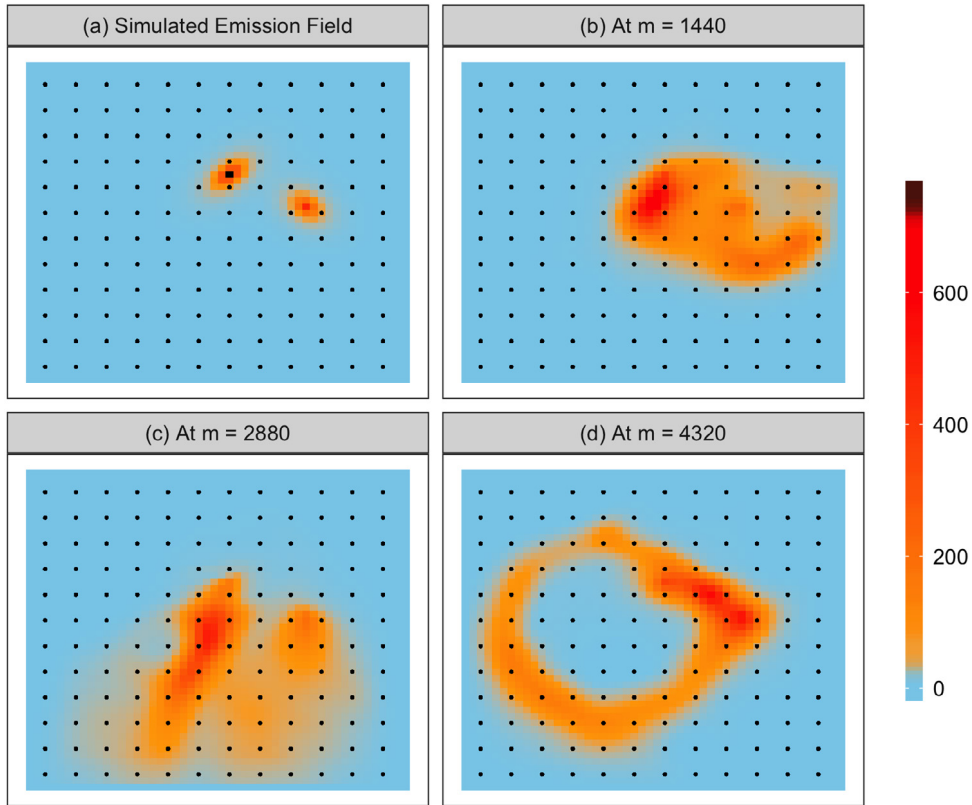
Fig. 2. Illustration of simulated data. (a) The simulation emission field $\{\theta_{(a,b)}\}$ defined on the grid $\mathbb{S}$ of 50×50 with $\delta_s = 3.0$ km. (b)–(d) The snapshots of simulated pollution concentration field $\{z_{(a,b)}^{(m)}\}$ captured at $m = 1440$, $m = 2880$, and $m = 4320$. The solid dots represent 144 monitoring stations, located over an equidistant grid every 12 km in both north and east directions, assumed in Scenarios 1–3.

Using the conjugate field $\{\chi_{i(a,b)}^{(m)}\}$ generated from the wind field and detector function $I[(a, b), m; (a_i, b_i), \{4261, \ldots, 4320\}]$, we generate a matrix $\boldsymbol{X}$ of size 144×2500 and set up the estimation model

$$y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i^B, \quad \varepsilon_i^B \overset{iid}{\sim} N(0, \sigma^2) \tag{22}$$

where $\boldsymbol{x}_i$ of length 2500 contains the information on backward flow from the ith monitoring station to each of 2500 grids. Our interest is to examine whether the backward estimation model in (22) can estimate the emission field, $\boldsymbol{\beta}$, using the forward model output $\tilde{w}_i^F$ corrupted by ε_i^F. As explained in Section 3.1, the discrepancy between $\tilde{\omega}_i^F$ and $\tilde{\omega}_i^B$ is only the discretization error which is often negligible.

We first conduct a simulation study under a strong assumption on the data generation scenario as follows.

Scenario 1 The simulated concentration field $z_{(a,b)}^{(m)}$ is generated from the discretized forward model in (7) with a fixed $\theta_{(a,b)}^{(m)} = \theta_{(a,b)}$. The measurement noise added on $\tilde{\omega}_i^F$ is assumed to be $\varepsilon_i^F \overset{iid}{\sim} N(0, \sigma^2)$ with $\sigma = 0.5$. The monitoring stations are located over an equidistant grid every 12 km distance, in both north and east directions, as depicted in Fig. 2(a)–(d) by the solid dots.

We fit our Bayesian spatial ridge model described in Section 3.2 to the simulation data to estimate the emission field $\boldsymbol{\beta}$. The hyperparameters are set as $(c_1, d_1) = (1.00, 1.78)$ as in Park and Casella (2008), $(c_2, d_2) = (1.00, 1.78)$, and $\mu_\sigma = 1.0$. Fig. 3(b) shows the estimated emission field under Scenario 1, which is the posterior means of $\{\beta_j\}$ obtained from 1000 iterations of the MCMC after 1000 iterations burn-in. The estimated field is slightly flatter than the true emission field in Fig. 3(a), but it successfully finds the overall location of the mode and yields a reasonable spatial structure given the small n and large p setting.

The Bayesian inference also enables the uncertainty quantification of the estimated field by finding the posterior standard deviation of the MCMC samples as shown in Fig. 3(c). By dividing the posterior standard deviation by the posterior mean, we compute the coefficient of variation (CV), which gives the uncertainty of the estimated emission field as the relative precision. Fig. 3(d) shows the calculated CV; the blanked areas have CVs larger than 1.00 because of too small posterior means. From
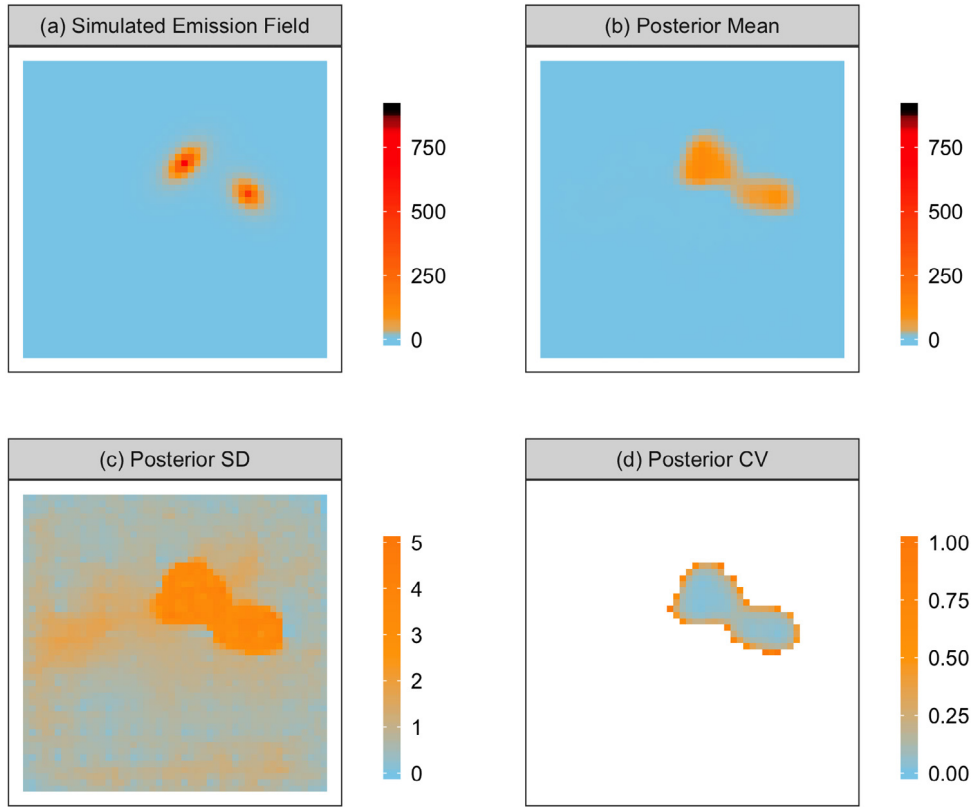
Fig. 3. Results from the MCMC samples of $\{\beta_j\}$, equivalently those of $\{\theta_{(a,b)}\}$, with the Bayesian spatial ridge in Scenario 1. (a) The simulation emission field; (b) the posterior means as the estimated emission field; (c) the posterior standard deviations; and (d) the coefficient of variation (CV), where the white spaces represent the locations whose mean estimates that are too close to zero result in a CV higher than 100%.
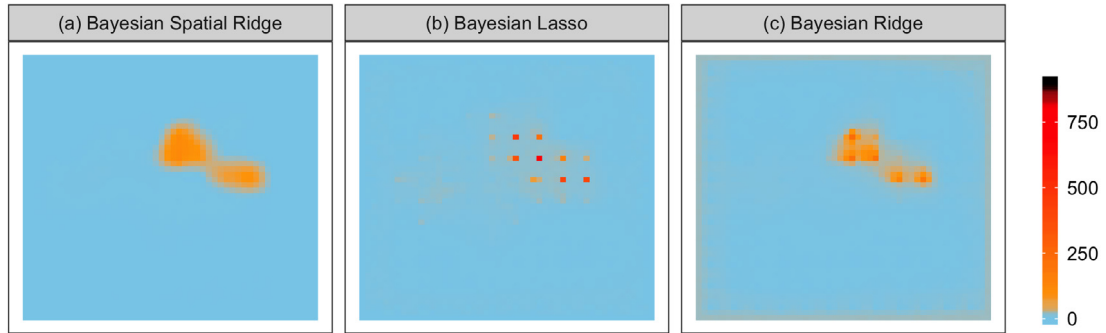
Fig. 4. Estimated emission fields in Scenario 1 with the different methods; (a) the proposed method (Bayesian spatial ridge), (b) Bayesian lasso, and (c) Bayesian ridge. Each plot shows the posterior mean of emission field obtained from 1000 iterations of the MCMC after 1000 burn-in iterations.

the figure, we have confirmed that the non-blanked areas with sufficiently high values of the posterior mean (i.e., the areas with meaningful pollution contribution) have reasonably small CVs, with an inter-quartile range of [0.09, 0.32].

Fig. 4 shows the estimated emission fields from our proposed method (Bayesian spatial ridge), Bayesian lasso, and Bayesian ridge under Scenario 1. The Bayesian ridge gives more smooth estimated field than the Bayesian lasso, which results in a better model fit because the true emission field is assumed to be smooth. Due to the spatial structure imposed by the function h_2, our Bayesian spatial ridge gives the estimated field more smooth than the other two methods that do not have spatial considerations. We also considered the Bayesian fused lasso model, but we found that the approach is not feasible as previously explained in Section 3.5.

For numerical evaluation, we replicate simulations 20 times under Scenario 1. Since $\tilde{\omega}_i^F$ is a deterministic function of $\{\theta_{(a,b)}\}$, the wind data, and the fixed values of κ_0 and η_0, only ε_i^F needs to be generated and added to y_i for each replicate.

Table 1
The summarized root mean integrated squared errors (rMISE) of pollution estimates for different methods, where the numbers in parenthesis present the standard deviation calculated from 20 replicates .

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Station locations	Regular	Regular	Regular	Irregular
Emission $\theta_{(a,b)}^{(m)}$	Constant	Varying	Constant	Constant
Noise ε_i	Independent	Independent	Spatially cor.	Independent
Bayesian Spatial Ridge	19.77	19.76	20.01	15.91
	(0.03)	(0.03)	(0.20)	(0.85)
Bayesian Lasso	33.34	33.32	34.11	33.91
	(2.41)	(1.64)	(2.02)	(4.95)
Bayesian Ridge	23.21	23.24	23.51	22.51
	(0.05)	(0.05)	(0.47)	(0.70)

We calculated the root mean integrated squared error (rMISE), $\sqrt{\sum_{j=1}^{p}(\hat{\beta}_j - \beta_j)^2/p}$, for the estimators $\hat{\boldsymbol{\beta}}$ for each method and replicate. The leftmost column of Table 1 shows the rMISE for Scenario 1 averaged over 20 simulations along with the corresponding standard deviations. The smaller rMISE value of our proposed method coincides with its visibly better model fit shown in Fig. 4.

4.2. Additional scenarios

The assumption regarding the nature of the pollution generation and its measurements in Section 4.1 is relatively restrictive. We conduct a simulation study under the following additional scenarios to further examine the behavior of the proposed method. The simulation settings used in Scenario 1 remain unchanged unless otherwise noted.

Scenario 2 The simulated concentration field $z_{(a,b)}^{(m)}$ is generated from the discretized forward model in (7) with varying $\theta_{(a,b)}^{(m)} = \theta_{(a,b)}U_{a,b}$, where $U_{a,b} \overset{iid}{\sim} \text{Unif}(0.5, 1.5)$.

Scenario 3 The measurement noise added on $\tilde{\omega}_i^F$ is assumed to be generated from a Gaussian process with the exponential correlation, $\text{cor}(\varepsilon_i^F, \varepsilon_{i'}^F) = \exp(-d_{ii'}/10)$ and $\text{var}(\varepsilon_i^F) = 0.5$, where $d_{ii'}$ is the distance between $\boldsymbol{s}_i$ and $\boldsymbol{s}_{i'}$.

Scenario 4 The monitoring stations are located randomly over the spatial region. The 144 pixels are randomly drawn from the total of 2500 pixels, as shown in Fig. B.5 of Appendix B.

These additional scenarios are chosen from the statistical modeling perspective and practical consideration. Scenario 2 can show how the method is affected by the variation in emission field propagated over the dispersion process. In this scenario, we keep using the conjugate field $\boldsymbol{x}_i$ from Scenario 1, and there is a discrepancy between $\tilde{\omega}_i^F$ and $\tilde{\omega}_i^B = \boldsymbol{x}_i^\top \boldsymbol{\beta}$ in addition to the (negligible) discretization error. Scenario 3 can provide an insight on the impact of error model misspecification. Independent normal error assumption for ε_i^B in (22) of Scenario 1 is violated because we introduce a spatial dependence in ε_i^F. Scenario 4 considers more realistic situation for the location of monitoring stations, which are often determined by external factors, such as administrative convenience. As the location of monitoring stations is changed from Scenario 1, the conjugate field and the values of $\boldsymbol{x}_i$ in (22) are recalculated accordingly. With the new values of $\boldsymbol{x}_i$, the estimation model is correctly specified for the simulation measurement $\boldsymbol{y}_i$, i.e., $\tilde{\omega}_i^F$ and $\tilde{\omega}_i^B$ have very similar values.

The results are summarized in Table 1. In comparison, our proposed Bayesian spatial ridge model performs consistently better than all other models we considered. Scenarios 2 and 3 suggest that the backward estimation models are robust to misspecification in the (uniform) random variation in source emission and to the error misspecification, respectively. Comparing Scenarios 1 and 4, Bayesian spatial ridge and Bayesian ridge models perform better with irregularly located monitoring stations than the regularly-located stations. Our conjecture on this phenomena is that the randomly generated points in Scenario 4 may provide better information on the spatial dependence of the emission surface by providing more diverse cases of spatial distances, while the regularly spaced points in Scenario 1 only provide the limited cases.

We note some limitations of our simulation studies. First, Scenarios 2 and 3 are based on the model misspecification only up to moderate extent. With much larger variation in $\theta_{(a,b)}$ or more critical dependence structure in ε_i values, the backward estimation approach with any regularized model considered here would not work well. Second, our simulation suggests an interesting finding in Scenario 4, but does not fully reveal the underlying mechanism. This topic is interesting and deserves a further study. Third, as we pointed out earlier in this section, the data for true pollution sources are difficult to obtain. Our simulation study suggests that our methodology can provide a helpful reference, but should be used with a reservation when used to estimating the real pollution levels.

5. Conclusion

In this study we have proposed a physics-motivated statistical framework for detecting the spatial locations of sources of pollutant emission. Our approach integrates the fundamental convection–diffusion equations and a ridge-regression type shrinkage estimation procedure. We also have proposed a sampling strategy to efficiently handle a few thousand grid points, while incorporating spatial dependence and non-negativity. The simulation example shows that the approach can accurately locate the sources of the pollutant emissions based on a limited number of observation data.

The current approach can be extended and improved in several different directions. First, one can incorporate parametric uncertainty in the physics model, by treating the diffusion and decay parameter as unknowns. Some prior knowledge on these values may be still needed to avoid identifiability issues. Second, it may be interesting to incorporate some structural model errors (a data-model discrepancy), in the form of spatially-correlated error terms in the regression model. However, this may also cause identifiability issues, as pointed by the computer model calibration literature (e.g., Brynjarsdóttir and O'Hagan, 2014). Finally, the physics model component can be modified for other related problems in environmental sciences, such as ocean pollutant detection (Solberg, 2012) or wildfire modeling problems (Baptiste Filippi et al., 2009).

Appendix A. MCMC Implementation

This section provides details of MCMC implementation for three alternative methods in Section 3.4 of the main text.

A.1. Bayesian lasso with non-negativity constraints

The Bayesian lasso with the non-negativity constraint is written as

$$\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim \mathrm{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}),$$

$$f(\boldsymbol{\beta}|\sigma^2, \boldsymbol{T}) \propto \mathrm{N}(\boldsymbol{\beta}; \boldsymbol{0}, \sigma^2\boldsymbol{T}^{-1}) \prod_{j=1}^{p} \mathbb{I}(\beta_j \geq 0), \tag{A.1}$$

$$f(\tau_j^2|\lambda_1) = \frac{\lambda_1^2}{2} e^{-\frac{\lambda_1^2 \tau_j^2}{2}}, \quad f(\sigma^2) \propto \sigma^{-2}, \quad \lambda_1^2 \sim \mathrm{Gamma}(c_1, d_1) \tag{A.2}$$

where $\boldsymbol{T}$ is a diagonal matrix whose jth diagonal element is τ_j^{-2} and Gamma(a, b) denotes the Gamma distribution whose mean is a/b.

Step 1. Metropolis updating for β_j

(a) For each $j = 1, \ldots, p$, propose b_j from $\log\mathrm{N}(\log\beta_j, s_\sigma^2)$ with step-size s_σ, then update $\beta_j = b_j$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \boldsymbol{b}^\top\boldsymbol{T}\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\beta}^\top\boldsymbol{T}\boldsymbol{\beta}\right\}\right]} \frac{\log\mathrm{N}(\beta_j; \log b_j, s_\sigma^2)}{\log\mathrm{N}(b_j; \log\beta_j, s_\sigma^2)}.$$

where $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, b_j, \beta_j, \ldots, \beta_p)^\top$.

(b) For each neighbors (j, j_{x+}) or (j, j_{y-}), propose $\boldsymbol{b}$ where β_j and $\beta_{j_{x+}}$ are swapped, i.e., $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j_{x+}}, \beta_{j+1}, \ldots, \beta_{j_{x+}-1}, \beta_j, \beta_{j_{x+}+1}, \ldots, \beta_p)^\top$. Then, update $\boldsymbol{\beta} = \boldsymbol{b}$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \boldsymbol{b}^\top\boldsymbol{T}\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\beta}^\top\boldsymbol{T}\boldsymbol{\beta}\right\}\right]}.$$

Step 2. Gibbs sampling for σ^2

Draw σ^2 from the full conditional distribution,

$$\sigma^2|\boldsymbol{\beta}, \boldsymbol{T}, \lambda_1 \sim \mathrm{Inv\text{-}Gamma}\left(\frac{n+p-1}{2}, \frac{1}{2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\beta}^\top\boldsymbol{T}\boldsymbol{\beta}\right\}\right).$$

Step 3. Gibbs sampling for τ_j^{-2}

For each $j = 1, \ldots, p$, draw τ_j^{-2} from the full conditional distribution,

$$\tau_j^{-2}|\boldsymbol{\beta}, \sigma^2, \lambda_1 \sim \text{Inv-Gaussian}\left(\sqrt{\frac{\lambda_1^2\sigma^2}{\beta_j^2}}, \lambda_1^2\right).$$

Step 4. Gibbs sampling for λ_1^2

Draw λ_1^2 from the full conditional distribution,

$$\lambda_1^2|\boldsymbol{\beta}, \sigma^2, \boldsymbol{T} \sim \text{Gamma}\left(c_1 + p, d_1 + \frac{1}{2}\sum_{j=1}^{p}\tau_j^2\right).$$

A.2. Bayesian ridge with non-negativity constraints

The Bayesian ridge with the non-negativity constraint is written as

$$\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}, \sigma^2 \sim N(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}),$$

$$f(\boldsymbol{\beta}|\sigma^2, \lambda_1) \propto N(\boldsymbol{\beta}; \boldsymbol{0}, \sigma^2(\lambda_1\boldsymbol{I})^{-1})\prod_{j=1}^{p}\mathbb{I}(\beta_j \geq 0),$$

$$\sigma^2 \sim \text{Inv-Gamma}(a_\sigma, b_\sigma), \quad \lambda_1 \sim \text{Gamma}(c_1, d_1)$$

Step 1. Metropolis updating for β_j

(a) For each $j = 1, \ldots, p$, propose b_j from $\log N(\log\beta_j, s_\sigma^2)$, then update $\beta_j = b_j$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \lambda_1\boldsymbol{b}^\top\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1\boldsymbol{\beta}^\top\boldsymbol{\beta}\right\}\right]}\frac{\log N(\beta_j; \log b_j, s_\sigma^2)}{\log N(b_j; \log\beta_j, s_\sigma^2)}.$$

where $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, b_j, \beta_j, \ldots, \beta_p)^\top$.

(b) For each neighbors (j, j_{x+}) or (j, j_{y-}), propose $\boldsymbol{b}$ where β_j and $\beta_{j_{x+}}$ are swapped, i.e., $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j_{x+}}, \beta_{j+1}, \ldots, \beta_{j_{x+}-1}, \beta_j, \beta_{j_{x+}+1}, \ldots, \beta_p)^\top$. Then, update $\boldsymbol{\beta} = \boldsymbol{b}$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}) + \lambda_1\boldsymbol{b}^\top\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1\boldsymbol{\beta}^\top\boldsymbol{\beta}\right\}\right]}.$$

Step 2. Gibbs sampling for σ^2

Draw σ^2 from the full conditional distribution,

$$\sigma^2|\boldsymbol{\beta}, \lambda_1 \sim \text{Inv-Gamma}\left(a_\sigma + \frac{n+p}{2}, b_\sigma + \frac{1}{2}\left[(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda_1\boldsymbol{\beta}^\top\boldsymbol{\beta}\right]\right).$$

Step 3. Gibbs sampling for λ_1

Draw λ_1 from the full conditional distribution,

$$\lambda_1|\boldsymbol{\beta}, \sigma^2 \sim \text{Gamma}\left(c_1 + \frac{p}{2}, d_1 + \frac{1}{2\sigma^2}\boldsymbol{\beta}^\top\boldsymbol{\beta}\right).$$

A.3. Bayesian fused lasso with non-negativity constraints

We ran the following Gibbs sampler of Kyung et al. (2010, p. 402), with modification for non-negativity constraints in the pollution source identification problem. However, we could not get the posterior mean of $\boldsymbol{\beta}$ which can be meaningfully interpreted as parameter estimates of our interest.

Step 1. Metropolis updating for β_j

(a) For each $j = 1, \ldots, p$, propose b_j from $q(b_j|\beta_j) \sim \text{logN}(\log \beta_j, s_\sigma^2)$. Then, update $\beta_j = b_j$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{Xb})^\top(\boldsymbol{y} - \boldsymbol{Xb}) + \boldsymbol{b}^\top(\boldsymbol{T} + \boldsymbol{W})\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X\beta})^\top(\boldsymbol{y} - \boldsymbol{X\beta}) + \boldsymbol{\beta}^\top(\boldsymbol{T} + \boldsymbol{W})\boldsymbol{\beta}\right\}\right]} \frac{\text{logN}(\beta_j; \log b_j, s_\sigma^2)}{\text{logN}(b_j; \log \beta_j, s_\sigma^2)}.$$

where $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, b_j, \beta_{j+1}, \ldots, \beta_p)^\top$, $\boldsymbol{T}$ is a diagonal matrix whose jth diagonal element is τ_j^{-2} and $\boldsymbol{W} = \{W_{j,l}, j, l = 1, \ldots, p\}$ is set as

$$W_{j,j} = \omega_{j_{x-}}^{-2} + \omega_{j_{y+}}^{-2} + \omega_j^{-2}[\mathbb{I}(j \in \Omega_{x+}) + \mathbb{I}(j \in \Omega_{y-})] \text{ and}$$

$$W_{j,l} = \begin{cases} -\omega_l^{-2} & \text{if } l = j_{x-} \text{ or } j_{y+}, \\ -\omega_j^{-2} & \text{if } l = j_{x+} \text{ or } j_{y-}, \\ 0 & \text{otherwise.} \end{cases}$$

(b) For each neighbors (j, j_{x+}) or (j, j_{y-}), propose $\boldsymbol{b}$ where β_j and $\beta_{j_{x+}}$ are swapped, i.e., $\boldsymbol{b} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j_{x+}}, \beta_{j+1}, \ldots, \beta_{j_{x+}-1}, \beta_j, \beta_{j_{x+}+1}, \ldots, \beta_p)^\top$. Then, update $\boldsymbol{\beta} = \boldsymbol{b}$ with the acceptance probability of $\min(1, \alpha)$ where

$$\alpha = \frac{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{Xb})^\top(\boldsymbol{y} - \boldsymbol{Xb}) + \boldsymbol{b}^\top(\boldsymbol{T} + \boldsymbol{W})\boldsymbol{b}\right\}\right]}{\exp\left[-\frac{1}{2\sigma^2}\left\{(\boldsymbol{y} - \boldsymbol{X\beta})^\top(\boldsymbol{y} - \boldsymbol{X\beta}) + \boldsymbol{\beta}^\top(\boldsymbol{T} + \boldsymbol{W})\boldsymbol{\beta}\right\}\right]}.$$

Step 2. Gibbs sampling for σ^2

Update σ^2 from the full conditional distribution,

$$\sigma^2|\boldsymbol{\beta}, \boldsymbol{T}, \boldsymbol{W}, \lambda_1, \lambda_2 \sim \text{Inv-Gamma}\left(a_\sigma + \frac{n+p}{2}, \ b_\sigma + \frac{1}{2}\left[(\boldsymbol{y} - \boldsymbol{X\beta})^\top(\boldsymbol{y} - \boldsymbol{X\beta}) + \boldsymbol{\beta}^\top(\boldsymbol{T} + \boldsymbol{W})\boldsymbol{\beta}\right]\right).$$

Step 3. Gibbs sampling for τ_j^{-2}

For each $j = 1, \ldots, p$, update τ_j^{-2} from the suggested sampling distribution,

$$\tau_j^{-2}|\boldsymbol{\beta}, \sigma^2, \boldsymbol{W}, \lambda_1, \lambda_2 \sim \text{Inv-Gaussian}\left(\sqrt{\frac{\lambda_1^2 \sigma^2}{\beta_j^2}}, \lambda_1^2\right).$$

Step 4. Gibbs sampling for ω_j^{-2}

For each j such that $\mathbb{I}(j \in \Omega_{x+}) + \mathbb{I}(j \in \Omega_{y-}) > 0$, update ω_j^{-2} from the suggested sampling distribution,

$$\omega_j^{-2}|\boldsymbol{\beta}, \sigma^2, \boldsymbol{T}, \boldsymbol{\omega}_{-j}, \lambda_1, \lambda_2 \sim \text{Inv-Gaussian}\left(\sqrt{\frac{\lambda_2^2 \sigma^2}{(\beta_{j_{x+}} - \beta_j)^2 + (\beta_{j_{y-}} - \beta_j)^2}}, \lambda_2^2\right).$$

Step 5. Gibbs sampling for λ_1^2

Update λ_1^2 from the suggested sampling distribution,

$$\lambda_1^2|\boldsymbol{\beta}, \sigma^2, \boldsymbol{T}, \boldsymbol{W}, \lambda_2 \sim \text{Gamma}\left(c_1 + p, d_1 + \frac{1}{2}\sum_{j=1}^{p} \tau_j^2\right).$$

Step 6. Gibbs sampling for λ_2^2

Update λ_2^2 from the suggested sampling distribution,

$$\lambda_2^2|\boldsymbol{\beta}, \sigma^2, \boldsymbol{T}, \boldsymbol{W}, \lambda_1 \sim \text{Gamma}\left(c_2 + p_2, d_2 + \frac{1}{2}\sum_{j \in \Omega_2} \omega_j^2\right).$$

where $\Omega_2 = \{j : \mathbb{I}(j \in \Omega_{x+}) + \mathbb{I}(j \in \Omega_{y-}) > 0\}$ and $p_2 = |\Omega_2|$.

A.3.1. Factorization of $(\boldsymbol{T} + \boldsymbol{W})$ in Bayesian fused lasso

This section introduces a toy example to compare the factorization of covariance matrix of $\boldsymbol{\beta}$ in Bayesian lasso and Bayesian fused lasso. For simplicity, we assume $p = 2$, i.e., $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top$ and disregard the nonnegative constraints which we imposed in the pollution detection problem.

Bayesian lasso

In this simple setting, the prior distributions of the Bayesian lasso in Eqs. (A.1) and (A.2) are written by

$$
\begin{aligned}
& f(\boldsymbol{\beta}, \tau_1^2, \tau_2^2 | \lambda_1, \sigma^2) \\
&= |2\pi\sigma^2 \boldsymbol{T}^{-1}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\boldsymbol{\beta}^\top \boldsymbol{T}\boldsymbol{\beta}\right) \cdot \prod_{j=1}^{p} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2 \tau_j^2}{2}\right) \\
&= (2\pi\sigma^2)^{-1} \left| \begin{matrix} \tau_1^{-2} & 0 \\ 0 & \tau_2^{-2} \end{matrix} \right|^{\frac{1}{2}} \exp\left(-\frac{\beta_1^2}{2\sigma^2\tau_1^2} - \frac{\beta_2^2}{2\sigma^2\tau_2^2}\right) \cdot \prod_{j=1}^{2} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2 \tau_j^2}{2}\right) \\
&= \prod_{j=1}^{2} (2\pi\sigma^2\tau_j^2)^{-\frac{1}{2}} \exp\left(-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right) \cdot \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2 \tau_j^2}{2}\right) \\
&= \prod_{j=1}^{2} \left[\sqrt{\frac{\eta}{2\pi}} (\tau_j^{-2})^{-\frac{3}{2}} \exp\left(-\frac{\eta(\tau_j^{-2} - \mu_j)^2}{2\mu_j^2 \tau_j^{-2}}\right) \tau_j^{-4} \right] \cdot \left[\frac{\lambda_1}{2\sigma} \exp\left(-\frac{\lambda_1 |\beta_j|}{\sigma}\right) \right]
\end{aligned}
\tag{A.3}
$$

where $\mu_j = \sqrt{\lambda_1^2 \sigma^2 / \beta_j^2}$ and $\eta = \lambda_1^2$.

When $1/\tau_j^2$ follows the inverse-Gaussian distribution with parameters μ_j and η, the density function of τ_j^2 is written as the first bracket of the last line above,

$$
f(\tau_j^2 | \beta_j, \lambda_1, \sigma^2) = \sqrt{\frac{\eta}{2\pi}} (\tau_j^{-2})^{-\frac{3}{2}} \exp\left(-\frac{\eta(\tau_j^{-2} - \mu_j)^2}{2\mu_j^2 \tau_j^{-2}}\right) \tau_j^{-4}.
$$

By integrating out τ_j^2 with $f(\tau_j^2 | \beta_j, \lambda_1, \sigma^2)$, for $j = 1, 2$, from $f(\boldsymbol{\beta}, \tau_1, \tau_2 | \lambda_1, \sigma^2)$, we can derive

$$
f(\beta_j | \lambda_1, \sigma^2) = \frac{\lambda_1}{2\sigma} \exp\left(-\frac{\lambda_1 |\beta_j|}{\sigma}\right), \quad j = 1, 2,
$$

which is the density function of the Laplace distribution with parameters 0 and σ/λ_1.

Bayesian fused lasso

The prior distributions of the Bayesian fused lasso, in Eqs. (16)–(18) of our main text, are written by

$$
\begin{aligned}
& f(\boldsymbol{\beta}, \tau_1^2, \tau_2^2, \omega_1^2 | \lambda_1, \lambda_2, \sigma^2) \\
&= |2\pi\sigma^2 (\boldsymbol{T} + \boldsymbol{W})^{-1}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\boldsymbol{\beta}^\top (\boldsymbol{T} + \boldsymbol{W})\boldsymbol{\beta}\right) \cdot \prod_{j=1}^{2} f(\tau_j^2 | \lambda_1) \cdot f(\omega_1^2 | \lambda_2) \\
&= (2\pi\sigma^2)^{-1} |\boldsymbol{T} + \boldsymbol{W}|^{\frac{1}{2}} \exp\left(-\frac{\beta_1^2}{2\sigma^2\tau_1^2} - \frac{\beta_2^2}{2\sigma^2\tau_2^2} - \frac{(\beta_1 - \beta_2)^2}{2\sigma^2\omega_1^2}\right) \cdot \prod_{j=1}^{2} f(\tau_j^2 | \lambda_1) \cdot f(\omega_1^2 | \lambda_2) \\
& \quad \text{where } |\boldsymbol{T} + \boldsymbol{W}| = \left| \begin{matrix} \tau_1^{-2} + \omega_1^{-2} & -\omega_1^{-2} \\ -\omega_1^{-2} & \tau_2^{-2} + \omega_1^{-2} \end{matrix} \right| = \tau_1^{-2}\tau_2^{-2} + \omega_1^{-2}(\tau_2^{-2} + \tau_2^{-2}) \\
&= |\boldsymbol{T} + \boldsymbol{W}|^{\frac{1}{2}} \left[\prod_{j=1}^{2} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{\beta_j^2}{2\sigma^2\tau_j^2}\right) f(\tau_j^2 | \lambda_1) \right] \left[\exp\left(-\frac{(\beta_1 - \beta_2)^2}{2\sigma^2\omega_1^2}\right) f(\omega_1^2 | \lambda_2) \right]
\end{aligned}
\tag{A.4}
$$

Eq. (A.4) looks similar to (A.3) of Bayesian lasso. However, $|\boldsymbol{T} + \boldsymbol{W}|$ is not partitioned into the functions of each τ_j^2, so finding the closed form of their marginal conditional distribution, $f(\tau_j^2 | \boldsymbol{\beta}, \lambda_1, \lambda_2, \sigma^2)$ is not straightforward. Because $|\boldsymbol{T} + \boldsymbol{W}|^{1/2} \neq \tau_1^{-1}\tau_2^{-1}\omega_1^{-1}$, with some algebras we can check that Eq. (A.4) is not identical to the product of the conditional
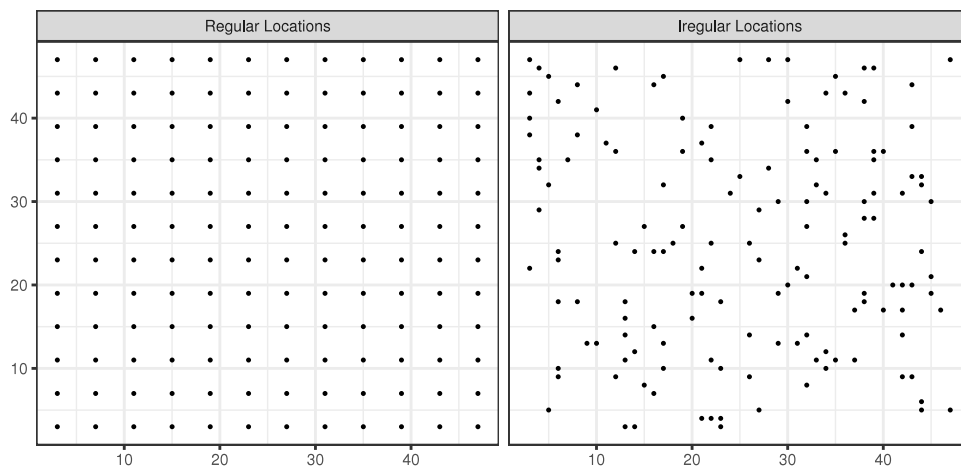
Fig. B.5. Vis-á-vis comparison of the locations of 144 monitoring stations in Scenarios 1–4 in Section 4. The left panel shows the monitoring stations located at every 12 km distance in north and east directions in Scenarios 1–3, and the right panel shows randomly selected 144 locations from the total of 2500 pixels in Scenario 4.

distributions of Kyung et al. (2010, p. 402),

$$
\left[\prod_{j=1}^{2} \sqrt{\frac{\eta}{2\pi}} \left(\tau_j^{-2}\right)^{-\frac{3}{2}} \exp\left(-\frac{\eta(\tau_j^{-2} - \mu_j)^2}{2\mu_j^2 \tau_j^{-2}}\right) \tau_j^{-4} \cdot \frac{\lambda_1}{2\sigma} \exp\left(-\frac{\lambda_1 |\beta_j|}{\sigma}\right) \right]
$$

$$
\cdot \sqrt{\frac{\zeta}{2\pi}} \left(\omega_1^{-2}\right)^{-\frac{3}{2}} \exp\left(-\frac{\zeta(\omega_1^{-2} - \nu)^2}{2\nu^2 \omega_1^{-2}}\right) \omega_1^{-4} \cdot \frac{\lambda_2}{2\sigma} \exp\left(-\frac{\lambda_2 |\beta_1 - \beta_2|}{\sigma}\right)
$$

$$
= \left(\tau_1^{-1} \tau_2^{-1} \omega_1^{-1}\right) \left[\prod_{j=1}^{2} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{\beta_j^2}{2\sigma^2 \tau_j^2}\right) f(\tau_j^2 | \lambda_1) \right] \left[\exp\left(-\frac{(\beta_1 - \beta_2)^2}{2\sigma^2 \omega_1^2}\right) f(\omega_1^2 | \lambda_2) \right]
$$

where $\mu_j = \sqrt{\lambda_1^2 \sigma^2 / \beta_j^2}$ for $j = 1, 2$, $\eta = \lambda_1^2$, $\nu = \sqrt{\lambda_2^2 \sigma^2 / (\beta_1 - \beta_2)^2}$, and $\zeta = \lambda_2^2$.

Even in this simple setting of $p = 2$, it is not easy to derive the joint conditional distribution of τ_1^2 and τ_2^2, i.e., $f(\tau_1^2, \tau_2^2 | \boldsymbol{\beta}, \lambda_1, \lambda_2, \sigma^2)$, which would be used to marginalize out the parameters to get the distributions of $f(\boldsymbol{\beta} | \lambda_1, \sigma^2, \sigma^2)$. Thus, it is unclear that Kyung et al. (2010)'s sampling method results in a posterior distribution corresponding to the Bayesian fused lasso model, as described in Eqs. (16)–(18) in our main text. For a similar reason, it is unclear that the posterior mode of their model is the solution of the original fused lasso of Tibshirani et al. (2005).

Appendix B. Figure for Scenario 4

See Fig. B.5.

References

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2014. Hierarchical Modeling and Analysis for Spatial Data, Second Edition. CRC Press.

Baptiste Filippi, J., Bosseur, F., Mari, C., Lac, C., Le Moigne, P., Cuenot, B., Veynante, D., Cariolle, D., Balbi, J.-H., 2009. Coupled atmosphere-wildland fire modelling. J. Adv. Modelling Earth Syst. 1 (4), 11.

Biegler, L., Biros, G., Ghattas, O., Heinkenschloss, M., Keyes, D., Mallick, B., Marzouk, Y., Tenorio, L., van Bloemen Waanders, B., Willcox, K., 2011. Large-Scale Inverse Problems and Quantification of Uncertainty. John Wiley & Sons, West Sussex, United Kingdon.

Brook, D., 1964. On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. Biometrika 51 (3), 481–483.

Brynjarsdóttir, J., O'Hagan, A., 2014. Learning about physical parameters: The importance of model discrepancy. Inverse Problems 30 (11), 114007.

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system. Appl. Mech. Rev. 59, 51–77.

Chan, S.T., Leach, M.J., 2007. A validation of FEM3MP with joint urban 2003 data. J. Appl. Meteorol. Climatol. 46 (12), 2127–2146.

Chang, W., Haran, M., Olson, R., Keller, K., 2015. A composite likelihood approach to computer model calibration with high-dimensional spatial data. Statist. Sinica 243–259.

Chang, C.-J., Joseph, V.R., 2014. Model calibration through minimal adjustments. Technometrics 56 (4), 474–482.

Christensen, W.F., Gunst, R.F., 2004. Measurement error models in chemical mass balance analysis of air quality data. Atmos. Environ. 38 (5), 733–744.

Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., Brode, R.W., 2005. AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. J. Appl. Meteor. 44 (5), 682–693.

Fast, J.D., Gustafson, W.I., Easter, R.C., Zaveri, R.A., Barnard, J.C., Chapman, E.G., Grell, G.A., Peckham, S.E., 2006. Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of houston using a fully coupled meteorology-chemistry-aerosol model. J. Geophys. Res.: Atmos. 111.

Higdon, D., Gattiker, J., Williams, B., Rightley, M., 2008. Computer model calibration using high-dimensional output. J. Amer. Statist. Assoc. 103 (482), 570–583.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67.

Hwang, Y., Barut, E., Yeo, K., 2018. Statistical-physical estimation of pollution emission. Statist. Sinica 28, 921–940.

Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63, 425–464.

Klein, L.J., Marianno, F.J., Albrecht, C.M., Freitag, M., Lu, S., Hinds, N., Shao, X., Rodriguez, S.B., Hamann, H.F., 2015. PAIRS: A scalable geo-spatial data analytics platform. In: 2015 IEEE International Conference on Big Data. IEEE, pp. 1290–1298.

Kyung, M., Gill, J., Ghosh, M., Casella, G., 2010. Penalized regression, standard errors, and Bayesian lassos. Bayesian Anal. 5, 369–412.

Moin, P., 2010. Fundamentals of engineering numerical analysis. Cambridge University Press.

Park, T., Casella, G., 2008. The Bayesian lasso. J. Amer. Statist. Assoc. 103 (482), 681–686.

Park, E.S., Guttorp, P., Henry, R.C., 2001. Multivariate receptor modeling for temporally correlated data by using MCMC. J. Amer. Statist. Assoc. 96 (456), 1171–1183.

Reid, J.S., Hyer, E.J., Johnson, R.S., Holben, B.N., Yokelson, R.J., Zhang, J., Campbell, J.R., Christopher, S.A., Girolamo, L.D., Giglio, L., Holz, R.E., Kearney, C., Miettinen, J., Reid, E.A., Turk, F.J., Wang, J., Xian, P., Zhao, G., Balasubramanian, R., Chew, B.N., Janjai, S., Lagrosas, N., Lestari, P., Lin, N.-H., Mahmud, M., Nguyen, A.X., Norris, B., Oanh, N.T., Oo, M., Salinas, S.V., Welton, E.J., Liew, S.C., 2013. Observing and understanding the Southeast Asian aerosol system by remote sensing: An initial review and analysis for the Seven Southeast Asian Studies (7SEAS) program. Atmos. Res. 122 (Supplement C), 403–468.

Solberg, A.H.S., 2012. Remote sensing of ocean oil-spill pollution. Proc. IEEE 100 (10), 2931–2945.

Strikwerda, J.C., 2004. Finite Difference Schemes and Partial Differential Equations. SIAM.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society, Series B. 58 (1), 267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. J. R. Statist. Soc. Ser. B 67 (1), 91–108.

Tibshirani, R.J., Taylor, J., 2011. The solution path of the generalized lasso. Ann. Statist. 39 (3), 1335–1371.

Williams, B., Christensen, W.F., Reese, C.S., 2011. Pollution source direction identification: embedding dispersion models to solve an inverse problem. Environmetrics 22 (8), 962–974.

World Health Organization, 2005. WHO air quality guidelines for paticularte matter, ozone, nitrogen dioxide and sulfur dioxide–global update (WHO/SDE/PHE/OEH/06.02). Air Quality Guidelines - Global Update 2005. World Health Organization.