ST 533: Applied Spatial Statistics

Mid-term Exam 2

Implementation of Stochastic Partial Differential Equation (SPDE) approach for Spatial Modeling

Presented by:

Md Mehedi Hasnat

Zahid Rangwala

Richard Watson

SPDEs

$$Df = \epsilon$$

D is a differential operator
Ex.

- D = d / dx + d^2 / dx^2
- Df = df / dx + d^2 f / dx^2

- f is a stochastic process, that is a solution for the SPDE
- **Interpretation**: There exists a function, f, for which the differential is ϵ

ϵ is a stochastic process

- Usually the white noise process
- Completely uncorrelated
- $\epsilon \sim N(0,\boldsymbol{\sigma})$

Why should you care? **It turns out that the covariance of the solution, f, are induced by the choice of D. We can abuse this by choosing D that induces the covariance function we want and use low cost computational methods to approximate the covariance matrix.**

Richard

Solving SPDEs: Finite Element Method(FEM)

What we need:

- A triangulation of the domain, or mesh, with n vertices
- Basis functions, $\boldsymbol{\psi}$, defined on each vertex or point in mesh

Then, the solution to the spde, f, is:

$$f(s) = \sum_{k=1}^{n} \psi_k(s) f_k$$

$$f_k \sim N(0, \Sigma)$$

The associated model is then:

$$\mathbf{Y} = \beta_0 + \mathbf{Af} + \epsilon$$

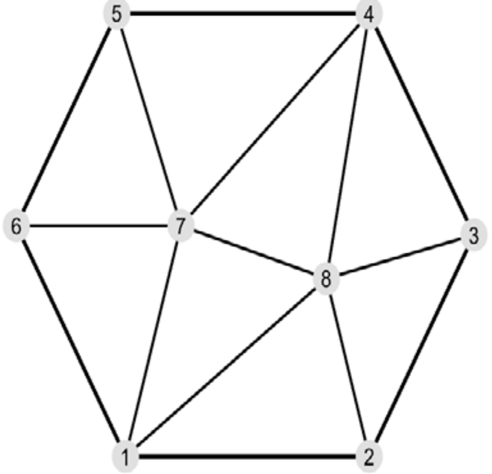
$$\epsilon \sim N(0, \sigma^2)$$

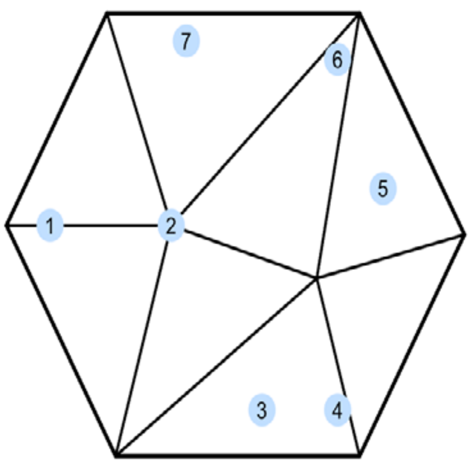
Where,

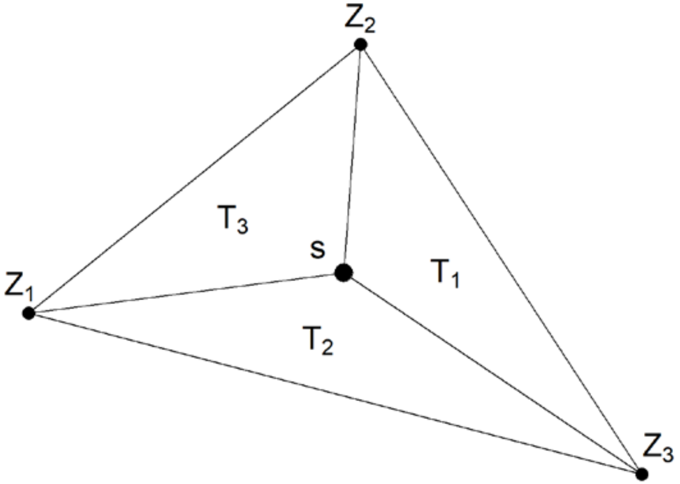
- **Y** is the vector of responses
- **A** is the projector matrix
- **f** is the vector of f_k's
- $\boldsymbol{\epsilon}$ is the vector of observation noise

Richard

Triangulated Mesh







- The mesh is built by splitting the domain into triangles.
- For each vertex in the mesh, $\boldsymbol{\psi}$□(s), is defined
  - Ex. $\boldsymbol{\psi}_1(s) = T_1 / (T_1 + T_2 + T_3)$

Richard

Projector Matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1G} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2G} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & A_{n3} & \dots & A_{nG} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ A_{21} & A_{22} & 0 & \dots & A_{2G} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & A_{n3} & \dots & 0 \end{bmatrix}$$

- **A** is a n x G matrix, where n is the number of observations and G is the number of vertices in the mesh
- A_j□ = $\boldsymbol{\psi}$□(s_j)
- Each row sums to 1
- ***Note:*** this means that each row has at most 3 entries with every other entry equal to 0
- This makes **A** sparse and thus the covariance sparse.

Richard

Putting it all together

1. Define the differential operator that will induce the matern covariance
    a. Here $\boldsymbol{\nu}$ is a smoothing parameter and *d* is the dimension
2. Define model
3. Estimate ***Z*** using FEM
4. Redefine model in terms of SPDE model

$$D = (\kappa^2 - \Delta)^{\alpha/2}$$

$$\alpha = \nu + d/2$$

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z} + \epsilon$$

$$\mathbf{Z} = \mathbf{Az}$$

$$\mathbf{H} = (\mathbf{A}, \mathbf{X}) \quad \mathbf{w} = (\mathbf{z}, \beta)$$

$$\mathbf{Y} = \mathbf{Hw} + \epsilon$$

$$\mathbf{w} \sim N(0, (\Sigma, \Sigma_\beta))$$

Richard

Implementation

4 step procedure:

1. Building “Mesh”
2. Defining SPDE
3. “Stacking”/Joining training and test data
4. Model run, prediction and measure of effectiveness.

Implementation: Building "Mesh"

"**Mesh**" is referred to the two-dimensional domain on which we do the analysis and prediction.

Important parameters within "Mesh" building function:

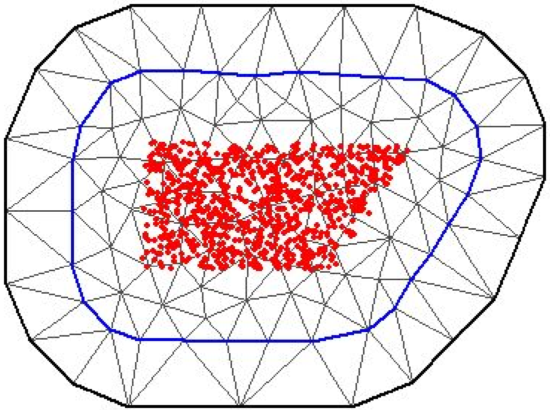
Cutoff - used to avoid building many small triangles around clustered input locations.

Offset - species the size of the inner and outer extensions around the data locations.

max.edge - species the maximum allowed triangle edge lengths in the inner domain and in the outer extension.

**Higher values in max.edge results in lower MSE with very low Coverage (as low as 7.9)*

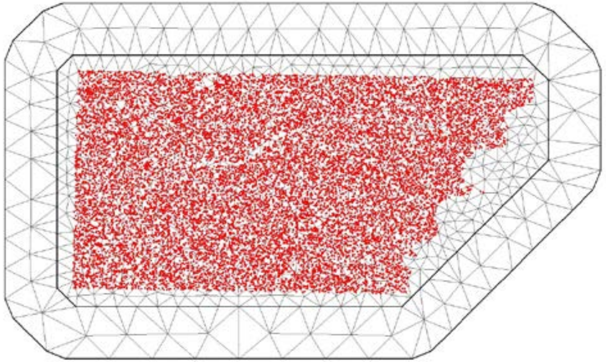
**Using too small values in max.edge takes a lot of time to build the mesh for large set of coordinates.*



cutoff = 0.01,

offset = c(0.015, 0.05),

max.edge = c(**0.05,0.5**)



cutoff = 0.01,

offset = c(0.015, 0.05),

max.edge = c(**0.015,0.15**)

Mehedi

Implementation: Defining SPDE

SPDE can be defined with or without the prior distributions.

Priors that can be included into the SPDE model:

sigma0 = field standard deviation

range0 = spatial range for theta = 0,

B.tau = matrix sorting spatial variance

B.kappa = matrix storing spatial scale parameter

SPDE without priors can be defined as: *inla.spde2.matern(mesh,* ***alpha*** *= 0.5),*

Where: α = v+d/2, d= dimensions, v = Matern Smoothness.

It is recommended to use the prior distributions for better results, however, it is very difficult to select proper prior distributions.

Implementation: Stacking

The Training dataset and the Testing dataset have to be stacked or joined together before passing them into the model.

```
dat_stack <- inla.stack(data  = list(evi = tdat1$Y),
                        A = list(A_dat,1,1),
                        effects = list(c(sindex, list(Intercept = 1)),
                                       list(lon = lon), list(lat = lat)),
                                       tag = "train_data")
```

```
stack_pred <- inla.stack(data = list(evi = NA),
     A = list(A_pred,1,1),
     effects = list(c(sindex, list (Intercept = 1)),
      list(lon = testsample$Longitude),list(lat = testsample$Latitude)),
                                                     tag = "predict")
```

```
join_stack <- inla.stack(dat_stack, stack_pred)
```

A represents the sparse matrix on the developed "Mesh" structure.

Implementation: Prediction and Measure of Effectiveness

The first order Matern with SPDE can be estimated as:

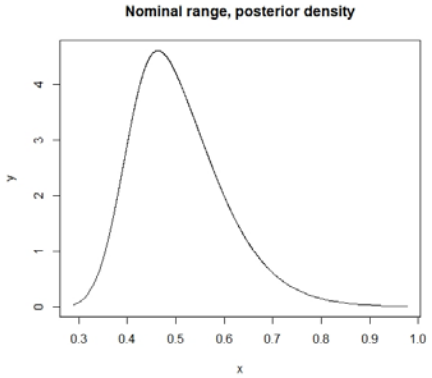
```
m1 <- inla(form, data = inla.stack.data(join_stack, spde = spde_model),
          family = "gaussian", control.predictor =
           list(A = inla.stack.A(join_stack), compute = TRUE),
          control.compute = list(cpo = TRUE, dic = TRUE))
```

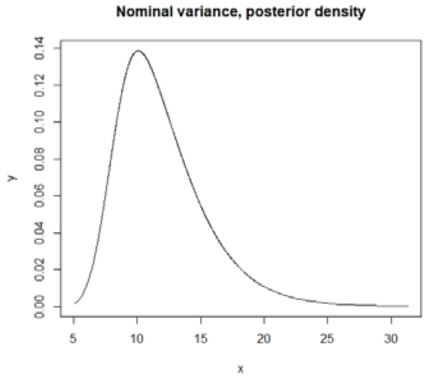
form represents the fitted model that includes the linear combination of lat, lon and spatial effects.

Summary of the model (summary(m1)) contains the predicted values for the data set.

--The parameter distributions are saved as “hyperparameter”

inla.spde2.results() can be used to extract and draw the posterior distribution of range and variance.





Mehedi

Results: Methodology

1. The training set was constructed with observations from $g \in \{1, 2, ..., 20\}$ to create 20 datasets of increasing size
2. Both the SPDE and the MLE-Kriging methods were both ran against the test set $g = 21$
3. Any jobs that ran over an hour were stopped

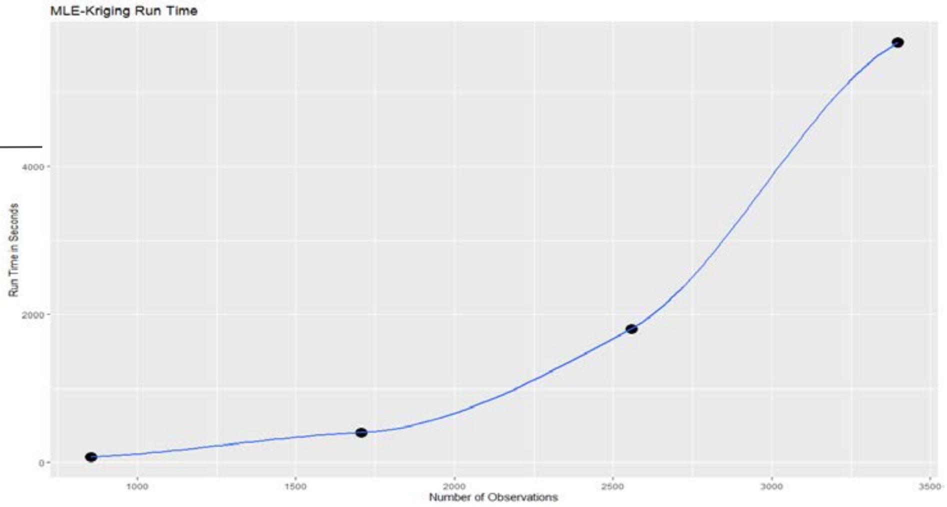
Results: MLE-Kriging

Computations took a long time, with the first set taking over 1 minute and the fourth taking over an hour and a half

Performance loss was exponential with every trial taking significantly longer than the previous

All performance metrics stayed relatively the same for trials 1-4 , with good coverage.

MSE stayed low for trials 1-4



MLE	OBS	Run Time Seconds	MSE	MAD	SD	Coverage	COR
1	854	75.77	0.0155	0.08010	0.12134	93.52428	0.5419
2	1705	405.32	0.0142	0.07525	0.12046	94.52055	0.6021
3	2557	1806.70	0.0134	0.07128	0.11614	93.64882	0.6234
4	3398	5673.85	0.0136	0.07039	0.11884	94.39601	0.6155

Zahid

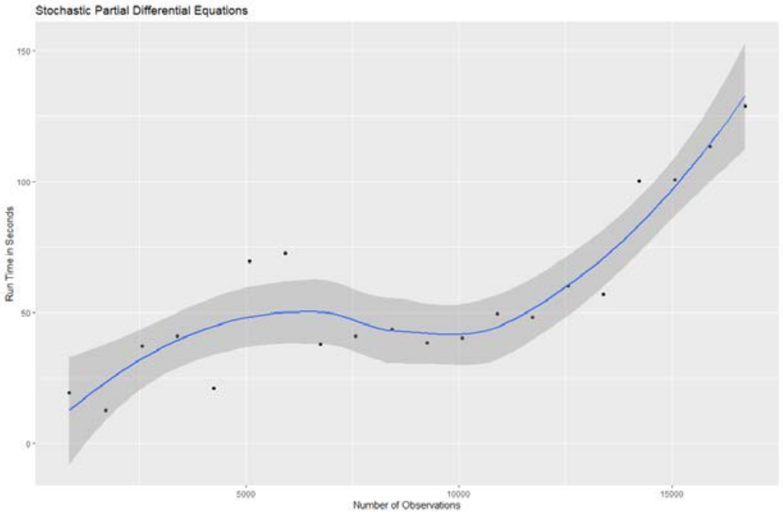
Results: SPDE

Computation time is much faster , SPDE can compute n = 13380 observations faster than MLE-Kriging can do n = 854.

SPDE’s run time is more random as it oscillates back and forth throughout the trials

MSE is notably higher compared to MLE-Kriging

Coverage and Correlation are notably lower compared to MLE-Kriging

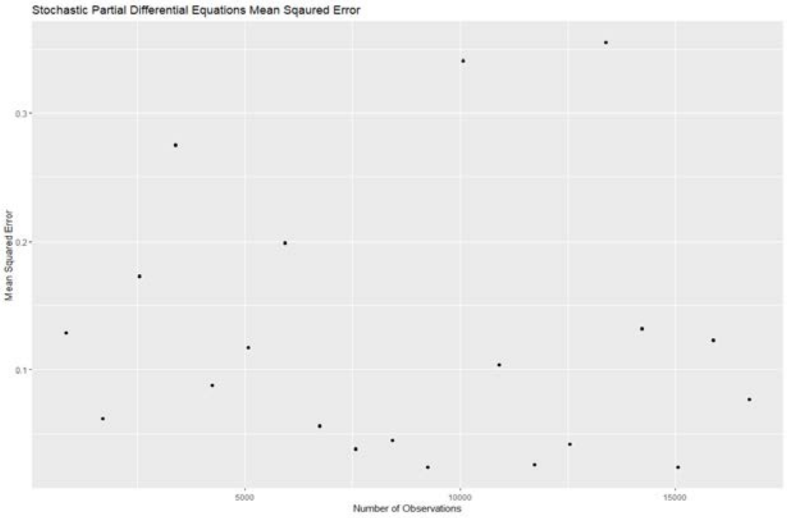


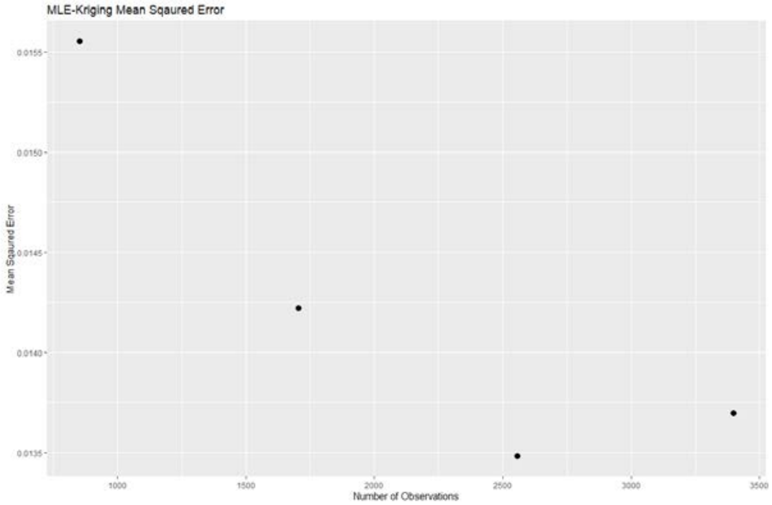
SPDE	OBS	Run Time Seconds	MSE	MAD	SD	Coverage	COR
1	853	19.40	0.129	0.103	0.161	42.7	0.205
2	1704	12.60	0.062	0.099	0.166	41.7	0.272
3	2555	37.10	0.173	0.1	0.095	38.2	0.21
4	3394	41.10	0.275	0.093	0.054	32.1	0.147
5	4247	21.10	0.088	0.099	0.103	34.3	0.221
6	5084	69.60	0.117	0.097	0.075	33.2	0.2
7	5928	72.60	0.199	0.098	0.095	30.3	0.223
8	6740	37.80	0.056	0.099	0.175	31.5	0.315
9	7567	40.97	0.038	0.095	0.14	33.5	0.366
10	8424	43.60	0.045	0.093	0.105	30.7	0.327
11	9241	38.5	0.024	0.095	0.17	31.5	0.514
12	10070	40.20	0.341	0.093	0.099	29.1	0.214
13	10898	49.50	0.104	0.094	0.065	28.6	0.229
14	11727	48.30	0.026	0.094	0.097	29	0.49
15	12552	60.00	0.042	0.094	0.172	34.4	0.265
16	13380	57.10	0.355	0.096	0.067	27.9	0.19
17	14218	100.20	0.132	0.091	0.054	27.5	0.211
18	15058	100.80	0.024	0.092	0.127	30	0.521
19	15883	113.40	0.123	0.094	0.052	26	0.216
20	16713	129.00	0.077	0.091	0.128	27.3	0.314

Zahid

Results: MLE vs SPDE

Mean Estimates	Intercept	Longitude	Latitude
MLE-Kriging	29.1599	.0138	-.6652
SPDE	.800	-.167	-.444





Zahid