Exploring a New Class of Non-statationary Spatial Gaussian Random Fields with Varying Local Anisotropy

Geir-Arne Fuglstad[1], Finn Lindgren[2], Daniel Simpson[1], and Håvard Rue[1]

[1]*Department of Mathematical Sciences, NTNU, Norway*
[2]*Department of Mathematical Sciences, University of Bath, United Kingdom*

Supplementary Material

The supplementary material contains investigation of inference on simulated data, the derivation of the discretization of the SPDE used in the paper, and the derivation of the marginal variance of the stationary solution.

S1 Inference on simulated data

We consider data generated from a known set of parameters. The prior used is improper, uniform on $(0, \infty)$ for γ and uniform on $\mathbb{R}$ for the rest of the parameters in $\mathbf{H}$. We investigate whether it is possible to estimate the stationary model with exactly observed data and whether the approximate estimation scheme performs well.

Example S1.1. Use the SPDE

$$u(\boldsymbol{s}) - \nabla \cdot \mathbf{H} \nabla u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in [0, 20] \times [0, 20], \tag{S1.1}$$

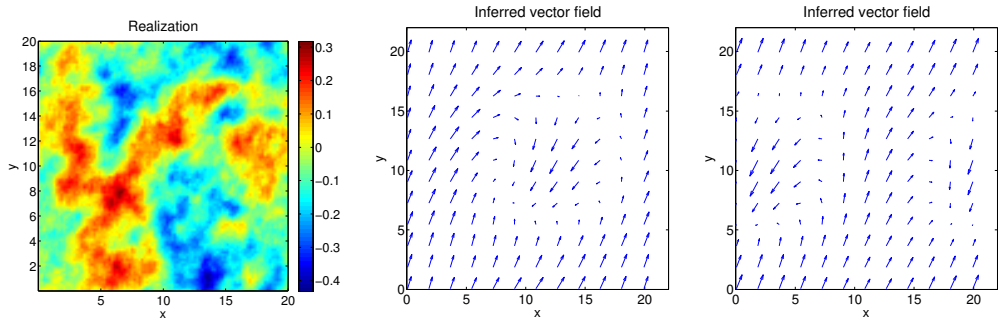
where $\mathcal{W}$ is a standard Gaussian white noise process and $\mathbf{H}$ is a 2×2 matrix, with periodic boundary conditions. Let $\mathbf{H} = 3\mathbf{I}_2 + 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$, with $\boldsymbol{v} = (1, \sqrt{3})/2$. Here $\mathbf{H}$ has an eigenvector $\boldsymbol{v}$ with eigenvalue 5 and an eigenvector orthogonal to $\boldsymbol{v}$ with eigenvalue 3. Construct the GMRF on a 100×100 grid.

One observation of the solution is shown in Figure 1(a). Assume that $\mathbf{H}$ is constant, but that its value is not known. Then using the decomposition from the previous sections one can write

$$\mathbf{H} = \gamma \mathbf{I}_2 + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix},$$

where γ, v_1 and v_2 are the parameters. Since the process is assumed to be exactly observed, we can use the distribution of $\boldsymbol{\theta}|\boldsymbol{u}$. This gives the posterior estimates shown in Table 1. From the table one can see that all the estimates are accurate to one digit, and within one standard deviation of the true value. This decomposition of $\mathbf{H}$ is invariant to the sign of $\boldsymbol{v}$, so there are two choices of parameters that mean the same.

The biases in the estimates were evaluated by generating 10000 datasets from the true model and estimating the parameters for each dataset. The estimated bias was



(a) Example S1.1. (b) Wrong maximum. (c) Wrong maximum.

Figure 1: (a) One realization of the solution of the SPDE in Example S1.1. (b) and (c) show two local maxima found for the vector field in Example S1.2. The vector field in (b) has a lower value for the posterior distribution of the parameters than the vector field in (c) and both have lower value than the actual maximum.

Table 1: Parameter estimates for Example S1.1.

Parameter	**True value**	**Estimate**	**Std.dev.**
γ	3	2.965	0.070
v_1	0.707	0.726	0.049
v_2	1.225	1.231	0.039

less than or equal to 0.1% of the true value for each parameter. Additionally, the sample standard deviations based on the estimation of the parameters for each of the 10000 datasets were 0.070, 0.050, and 0.039 for γ, v_1, and v_2, respectively. Each one corresponds well to the corresponding approximate standard deviation, computed via the observed information matrix, that is shown in Table 1.

It is then possible to estimate the model, but this is under the assumption that it is known beforehand that the model is stationary. The estimation is repeated for the more complex model developed in the previous sections that allows for significant non-stationarity controlled through a vector field. The intention is to evaluate whether the more complex model is able to detect that the true model is a stationary model, and if there are identifiability issues.

Example S1.2. Use the same SPDE and observation as in Example S1.1, but assume that it is not known that $\mathbf{H}$ is constant. Add the terms in the Fourier series corresponding to the next frequencies, $(k, l) = (0, 1)$, $(k, l) = (1, -1)$, $(k, l) = (1, 0)$ and $(k, l) = (1, 1)$. The observation is still assumed to be exact, but there are 16 additional parameters, 4 additional parameters for each frequency.

Two arbitrary starting positions are chosen for the optimization: $\gamma = 3.0$ and all other parameters at 0.1; $\gamma = 3.0$, $A^{(1)}_{0,0} = 0.1$, $A^{(2)}_{0,0} = 0.1$, and all other parameters equal to 0. For both starting points the optimization converges to non-global maximums. Pa-

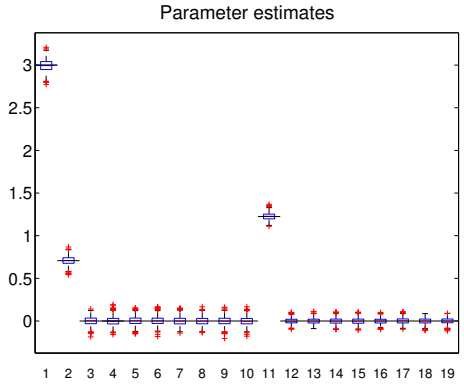


Figure 2: Boxplot of estimated parameters for 1000 simulated datasets in Example S1.2. Parameters 1, 2 and 11 corresponds to γ, v_1 and v_2, respectively. The horizontal lines through the boxes specify the true parameter values, which in most cases closely match the medians.

rameter estimates and approximate standard deviations are not given, but Figures 1(b) and 1(c) show the two vector fields found.

A third optimization is done with starting values close to the correct parameter values. This gives a vector field close to the actual one, with estimates for γ, $A^{(1)}_{0,0}$, and $A^{(2)}_{0,0}$ that agree with the ones in Example S1.1 to two digits. The other frequencies all had coefficients close to zero, with the largest having an absolute value of 0.058.

The results illustrate a difficulty with estimation caused by the the inherent non-identifiability of the sign of the vector field. The true vector field is constant, and each estimated vector field has large parts which have the correct appearance if one only considers the lines defined by the arrows and not the direction. The positions where the vector field is wrong are smaller areas where the vector field flips its direction. It is difficult to reverse this flipping as it requires moving through states with smaller likelihood, thus creating undesirable local maximums. One approach to improving the situation would be to force an apriori preference for vector fields without abrupt changes, to introduce a prior that forces higher frequencies of the Fourier basis to be less desirable. This is an issue that needs to be addressed, and is briefly discussed in the main paper.

By starting close to the true value, one can do repeated simulations of datasets and prediction of parameters to evaluate how well the non-stationary model captures the fact that the true model is stationary and to see if there is any consistent bias. 1000 datasets were simulated and the estimation of the parameters was done for each dataset with a starting value close to the true value. This gives the result summarized in the boxplot in Figure 2. There does not appear to be any significant bias and the parameters that give non-stationarity are all close to zero.

There are issues in estimating the anisotropy in the non-stationary model due to the

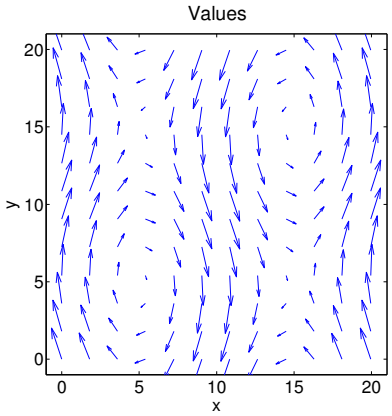


Figure 3: Vector field used in Example S1.3.

Table 2: Posterior inference on parameters in Example S1.3.

Parameter	**True value**	**Estimate**	**Std.dev.**
γ	0.5	0.5012	0.0081
β	5	5.014	0.084

non-identifiability of the sign of the vector field, but if one avoids the local maximums the estimated model is close to the true stationary model. There is a significant increase in computation time when increasing the parameter space from 3 to 19 parameters; the time required increases by a factor of approximately 10.

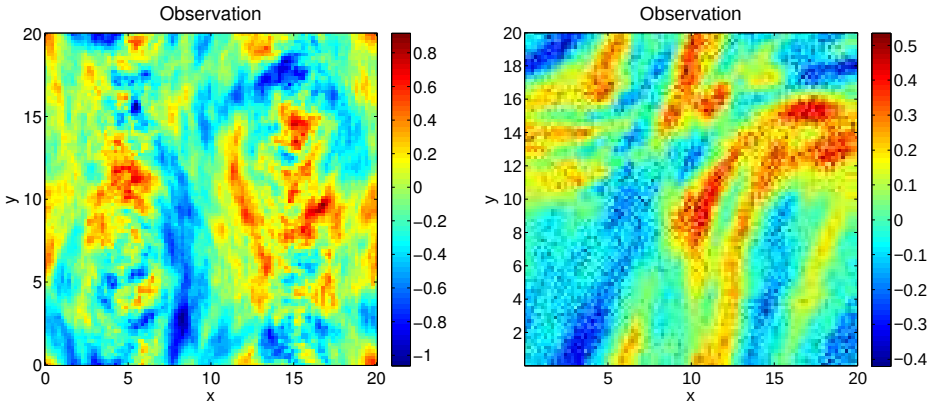
In situations where there is a physical explanation of the additional dependence in one direction, it would be desirable to do a simpler model with one parameter for the baseline isotropic effect and one parameter specifying the degree of anisotropy caused by a pre-defined vector field.

Example S1.3. Use a 100×100 grid of $[0, 20]^2$ and periodic boundary conditions for the SPDE in (S1.1). Let κ^2 be 1 and let $\mathbf{H}(\boldsymbol{s}) = \gamma \mathbf{I}_2 + \beta \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where $\boldsymbol{v}$ is the vector field in Figure 3.

Figure 4(a) shows one observation of the solution with $\gamma = 0.5$ and $\beta = 5$. One expects that it is possible to make accurate estimates about γ and β as the situation is simpler than in the previous example.

The estimated parameters are shown in Table 2. From the table one can see that the estimates for γ and β are quite accurate, to 2 digits. As in the previous example, the bias is estimated to be less than 0.02% for each each parameter, and the sample standard deviation from estimation over many datasets is 0.008 and 0.08 for γ and β, respectively.

Example S1.4. Use a 100×100 grid of $[0, 20]^2$ and periodic boundary conditions for the SPDE in (S1.1). Let κ^2 be 1 and let $\mathbf{H}(\boldsymbol{s}) = \mathbf{I}_2 + \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where $\boldsymbol{v}$ is the vector



(a) Example S1.3 (b) Example S1.4

Figure 4: (a) An observation of the SPDE in (S1.1) on a 100×100 regular grid of $[0, 20]^2$ with periodic boundary conditions, $\kappa^2 = 1$ and $\mathbf{H}(\boldsymbol{s}) = 0.5\mathbf{I}_2 + 5\boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where $\boldsymbol{v}$ is the vector field in Figure 3. (b) An observation of the SPDE in Example S1.4 with i.i.d. Gaussian white noise with precision 400.

field

$$\boldsymbol{v}(x, y) = \begin{bmatrix} 2 + \cos\left(\frac{\pi}{10}x\right) \\ 3 + 2\sin\left(\frac{\pi}{10}y\right) + \sin\left(\frac{\pi}{10}(x + y)\right) \end{bmatrix}.$$

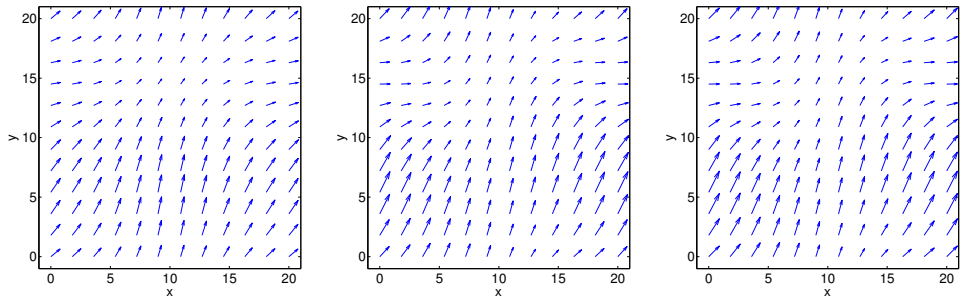
One observation with i.i.d. Gaussian noise with precision 400 is shown in Figure 4(b). Based on this realization it is desired to estimate the correct value of γ and the correct vector field $\boldsymbol{v}$ in the parametrization $\mathbf{H}(\boldsymbol{s}) = \gamma\mathbf{I}_2 + \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$. First use only the frequencies $(0, 0)$, $(0, 1)$, and $(1, 0)$. This gives the estimated vector field shown in Figure 5(a). Then use the frequencies $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. This gives the estimated vector field shown in Figure 5(b). The true vector field is shown in Figure 5(c).

The estimated vector fields are quite similar to the true vector field, and the γ parameter estimate was 1.14 in the first case and 1.09 in the latter case. There is a clear bias in the estimate of γ, to be expected as there is a need to compensate for the lacking frequencies. All parameter values were estimated, but are not shown. For the first case many parameters are more than two standard deviations from their correct values and, in the second case, this only happens for one parameter. If the difference between the true $\mathbf{H}$ and the estimated $\hat{\mathbf{H}}$ is calculated through

$$\frac{1}{100}\sqrt{\sum_{i=1}^{100}\sum_{j=1}^{100}\left|\left|\mathbf{H}(\boldsymbol{s}_{i,j}) - \hat{\mathbf{H}}(\boldsymbol{s}_{i,j})\right|\right|_2^2},$$

where $\boldsymbol{s}_{i,j}$ are the centres of the cells in the grid and $||\cdot||_2$ denotes the 2-norm, the first case gives 7.9 and the second case gives 1.5.

These examples focus on simple cases where specific issues can be highlighted. The



(a) $(0,0)$, $(0,1)$ and $(1,0)$ frequencies (b) $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$ frequencies (c) The true vector field

Figure 5: True vector field and inferred vector fields in Example S1.4. Each of the vector fields is scaled with a factor 0.3.

inherent challenges in estimating a spatially varying direction and strength are equally important in the more general setting where also κ and the baseline effect γ is allowed to vary. The estimation of the vector field presents an important component that must be dealt with in any inference strategy for the more general case.

S2 Derivation of precision matrix

S2.1 Formal equation

The SPDE is

$$(\kappa^2(\boldsymbol{s}) - \nabla \cdot \mathbf{H}(\boldsymbol{s}))\nabla u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in [0, A] \times [0, B], \tag{S2.1}$$

where A and B are strictly positive constants, κ^2 is a scalar function, $\mathbf{H}$ is a 2×2 matrix-valued function, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$, and $\mathcal{W}$ is a standard Gaussian white noise process. Here κ^2 is assumed to be a continuous, strictly positive function and $\mathbf{H}$ is assumed to be a continuously differentiable function which gives a positive definite matrix $\mathbf{H}(\boldsymbol{s})$ for each $\boldsymbol{s} \in [0, A] \times [0, B]$.

Periodic boundary conditions are used so that opposite sides of the rectangle $[0, A] \times [0, B]$ are identified. Thus values of κ^2 must agree on opposite edges, and the values of $\mathbf{H}$ and its first order derivatives must agree on opposite edges. Periodic boundary conditions are not essential to the methodology presented, but we avoid the issue of appropriate boundary conditions for now.

S2.2 Finite volume methods

In the discretization of the SPDE in (S2.1), a finite volume method is employed. Finite volume methods are useful for creating discretizations of conservation laws of the form

$$\nabla \cdot \boldsymbol{F}(\boldsymbol{x}, t) = f(\boldsymbol{x}, t),$$

where $\nabla\cdot$ is the spatial divergence operator. This equation relates the spatial divergence of the flux $\boldsymbol{F}$ and the sink-/source-term f. The main tool here is the use of the divergence theorem

$$\int_E \nabla \cdot \boldsymbol{F} \, \mathrm{d}V = \oint_{\partial E} \boldsymbol{F} \cdot \boldsymbol{n} \, \mathrm{d}\sigma, \tag{S2.2}$$

where $\boldsymbol{n}$ is the outer normal vector of the surface ∂E relative to E.

The main idea is to divide the domain of the SPDE in (S2.1) into smaller parts and consider the resulting "flow" between the different parts. A lengthy treatment of finite volume methods is not given, but a comprehensive treatment of the method for deterministic differential equations can be found in Eymard, Gallouët, and Herbin (2000).

S2.3 Derivation

To keep the calculations simple the domain is divided into a regular grid of rectangular cells. Use M cells in the x-direction and N cells in the y-direction. Then for each cell the sides parallel to the x-axis have length $h_x = A/M$ and the sides parallel to the y-axis have length $h_y = B/N$. Number the cells by (i, j), where i is the column of the cell (along the x-axis) and j is the row of the cell (along the y-axis). If the lowest row is 0 and the leftmost column is 0, then cell (i, j) is

$$E_{i,j} = [ih_x, (i+1)h_x] \times [jh_y, (j+1)h_y],$$

and the set of cells, $\mathcal{I}$, is

$$\mathcal{I} = \{E_{i,j} : i = 0, 1, \ldots, M-1, j = 0, 1, \ldots, N-1\}.$$

Figure 6 shows an illustration of the discretization of $[0, A] \times [0, B]$ into the cells $\mathcal{I}$.

Each cell has four faces, two parallel to the x-axis (top and bottom) and two parallel to the y-axis (left and right). Let the right face, top face, left face and bottom face of cell $E_{i,j}$ be denoted $\sigma^{\mathrm{R}}_{i,j}$, $\sigma^{\mathrm{T}}_{i,j}$, $\sigma^{\mathrm{L}}_{i,j}$ and $\sigma^{\mathrm{B}}_{i,j}$, respectively. Additionally, denote by $\sigma(E_{i,j})$ the set of faces of cell $E_{i,j}$.

For each cell $E_{i,j}$, $\boldsymbol{s}_{i,j}$ gives the centroid of the cell, and $\boldsymbol{s}_{i+1/2,j}$, $\boldsymbol{s}_{i,j+1/2}$, $\boldsymbol{s}_{i-1/2,j}$, and $\boldsymbol{s}_{i,j-1/2}$ give the centres of the faces of the cell. Due to the periodic boundary conditions, the i-index and j-index in $\boldsymbol{s}_{i,j}$ are modulo M and modulo N, respectively. Figure 7 shows one cell $E_{i,j}$ with the centroid and the faces marked on the figure. Further, let $u_{i,j} = u(\boldsymbol{s}_{i,j})$ for each cell and denote the area of $E_{i,j}$ by $V_{i,j}$. Since the grid is regular, all $V_{i,j}$ are equal to $V = h_x h_y$.

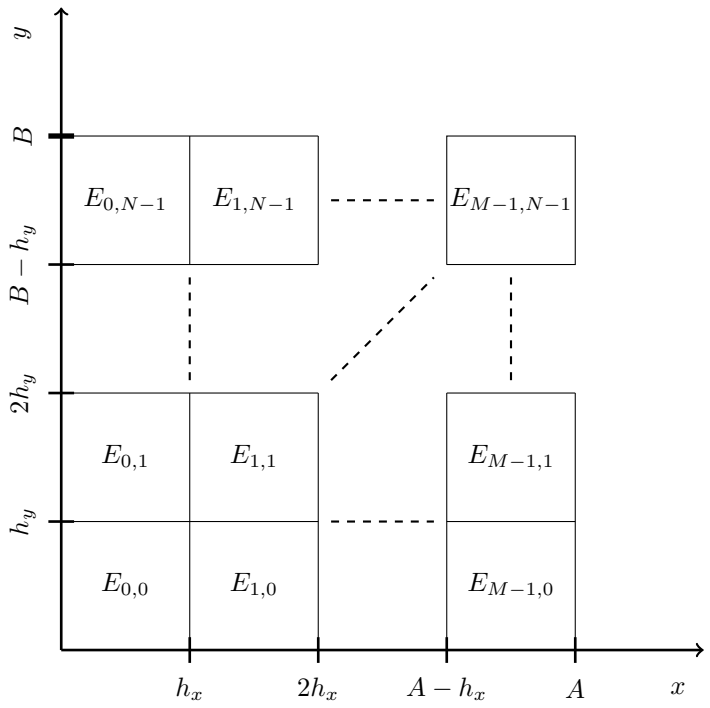


Figure 6: Illustration of the division of $[0, A] \times [0, B]$ into a regular $M \times N$ grid of rectangular cells.

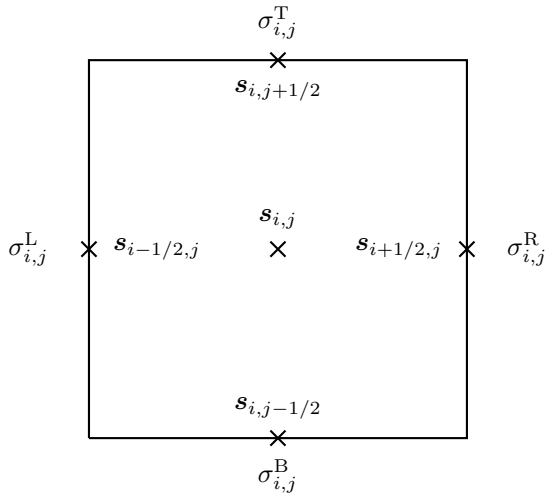


Figure 7: One cell, $E_{i,j}$, of the discretization with faces $\sigma^{\mathrm{R}}_{i,j}$, $\sigma^{\mathrm{T}}_{i,j}$, $\sigma^{\mathrm{L}}_{i,j}$ and $\sigma^{\mathrm{B}}_{i,j}$, centroid $\boldsymbol{s}_{i,j}$ and centres of the faces $\boldsymbol{s}_{i-1/2,j}$, $\boldsymbol{s}_{i,j-1/2}$, $\boldsymbol{s}_{i+1/2,j}$ and $\boldsymbol{s}_{i,j+1/2}$.

To derive the finite volume scheme, begin by integrating (S2.1) over a cell, $E_{i,j}$. This gives

$$\int_{E_{i,j}} \kappa^2(\boldsymbol{s})u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} - \int_{E_{i,j}} \nabla\cdot\mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} = \int_{E_{i,j}} \mathcal{W}(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}, \tag{S2.3}$$

where $\mathrm{d}\boldsymbol{s}$ is an area element. The integral on the right hand side is distributed as a Gaussian variable with mean 0 and variance V for each (i,j) (Adler and Taylor (2007, pp. 24–25)). Further, the integral on the right hand side is independent for different cells, because two different cells can at most share a common face. Thus (S2.3) can be written as

$$\int_{E_{i,j}} \kappa^2(\boldsymbol{s})u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} - \int_{E_{i,j}} \nabla\cdot\mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} = \sqrt{V}z_{i,j},$$

where $z_{i,j}$ is a standard Gaussian variable for each (i,j) and the Gaussian variables are independent.

By the divergence theorem, (S2.2), the second integral on the left hand side can be written as an integral over the boundary of the cell, so

$$\int_{E_{i,j}} \kappa^2(\boldsymbol{s})u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} - \oint_{\partial E_{i,j}} (\mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s}))^{\mathrm{T}}\boldsymbol{n}(\boldsymbol{s})\,\mathrm{d}\sigma = \sqrt{V}z_{i,j}, \tag{S2.4}$$

where $\boldsymbol{n}$ is the exterior normal vector of $\partial E_{i,j}$ with respect to $E_{i,j}$ and $\mathrm{d}\sigma$ is a line element. It is useful to divide the integral over the boundary in Equation (S2.4) into integrals over each face,

$$\int_{E_{i,j}} \kappa^2(\boldsymbol{s})u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} - \left(W_{i,j}^R + W_{i,j}^T + W_{i,j}^L + W_{i,j}^B\right) = \sqrt{V}z_{i,j}, \tag{S2.5}$$

where $W_{i,j}^{\mathrm{dir}} = \int_{\sigma_{i,j}^{\mathrm{dir}}} (\mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s}))^{\mathrm{T}}\boldsymbol{n}(\boldsymbol{s})\,\mathrm{d}\sigma$.

The first integral on the left hand side of (S2.5) is approximated by

$$\int_{E_{i,j}} \kappa^2(\boldsymbol{s})u(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} = V\kappa_{i,j}^2 u(\boldsymbol{s}_{i,j}) = V\kappa_{i,j}^2 u_{i,j}, \tag{S2.6}$$

where $\kappa_{i,j}^2 = \frac{1}{V}\int_{E_{i,j}} \kappa^2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}$. The function κ^2 is assumed to be continuous and $\kappa_{i,j}^2$ is approximated by $\kappa^2(\boldsymbol{s}_{i,j})$.

The second part of (S2.5) requires the approximation of the surface integral over each face of a given cell. The values of $\mathbf{H}$ are in general not diagonal, so it is necessary to estimate both components of the gradient on each face of the cell. For simplicity, it is assumed that the gradient is constant on each face and that it is identically equal to the value at the centre of the face. On a face parallel to the y-axis the estimate of the partial derivative with respect to x is simple since the centroid of each of the cells which share the face have the same y-coordinate. The problem is the estimate of the partial derivative with respect to y. The reverse is true for the top and bottom face of the cell.

Table 3: Finite difference schemes for the partial derivative with respect to x and y at the different faces of cell $E_{i,j}$.

Face	$\frac{\partial}{\partial x}u(s)$	$\frac{\partial}{\partial y}u(s)$
$\sigma^{\mathrm{R}}_{i,j}$	$\frac{u_{i+1,j}-u_{i,j}}{h_x}$	$\frac{u_{i,j+1}+u_{i+1,j+1}-u_{i,j-1}-u_{i+1,j-1}}{4h_y}$
$\sigma^{\mathrm{T}}_{i,j}$	$\frac{u_{i+1,j}+u_{i+1,j+1}-u_{i-1,j}-u_{i-1,j+1}}{4h_x}$	$\frac{u_{i,j+1}-u_{i,j}}{h_y}$
$\sigma^{\mathrm{L}}_{i,j}$	$\frac{u_{i,j}-u_{i-1,j}}{h_x}$	$\frac{u_{i-1,j+1}+u_{i,j+1}-u_{i-1,j-1}-u_{i,j-1}}{4h_y}$
$\sigma^{\mathrm{B}}_{i,j}$	$\frac{u_{i+1,j}+u_{i+1,j-1}-u_{i-1,j-1}-u_{i-1,j}}{4h_x}$	$\frac{u_{i,j}-u_{i,j-1}}{h_y}$

It is important to use a scheme which gives the same estimate of the gradient for a given face no matter which of the two neighbouring cells are chosen. For $\sigma^{\mathrm{R}}_{i,j}$, the approximation used is

$$\frac{\partial}{\partial y}u(\boldsymbol{s}_{i+1/2,j}) \approx \frac{1}{h_y}(u(\boldsymbol{s}_{i+1/2,j+1/2}) - u(\boldsymbol{s}_{i+1/2,j-1/2})).$$

where the values of u at $\boldsymbol{s}_{i+1/2,j+1/2}$ and $\boldsymbol{s}_{i+1/2,j-1/2}$ are linearly interpolated from the values at the four closest cells. More precisely, because of the regularity of the grid the mean of the four closest cells are used, giving

$$\frac{\partial}{\partial y}u(\boldsymbol{s}_{i+1/2,j}) \approx \frac{1}{4h_y}(u_{i+1,j+1} + u_{i,j+1} - u_{i,j-1} - u_{i+1,j-1}). \tag{S2.7}$$

This formula can be used for the partial derivative with respect to y on any face parallel to the y-axis by suitably changing the i and j indices. The partial derivative with respect to x on a face parallel to the y-axis can be approximated directly by

$$\frac{\partial}{\partial x}u(\boldsymbol{s}_{i+1/2,j}) \approx \frac{1}{h_x}(u_{i+1,j} - u_{i,j}). \tag{S2.8}$$

In more or less the same way the two components of the gradient on the top face of cell $E_{i,j}$ can be approximated by

$$\frac{\partial}{\partial x}u(\boldsymbol{s}_{i,j+1/2}) \approx \frac{1}{4h_x}(u_{i+1,j+1} + u_{i+1,j} - u_{i-1,j} - u_{i-1,j+1}),$$

$$\frac{\partial}{\partial y}u(\boldsymbol{s}_{i,j+1/2}) \approx \frac{1}{h_y}(u_{i,j+1} - u_{i,j}).$$

These approximations can be used on any side parallel to the x-axis by changing the indices appropriately.

The approximations for the partial derivatives on each face are collected in Table 3. Using these, one can find the approximations needed for the second part of (S2.5). It is helpful to write

$$W^{\mathrm{dir}}_{i,j} = \int_{\sigma^{\mathrm{dir}}_{i,j}} (\mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s}))^{\mathrm{T}}\boldsymbol{n}(\boldsymbol{s})\,\mathrm{d}\sigma = \int_{\sigma^{\mathrm{dir}}_{i,j}} (\nabla u(\boldsymbol{s}))^{\mathrm{T}}(\mathbf{H}(\boldsymbol{s})\boldsymbol{n}(\boldsymbol{s}))\,\mathrm{d}\sigma,$$

where the symmetry of $\mathbf{H}$ is used to avoid transposing the matrix. Assuming that the gradient is identically equal to the value at the centre of the face, one finds

$$W_{i,j}^{\mathrm{dir}} \approx \left(\nabla u(\boldsymbol{c}_{i,j}^{\mathrm{dir}})\right)^{\mathrm{T}} \int_{\sigma_{i,j}^{\mathrm{dir}}} \mathbf{H}(\boldsymbol{s})\boldsymbol{n}(\boldsymbol{s})\,\mathrm{d}\sigma,$$

where $\boldsymbol{c}_{i,j}^{\mathrm{dir}}$ is the centre of face $\sigma_{i,j}^{\mathrm{dir}}$.

Since the cells form a regular grid, $\boldsymbol{n}$ is constant on each face. If $\mathbf{H}$ be approximated by its value at the centre of the face, then

$$W_{i,j}^{\mathrm{dir}} \approx m(\sigma_{i,j}^{\mathrm{dir}}) \left(\nabla u(\boldsymbol{c}_{i,j}^{\mathrm{dir}})\right)^{\mathrm{T}} \left(\mathbf{H}(\boldsymbol{c}_{i,j}^{\mathrm{dir}})\boldsymbol{n}(\boldsymbol{c}_{i,j}^{\mathrm{dir}})\right), \tag{S2.9}$$

where $m(\sigma_{i,j}^{\mathrm{dir}})$ is the length of the face. Note that the length of the face is either h_x or h_y and that the normal vector is parallel to the x-axis or the y-axis.

If

$$\mathbf{H}(\boldsymbol{s}) = \begin{bmatrix} H^{11}(\boldsymbol{s}) & H^{12}(\boldsymbol{s}) \\ H^{21}(\boldsymbol{s}) & H^{22}(\boldsymbol{s}) \end{bmatrix},$$

then using Table 3 one finds the approximations

$$\begin{aligned}
\hat{W}_{i,j}^{\mathrm{R}} = {} & \\
& h_y \left[H^{11}(\boldsymbol{s}_{i+1/2,j}) \frac{u_{i+1,j} - u_{i,j}}{h_x} \right] + \\
& h_y \left[H^{21}(\boldsymbol{s}_{i+1/2,j}) \frac{u_{i,j+1} + u_{i+1,j+1} - u_{i,j-1} - u_{i+1,j-1}}{4h_y} \right],
\end{aligned}$$

$$\begin{aligned}
\hat{W}_{i,j}^{\mathrm{T}} = {} & \\
& h_x \left[H^{12}(\boldsymbol{s}_{i,j+1/2}) \frac{u_{i+1,j+1} + u_{i+1,j} - u_{i-1,j+1} - u_{i-1,j}}{4h_x} \right] + \\
& h_x \left[H^{22}(\boldsymbol{s}_{i,j+1/2}) \frac{u_{i,j+1} - u_{i,j}}{h_y} \right],
\end{aligned}$$

$$\begin{aligned}
\hat{W}_{i,j}^{\mathrm{L}} = {} & \\
& h_y \left[H^{11}(\boldsymbol{s}_{i-1/2,j}) \frac{u_{i-1,j} - u_{i,j}}{h_x} \right] + \\
& h_y \left[H^{21}(\boldsymbol{s}_{i-1/2,j}) \frac{u_{i,j-1} + u_{i-1,j-1} - u_{i-1,j+1} - u_{i,j+1}}{4h_y} \right],
\end{aligned}$$

$$\begin{aligned}
\hat{W}_{i,j}^{\mathrm{B}} = {} & \\
& h_x \left[H^{12}(\boldsymbol{s}_{i,j-1/2}) \frac{u_{i-1,j} + u_{i-1,j-1} - u_{i+1,j} - u_{i+1,j-1}}{4h_x} \right] + \\
& h_x \left[H^{22}(\boldsymbol{s}_{i,j-1/2}) \frac{u_{i,j-1} - u_{i,j}}{h_y} \right].
\end{aligned}$$

These approximations can be combined with the approximations in (S2.6) and inserted into (S2.5) to give

$$V\kappa_{i,j}^2 u_{i,j} - \left(\hat{W}_{i,j}^{\mathrm{R}} + \hat{W}_{i,j}^{\mathrm{T}} + \hat{W}_{i,j}^{\mathrm{L}} + \hat{W}_{i,j}^{\mathrm{B}}\right) = \sqrt{V} z_{i,j}.$$

Stacking the variables $u_{i,j}$ row-wise in a vector $\boldsymbol{u}$ gives the linear system of equations,

$$\mathbf{D}_V \mathbf{D}_{\kappa^2} \boldsymbol{u} - \mathbf{A}_{\mathbf{H}} \boldsymbol{u} = \mathbf{D}_V^{1/2} \boldsymbol{z}, \tag{S2.10}$$

where $\mathbf{D}_V = V\mathbf{I}_{MN}$, $\mathbf{D}_{\kappa^2} = \mathrm{diag}(\kappa_{0,0}^2, \ldots, \kappa_{M-1,0}^2, \kappa_{0,1}^2, \ldots, \kappa_{M-1,N-1}^2)$, $\boldsymbol{z} \sim \mathcal{N}_{MN}(\mathbf{0}, \mathbf{I}_{MN})$; $\mathbf{A}_{\mathbf{H}}$ is considered more closely in what follows.

The construction of $\mathbf{A}_{\mathbf{H}}$, which depends on the function $\mathbf{H}$, requires only that one write out the sum

$$\hat{W}_{i,j}^{\mathrm{R}} + \hat{W}_{i,j}^{\mathrm{T}} + \hat{W}_{i,j}^{\mathrm{L}} + \hat{W}_{i,j}^{\mathrm{B}}$$

and collects the coefficients of the different $u_{a,b}$ terms. This is not difficult, but requires many lines of equations. Therefore, only the resulting coefficients are given. Fix (i, j) and consider the equation for cell $E_{i,j}$. For convenience, let i_p and i_n be the column left and right of the current column, respectively and let j_n and j_p be the row above and below the current row, respectively. These rows and columns are 0-indexed and due to the periodic boundary conditions one has, for example, that column 0 is to the right of column $M - 1$. Further, number the rows and columns of the matrix $\mathbf{A}_{\mathbf{H}}$ from 0 to $MN - 1$.

For row $jM + i$ the coefficient of $u_{i,j}$ is given by

$$\begin{aligned}(\mathbf{A}_{\mathbf{H}})_{jM+i,jM+i} = \\ &- \frac{h_y}{h_x}\left[H^{11}(\boldsymbol{s}_{i+1/2,j}) + H^{11}(\boldsymbol{s}_{i-1/2,j})\right] \\ &- \frac{h_x}{h_y}\left[H^{22}(\boldsymbol{s}_{i,j+1/2}) + H^{22}(\boldsymbol{s}_{i,j-1/2})\right].\end{aligned}$$

The four closest neighbours have coefficients

$$\begin{aligned}(\mathbf{A}_{\mathbf{H}})_{jM+i,jM+i_p} &= \frac{h_y}{h_x} H^{11}(s_{i-1/2,j}) - \frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j+1/2}) - H^{12}(\boldsymbol{s}_{i,j-1/2})\right], \\ (\mathbf{A}_{\mathbf{H}})_{jM+i,jM+i_n} &= \frac{h_y}{h_x} H^{11}(\boldsymbol{s}_{i+1/2,j}) + \frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j+1/2}) - H^{12}(\boldsymbol{s}_{i,j-1/2})\right], \\ (\mathbf{A}_{\mathbf{H}})_{jM+i,j_nM+i} &= \frac{h_x}{h_y} H^{22}(\boldsymbol{s}_{i,j+1/2}) + \frac{1}{4}\left[H^{21}(\boldsymbol{s}_{i+1/2,j}) - H^{21}(\boldsymbol{s}_{i-1/2,j})\right], \\ (\mathbf{A}_{\mathbf{H}})_{jM+i,j_pM+i} &= \frac{h_x}{h_y} H^{22}(\boldsymbol{s}_{i,j-1/2}) - \frac{1}{4}\left[H^{21}(\boldsymbol{s}_{i+1/2,j}) - H^{21}(\boldsymbol{s}_{i-1/2,j})\right].\end{aligned}$$

Lastly, the four diagonally closest neighbours have coefficients

$$
\begin{aligned}
(\mathbf{A_H})_{jM+i,j_pM+i_p} &= +\frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j-1/2}) + H^{21}(\boldsymbol{s}_{i-1/2,j})\right], \\
(\mathbf{A_H})_{jM+i,j_pM+i_n} &= -\frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j-1/2}) + H^{21}(\boldsymbol{s}_{i+1/2,j})\right], \\
(\mathbf{A_H})_{jM+i,j_nM+i_p} &= -\frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j+1/2}) + H^{21}(\boldsymbol{s}_{i-1/2,j})\right], \\
(\mathbf{A_H})_{jM+i,j_nM+i_n} &= +\frac{1}{4}\left[H^{12}(\boldsymbol{s}_{i,j+1/2}) + H^{21}(\boldsymbol{s}_{i+1/2,j})\right].
\end{aligned}
$$

The rest of the elements of row $jM + i$ are 0.

Based on (S2.10) one can write $\boldsymbol{z} = \mathbf{D}_V^{-1/2}\mathbf{A}\boldsymbol{u}$, where $\mathbf{A} = \mathbf{D}_V\mathbf{D}_{\kappa^2} - \mathbf{A_H}$. This gives the joint distribution of $\boldsymbol{u}$,

$$
\begin{aligned}
\pi(\boldsymbol{u}) &\propto \pi(\boldsymbol{z}) \propto \exp\left(-\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}\right) \\
\pi(\boldsymbol{u}) &\propto \exp\left(-\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{D}_V^{-1}\mathbf{A}\boldsymbol{u}\right) \\
\pi(\boldsymbol{u}) &\propto \exp\left(-\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\mathbf{Q}\boldsymbol{u}\right),
\end{aligned}
$$

where $\mathbf{Q} = \mathbf{A}^{\mathrm{T}}\mathbf{D}_V^{-1}\mathbf{A}$. This is a sparse matrix with a maximum of 25 non-zero elements on each row, corresponding to the point itself, its 8 closest neighbours and the 8 closest neighbours of each of the 8 closest neighbours.

S3 Marginal variances with constant coefficients

Proposition S3.1. *Let u be a stationary solution of the SPDE*

$$
\kappa^2 u(x,y) - \nabla \cdot \mathbf{H}\nabla u(x,y) = \mathcal{W}(x,y), \qquad (x,y) \in \mathbb{R}^2, \tag{S3.1}
$$

where $\mathcal{W}$ is a standard Gaussian white noise process, $\kappa^2 > 0$ is a constant, $\mathbf{H}$ is a positive definite 2×2 matrix and $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$. Then u has marginal variance

$$
\sigma_m^2 = \frac{1}{4\pi\kappa^2\sqrt{\det(\mathrm{H})}}.
$$

Proof. Since the solution is stationary, Gaussian white noise is stationary, and the SPDE has constant coefficients, the SPDE is acting as a linear filter. Thus one can use spectral theory to find the marginal variance. The transfer function of the SPDE is

$$
g(\boldsymbol{w}) = \frac{1}{\kappa^2 + \boldsymbol{w}^{\mathrm{T}}\mathbf{H}\boldsymbol{w}}.
$$

The spectral density of a standard Gaussian white noise process on $\mathbb{R}^2$ is identically equal to $1/(2\pi)^2$ so it follows that the spectral density of the solution is

$$f_S(\boldsymbol{w}) = \left(\frac{1}{2\pi}\right)^2 \frac{1}{(\kappa^2 + \boldsymbol{w}^{\mathrm{T}}\mathbf{H}\boldsymbol{w})^2}.$$

From the spectral density it is only a matter of integrating the density over $\mathbb{R}^2$,

$$\sigma_m^2 = \int_{\mathbb{R}^2} f_S(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}.$$

The matrix $\mathbf{H}$ is (symmetric) positive definite and, therefore, has a (symmetric) positive definite square root, say $\mathbf{H}^{1/2}$. Use the change of variables $\boldsymbol{w} = \kappa\mathbf{H}^{-1/2}\boldsymbol{z}$ to find

$$\begin{aligned}
\sigma_m^2 &= \frac{1}{4\pi^2}\int_{\mathbb{R}^2} \frac{1}{(\kappa^2 + \kappa^2\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z})^2}\det(\kappa\mathbf{H}^{-1/2})\,\mathrm{d}\boldsymbol{z} \\
&= \frac{1}{4\pi^2\kappa^2\sqrt{\det(\mathbf{H})}}\int_{\mathbb{R}^2} \frac{1}{(1 + \boldsymbol{z}^{\mathrm{T}}\boldsymbol{z})^2}\,\mathrm{d}\boldsymbol{z} \\
&= \frac{1}{4\pi\kappa^2\sqrt{\det(\mathbf{H})}}.
\end{aligned}$$

□

References

Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. Springer Verlag.

Eymard, R., Gallouët, T., and Herbin, R. (2000). Finite Volume Methods. In Ciarlet, P.G. and Lions, J.L., editors, *Solution of Equations in $\mathbb{R}^n$, Techniques of Scientific Computing*, 713–1018. Elsevier.