Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Mathematical Sciences

André Egeli

Probabilistic Forecasting of Solar Radiation in Norway Using Spatial Varying Coefficients for NWP Forecasts

Master's thesis in MTFYMA
Supervisor: Ingelin Steinsland
May 2023

NTNU
Norwegian University of
Science and Technology

André Egeli

Probabilistic Forecasting of Solar Radiation in Norway Using Spatial Varying Coefficients for NWP Forecasts

Master's thesis in MTFYMA
Supervisor: Ingelin Steinsland
May 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Abstract

This thesis introduces and evaluates a model for post-processing of solar radiation forecasts, utilizing a dataset comprising 2.5 years of NWP data and in situ ground observations from 15 locations in Norway. The proposed model is a Spatially varying coefficient model (SVCM) with a Beta distributed likelihood scaled by the physical constraints of solar radiation. We use a Bayesian hierarchical framework and employ a sliding window training period of 20 days to estimate the model parameters. To evaluate the model's probabilistic calibration, we utilized the probability integral transform (PIT) histogram, and for predictive accuracy, the continuously ranked probability score (CRPS). Additionally, the deterministic accuracy of the post-processed forecasts was evaluated using the root-mean-squared error (RMSE).

Our findings demonstrate that the SVCM effectively enhances predictive performance on validation locations beyond the training sample, indicating the model's generalizability to unseen locations within the domain. Specifically, when compared to the original NWP forecasts, the CRPS increase is 35.6% on the training locations and 20.9% on the validation locations. These improvements highlight the potential of the SVCM based on the scaled Beta distribution for improving solar radiation forecasts in Norway and contributing to the field of solar radiation prediction.

Preface

This thesis is a part of the course TMA4900 Master's thesis in Industrial Mathematics at the Norwegian University of Science and Technology (NTNU), Department of Mathematical Sciences. This work was carried out during the spring of 2023.

The main part of this work has been to develop and test different methods for ensemble post-processing of solar radiation in Norway, and I would like to thank John Bjørnas Bremnes from the Norwegian Meteorological Institute for the data necessary to develop and test the method used in this thesis.

I would also like to thank my supervisor, Professor Ingelin Steinsland, for our weekly meetings and guidance with model selection, evaluation metrics, and other valuable feedback.

This thesis is a continuation of the work done in the specialization project by Egeli (2022) in the course TMA4500, and therefore the datasets and motivation have large similarities. Because of this, the first part of the introduction, Chapters 2.1-2.4 and the entire Chapter 3 might have some resemblance.

Trondheim, June 2023

André Egeli

Table of Contents

Chapter 1

Introduction

Solar radiation forecasting has traditionally not been a focus of the meteorological community in Norway. While global solar radiation maps are available, they only cover the southern part of Norway, and the Norwegian Meteorological Institute does not offer solar radiation forecasts to the public. However, there is increasing interest in solar radiation forecasts in Norway due to the potential for solar energy to play a significant role in the country's energy mix. Multiconsult has conducted a study that suggests solar energy could be as significant for Norway as hydropower currently is (Cato, 2022). Accurate solar radiation forecasts are crucial for predicting solar panel power output and enabling grid operators to balance energy generation and consumption. Solar forecasting is also considered an important aspect of integrating solar energy into power grids (Zwane et al., 2022).

Solar radiation refers to the energy emitted by the sun in the form of electromagnetic waves. This energy travels through space and reaches the Earth's atmosphere, where it interacts with the atmosphere, clouds, and the Earth's surface (US Department of Energy, 2017). The property of our interest is the total solar radiation that reaches a horizontal plane on the earth's surface, known as Global Horizontal Irradiance (GHI). GHI is the sum of direct normal radiation (DNI), diffuse radiation (DHI), and ground-reflected radiation (HOMER, 2017). GHI is predictable due to its dependence on the sun's angle in the sky, time of year, and the amount of diffuse and reflected radiation. By taking these factors into account, it is possible to calculate the physical constraints of solar radiation (C. N. Long, 2010).

In a recently published paper by Gneiting et al. (2023), it is stated that the future of solar forecasting is probabilistic. Probabilistic forecasting is a technique that summarizes future events by assigning a probability distribution to potential outcomes. This technique generates a complete set of probabilities that is called a probabilistic forecast. According to Raftery et al. (2005), the goal of probabilistic forecasting is to produce a probability density function (pdf) that is sharp and calibrated. The same paper defines calibration as the consistency between the forecast pdf and the actual outcomes, and sharpness as the concentration of the forecast pdf. Specifically, a probabilistic forecast is calibrated if an event with probability p of occurring occurs a proportion of p times on average, and sharp if the prediction intervals are narrower than those obtained using simple methods such as climatological data.

Egeli (2022) provides a detailed analysis of the three different methods of forecasting solar radiation. They conclude that for Norway, the only viable method is Numerical Weather Prediction (NWP). According to NOAA, newer NWP systems incorporate ensemble prediction to improve the accuracy and reliability of NWP forecasts. This involves running the same model multiple times with varying initial conditions. However, ensemble-generated forecasts can be unreliable due to their tendency to be under-dispersive and biased (Gneiting et al., 2005). Under-dispersive forecasts are those in which the actual observation falls outside the range predicted by the forecast, while biased forecasts contain systematic errors (Wang et al., 2018). Therefore, statistical post-processing is crucial to obtain reliable solar radiation forecasts. The selection and application of post-processing methods depend on the specific context and requirements.

Egeli (2022) investigated the usage of Bayesian Model Averaging (BMA) to combine several ensemble forecasts into a probabilistic forecast. Specifically, they considered 12-hour ahead forecasts of hourly irradiance at weather stations in Norway based on the operational high-resolution NWP model run by the MEPS (MetoCoOp Ensemble Prediction System). They modeled each ensemble member as following a Beta distribution, scaled by the physical limits of solar radiation, to generate probabilistic forecasts. However, post-processed probabilistic forecasts were only made for locations where observational data exists, which in Norway is a limited number of weather stations (Bakketun and Kristiansen, 2018). To ensure grid reliability, solar radiation forecasts must exist in the locations where solar panels are placed, not where the weather stations are located (Yang et al., 2022).

To address this challenge, spatial interpolation techniques of post-processed forecasts are needed for locations where observational data is not available. There exist several applicable methods, from the simplest methods using a weighted

average of the closest observations, to more advanced methods with estimating parameters and varying them spatially (Gaetan and Guyon, 2010). The model utilized in this thesis is a Bayesian hierarchical model with a Beta distributed likelihood, scaled by the physical constraints of solar radiation. The latent field is the spatially varying coefficients. The explanatory variable for the model is the mean of the ensemble members, as it simplifies the analysis and interpretation of the parameters and is often considered a more reliable and skilled estimate compared to the members of the ensemble (Christiansen, 2019). The ensemble mean is hereafter referred to as the NWP forecast.

Spatial varying coefficient models (SVCM) are a class of statistical models used to analyze spatially referenced data, where the coefficients in the regression models are allowed to vary in space (Gelfand et al., 2003). Spatial varying coefficient models have been widely applied in a large range of fields, including geography (Su et al., 2017), ecology (Finley, 2011), econometrics (Bitter et al., 2007), and environmental science (Waller et al., 2007). Finally, and of great importance to this research, spatially varying coefficients have been used in the literature for weather variables, among them simulations from a hydrological model (Roksvåg et al., 2022).

For inference, a Bayesian framework with a hierarchical model is used (Blangiardo and Cameletti, 2015). On the first level, the observational likelihood is a Beta distribution with a logit link function. On the second level, the latent model is the spatially varying parameters, modeled as Gaussian Random Fields (GRFs). At the third and last level, the model parameters need suitable priors. To estimate these parameters, a sliding window methodology is employed, with a training period of 20 days. The parameters are then estimated by the Stochastic partial differential equation (SPDE) approach of Lindgren et al. (2011), which employs an approximation of a Gaussian random field with parameters comparable to those of a Matérn covariance function. The software used in this study is R-INLA, which performs fast approximate inference for Latent Gaussian Models (LGMs) (Gómez-Rubio, 2020).

To evaluate the probabilistic solar radiation forecasts generated, we focus on the measures of sharpness and calibration. The Probability Integral Transform (PIT) histogram is used to assess the calibration of the forecasts (Raftery et al., 2005). The PIT is a measure of the accuracy of the predictions of a model, by comparing the predicted probabilities of an event occurring with the actual frequency of that event (Baran and Lerch, 2016). The PIT histogram is a graphical representation of the PIT values and can be used to identify areas where the model may need to be improved, as well as to evaluate the reliability of the predictions (Lauret et al., 2019).

To quantitatively compare and rank the models, we use the Continuous Ranked Probability Score (CRPS). CRPS is a statistical measure that evaluates the accuracy of NWP forecasts by measuring the difference between the predicted probabilities of an event occurring and the observed frequency of that event. Therefore, it is a method to assess both calibration and sharpness and is commonly used by the weather forecasting community (Lauret et al., 2019). Furthermore, the Continuous Ranked Probability Skill Score (CRPSS) is used to compare how well a model performs over a baseline model, where the baseline model is a simple model like the raw forecast or the climatology. The CRPSS then represents the increase, or decrease, in the accuracy of the post-processed models compared to the baseline model. Lastly, root-mean-squared error (RMSE) is used to measure the deterministic accuracy of a point forecast.

The main goal of this thesis is to propose a method for probabilistic solar radiation forecasts. The proposed method relies on NWP forecasts and ground observations and utilizes the physical constraints of solar radiation to scale the possible domain. The model can be applied both for locations with and without ground observations. It is outside the scope to model the dependency in solar radiation forecast errors between locations.

This work presents a spatial statistical method for probabilistic solar radiation forecasting, utilizing a dataset comprising 2.5 years of NWP data and in situ ground observations from 15 locations in Norway. The proposed model is a spatially varying coefficient model (SVCM) that leverages a sliding window training period of 20 days to estimate the parameters. In Chapter 2, we provide an overview of the theory relevant to solar radiation, probabilistic forecasting, the model used, and the inference methods. In Chapter 3, we present and visualize the data and conduct a brief exploratory analysis of the NWP forecasts. In Chapter 4, we describe the methodology behind the post-processing methods and provide an overview of the model. The results of the post-processed forecasts are presented in Chapter 5. Finally, in Chapter 6, we discuss the results and provide a conclusion.

Chapter 2

Background

This Chapter provides an overview of the theories relevant to solar radiation, probabilistic forecasting, spatial statistics and the inference method. We first introduce the physical quantity of solar radiation. Then we present the theories behind probabilistic forecasting and the scoring rules used to evaluate the models. Finally, we outline the theory behind the spatial models and the inference methods for estimating the parameters in the spatially varying coefficient model. For notation, we use $x_t(s)$ to denote the NWP forecast at location s, forecasted for time t, and $y_t(s)$ to denote the corresponding observation at location s measured at time t. In this Section, we omit the notation for t as we primarily focus on a single day. However, some equations involve multiple training days.

2.1 Solar radiation

One way to describe the position of the sun is by using a horizontal coordinate system that includes two angles: the zenith angle (θ) and the azimuth angle (ϕ). The zenith angle represents the angle at which the sun is shining on the observer, with 0° directly overhead. On the other hand, the azimuth angle represents the direction from which the sun is shining, with 0° from the north and 180° from the south. A reference system for these angles can be seen in Figure 2.1. To calculate these angles for a specific time and location, NREL's SPA algorithm is commonly used (Reda and Andreas, 2004). In particular, sunset and sunrise correspond to when the zenith angle reaches 90 degrees.

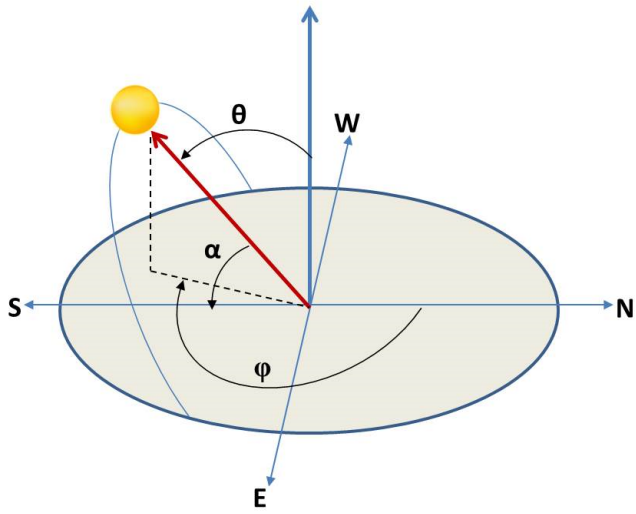


Figure 2.1: The Figure illustrates the solar position relative to the observer. θ is the zenith angle and ϕ is the azimuth angle.

The Baseline Surface Radiation Network (BSRN) has created a toolkit that can determine the physical and exceedingly rare constraints of solar radiation based on the latitude of the observation station, the time of day, and the date (C. N. Long, 2010). These limitations represent the extreme minimum and maximum amounts of solar radiation, which account for all sources. The physical limit can be determined using the following equation, where $-4W/m^2$ denotes radiative cooling at night, μ_0 is the measured zenith angle in radians, and S_a is the extraterrestrial radiation, which averages at 1361 W/m^2. This limit is known as the physical possible limit (PPL).

$$PPL: \quad -4W/m^2 < GHI < S_a \cdot 1.5\mu_0^{1.2} + 100W/m^2 \tag{2.1}$$

The extremely rare limits are less frequently reached by observations and should be carefully examined.

$$ERL: \quad -2W/m^2 < GHI < S_a \cdot 1.2\mu_0^{1.2} + 50W/m^2 \tag{2.2}$$

The extremely rare limit is called ERL. This paper will make frequent use of

these limits to scale NWP forecasts and observations within the range of 0 to 1. For simplicity, we define the lower limit as 0, making the scaling parameter equivalent to the upper limit of solar radiation, PPL.

2.2 Probabilistic forecast

"Forecasts characterize and reduce but generally do not eliminate the uncertainty. Consequently, forecasts should be probabilistic in nature, taking the form of probability distributions over future events." (Dawid, 1984)

As previously stated, the main objective of probabilistic forecasting is to optimize the forecast pdfs' sharpness while ensuring calibration. To measure sharpness and calibration, we rely on two metrics: The PIT histogram and the CRPS. The PIT histogram is a graphical tool that helps evaluate calibration, whereas the CRPS is a proper scoring rule that accounts for both sharpness and calibration concurrently (Lauret et al., 2019). Proper scoring rules are metrics that maximize the expected score when the forecasted distribution equals the true distribution (Gneiting et al., 2007). For assessing deterministic point forecasts, we use RMSE, which quantifies the deviation between a predicted value and an observed value and is not a probabilistic scoring rule.

To evaluate the performance of these metrics, we will employ cross-validation. Specifically, we will calculate PIT histograms, CRPS, and RMSE for each fold of the cross-validation. By utilizing cross-validation, we can assess the model's performance on unseen data and evaluate its generalization ability. The resulting scores will provide insight into the model's accuracy across different folds of the data.

2.2.1 Probability Integral Transform histogram

The Probability Integral Transform (PIT) histogram is a common technique for assessing forecast calibration. According to Baran and Lerch (2016), the PIT is defined as the predictive cumulative density function's value computed at the verifying observations. Calibrated probabilistic forecasts generate a PIT histogram that closely approximates a uniform distribution, indicating that values modeled from a continuous distribution can be transformed into random variables with a standard uniform distribution.

The PIT histogram's shape offers insight into the forecast's calibration. When the

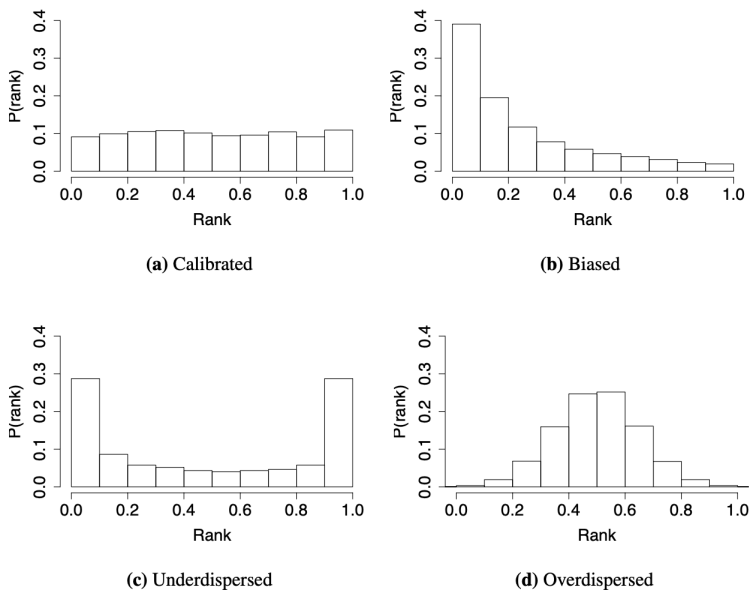


Figure 2.2: Figure of the different types of PIT-histograms. (Kleiven, 2017)

predicted interval is excessively broad, the PIT histogram exhibits a hump shape, referred to as overdispersion. Conversely, if the predicted interval is too narrow, the PIT histogram displays a U-shape, called underdispersion. In addition, biased forecasts can shift the histograms to the left or right (Kleiven, 2017). Figure 2.2 provides examples of these various scenarios.

It is important to bear in mind the difficulty of evaluating probabilistic forecasts, especially by utilizing the PIT-histogram. Hamill (2001) made an example where all probabilistic forecasts were biased, and still the PIT values were uniform. "His example aimed to show that the uniformity of the PIT values is a necessary but not a sufficient condition for the forecaster to be ideal." (Gneiting et al., 2007). Therefore, we need a new and sufficient metric to assess the probabilistic forecasts.

2.2.2 Continuous Ranked Probability Score

"Scoring rules assign numerical scores to probabilistic forecasts and form attractive summary measures of predictive performance, in that they address calibration

and sharpness simultaneously." (Gneiting et al., 2007)

Proper scoring rules play a vital role in assessing the accuracy of probabilistic forecasts. They work by maximizing the expected score when the issued forecast is the true distribution of the quantity being forecasted. For more information on proper scoring rules, we refer to Gneiting and Raftery (2007), which this Section is built on. The Continuous Ranked Probability Score (CRPS) is a commonly used proper scoring rule that considers the entire distribution of predicted probabilities and is particularly useful in meteorology (Jordan et al., 2017). CRPS takes into account the entire distribution of predicted probabilities rather than just the point estimate, allowing for a direct comparison of forecasted probabilities to the observations. This is in accordance with the prequential principle from Dawid (1984), where it is stated that the assessment of probabilistic forecasts needs to be based on the predictive distributions and the observations only. If F is the predictive cumulative density function and y is the corresponding observation, the CRPS is defined as:

$$CRPS(F, y) = \int_{\mathbb{R}} (F(x) - 1(x \geq y))^2 dx, \tag{2.3}$$

Gneiting and Raftery (2007) showed an alternative representation of the CRPS as:

$$CRPS(F, y) = E_F|X - x| - \frac{1}{2} E_F|X - X^{'}|, \tag{2.4}$$

where X and X' are independent copies of a random variable with cumulative distribution function F and finite first moment. This alternative CRPS representation reduces to Mean Absolute Error (MAE) if F is a point forecast.

The CRPS is expressed in the same unit as the observed variable and measures the difference between the predicted and observed cumulative distributions, with smaller values indicating better performance. Closed-form solutions exist for calculating the CRPS for many distributions, but not for all (Zamo and Naveau, 2018). In such cases, we turn to simulation-based approaches to approximate the score. To calculate the CRPS for simulated samples from a distribution, we use the following method:

$$CRPS \approx \frac{1}{M} \sum_{m=1}^{M} (F_m(y) - [y \geq x_m])^2 \tag{2.5}$$

Here, M is the number of simulated samples, F_m is the cumulative density function for the mth sample, y is the observed value, and x_m is the predicted value for the mth sample.

We rank competing forecasting models based on its average over N days for a given location, where F_t represents the predictive cdf of y_t for each day (Gneiting and Katzfuss, 2014):

$$CRPS = \frac{1}{N} \sum_{t=1}^{N} CRPS(F_t, y_t) \tag{2.6}$$

We also compare the performance of different models to a baseline model using the Continuous Rank Probability Skill Score (CRPSS) (Lauret et al., 2019). The CRPSS is defined as:

$$CRPSS = \left(1 - \frac{CRPS_i}{CRPS_0}\right) \cdot 100 \tag{2.7}$$

where $CRPS_0$ is the average CRPS for a baseline model, and $CRPS_i$ is the average CRPS for model i. The CRPSS indicates the percentage increase or decreases in accuracy compared to the baseline model. A value of 100 represents perfect skill compared to the baseline model, while a value of 0 indicates no skill compared to the baseline model. Negative values indicate negative skill compared to the baseline model.

2.2.3 Root-Mean Squared Error

The Root-mean squared error (RMSE) can be used to measure the accuracy of a point forecast. It is calculated as the square root of the mean squared error between the predicted value, $\hat{y}$, and the true value, y. In the context of solar radiation forecasting at a single location, if the observed solar radiation at time t is y_t and the corresponding point forecast is $\hat{y}_t$, the RMSE can be calculated over a set of N observations as follows:

$$RMSE = \sqrt{\frac{\sum_{t=i}^{N} (\hat{y}_t - y_t)^2}{N}} \tag{2.8}$$

RMSE is expressed in the same units as the predicted and true values and a smaller value indicates a more accurate forecast.

2.2.4 Cross-validation

Cross-validation (CV) is a resampling method, where it uses different portions of the data to test and train a model (Bivand et al., 2008). It is used in conjunction with our other quantitive metrics, namely CRPS and RMSE, to assess predictive performance on untrained data. There exist mainly two types of cross-validation suitable for our task: K-fold cross-validation and leave-one-out cross-validation (LOOCV). The K in k-fold cross-validation refers to the number of groups that a given data sample is split into. A figure of a 5-fold CV can be seen in Figure 2.3.

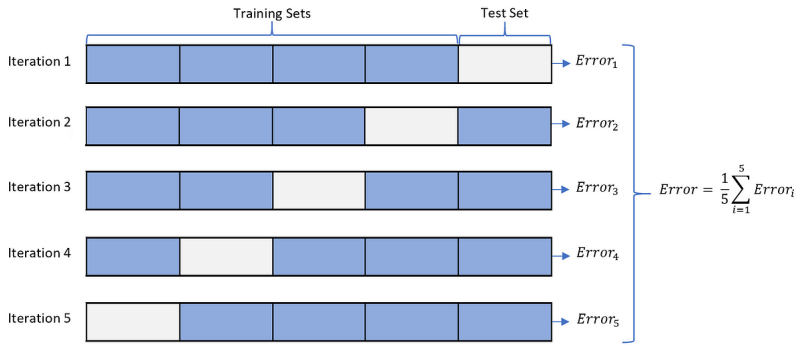


Figure 2.3: The figure displays a 5-fold cross-validation technique, where the dataset is iterated five times and tested on a new sample in each iteration. The resulting error is calculated as the mean of the five iterations.

Cross-validation is not an evaluation metric in itself but is rather a method to get an unbiased skill score based on the data available. The formula for RMSE in K-fold CV is:

$$MSE_{CV} = \sum_{k=1}^{K} MSE_i, \tag{2.9}$$

where i relates to split number i. If K is equal to the sample size, it is called LOOCV.

2.3 Beta distribution

The Beta distribution is a commonly used probability distribution in cases where a random variable has values limited to the domain between 0 and 1. It is frequently utilized for modeling percentages or fractions (Abonazel et al., 2022). This Section is based on the presentation of Beta regression by Ferrari and Cribari-Neto (2004).

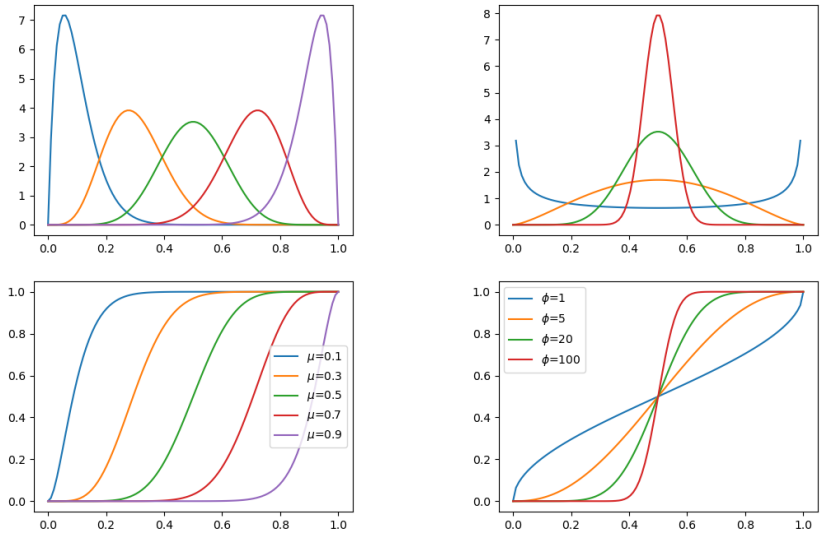
The probability density function (pdf) of the Beta distribution is given by the following formula, where $0 \leq x \leq 1$ and $\alpha, \beta > 0$ are the shape parameters:

$$g(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \tag{2.10}$$

Here, $\Gamma(\cdot)$ is the Gamma function, which can be defined as $\Gamma(n) = (n-1)!$. To provide a clearer understanding of the Beta distribution, we can adopt the approach used in Beta regression and reparameterize the distribution to a mean parameter μ and a precision parameter ϕ. As a result, we obtain the following probability density function:

$$g(x; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \tag{2.11}$$

Here, μ and ϕ represent the mean parameter and precision parameter, respectively, with $\mu = \frac{\alpha}{\alpha+\beta}$ and $\phi = \alpha + \beta$. To convert these parameters back to the shape parameters α and β, it is a simple matter of using the formulas $\alpha = \mu\phi$ and $\beta = \phi(1-\mu)$. The precision parameter ϕ indicates the distribution's variability, with larger values corresponding to smaller variance. A figure depicting the Beta probability density function and cumulative distribution function for various values of μ and ϕ is available in Figure 2.4. This Figure also depicts the flexibility of the Beta distribution.



(a) ϕ fixed at 20

(b) μ fixed at 0.5

Figure 2.4: The upper and lower figures show the probability density function (pdf) and the cumulative distribution function (CDF) for a Beta distribution with different values for the parameters μ and ϕ. In the left figure, the parameter ϕ is fixed at 20 while the parameter μ has five different values. In the right figure, the parameter μ is fixed at 0.5 while the parameter ϕ has four different values. The pdf and CDF illustrate the shape of the Beta distribution for each combination of μ and ϕ values.

The Beta distribution provides the mean and variance of a random variable X as:

$$E(X) = \mu$$
$$Var(X) = \frac{\mu(1-\mu)}{1+\phi}$$

The variance function demonstrates that the precision parameter ϕ decreases

the variance for fixed mean values. Therefore, ϕ is referred to as a precision parameter.

The mean μ of the Beta distribution must be bounded between 0 and 1 since it represents the distribution's mean. To guarantee that μ stays within the domain, we can use a linear predictor, and find the mean through a link function. For N explanatory variables x_i, we assume that μ can be expressed as:

$$\eta = \alpha + \sum_{i=1}^{N} \beta_i x_i = g(\mu), \tag{2.12}$$

where α is the intercept and $\beta_i, i \in [1, N]$ is the coefficients for the explanatory variables. Furthermore, $g(\cdot)$ represents a link function that defines the relationship between the mean μ and the linear predictor η. The logit function is a common choice for the regular Beta regression model because it confines the domain between 0 and 1.

It is possible to adjust the range of the Beta distribution by using a linear transformation. By substituting the variable x in the standard form of the Beta distribution with a new variable y defined on the desired range $[a, c]$ using the transformation $y = x(c - a) + a$, the distribution can be scaled. The resulting scaled Beta distribution has a probability density function (pdf) given by:

$$f(y; \mu, \phi, a, c) = \frac{f(x; \mu, \phi)}{c - a}, \tag{2.13}$$

where $f(x; \mu, \phi)$ is the pdf of the standard Beta distribution and x can be obtained from y through the transformation $x = \frac{y-a}{c-a}$. Note that the mean, μ, of the Beta distribution should also be within the range [a, c]. To achieve this, a linear transformation can be applied to μ resulting in a new variable, μ_t, that represents the transformed μ bounded within the range [a, c]. This is done using the transformation $\mu_t = \mu(c - a) + a$. The mean, μ_t, for the scaled Beta distribution, can be calculated as:

$$\mu_t = \mu(c - a) + a = \text{logit}^{-1}(\eta)(c - a) + a \tag{2.14}$$

where logit^{-1} is the inverse of the logit function, and η is the linear predictor obtained from a Beta regression model that links the mean μ to a set of explanatory variables $x_i, i \in [1, N]$.

Since the parameters α and $\beta_i, i \in [1, N]$ are converted through a logit link function, the interpretation of these parameters can be difficult. For simplicity assume one intercept parameter α and one slope parameter β linked to an observation x, such that the mean is defined as $\mu = \text{logit}^{-1}(\alpha + \beta x)$. Notably, α represents the mean when $x = 0$, i.e. $\mu = \text{logit}^{-1}(\alpha)$. However, since it is transformed non-linearly through the logit link function, α represents the mean in a non-linear way. For $x = 0$ and negative α, the mean approaches 0 when $\alpha \to -\infty$, and 1 when $\alpha \to \infty$. For $\alpha = 0$ and $x = 0$, the mean is 0.5.

The slope parameter β is more difficult to interpret as it also relies on the corresponding x value. Figure 2.5 presents various examples of the mean μ with different α and β values. Three plots are made, each with a fixed α and four different β values. Note that the mean μ equals 0.5 when the linear predictor η equals 0, and the shape of the Beta distribution is based on the precision parameter ϕ as depicted in Figure 2.4b.

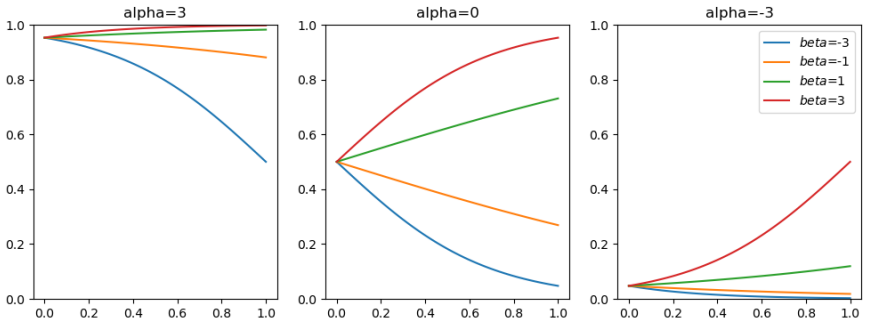


Figure 2.5: The Figure displays 12 μ outcomes, given x, for varying α and β combinations. The mean μ is presented on the y-axis, calculated as $\mu = \text{logit}^{-1}(\alpha + \beta x)$, while the x-axis is the x-value. α is fixed at 3, 0, and -3 in the left, middle, and right panels, respectively. Four β values (-3, -1, 1, and 3) are denoted by blue, orange, green, and red, respectively. Notably, β is more significant with smaller α, and vice versa.

When both α and β are 0, the mean μ is 0.5, and the value x has no impact on the model. When α is large and β is small, the intercept dominates and x has little influence on the mean μ. Conversely, if α is small and β is large, x has a significant impact on the mean μ.

2.4 Bayesian Model Averaging

Bayesian model averaging (BMA) is a widely used technique for post-processing ensemble forecasts, first introduced by Raftery et al. (2005). This method combines multiple probability density function (pdf) forecasts to generate a probabilistic forecast. In BMA, each member of the ensemble is treated as a statistical model represented by the pdf $g(y|x_m)$. The resulting BMA pdf is obtained by taking a weighted average of the individual pdfs, where the weights w_m satisfy the condition $\sum_{m=1}^{M} w_m = 1$. Assuming an ensemble member x_m with its associated pdf $g(y|x_m)$, the resulting BMA forecast can be expressed as:

$$f(y|x_1, x_2, ..., x_M) = \sum_{m=i}^{M} w_m g(y|x_m), \tag{2.15}$$

The weights can be estimated using a training period, typically by utilizing a sliding window methodology. In short, the N days preceding each prediction day i are used as the training period, covering the time interval from day $t = i - N$ to day $t = i - 1$. On the other hand, all ensemble members can be considered equally important by setting $w_m = 1/M$ for all m. When applying BMA to ensemble weather forecasting, it is commonly assumed that the ensemble members are exchangeable, meaning that they are equally plausible future scenarios (Broecker and Kantz, 2011). In this case, the parameters for the distributions $g(y|x_m)$ are the same for all ensemble members. For a more comprehensive introduction to BMA, we recommend the work of Raftery et al. (2005).

2.5 Spatial statistics

Spatial statistics is a large domain and is covered in many books. This Section and the two next Sections about spatial statistics, Bayesian hierarchical models, and Gaussian random fields are mainly based on the books by Blangiardo and Cameletti (2015) and Gaetan and Guyon (2010).

Spatial data are data with known spatial locations that can be modeled as realizations of a stochastic process indexed by space. The stochastic process is denoted as:

$$Y(s) = y(s), s \in D, \tag{2.16}$$

where D is a subset of $\mathcal{R}^d$. In spatial statistics, we usually divide them into three main types: Area or lattice data, point-referenced data, and spatial point patterns. With observations taken from a specific location, we deal with point-referenced data. Here we let $y(s)$ be a random outcome at a specific location, and the spatial index s can vary continuously in the fixed domain D. The location s is typically represented as a two-dimensional vector with latitude and longitude. The actual data are a collection of observations represented by $y = (y(s_1), \ldots, y(s_n))$, where the set $(s_1, \ldots, s_n)$ indicates the locations at which the measurements were taken. We are interested in predicting the outcome at unobserved locations within the domain D.

2.6 Bayesian hierarchical modelling

Given three random variables U, V and W, we can always decompose the joint distribution of the triplet (U, V, W) by successive conditioning:

$$[U, V, W] = [W|U, V][V|U][U] \tag{2.17}$$

When examining a complex process, one approach is to create a hierarchy of underlying processes or distributions to model it, rather than studying it directly (Banerjee et al., 2014). For instance, the process governing solar radiation forecasts (ξ) may depend on specific parameters (θ) that represent the spatial correlation between locations. Both ξ and θ are stochastic variables with prior and posterior distributions. A common Bayesian model for this scenario is a three-stage hierarchical model. The first stage is the observation likelihood, the second stage is the latent or process model distribution $\pi(\xi|\theta)$, and the third stage is the model parameters $\pi(\theta)$. In the formulation from Equation 2.17, we define W as the data y, V as the process ξ, and U as some (hyper)parameters θ. By using Bayes' formula with the observations y and some priors, we can infer about ξ and θ.

2.7 Gaussian random fields

Gaussian Random Fields (GRF) are often used to model environmental data or phenomena that are continuous in space (Roksvåg et al., 2020). A GRF models the spatial dependency between locations. If we have a continuous field $\{\xi(s); s \in D\}$, it is a GRF if it follows a multivariate Gaussian distribution:

$$(\xi(s_1), ..., \xi(s_n))^T \sim N(\mu, \Sigma), \tag{2.18}$$

where $s_i, i \in [1, N]$ is the locations, μ is a vector of expected values and Σ is the covariance matrix. The covariance matrix defines the spatial dependency structure of the variable of interest and is established through a covariance function $C(\cdot, \cdot)$ such that:

$$\Sigma_{ij} = \text{Cov}(\xi(s_i), \xi(s_j)) = C(s_i, s_j) \tag{2.19}$$

The covariance function $C(\cdot, \cdot)$ typically has two parameters: a marginal variance parameter σ^2 and a range parameter ρ, which describe the underlying spatial field. The marginal variance provides insights into the degree of spatial variation in the process being considered, whereas the range parameter determines the rate of decay in covariance between two spatial locations. Specifically, the range parameter indicates the distance at which the correlation between two spatial locations becomes almost negligible. A Gaussian Random Field (GRF) is considered stationary when the marginal variance and range parameters remain constant throughout the spatial domain.

To create a hierarchical framework as outlined in the preceding Section, three steps must be followed. Firstly, a distribution is defined, characterized by a set of parameters. This set of parameters includes the second step, defining a latent GRF denoted $\xi(s)$ that accounts for spatial correlation via the covariance function $C(\cdot, \cdot)$. The third and final step involves specifying the prior distributions $\pi(\theta)$ for the hyperparameters. While parameter estimation in this framework can be computationally expensive, later sections will demonstrate the availability of computationally efficient methods.

2.8 Spatially Varying Coefficient

Varying coefficient models, also referred to as varying effects models, are a type of regression model that allows for the regression coefficients to vary with one or more covariates (Hastie and Tibshirani, 1993). These models are particularly useful when the relationship between the response variable and the covariates varies across different levels of the covariates. In contrast to standard regression models, varying coefficient models model the regression coefficients as smooth functions of the covariates, rather than fixed constants. This flexibility allows for the modeling of complex relationships between variables.

The general spatially varying coefficient model (SVCM) was first introduced by Gelfand et al. (2003), and this section is based on that paper. SVCM is an extension of varying coefficient models that incorporate spatial dependence. In these models, the coefficients of the covariates are allowed to vary across space, meaning that their effect on the response variable can change depending on the location. The spatially varying coefficient surface is modeled as a realization from a spatial process located in a spatial region $D \subseteq \mathcal{R}^2$. A stationary specification is commonly used, which allows for the modeling of the desired level of smoothness in the process realization by selecting an appropriate covariance function.

Mathematically, for a single covariate $x(s)$ which varies over space, the Gaussian stationary spatial process is modeled as

$$Y(s) = \mu(s) + W(s) + \epsilon(s), \tag{2.20}$$

where s is the location, $\mu(s) = x(s)^T\beta$ and $\epsilon(s)$ is a white noise process. W(s) are viewed as spatial random effects. The general specification of this model is to split the intercept and slope to a general term and a spatially varying term, resulting in the equation:

$$Y(s) = \beta_0 + \beta_1 x(s) + \beta_0(s) + \beta_1(s)x(s) + \epsilon(s) \tag{2.21}$$

The general spatially varying coefficient model adopts a Bayesian approach for the modeling framework. This approach is particularly advantageous as it allows for inference on random spatial effects. By using the Bayesian framework, a posterior for the spatial coefficient process at both observed and unobserved locations, as well as posterior distributions for all model parameters, can be obtained.

A spatially varying coefficient can be seen as a realization of a Gaussian random field (GRF) with a Matérn covariance function. The spatially varying coefficient has the form:

$$\beta \sim \mathcal{N}(\mu, \tau^{-1}R(\phi)), \tag{2.22}$$

where μ is a constant mean and $R(\phi)_{ij} = (C(||i-j||))$. $C(\cdot)$ is a Matérn covariance function with a fixed smoothness ν, characterized by two parameters, namely the marginal variance σ and a range parameter ρ:

$$Cov(s_i, s_j) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa||s_i - s_j||)^\nu K_\nu(\kappa||s_i - s_j||), \tag{2.23}$$

where $||s_i - s_j||$ is the Euclidian distance between two locations $s_i, s_j \in D$, $\Gamma(\cdot)$ is the Gamma function, σ^2 is the marginal variance, K_ν denotes the modified Bessel function of the second kind and order $\nu > 0$. ν measures the smoothness of the process and is usually kept fixed. The parameter $\kappa > 0$ is a scaling parameter related to the range ρ, i.e. the distance at which the spatial correlation becomes almost null. Typically, this value is set at $\rho = \frac{\sqrt{8\nu}}{\kappa}$, and is defined such that the spatial correlation between two locations has dropped to 0.1 (Roksvåg et al., 2020).

The spatial varying coefficient model can be extended to include generalized linear models, which are called generalized spatially varying coefficient models (GSVCM). Instead of assuming a Gaussian distribution on the response variable, GSVCM replaces it with a distribution from an exponential family. In GSVCM, the linear predictor η, which is linked to the mean response μ via the link function $g(\cdot)$, takes the form:

$$\eta_i = \beta_0(s_i) + \sum_{j=1}^{p} \beta_j(s_i) X_{i,j} \tag{2.24}$$

Here, $\beta_j(s_i)$ is the varying regression coefficient, which can be considered a stochastic process on the effect modifier domain. When the coefficients vary continuously in space, a Gaussian random field (GRF) with a specific covariance function can be assumed.

Suppose a spatially varying coefficient model is expressed with a Beta distribution and a logit link function. Then, the spatially varying coefficient model for a beta-distributed response variable with a logit link function for the mean is given by:

$$Y_i \sim \text{Beta}(\mu_i, \phi) \tag{2.25}$$

$$g(\mu_i) = \eta_i = \beta_0(s_i) + \sum_{j=1}^{p} \beta_j(s_i) X_{i,j} \tag{2.26}$$

where $\beta_0(s)$ and $\beta_j(s)$ are the spatially varying intercept and slope coefficients respectively, and $g(\cdot)$ represents the link function.

2.9 SPDE approach to spatial modeling

As noted in section 2.7, inference in Bayesian hierarchical models has traditionally been very computationally expensive, but Lindgren et al. (2011) came up with a computationally effective method called the stochastic partial differential equation (SPDE) approach. "A key idea of this approach is to construct continuously indexed approximations of GRFs by solving SPDEs." (Haug et al., 2020) It consists of performing the computations using a Gaussian Markov random field (GMRF) representation of the GRF. GMRFs are characterized by sparse precision matrices and this feature allows us to implement computationally efficient numerical methods.

The GMRF is a discretely indexed spatial random process, and the starting point is the linear fractional SPDE:

$$(\kappa^2 - \Delta)^{\alpha/2}(\tau\xi(s)) = W(s), \tag{2.27}$$

where $s \in \mathcal{R}^d$, Δ is the Laplacian, α controls the smoothness, $\kappa > 0$ is the scale parameter, τ controls the variance, and $W(s)$ is a Gaussian spatial white noise process. The exact and stationary solution to this SPDE is the stationary GRF $\xi(s)$ with a Matérn covariance function given by Equation 2.23.

The SPDE approach uses a mesh to calculate spatial dependency between observations. A good mesh is crucial for accurate results and is constructed using the location of study points, the domain extent, and a set of polygons defining the region of interest (Røste, 2020).

The solution to the SPDE – represented by stationary and isotropic Matérn GRF $\xi(s)$ – can be approximated using the finite element method through a basis function representation defined on the meshes in the domain D:

$$\xi(s) = \sum_{g=1}^{G} \phi_g(s)\hat{\xi}_g, \tag{2.28}$$

where G is the total number of vertices of the mesh, ϕ_g is the set of basis functions and $\hat{\xi}_g$ are zero mean Gaussian distributed weights. The basis functions $\{\phi_g(s)\}$ are chosen to have value 1 at vertex g and 0 on all the other vertices. A more detailed explanation of the link between the Matérn covariance function and the SPDE approach can be found in Lindgren et al. (2011).

2.10 Inference

There are two main paradigms for estimating parameters in statistical inference. The classical approach, also called frequentist, involves estimating a fixed vector of unknown parameters $(\theta_1, ..., \theta_n)$ using methods such as maximum likelihood estimation. In contrast, the Bayesian approach treats all unknown quantities as random variables, intending to calculate the joint posterior distribution. In this approach, the parameter vector θ is also treated as a random variable, and a prior probability distribution $\pi(\theta)$ is assigned to it. The posterior distribution is then calculated using Bayes' rule:

$$\pi(\theta|y) = \frac{\pi(y|\theta)\pi(\theta)}{\pi(y)}, \tag{2.29}$$

where $\pi(y|\theta)$ is the likelihood of the data y given the parameters θ, $\pi(y)$ is the marginal likelihood and $\pi(\theta)$ is the prior distribution. By using the joint posterior distribution, it is possible to derive marginal distributions for each element in the parameter vector. This thesis utilizes Bayesian methods for inference.

2.10.1 Prior selection

The prior distribution in Bayesian models represents our beliefs or knowledge about the parameters before observing the data. It plays a crucial role in Bayesian inference, as it can strongly influence the posterior distribution and subsequently the estimates of the parameters (Krainski et al., 2018).

The choice of prior can reflect a wide range of beliefs, from strong prior information to complete ignorance about the parameters. If prior information is available, incorporating it into the model can improve the accuracy of the parameter estimates and make them more robust to small sample sizes. However, if the prior is misspecified or has a large influence on the posterior, it can lead to biased or unreliable results (Blangiardo and Cameletti, 2015). For the varying coefficient models, the prior selection is especially important as common choices of the prior might lead to overfitting (Franco-Villoria et al., 2019).

The penalizing complexity (PC) priors were first introduced by Simpson et al. (2017). Overfitting occurs when a model becomes too complex and flexible, resulting in a poor fit for new data. PC priors counteract this by penalizing the increase in complexity that occurs when a model deviates from a simpler, less flexible base model. The PC priors are specifically designed for modeling

precisions, which are the inverses of variances. For example, a Gaussian effect with a precision τ and zero mean can be represented as $\mathcal{N}(0, \tau^{-1})$. The PC prior to this precision is constructed with the density function

$$\pi(\tau) = \frac{\lambda}{2}\tau^{-\frac{3}{2}} exp(-\lambda\tau^{-\frac{1}{2}}), \quad \tau > 0, \lambda > 0 \tag{2.30}$$

The parameter λ determines how much penalty is given to deviations from a simpler model. To determine the value of λ, two quantities need to be specified: a quantile u and a probability α. The probability of the standard deviation σ exceeding the threshold u is set equal to α, expressed concisely as $\text{Prob}(\sigma > u) = \alpha$. The value of λ can then be found from $\lambda = -ln(\alpha)/u$.

Furthermore, Fuglstad et al. (2019) proposed using a joint informative PC prior to the range and marginal variance parameters of Gaussian random fields (GRFs). This approach assumes that the GRF is defined on a two-dimensional space $D \in \mathcal{R}^2$ with a Matérn covariance function characterized by the parameters ρ, σ, and ν. The joint PC prior to these parameters is:

$$\pi(\sigma, \rho) = \frac{d}{2}\tilde{\lambda_1}\tilde{\lambda_2}\rho^{-d/2-1}\exp(-\tilde{\lambda_1}\rho^{-d/2} - \tilde{\lambda_2}\sigma), \quad \sigma > 0, \rho > 0 \tag{2.31}$$

where the tail probabilities $\text{Prob}(\rho < \rho_0) = \alpha_1$ and $\text{Prob}(\sigma > \sigma_0) = \alpha_2$ are achieved by:

$$\tilde{\lambda_1} = -\log(\alpha_1)\rho_0^{d/2}$$
$$\tilde{\lambda_2} = -\frac{\log(\alpha_2)}{\sigma_0}$$

This prior is weakly informative and penalizes complexity by shrinking the range toward infinity and the marginal variance toward zero.

2.10.2 INLA

The Integrated Nested Laplace Approximation (INLA) was introduced by Rue et al. (2009). The INLA approach provides approximate Bayesian inference for a subclass of latent Gaussian models (LGMs). In particular, they focus on estimating the posterior marginals of the model parameters. Hence, instead of estimating a highly multivariate joint posterior distribution $\pi(\theta|y)$, they focus on obtaining

approximations to univariate posterior distributions $\pi(\theta_i|y)$. INLA uses a direct numerical calculation to approximate the posterior distribution. For more elaboration in INLA, we refer to Blangiardo and Cameletti (2015) and Krainski et al. (2018) which this introduction is based upon.

INLA can only be used on latent Gaussian models which represent a subclass of structured additive regression models (Blangiardo and Cameletti, 2015). In general, these models have a three-stage hierarchical structure: A likelihood model, a latent Gaussian field, and a vector of hyperparameters. Firstly we have observations $y = (y_1, ..., y_n)$ whose distribution is in the exponential family, and a mean μ_i are linked to the linear predictor η_i using an appropriate link function. The distribution of y depends on some vector of hyperparameters θ_1 and is assumed to be conditionally independent given the latent Gaussian field ξ. The vector of latent effects ξ is assumed to be a GMRF with zero mean and a precision matrix $Q(\theta_2)$. We denote the hyperparameters $\theta = (\theta_1, \theta_2)$. We can write the likelihood as:

$$\pi(y|\xi, \theta_1) = \prod_{i \in I} \pi(y_i|\xi_i, \theta_1) \tag{2.32}$$

The latent field is the second level in the model and is characterized by a multivariate Normal distribution

$$\xi|\theta_2 \sim \text{MVNormal}(0, Q^{-1}(\theta_2)) \tag{2.33}$$

where $Q^{-1}(\theta_2)$ is the precision matrix of the latent Gaussian field. Finally, at the last level of the hierarchical structure, appropriate priors are assigned to the hyperparameters θ.

INLA aims to approximate the posterior marginals of the model effects and hyperparameters. By exploiting the computational properties of GMRF, and using Laplace approximation, the joint posterior distribution for both the latent field ξ and the hyperparameters θ is expressed as:

$$\pi(\xi, \theta|y) \propto \pi(\theta)\pi(\xi|\theta) \prod_{i \in I} \pi(y_i|\xi_i, \theta) \tag{2.34}$$

$$\propto \pi(\theta)|Q(\theta)|^{1/2} \text{exp}\left(-\frac{1}{2}\xi^T Q(\theta)\xi + \sum_{i \in I} \log(\pi(y_i|\xi_i, \theta)) \right) \tag{2.35}$$

Here $Q(\theta)$ represents the precision matrix of the latent effects. We can now calculate the marginal posterior pdfs for each element of the latent Gaussian field and the hyperparameters:

$$\pi(\xi_i|y) = \int \pi(\xi_i|\theta, y)\pi(\theta|y)d\theta, \tag{2.36}$$

and

$$\pi(\theta_j|y) = \int \pi(\theta|y)d\theta_{-j}, \tag{2.37}$$

The procedure to calculate these marginals can be found in Blangiardo and Cameletti (2015).

Chapter 3

Forecasts and observations of solar radiation in Norway

In this chapter, we provide an overview of the case study, including the format and locations, as well as the observations and forecasts. We present visual plots of both the observations and NWP forecasts and discuss the pre-processing of the data.

3.1 Overview

The current study utilizes data from two distinct sources for observations and one source for NWP forecasts. The observations were obtained from LMT (landbruks-meterologisk tjeneste) and Eklima, both of which measure the total radiation in W/m^2 hitting the ground. A total of 40 observation stations were randomly selected from the public catalogs available from LMT and Eklima. More specifically, the data was obtained from 18 stations from Eklima (https://seklima.met.no) and 22 stations from LMT (https://lmt.nibio.no), and the locations are depicted in Figure 3.1.

The MEPS (MetoCoOp Ensemble Prediction System) provided the forecasts for this project, and the data was obtained through Thredds, which is the API for data retrieval offered by the Norwegian Meteorological Institute. The MEPS forecasts retrieved consist of six ensemble members and are generated daily at 00:00

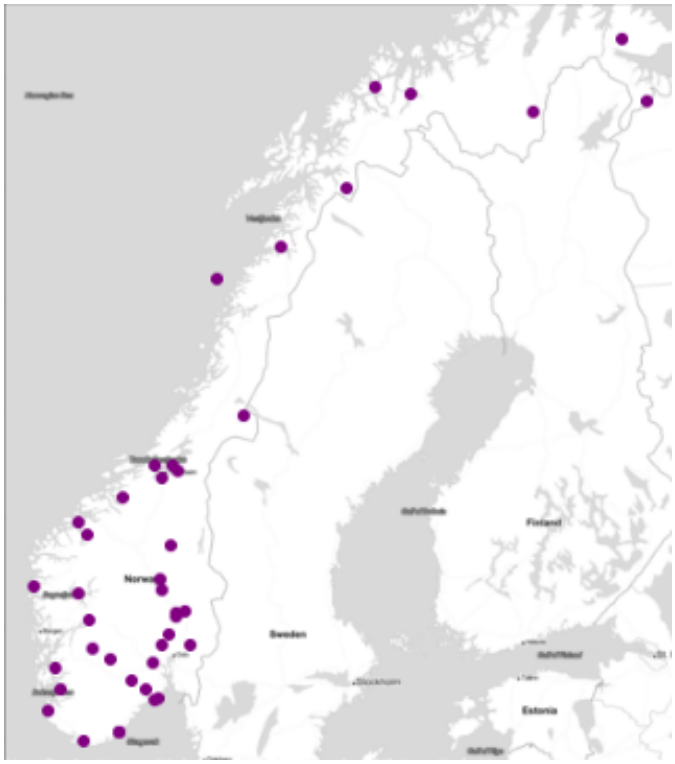


Figure 3.1: A Visualization of the 40 weather stations where the data have been retrieved from.

UTC, covering a lead time of 54 hours from January 1st, 2020, to June 1st, 2022. The notation used in this study represents an ensemble member m at location s, issue time i, and lead time l denoted as $x_{iml}(s)$, with the corresponding ground observation $y_{(i+l)}(s)$. However, our focus is solely on lead time 12, simplifying the notation as $t = i + l$. Therefore, the ensembles can be denoted as $x_{tm}(s)$ and the observations as $y_t(s)$. Furthermore, we adopt the mean of the six ensemble members, represented as $\hat{x_t(s)} = \frac{1}{M}\sum_{m=1}^{M} x_{tm}(s)$, as the forecast value in our model. To streamline notation, we set $x_t(s)$ to denote the mean of the ensembles, representing the NWP forecast at location s and time t.

3.2 Observations

LMT is an initiative led by Nibio (https://www.nibio.no), which focuses on collecting meteorological data for research and warning services in Norwegian agricultural and horticultural regions. Global horizontal irradiance was measured using pyranometers, which are specialized devices that quantify solar radiation on a flat surface in units of W/m^2. Two types of pyranometers, namely CM11 and CM3, were employed. Figure 3.2 illustrates a photograph of the CM11 pyranometer.

Figure 3.2: An image of a CM11 pyranometer from Kipp&Zonen.

Eklima is a service provided by the Norwegian climate service center (KSS) that collects and organizes climate and hydrological data for use in climate adaptation and research. The observations from Eklima are collected using the same method as the observations from LMT, as all the observation stations used in this study are owned by Nibio. The observations have an hourly resolution and can be used to visualize daily and seasonal trends.

The observations exhibit a significant skew towards zero values, which can be attributed to various factors. One possible explanation is the region's geographic location, which may experience fewer sun hours and longer nights, resulting in lower solar radiation levels. This trend may be particularly evident during the winter months. Furthermore, a majority of the observations fall below the 800 GHI threshold. A relative frequency histogram of the observations can be seen in Figure 3.3.

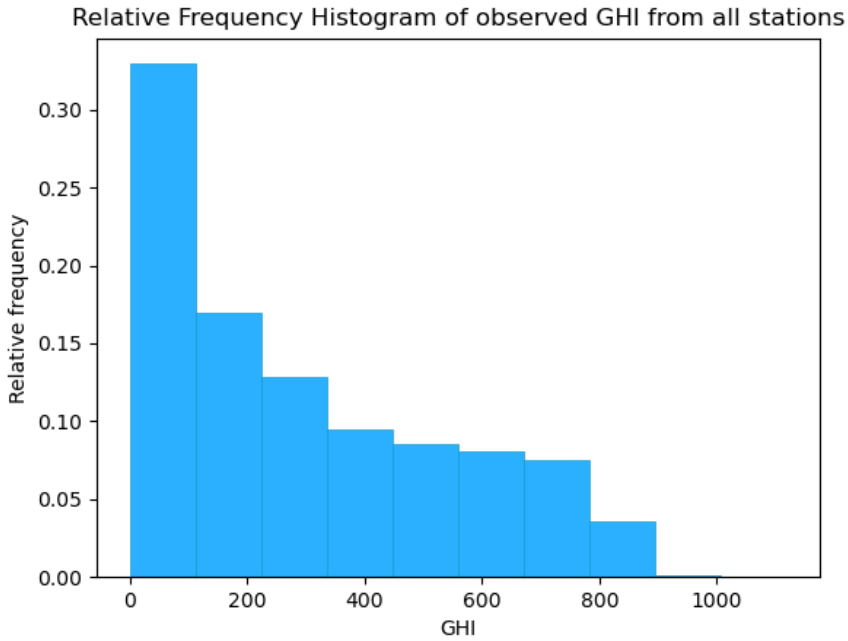


Figure 3.3: A histogram of measured global horizontal irradiance from all stations, measured at 12.00 each day.

Our dataset comprises a total of 338820 observations, and we conducted a thorough analysis of the entire dataset to identify any observations that may fall outside the theoretically possible range, as described in Section 2.1. Figure 3.4 presents the outcomes of this analysis, indicating that certain observations may be invalid. Out of the total observations, 33 were deemed to be outside the physically possible range and were removed, whereas 528 observations were identified to be beyond the extremely rare limits. The proportion of such observations is 0.156% of the dataset, which is considered to be negligible and hence retained for further analysis.

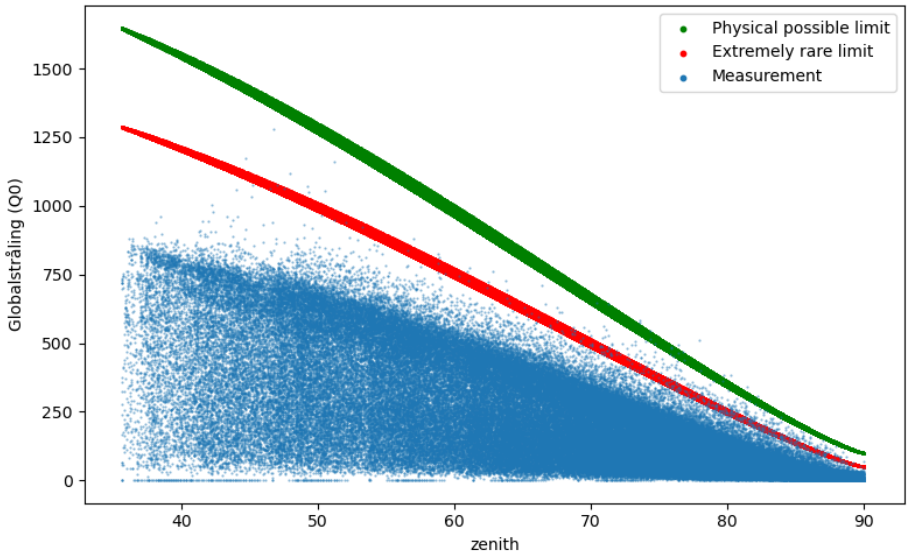


Figure 3.4: The BSRN check of physical and extremely rare limit for the observations measured by stations owned by Nibio. The blue dots are observations, while the red and green line is respectively the extremely rare and physical possible limit.

3.3 Ensemble forecasts

The ensemble forecast used in this research is obtained from MEPS, which is based on the HARMONIE-AROME model's ensemble version (Frogner et al., 2019). The HARMONIE-AROME model comprises two primary radiation schemes: short-wave and long-wave. The short-wave scheme is primarily responsible for simulating global horizontal irradiance, and the physical mechanisms underlying these schemes are described in ECMWF (2016). The ensemble comprises five members in addition to a control ensemble. The control ensemble is generated without any model perturbations, as specified in WMO (2012).

In Figure 3.5, a plot showing the relationship between the zenith angle and the mean of the ensembles is presented, together with the physical possible limit and the extremely rare limit. A crucial observation is that solar radiation forecasts consistently fall below the physical and extremely rare limits.

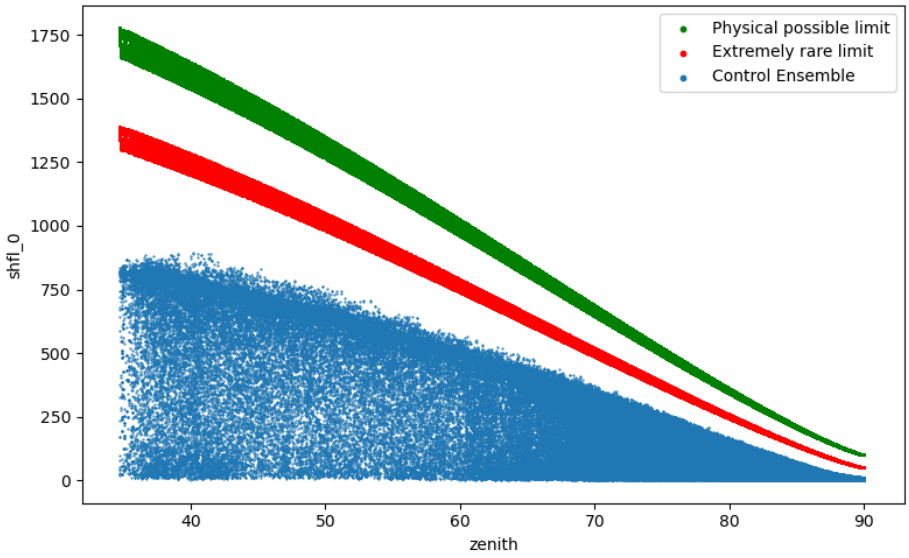
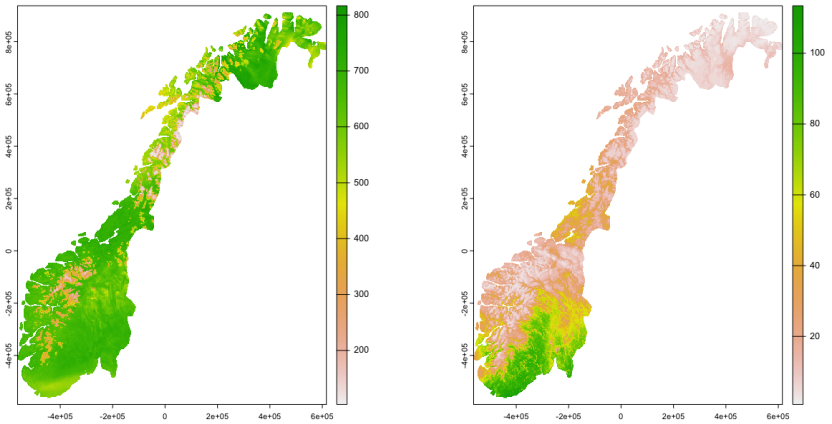


Figure 3.5: The BSRN check of physical and extremely rare limit for the mean of the ensembles made by MEPS. The blue dots are forecasted values, while the red and green line is respectively the extremely rare and physical possible limit.

This study utilizes the NWP forecast obtained from MEPS for two main objectives. Firstly, the forecasts are employed for statistical post-processing to enhance forecast accuracy by addressing systematic biases and errors. This is accomplished through the development of statistical models. Secondly, calibrated and probabilistic weather forecasts are generated using the NWP forecast to produce spatial maps across the desired domain. This involves extracting the forecasts for the entire domain of interest. This is a large amount of data, as the data comes in a 2.5x2.5km gridded format, and our study case is Norway. In total, the bounding box is 600x700 pixels large, resulting in 420 000 forecasted values for a single lead time. It is important to note that due to memory limitations, the full extent of the data is only used on selected days for visualization purposes. In Figure 3.6, the spatial extent of the NWP forecast can be seen with both a sunny day and a winter day.



(a) Summer and sunny day.

(b) Winter day.

Figure 3.6: The Figure illustrates the geographical coverage of the NWP forecasts for a summer day (6th of June 2022) to the left and a winter day (2nd of February 2022) to the right. The contrast between the two maps is striking, as there is a significant difference in solar radiation forecasts between the two seasons. Specifically, during winter in the northern regions, the solar radiation forecast is almost 0.

3.4 Pre-processing

The occurrence of missing values is a common issue in ground observations and can be caused by instrument malfunctions or poor measurements. In our dataset, we observed that some locations had a large portion of missing values. To handle this, we removed locations with more than 5 missing values. This resulted in a final dataset of 15 locations, and a list of these locations and their metadata can be found in the Appendix.

Furthermore, we encountered a difference in time zones when retrieving the observations and forecasts from different sources. The forecasts from MET were in Coordinated Universal Time (UTC), whereas the observations from LMT and Eklima were in Central European Time (CET), with Eklima not observing daylight saving time. The forecasts from MEPS were accumulated over the previous hour, while the observations from LMT were averaged values for the first hour after the given observation time. To ensure consistency, we converted all observa-

tions and forecasts to UTC. We also transformed the data to the same Coordinate reference system (CRS), and use UTM zone 33N on a kilometer scale.

Since measuring solar radiation at night is not informative, we only considered solar radiation at 12.00 UTC, representing the average solar radiation between 11.00-12.00 UTC. Additionally, there were some missing values in the ensembles for certain locations, which we addressed by considering only those dates where all 15 locations had no missing values. Consequently, we obtained a dataset with 775 rows of observations and forecasts for each location, which we intend to use for both statistical post-processing of the forecasts and creating spatial maps of calibrated weather forecasts.

15 locations is a low number when the goal is to create spatial maps, especially when the extent is as large as Norway. Luckily, we can use the locations that are not trained for validation purposes, resulting in 15 locations for model training and 25 locations for model validation. Figure 3.7 presents a map of these locations, with the 15 training locations in red and the 25 validation locations in blue. The training locations cover the entire domain, which is a favorable characteristic, but both sets of samples are somewhat limited in the northern region. It is crucial to acknowledge that this could potentially impede model training and introduce biases.

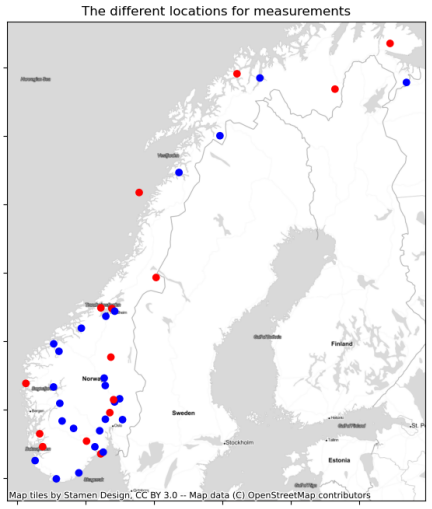


Figure 3.7: A figure showing the spatial location of the weather stations utilized in this study, where the red dots represents the locations used for model training, and the blue dots represent the locations used for model validation.

3.5 Forecast errors

The NWP forecasts are issued at noon each day and we are interested in how well they forecast the weather for the same day, i.e. 12 hours into the future. In this subsection, we investigate the errors of the NWP forecast to identify if there exists locational bias. A forecast is biased if there is a consistent difference between the observations and the forecasts, which can occur when the forecasts tend to be too high or too low.

To quantify the error, we use the following formula:

$$err = y_t(s) - x_t(s) \tag{3.1}$$

where $y_t(s)$ is the observation at location s at time t, and $x_t(s)$ is the corresponding NWP forecast.

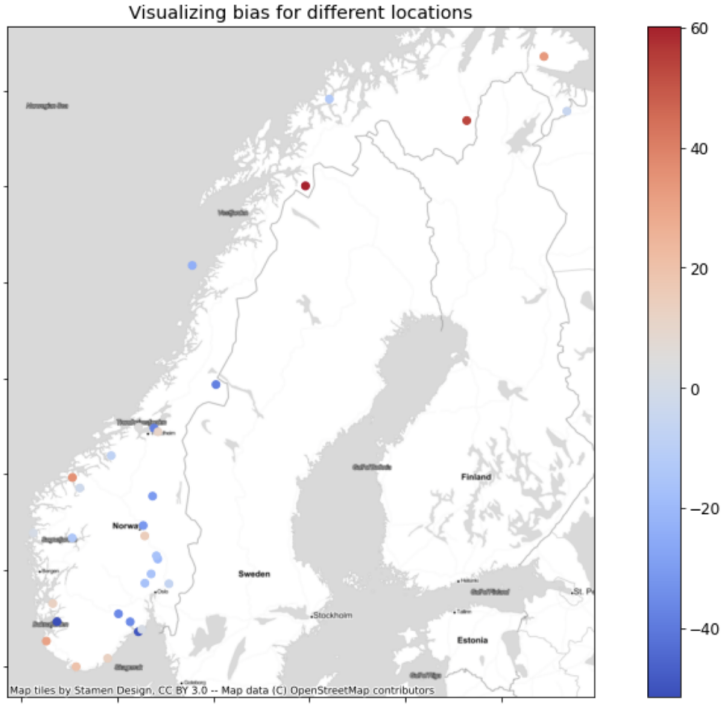


Figure 3.8: A map of Norway showing locational bias. The color represents the bias, and the dot is the location of the bias.

Figure 3.8 depicts a map of Norway illustrating bias, indicated by color marks representing positive (red) or negative (blue) bias. These colors represent the average bias for each location. Although there are areas of dark red in the north and dark blue in the south, suggesting some regional variations, the visualization does not reveal a distinct spatial bias pattern. Both positive and negative biases are observed in the northern and southern regions of Norway, making it inconclusive to draw any definitive conclusions from this visualization.

Chapter 4

A spatially varying coefficient model for solar radiation forecast

In this Chapter, we suggest a model based on the NWP forecasts and the theory described in Chapter 2. The model is a Spatially varying coefficients model (SVCM) with a scaled Beta distributed likelihood. In this section, we go into details of the hierarchical model and the prior specification, before we define the evaluation procedure. Finally, we write about the software utilized in this study. Throughout this Chapter, we let s denote the spatial location and t denote the time.

4.1 Mathematical Formulation

Assume that we have an NWP forecast of solar radiation at location s and time t denoted $x_t(s)$, and corresponding observation $y_t(s)$. We want to predict for day i based on the NWP forecast on day i, and the parameters estimated. We let $z_i(s)$ denote the probabilistic post-processed forecast at location s on day i, and it is given by the following formulation:

$$\pi(z_i(s)|\xi_i, \phi_i) \sim Beta(\mu_i(s), \phi_i) \tag{4.1}$$
$$g(\mu_i(s)) = \alpha_i(s) + \beta_i(s) \cdot x_i(s) \tag{4.2}$$
$$\boldsymbol{\alpha} \sim N(0, \Sigma_\alpha) \tag{4.3}$$
$$\boldsymbol{\beta} \sim N(0, \Sigma_\beta) \tag{4.4}$$

Here $\xi_i = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ is the latent field for day i with $\boldsymbol{\alpha} = (\alpha(s_1), ..., \alpha(s_n))^T$ and $\boldsymbol{\beta} = (\beta(s_1), ..., \beta(s_n))^T$ as the spatially varying coefficients. The response variable is assumed to follow a Beta distribution, with a logit link function $g(\cdot)$. We model the spatially varying coefficients as Gaussian Random Fields (GRFs) with a Matérn covariance matrix defined in Equation 2.23.

4.2 Hyperparameters

There are in total five model parameters in the model formulation from Section 4.1, and all of these need prior distributions as we use a Bayesian framework for inference. Since we use the sliding window methodology for inference, we fix the precision parameter ϕ throughout the entire domain. In R-INLA, The precision parameter ϕ is represented as

$$\phi = h_i exp(\gamma), \tag{4.5}$$

where $h = (h_i) > 0$ is a fixed scaling, and the prior is defined on γ.

To define the GRFs with the Matérn covariance matrix, we need to specify a range parameter ρ and a marginal variance parameter σ. The spatially varying intercept has range parameter ρ_α and marginal variance parameter σ_α, whereas the spatially varying slope has range parameter ρ_β and marginal variance parameter σ_β.

To determine the prior distributions for the spatially varying parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we must select priors for both the range and the marginal variance parameters. We use the same prior distributions for both the intercept and the slope parameters. Specifically, we employ joint informative PC priors, as described in Section 2.10.1.

To define the prior for the spatial range, we consider the distance between the measurement stations. We set the range prior such that the probability of a

range exceeding 500 km is 0.5. This choice is informed by the properties of the geographical region under study.

Selecting a prior for the marginal variance parameter σ^2 is more challenging. However, based on previous work by Egeli (2022), we deem a variance of approximately 10 to be reasonable for the parameters. Therefore, we set the prior for the marginal variance such that the probability of exceeding 10 is 0.5.

In formulas, it looks like this:

$$\begin{aligned} \text{Prob}(\rho_\alpha < 500km) = 0.5, \quad \text{Prob}(\sigma_\alpha > 10) = 0.5 \\ \text{Prob}(\rho_\beta < 500km) = 0.5, \quad \text{Prob}(\sigma_\beta > 10) = 0.5 \end{aligned}$$

4.3 Sliding Window

The sliding window methodology is utilized to estimate the parameters of the latent field. Specifically, the N days preceding each prediction day i are utilized as the training period, covering the time interval from day $t = i - N$ to day $t = i - 1$. This approach is particularly suitable considering the significant temporal variation observed in solar radiation, with higher values during summer months and lower values during winter months. By implementing the sliding window approach, the model can effectively account for seasonal variations without requiring the incorporation of separate temporal parameters.

4.4 Inference with the SPDE approach

The full model can be specified in a hierarchical framework with three levels as defined in Section 2.6. This Section has taken inspiration from Haug et al. (2020) and Roksvåg et al. (2022), as there are certain similarities with the model setup and inference method.

In the first level, we have the observational likelihood given by the observations y, which are assumed to be conditionally independent given the latent field ξ and some hyperparameters θ_1:

$$\pi(\boldsymbol{Y}|\boldsymbol{\xi},\theta_1) = \prod_{t=i-N}^{i-1}\prod_{s=1}^{S} \text{Beta}(\mu_t(s),\phi_t) \tag{4.6}$$

$$g(\mu_t(s)) = \alpha_t(s) + \beta_t(s)\cdot x_t(s) \tag{4.7}$$

Here $\boldsymbol{Y}$ is a vector containing all observations $\{y_t(s)\}$ from all locations s and training days t from $i-N$ to $i-1$. $\boldsymbol{\xi}$ is a vector containing all latent variables, in particular, the spatially varying parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, such that $\boldsymbol{\xi} = (\{\alpha_t(s)\},\{\beta_t(s)\})$. Finally, $g(\cdot)$ is the logit link function.

The second level of the hierarchical model is the latent field ξ. Since we model it as a latent Gaussian model, ξ is a Gaussian Markov random field given by:

$$\pi(\xi|\theta_2) \sim N(0, Q^{-1}(\theta_2)), \tag{4.8}$$

where θ_2 are some hyperparameters and the precision matrix $Q^{-1}(\theta_2)$ is sparse.

Finally, at the last level of the hierarchical model, the hyperparameters $\boldsymbol{\theta} = (\theta_1,\theta_2)$ have some prior distributions. In the Bayesian setting, independent priors are chosen for the parameters, so $\pi(\boldsymbol{\theta}) = \pi(\rho_\alpha,\sigma_\alpha)\pi(\rho_\beta,\sigma_\beta)\pi(\gamma)$.

In this study, we utilize R-INLA to perform inference. Given that our latent field $\boldsymbol{\xi}$ is Gaussian, the requirements for employing R-INLA are met for our spatially varying coefficient model. To ensure rapid and accurate approximation, we adopt the SPDE approach, which is discussed in detail in Section 2.9. This approach is based on the observation that a GRF with a Matérn covariance matrix can be expressed as the solution to a stochastic partial differential equation.

To perform inference, it is necessary to triangulate the domain. This work utilizes standard settings in R-INLA for the mesh generation, and Figure 4.1 presents the mesh, with the training locations represented as black dots.

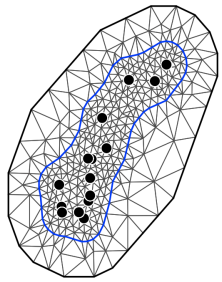


Figure 4.1: The mesh used in this study. The main region is inside the blue line, and the outer region is between the blue line and the outer black line. The black triangles are the mesh and the black dots are the locations used for training purposes.

4.5 Model Interpretation

The interpretation of the model is challenging due to the presence of many parameters and a nonlinear transformation due to the logit link function from Equation 4.1. In this section, we aim to explain how the parameters relate to the prediction by providing a detailed analysis of a particular day.

Assuming that we have employed the sliding window methodology, and have successfully estimated all the parameters required to predict solar radiation for a given day. These parameters include the spatially varying parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, as well as the precision parameter ϕ. The prediction is assumed to follow a Beta distribution, scaled by the physical limits of solar radiation, as explained in Section 2.1. The mean parameter $\boldsymbol{\mu}$ is spatially varying and is determined by the transformation $\mu_t(s) = \text{logit}^{-1}(\alpha_t(s) + \beta_t(s) \cdot x_t(s))$, where logit^{-1} is the logistic function.

Having estimated all relevant parameters and established their relationship to the prediction, we can now examine the model using a specific day and location. Specifically, we will analyze the prediction for Alvdal on June 30th, 2021. For this day and location, the posterior mean of the intercept and slope parameters α and β are estimated to be -2.02 and 3.73, respectively. The NWP forecast for Alvdal on this day is 0.274 on the unit scale, and the mean parameter μ in the Beta distribution is obtained through the transformation $\text{logit}^{-1}(\alpha + \beta \cdot x)$, yielding a value of 0.270 for μ.

To further elucidate the relationship between μ and the NWP forecast range, we

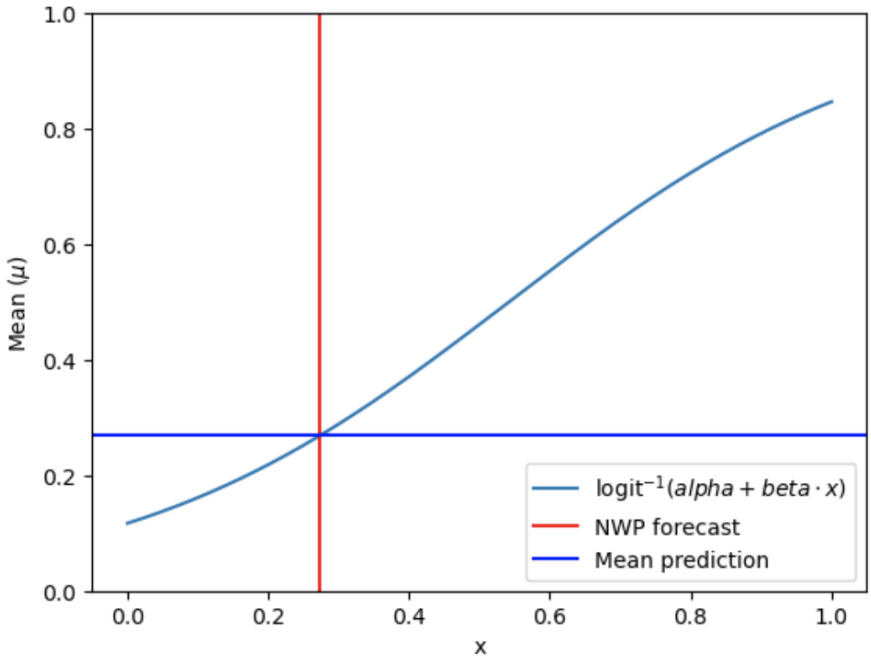


Figure 4.2: The x-axis represents the possible domain of the NWP forecast, and the y-axis represents the mean parameter μ after the logit transformation. The light blue line represents the formula $y = \text{logit}^{-1}(\alpha + \beta \cdot x)$, the red vertical line represents the NWP forecast for Alvdal 30th of June 2021, and the blue horizontal line represents the corresponding mean parameter μ after the logit transformation.

present Figure 4.2, which illustrates how μ varies across the possible range of NWP forecasts, given the aforementioned values of α and β.

We can see from Figure 4.2, that the mean parameter μ is the point of intersection between the logistic function with these parameter values and the NWP forecast. As we can see, this transforms the NWP forecast to the mean value μ in an almost linear way, but the possible domain is decreased from $[0, 1]$ to $[0.12, 0.85]$.

Furthermore, we have also estimated a precision parameter ϕ. This parameter is constant throughout space, and on this specific day, the posterior mean of ϕ is 21.266. For an easier interpretation of the results and illustrational purposes, we plug in the posterior mean estimates, which yields the distribution on the unit

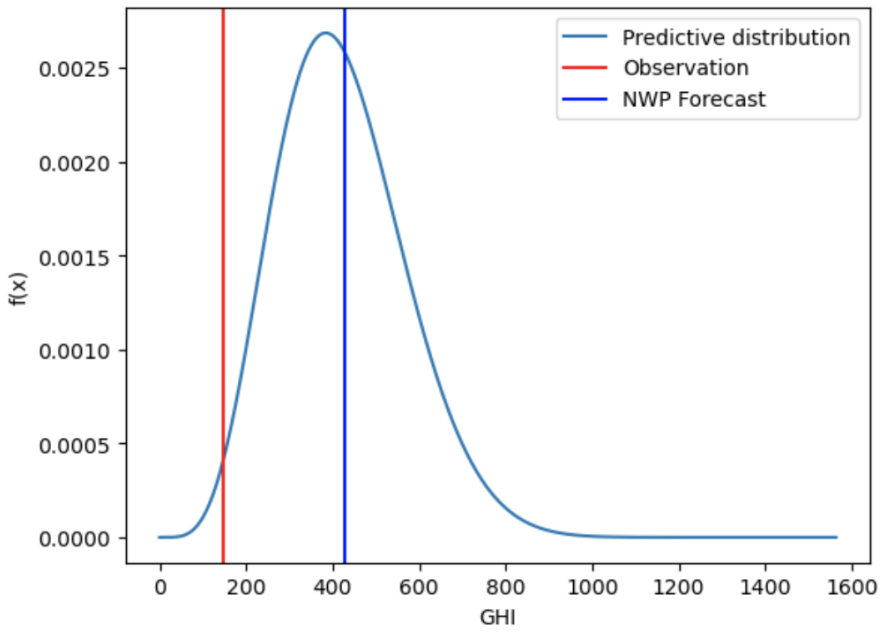


Figure 4.3: The x-axis represents Global Horizontal Irradiance (GHI) measured in W/m^2, and the y-axis represents the pdf of the distribution. The light blue line is the predictive distribution, the red vertical line is the observation at Alvdal 30th of June 2021, and the blue vertical line is the corresponding NWP forecast.

scale to be $\pi(z) \sim \text{Beta}(0.270, 21.266)$. We can transform this to the original domain of solar radiation by multiplying the resulting pdf with the physical limitations of solar radiation in Alvdal on the 30th of June 2021, which reads $1563 W/m^2$. The resulting predictive distribution on the original domain can be seen in Figure 4.3, and it is this distribution we use to evaluate the post-processed forecast.

4.6 Evaluation of Results

The evaluation methodology utilizes two distinct methods to objectively assess the performance of the models. The first method involves training the model on 15 reliable locations and validating it on the remaining 25 locations as described in

Section 3.4. This approach allows for the evaluation of the model's predictions on new and unseen locations. The use of a separate validation set helps in detecting potential overfitting and enables the adjustment of the model accordingly.

The second method involves Leave-One-Out Cross Validation (LOOCV) on the 15 reliable locations, which provides an unbiased evaluation of the model's performance. LOOCV is particularly useful when the dataset is small, as you can train the model on many locations and still get an unbiased evaluation of the model.

In both evaluation methods, the Root Mean Squared Error (RMSE) and Continuous Ranked Probability Score (CRPS) serve as objective performance metrics. RMSE is a commonly used metric in regression analysis and provides a measure of the average deviation between the predicted and actual values. RMSE is calculated using Equation 2.8. CRPS, on the other hand, is based on a predictive distribution and provides a more comprehensive measure of the model's predictive performance. CRPS for the predictive distribution is calculated using Equation 2.4.

The evaluation of model performance against a baseline model is a crucial step in assessing the effectiveness of the proposed model. In this study, the baseline model is the original NWP forecast. To evaluate the performance of the baseline model, we rely on the same metrics, namely RMSE and CRPS. We calculate RMSE by Equation 2.8. For CRPS we refer to Section 2.2.2 which highlights that when the CRPS metric is used on a point forecast, it reduces to Mean Absolute Error (MAE). Therefore, we use MAE as the baseline performance metric for the NWP forecast. To establish the effectiveness of the proposed model relative to this baseline model, we employ Equation 2.7 to calculate a percentage increase or decrease in predictive skill.

Calibration is evaluated using PIT-histograms as explained in Section 2.2.1. Evaluating calibration for the model is straightforward as the prediction is a distribution. However, evaluating the calibration for the NWP forecast is more challenging since it does not make sense to evaluate calibration on a point forecast. Consequently, we use the original raw ensembles before averaging them, and the calibration assessment is conducted by using the empirical cumulative distribution function (CDF), as described in Section 4.1 of Egeli (2022).

Finally, we compare the performance of the spatial model on the 15 training locations with a local Bayesian Model Averaging (BMA) method. This is briefly introduced in Section 2.4 and thoroughly evaluated in Egeli (2022). This comparison allows us to assess the degree to which the spatial model performs compared to a local model, providing a practical measure of its performance. The evalu-

ation is conducted using the same two metrics: RMSE and CRPS. By considering both RMSE and CRPS, we gain a comprehensive understanding of the spatial model's performance compared to the local BMA method.

4.7 Software

This study utilizes several packages in both R and Python and the usage of which depends on the application. All pre-processing of the data is done in Python using the Pandas package (pandas development team, 2023). Inference and simulations for the spatially varying coefficient model are done using the R-package inlabru (Bachl et al., 2019). Inlabru facilitates spatial modeling using integrated nested Laplace approximation via the R-INLA package (Rue et al., 2009), which can be downloaded from www.r-inla.org. The visualizations are made by different libraries in Python and R, where the main package in Python is Matplotlib (Hunter, 2007). Most of the spatial maps, including the maps of the spatial parameters, the original NWP forecast, and the post-processed predictions, are made with the two R-packages Terra and ggplot2 (Wickham, 2016).

Chapter 5

Results

This section aims to evaluate and compare the performance of the model introduced in Section 4.1. Firstly, we showcase a few examples to illustrate how the coefficients change spatially under various conditions of season and weather. Next, we conduct a prior sensitivity analysis to evaluate the significance of prior specifications. Furthermore, we evaluate the model's performance using quantitative metrics such as RMSE and CRPS on the training and validation sets. We then assess calibration using the PIT-histogram, which is based on the results of Cross-validation. Finally, we investigate the performance of this spatial model versus previous local models using the same metrics.

5.1 Case study: differences in space and time

In this section, we show the predictions and the model parameters during various seasons of the year. The amount of solar radiation varies greatly over time, with higher values during summer months and lower values during winter months. This effect is particularly pronounced in the North, where the phenomenon of the midnight sun occurs in the summer and polar night occurs in the winter. Therefore, it is crucial to investigate the model's performance across different seasons and to examine how the model's parameters vary over time. Additionally, it is essential to assess the level of confidence we can place in the model's predictions.

5.1.1 Summer day

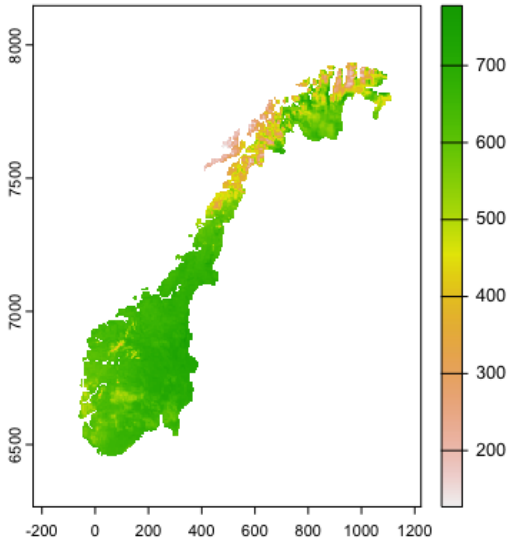
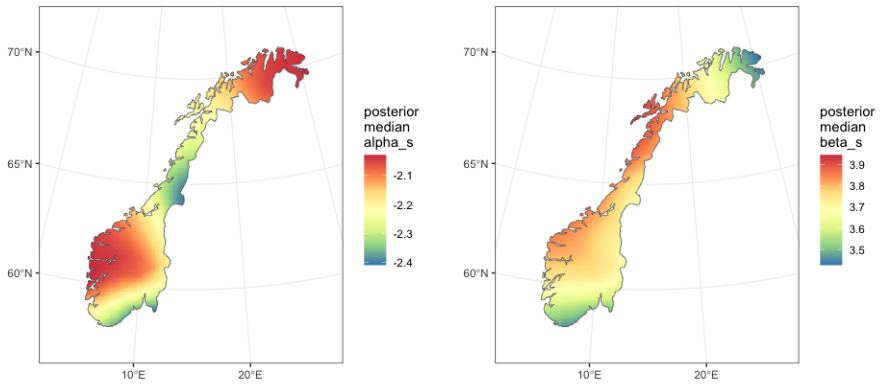


Figure 5.1: The NWP forecast of solar radiation for 30th of June 2021 at 12.00, issued at noon the same day. The color represents solar radiation measured in W/m^2.

The date chosen for this section is the 30th of June 2021. On this particular day, the weather was warm, and there were very few clouds in the entire domain, making it an ideal day to examine how the model's parameters vary spatially. A figure of the forecasted solar radiation for 12.00 UTC issued at noon the same day can be seen in Figure 5.1. The model is trained from the 9th of June 2021 to the 29th June of 2021, and the prediction is based on Equation 4.1. Figure 5.2 shows the median spatial variation of the parameters α and β. The β-parameter has higher values along the west coast and in the middle of Norway, while it is lower in the north and south. This suggests that the forecast is more precise in these areas since the forecasted value is given a larger weight, as explained in Section 2.3. However, the parameters are relatively smooth across the domain, mainly because we intentionally defined the priors to make the parameters smooth. While the parameters do not vary too much, this is relative, as there may still be substantial

differences when they are transformed through the logit link function.

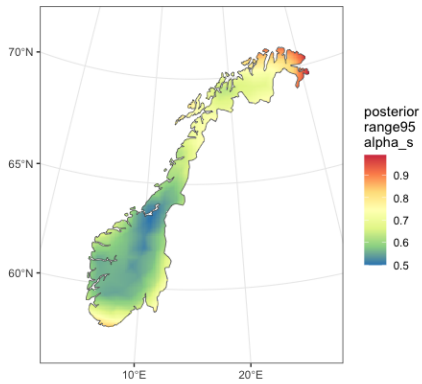


(a) The intercept parameter α (b) The slope parameter β

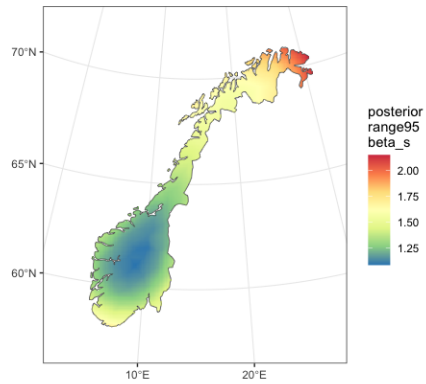
Figure 5.2: The Figure illustrates the posterior median of the intercept parameter α and the slope parameter β on the sunny day, 30th June 2021.

Each parameter has an associated uncertainty, as shown by Equation 2.22. The 95% confidence interval provides a measure of this uncertainty, representing the range of expected values. Figure 5.3 illustrates the 95% confidence interval of the spatially varying parameters. Interestingly, the parameters have higher uncertainty in the northern and southern regions compared to the central part of Norway, with the northern region exhibiting the highest uncertainty. This observation can be attributed to the limited number of training locations in the north, which contributes to increased parameter uncertainty.

Finally, the probabilistic forecast for solar radiation follows a Beta distribution, and the predictions are found by combining the parameters with the NWP forecast. This distribution is then transformed back to the original domain, and the results for the 30th of June 2021 are shown in Figure 5.4. It is challenging to discern any direct relationship between the parameter estimates and the predictions, as the latter is heavily influenced by the NWP forecast. The standard deviation of the predictions is also of interest, and it is evident that the spatial patterns have higher confidence in the predictions close to the locations it is trained on.

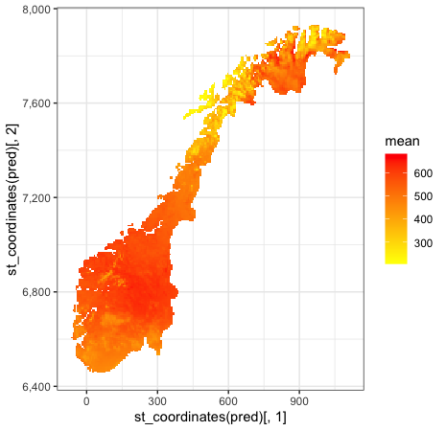


(a) The intercept parameter α's range

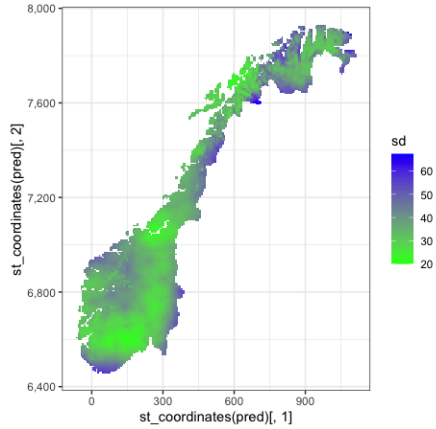


(b) The intercept parameter β's range

Figure 5.3: The Figure illustrates the range of the intercept parameter α and the slope parameter β's 95% confidence interval on the day 30th June 2021.



(a) The predictions for 30th June 2021



(b) The standard error for 30th June 2021

Figure 5.4: The Figure illustrates the predictions and the standard error for 30th June 2021.

5.1.2 Cloudy day

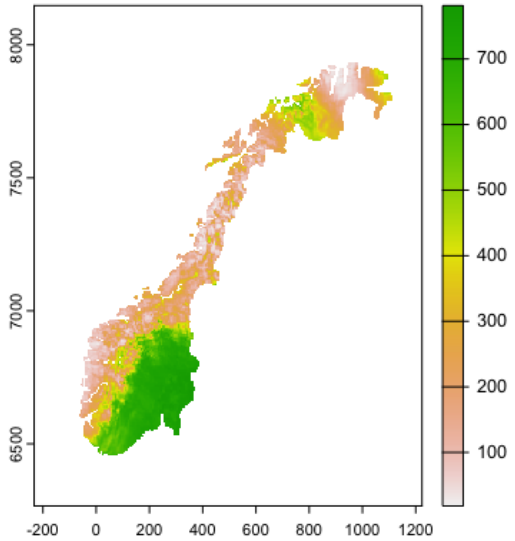
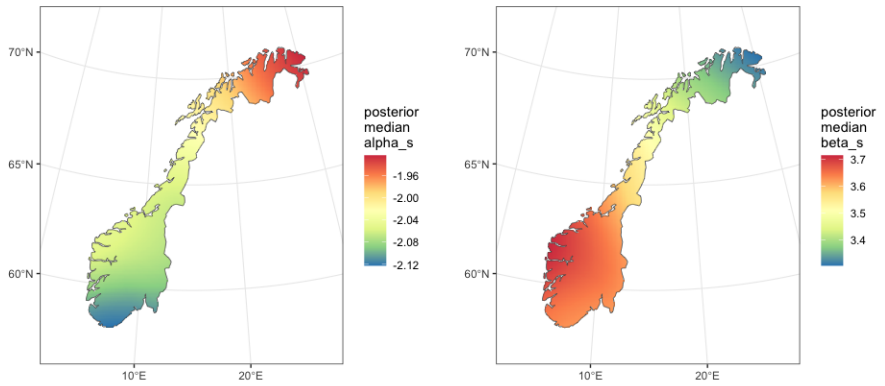


Figure 5.5: The NWP forecast of solar radiation for 13th of June 2021 at 12.00, issued at noon the same day. The color represents solar radiation measured in W/m^2.

On the 13th of June 2021, the weather conditions were observed to be cloudy in the western and middle regions of Norway, while the southern and eastern regions had minimal cloud cover. The solar radiation forecasted for 12.00 UTC this day, issued at noon can be seen in Figure 5.5. The model is trained from the 24th of May 2021 to the 12th of June 2021, and the prediction is based on Equation 4.1. The posterior median spatially varying parameters, α, and β, are shown in Figure 5.6. These parameters vary linearly based on latitude, with the slope parameter β having a higher value in the south, indicating its greater significance for the prediction. Conversely, the intercept α has a larger value in the north, indicating its greater importance in the model's output in this area.



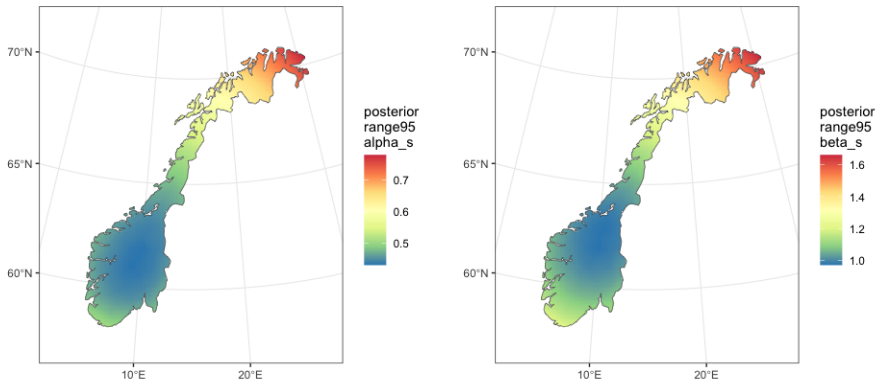
(a) The intercept parameter α (b) The slope parameter β

Figure 5.6: The Figure illustrates the posterior median of the intercept parameter α and the slope parameter β on the cloudy day, 13th of June 2021.

Figure 5.7 displays the 95% confidence interval of the spatially varying parameters α and β. The range of these parameters exhibits a similar spatial distribution as those observed for the sunny day discussed in the previous subsection.

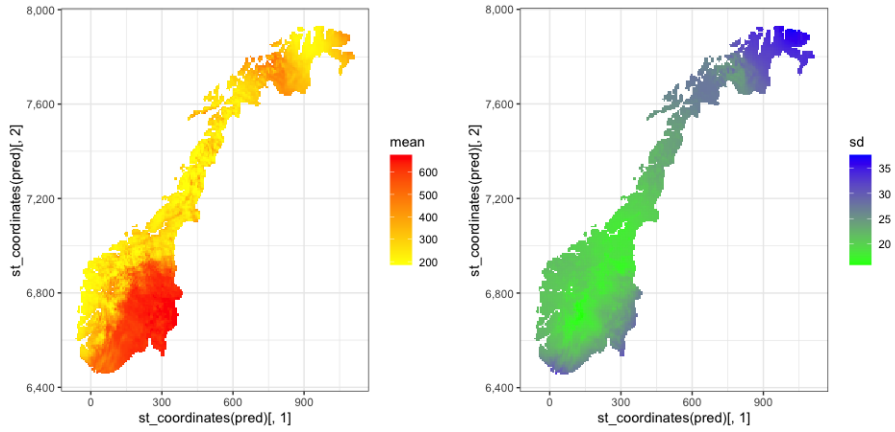
Finally, the parameters together with the weather forecast define a probabilistic forecast based on the Beta distribution. This distribution is scaled back to the original domain, and the result for the 13th of June 2021 can be seen in figure 5.8. Again, it is difficult to see any direct link between the parameter values and the predictions.

The observant reader might have observed the similarity in the parameters estimated for both the cloudy and the sunny day. It should be noted that the parameters were not trained on individual days but rather on a sliding window over 20 days, and therefore the forecast values of the current day do not alter the parameter estimates.



(a) The intercept parameter α's range (b) The intercept parameter β's range

Figure 5.7: The Figure illustrates the range of the intercept parameter α and the slope parameter β's 95% confidence interval on the cloudy day, 13th of June 2021.



(a) The predictions for 13th June 2021 (b) The standard error for 13th June 2021

Figure 5.8: The Figure illustrates the predictions and the standard error for the cloudy day, 13th of June 2021.

5.1.3 Winter day

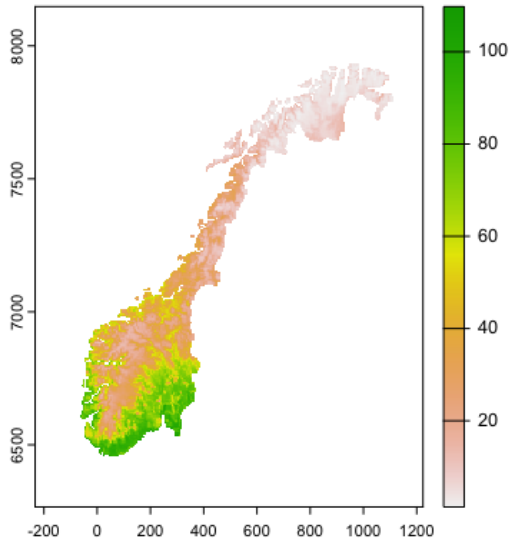
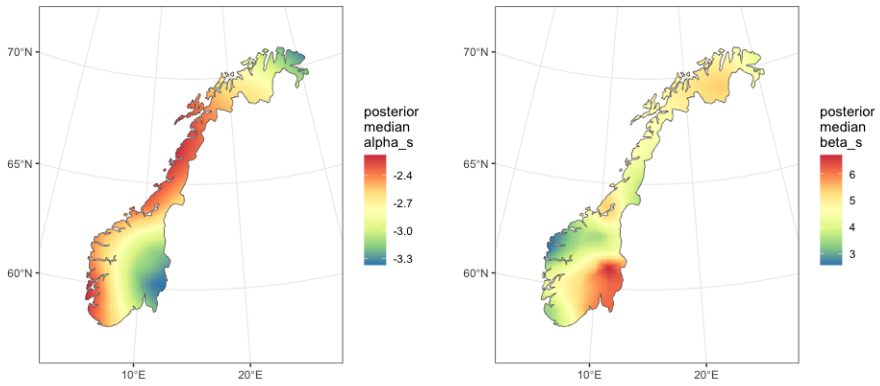


Figure 5.9: The NWP forecast of solar radiation for the 5th of February 2021 at 12.00, issued at noon the same day. The color represents solar radiation measured in W/m^2.

For this Section, we have chosen the date 5th of February 2021. On this day, the northern part of Norway experienced minimal solar radiation, making it a relatively dark day. The mountainous regions of Norway also received a small amount of solar radiation, and it is unclear whether this was due to clouds or the steep solar angle. The forecasted solar radiation at 12.00, issued at noon the same day can be seen in Figure 5.9.

The median values of the spatially varying parameters α and β are displayed in Figure 5.10. In contrast to the sunny and the cloudy day, the parameters exhibit significant spatial variations. A high slope parameter is assigned to the eastern part of Norway, indicating that the forecasted value has a considerable weight in the prediction for this region. Conversely, a lower slope parameter is assigned to the west coast. Notably, the parameters have a low value in the northern part of

Norway, possibly indicating that solar radiation prediction in this region during the winter season is relatively straightforward, with a high skill in predicting close to zero radiation values.

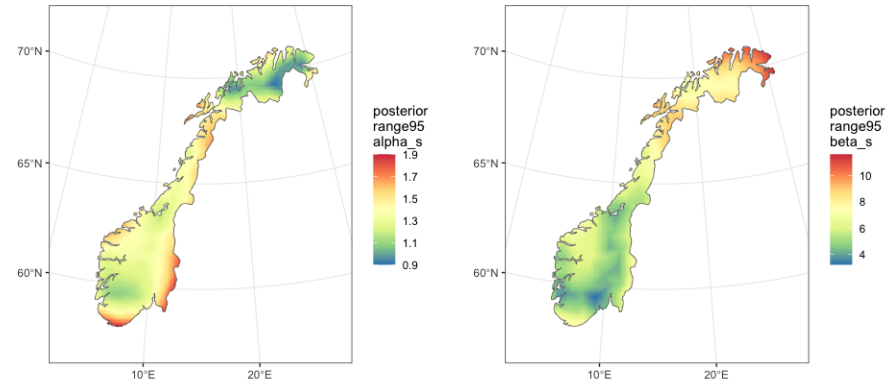


(a) The intercept parameter α (b) The slope parameter β

Figure 5.10: The Figure illustrates the posterior median of the intercept parameter α and the slope parameter β on the winter day, 5th of February 2021.

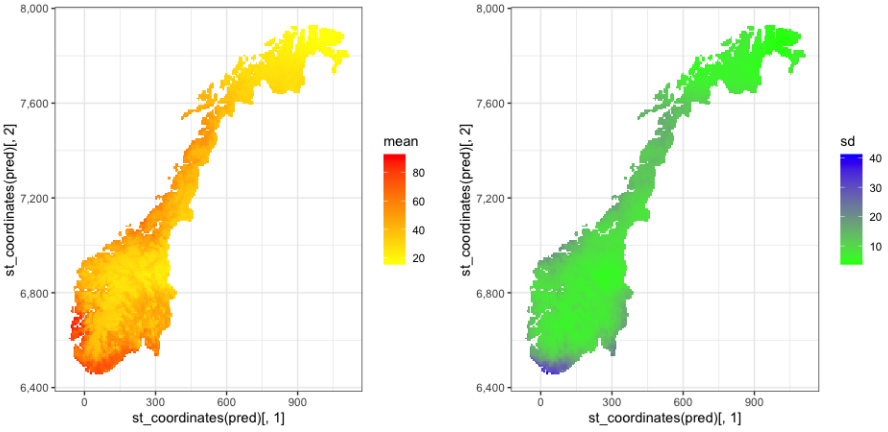
The parameters also have uncertainty associated with them. The 95% confidence interval of the parameters can be seen in Figure 5.11. Compared to the two previous scenarios, the confidence interval for the slope parameter β is wider in this case. This observation can be explained by the fact that the forecasted values in the northern areas are close to zero, resulting in β times the forecast being close to 0 as well.

Finally, the parameters together with the weather forecast define a probabilistic forecast based on the Beta distribution. This distribution is scaled back to the original domain, and the result for the 5th of February 2021 can be seen in figure 5.12. Interestingly, the predictions display similar spatial variability as the original NWP forecast depicted in Figure 5.9, suggesting that the spatially varying coefficients do not play a prominent role in this particular scenario. Moreover, the standard deviation also shares a resemblance with the predictions, exhibiting lower values in the north and higher values in the south.



(a) The intercept parameter α's range (b) The intercept parameter β's range

Figure 5.11: The Figure illustrates the range of the intercept parameter α and the slope parameter β's 95% confidence interval on the winter day, 5th of February 2021.



(a) The predictions for 5th of February 2021 (b) The standard error for 5th of February 2021

Figure 5.12: The Figure illustrates the predictions and the standard error for the winter day, 5th of February 2021.

5.2 Prior sensitivity

A prior sensitivity analysis is conducted to investigate the sensitivity of the model to variations in the prior distribution. The objective of this analysis is to determine the robustness of the model's conclusions to reasonable changes in the prior distribution, by systematically varying the prior distribution and examining how much the posterior distribution changes in response.

To carry out the prior sensitivity analysis, two days are selected as sufficient for the analysis. Initially, the range parameter is varied while keeping the variance fixed, followed by a variation of the variance parameter with the range parameter held constant. The evaluation metrics employed for the validation dataset are the Continuous Ranked Probability Score (CRPS) and Root Mean Square Error (RMSE), in addition to visual inspection of the model output and parameters.

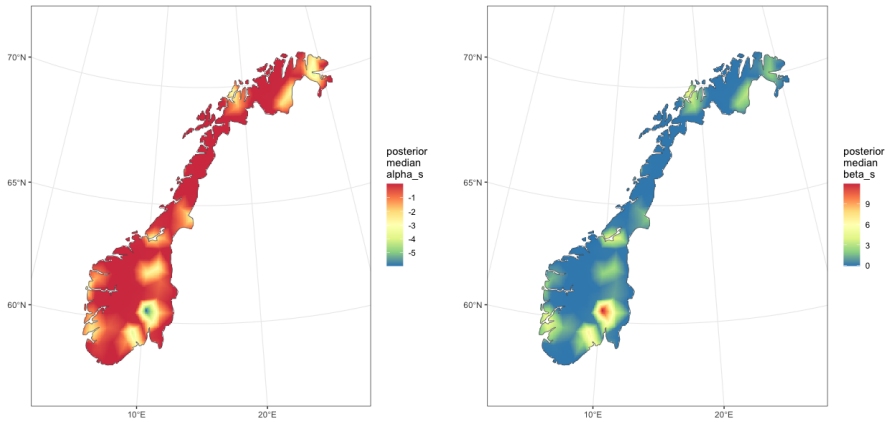
The range parameter is a user-specified parameter that is defined as $\text{Prob}(\rho < \rho_0) = \alpha$, where ρ_0 and α are the parameters of interest. It denotes the probability that the spatial range exceeds ρ_0 and is usually expressed in kilometers. In this study, we hold α constant at 0.5 while varying ρ_0 between 10 and 10000 km.

A small value of ρ_0 restricts the spatial correlation and limits the availability of information to only areas very close to the locations the model is trained on, resulting in poor predictive performance. This issue is depicted in Figure 5.13, which illustrates the spatially varying parameters with a small range parameter. As noted in Section 2.3, when both α and β are 0, the prediction is 0.5, which is observed to be the case for most parts of Norway in this plot. The poor predictive performance underscores the need for a larger range parameter.

Conversely, setting the range parameter to a large value imposes a highly restrictive constraint by enforcing uniformity throughout the entire domain. This is evident in Figure 5.14, which displays the spatially varying coefficients for the range parameter ρ set to 10000. The parameter values remain nearly constant throughout the domain, as indicated by the color legend. Moreover, the large intercept value α relative to the slope β implies that the post-processed forecast assigns limited weight to the NWP forecast.

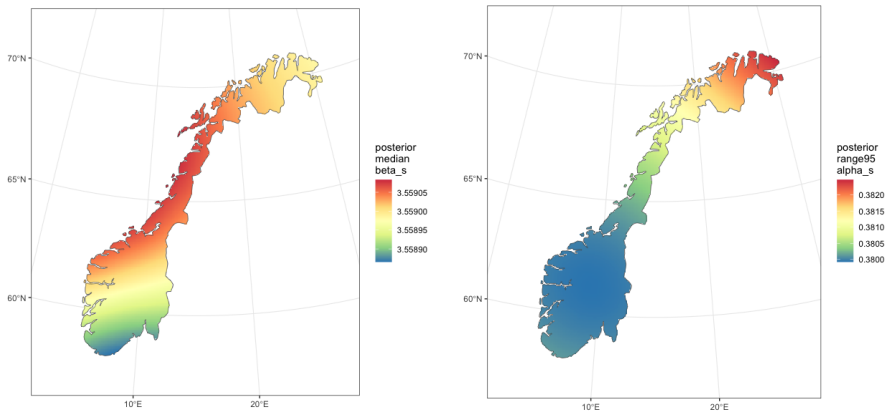
The inspection of the coefficients when setting the range parameter too small and large values provides conclusive evidence of the significance of prior specification. This is reinforced by the quantitative measures of the out-of-sample locations' Root Mean Square Error (RMSE) and Continuous Ranked Probability Score (CRPS), as shown in Table 5.1.

Although the marginal variance parameter σ is an essential parameter, its effects



(a) The intercept parameter α (b) The slope parameter β

Figure 5.13: The Figure illustrates the geographical coverage of the two parameters when the range is specified at a small value, in this case, $\text{Prob}(\rho < 10km) = 0.5$.



(a) The intercept parameter α (b) The slope parameter β

Figure 5.14: The Figure illustrates the two spatially varying coefficients when the range is specified at a large value, in this case, $\text{Prob}(\rho < 10000km) = 0.5$.

ρ	σ	CRPS	RMSE
10	10	55.68	95
100	10	51.09	90.41
500	10	51.11	89.93
1000	10	51.10	89.94
10000	10	49.61	88.4
500	0.1	50.16	90.03
500	1	50.96	90.03
500	10	50.24	88.85
500	100	52.31	91.67
500	1000	52.20	91.50

Table 5.1: A Table showing CRPS and RMSE values for different combinations of the range parameter ρ and the marginal variance parameter σ.

are not as evident in the visual inspection as those of the range parameter ρ. Therefore, we turn to the quantitive metrics shown in Table 5.1.

The performance of a small range parameter value is inadequate, thus emphasizing the necessity for a larger range parameter. However, Table 5.1 presents unexpected results, indicating that the largest range parameter value leads to the smallest CRPS and RMSE estimates for the present test case. Nevertheless, a more conservative range parameter appears to perform better overall. In contrast, the marginal variance parameter does not appear to be crucial, provided it is not excessively large. In the test scenario, a marginal variance parameter value of 10 exhibited the best performance, and this value was also selected for the main results in the subsequent section.

5.3 Evaluation

This section employs the evaluation procedure described in Section 4.6. Initially, the results from Leave-One-Out Cross Validation (LOOCV) are examined to assess the model's performance. Subsequently, a detailed analysis is conducted, comparing the outcomes obtained from the training and validation locations. The evaluation is based on quantitative metrics, namely Root Mean Squared Error (RMSE) and Continuous Ranked Probability Score (CRPS). Additionally, the calibration of the model is assessed through the examination of Probability Integral Transform (PIT) histograms. Lastly, the performance of the spatial model is compared to a local Bayesian Model Averaging (BMA) method, utilizing

RMSE and CRPS as comparative measures.

5.3.1 Leave-one-out Cross-validation

We evaluated the performance of the spatial model using leave-one-out cross-validation (LOOCV) on the 15 locations with good-quality data. The evaluation involved comparing the model's predictions to the actual observations on each location using two performance metrics: the continuous ranked probability score (CRPS) and the root mean squared error (RMSE).

Based on the evaluation results, the average CRPS was 58.73, which is 30% better than the mean absolute error (MAE) of the raw NWP forecast, which was 83.4. This indicates that the model can post-process the NWP forecasts to create more accurate probabilistic forecasts.

In terms of RMSE, the model outperforms the NWP forecasts. The RMSE decreased from 108.5 on the original NWP forecast to 103.65 for the model, indicating that the model's predicted values are on average closer to the actual observations. It is important to note that the RMSE is a measure of the average deviation between predicted and actual values, and does not take into account the uncertainty of the predictions.

Overall, the evaluation results suggest that the model can make more accurate predictions than the original NWP forecasts, both in terms of probabilistic forecasts and average deviation between predicted and actual values.

5.3.2 Training and validation locations

It is important to investigate the performance of the model in both the training locations and the validation locations. Evaluating the model on the training locations indicates how well the model is fitting the locations it was trained on. However, this can lead to overfitting, where the model performs well in the training locations but poorly in new, unseen locations. Evaluating the model on the validation locations provides an independent assessment of how well the model can generalize to new areas. This is important for the spatial model, where the goal is to make accurate predictions on new locations that were not used during model training.

Based on Figure 5.15, it is evident that the model has a superior predictive performance compared to the NWP forecasts on the training locations. The

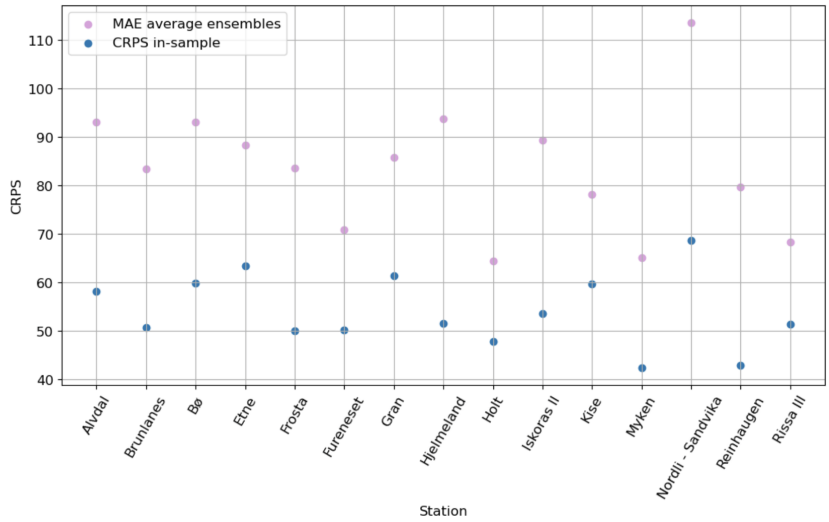


Figure 5.15: The Figure displays the average CRPS for each training location.

CRPS values for all locations are lower, and in Hjelmeland, it is almost twice as good as the original NWP forecast. The increase in predictive accuracy can be quantified by computing the CRPSS values using Equation 2.7, which results in an average increase of 35.6%. This indicates a substantial improvement in the model's performance.

Root mean squared error (RMSE) is a commonly used metric to evaluate the accuracy of point forecasts. Although it is not the primary metric for probabilistic forecasts, the model needs to perform well in terms of RMSE as it indicates predictions close to the observed values. The RMSE values for the training locations are presented in Figure 5.16. The figure shows that in most locations, the model performs better than the NWP forecast in terms of deterministic accuracy. However, the improvement in RMSE is not as evident as for the CRPS metric. On average, the model improves the RMSE by 11.2%.

However, it is important to note that the model is trained on the training locations, and a good spatial model needs to perform well on locations that it has not been trained on. The predictive performance of the model on the validation locations is shown in Figure 5.17, and the performance increase in CRPS is no longer as clear as it was for the training locations. In Landvik, the CRPS for the model is worse than the MAE for the NWP forecast. When comparing the model to the MAE of the NWP forecasts, the CRPSS values result in a predictive

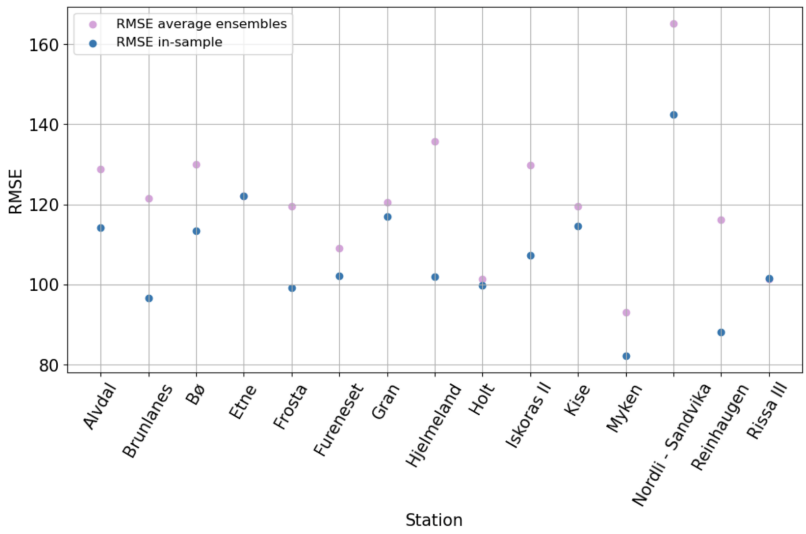


Figure 5.16: The Figure displays the average RMSE for each training location.

performance increase of 20.9%, which is still a considerable improvement.

Figure 5.18 displays the RMSE values for the validation locations. Our findings indicate that the model's performance varies significantly across locations, with certain locations exhibiting improvements in the RMSE and others displaying a deterioration of the RMSE. Specifically, the locations of Landvik, Lyngdal, and Pasvik-Svanvik manifest an RMSE that is worse by 30%, 26%, and 34%, respectively. Consequently, the overall average RMSE experiences a decrease of 3.4%. These outcomes suggest that while the model enhances predictive accuracy for the validation locations, there is no discernible improvement in the model's deterministic accuracy.

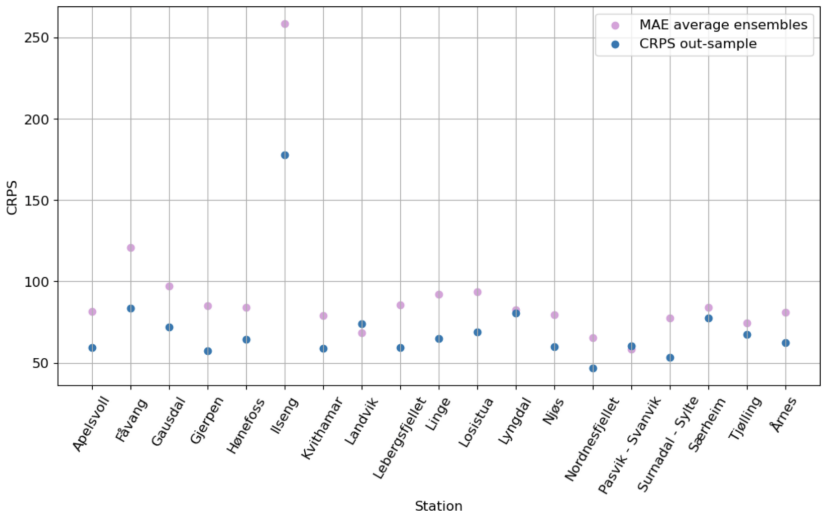


Figure 5.17: The Figure displays the average CRPS for each validation location.

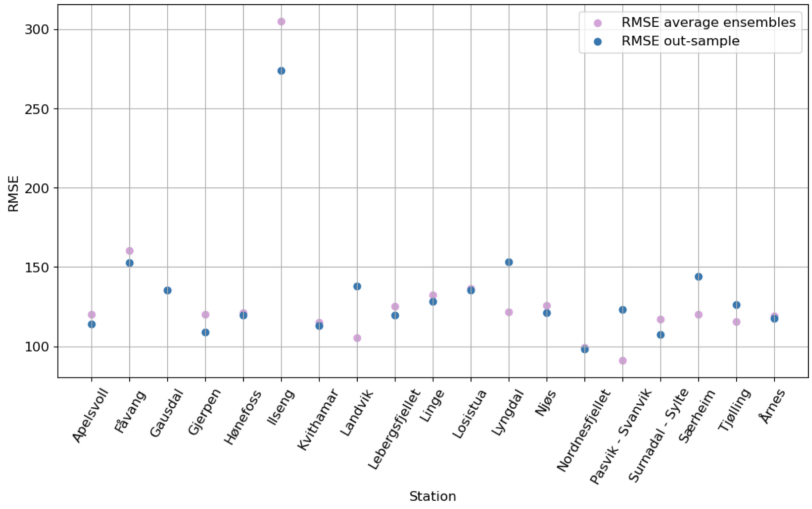


Figure 5.18: The Figure displays the average RMSE for each validation location.

5.3.3 Evaluating calibration with PIT-histograms

To evaluate calibration, the probability integral transform (PIT) histogram is utilized. The PIT values used in the histogram are obtained from the validation locations and can be seen in Figure 5.19. The histogram closely resembles a uniform distribution, indicating that the model has good spatial calibration. As mentioned in Section 1, ensemble forecasts are often uncalibrated. To further illustrate this point, the PIT-histograms based on the raw ensembles from the 15 locations used in model training can be seen in Figure 5.20, and the histogram indicates that the raw ensembles are uncalibrated.

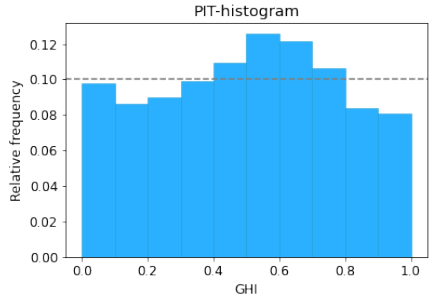


Figure 5.19: PIT-histogram for the spatial model based on the 15 training locations.

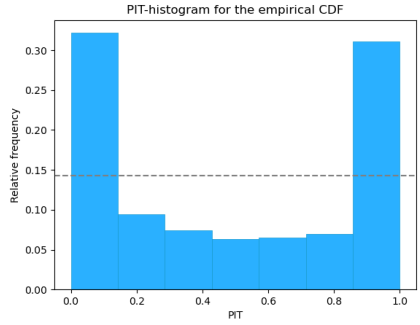


Figure 5.20: PIT-histogram for the raw ensemble.

5.3.4 Spatial versus local models

Egeli (2022) used Bayesian Model Averaging (BMA) to post-process the ensemble forecast locally. I.e. they trained and evaluated the model on a specific location, without making the model generalizable to other locations. Therefore, we can analyze and compare the model performance on the local BMA model from the previous study with the Generalized spatially varying coefficient model (GSVCM) used in this study, by investigating the training location's CRPS and RMSE scores. Both of the models are using a Beta distribution scaled by the physical constraints of solar radiation and a training period of 20 days.

It is clear from both figures of CRPS (Figure 5.21) and RMSE (Figure 5.22) that the spatially varying coefficient model is superior in terms of both predictive accuracy and deterministic accuracy.

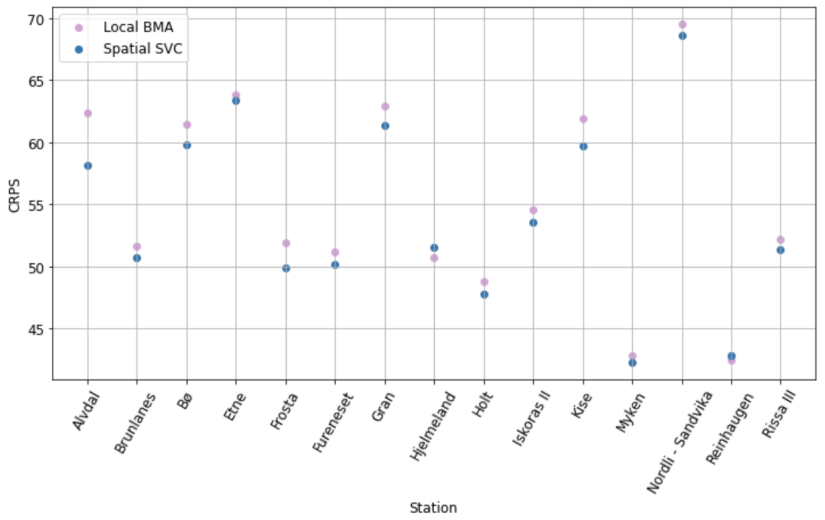


Figure 5.21: The Figure displays the average CRPS for each location the model is trained on. The blue dots are from the spatially varying coefficient model, while the pink dots are from the local BMA model.

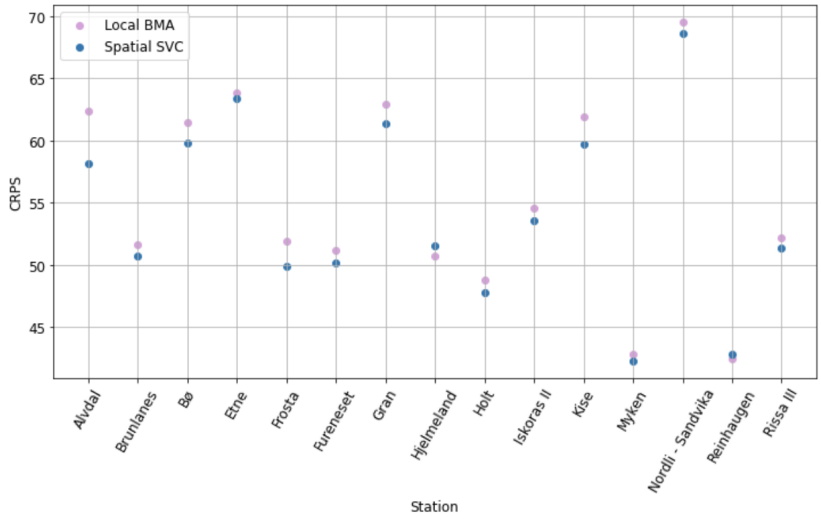


Figure 5.22: The Figure displays the average RMSE for each location the model is trained on. The blue dots are from the spatially varying coefficient model, while the pink dots are from the local BMA model.

5.3.5 Proximity to training locations

Furthermore, we can look at the performance of the validation locations, and compare it with the distance to the training locations. It is natural to assume better performance in the locations closer to the training locations, and it will be interesting to validate if this assumption is correct. A table with all the locations in the validation set, its closest location from the training set, and the distance between these two can be found in the Appendix. From the previous section, we remember the poor performance of the model in the locations Landvik, Lyngdal, and Pasvik - Svanvik. They are all over 200 km from the closest training location, and this can be a reason for the poor performance. When investigating the best-performing validation locations, We see Gjerpen, Ilseng, Lebergsfjellet, and Surnadal - Sylte as all with over 30% improvement in the CRPS compared to the NWP forecast. Both Gjerpen and Ilseng are close to a training location, but the two others are far away, with Lebergsfjellet being over 350km away from the closest training location. Therefore, it is inconclusive whether closer proximity to training locations leads to better or worse performance.

Chapter 6

Discussion and conclusion

This thesis focuses on the development and evaluation of a spatially varying coefficient model (SVCM) for post-processing solar radiation forecasts from NWP models. The SVCM utilizes a Beta distributed likelihood scaled by the physical constraints of solar radiation. For the case study conducted in Norway, the SVCM was trained at 15 distinct locations using a sliding window of 20 days. The model was evaluated on training and validation locations, as well as through cross-validation on locations with high-quality data to obtain an unbiased estimate. We also compared the model's performance on a previous study that utilized Bayesian Model Averaging (BMA) to locally post-process the NWP forecasts.

Continuous Ranked Probability Score (CRPS) was used to measure the predictive accuracy of the SVCM. The findings indicate a significant enhancement in the predictive accuracy of the model, achieving a 34.6% improvement in CRPS in the training locations and a 20.9% improvement in the validation locations. The model also outperformed the NWP forecasts based on the root mean squared error (RMSE) on the training locations, but failed to reduce the errors on the validation locations. To ensure calibration, a Probability integral transform (PIT) histogram was produced, and its visual inspection confirmed that the SVCM model improved calibration compared to the original ensemble forecasts. Additionally, the performance of the SVCM model was compared to a local BMA model previously utilized on this dataset by Egeli (2022). The Figures 5.21 and 5.22 demonstrate that the SVCM model improves the accuracy both in terms of CRPS and RMSE compared to the local BMA model.

The performance of the model on the validation locations, as measured by the

deterministic accuracy through RMSE, was not superior to the raw ensembles. However, RMSE is not the primary metric for probabilistic forecasting. As mentioned in the introduction, the objective of probabilistic forecasting is to ensure calibration while maintaining sharpness, and RMSE does not fulfill either of these objectives. The results obtained from the PIT-histogram indicate that the model enhances the calibration of predictive distributions, while the CRPS estimation demonstrates that the model improves the predictive distribution concerning both calibration and sharpness.

When investigating why the spatial model performs better than the local BMA model from Egeli (2022), we can highlight some possible reasons. Firstly, the spatial model benefits from the spatial correlation among different locations, allowing it to learn from the nearby locations. This is not possible in the local model, which only considers the data from a single location. Secondly, the spatial model has a larger training set since it utilizes data from multiple locations. This can help to reduce overfitting and improve the model's generalization performance. Lastly, the local BMA model employs all six ensembles per day with equal weighting, while the spatial model takes the average of the six ensembles as input. Christiansen (2019) investigated the performance of the ensemble average compared to individual members, and concluded that the ensemble average has better performance. Overall, these factors may contribute to the superior performance of the spatial model compared to the local BMA model.

This study employs two parameters, the spatially varying intercept α and the spatially varying slope β, for spatial postprocessing of NWP forecast. In contrast, other spatially varying coefficient models, such as those implemented in Roksvåg et al. (2022) and Meehan et al. (2019), utilize multiple parameters, such as a fixed intercept and a spatiotemporal parameter. We do not employ a fixed intercept in this study due to the significant spatial variations observed in the bias plot from Figure 3.8. Consequently, a fixed intercept for all locations would probably not enhance the overall spatial model. Conversely, the addition of a spatiotemporal parameter may be advantageous as the prior day is likely more influential in forecasting than the preceding 19 days. This presents a potential avenue for future model improvements.

Bibliography

Mohamed R Abonazel, Zakariya Yahya Algamal, Fuad A Awwad, and Ibrahim M Taha. A new two-parameter estimator for beta regression model: method, simulation, and application. *Front. Appl. Math. Stat. 7: 780322. doi: 10.3389/fams*, 2022.

Fabian E. Bachl, Finn Lindgren, David L. Borchers, and Janine B. Illian. inlabru: an R package for bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution*, 10:760–766, 2019. doi: 10.1111/2041-210X.13168.

Åsmunnd Bakketun and Jørn Kristiansen. Strålinn: Evaluation of predicted shortwave radiation. *MET*, 8, 2018.

Sudipto Banerjee, Bradley P Carlin, and Alan E Gelfand. *Hierarchical modeling and analysis for spatial data*. CRC press, 2014.

Sándor Baran and Sebastian Lerch. Mixture emos model for calibrating ensemble forecasts of wind speed. *Environmetrics*, 27(2):116–130, 2016.

Christopher Bitter, Gordon F Mulligan, and Sandy Dall'erba. Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of geographical systems*, 9:7–27, 2007.

Roger S Bivand, Edzer J Pebesma, Virgilio Gomez-Rubio, and Edzer Jan Pebesma. *Applied spatial data analysis with R*, volume 747248717. Springer, 2008.

Marta Blangiardo and Michela Cameletti. *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons, 2015.

Jochen Broecker and Holger Kantz. The concept of exchangeability in ensemble forecasting. *Nonlinear Processes in Geophysics*, 18(1):1–5, 2011.

E. G. Dutton C. N. Long. Bsrn global network recommended qc tests, v2.0. 2010.

A. Mørk Cato. Sol kan bli like stort som vannkraft i norge. 2022.

Bo Christiansen. Analysis of ensemble mean forecasts: The blessings of high dimensionality. *Monthly Weather Review*, 147(5):1699–1712, 2019.

A Philip Dawid. Present position and potential developments: Some personal views statistical theory the prequential approach. *Journal of the Royal Statistical Society: Series A (General)*, 147(2):278–290, 1984.

ECMWF. Part iv: Physical processes, in: Ifs documentation–cy41r2, 2016.

André Egeli. Probabilistic forecasting of solar radiation in norway using bayesia model averaging for ensemble forecasts. 2022.

Silvia Ferrari and Francisco Cribari-Neto. Beta regression for modelling rates and proportions. *Journal of applied statistics*, 31(7):799–815, 2004.

Andrew O Finley. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods in ecology and evolution*, 2(2):143–154, 2011.

Maria Franco-Villoria, Massimo Ventrucci, and Håvard Rue. A unified view on bayesian varying coefficient models. 2019.

Inger-Lise Frogner, Ulf Andrae, Jelena Bojarova, Alfons Callado, PAU Escribà, Henrik Feddersen, Alan Hally, Janne Kauhanen, Roger Randriamampianina, Andrew Singleton, et al. Harmoneps—the harmonie ensemble prediction system. *Weather and Forecasting*, 34(6):1909–1937, 2019.

Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Constructing priors that penalize the complexity of gaussian random fields. *Journal of the American Statistical Association*, 114(525):445–452, 2019.

Carlo Gaetan and Xavier Guyon. *Spatial statistics and modeling*, volume 90. Springer, 2010.

Alan E Gelfand, Hyon-Jung Kim, CF Sirmans, and Sudipto Banerjee. Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98(462):387–396, 2003.

Tilmann Gneiting and Matthias Katzfuss. Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1:125–151, 2014.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

Tilmann Gneiting, Adrian E Raftery, Anton H Westveld, and Tom Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133(5):1098–1118, 2005.

Tilmann Gneiting, Fadoua Balabdaoui, and Adrian E Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, 2007.

Tilmann Gneiting, Sebastian Lerch, and Benedikt Schulz. Probabilistic solar forecasting: Benchmarks, post-processing, verification. *Solar Energy*, 252:72–80, 2023.

Virgilio Gómez-Rubio. *Bayesian inference with INLA*. CRC Press, 2020.

Thomas M Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560, 2001.

Trevor Hastie and Robert Tibshirani. Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(4):757–779, 1993.

Ola Haug, Thordis L Thorarinsdottir, Sigrunn H Sørbye, and Christian LE Franzke. Spatial trend analysis of gridded temperature data at varying spatial scales. *Advances in Statistical Climatology, Meteorology and Oceanography*, 6 (1):1–12, 2020.

HOMER. Global Horizontal Irradiance (GHI). https://www.homerenergy.com/products/pro/docs/3.11/global_horizontal_irradiance_ghi.html, 2017. Online; accessed 2022-12-01.

J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

Alexander Jordan, Fabian Krüger, and Sebastian Lerch. Evaluating probabilistic forecasts with scoringrules. *arXiv preprint arXiv:1709.04743*, 2017.

Andreas Kleiven. Bayesian model averaging using varying coefficient regression and climatology cumulative probability regression-a case study of postprocessing hydrological ensembles from osali. Master's thesis, NTNU, 2017.

Elias T Krainski, Virgilio Gómez-Rubio, Haakon Bakka, Amanda Lenzi, Daniela Castro-Camilo, Daniel Simpson, Finn Lindgren, and Håvard Rue. *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. CRC press, 2018.

Philippe Lauret, Mathieu David, and Pierre Pinson. Verification of solar irradiance probabilistic forecasts. *Solar Energy*, 194:254–271, 2019.

Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.

Timothy D Meehan, Nicole L Michel, and Håvard Rue. Spatial modeling of audubon christmas bird counts reveals fine-scale patterns and drivers of relative abundance trends. *Ecosphere*, 10(4):e02707, 2019.

NOAA. Numerical weather prediction (weather models). https://www.weather.gov/media/ajk/brochures/NumericalWeatherPrediction.pdf. Accessed: 2023-05-19.

The pandas development team. pandas-dev/pandas: Pandas, April 2023. URL https://doi.org/10.5281/zenodo.7794821. If you use this software, please cite it as below.

Adrian E Raftery, Tilmann Gneiting, Fadoua Balabdaoui, and Michael Polakowski. Using bayesian model averaging to calibrate forecast ensembles. *Monthly weather review*, 133(5):1155–1174, 2005.

Ibrahim Reda and Afshin Andreas. Solar position algorithm for solar radiation applications. *Solar energy*, 76(5):577–589, 2004.

Thea Roksvåg, Ingelin Steinsland, and Kolbjørn Engeland. Estimation of annual runoff by exploiting long-term spatial patterns and short records within a geostatistical framework. *Hydrology and Earth System Sciences*, 24(8):4109–4133, 2020.

Thea Roksvåg, Ingelin Steinsland, and Kolbjørn Engeland. A geostatistical spatially varying coefficient model for mean annual runoff that incorporates process-based simulations and short records. *Hydrology and Earth System Sciences*, 26(20):5391–5410, 2022.

Julie Røste. The importance of mesh resolution when using the spde approach. Master's thesis, NTNU, 2020.

Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71 (2):319–392, 2009.

Daniel Simpson, Håvard Rue, Andrea Riebler, Thiago G Martins, and Sigrunn H Sørbye. Penalising model component complexity: A principled, practical approach to constructing priors. 2017.

Shiliang Su, Chaoran Lei, Anyun Li, Jianhua Pi, and Zhongliang Cai. Coverage inequality and quality of volunteered geographic features in chinese cities: Analyzing the associated local characteristics using geographically weighted regression. *Applied geography*, 78:78–93, 2017.

US Department of Energy. Solar radiation basics). https://www.energy.gov/eere/solar/solar-radiation-basics, 2017. Online; accessed 2022-12-01.

Lance A Waller, Li Zhu, Carol A Gotway, Dennis M Gorman, and Paul J Gruenewald. Quantifying geographic variations in associations between alcohol distribution and violence: a comparison of geographically weighted regression and spatially varying coefficient models. *Stochastic Environmental Research and Risk Assessment*, 21:573–588, 2007.

Jingzhuo Wang, Jing Chen, Jun Du, Yutao Zhang, Yu Xia, and Guo Deng. Sensitivity of ensemble forecast verification to model bias. *Monthly Weather Review*, 146(3):781–796, 2018.

Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL https://ggplot2.tidyverse.org.

WMO. Guidelines on ensemble prediction systems and forecasting. *World Meteorological Organization Weather Climate and Water*, 1091, 2012.

Dazhi Yang, Wenting Wang, Christian A Gueymard, Tao Hong, Jan Kleissl, Jing Huang, Marc J Perez, Richard Perez, Jamie M Bright, Xiang'ao Xia, et al. A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. *Renewable and Sustainable Energy Reviews*, 161:112348, 2022.

Michaël Zamo and Philippe Naveau. Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. *Mathematical Geosciences*, 50(2):209–234, 2018.

Noshipo Zwane, Henerica Tazvinga, Christina Botai, Miriam Murambadoro, Joel Botai, Jaco D Wit, Brighton Mabasa, Siphamanadla Daniel, and Tafadzwanashe Mabhaudi. A bibliometric analysis of solar energy forecasting studies in africa. *Energies*, 15, 2022.

Appendix A

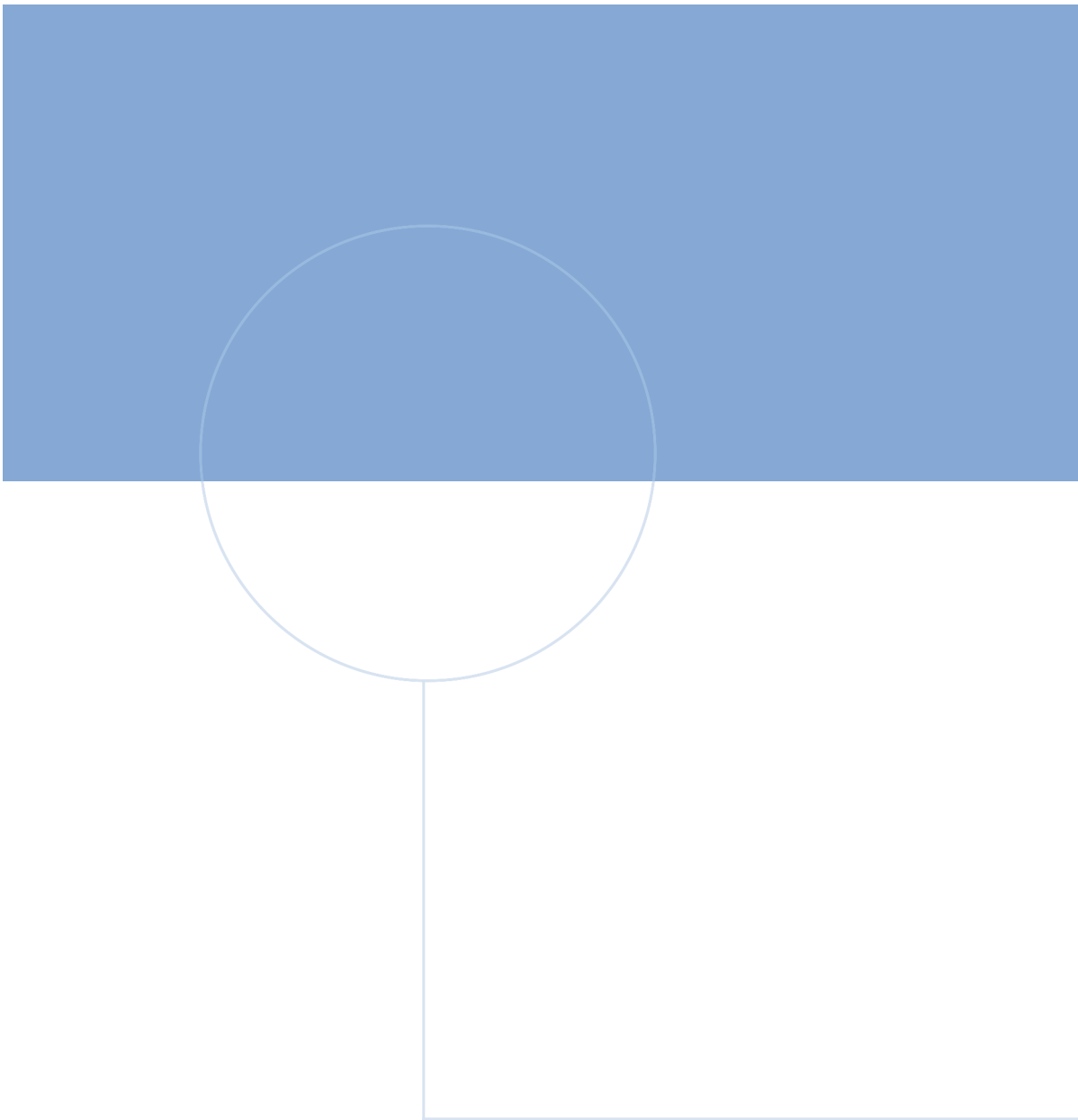
All locations

Navn	Latitude	Longitude	Altitude	Training Location
Alvdal	62.10944	10.62687	478	X
Apelsvoll	60.70024	10.86952	262	
Brunlanes	58.9878	9.96187	20	X
Bø	59.4175	9.02859	105	X
Etne	59.6625	5.95383	8	X
Frosta	63.56502	10.69298	18	X
Fureneset	61.29272	5.04428	12	X
Fåvang	61.45822	10.1872	184	
Gausdal	61.22468	10.25878	375	
Gjerpen	59.22684	9.57805	41	
Gran	60.35575	10.55906	245	X
Hjelmeland	59.22995	6.14992	43	X
Hokksund	59.76152	9.89166	15	
Holt	69.65381	18.90946	12	X
Hønefoss	60.14032	10.2661	126	
Ilseng	60.80264	11.20298	182	
Kise	60.77324	10.80569	129	X
Kvithamar	63.48795	10.87994	28	
Landvik	58.340071	8.522554	10	
Særheim	58.7605	5.6508	87	
Midtstova	60.6563	7.2755	1162	
Njøs	61.1792	6.8608	45	
Hansbu	60.0825	7.4247	1160	
Årnes	60.1268	11.3933	160	
Nordli - Sandvika	64.4595	13.5978	420	X
Skjetlein	63.3403	10.2973	48	
Rissa III	63.5858	9.9705	23	X
Lebergsfjellet	62.5158	6.8717	625	
Surnadal - Sylte	62.9853	8.6895	5	
Linge	62.288	7.2173	34	
Møsstrand II	59.8397	8.1785	977	
Myken	66.7628	12.486	17	X
Tjølling	59.0467	10.125	19	
Losistua	68.1905	17.7905	740	
Valnesfjord	67.2763	15.1018	20	
Lyngdal	58.134	7.0452	6	
Landvik	58.34	8.5225	6	
Nordnesfjellet	69.5575895	20.4152984	697	
Pasvik - Svanvik	69.4552	30.041	27	
Reinhaugen	70.3357	28.9648	470	X
Iskoras II	69.3003	25.346	591	X

Appendix B

Validation locations with distance to nearest training location

Validation location	Closest Training location	Distance (meters)
Apelsvoll	Kise	18079.12
Tjølling	Brunlanes	22181.04
Kvithamar	Frosta	28346.89
Ilseng	Kise	44731.64
Hønefoss	Gran	58300.66
Gjerpen	Brunlanes	67168.95
Skjetlein	Rissa Iii	71170.15
Årnes	Gran	106117.72
Særheim	Hjelmeland	115672.01
Gausdal	Kise	120212.9
Hokksund	Bø	122288.81
Møsstrand Ii	Bø	132651.09
Fåvang	Alvdal	160956.64
Nordnesfjellet	Holt	170423.91
Hansbu	Etne	188379.28
Njøs	Fureneset	203912.21
Surnadal - Sylte	Rissa Iii	206030.11
Landvik	Brunlanes	211887.46
Lyngdal	Hjelmeland	255015.97
Midtstova	Etne	266620.15
Pasvik - Svanvik	Reinhaugen	309338.3
Valnesfjord	Myken	325929.7
Linge	Fureneset	336842.94
Lebergsfjellet	Fureneset	353538.6
Losistua	Holt	469931.13



NTNU

Norwegian University of
Science and Technology