Supplementary Material for
"A Mechanistic Model of Annual Sulfate Concentrations in the U.S."

October 4, 2020

Contents

1 Numerical Methods for Approximating Continuous Processes in Discrete Space

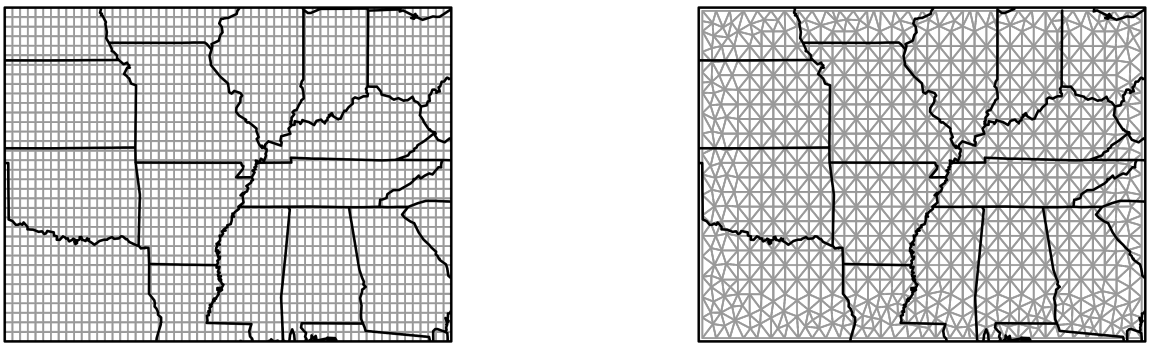
Consider the continuous space-time stochastic process introduced in Section 3.1,

$$dy_{\boldsymbol{s}}(t) = \Big(-\mathcal{A}_{\boldsymbol{s}}(\boldsymbol{\theta}) y_{\boldsymbol{s}}(t) + m_{\boldsymbol{s}}(\boldsymbol{\theta}) \Big) dt + \mathcal{B}_{\boldsymbol{s}}(\boldsymbol{\theta})\, \xi(\boldsymbol{s}, t), \tag{SM.1}$$

where $y_{\boldsymbol{s}}(t)$ ($\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^2$, $t \geq 0$) is the quantity of interest, $\mathcal{A}_{\boldsymbol{s}}$ is a linear operator, $m_{\boldsymbol{s}}$ denotes a source/sink at location $\boldsymbol{s}$, and $\mathcal{B}_{\boldsymbol{s}}\, \xi(\boldsymbol{s}, t)$ is a space-time white noise process, with process variance defined by real-valued function $\mathcal{B}_{\boldsymbol{s}}$. For convenience, we will assume that this process has constant variance, i.e., $\mathcal{B}_{\boldsymbol{s}} \equiv \sigma$ for all $\boldsymbol{s}$. As discussed in our manuscript, it is often useful to define a dynamic process as an SPDE in continuous space. However, the theoretical and computational limitations associated with working in continuous space-time necessitates an approximation of (SM.1) in discrete space. This supplementary document provides a brief overview of how we implement such an approximation. Particular attention is focused on the sulfate model presented in the manuscript (see Sections 2 and 4 of the manuscript for details).

1.1 Step One: Discretization

The numerical representation of (SM.1) in discrete space occurs in two steps. First, the continuous surface $\mathcal{D} \subset \mathbb{R}^2$ must be appropriately *discretized* to some finite collection of points, $\mathcal{S} = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k\}$. Second, the mathematical operator $\mathcal{A}$ is *approximated with a matrix operator*, $\boldsymbol{A}$, on the set of discretized points, $\mathcal{S}$. The resulting stochastic process, defined on discrete space $\mathcal{S}$, is a computationally convenient representation of the original dynamic process defined in continuous space.



(a) A 58×35 rectangular grid.

(b) Mesh created with Delaunay triangulation.

Figure 1: Example discretizations of the surface, $\mathcal{D}$, considered in the sulfate analysis, including (a) a simple rectangular grid, and (b) a more complex Delaunay triangulation (created with the `R-INLA` package [3]).

The discretization step is conceptually simple. A grid or polygonal *mesh* is overlaid on top of the continuous surface, $\mathcal{D}$, and the process dynamics are defined with respect to the discrete set of mesh cells. The structure of the mesh may vary, from a simple rectangular grid (Figure 1a) to more complex structures, such as those generated via Delaunay triangulation (Figure 1b). The choice of mesh structure is often dictated by the complexity of the process dynamics, the boundary

of the surface $\mathcal{D}$, and the chosen numerical approximation scheme (for example, finite element method implementations often use a triangulation method). Because of the variety of discretization schemes, and the large body of literature discussing their relative merits [2, 5, 1], we encourage the interested reader to explore implementation details tailored to their problem of interest. For now, we emphasize that the discretization step restricts the process of interest to some discrete set of points, $\mathcal{S} = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$, corresponding to mesh vertices and/or centroids.

1.2 Step Two: Numerical Approximation of $\mathcal{A}$

After a suitable mesh has been chosen, process dynamics are approximated on the discretized space, $\mathcal{S}$. A variety of numerical schemes have been developed to facilitate this approximation, including the finite difference method (FDM), finite volume method (FVM), and finite element method (FEM). Again, the literature and implementation for each of these methods are too extensive and problem-specific to permit a comprehensive introduction to these methods, and we encourage the interested reader to consult one of the following references for implementation details [2, 5, 1]. We instead emphasize that each method attempts to approximate a continuous mathematical operator, $\mathcal{A}$, with a discrete operator, $\boldsymbol{A}$, defined on the discretized space, $\mathcal{S}$. These approximations may occur via simple comparisons of neighboring values (such as first and second differences), or via the construction of appropriate basis functions; the implementation is dependent on the problem of interest and choice of numerical method.

As an example, consider the operator,

$$\mathcal{A} = \gamma\Delta + \alpha\, \boldsymbol{v}_i(t) \cdot \nabla - \delta, \tag{SM.2}$$

corresponding to the advection-diffusion process used to model sulfate concentration in Section 2.1 (Equation (1) of the manuscript). Here, $\gamma\Delta$ denotes homogeneous spatial diffusion with rate γ (note, $\Delta \equiv \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$); advection due to wind is defined by the advective derivative $\boldsymbol{v}_i(t) \cdot \nabla \equiv v_{x_1}\frac{\partial}{\partial x_1} + v_{x_2}\frac{\partial}{\partial x_2}$, where $\boldsymbol{v}_i(t) = (v_{x_1}, v_{x_2})'$ is the velocity vector at time t; and atmospheric deposition occurs at rate δ. A simple numerical approximation of this operator might be constructed in the following way.

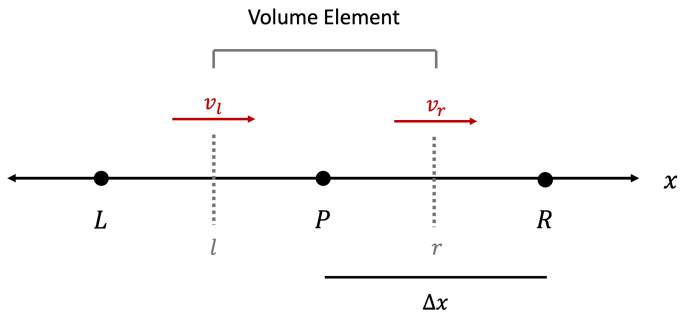


Figure 2: Example discretization (regular grid, one dimension) at three points: L, P, and R. The process $Y(x,t)$ evaluated at these points is used for the second-order central difference approximation to diffusion and the FVM upwind scheme for advection-diffusion.

Let $\mathcal{A}_1 \equiv \gamma\Delta$ denote homogeneous diffusion with rate γ, which for a process $Y(x,t)$ defined in

one-dimensional space x can be equivalently written as,

$$\mathcal{A}_1\, Y = \frac{\partial}{\partial x}\Big(\gamma\, \frac{\partial Y}{\partial x}\Big). \tag{SM.3}$$

How might we approximate this operator on some set of regular discrete points, $\mathcal{S}$? Consider three points, P, $R = P + \Delta x$, and $L = P - \Delta x$, and assume the process Y can be evaluated at each point at time t: i.e., $Y(P,t)$, $Y(R,t)$, and $Y(L,t)$. These correspond to three discrete points in our discretization, $\mathcal{S}$ (Figure 2). A second-order difference approximation of $\mathcal{A}_1$ can be constructed by first recognizing that the first derivative, $\frac{\partial Y(P,t)}{\partial x}$, can be approximated with simple linear differences,

$$\Big[\frac{\partial Y(P,t)}{\partial x}\Big]_R \approx \frac{Y(P+\Delta x,t) - Y(P,t)}{\Delta x} = \frac{Y(R,t) - Y(P,t)}{\Delta x} \tag{SM.4}$$

and

$$\Big[\frac{\partial Y(P,t)}{\partial x}\Big]_L \approx \frac{Y(P,t) - Y(P-\Delta x,t)}{\Delta x} = \frac{Y(P,t) - Y(L,t)}{\Delta x}. \tag{SM.5}$$

Then, using (SM.4) and (SM.5), we have

$$\begin{aligned}
\mathcal{A}_1\, Y(P,t) &\equiv \frac{\partial}{\partial x}\big(\gamma\, \frac{\partial Y(P,t)}{\partial x}\big) && \text{(SM.6)}\\
&\approx \big(\gamma\, \big[\frac{\partial Y(P,t)}{\partial x}\big]_R - \gamma\, \big[\frac{\partial Y(P,t)}{\partial x}\big]_L\big)/\Delta x && \text{(SM.7)}\\
&= \gamma\, \frac{Y(R,t) - 2Y(P,t) + Y(L,t)}{(\Delta x)^2}. && \text{(SM.8)}
\end{aligned}$$

Repeating this process for all points in $\mathcal{S}$, we can now approximate the operator $\mathcal{A}_1\, Y$ in discrete space as,

$$\boldsymbol{A}_1\, \boldsymbol{Y_t} = \gamma \begin{pmatrix} \ddots & \ddots & \ddots & & \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ Y(L,t) \\ Y(P,t) \\ Y(R,t) \\ \vdots \end{pmatrix}. \tag{SM.9}$$

Note that $\boldsymbol{A}_1$ is a sparse, banded matrix, which permits computationally efficient matrix operations [4] when working with the discrete space model.

The above discretization of a diffusion operator can be derived using either the finite difference method or the finite volume method [5]. However, the addition of advection due to wind necessitates a slightly more complex discretization scheme, as centered difference approximations become increasingly biased towards upstream process values as advective forces grow large [5]. Instead, we use an upwind differencing scheme to approximate advection-diffusion with the FVM. In short, for our three points, P, $R = P + \Delta x$, and $L = P - \Delta x$, we evaluate (or estimate) the wind velocity $v(x,t)$ along volume element faces l and r (i.e., $v_l \equiv v(l,t)$ and $v_r \equiv v(r,t)$; see Figure 2). Then, the operator

$$\mathcal{A}_2\, Y \equiv \gamma\Delta + \alpha\, \boldsymbol{v}_i(t) \cdot \nabla = \frac{\partial}{\partial x}\Big(\gamma\, \frac{\partial Y}{\partial x}\Big) + \alpha\, v(x,t)\frac{\partial Y}{\partial x} \tag{SM.10}$$

can be approximated with

$$\boldsymbol{A}_2\,\boldsymbol{Y_t} = \gamma \begin{pmatrix} \ddots & \ddots & \ddots & & \\ & a_L & -a_P & a_R & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ Y(L,t) \\ Y(P,t) \\ Y(R,t) \\ \vdots \end{pmatrix}, \tag{SM.11}$$

where a_L, a_P, and a_R are given in Table 1. Additional details of the FVM upwind differencing scheme can be found in Chapter 5 of Versteeg and Malalasekera [5].

Table 1: Matrix entries for the FVM upwind difference approximation of an advection-diffusion operator.

	Matrix Entries		
Edge Velocities	a_L	a_R	a_P
$v_l > 0,\ v_r > 0$	$\gamma + \alpha\ v_l$	γ	$2\gamma + \alpha(v_r - v_l)$
$v_l < 0,\ v_r < 0$	γ	$\gamma - \alpha\ v_r$	$2\gamma + \alpha(v_r - v_l)$

Finally, including atmospheric deposition in the process (i.e., $\mathcal{A} = \gamma\Delta + \alpha\ \boldsymbol{v}_i(t) \cdot \nabla - \delta$) only requires the subtraction of the deposition rate δ from the matrix diagonal of $\boldsymbol{A}_2$, i.e., $\mathcal{A} \approx \boldsymbol{A} = \boldsymbol{A}_2 - \delta\boldsymbol{I}$. Thus, the advection-diffusion-deposition process found in our manuscript has been approximated in one dimension; the extension of these methods to two dimensions is straightforward [5].

1.3 Atmospheric Sulfate Process

In summary, the approximation of a continuous process in discrete space occurs by first discretizing continuous space $\mathcal{D}$ to discrete space $\mathcal{S}$, and then approximating the continuous operator $\mathcal{A}$ with its discrete representation, $\boldsymbol{A}$, as determined by some numerical scheme such as the FDM, FVM, or FEM. In our manuscript, the continuous process was the region of the continental USA shown in Figure 3. We discretized this space with a 116×70 rectangular grid, represented with black dots in Figure 3. Then, using this discretization, we approximated the advection-diffusion-deposition process, $\mathcal{A} = \gamma\Delta + \alpha\ \boldsymbol{v}_i(t) \cdot \nabla - \delta$, with the $\mathbb{R}^2$ extension of the matrix $\boldsymbol{A}$ in the previous section. Finally, we let $\boldsymbol{Y_t}$ denote the vector of sulfate values evaluated at each point in $\mathcal{S}$ (Figure 3), $\boldsymbol{m}$ a vector of sulfate sources corresponding to each point in $\mathcal{S}$, and $\boldsymbol{W_t}$ a vector of independent Brownian motions located at each point in $\mathcal{S}$. Thus, each continuous element in Equation (SM.1) has a discrete counterpart, and our continuous space-time sulfate process can now be approximated with the multivariate Ornstein-Uhlenbeck (discrete space) process,

$$d\boldsymbol{Y_t} = \big(-\boldsymbol{A}\boldsymbol{Y_t} + \boldsymbol{m}\big)dt + \sigma\ d\boldsymbol{W_t}. \tag{SM.12}$$

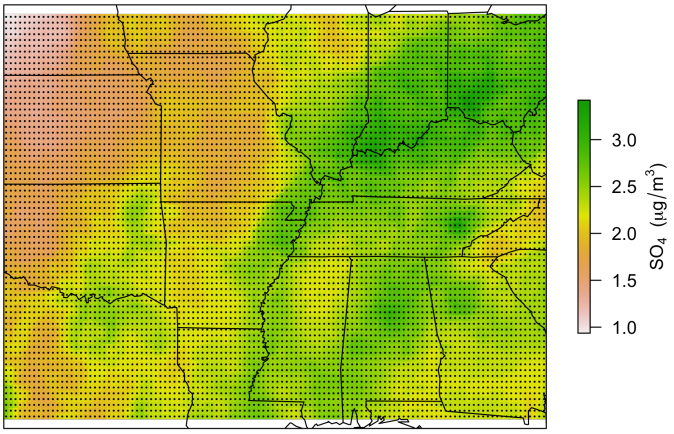


Figure 3: Spatial region $\mathcal{D}$ considered in the sulfate analysis, with the overlaid 116×70 rectangular grid, $\mathcal{S}$, used in the discrete approximation.

References

[1] Claes Johnson. *Numerical solution of partial differential equations by the finite element method*. Dover Publications Inc., New York, 2009.

[2] Hans Petter Langtangen and Svein Linge. *Finite difference computing with PDEs: A modern software approach*. Springer International Publishing, 2017.

[3] Finn Lindgren and Havard Rue. Bayesian spatial modelling with r-inla. *Journal of Statistical Software, Articles*, 63(19):1–25, 2015.

[4] Havard Rue and Leonhard Held. *Gaussian Markov random fields: Theory and methods*. Chapman and Hall/CRC, New York, 2005.

[5] H. K. Versteeg and W. Malalasekera. *An introduction to computational fluid dynamics: The finite volume method*. Pearson Education Ltd, Harlow, England, 2007.