INLA analysis of Swedish Temperature Data

Wan Kuan Teh and Wanru Cheng

January 2025

1 Introduction

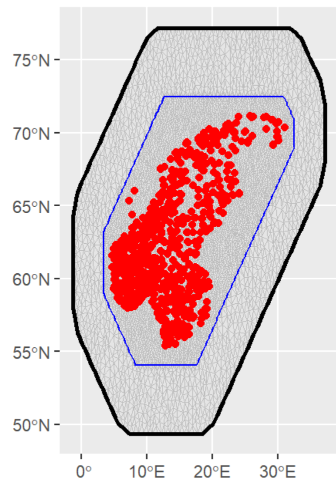
In this project, we will construct a spatial model of average monthly temperatures in Norway and Sweden during September of 2022. With observations of weather stations and data including:

- Latitude and Longitude for each location
- Elevation in meters
- Distance to closest coast
- Distances to the Swedish and Norwegian coast

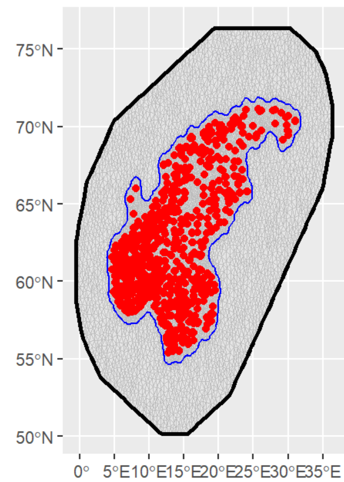
We considered different models, mesh constructions, covariates and have a full investigation into this project.

2 Mesh Construction

First, we start by refining the mesh, avoiding close points, adjusting interior and buffer sizes. Then, adding the non-convex boundary to further refine the mesh. A before and after comparison is shown below.



(a) Before refining mesh



(b) After refining mesh

3 Analysis on Covariates and Model construction

To investigate more models that may fit the senario, the linearities of each covariate should be well considered. First, we do a simple theoretical analysis and then analyze the linearity of covariates using the output of a Generalized Additive Model (GAM).

•Latitude and Longitude for each location

To start with,the Latitude is a linear covariate in that Latitude often has a direct linear relationship with temperature (toward the poles), especially in larger spatial regions. Besides, as for longitude, unless correlated with other factors like continentality or topography, is unlikely to have a strong linear or non-linear direct effect on temperature.

• Elevation in meters

Elevation has a well-documented relationship with temperature due to the temperature lapse rate, where temperature typically decreases as elevation increases. However, this relationship can vary across elevation ranges, making it potentially non-linear, for example, local effects such as valleys or ridges may introduce complex variations.

• Distance to closest coast

Coastal areas experience moderated temperatures due to the thermal properties of water, while inland areas experience greater temperature fluctuations. This effect decreases non-linearly as distance from the coast increases.

• Distances to the Swedish and Norwegian coast

Similar to the general distance-to-coast effect, this is likely non-linear due to diminishing moderation effects as distance increases.

Additionally, we have performed analysis using GAM model to define the linearities of all the covariates, the results are displayed in taable 1

	Lat	Lon	Elevation	Coast	CoastSE	CoastNO
EDF	8.462	2.521	3.870	7.327	7.942	8.100
p-value	2e-16	2e-16	2e-16	2e-6	2e-16	3.75e-16

Table 1: The result from GAM, for judging the linearities of covariates

From the table, it can be concluded all covariates should be modeled as non-linear covariates since EDF are all bigger than 1, and every smooth term is significant, since the p-values are all very small. However, from the figure2 below,the linearity is presented in the relationship between temperature and elevation as well ad temperature between latitude (in figure 4 and figure 3). So later on we will compare models with Elevation and Latitude being linear and non-linear separately.

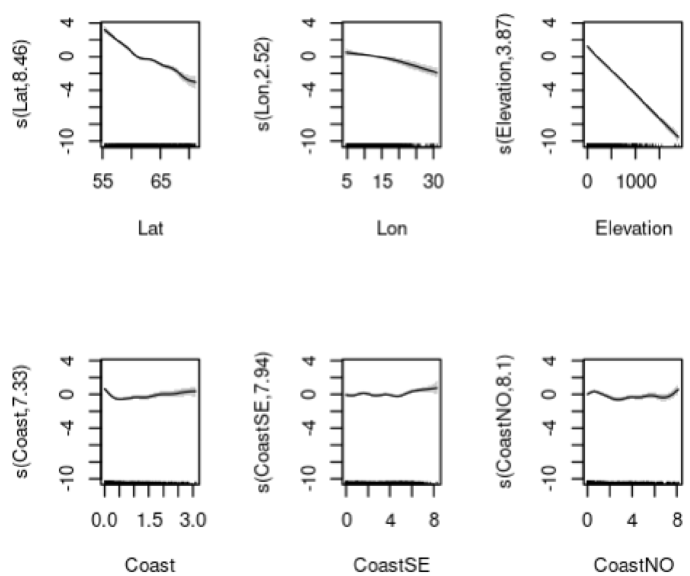


Figure 2: different covariates versus temperature

Next, we will include all the non-linear effects of covariates into the model using Random walk model. We first tried SPDE models, which failed many times because it is computational intensive. However, we successfully sun two sets of model structure as follows, each with elevation modeled as linear and non-linear:

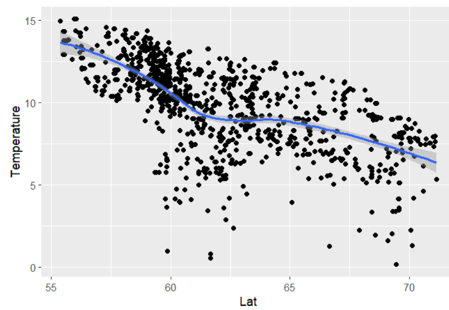


Figure 3: Lat modeled as linear

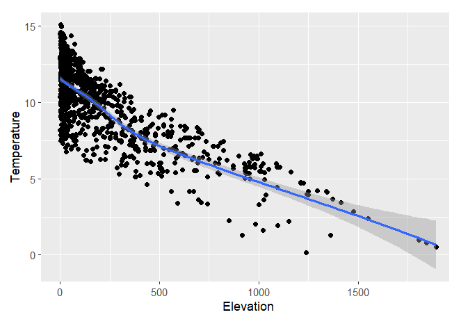


Figure 4: Elevation modeled as linear

4 Model Results and Validation

The qualitative results are listed below for the models, the DIC and MAIC metrics(The Elevation is modeled as linear). From the list above, it turned out

Model's covariates	DIC ↓	WAIC↓
Elevation(linear)	1881. 53	1858.27
CoastSE(Nonlinear)	3723.32	3730.30
CoastNO(Nonlinear)	3723.88	3726.21
Coast(Nonlinear)	3680.66	3685.05
Lat(Linear)	3705.31	3698.99
Ele+Lat	1912.91	1885.04
ELe+Co	1913.09	1893.33
Lat+ELe+Co	1879.15	1852.75
Ele+CoSE	1927.79	1909.73
Lat+Ele+CoSE	1891.07	1865.39
ELe+CoNO	1910.05	1888.25
Lat+ELe+CoNO	1881.90	1852.77
Ele+CoNO+Co	1912.72	1892.98
Lat+Ele+CoNO+Co	1870.80	1843.86
Ele+CoNO+CoSE	1915.55	1899.74
Lat+Ele+CoNO+CoSE	1863.05	1839.73
Ele+CoNO+COSE+Co	1914.63	1895.33
Lat+Ele+CoNO+COSE+Co	1859.37	1834.26

Table 2: DIC and WAIC model selection

the best model to be selected is:

$$Temperature \sim Lat + Ele + CoNO + COSE + Co$$

The model result is listed below:
The marginal posterior is below:
Range:
Variance:
Result:

5 Final Analysis and Conclusion

Spatial Effects and Covariates

The model results highlight the significant influence of both spatial effects and covariates on temperature patterns. The spatial components (CONO, COSE, Co) capture the inherent spatial autocorrelation in temperature data, indicating that temperatures at nearby locations tend to be similar. Additionally, the covariates Lat and Ele contribute significantly to explaining temperature

```
Time used:
    Pre = 0.817, Running = 52.2, Post = 0.135, Total = 53.2
Fixed effects:
                   mean    sd 0.025quant 0.5quant 0.975quant   mode kld
(Intercept)      19.678 0.539     18.621   19.677     20.741 19.677   0
Lat_binned       -0.459 0.017     -0.491   -0.458     -0.426 -0.458   0
Elevation_binned -0.006 0.000     -0.006   -0.006     -0.006 -0.006   0

Random effects:
  Name    Model
    spatial.field SPDE2 model
   CoastNO_binned RW1 model
   Coast_binned RW1 model
   CoastSE_binned RW1 model

Model hyperparameters:
                                           mean     sd 0.025quant 0.5quant 0.975quant    mode
Precision for the Gaussian observations   4.416  0.244      3.952    4.410       4.91   4.401
Theta1 for spatial.field                 -1.718  0.201     -2.128   -1.713      -1.34  -1.690
Theta2 for spatial.field                  0.988  0.185      0.638    0.984       1.37   0.963
Precision for CoastNO_binned            161.868 83.702     42.835  146.589     360.46 112.080
Precision for Coast_binned                7.480  3.088      3.110    6.930      15.04   5.939
Precision for CoastSE_binned             21.708 12.284      7.163   18.787      53.71  14.275

Deviance Information Criterion (DIC) ...............: 1859.37
Deviance Information Criterion (DIC, saturated) ....: 1442.91
Effective number of parameters .....................: 271.56

Watanabe-Akaike information criterion (WAIC) ...: 1834.26
Effective number of parameters .................: 207.48

Marginal log-Likelihood:  -1219.30
 is computed
Posterior summaries for the linear predictor and the fitted values are computed
(Posterior marginals needs also 'control.compute=list(return.marginals.predictor=TRUE)')
```

Figure 5: Model Results

```
> head(p_theta)
# A tibble: 6 × 3
  name                                        x        y
  <chr>                                   <dbl>    <dbl>
1 Precision for the Gaussian observations  3.47 0.000219
2 Precision for the Gaussian observations  3.58 0.00192
3 Precision for the Gaussian observations  3.71 0.0161
4 Precision for the Gaussian observations  3.87 0.124
5 Precision for the Gaussian observations  3.95 0.269
6 Precision for the Gaussian observations  4.02 0.469
>
```

Figure 6: Precision of Gaussian Obvs

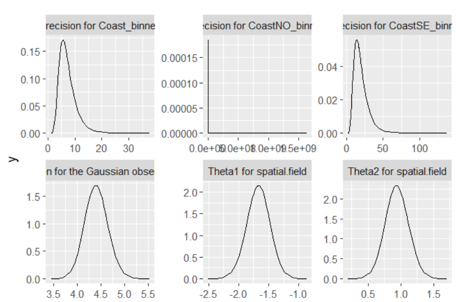


Figure 7: Marginal Posterior

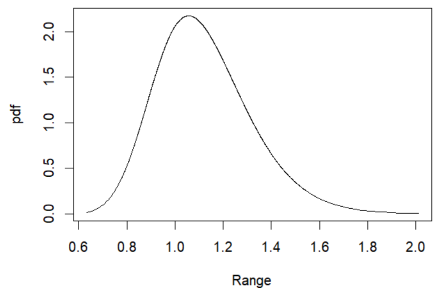


Figure 8: Range

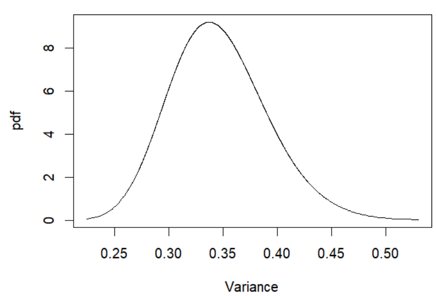


Figure 9: Variance

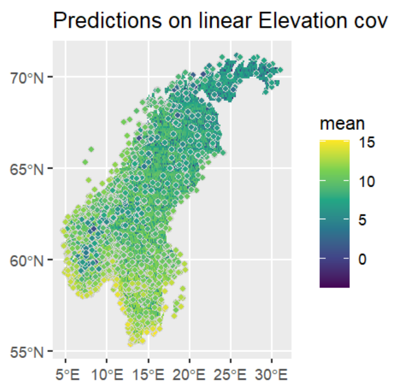


Figure 10: Prediction

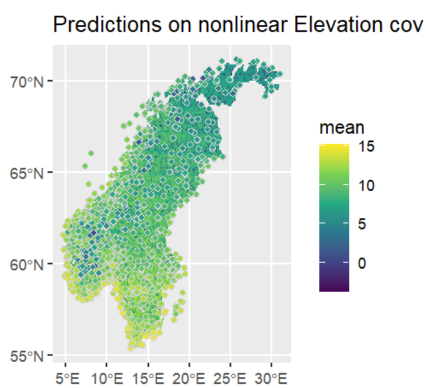


Figure 11: Prediction

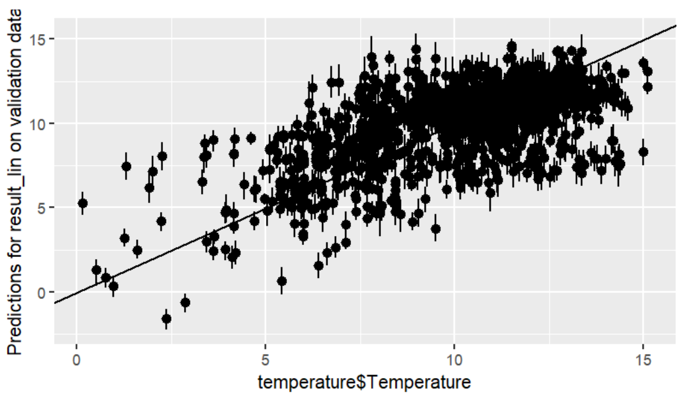


Figure 12: Prediction

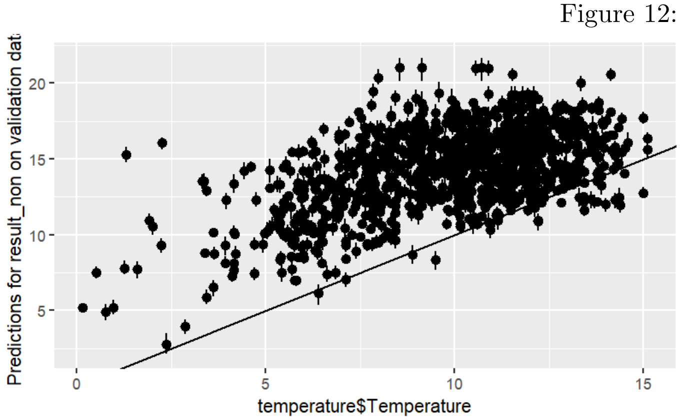


Figure 13: Prediction

variation. Latitude exhibits a clear negative effect, with temperatures decreasing as we move northward, which aligns with expectations. Elevation also shows a negative effect, reflecting the decrease in temperature with increasing altitude.

Marginal Posterior Distributions

The marginal posterior distributions of the hyperparameters provide insights into the model's uncertainty and behavior. The well-defined peaks and relatively narrow distributions for most hyperparameters suggest that the model has reasonably precise estimates. However, some distributions (e.g., "Precision for CoastSE") exhibit skewness, warranting further investigation to understand potential implications.

Visualization of Spatial Patterns

The plots of the spatial effects and covariate effects illustrate the spatial patterns in temperature variation. The latitude and elevation plots clearly show the decreasing trends in temperature with increasing latitude and elevation, respectively. The range and variance plots provide a comprehensive view of the spatial distribution of these metrics, highlighting areas of high and low variability.

Conclusion

Our analysis demonstrates the importance of considering both spatial effects and covariates when modeling temperature patterns. The selected model effectively captures the spatial autocorrelation and the influence of latitude and elevation on temperature variation. The marginal posterior distributions provide insights into the model's uncertainty, while the visualization of spatial patterns aids in understanding the complex interplay of factors influencing temperature. Future work could involve exploring additional covariates, refining the spatial components, and conducting more detailed sensitivity analyses to further enhance the model's accuracy and interpretability.