Bayesian modeling of monthly precipitation via integration of measurements and meteorological model output

Atli Norðmann Sigurðarson

Faculty of Physical Sciences
University of Iceland
2013

BAYESIAN MODELING OF MONTHLY PRECIPITATION VIA INTEGRATION OF MEASUREMENTS AND METEOROLOGICAL MODEL OUTPUT

Atli Norðmann Sigurðarson

60 ECTS thesis submitted in partial fulfillment of a
Magister Scientiarum degree in Mathematics

Advisor
Birgir Hrafnkelsson

Faculty Representative
Jóhanna Jakobsdóttir

M.Sc. committee
Birgir Hrafnkelsson
Sven Þ. Sigurðsson

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, May 2013

Bayesian modeling of monthly precipitation via integration of measurements and meteorological model output
Bayesian modeling of monthly precipitation
60 ECTS thesis submitted in partial fulfillment of a M.Sc. degree in Mathematics

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Hjarðarhagi 2-6
107, Reykjavik, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:
Atli Norðmann Sigurðarson, 2013, Bayesian modeling of monthly precipitation via integration of measurements and meteorological model output, M.Sc. thesis, Faculty of Physical Sciences, University of Iceland.

Printing: Háskólaprent, Fálkagata 2, 107 Reykjavík
Reykjavik, Iceland, May 2013

Abstract

The object of this study is to propose a Bayesian hierarchical model for observed monthly precipitation that incorporates predictors based on an output from a meteorological model. The observed data come from forty sites across Iceland. Each month is modeled separately with a Gaussian field with Matérn correlation function and measurement error at the data level. The location and scale parameters at the latent level are also modeled with Gaussian fields with Matérn correlation function. The observed data were collected and corrected for wind, wetting and evaporation loss by the Icelandic Meteorological Office (IMO). The IMO also provided an output from a linear model of orographic precipitation defined on a kilometer by kilometer grid over Iceland. The output from this model is used to construct predictors for the latent parameters on the grid. These predictors are then projected onto each of the observed sites for each month and incorporated into the Bayesian hierarchical model. Markov chain Monte Carlo (MCMC) was used for posterior inference for the parameters. Bayesian kriging is used to predict the latent parameters on the grid.

The model developed is quite detailed and accurate in describing the monthly precipitation even though, due to computational difficulties, some parameters had to be estimated outside the MCMC scheme. The results indicate that the output from the meteorological model is well suited to describing the mean monthly precipitation field over Iceland. The meteorological model is not as accurate in describing the variability of the monthly precipitation and generally underestimates it in the North-East part of Iceland and overestimates it in the South-West corner. The results also indicate that no long-term temporal trend is present in the monthly precipitation over the period examined.

Útdráttur

Markmið þessa verkefnis er að setja fram lagskipt Bayesískt líkan fyrir mælingar á mánaðarlegri úrkomu sem notar skýribreytur unnar úr útkomu veðurfræðilíkans. Mældu gögnin koma frá fjörutíu veðurstöðvum sem dreifast yfir Ísland. Hver mánuður er greindur út af fyrir sig með því að lýsa gögnunum með Gaussískum fleti sem lýst er með Matérn fylgnifalli ásamt mæliskekkju. Stikunum fyrir meðalgildi og dreifni mánaðarlegrar úrkomu, í öðru lagi líkansins, er einnig lýst með Gaussískum fleti með Matérn fylgnifalli. Gögnunum var safnað og þau leiðrétt fyrir vind, vætun og uppgufun af Veðurstofu Íslands (VÍ). Frá VÍ fengust einnig hermd úrkomugögn úr veðurfræðilíkani fyrir sérhvern punkt á kílómetra sinnum kílómetra neti yfir Íslandi. Hermda úrkoman úr þessu líkani er notuð til að útbúa skýribreytur fyrir Gaussísku fletina sem lýsa meðalgildi og dreifni mánaðarlegrar úrkomu á áðurnefndu neti. Þessum skýribreytum er varpað niður á veðurstöðvarnar fyrir hvern mánuð og þær innleiddar inn í lagskipta Bayesíska líkanið. Markov keðju Monte Carlo hermun er notuð til að fá eftirámat á stikum líkansins. Bayesíska útgáfan af aðferð Krige er notuð til að meta Gaussísku fleti meðalgildis og dreifni, í öðru lagi líkansins, í öllum punktum á netinu.

Líkanið sem sett var fram er margþætt og lýsir á sannfærandi hátt mánaðarlegri úrkomu, þrátt fyrir að nokkra stika hafi þurft að meta fyrir utan Markov keðju Monte Carlo reikniritið. Niðurstöðurnar sýna að hermdu úrkomugögnin úr veðurfræðilíkaninu lýsa vel meðalgildi mánaðarlegrar úrkomu á Íslandi. Veðurfræðilíkanið er ekki eins nákvæmt þegar kemur að dreifni mánaðarlegrar úrkomu. Það vanmetur almennt dreifnina á Norðausturlandi en ofmetur hana á Suðvesturhorninu. Niðurstöðurnar sýna einnig að engin langtímaleitni er til staðar í mánaðarlegri úrkomu yfir tímabilið sem skoðað var.

Contents

List of Figures

List of Tables

Acknowledgments

This M.Sc. project was carried out at the Department of Mathematics, which is under the Faculty of Physical Sciences, at the University of Iceland and was partially funded by the University of Iceland Research Fund.

First of all, I would like to specially thank my advisor, Birgir Hrafnkelsson, for his guidance, support and patience throughout this project. I also thank my co-advisor, Sven Þ. Sigurðsson, for proofreading the thesis and providing insightful comments.

I would like to thank Philippe Crochet at the Icelandic Meteorological Office for providing the corrected rain gauge data and the output from the linear model of orographic precipitation.

A special thanks go to my friend and fellow student, Óli Páll Geirssons, for the use of his method to project the covariates onto the rain gauge sites.

I want to express my gratitude to my family, which has always supported me and taken interest in my studies. Finally, I would like to thank my girlfriend, Elín, for her loving support and encouragement over the last months of this project.

Reykjavík, May 2013.

Atli Norðmann Sigurðarson.

1 Introduction

1.1 Overview of the project

There are two main goals of this project. The first one is to propose a realistic Bayesian hierarchical model for the accumulated monthly precipitation in Iceland. The second one is to incorporate an output from a linear model of orographic precipitation, into the Bayesian model, both to benefit the statistical model and to assess the accuracy of the meteorological model.

The Bayesian hierarchical model incorporates all available scientific knowledge about precipitation as well as all reasonable assumptions that could be made. To that end the model assumes a complex distribution for the data which incorporates both a spatial and temporal component in the mean structure and a complex correlation structure, which incorporates site specific variance, correlation between data and measurement error. Of course, due to the problem of dimensionality and other problems associated with simulations, the model presented in this thesis is not as complete or complex as one would want, but it is a realistic model and more elaborate than most statistical models generally used in analysis of climatological data of this type.

The linear model of orographic precipitation is a purely meteorological model. It does not include a statistical component or incorporate observed data. This model, explained in [1], was found to be in good agreement with observations. It is a part of this project to get a more detailed comparison between the output from the meteorological model and the observed data.

1.2 Possible applications

Good understanding of the distributional properties of precipitation is very important for understanding and modeling the hydrological cycle. Associated with the hydrological cycle are phenomena such as floods, droughts, landslides and avalanches which are of vital importance for public safety. A good understanding of the cycle

aids in the understanding and explanation of these phenomena. It can also help with predicting reservoir levels in dams linked to hydrological power stations. A better knowledge of the hydrological cycle is also important in various scientific fields such as hydrology, glaciology, climatology and meteorology. A realistic statistical model adds to the understanding of the cycle and related phenomena as well as providing information about regional climate and climate change.

Detailed meteorological models and their development play an important role in tasks such as weather forecasts and describing the distribution of meteorological variables over various timescales. It is therefore important that these models are as accurate as possible. A comparison between the output from such models and observed values and an assessment of their accuracy can give information about how to best implement these meteorological models and aid in the development of more sophisticated models.

1.3 Literature review

Since Wikle et al. [2] proposed the use of hierarchical spatio-temporal models for the analysis of environmental data distributed in time and space, these models have been developed for usage in environmental sciences to describe various processes.

Many climatological models are being used as well to describe the distribution of variables such as precipitation and temperature. The output of these climatological models are often combined (see e.g. [3]), but there is a reason to expect that not all of these models can be trusted equally well and that some should receive more weight than others in the combination of their results. Therefore, a comparison between the output from these models and observed data with statistical methods is important. Tebaldi and Sansó [4] tackle this problem and show that the Bayesian approach is well suited to quantifying the uncertainty that shows up in the varying results of the various models. Their study focused on analyzing past and future trends jointly for precipitation and temperature. A number of general circulation models were used in the study, which they regarded as a biased and noisy versions of the true climate process. Their approach provides an estimate of the overall bias for an ensemble of model simulations. Often only a simple statistical approach is used for comparing outputs from climatological models and observed data. In [5] and [1] an inspection by eye and a simple regression analysis, which does not take into account any spatial correlation, is used. The Bayesian model presented here is well suited to describing the variable being analyzed and thus well suited to comparing the observed data to the output from the meteorological model. It not only allows for site by site comparison as in [5] and [1] but also evaluates the distributional properties of the whole field. The method presented here could also be used to compare observed data

to outputs from meteorological models other than the linear model of orographic precipitation used here.

Due to sparsity of rain gauge sites, good methods for spatial predictions are important. For such predictions several kinds of information sources can be used [6]. Digital maps of climate parameters at a fine resolution are required for example to run models of ecosystem carbon and water cycling [7]. Various means have been proposed for these spatial predictions. The ordinary kriging method is the most widely used but there are also other methods such as elevationally detrended kriging, elevational cokriging and the PRISM (Precipitation-elevation Regression on Independent Slopes Model) method developed in [8]. Kriging methods benefit greatly from a set of highly relative covariates on a fine grid. In this project we have access to such covariates and are therefore in the position to obtain good spatial predictions.

Combining statistical methods and meteorological model output is an active area of research. Cooley et al. [9] presented a method to predict the distributional properties of extreme precipitation by using a geographical covariate (elevation) and a climatological covariate (mean precipitation) on a grid. We use a similar approach to theirs, in modeling monthly precipitation, in that we use spatial latent fields for the location and scale parameters, which incorporate both geographical and climatological covariates. The advantage of our method over theirs lies in the quality of our covariate, which takes into account more complex factors such as topology, airflow dynamics, condensed water advection and downslope evaporation and is available on a finer grid.

The novelty of our approach lies in incorporating a meteorological covariate, based on a linear model of orographic precipitation (see [1]), into a statistical model. Also, we assume a correlation structure, as well as measurement error, at the data level and include a spatially varying variance at the latent level. Finally, distributional assumptions of the observations are handled with a Box-Cox transformation (see [10]) to achieve normality.

1.4 Structure of the thesis

The thesis is structured in the following way. In Chapter 2 the observed precipitation data and the output from the linear model of orographic precipitation are presented. The data from rain gauge sites are examined and its distributional properties explored. The output from the meteorological model is presented and it is explained how it was used to create covariates for the rain gauge sites. In Chapter 3 a novel Bayesian hierarchical model for monthly precipitation is introduced. The motivation behind the data distribution and the prior distributions is given and their details

described. There is a section about the derivation of the conditional posterior distribution of the parameters in the model, which are used for posterior inference. The chapter concludes with a discussion about the complications of the model and how the sampling scheme was finally set up. In Chapter 4 we present the main results of the Bayesian hierarchical model, predict the location and scale parameters on a 1-km grid over Iceland and compare the output from the aforementioned meteorological model to these predictions. Chapter 5 concludes the thesis with some discussion about the model and its implementation and the comparison with the meteorological model. Finally, there is a section about ideas on future studies.

2 The Data

2.1 Rain gauge data

The data analyzed in this project were supplied by Philippe Crochet at the Icelandic Meteorological Office (IMO). They consists of accumulated monthly precipitation at 40 different sites, corrected for wind, wetting and evaporation loss, over the period from 1958 to 2006. The correction method can be found in [1]. The sites are mostly located at low elevation and scattered across Iceland. Most are close to the coast but some are more inland (cf. Fig. 2.1). The decision to work with this corrected data, as opposed to completely raw rain gauge data, is natural since the intention is to compare the data to an output from a linear model of orographic precipitation developed in [1]. The simulated precipitation output from this model is not subject to errors due to wind, wetting and evaporation and therefore it should be compared to data corrected for such errors. The output from the linear model of orographic precipitation was also provided by Philippe Crochet at the IMO.

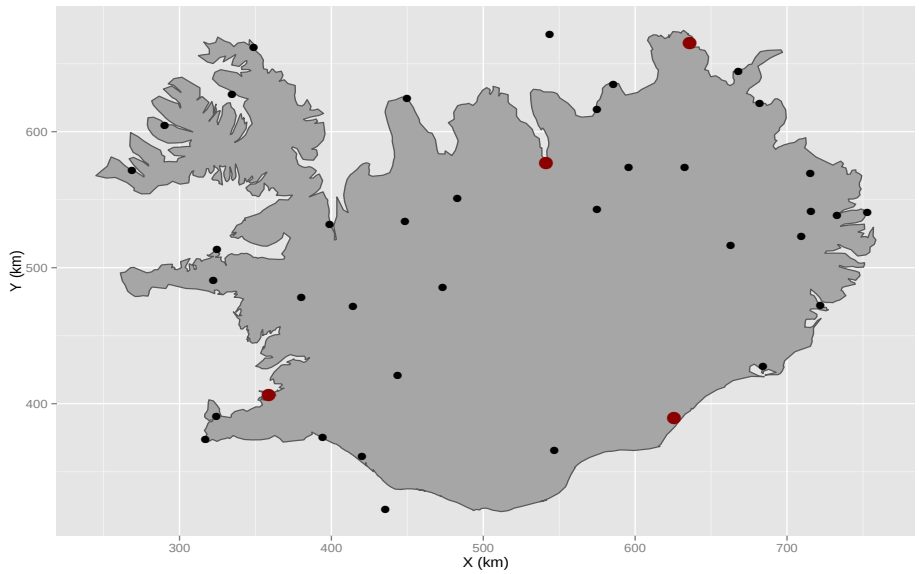

Figure 2.1: Location of rain gauge sites across Iceland.

The IMO also provided some site specific data, such as latitude, longitude and height above sea as well as xy-coordinates on the ISN93 grid for Iceland. The xy-coordinates as well as height above sea were used as covariates in the model.

A decision was made early to analyze each month separately. This removes the necessity of incorporating temporal correlation in the model and also alleviates the problem of dimensionality. A first look at the data suggests that there is substantial difference between months and also between stations (cf. Fig. 2.2). It appears

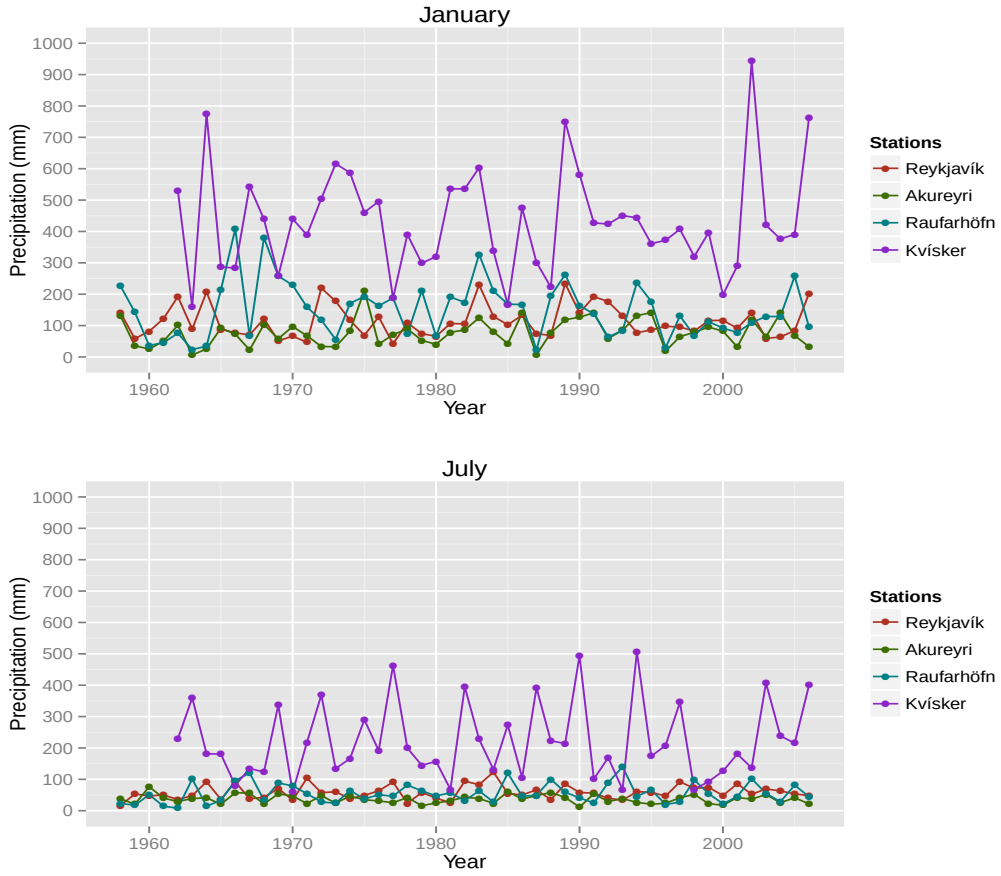

Figure 2.2: Time series for selected stations in January and July.

necessary to incorporate site specific mean and variance in the model (cf. Fig. 2.3). Figure 2.2 does not show a clear temporal trend so there does not seem to be a reason to incorporate that in the model. Figure 2.3 indicates that at very few stations the precipitation follows a symmetric distribution so it doesn't seem plausible to assume a normal distribution for the data. The data also exhibits a clear spatial correlation (cf. Fig. 2.4) which will be incorporated in the model.

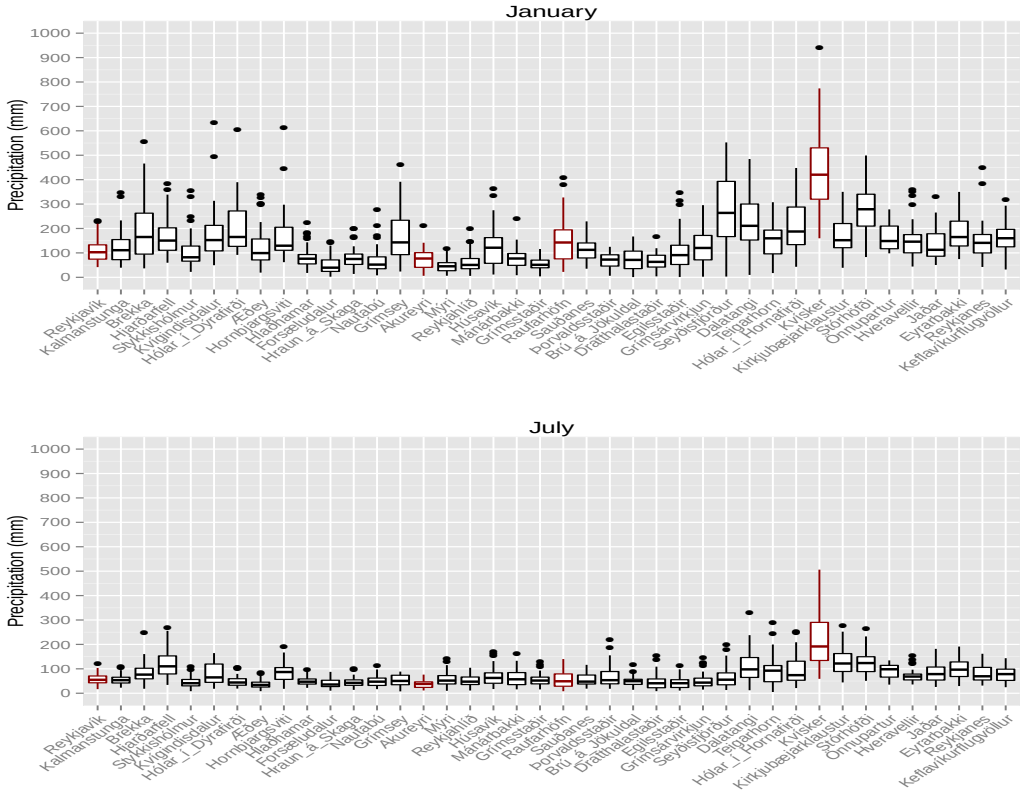

Figure 2.3: Boxplot of the data in January and July.

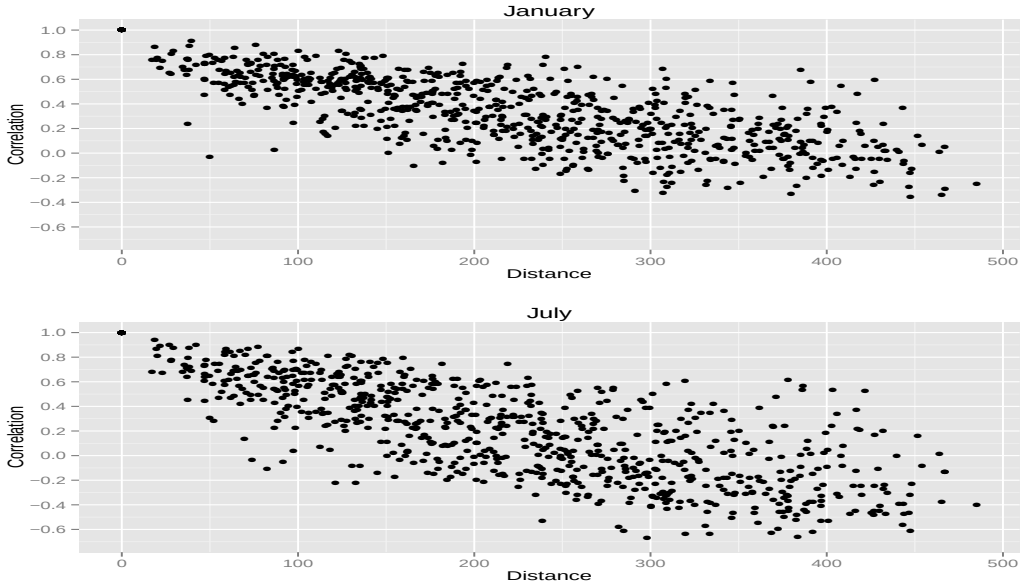

Figure 2.4: Correlation of the data in January and July as a function of distance between sites.

2.2 Distributional properties

In general when modeling data, it is convenient to assume normality for the observed values. Of course this assumption should be evaluated and tested. To get an idea about the data distribution normal and log-normal quantile-quantile plots were made for the data. The Anderson-Darling test for normality (see e.g. [11] and [12]) was also performed on both the data and on the log transformed data. Each station for each month was examined and although in some cases a normal, or log-normal, distribution might be adequate, neither of these were deemed good enough to be used as the distributional form for all the data.

Figure 2.5 shows the Q-Q normal plot for the data from selected stations in January and 2.6 show the Q-Q normal plot for the natural logarithm of the data from the same stations in January. As can can be seen in the figures neither the normal distribution nor the log-normal distribution fit the data adequately.

Figure 2.7 shows the p-values from the Anderson-Darling test for the data and the natural logarithm of the data from all the stations in January as well as the Q-Q uniform plot for the p-values. If the composite hypothesis of normality were true the p-values from the Anderson-Darling test would roughly follow a uniform distribution. It is clear from Figure 2.7 that this is not the case.

Since transforming the data with the natural logarithm was unsuccessful other transformations had to be considered. To that end it was decided to use the Box-Cox transformation [10] which is given by

$$\tilde{y} = \begin{cases} \dfrac{y^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln(y) & \text{if } \lambda = 0. \end{cases} \tag{2.1}$$

The problem then arose, which value of the Box-Cox parameter λ would be suitable. It was preferred, for simplicity, to use the same value of λ for all the sites in all the months. To determine the most suitable value of λ, the Box-Cox log-likelihood function for the power transformation was plotted for the data at each site in each month, for values of λ in the interval $(-2, 2)$. Figure 2.8 shows these plots for sites no. 1-4 in January and sites no. 13-16 in July. All of the figures showed the top of a concave function, and therefore there was no need to consider values of λ outside of the interval $(-2, 2)$ [13]. The value $\lambda = 0.4$ was in the 95% confidence interval for most of the stations for all of the months. The log-likelihood function for λ was also examined when using the data from all the sites together in each month and finally it was examined when all the months had been taken together. The value for λ which then maximized the function was $\lambda = 0.4$. Figure 2.9 shows the log-likelihood function for the power transformation when all the sites in each month and all the months had been taken together. Based on this aforementioned examination of the

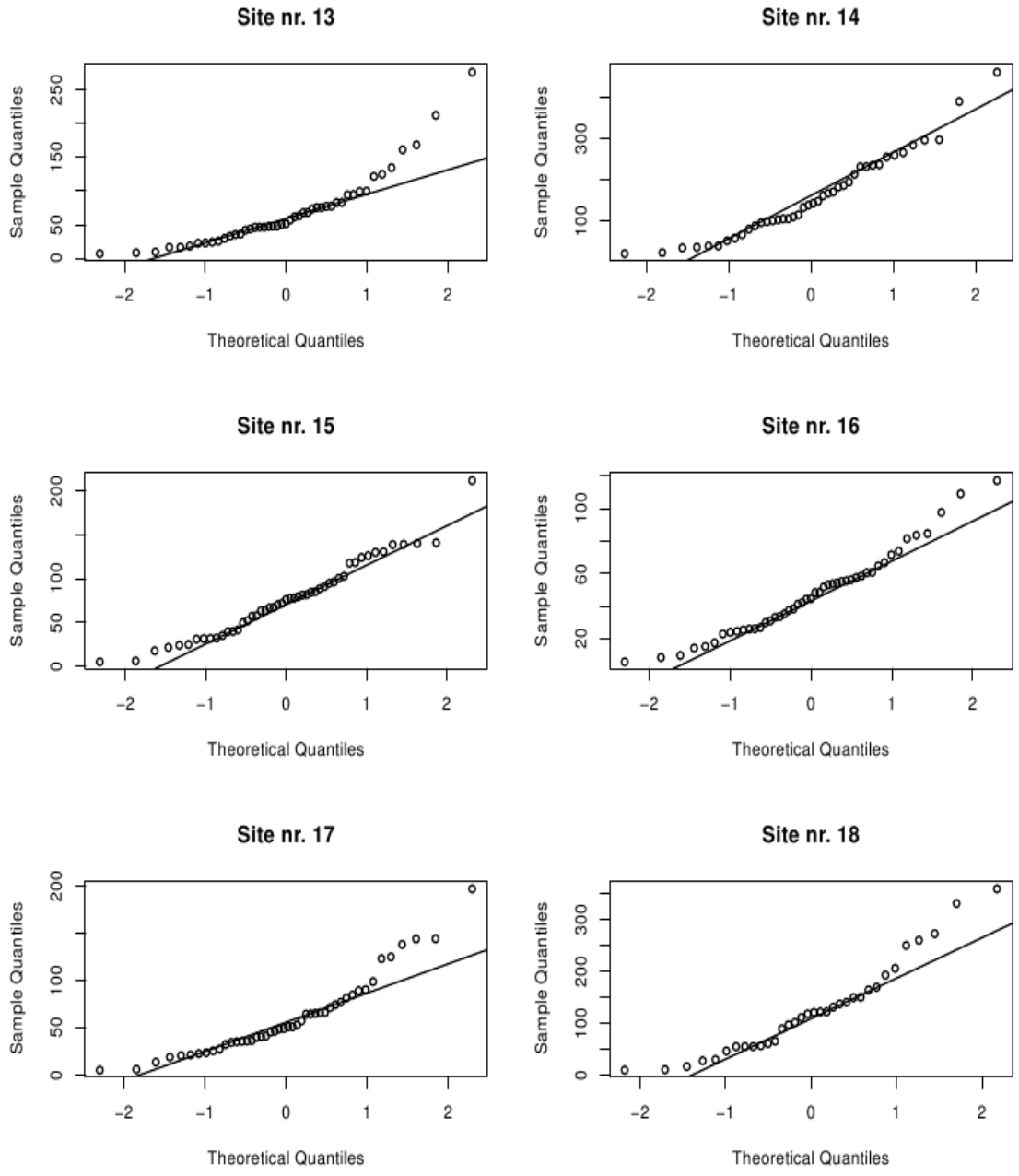

Figure 2.5: Q-Q plot of the data from selected sites in January.

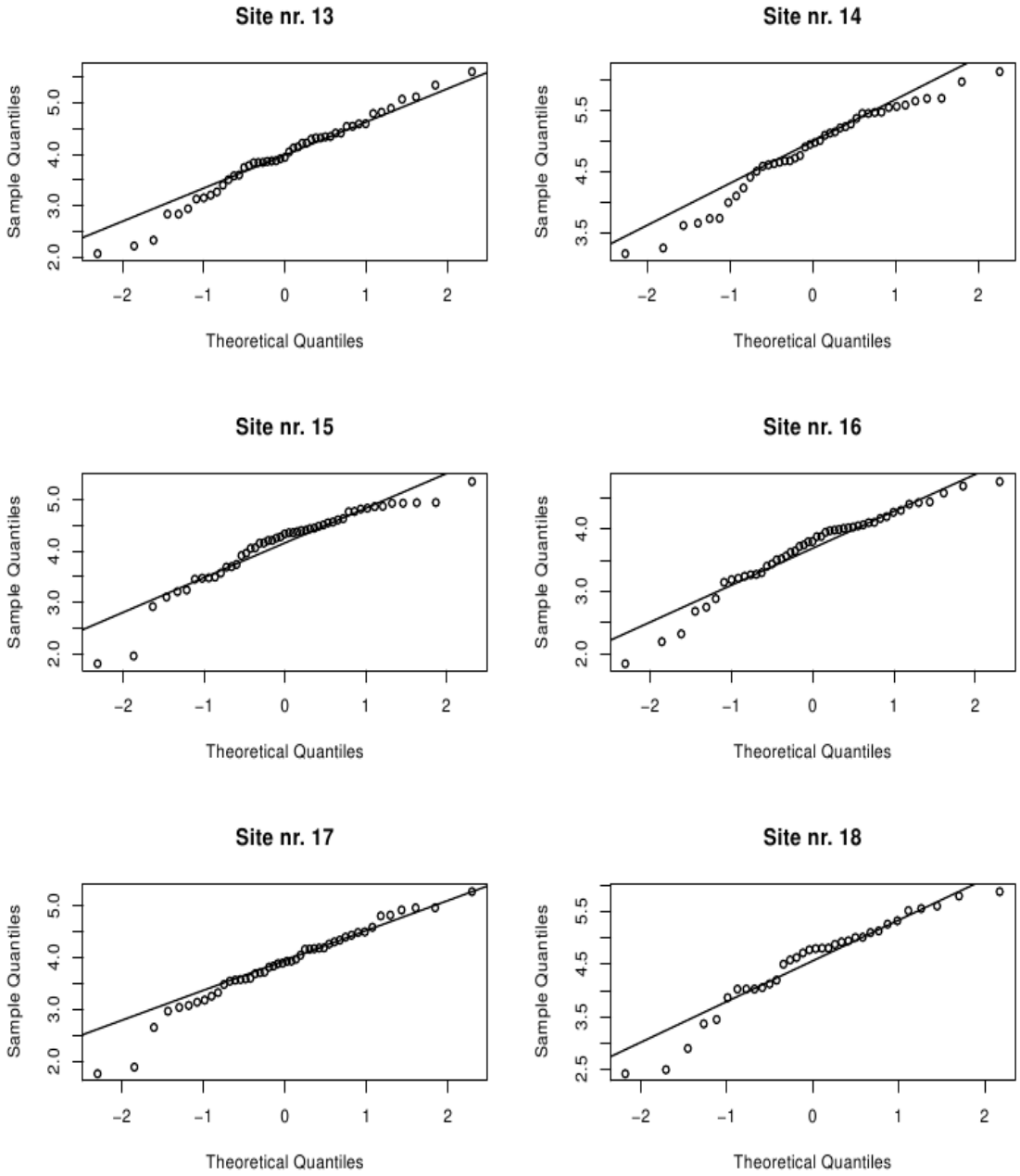

Figure 2.6: Q-Q plot of the natural logarithm of the data from selected sites in January.

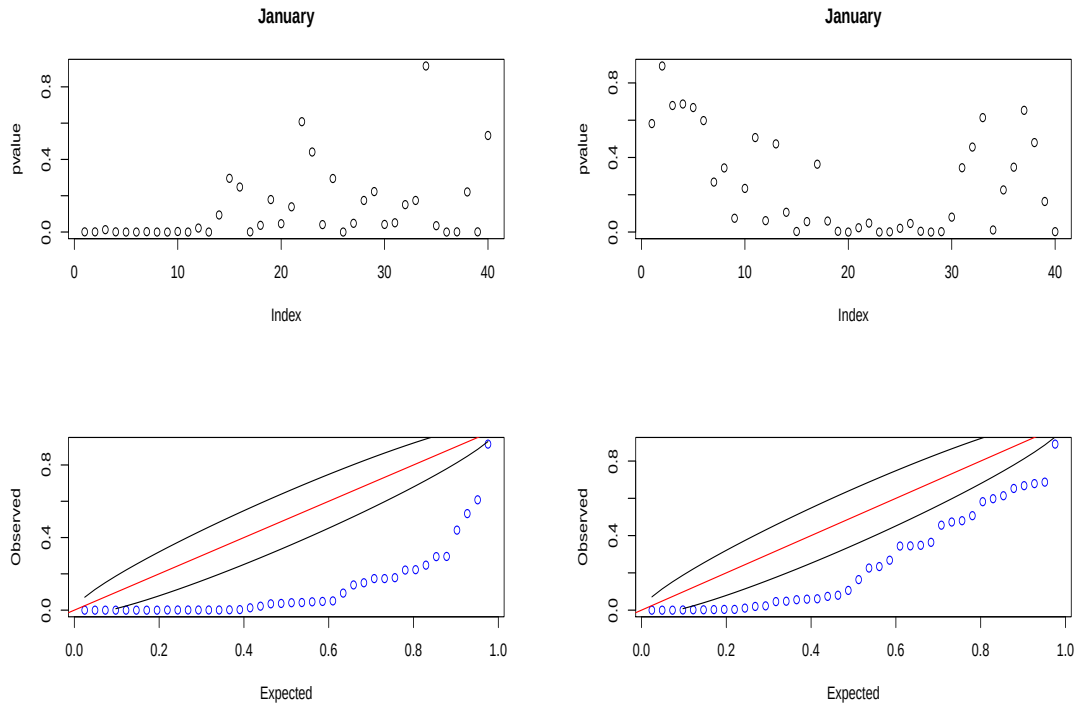

Figure 2.7: p-values from the Anderson-Darling test for the data (left) and the natural logarithm of the data (right) from all the stations in January.

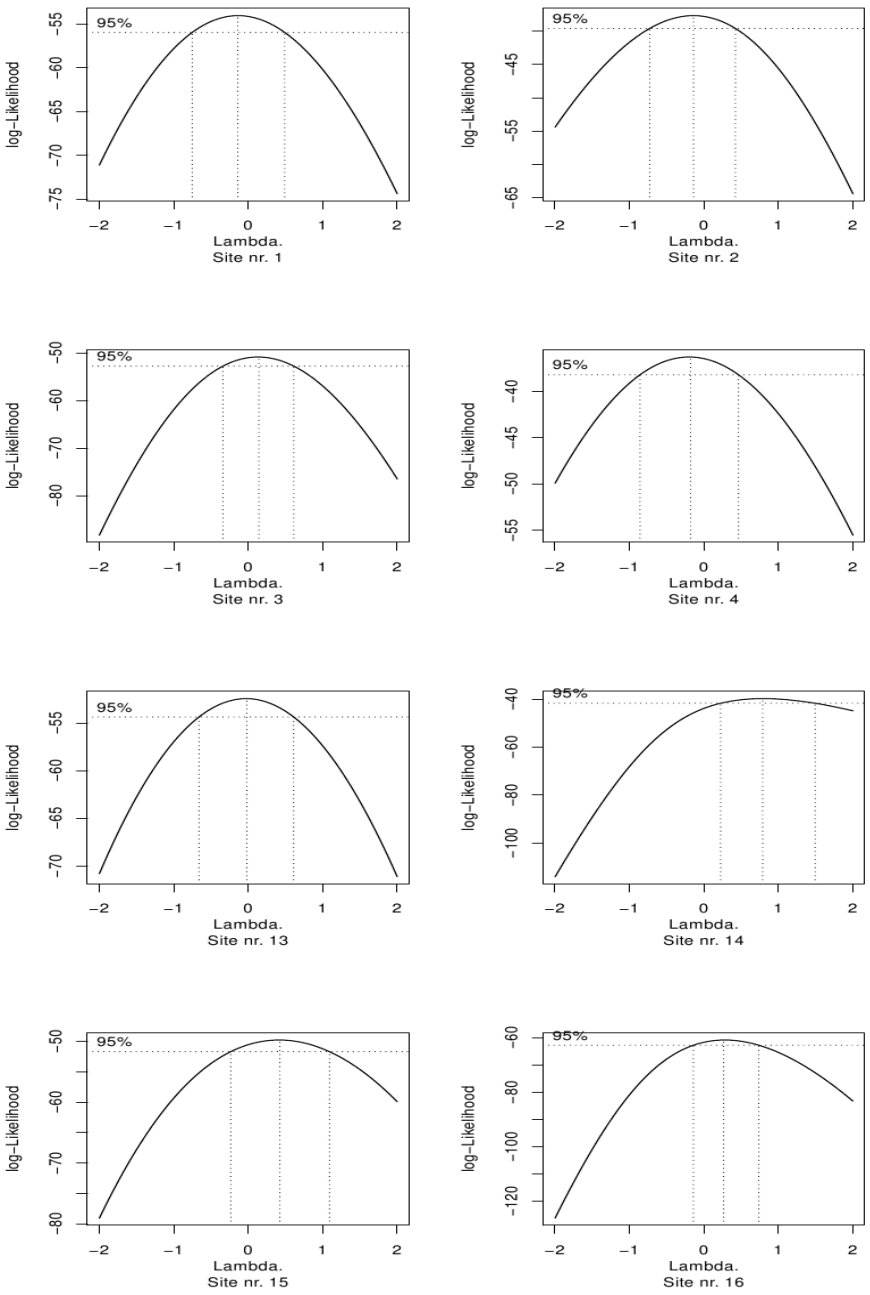

Figure 2.8: The log-likelihood function for the Box-Cox parameter λ for sites no. 1-4 in January (upper) and sites no. 13-16 in July (lower).

data it was decided to use the value $\lambda = 0.4$ to transform the data for all the stations in each month.

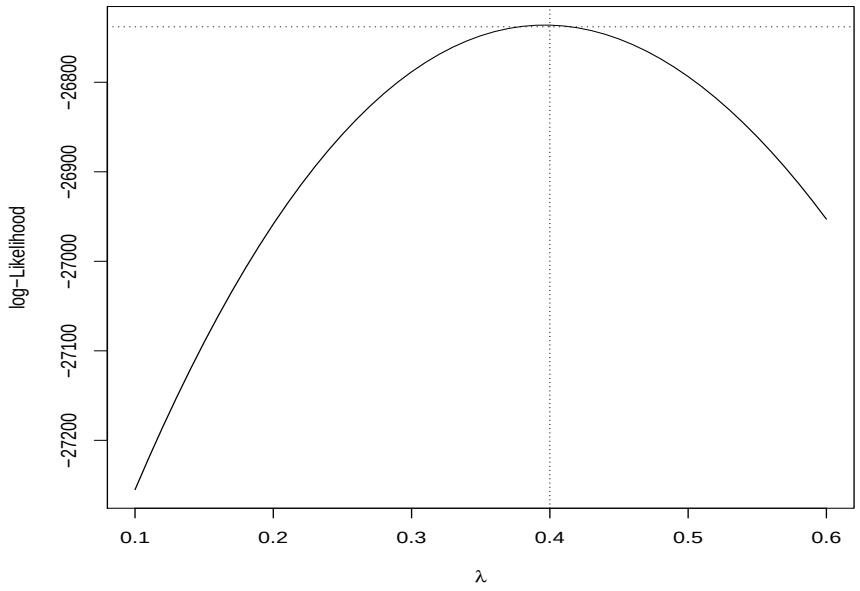

Figure 2.9: The log-likelihood function for the Box-Cox parameter λ with all the sites in each month and all the months taken together.

Figure 2.10 shows the normal quantile-quantile plot for the data from the same sites in January, as before, after the Box-Cox transformation with $\lambda = 0.4$. The figure clearly shows that the transformed data is closer to following a normal distribution than the untransformed data or the data transformed with the natural logarithm. Figure 2.11 shows the Anderson-Darling p-values for the data from all the stations in January and July after the Box Cox transformation. For July the fit is very good and it is save to assume that the transformed data for July is adequately described by the normal distribution. The fit for January is not as good as for July, and this was also the case with a few other months. However, most of them showed a good fit, and in any case a much better fit than with the untransformed data or the data transformed with the natural logarithm. Since it was desired to use the same value of λ for each month it was expected that not all of them would exhibit a perfect fit. The Box-Cox transformed data fitted the normal distribution much better than both the untransformed data and the data transformed with the natural logarithm.

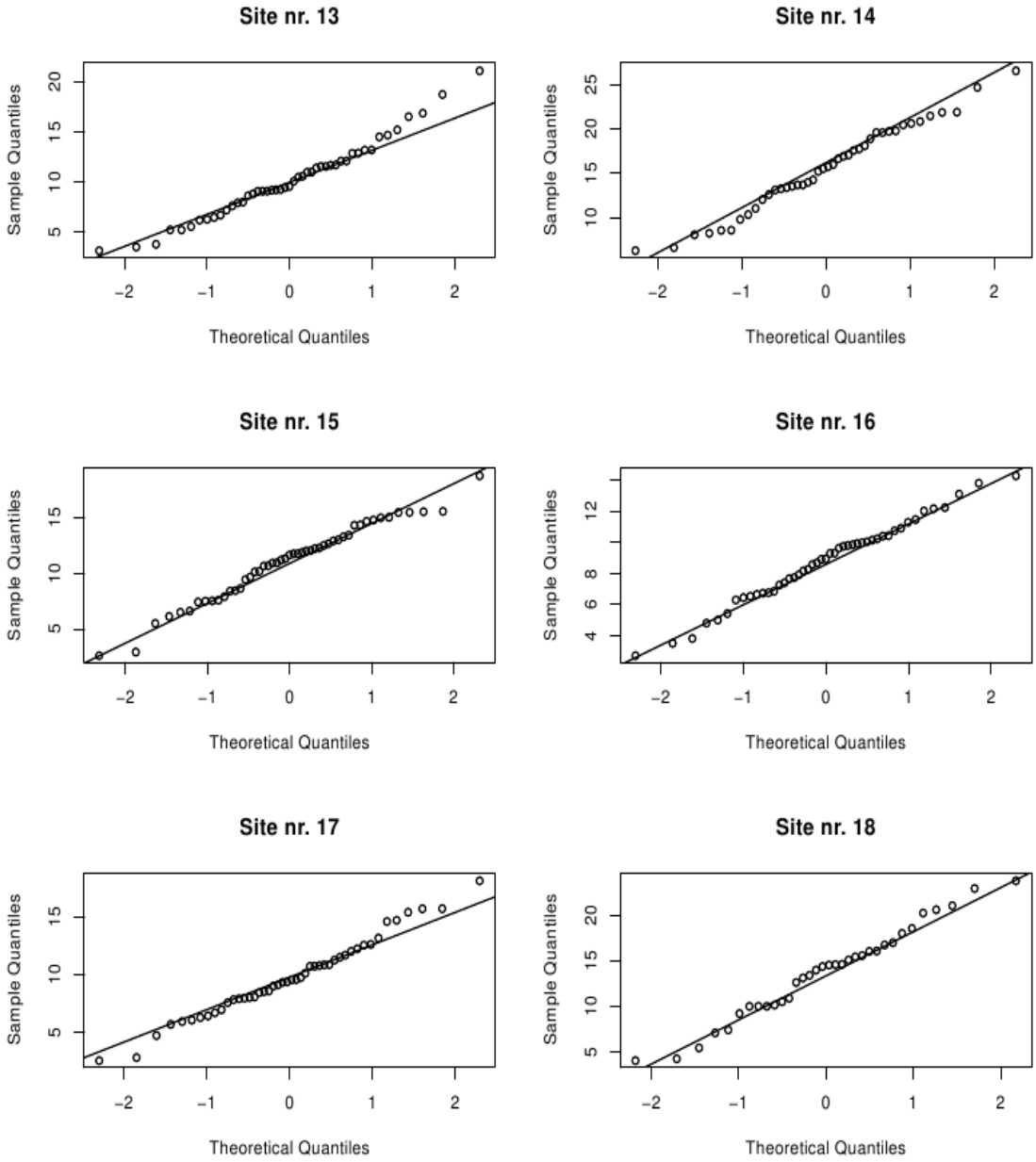

Figure 2.10: Q-Q plot of the Box-Cox transformed data, with parameter $\lambda = 0.4$, from selected sites in January.

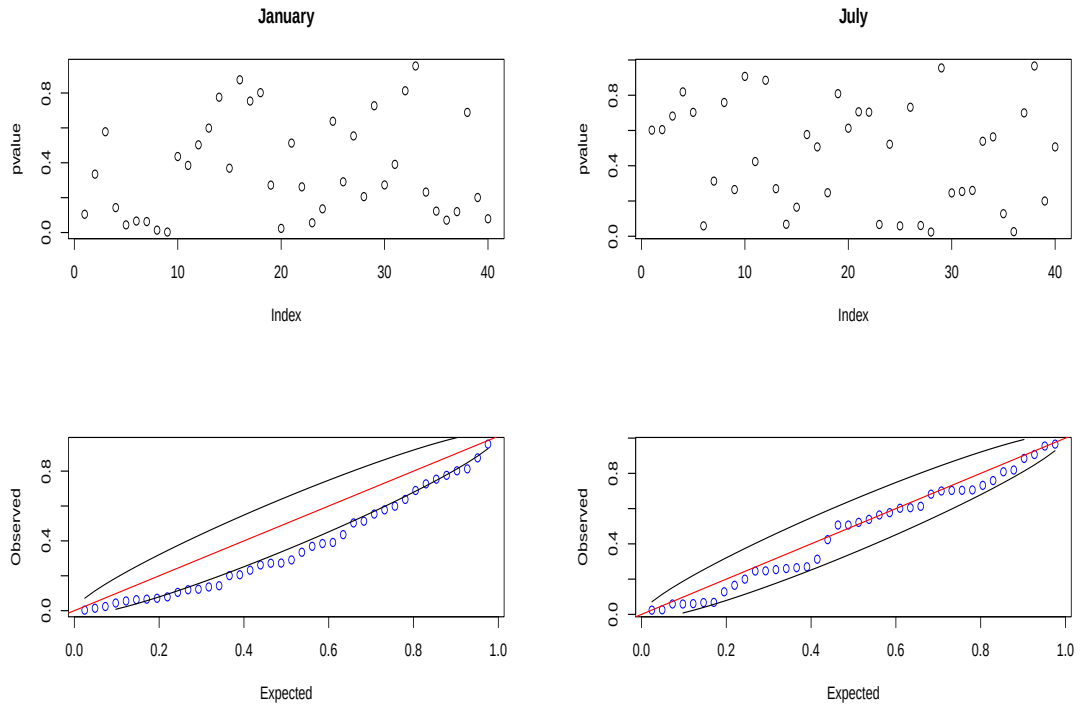

Figure 2.11: p-values from the Anderson Darling test for the transformed data from all the sites in January (left) and July (right).

2.3 Meteorological covariates

As already mentioned, one of the goals of this project was to compare the corrected rain gauge data with the simulated precipitation output from the meteorological model developed in [1]. This output was provided by Philippe Crochet at the IMO. The simulated precipitation consisted of gridded simulated daily values on a 1-km grid over Iceland for the period 1958-2002.

The simulated precipitation is contained in 6 netCDF files. Since the data being analyzed is monthly accumulated precipitation the simulated daily values in the netCDF files needed to be added together within each month to be comparable to the data. This was done for all the years 1958-2002 and all the months, taking special care of leap years by adding over 29 days for February of those years. The monthly accumulated simulated precipitation was stored in 12 arrays, one for each month. Each of these arrays contained the simulated monthly value for one month in each of the grid points on the 1-km grid over Iceland, over the whole period 1958-2002. Since the observed data was to be transformed with the Box-Cox transformation with parameter $\lambda = 0.4$ (2.1), the simulated monthly values were transformed with the same transformation, to be comparable to the data. The intention was to use a statistic based on the simulated precipitation as a covariate in the spatial mean and variance of the data. To that end the mean and variance of the transformed simulated monthly values were calculated for each grid point over the period 1958-2002 and stored in an array for each month.

Figures 2.13 to 2.16 show the meteorological covariate used for the spatial mean on the 1-km grid over Iceland for all the months. They show the mean in each grid point transformed back with the inverse of the Box-Cox transformation to have the scale in millimeters and to be comparable to the median predicted precipitation from the model (see Chapter 4, Figures 4.7 to 4.18). Figures 2.17 to 2.20 show the meteorological covariate used for the spatial variance on the 1-km grid over Iceland for all the months. They show the natural logarithm of the square root of the variance in each grid point since that was how the covariate was used in the model, as the mean of the parameter $\vec{\tau}$ (see Section 3.2).

With the mean and variance of the Box-Cox transformed simulated monthly precipitation available in each grid point it was possible to construct a covariate based on those values for each of the 40 sites across Iceland. To accomplish that we applied a method developed by doctoral student, Óli Páll Geirsson, for his Ph.D. project. This method is such that for each site we find which grid points are contained in a circle around the site with radius $r = 1$ km. These grid points are then given weights according to the $1 - F_{Beta}(d/r \,|\, \alpha = 10, \beta = 10)$ function, where F_{Beta} is the cumulated distribution function of the Beta distribution, d is the distance to the grid point from the site and α and β are the parameters in the Beta distribution (cf.

Fig. 2.12). Finally, the weighted mean of the values in the grid points is calculated, with the given weights.

The values $r = 1$ and $\alpha = \beta = 10$ were found by minimizing the mean square error of the variables produced and the monthly mean and variance of the data. The constraint $\alpha = \beta$ was imposed in the minimization.

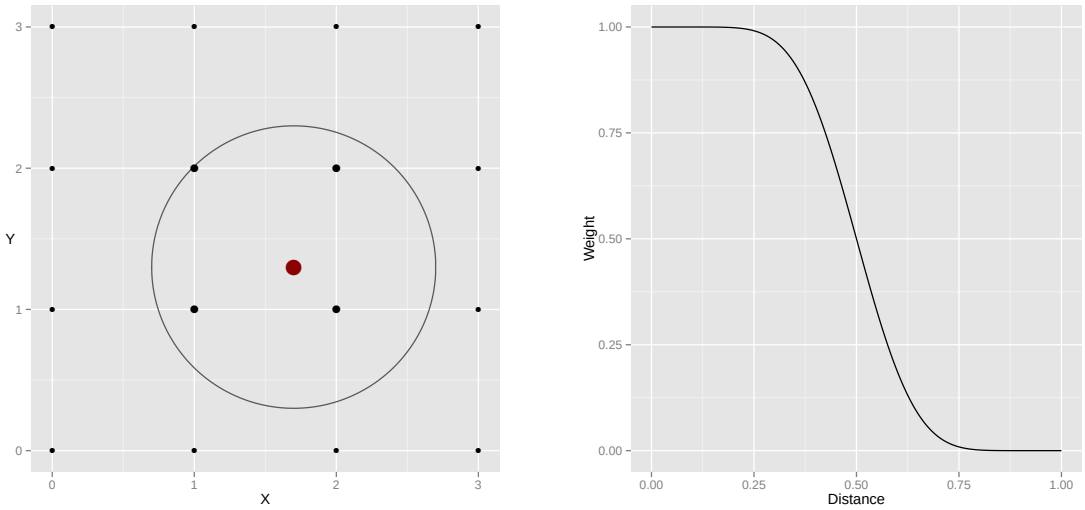

Figure 2.12: Construction of the meteorological covariates for the rain gauge sites.

After applying this method, covariates, based on the transformed simulated monthly precipitation, were readily available for the spatial mean and variance of the transformed data for each of the 40 rain gauge sites.

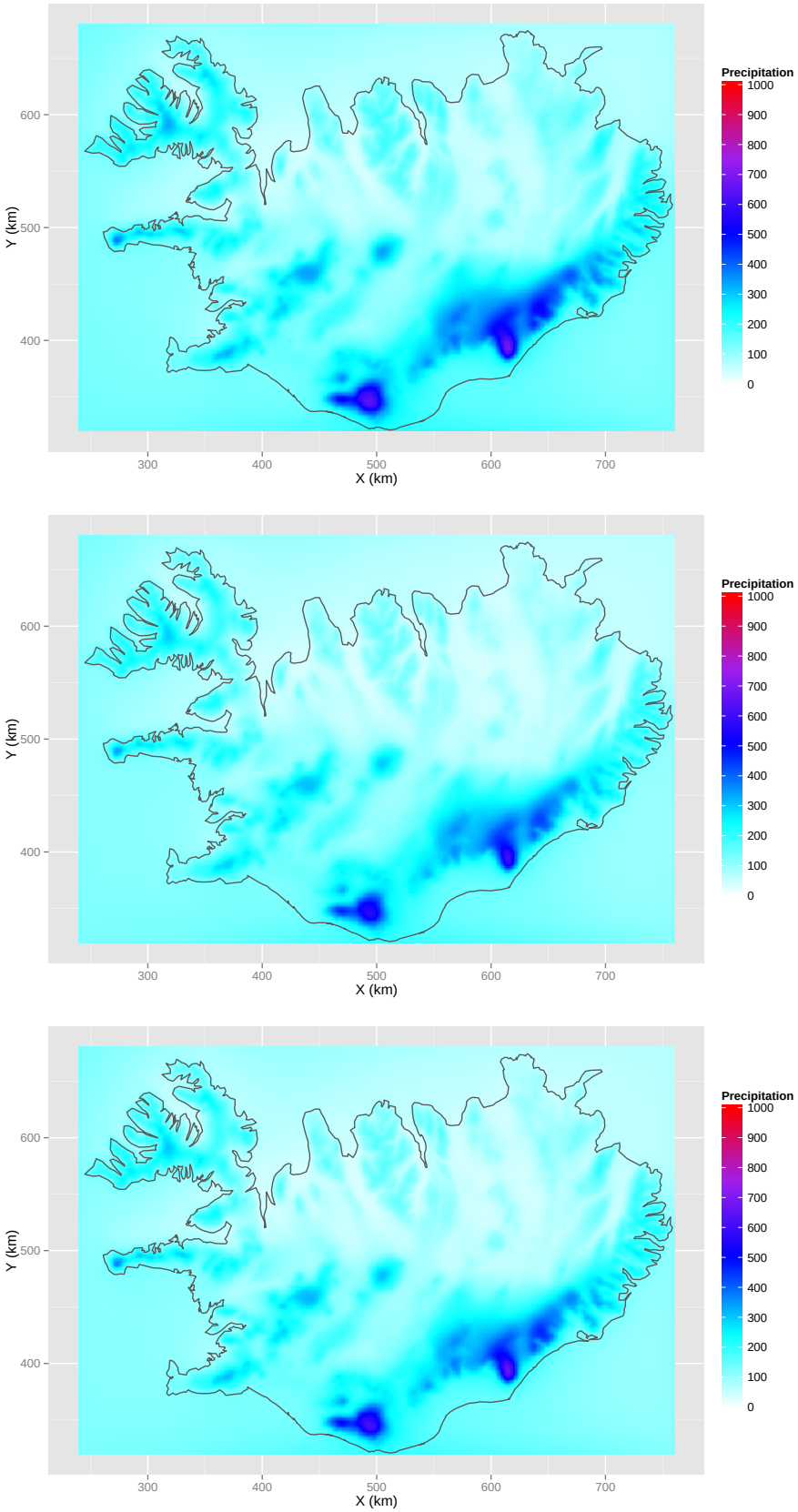

Figure 2.13: Meteorological covariate for the spatial mean on the 1-km grid over Iceland for January-March.

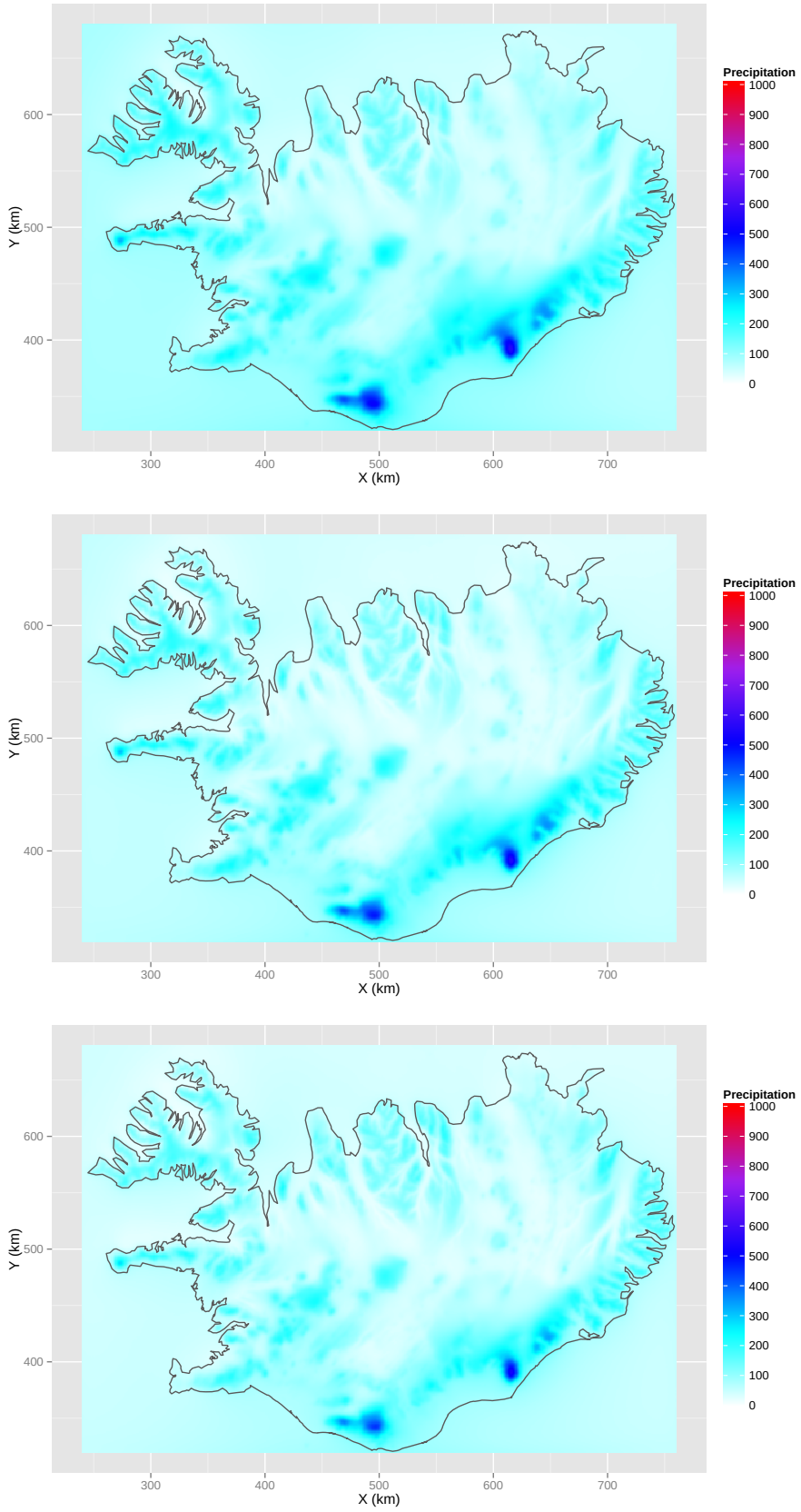

Figure 2.14: Meteorological covariate for the spatial mean on the 1-km grid over Iceland for April-June.

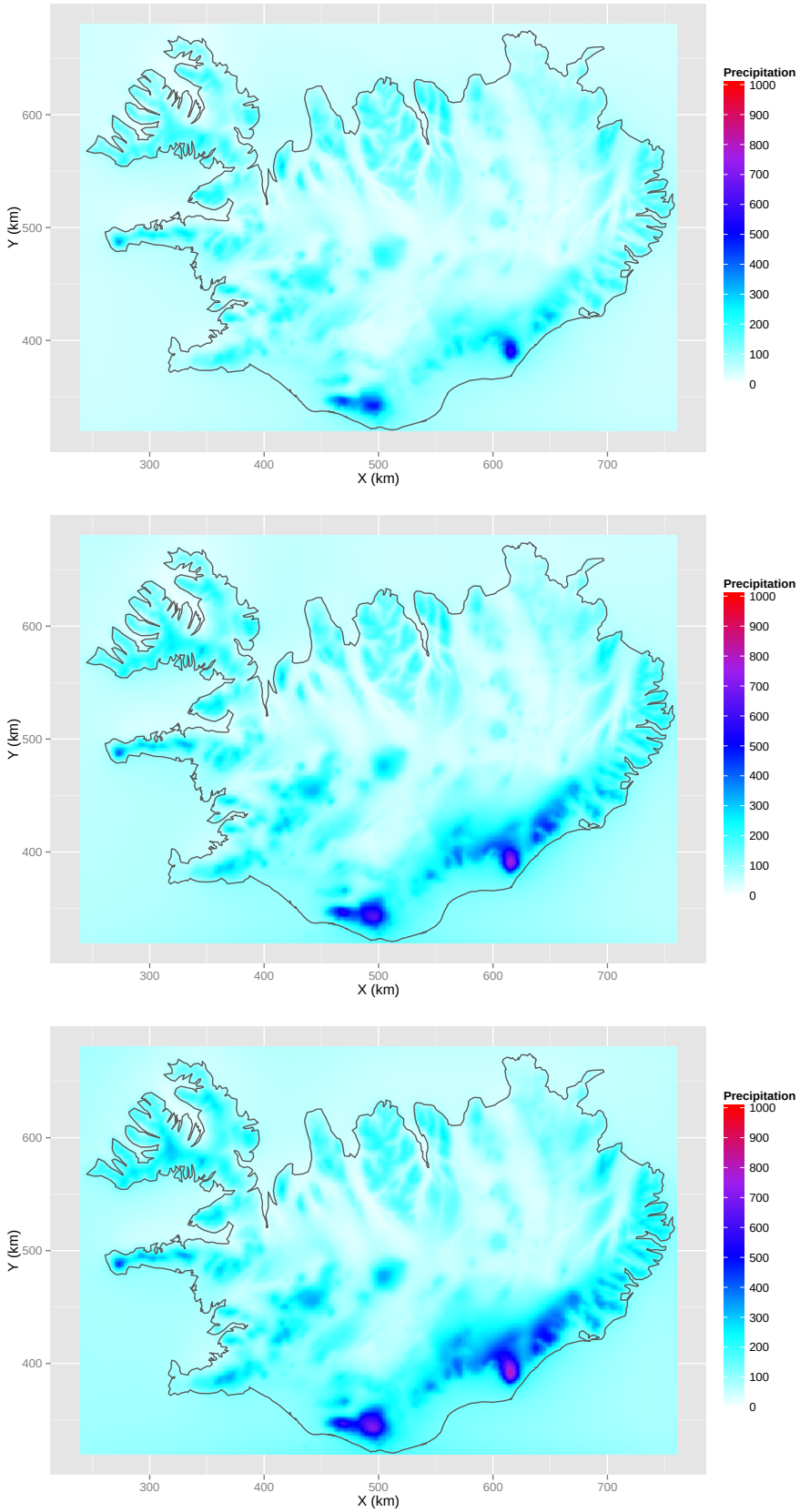

Figure 2.15: Meteorological covariate for the spatial mean on the 1-km grid over Iceland for July-September.

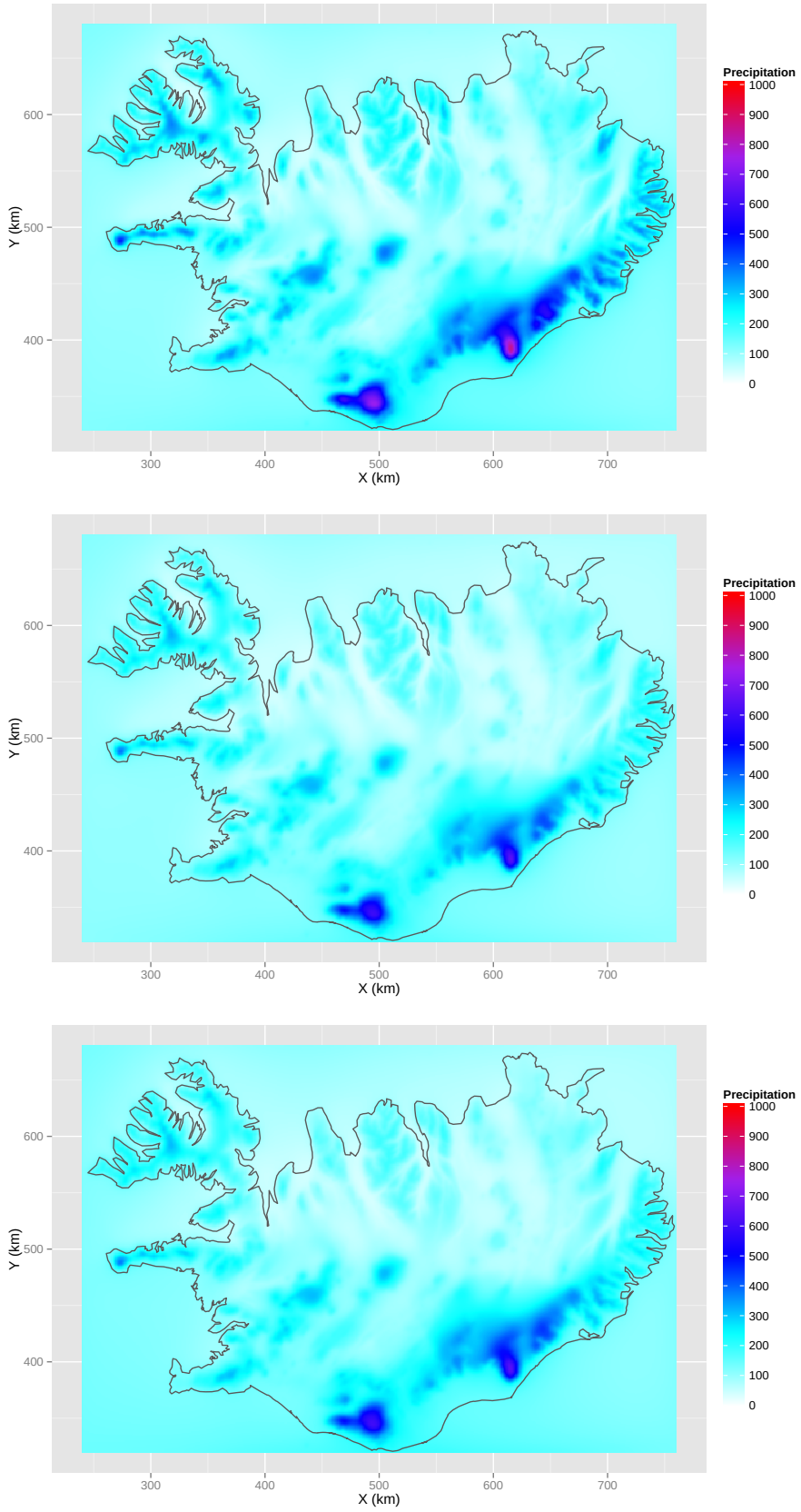

Figure 2.16: Meteorological covariate for the spatial mean on the 1-km grid over Iceland for October-December.

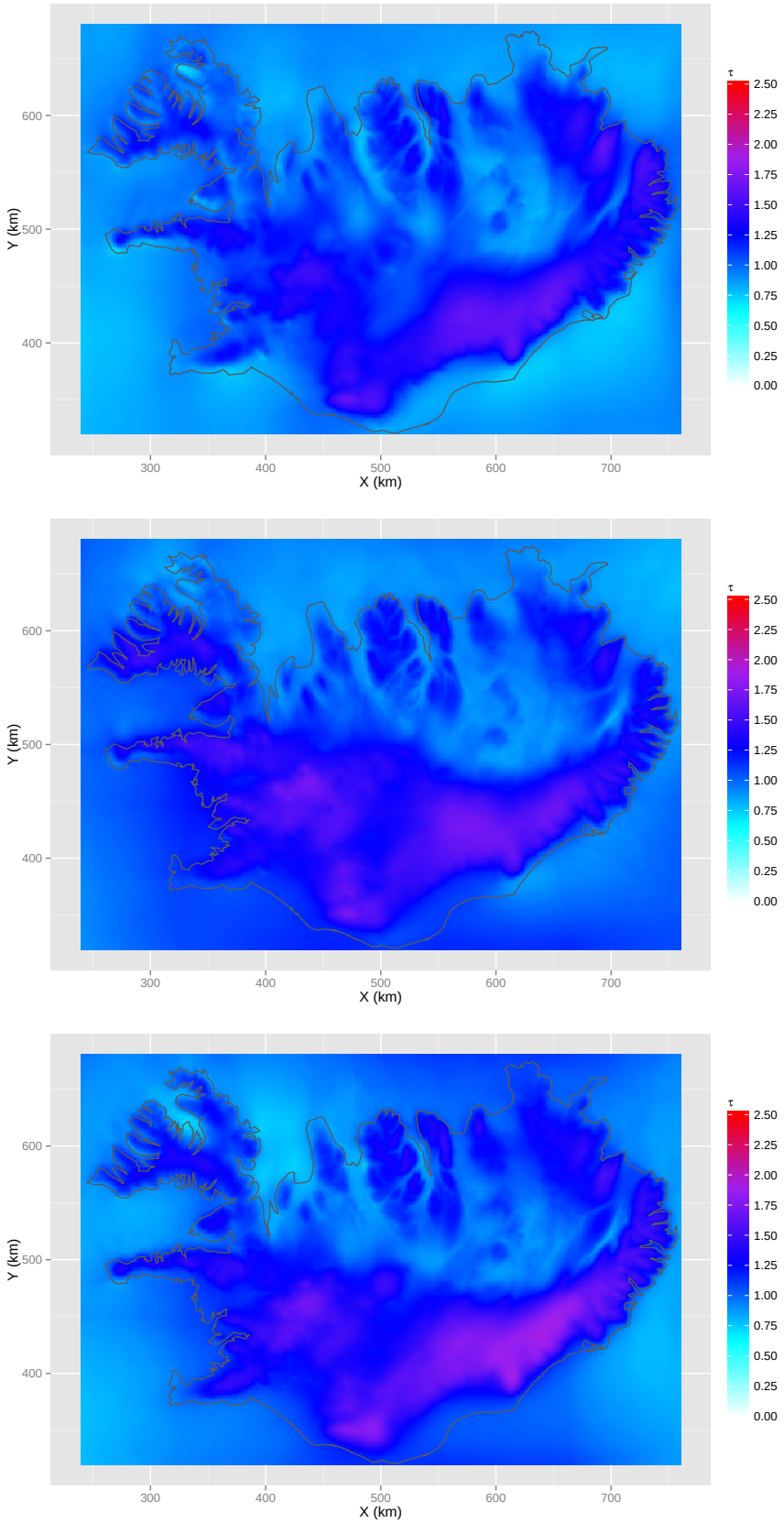

Figure 2.17: Meteorological covariate for the spatial variance on the 1-km grid over Iceland for January-March.

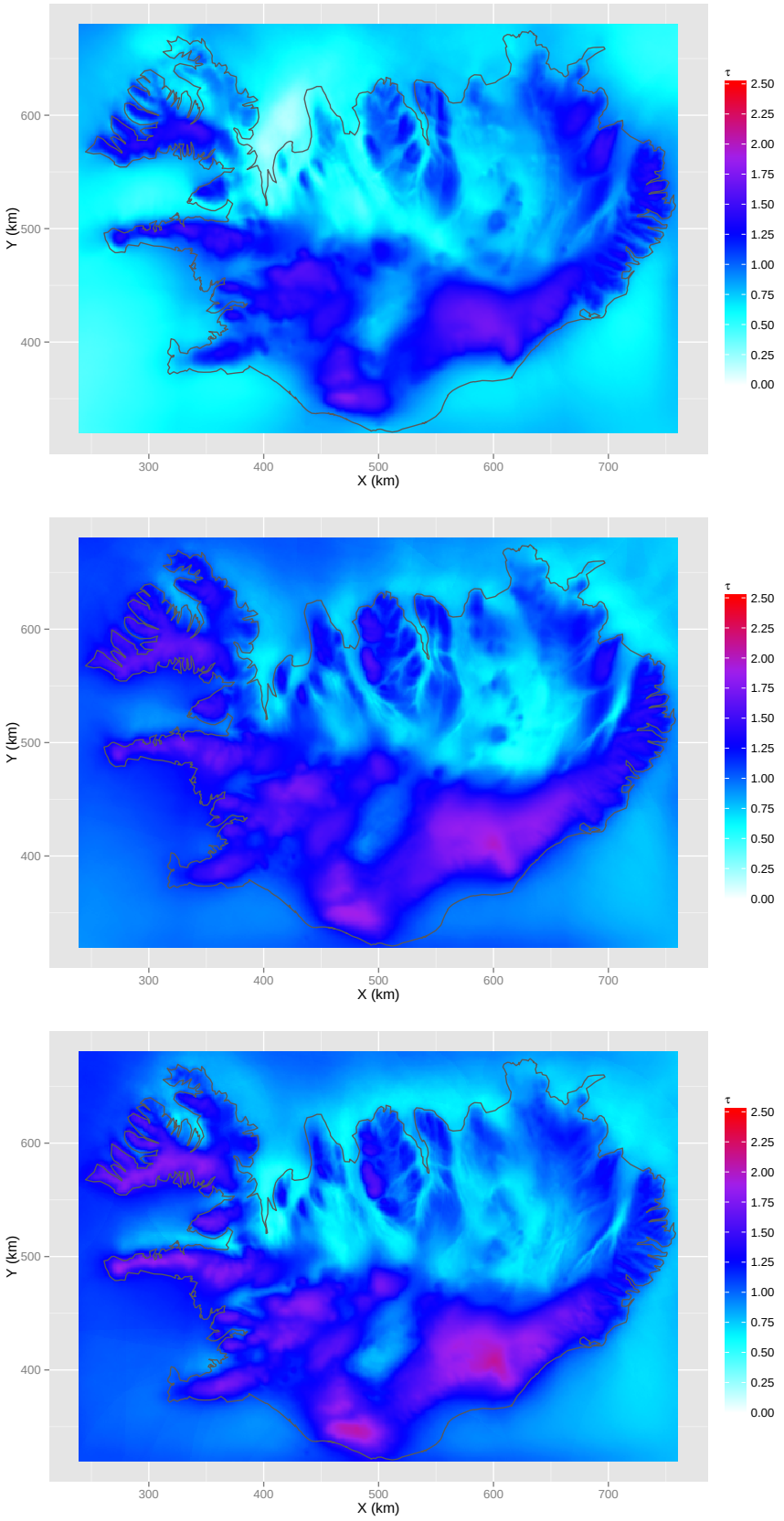

Figure 2.18: Meteorological covariate for the spatial variance on the 1-km grid over Iceland for April-June.

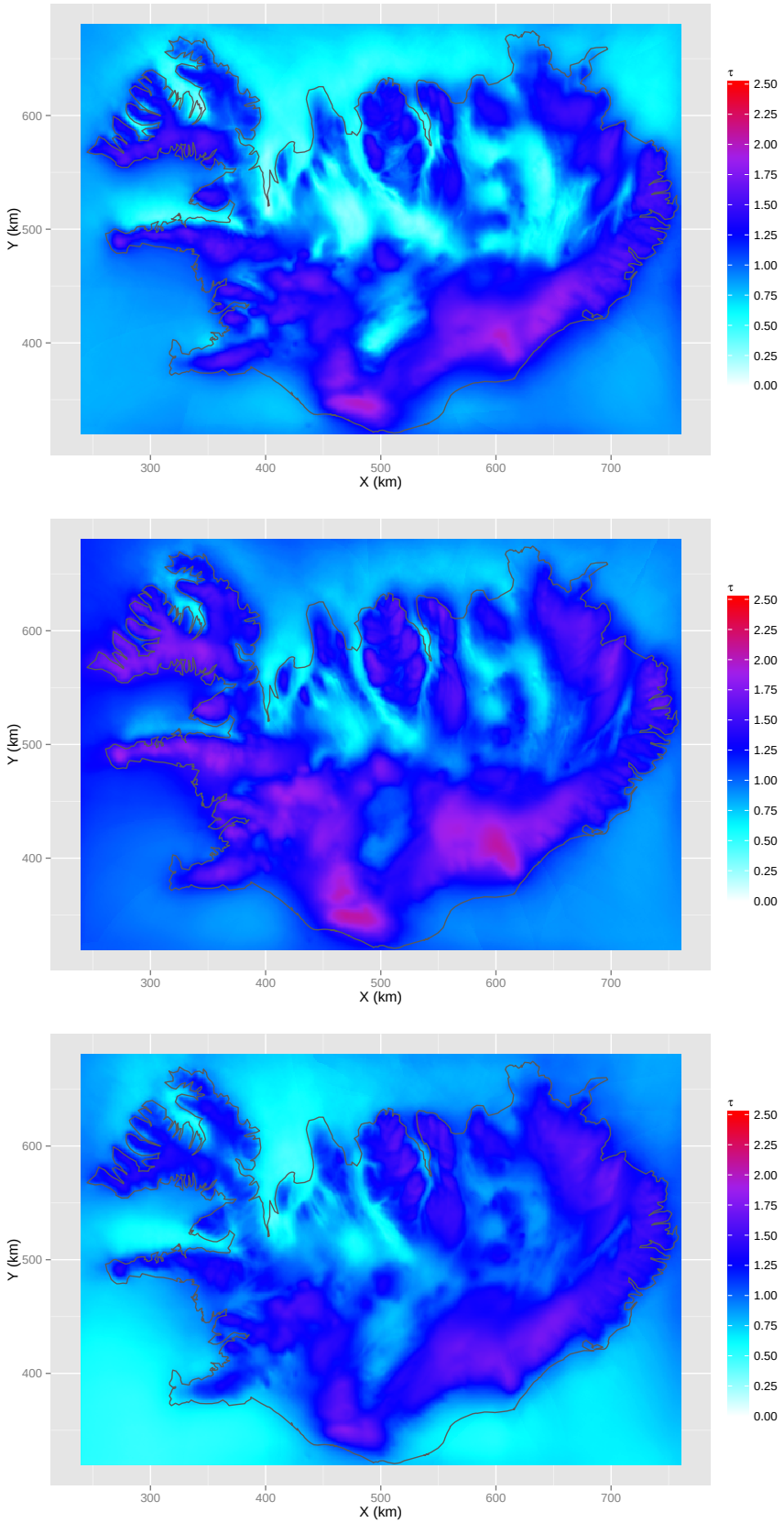

Figure 2.19: Meteorological covariate for the spatial variance on the 1-km grid over Iceland for July-September.

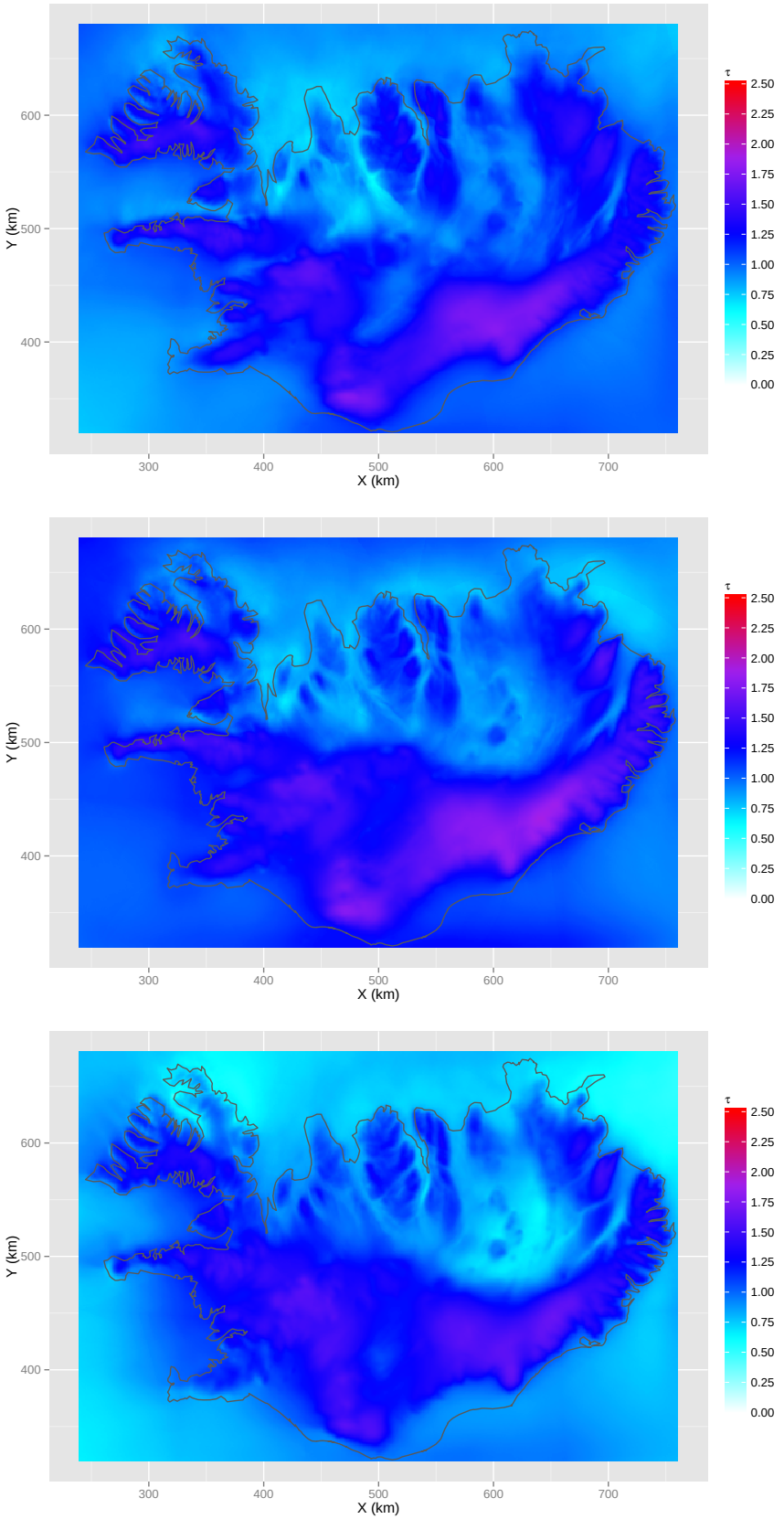

Figure 2.20: Meteorological covariate for the spatial variance on the 1-km grid over Iceland for October-December.

3 Model

3.1 Motivation

Examination of the raw data and some consideration motivated a large part of the model that was proposed. The spatial structure of the data called for a three level Bayesian hierarchical model with data level, latent level and hyperparameter level.

First, by looking at Figure 2.3 it looks like sites located close to one another have a similar precipitation distribution. It is also clear from Figure 2.4 that the data is spatially correlated. One is therefore inclined to use a spatial correlation structure at the data level. Due to how scattered the correlation is as a function of distance there seems to be a need to include a measurement error as well. Both, a spatial correlation structure and measurement error were incorporated at the data level in the model.

It is apparent that each rain gauge site should have its own mean and variance at the latent level. Also since there seems to be a spatial correlation in the data, such a correlation structure should be incorporated for the spatial mean and variance at the latent level, as well as on the data level. To try to explain the spatial difference between stations it was decided to use the centered xy-coordinates, from the ISN93 grid over Iceland, and height above sea level (in 100 m.) as covariates for the spatial mean, along with the meteorological covariate discussed in Section 2.3. These covariates were only used for the spatial mean at the latent level. They were not used for the spatial variance as less is known about how it might be affected by these covariates and also to get a better comparison between the data and the meteorological covariate discussed in Section 2.3.

Recall from Section 2.1, that it was decided to analyze each month separately and to ignore any temporal correlation. Still the model should be able to account for the fact that some years are dryer, or wetter, than others. Therefore, it was decided that although there would be no temporal correlation, each month would have its own time main effect. According to the aforementioned lack of correlation these parameters should be independent and since there was no apparent temporal trend in the data these parameters should not contain a trend.

With all the above considered, the model in the next section was proposed.

3.2 A Bayesian hierarchical model

After preliminary data manipulation, the data for each month was contained in a 49x40 matrix (49 year and 40 sites). The matrix was then transformed with the Box-Cox transformation with parameter $\lambda = 0.4$ (2.1). After the transformation the matrix was turned into a data vector where the first 40 entries were the data from all the 40 stations for year no. 1, the next 40 entries were the data from all the stations for year no. 2. etc. In what follows this data vector will be denoted by $\vec{\mathbf{y}}$.

The following Bayesian hierarchical model was applied to analyze the Box-Cox transformed monthly accumulated precipitation.

$$
\begin{aligned}
\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa &\sim N(Z_1\vec{\boldsymbol{\alpha}} + Z_2\vec{\boldsymbol{\gamma}},\, Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}})) \\
\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi &\sim N(X\vec{\boldsymbol{\beta}},\, \sigma_\alpha^2 R_\nu(\phi)) \\
\vec{\boldsymbol{\gamma}} \,|\, \sigma_\gamma^2 &\sim N(0,\, \sigma_\gamma^2 I) \\
\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho &\sim N(V\eta,\, \sigma_\tau^2 R_\nu(\rho)) \\
\psi \,|\, \mu_\psi, \sigma_\psi^2 &\sim \text{Log-}N(\mu_\psi, \sigma_\psi^2) \\
\kappa \,|\, a_\kappa, b_\kappa &\sim \text{Unf}(a_\kappa, b_\kappa) \\
\vec{\boldsymbol{\beta}} \,|\, \mu_\beta, \Sigma_\beta &\sim N(\mu_\beta,\, \Sigma_\beta) \\
\sigma_\alpha^2 \,|\, \nu_\alpha, s_\alpha^2 &\sim \text{Inv-}\chi^2(\nu_\alpha, s_\alpha^2) \\
\phi \,|\, \mu_\phi, \sigma_\phi^2 &\sim \text{Log-}N(\mu_\phi, \sigma_\phi^2) \\
\sigma_\gamma^2 \,|\, \nu_\gamma, s_\gamma^2 &\sim \text{Inv-}\chi^2(\nu_\gamma, s_\gamma^2) \\
\eta \,|\, \mu_\eta, \sigma_\eta^2 &\sim N(\mu_\eta,\, \sigma_\eta^2) \\
\sigma_\tau^2 \,|\, \nu_\tau, s_\tau^2 &\sim \text{Inv-}\chi^2(\nu_\tau, s_\tau^2) \\
\rho \,|\, \mu_\rho, \sigma_\rho^2 &\sim \text{Log-}N(\mu_\rho, \sigma_\rho^2).
\end{aligned}
$$

The spatial dependence of the transformed rain gauge data is modeled with a Gaussian field with Matérn correlation, $R_\nu(\psi)$, with smoothness parameter $\nu = 1.5$ and range parameter ψ (see e.g. [14]). The measurement error is a scalar multiple of the point estimate of each sites' median times the model variance $e^{2\tau_i}$, κ is the scalar and the matrix E is 1960x1960 dimensional and contains the 40 point estimates, repeated 49 times, on the diagonal, if we assume that no measurements are missing. The matrix Z_1 and Z_2 are index matrices. If we again assume that no measurements are missing then Z_1 is a 1960x40 dimensional, Z_2 is 1960x49 dimensional and Z_1

and Z_2 have the following forms.

$$Z_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \qquad Z_2 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Still under the assumption that no measurements are missing, the matrix Q is a 1960x1960 dimensional diagonal matrix. It contains the 40 dimensional vector $\exp(\vec{\boldsymbol{\tau}}) = \vec{\boldsymbol{\sigma}}$ on the diagonal, repeated 49 times, where $\vec{\boldsymbol{\sigma}} = (\sigma_1, \dots, \sigma_{40})$ is the spatial standard deviation of the transformed rain gauge data.

The latent field $\vec{\boldsymbol{\alpha}}$ is a spatial mean and the latent field $\vec{\boldsymbol{\tau}}$ is the natural logarithm of the square root of the spatial variance. The latent process $\vec{\boldsymbol{\gamma}}$ is the time main effect.

The latent field $\vec{\boldsymbol{\alpha}}$ is modeled as a Gaussian random field with Matérn correlation with smoothness parameter $\nu = 1.5$ and range parameter ϕ. The 40x4 dimensional matrix X is a matrix of covariates. $X_{\cdot 1}$ is the meteorological covariate for the spatial mean, discussed in Section 2.3, $X_{\cdot 2}$ is the centered x-coordinate from the ISN93 grid, $X_{\cdot 3}$ is the centered y-coordinate from the ISN93 grid and $X_{\cdot 4}$ is height above sea level in 100 m.

The latent field $\vec{\boldsymbol{\tau}}$ is modeled as a Gaussian random field with Matérn correlation with smoothness parameter $\nu = 1.5$ and range parameter ρ. The 40x1 dimensional matrix V contains the natural logarithm of the square root of the meteorological covariate for the spatial variance, see Section 2.3.

The latent process $\vec{\boldsymbol{\gamma}}$ is modeled as an independent Gaussian process with mean zero and variance σ_γ^2.

The hyperparameters for the latent fields $\vec{\boldsymbol{\alpha}}$ and $\vec{\boldsymbol{\tau}}$ and the latent process $\vec{\boldsymbol{\gamma}}$ are given prior distribution as listed above. Other parameters in the model are set as suitable constants.

3.3 Posterior distributions

Let $\vec{\boldsymbol{\theta}} = (\vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \sigma_\eta^2, \psi, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2)$. Here Markov Chain Monte Carlo (MCMC) based on the Gibbs sampler (see e.g. [15]) is used to sample from the posterior distribution of $\vec{\boldsymbol{\theta}}$.

The prior distribution of $\vec{\boldsymbol{\theta}}$ is

$$\begin{aligned}
p(\vec{\boldsymbol{\theta}}) &= p(\vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \\
&= p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \cdot p(\vec{\boldsymbol{\gamma}} \,|\, \vec{\boldsymbol{\tau}}, \psi, \kappa, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \\
&\quad \cdot p(\vec{\boldsymbol{\tau}} \,|\, \psi, \kappa, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \cdot p(\psi \,|\, \kappa, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \\
&\quad \cdot p(\kappa \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \cdot p(\vec{\boldsymbol{\beta}} \,|\, \sigma_\alpha^2, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \cdot p(\sigma_\alpha^2 \,|\, \phi, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \\
&\quad \cdot p(\phi \,|\, \sigma_\gamma^2, \eta, \sigma_\tau^2, \rho) \cdot p(\sigma_\gamma^2 \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\eta \,|\, \sigma_\tau^2, \rho) \cdot p(\sigma_\tau^2 \,|\, \rho) \cdot p(\rho) \\
&= p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\vec{\boldsymbol{\gamma}} \,|\, \sigma_\gamma^2) \cdot p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\psi) \cdot p(\kappa) \cdot p(\vec{\boldsymbol{\beta}}) \cdot p(\sigma_\alpha^2) \cdot p(\phi) \\
&\quad \cdot p(\sigma_\gamma^2) \cdot p(\eta) \cdot p(\sigma_\tau^2) \cdot p(\rho).
\end{aligned}$$

The posterior distribution of $\vec{\boldsymbol{\theta}}$ is

$$p(\vec{\boldsymbol{\theta}} \,|\, \vec{\mathbf{y}}) \propto p(\vec{\boldsymbol{\theta}}) p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\theta}})$$

where

$$p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\theta}}) = p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) = N(Z_1 \vec{\boldsymbol{\alpha}} + Z_2 \vec{\boldsymbol{\gamma}},\, Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E) Q(\vec{\boldsymbol{\tau}})).$$

Therefore the posterior distribution of $\vec{\boldsymbol{\theta}}$ is given by

$$\begin{aligned}
p(\vec{\boldsymbol{\theta}} \,|\, \vec{\mathbf{y}}) &\propto p(\vec{\boldsymbol{\theta}}) p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\theta}}) \\
&= p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\vec{\boldsymbol{\gamma}} \,|\, \sigma_\gamma^2) \cdot p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\psi) \cdot p(\kappa) \\
&\quad \cdot p(\vec{\boldsymbol{\beta}}) \cdot p(\sigma_\alpha^2) \cdot p(\phi) \cdot p(\sigma_\gamma^2) \cdot p(\eta) \cdot p(\sigma_\tau^2) \cdot p(\rho) \\
&= N(\vec{\mathbf{y}} \,|\, Z_1 \vec{\boldsymbol{\alpha}} + Z_2 \vec{\boldsymbol{\gamma}},\, Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E) Q(\vec{\boldsymbol{\tau}})) \cdot N(\vec{\boldsymbol{\alpha}} \,|\, X\vec{\boldsymbol{\beta}},\, \sigma_\alpha^2 R_\nu(\phi)) \\
&\quad \cdot N(\vec{\boldsymbol{\gamma}} \,|\, 0,\, \sigma_\gamma^2 I) \cdot N(\vec{\boldsymbol{\tau}} \,|\, V\eta,\, \sigma_\tau^2 R_\nu(\rho)) \cdot \text{Log-}N(\psi \,|\, \mu_\psi, \sigma_\psi^2) \\
&\quad \cdot \text{Unf}(\kappa \,|\, a_\kappa, b_\kappa) \cdot N(\vec{\boldsymbol{\beta}} \,|\, \mu_\beta,\, \Sigma_\beta) \cdot \text{Inv-}\chi^2(\sigma_\alpha^2 \,|\, \nu_\alpha, s_\alpha^2) \cdot \text{Log-}N(\phi \,|\, \mu_\phi, \sigma_\phi^2) \\
&\quad \cdot \text{Inv-}\chi^2(\sigma_\gamma^2 \,|\, \nu_\gamma, s_\gamma^2) \cdot N(\eta \,|\, \mu_\eta,\, \sigma_\eta^2) \cdot \text{Inv-}\chi^2(\sigma_\tau^2 \,|\, \nu_\tau, s_\tau^2) \cdot \text{Log-}N(\rho \,|\, \mu_\rho, \sigma_\rho^2).
\end{aligned}$$

The Gibbs sampler consists of the posterior conditional distributions of $\vec{\boldsymbol{\alpha}}$, $\vec{\boldsymbol{\gamma}}$, $\vec{\boldsymbol{\tau}}$, ψ, κ, $\vec{\boldsymbol{\beta}}$, σ_α^2, ϕ, σ_γ^2, η, σ_τ^2 and ρ.

These conditional distributions are

$$
\begin{aligned}
p(\vec{\boldsymbol{\alpha}} \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \\
p(\vec{\boldsymbol{\gamma}} \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\gamma}} \,|\, \sigma_\gamma^2) \\
p(\vec{\boldsymbol{\tau}} \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \\
p(\psi \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\psi) \\
p(\kappa \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\kappa) \\
p(\vec{\boldsymbol{\beta}} \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\vec{\boldsymbol{\beta}}) \\
p(\sigma_\alpha^2 \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\sigma_\alpha^2) \\
p(\phi \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\phi) \\
p(\sigma_\gamma^2 \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\gamma}} \,|\, \sigma_\gamma^2) \cdot p(\sigma_\gamma^2) \\
p(\eta \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\eta) \\
p(\sigma_\tau^2 \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\sigma_\tau^2) \\
p(\rho \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\rho).
\end{aligned}
$$

A detailed description of the conditional distributions is given below.

In what follows let

$$
\begin{aligned}
\Sigma_y &= Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}}), \\
\Sigma_\alpha &= \sigma_\alpha^2 R_\nu(\phi) \quad \text{and} \\
\Sigma_\tau &= \sigma_\tau^2 R_\nu(\rho).
\end{aligned}
$$

The conditional distribution of $\vec{\boldsymbol{\alpha}}$ is such that

$$
\begin{aligned}
p(\vec{\boldsymbol{\alpha}} \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \\
&= N(\vec{\mathbf{y}} \,|\, Z_1\vec{\boldsymbol{\alpha}} + Z_2\vec{\boldsymbol{\gamma}},\, \Sigma_y) \cdot N(\vec{\boldsymbol{\alpha}} \,|\, X\vec{\boldsymbol{\beta}}, \Sigma_\alpha) \\
&\propto \exp\left(-\frac{1}{2}((\vec{\mathbf{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}})^T\Sigma_y^{-1}(\vec{\mathbf{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}}))\right) \\
&\quad \cdot \exp\left(-\frac{1}{2}((\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})^T\Sigma_\alpha^{-1}(\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}}))\right) \\
&\propto \exp\Bigg(-\frac{1}{2}(\vec{\boldsymbol{\alpha}}^T(Z_1^T\Sigma_y^{-1}Z_1 + \Sigma_\alpha^{-1})\vec{\boldsymbol{\alpha}} \\
&\quad - \vec{\boldsymbol{\alpha}}^T(Z_1^T\Sigma_y^{-1}\vec{\mathbf{y}} - Z_1^T\Sigma_y^{-1}Z_2\vec{\boldsymbol{\gamma}} + \Sigma_\alpha^{-1}X\vec{\boldsymbol{\beta}}))\Bigg).
\end{aligned}
$$

It follows that

$$
p(\vec{\boldsymbol{\alpha}} \,|\, \text{rest}) = N(\vec{\boldsymbol{\alpha}} \,|\, \mu_{\alpha,\text{post}}, \Sigma_{\alpha,\text{post}})
$$

where

$$\mu_{\alpha,\text{post}} = (Z_1^T\Sigma_y^{-1}Z_1 + \Sigma_\alpha^{-1})^{-1}(Z_1^T\Sigma_y^{-1}\vec{\mathbf{y}} - Z_1^T\Sigma_y^{-1}Z_2\vec{\boldsymbol{\gamma}} + \Sigma_\alpha^{-1}X\vec{\boldsymbol{\beta}}),$$
$$\Sigma_{\alpha,\text{post}} = (Z_1^T\Sigma_y^{-1}Z_1 + \Sigma_\alpha^{-1})^{-1}.$$

Similarly it can be seen that

$$p(\vec{\boldsymbol{\gamma}} \,|\, \text{rest}) = N(\vec{\boldsymbol{\gamma}} \,|\, \mu_{\gamma,\text{post}}, \Sigma_{\gamma,\text{post}}),$$

where

$$\mu_{\gamma,\text{post}} = (Z_2^T\Sigma_y^{-1}Z_2 + \sigma_\gamma^{-2}I)^{-1}(Z_2^T\Sigma_y^{-1}\vec{\mathbf{y}} - Z_2^T\Sigma_y^{-1}Z_1\vec{\boldsymbol{\alpha}}),$$
$$\Sigma_{\gamma,\text{post}} = (Z_2^T\Sigma_y^{-1}Z_2 + \sigma_\gamma^{-2}I)^{-1}.$$

It is also apparent by similar reasoning that

$$p(\vec{\boldsymbol{\beta}} \,|\, \text{rest}) = N(\vec{\boldsymbol{\beta}} \,|\, \mu_{\beta,\text{post}}, \Sigma_{\beta,\text{post}}),$$

where

$$\mu_{\beta,\text{post}} = (X^T\Sigma_\alpha^{-1}X + \Sigma_\beta^{-1})^{-1}(X^T\Sigma_\alpha^{-1}\vec{\boldsymbol{\alpha}} + \Sigma_\beta^{-1}\mu_\beta),$$
$$\Sigma_{\beta,\text{post}} = (X^T\Sigma_\alpha^{-1}X + \Sigma_\beta^{-1})^{-1},$$

and that

$$p(\eta \,|\, \text{rest}) = N(\eta \,|\, \mu_{\eta,\text{post}}, \sigma^2_{\eta,\text{post}}),$$

where

$$\mu_{\eta,\text{post}} = (V^T\Sigma_\tau^{-1}V + \sigma_\eta^{-2})^{-1}(V^T\Sigma_\tau^{-1}\vec{\boldsymbol{\tau}} + \sigma_\eta^{-2}\mu_\eta),$$
$$\sigma^2_{\eta,\text{post}} = (V^T\Sigma_\tau^{-1}V + \sigma_\eta^{-2})^{-1}.$$

We also have that

$$\begin{aligned}
p(\sigma_\alpha^2 \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\sigma_\alpha^2) \\
&= N(\vec{\boldsymbol{\alpha}} \,|\, X\vec{\boldsymbol{\beta}}, \sigma_\alpha^2 R_\nu(\phi)) \cdot \text{Inv-}\chi^2(\sigma_\alpha^2 \,|\, \nu_\alpha, s_\alpha^2) \\
&\propto (\sigma_\alpha^2)^{-n_\alpha/2} \exp\left(-\frac{1}{2}((\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})^T(\sigma_\alpha^2 R_\nu(\phi))^{-1}(\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}}))\right) \\
&\quad \cdot (\sigma_\alpha^2)^{-(\nu_\alpha/2+1)} \exp\left(-\frac{\nu_\alpha s_\alpha^2}{2\sigma_\alpha^2}\right) \\
&= (\sigma_\alpha^2)^{-((\nu_\alpha+n_\alpha)/2+1)} \exp\left(-\frac{1}{2\sigma_\alpha^2}(\nu_\alpha s_\alpha^2 + (\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})^T R_\nu(\phi)^{-1}(\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}}))\right).
\end{aligned}$$

It can therefore be seen that

$$p(\sigma_\alpha^2 \,|\, \text{rest}) = \text{Inv-}\chi^2(\sigma_\alpha^2 \,|\, \nu_{\alpha,\text{post}}, s^2_{\alpha,\text{post}}),$$

where

$$\begin{aligned}
\nu_{\alpha,\text{post}} &= \nu_\alpha + n_\alpha, \\
s^2_{\alpha,\text{post}} &= \frac{\nu_\alpha s_\alpha^2 + (\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})^T R_\nu(\phi)^{-1}(\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})}{\nu_\alpha + n_\alpha} \quad \text{and} \\
n_\alpha &= 40.
\end{aligned}$$

Similarly we have that

$$\begin{aligned}
p(\sigma_\tau^2 \,|\, \text{rest}) &= \text{Inv-}\chi^2(\sigma_\tau^2 \,|\, \nu_{\tau,\text{post}}, s^2_{\tau,\text{post}}) \quad \text{and} \\
p(\sigma_\gamma^2 \,|\, \text{rest}) &= \text{Inv-}\chi^2(\sigma_\gamma^2 \,|\, \nu_{\gamma,\text{post}}, s^2_{\gamma,\text{post}}),
\end{aligned}$$

where

$$\begin{aligned}
\nu_{\tau,\text{post}} &= \nu_\tau + n_\tau, \\
s^2_{\tau,\text{post}} &= \frac{\nu_\tau s_\tau^2 + (\vec{\boldsymbol{\tau}} - V\eta)^T R_\nu(\rho)^{-1}(\vec{\boldsymbol{\tau}} - V\eta)}{\nu_\tau + n_\tau}, \\
n_\tau &= 40, \\
\nu_{\gamma,\text{post}} &= \nu_\gamma + n_\gamma, \\
s^2_{\gamma,\text{post}} &= \frac{\nu_\gamma s_\gamma^2 + \vec{\boldsymbol{\gamma}}^T\vec{\boldsymbol{\gamma}}}{\nu_\gamma + n_\gamma} \quad \text{and} \\
n_\gamma &= 49.
\end{aligned}$$

As can be seen above, the parameters already mentioned all have a conjugate prior distribution and therefore their conditional posterior distributions all have known form and can easily be sampled from. That is not the case with the parameters which follow. Their posterior conditional distribution are not of any known distributional form. Therefore, to sample from their distribution the Metropolis Hastings algorithm (see e.g. [15]) is used or other methods explained later in this text.

The conditional distribution of $\vec{\boldsymbol{\tau}}$ is such that

$$\begin{aligned}
p(\vec{\boldsymbol{\tau}} \,|\, \text{rest}) &\propto p(\vec{\mathbf{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\vec{\boldsymbol{\tau}} \,|\, \beta, \sigma_\tau^2, \rho) \\
&= N(\vec{\mathbf{y}} \,|\, Z_1\vec{\boldsymbol{\alpha}} + Z_2\vec{\boldsymbol{\gamma}},\, Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}})) \cdot N(\vec{\boldsymbol{\tau}} \,|\, V\eta, \Sigma_\tau) \\
&\propto |Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}})|^{-\frac{1}{2}} \\
&\quad \cdot \exp\left(-\frac{1}{2}((\vec{\mathbf{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}})^T(Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}}))^{-1}(\vec{\mathbf{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}}))\right) \\
&\quad \exp\left(-\frac{1}{2}((\vec{\boldsymbol{\tau}} - V\eta)^T\Sigma_\tau^{-1}(\vec{\boldsymbol{\tau}} - V\eta))\right).
\end{aligned} \tag{3.1}$$

The conditional distribution of (ψ, κ) is given by

$$\begin{aligned} p(\psi, \kappa \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{y}} \,|\, \vec{\boldsymbol{\alpha}}, \vec{\boldsymbol{\gamma}}, \vec{\boldsymbol{\tau}}, \psi, \kappa) \cdot p(\psi) \cdot p(\kappa) \\ &= N(\vec{\boldsymbol{y}} \,|\, Z_1\vec{\boldsymbol{\alpha}} + Z_2\vec{\boldsymbol{\gamma}}, \Sigma_y) \cdot \text{Log-}N(\psi \,|\, \mu_\psi, \sigma_\psi^2) \cdot \text{Unf}(\kappa \,|\, a_\kappa, b_\kappa) \\ &\propto |Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}})|^{-\frac{1}{2}} \\ &\quad \cdot \exp\left(-\frac{1}{2}((\vec{\boldsymbol{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}})^T(Q(\vec{\boldsymbol{\tau}})(R_\nu(\psi) + \kappa E)Q(\vec{\boldsymbol{\tau}}))^{-1}(\vec{\boldsymbol{y}} - Z_1\vec{\boldsymbol{\alpha}} - Z_2\vec{\boldsymbol{\gamma}}))\right) \\ &\quad \cdot \psi^{-1} \exp\left(-\frac{1}{2\sigma_\psi^2}(\log(\psi) - \mu_\psi)^2\right). \end{aligned} \tag{3.2}$$

Finally

$$\begin{aligned} p(\phi \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\alpha}} \,|\, \vec{\boldsymbol{\beta}}, \sigma_\alpha^2, \phi) \cdot p(\phi) \\ &= N(\vec{\boldsymbol{\alpha}} \,|\, X\vec{\boldsymbol{\beta}}, \sigma_\alpha^2 R_\nu(\phi)) \cdot \text{Log-}N(\phi \,|\, \mu_\phi, \sigma_\phi^2) \\ &\propto |\sigma_\alpha^2 R_\nu(\phi)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}((\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}})^T(\sigma_\alpha^2 R_\nu(\phi))^{-1}(\vec{\boldsymbol{\alpha}} - X\vec{\boldsymbol{\beta}}))\right) \\ &\quad \cdot \phi^{-1} \exp\left(-\frac{1}{2\sigma_\phi^2}(\log(\phi) - \mu_\phi)^2\right) \end{aligned} \tag{3.3}$$

and

$$\begin{aligned} p(\rho \,|\, \text{rest}) &\propto p(\vec{\boldsymbol{\tau}} \,|\, \eta, \sigma_\tau^2, \rho) \cdot p(\rho) \\ &= N(\vec{\boldsymbol{\tau}} \,|\, V\eta, \sigma_\tau^2 R_\nu(\rho)) \cdot \text{Log-}N(\rho \,|\, \mu_\rho, \sigma_\rho^2) \\ &\propto |\sigma_\tau^2 R_\nu(\rho)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}((\vec{\boldsymbol{\tau}} - V\eta)^T(\sigma_\tau^2 R_\nu(\rho))^{-1}(\vec{\boldsymbol{\tau}} - V\eta))\right) \\ &\quad \cdot \rho^{-1} \exp\left(-\frac{1}{2\sigma_\rho^2}(\log(\rho) - \mu_\rho)^2\right) \end{aligned} \tag{3.4}$$

3.4 Execution

As mentioned in the previous section the Gibbs sampler was used to draw samples from the joint posterior distribution of $\vec{\boldsymbol{\theta}}$. However, due to problems with dimensionality and correlation between parameters, not all the unknown parameters could be included in the sampler. To start with, it was not feasible to include the parameters ψ and κ because that would have made it necessary to update and invert the 1960x1960 (ignoring missing data) dimensional matrix $(R_\nu(\psi) + \kappa E)$. Even though this is a block diagonal matrix of 49 blocks of size 40x40, each step would have been

too time consuming for the sampler to be practical . Instead these parameters were estimated by minimizing the negative of the natural logarithm of their posterior conditional distribution (3.2) with other parameters set equal to estimates which are explained below.

The parameter ρ was introduced relatively late in the development of the model. Since it was already needed to use the Metropolis Hastings algorithm for the parameter $\vec{\boldsymbol{\tau}}$ and since ρ turned out to be highly correlated with τ they would have needed to be sampled together by using a block updating scheme. Instead of doing that it was decided to estimate this parameter similarly to what was done with ψ and κ. Therefore, ρ was estimated by minimizing the negative of the natural logarithm of it's posterior conditional distribution (3.4) with other parameters set equal to estimates which are explained below.

Now, the parameter $\vec{\boldsymbol{\tau}}$ was a bit troublesome to deal with. It is multidimensional and needed the Metropolis Hastings algorithm to be sampled from. To start with the proposed value of $\vec{\boldsymbol{\tau}}$, $\vec{\boldsymbol{\tau}}^*$, was sampled from independent normal distributions with the previous value of $\vec{\boldsymbol{\tau}}$ as the mean. This resulted in a poor proposed value, firstly because the acceptance ratio was difficult to tweak and secondly, the sample produced was highly correlated and therefore converged quite slowly towards the target distribution. To get a better proposed value of $\vec{\boldsymbol{\tau}}$ we estimated the covariance matrix of the posterior distribution of $\vec{\boldsymbol{\tau}}$, Σ_τ, and used a multivariate normal distribution with the previous value of $\vec{\boldsymbol{\tau}}$ as the mean and the estimated covariance matrix. This improved the sampling scheme and the resulting chain for $\vec{\boldsymbol{\tau}}$ converged reasonably fast. The estimate of the covariance matrix was obtained by minimizing the negative of the natural logarithm of the posterior conditional distribution of $\vec{\boldsymbol{\tau}}$, finding the Hessian matrix at this minimum and then inverting it. This was done with other parameters set equal to estimates which are explained below.

All parameters, other than ψ, κ and ρ, were included in the Gibbs sampler. In each run, four chains were sampled, all with 10,000 iterations with 3,000 as burn-in. The parameters $\vec{\boldsymbol{\alpha}}$, $\vec{\boldsymbol{\gamma}}$ and $\vec{\boldsymbol{\tau}}$ were sampled using vector updating. Two Metropolis Hastings steps were needed. $\vec{\boldsymbol{\tau}}$ was sampled by using a proposal from a multivariate normal distribution with the aforementioned estimated covariance matrix. The parameter ϕ was sampled simply by using a proposal from a normal distribution with the previous value of ϕ as the mean and the variance tweaked to get a reasonable acceptance rate.

Since the parameters ψ, κ, and ρ needed to be estimated outside of the Gibbs sampler it was decided to run the sampler a few times and update their estimates after each run. After some consideration it was decided to use the following scheme to get reasonable estimates of all the parameters.

Sampling scheme

1. Estimate the parameters.
   First ψ and κ were estimated. For that, estimates for the parameters $\vec{\boldsymbol{\alpha}}$, $\vec{\boldsymbol{\gamma}}$ and $\vec{\boldsymbol{\tau}}$ were needed. $\vec{\boldsymbol{\alpha}}$ was estimated with the sample median for each station, $\vec{\boldsymbol{\gamma}}$ was estimated with the sample median for each year minus the sample median of all the data and $\vec{\boldsymbol{\tau}}$ was estimated with the natural logarithm of the sample standard deviation for each station.
   Next ρ was estimated. For that, estimates for the parameters η and σ_τ^2, as well as $\vec{\boldsymbol{\tau}}$, were needed. An estimate for η was found by linear regression on the estimate of $\vec{\boldsymbol{\tau}}$, previously found, and the matrix V and σ_τ^2 was estimated with the variance of the estimate of $\vec{\boldsymbol{\tau}}$. This first estimate for ρ was quite different for different months, with values ranging from around 0.3 to around 23. It was decided to use the maximum of these values as a first estimate for all the months since a value of around 23 seems more reasonable as the range of the correlation than a smaller value.
   Finally an estimate Σ_τ was obtained by using all the other aforementioned estimates.

2. Run the Gibbs sampler with the values of ψ, κ and ρ fixed as the estimates obtained in the previous step and the estimate of Σ_τ as the covariance matrix of the proposal distribution for $\vec{\boldsymbol{\tau}}$.

3. Update the estimates of ψ, κ, ρ and Σ_τ by using posterior medians, obtained from the Gibbs sampler, as point estimates of other parameters.

4. Run the Gibbs sampler with the updated estimates of ψ, κ, ρ and Σ_τ.

5. Update the estimates of the parameters again.

6. Examine which of the $\vec{\boldsymbol{\beta}}$ parameters, corresponding to the covariates for the spatial mean, contain zero in their 95% posterior intervals and remove the ones which do.
   It turned out that all the $\vec{\boldsymbol{\beta}}$ parameters contained zero in their 95% posterior interval except for β_1, which corresponds to the meteorological covariate discussed in Section 2.3.

7. Run the Gibbs sampler for the last time with the remaining covariate, using the updated estimates of ψ, κ, ρ and Σ_τ.

8. Get the final estimates of the values of the parameters ψ, κ and ρ by using posterior medians, obtained from the final run of the Gibbs sampler, as point estimates of other parameters.

This sampling scheme resulted in convincing samples of the parameters in the Gibbs sampler and reasonable point estimates of the parameters ψ, κ and ρ.

4 Results

4.1 Description

In this section the main results from the Bayesian hierarchical model are presented with a short description. A more involved discussion about the results can be found in Section 4.2.

As was said in Section 3.4, it was examined which covariates could be removed from the spatial mean field, $\vec{\boldsymbol{\alpha}}$, before the last run of the Gibbs sampler. It turned out that all the covariates, except for the meteorological covariate discussed in Section 2.3, could be removed for all the months, with the exception of July, where the parameter corresponding to height above sea level turned out to have most of its posterior probability mass above zero. Despite this, it was decided to remove the three covariates from the model for all the months. Tables 4.2 and 4.3 contain the posterior median, 2.5% quantile and 97.5% quantile for the parameters β_1, β_2, β_3 and β_4, corresponding to the covariates in the spatial mean field, $\vec{\boldsymbol{\alpha}}$, for all the months before being removed from the model. After removing the three covariates, the Gibbs sampler was run for the final time with the meteorological covariate only.

The figures and tables that follow show various statistics and estimates for the parameters of the model after the final run of the Gibbs sampler. Figures 4.1 and 4.2 show the posterior density and the mixture of the 4 chains, based on the Gibbs sampler, for the remaining covariate effect $\beta = \beta_1$ in the spatial mean field. Figures 4.3 and 4.4 show the same for the parameter η, corresponding to the meteorological covariate in the spatial field for $\vec{\boldsymbol{\tau}}$. Table 4.1 contains the final estimates of the parameters ψ, κ and ρ based on point estimates of other parameters obtained from the final run of the Gibbs sampler. Tables 4.4 and 4.5 show posterior statistics for all the univariate parameters, included in the Gibbs sampler, for all the months, after the final run of the sampler. Figure 4.5 shows the posterior median along with the 95% posterior interval for the parameter $\vec{\boldsymbol{\gamma}}$ for each month and Figure 4.6 shows the same for all the months taken together in correct temporal order. Figures 4.7 to 4.18 show the estimated median precipitation field over Iceland along with the corresponding meteorological covariate and the difference between the two (posterior estimate minus covariate). The estimated median precipitation is based

on the posterior median of $\vec{\boldsymbol{\alpha}}$, the covariate at unobserved sites and the kriging method explained in [16]. Both the posterior median of $\vec{\boldsymbol{\alpha}}$ and the covariate at unobserved sites are transformed with the inverse of the Box-Cox transformation (2.1) so that the scale of the figures can be in millimeters. Figures 4.19 to 4.30 show the estimated $\vec{\boldsymbol{\tau}}$ field over Iceland along with the corresponding meteorological covariate and the difference between the two (posterior estimate minus covariate). The figures for the estimated $\vec{\boldsymbol{\tau}}$ field are made the same way as the corresponding figures for the median precipitation, except that neither the posterior median of $\vec{\boldsymbol{\tau}}$ nor the covariate at unobserved sites are transformed in any way.

Table 4.1: Final estimates of the parameters ψ, κ and ρ. Note that ψ and ρ are in km.

Month	ψ	κ	ρ
January	139.625	0.022	199.460
February	133.647	0.018	309.492
March	102.222	0.022	176.148
April	85.492	0.030	234.135
May	150.796	0.026	247.693
June	143.506	0.023	158.264
July	97.881	0.022	101.621
August	113.890	0.014	274.103
September	113.008	0.014	140.453
October	95.611	0.012	249.614
November	120.915	0.021	230.886
December	113.152	0.021	165.616

Table 4.2: Parameters β_1, β_2, β_3 and β_4 in the spatial mean field $\vec{\alpha}$ for January-June.

January	2.5% quantile	Median	97.5% quantile
β_1	1.001	1.074	1.149
β_2	-0.005	-0.001	0.004
β_3	-0.006	0.001	0.008
β_4	-0.648	-0.254	0.143
February	2.5% quantile	Median	97.5% quantile
β_1	1.000	1.079	1.158
β_2	-0.006	-0.001	0.004
β_3	-0.005	0.001	0.008
β_4	-0.440	-0.065	0.305
March	2.5%-quant	median	97.5%-quant
β_1	0.983	1.054	1.124
β_2	-0.006	-0.001	0.003
β_3	-0.004	0.002	0.009
β_4	-0.469	-0.094	0.288
April	2.5%-quant	median	97.5%-quant
β_1	0.979	1.043	1.107
β_2	-0.004	-0.001	0.003
β_3	-0.006	-0.001	0.005
β_4	-0.479	-0.140	0.195
May	2.5%-quant	median	97.5%-quant
β_1	0.898	0.980	1.062
β_2	-0.005	-0.001	0.003
β_3	-0.007	-0.002	0.003
β_4	-0.382	-0.099	0.182
June	2.5%-quant	median	97.5%-quant
β_1	0.916	0.991	1.065
β_2	-0.006	-0.002	0.001
β_3	-0.008	-0.002	0.003
β_4	-0.076	0.220	0.510

Table 4.3: Parameters β_1, β_2, β_3 and β_4 in the spatial mean field $\vec{\alpha}$ for July-December.

July	2.5%-quant	median	97.5%-quant
β_1	0.914	0.976	1.039
β_2	-0.005	-0.001	0.002
β_3	-0.008	-0.003	0.002
β_4	0.052	0.348	0.645
August	**2.5%-quant**	**median**	**97.5%-quant**
β_1	0.892	0.965	1.038
β_2	-0.006	-0.002	0.002
β_3	-0.006	-0.000	0.005
β_4	-0.040	0.274	0.591
September	**2.5%-quant**	**median**	**97.5%-quant**
β_1	0.983	1.045	1.106
β_2	-0.007	-0.003	0.001
β_3	-0.003	0.003	0.009
β_4	-0.365	-0.038	0.285
October	**2.5%-quant**	**median**	**97.5%-quant**
β_1	0.989	1.051	1.113
β_2	-0.006	-0.001	0.003
β_3	-0.005	0.001	0.006
β_4	-0.488	-0.155	0.182
November	**2.5%-quant**	**median**	**97.5%-quant**
β_1	1.025	1.094	1.162
β_2	-0.005	-0.001	0.004
β_3	-0.004	0.002	0.009
β_4	-0.618	-0.259	0.101
December	**2.5%-quant**	**median**	**97.5%-quant**
β_1	1.012	1.078	1.142
β_2	-0.005	-0.000	0.004
β_3	-0.008	-0.002	0.005
β_4	-0.613	-0.226	0.163

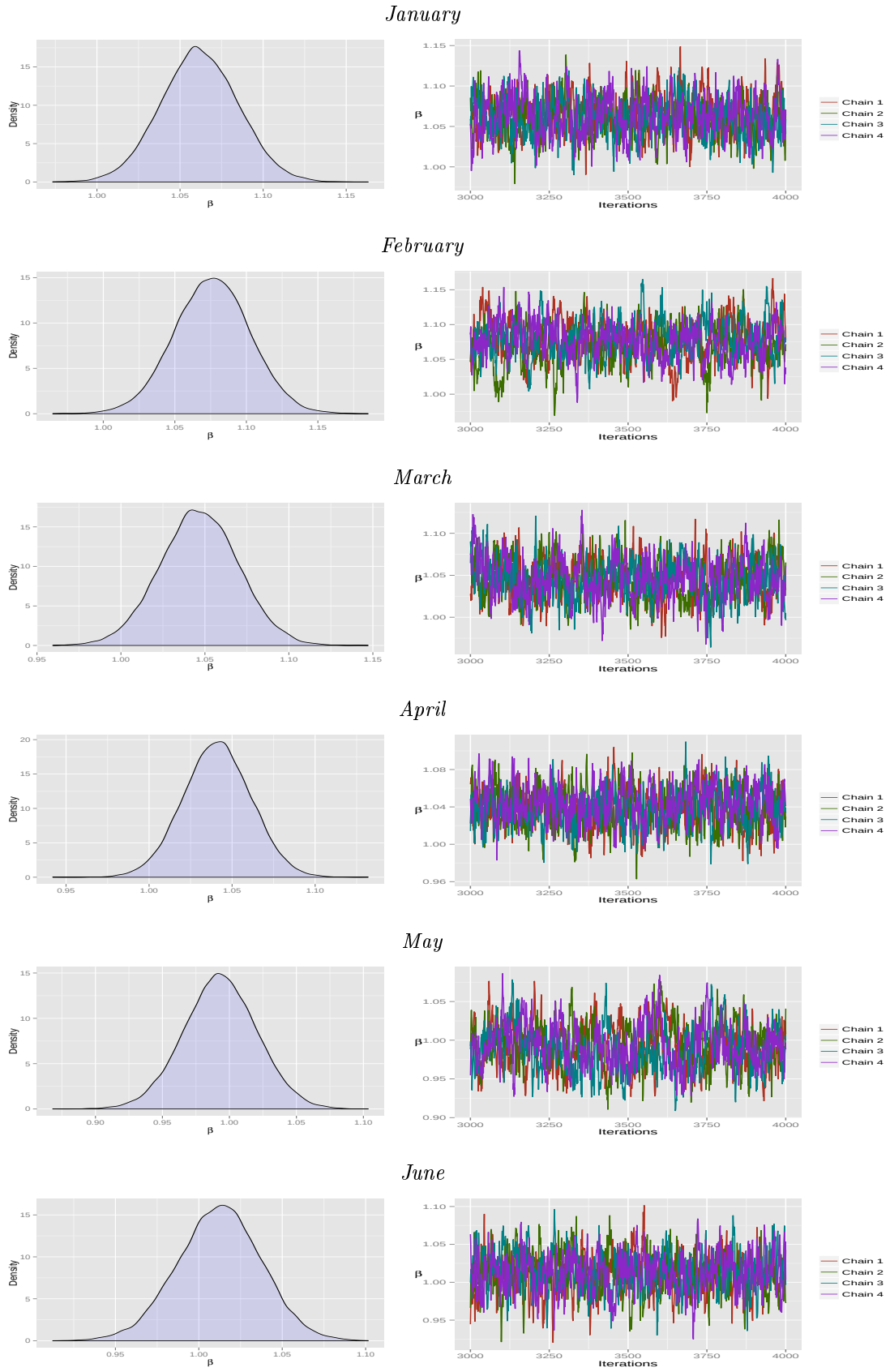

Figure 4.1: Posterior density and the MCMC chains for the parameter β for January-June.

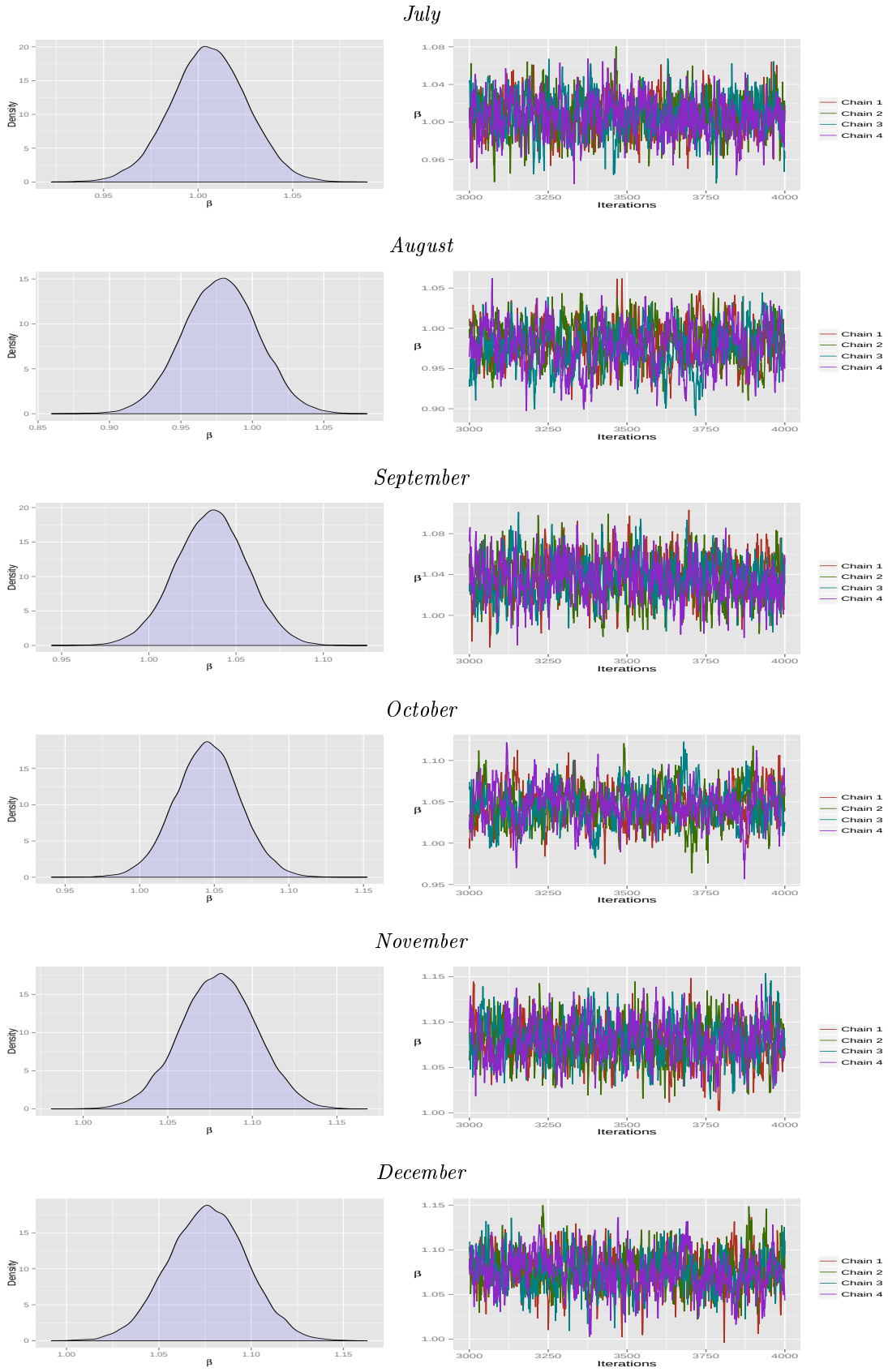

Figure 4.2: Posterior density and the MCMC chains for the parameter β for July-December.

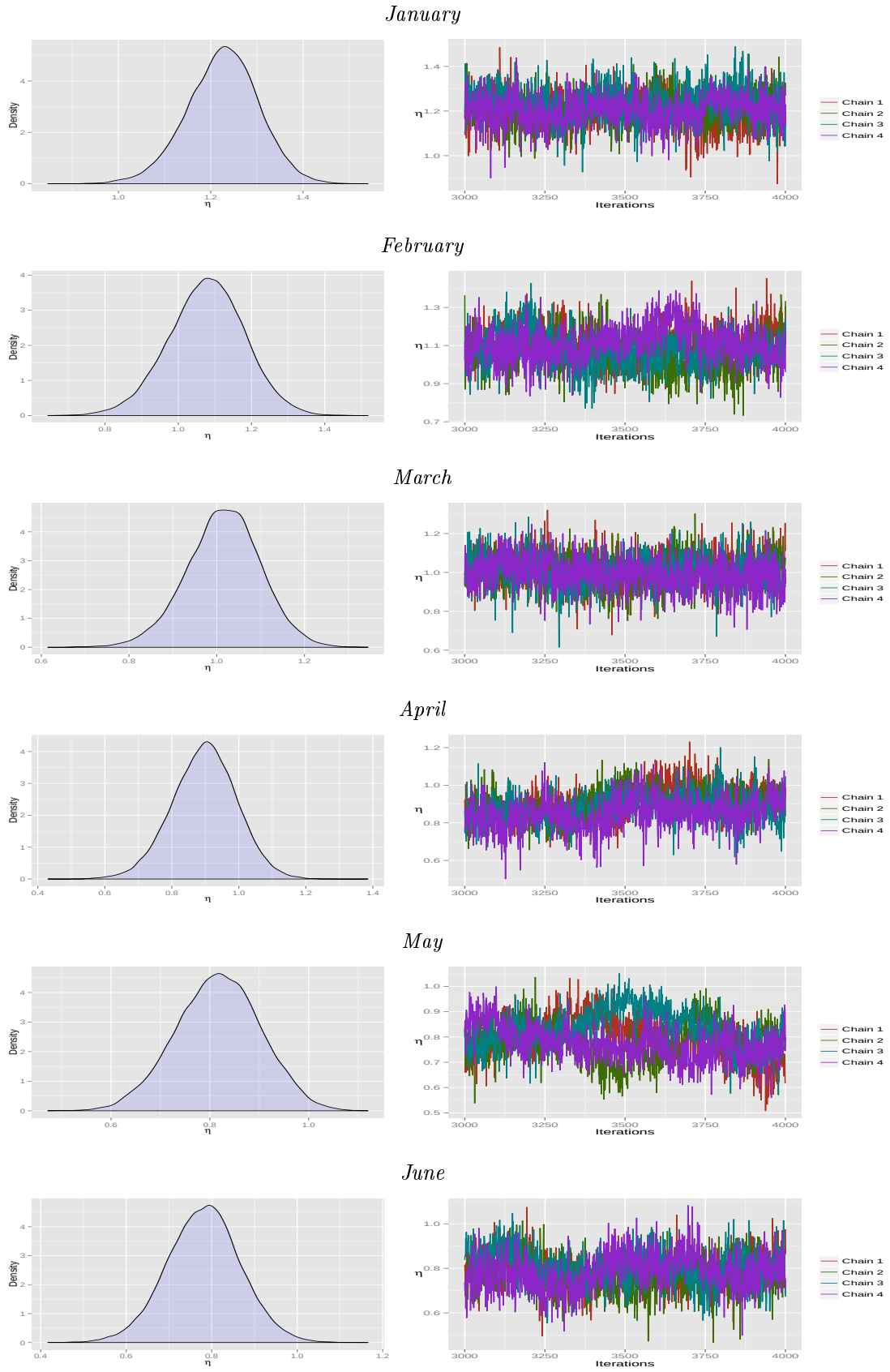

Figure 4.3: Posterior density and the MCMC chains for the parameter η for January-June.

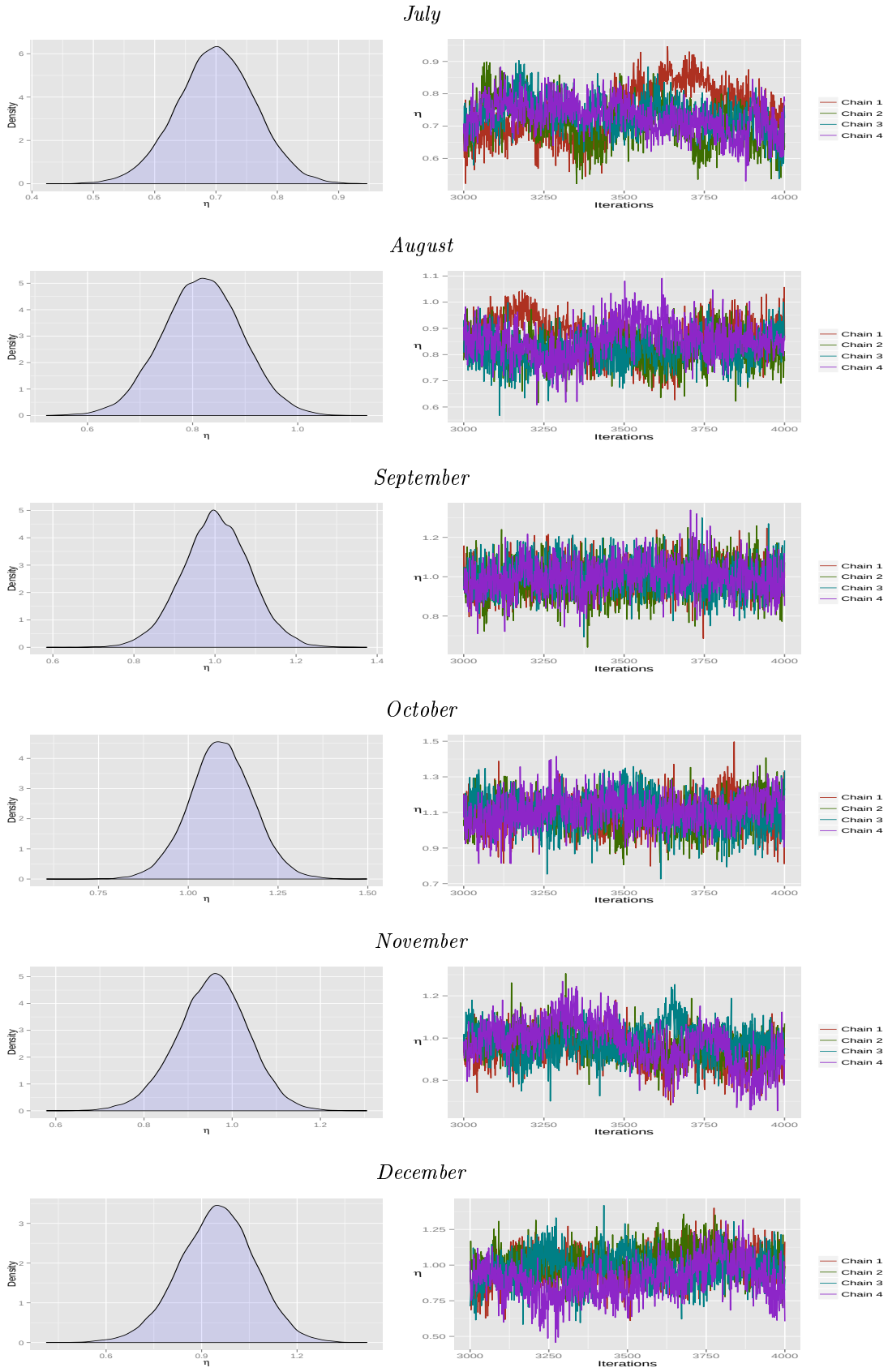

Figure 4.4: Posterior density and the MCMC chains for the parameter η for July-December.

Table 4.4: Posterior statistics for the parameters β, η, σ_α^2, σ_γ^2, σ_τ^2 and ϕ for January-June.

January	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.06	0.02	1.02	1.05	1.08	1.11
η	1.23	0.08	1.07	1.18	1.28	1.37
σ_α^2	2.64	0.69	1.71	2.25	3.11	4.36
σ_γ^2	2.29	1.20	0.67	1.61	3.13	5.31
σ_τ^2	0.05	0.02	0.02	0.04	0.06	0.10
ϕ	0.49	2.35	0.03	0.18	1.34	8.56
February	**Median**	**Stand. dev.**	**2.5%-quantile**	**25%-quantile**	**75%-quantile**	**97.5%-quantile**
β	1.08	0.03	1.02	1.06	1.09	1.13
η	1.08	0.10	0.87	1.01	1.15	1.28
σ_α^2	2.07	0.54	1.32	1.76	2.44	3.43
σ_γ^2	5.02	1.78	2.43	3.97	6.23	9.39
σ_τ^2	0.08	0.03	0.04	0.06	0.10	0.15
ϕ	0.47	1.91	0.03	0.18	1.28	6.84
March	**Median**	**Stand. dev.**	**2.5%-quantile**	**25%-quantile**	**75%-quantile**	**97.5%-quantile**
β	1.05	0.02	1.00	1.03	1.06	1.09
η	1.02	0.08	0.84	0.96	1.07	1.18
σ_α^2	2.23	0.59	1.42	1.90	2.65	3.70
σ_γ^2	3.62	1.25	1.80	2.89	4.48	6.64
σ_τ^2	0.08	0.03	0.04	0.07	0.10	0.15
ϕ	0.46	1.78	0.02	0.17	1.26	6.44
April	**Median**	**Stand. dev.**	**2.5%-quantile**	**25%-quantile**	**75%-quantile**	**97.5%-quantile**
β	1.04	0.02	1.00	1.03	1.05	1.08
η	0.90	0.10	0.70	0.83	0.96	1.08
σ_α^2	1.75	0.46	1.11	1.49	2.06	2.89
σ_γ^2	1.79	0.65	0.86	1.42	2.24	3.38
σ_τ^2	0.06	0.02	0.03	0.05	0.08	0.12
ϕ	0.46	1.89	0.02	0.17	1.23	6.95
May	**Median**	**Stand. dev.**	**2.5%-quantile**	**25%-quantile**	**75%-quantile**	**97.5%-quantile**
β	0.99	0.03	0.94	0.97	1.01	1.05
η	0.82	0.09	0.65	0.76	0.88	0.98
σ_α^2	1.20	0.34	0.74	1.01	1.43	2.06
σ_γ^2	3.08	1.13	1.38	2.41	3.85	5.83
σ_τ^2	0.07	0.02	0.04	0.06	0.09	0.13
ϕ	0.47	2.38	0.02	0.17	1.33	8.74
June	**Median**	**Stand. dev.**	**2.5%-quantile**	**25%-quantile**	**75%-quantile**	**97.5%-quantile**
β	1.01	0.02	0.96	1.00	1.03	1.06
η	0.78	0.09	0.61	0.72	0.84	0.95
σ_α^2	1.51	0.41	0.95	1.28	1.79	2.52
σ_γ^2	1.54	0.71	0.57	1.13	2.03	3.28
σ_τ^2	0.10	0.03	0.06	0.08	0.12	0.18
ϕ	0.49	2.17	0.03	0.18	1.40	8.17

Table 4.5: Posterior statistics for the parameters β, η, σ_α^2, σ_γ^2, σ_τ^2 and ϕ for July-December.

July	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.01	0.02	0.97	0.99	1.02	1.04
η	0.70	0.06	0.57	0.66	0.74	0.83
σ_α^2	1.77	0.46	1.14	1.51	2.09	2.92
σ_γ^2	0.84	0.39	0.33	0.62	1.11	1.86
σ_τ^2	0.05	0.02	0.03	0.04	0.07	0.11
ϕ	0.50	2.39	0.03	0.18	1.41	9.01
August	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	0.98	0.03	0.93	0.96	1.00	1.03
η	0.82	0.08	0.67	0.77	0.87	0.97
σ_α^2	1.80	0.48	1.13	1.52	2.13	3.00
σ_γ^2	3.53	1.17	1.82	2.84	4.33	6.39
σ_τ^2	0.08	0.03	0.04	0.06	0.09	0.14
ϕ	0.48	2.09	0.03	0.17	1.31	7.86
September	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.04	0.02	1.00	1.02	1.05	1.08
η	1.00	0.08	0.84	0.95	1.06	1.17
σ_α^2	1.80	0.48	1.15	1.53	2.14	3.00
σ_γ^2	1.76	0.64	0.84	1.39	2.21	3.31
σ_τ^2	0.08	0.02	0.05	0.06	0.09	0.14
ϕ	0.43	1.42	0.03	0.17	1.12	5.11
October	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.05	0.02	1.00	1.03	1.06	1.09
η	1.09	0.09	0.92	1.03	1.15	1.27
σ_α^2	1.78	0.47	1.12	1.51	2.10	2.95
σ_γ^2	3.52	1.14	1.89	2.88	4.33	6.33
σ_τ^2	0.07	0.02	0.04	0.06	0.09	0.13
ϕ	0.44	1.81	0.02	0.16	1.19	6.30
November	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.08	0.02	1.04	1.06	1.10	1.12
η	0.96	0.08	0.79	0.90	1.01	1.11
σ_α^2	2.15	0.58	1.37	1.83	2.55	3.61
σ_γ^2	2.47	1.12	0.88	1.82	3.25	5.22
σ_τ^2	0.05	0.02	0.03	0.04	0.07	0.11
ϕ	0.48	2.78	0.02	0.18	1.38	10.42
December	Median	Stand. dev.	2.5%-quantile	25%-quantile	75%-quantile	97.5%-quantile
β	1.08	0.02	1.03	1.06	1.09	1.12
η	0.95	0.12	0.71	0.87	1.03	1.17
σ_α^2	2.42	0.62	1.56	2.07	2.86	3.94
σ_γ^2	2.23	0.86	1.01	1.74	2.82	4.37
σ_τ^2	0.12	0.04	0.07	0.10	0.14	0.22
ϕ	0.46	1.86	0.03	0.17	1.23	6.54

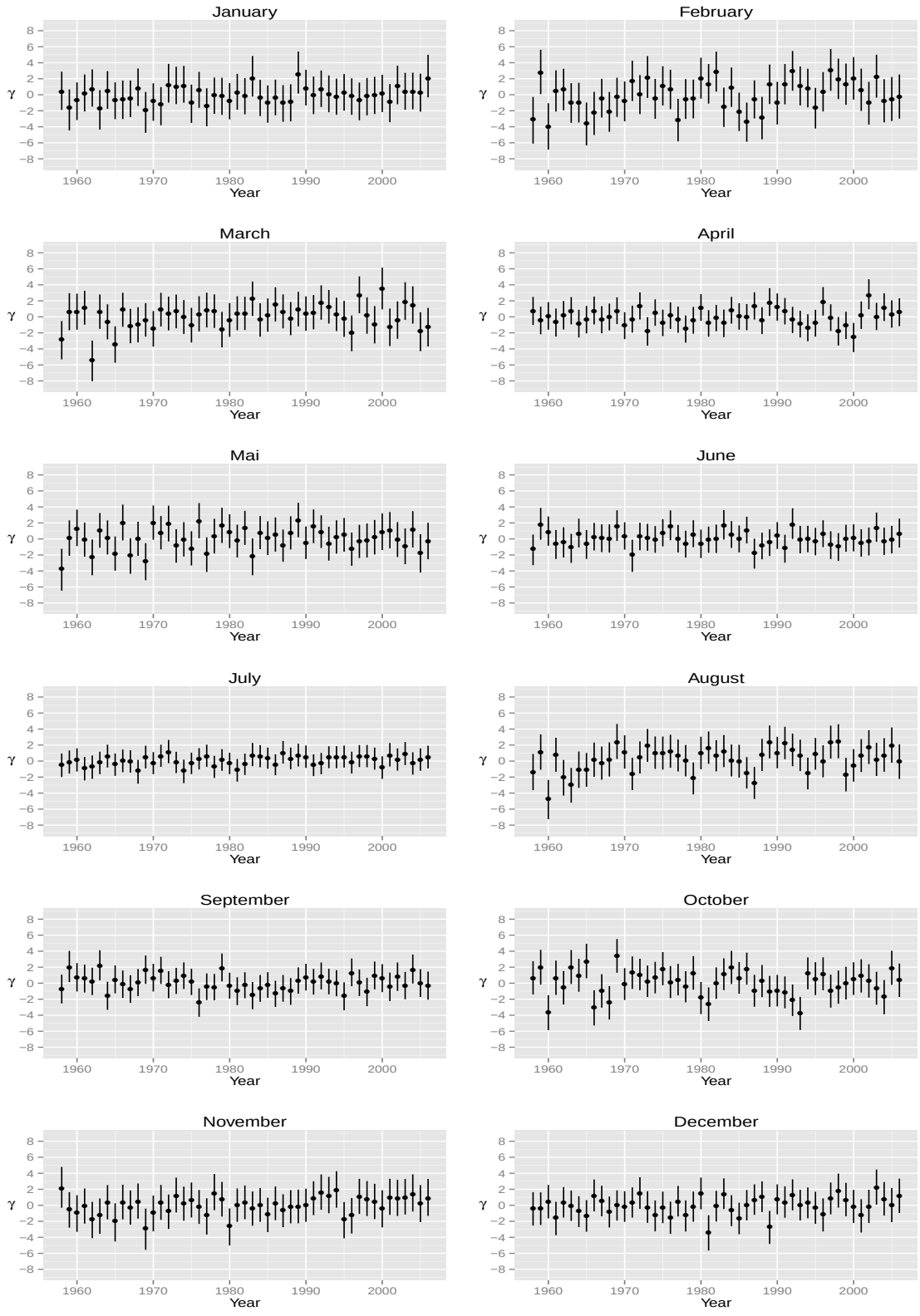

Figure 4.5: Posterior median and 95% posterior interval for the parameter $\vec{\gamma}$ for all months.

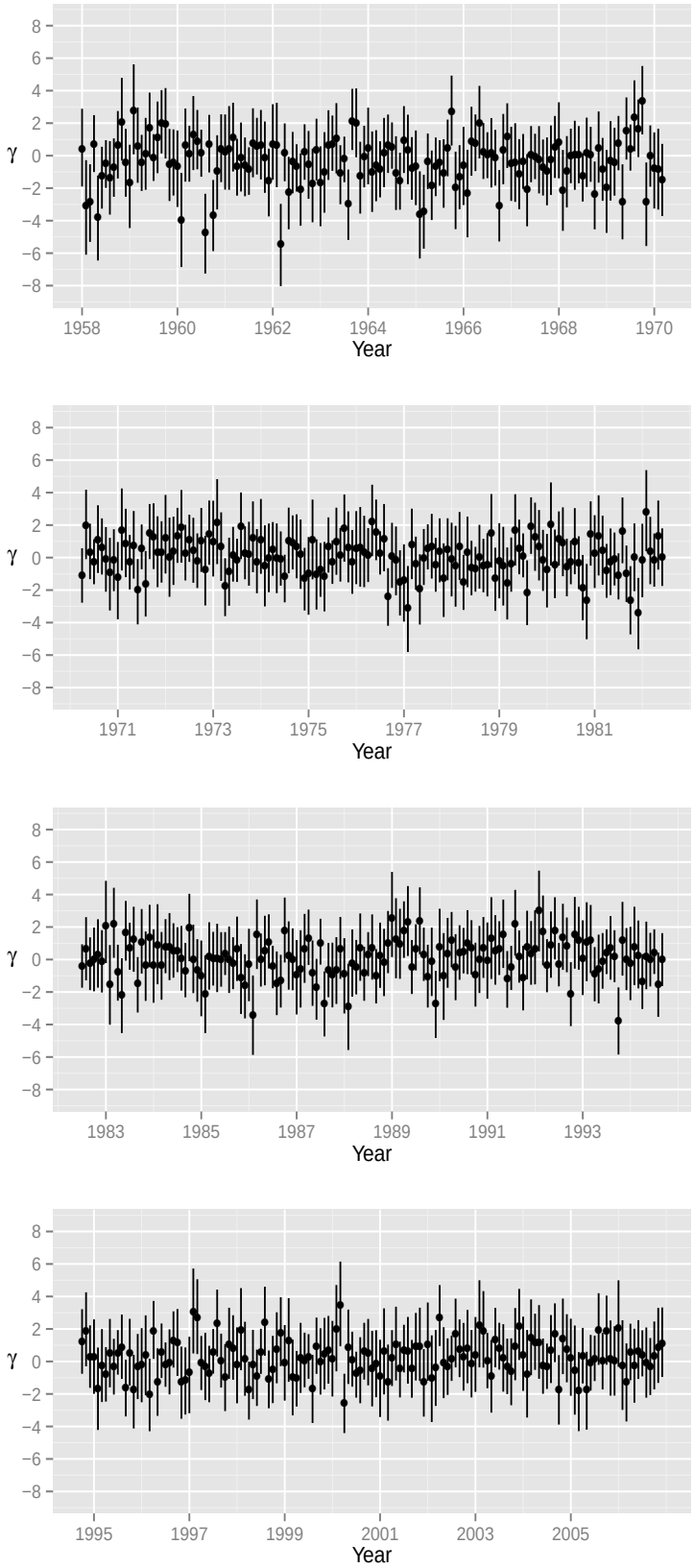

Figure 4.6: Posterior median and 95% posterior interval for the parameter $\vec{\gamma}$ for all the months taken together in correct temporal order.

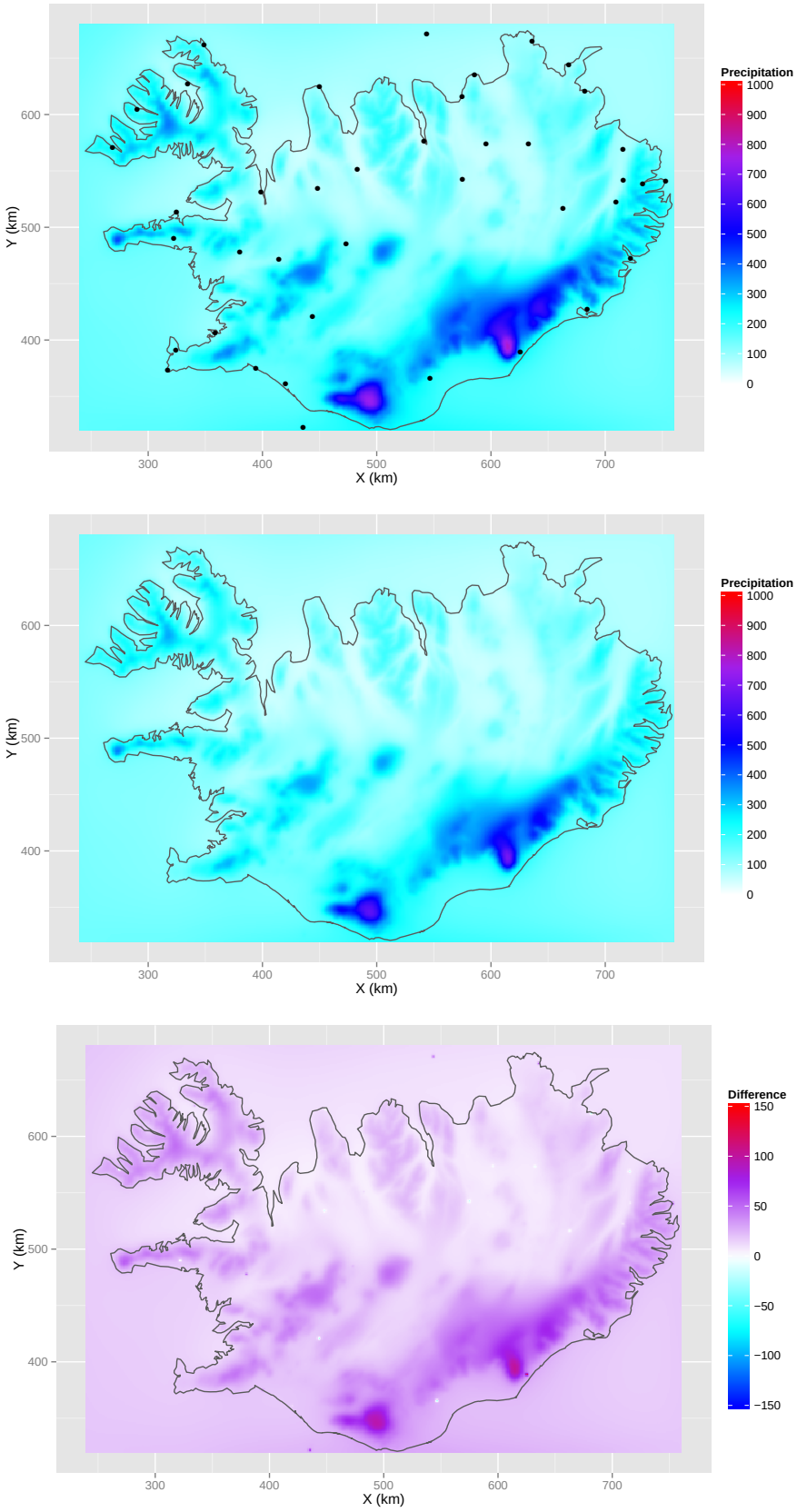

Figure 4.7: Spatial field for median precipitation in January. Posterior estimate, meteorological covariate and the difference between them.

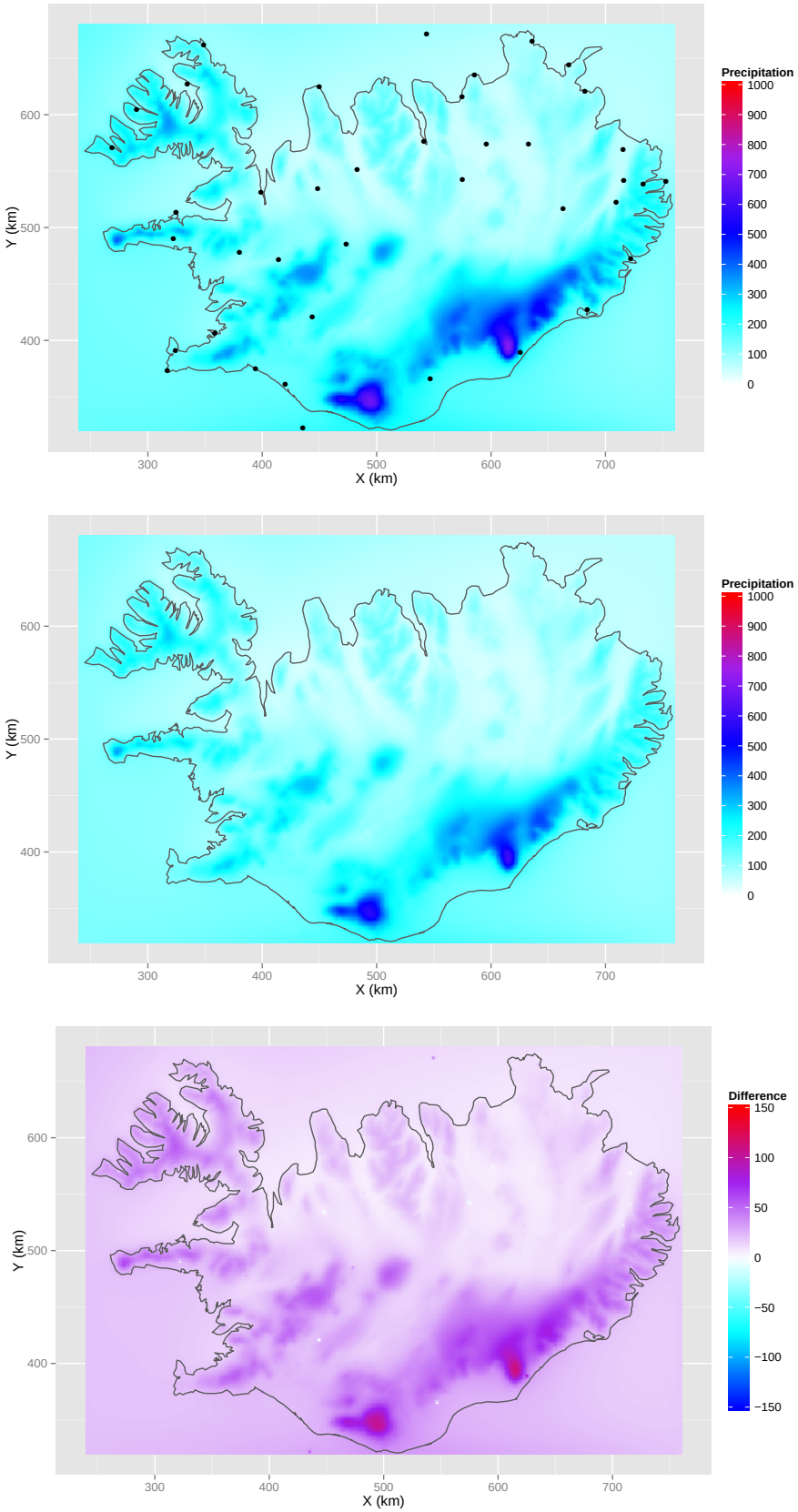

Figure 4.8: Spatial field for median precipitation in February. Posterior estimate, meteorological covariate and the difference between them.

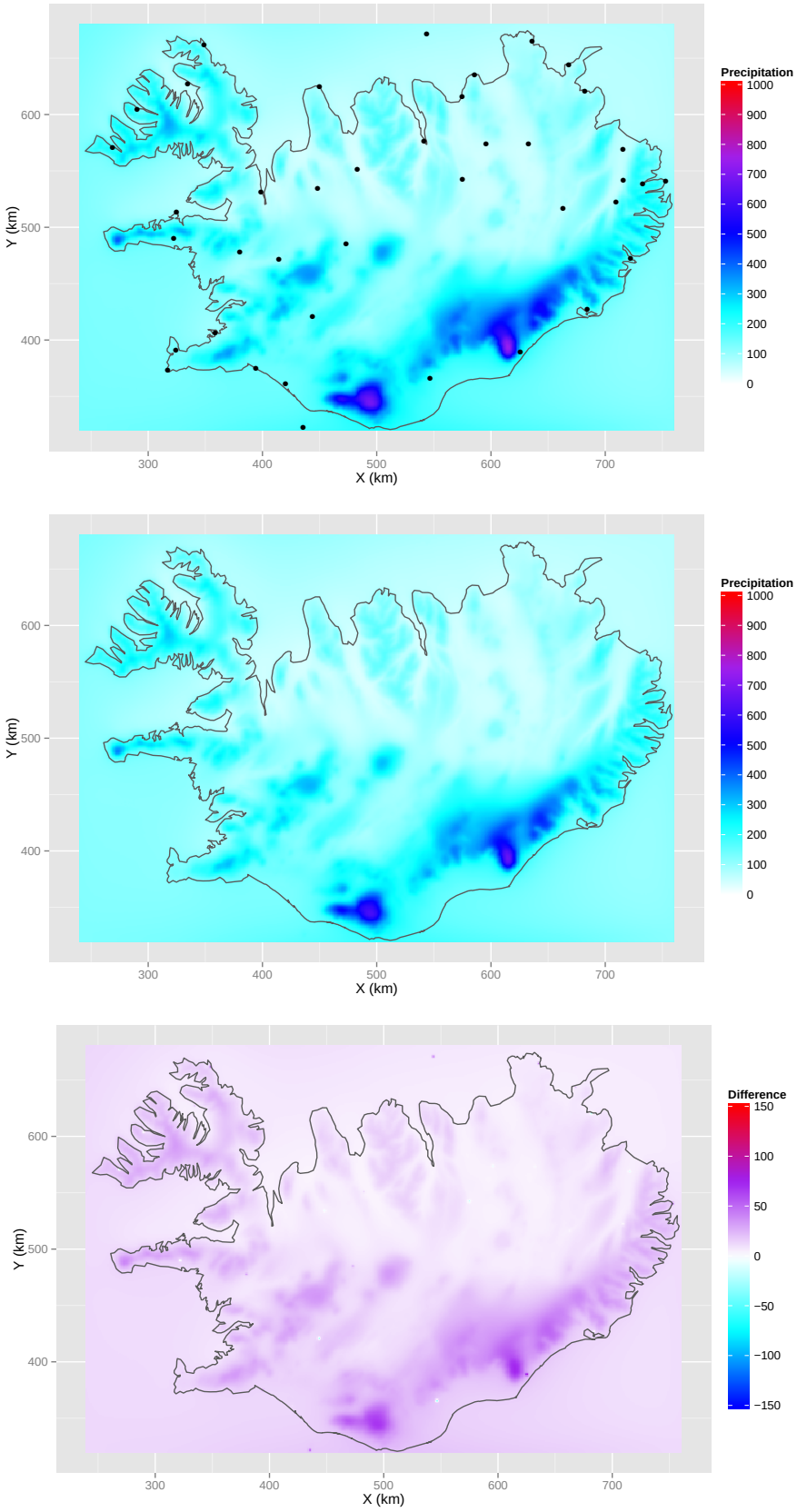

Figure 4.9: Spatial field for median precipitation in March. Posterior estimate, meteorological covariate and the difference between them.

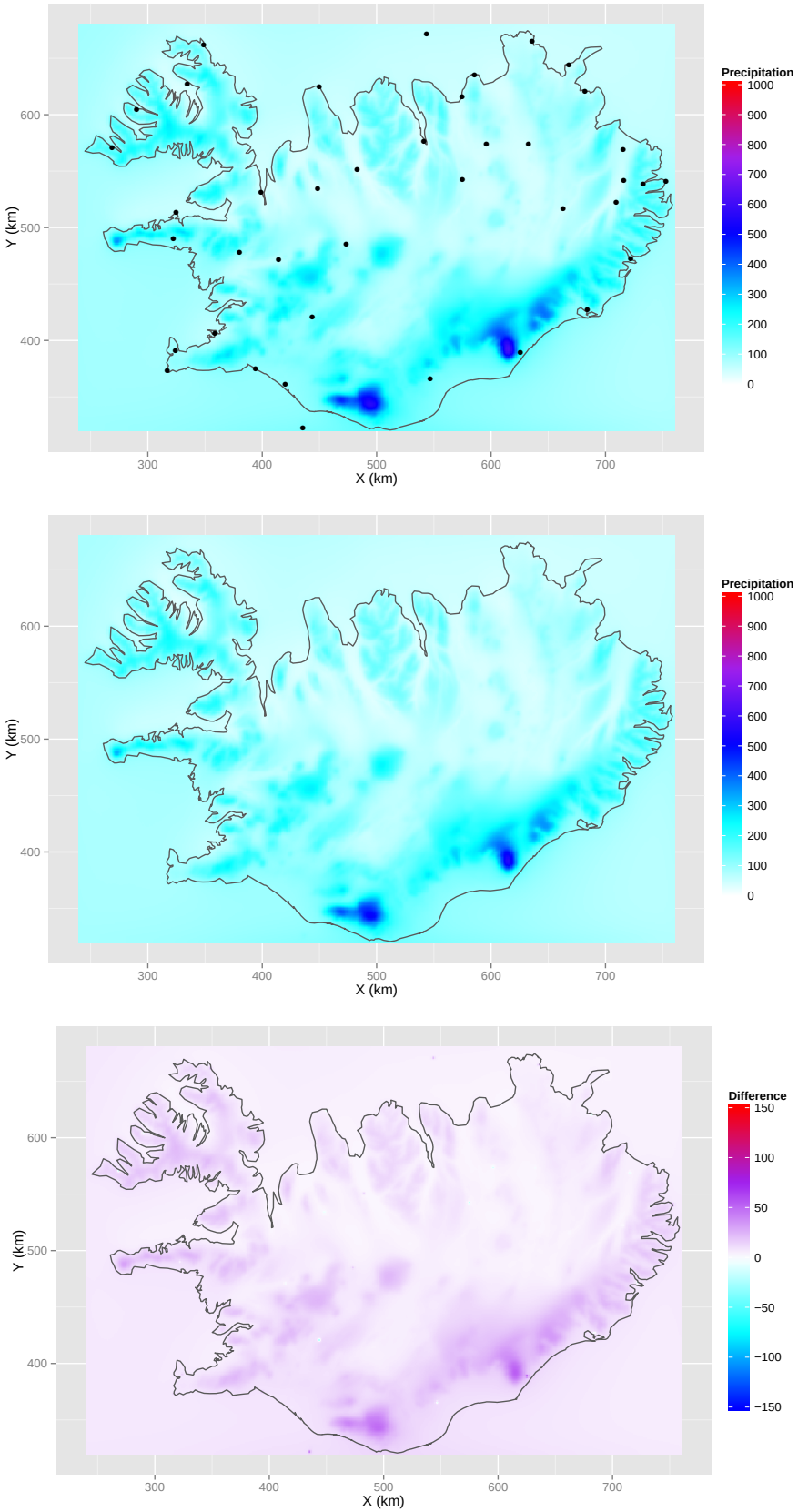

Figure 4.10: Spatial field for median precipitation in April. Posterior estimate, meteorological covariate and the difference between them.

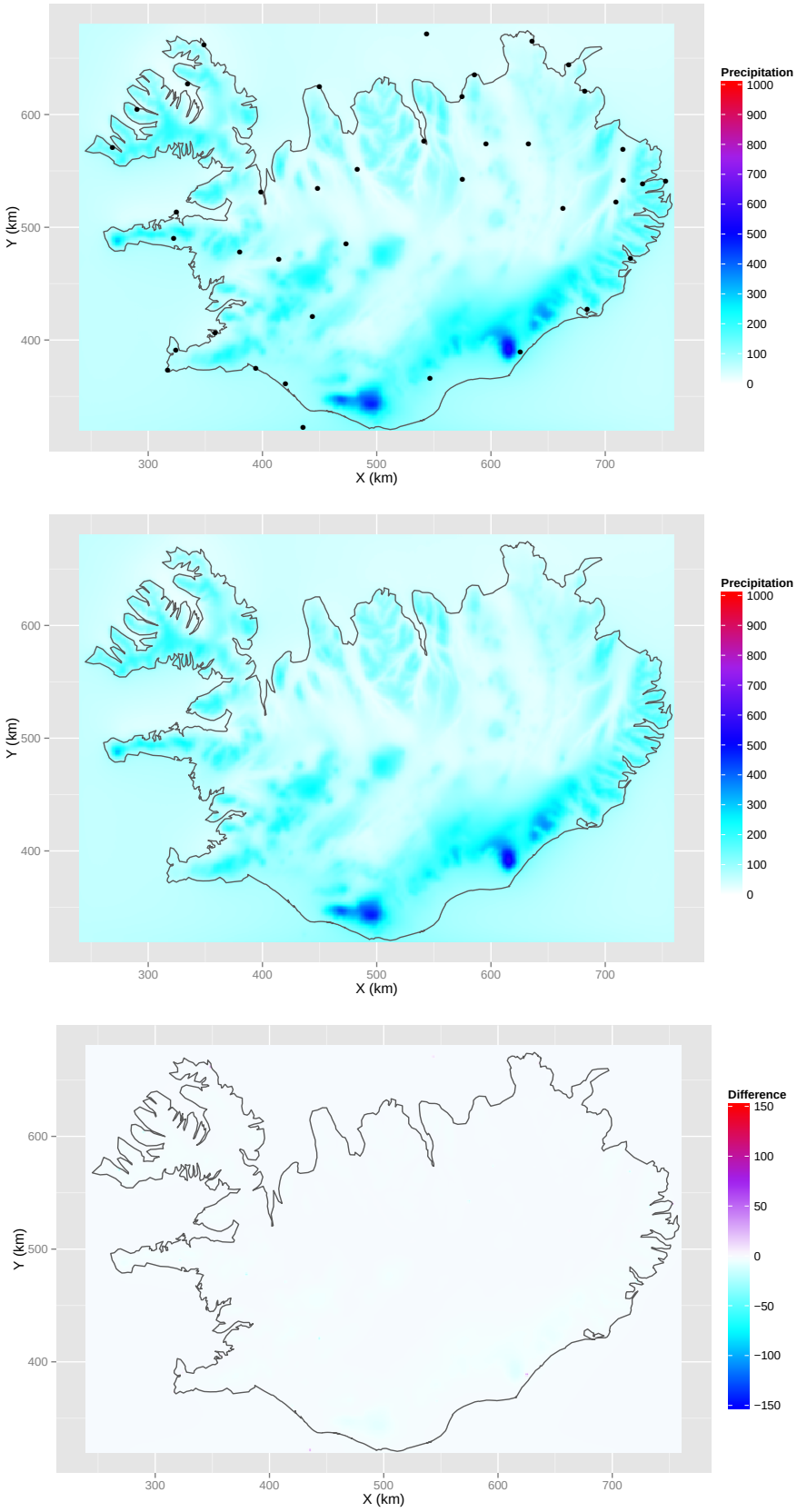

Figure 4.11: Spatial field for median precipitation in May. Posterior estimate, meteorological covariate and the difference between them.

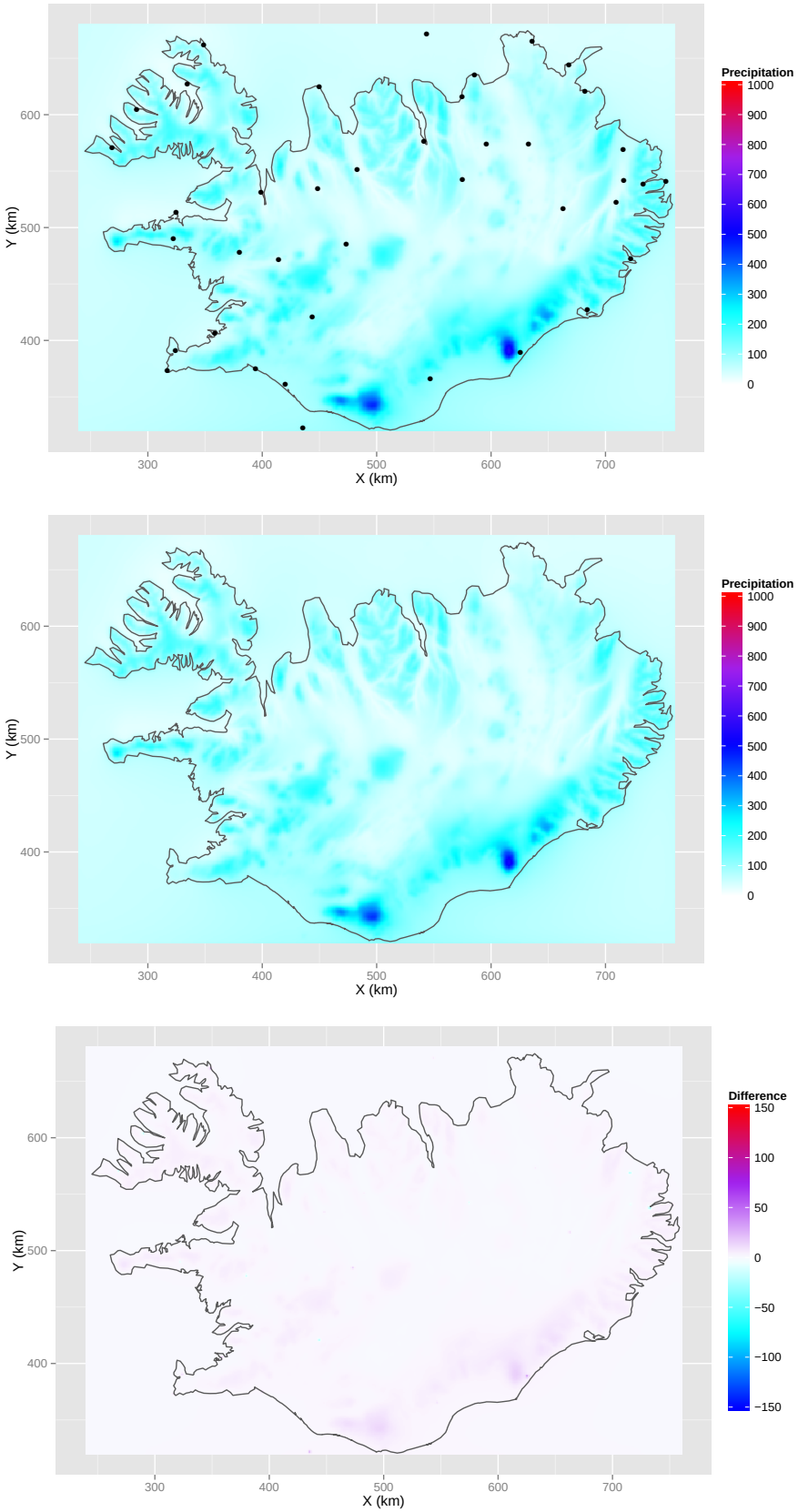

Figure 4.12: Spatial field for median precipitation in June. Posterior estimate, meteorological covariate and the difference between them.

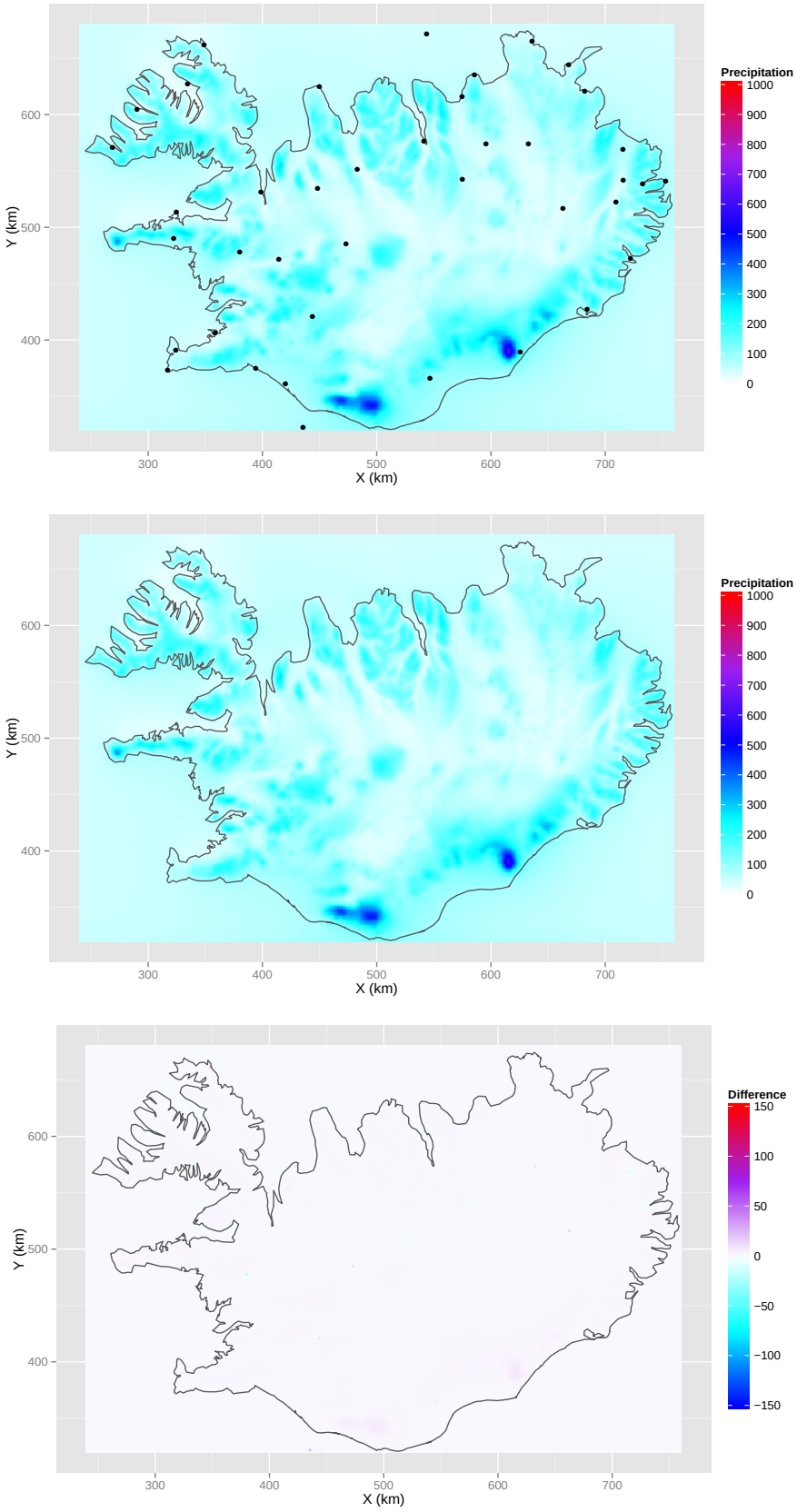

Figure 4.13: Spatial field for median precipitation in July. Posterior estimate, meteorological covariate and the difference between them.

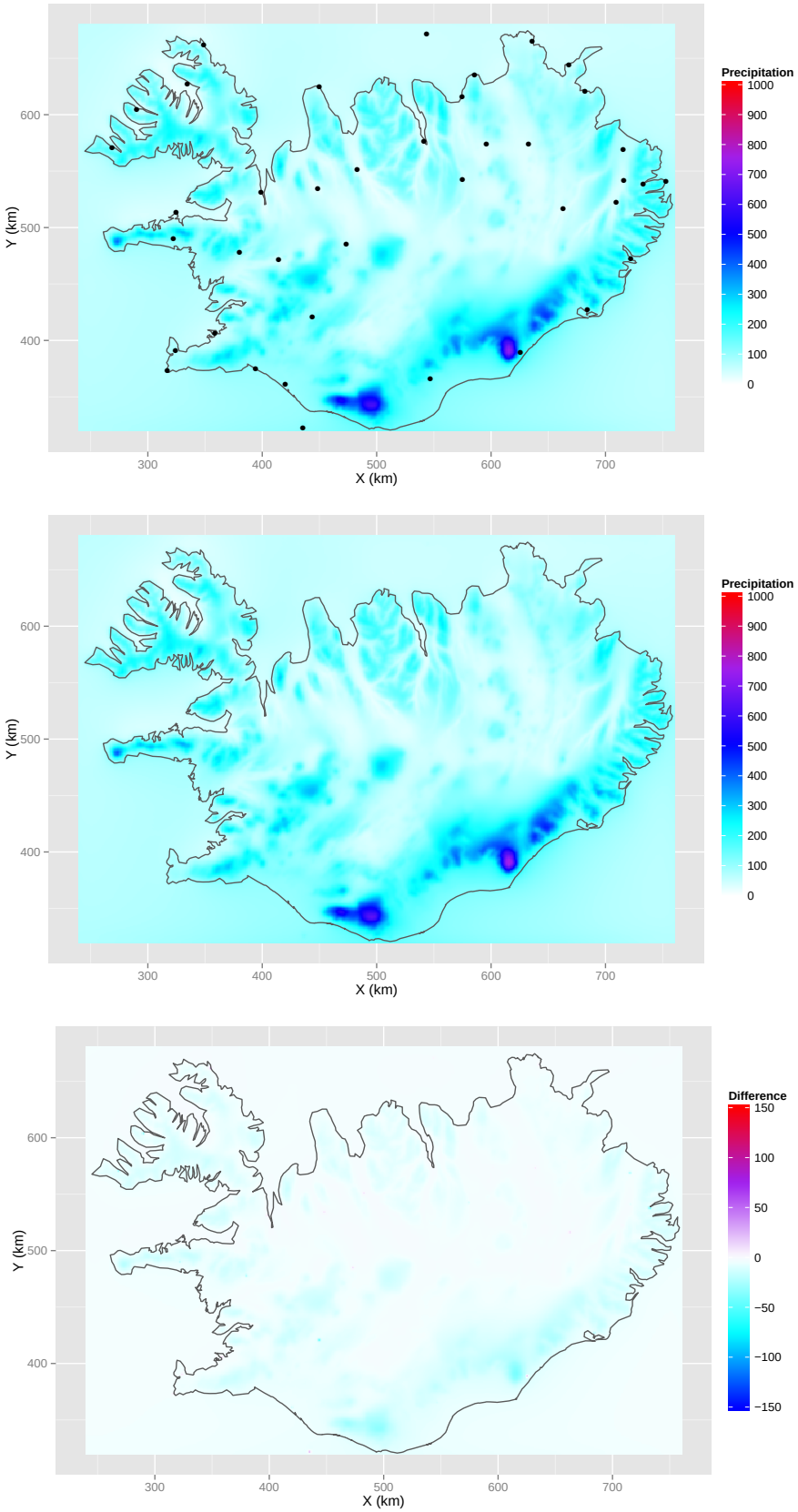

Figure 4.14: Spatial field for median precipitation in August. Posterior estimate, meteorological covariate and the difference between them.

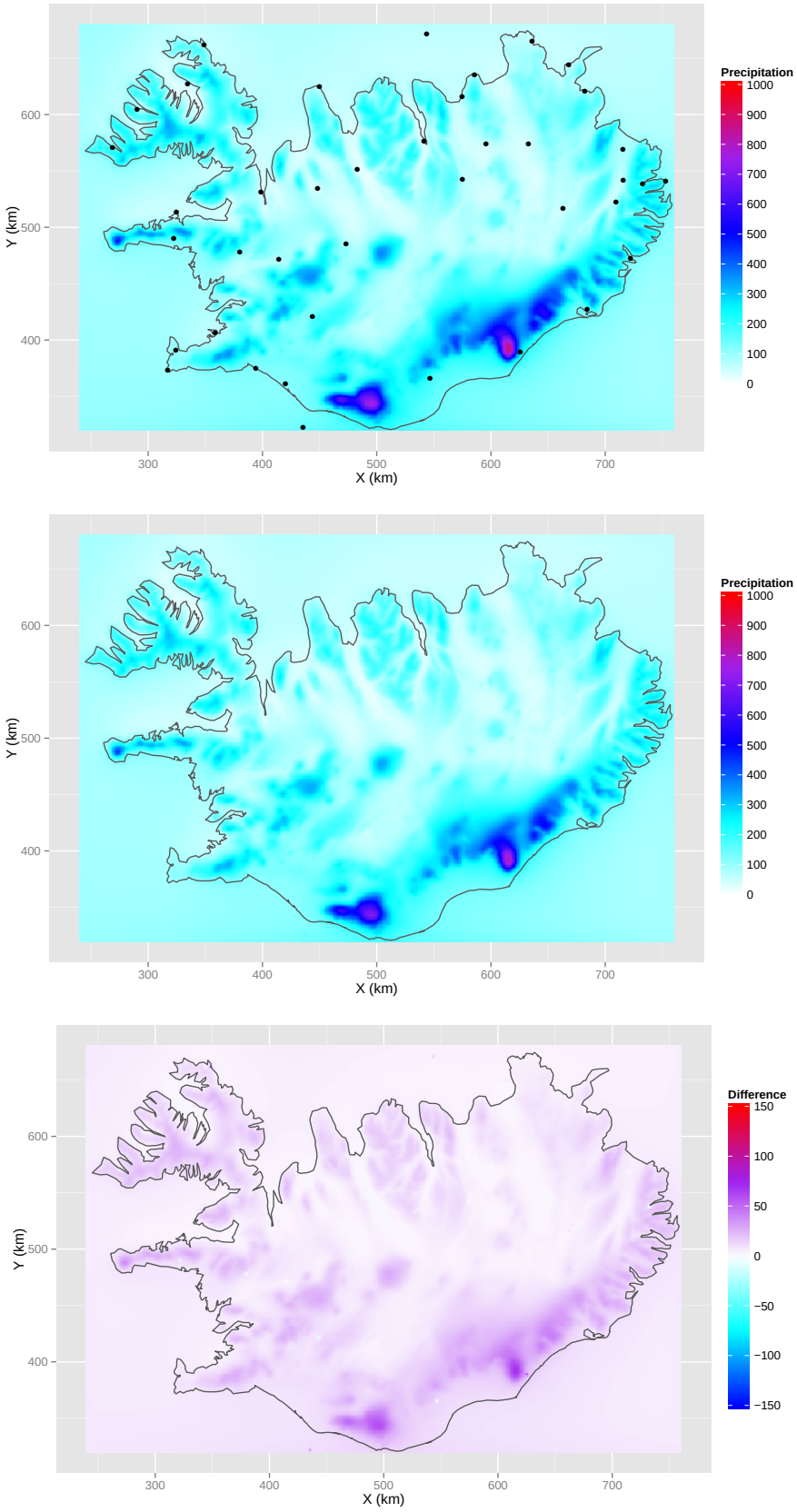

Figure 4.15: Spatial field for median precipitation in September. Posterior estimate, meteorological covariate and the difference between them.

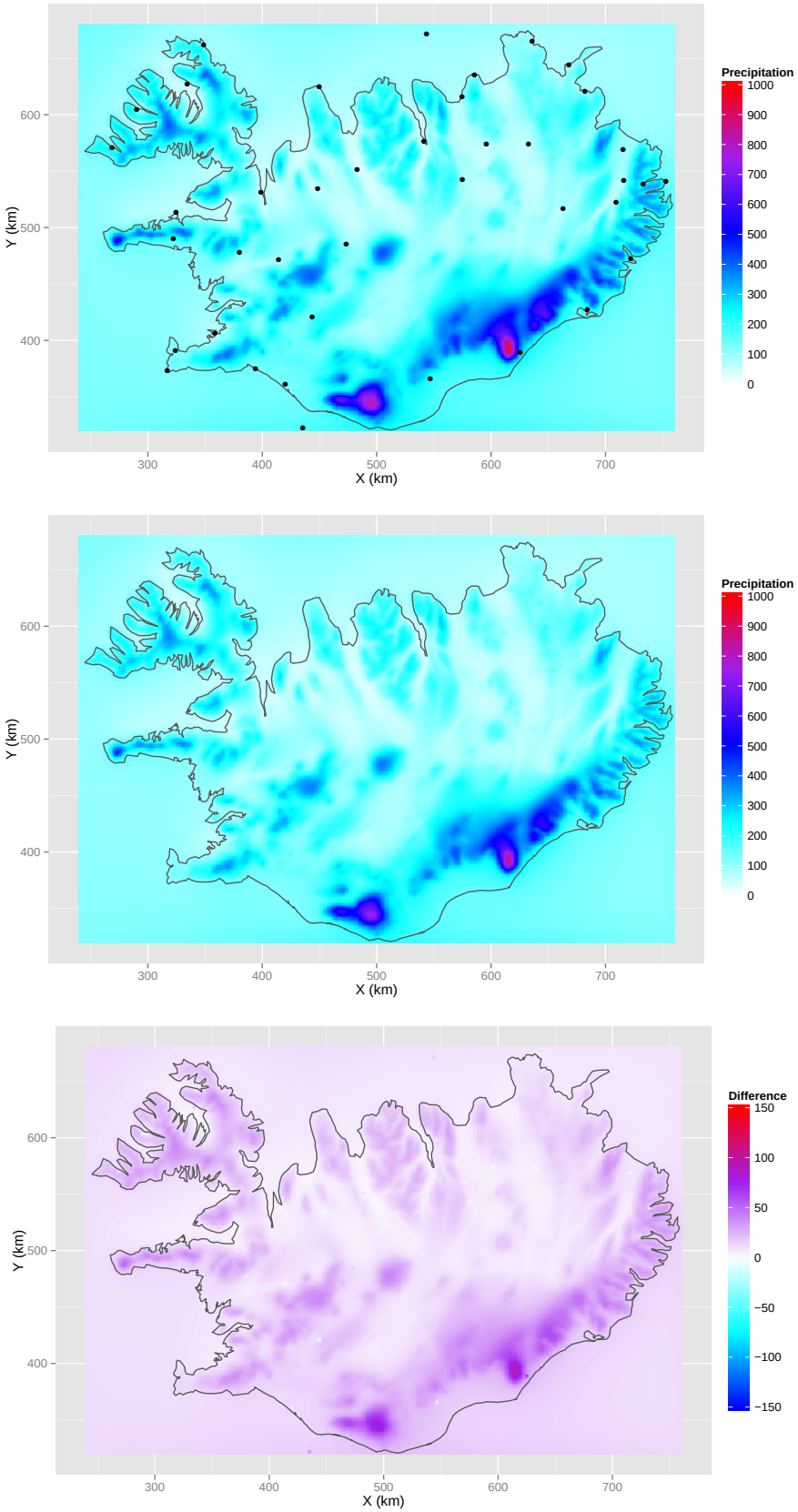

Figure 4.16: Spatial field for median precipitation in October. Posterior estimate, meteorological covariate and the difference between them.

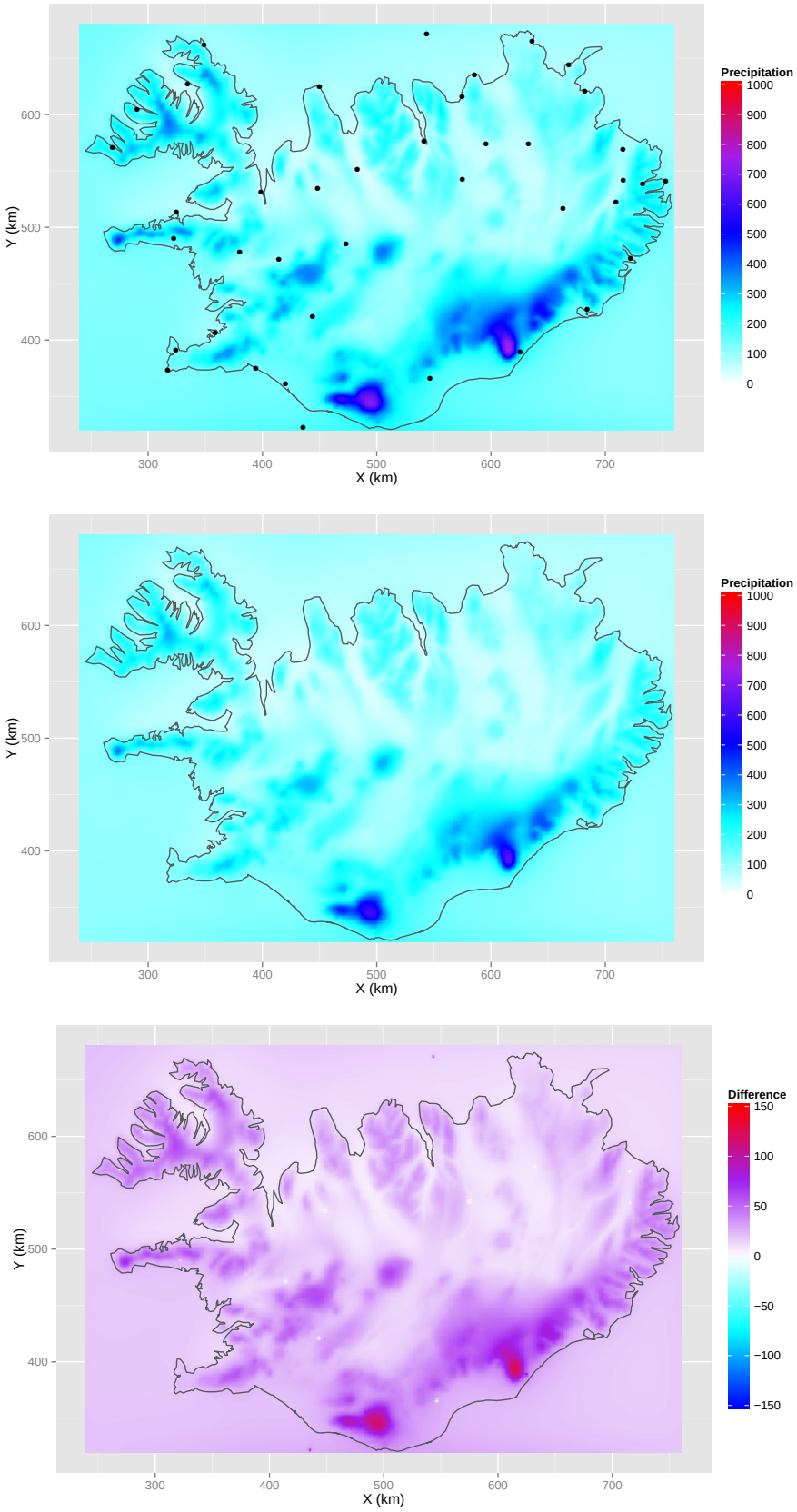

Figure 4.17: Spatial field for median precipitation in November. Posterior estimate, meteorological covariate and the difference between them.

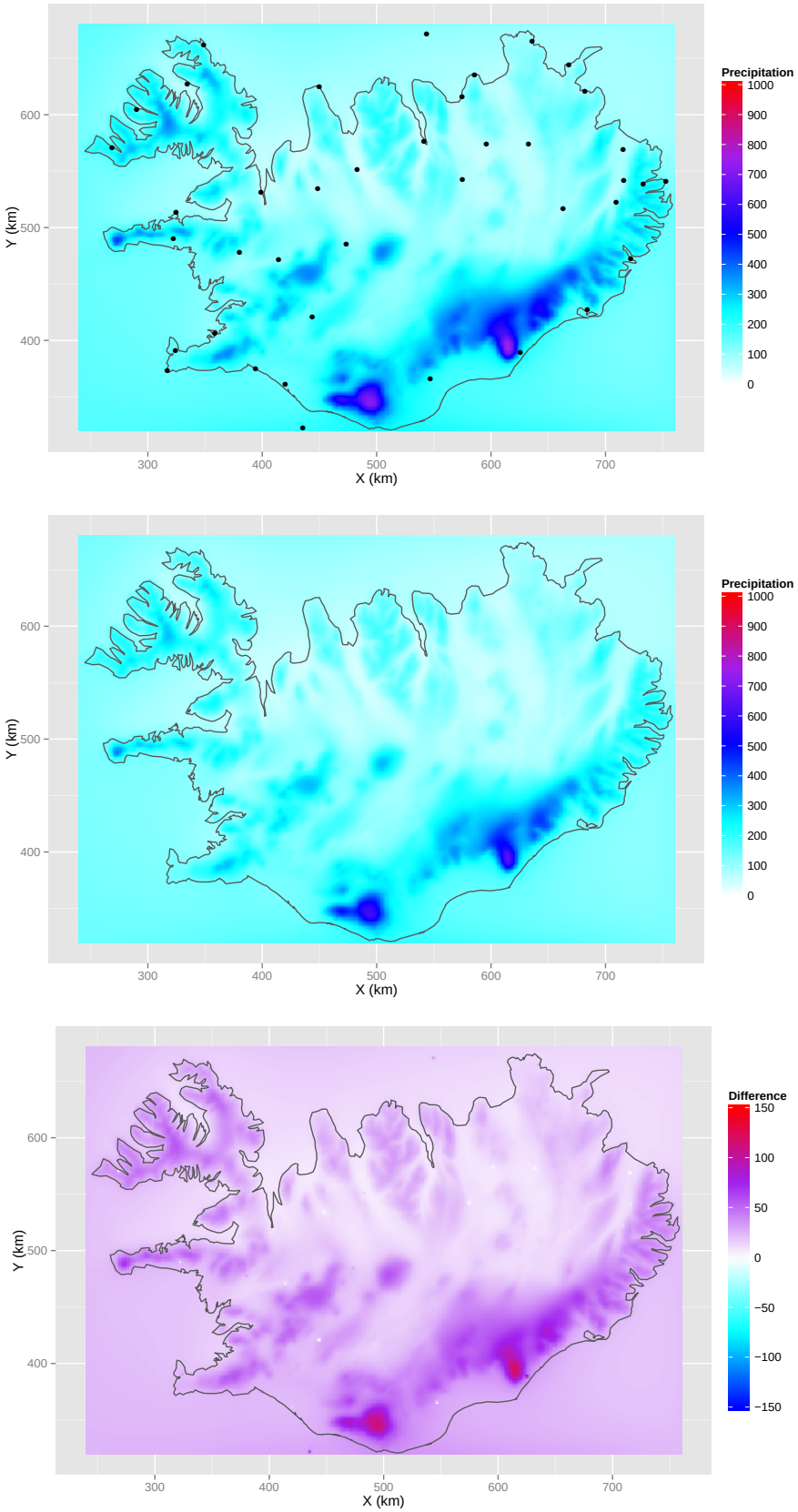

Figure 4.18: Spatial field for median precipitation in December. Posterior estimate, meteorological covariate and the difference between them.

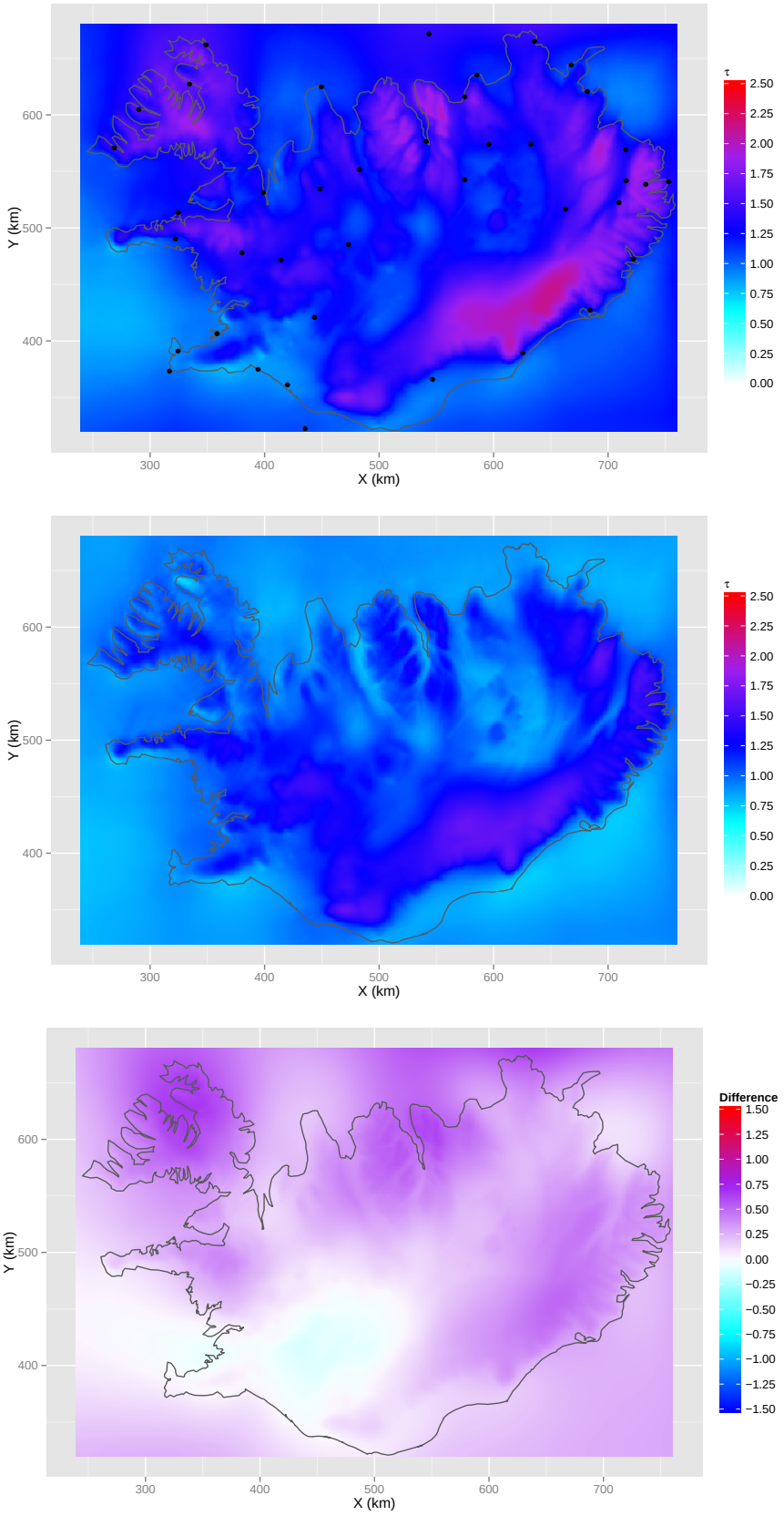

Figure 4.19: Spatial field for $\vec{\tau}$ in January. Posterior estimate, meteorological covariate and the difference between them.

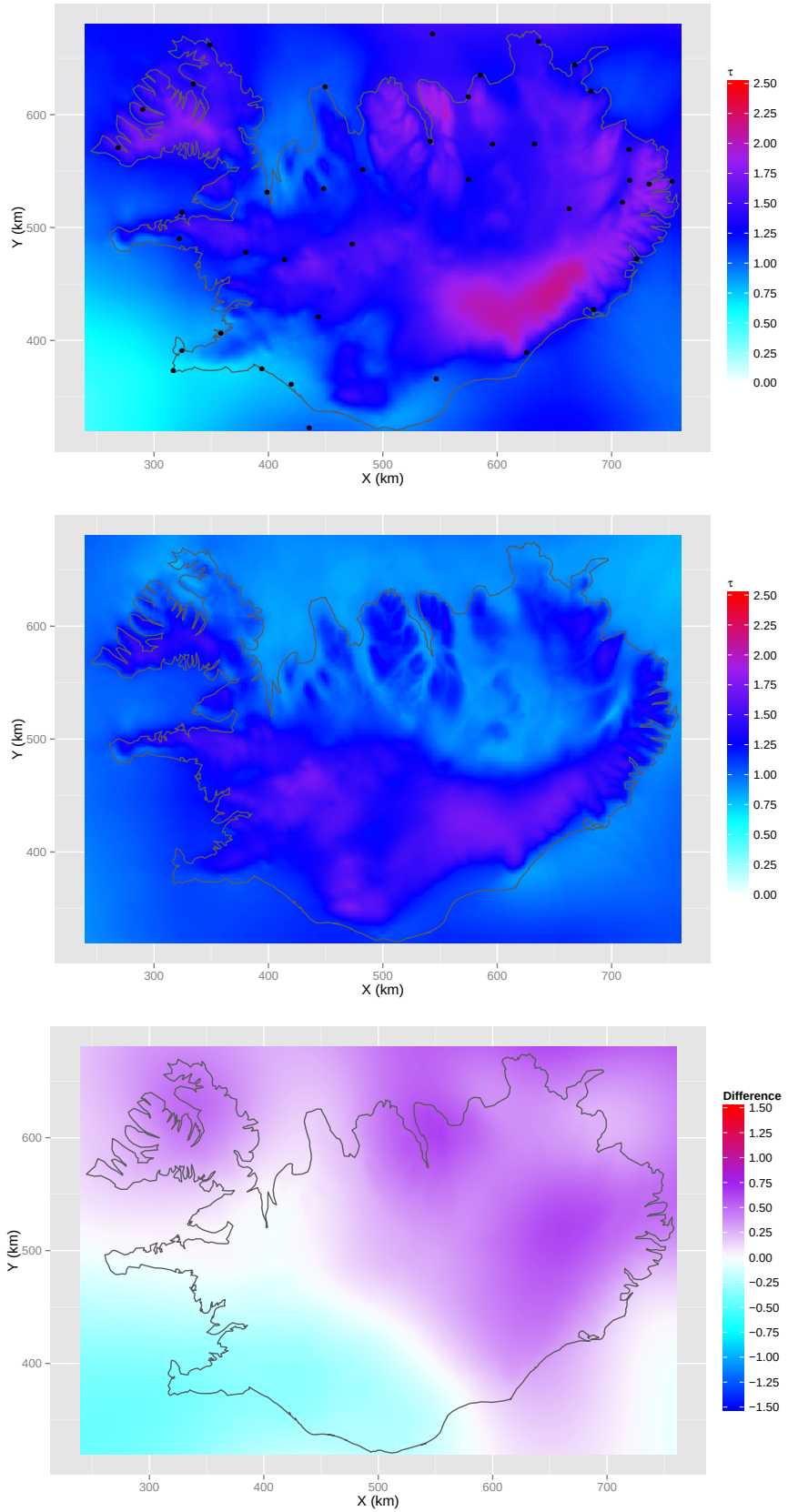

Figure 4.20: Spatial field for $\vec{\tau}$ in February. Posterior estimate, meteorological covariate and the difference between them.

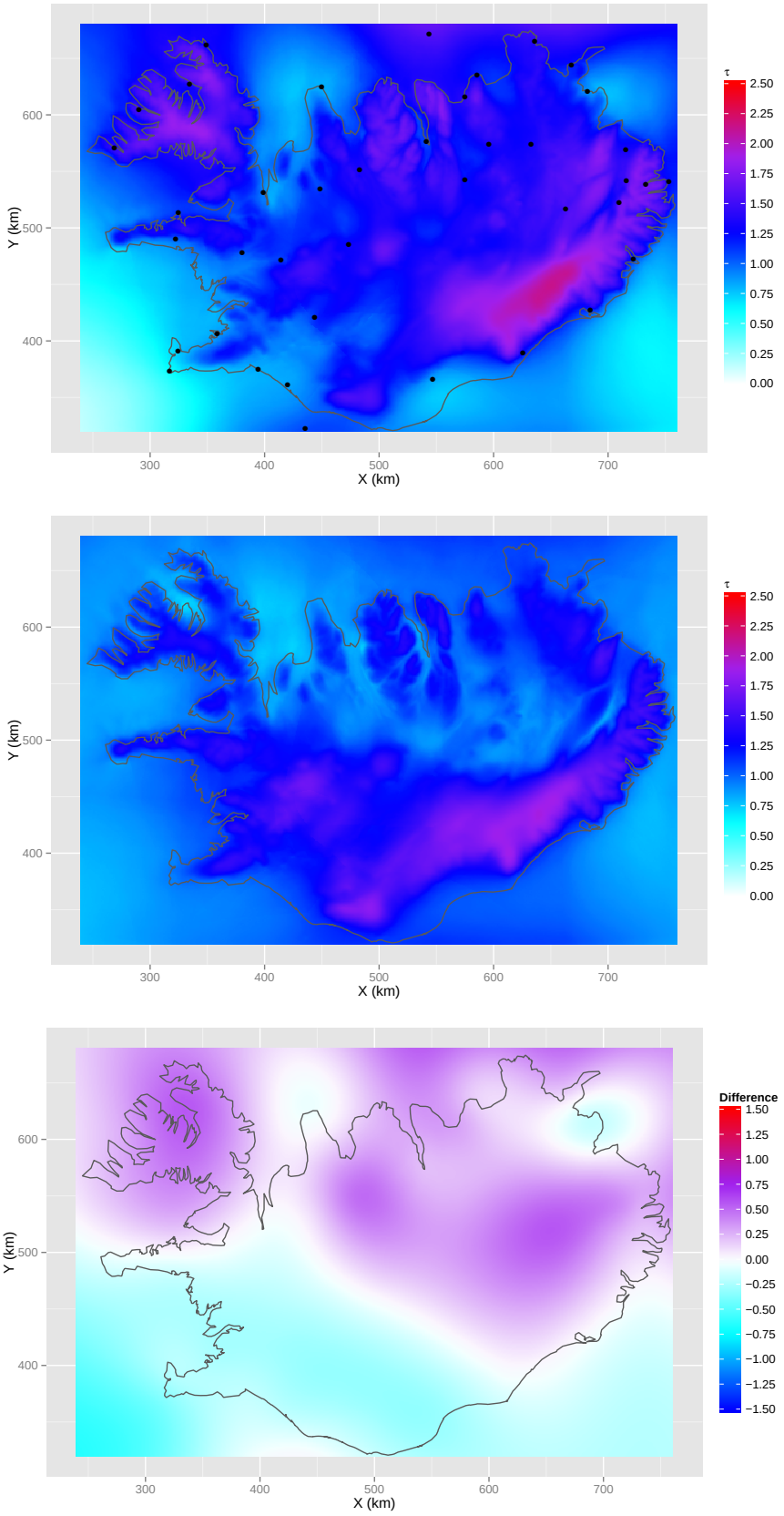

Figure 4.21: Spatial field for $\vec{\tau}$ in March. Posterior estimate, meteorological covariate and the difference between them.

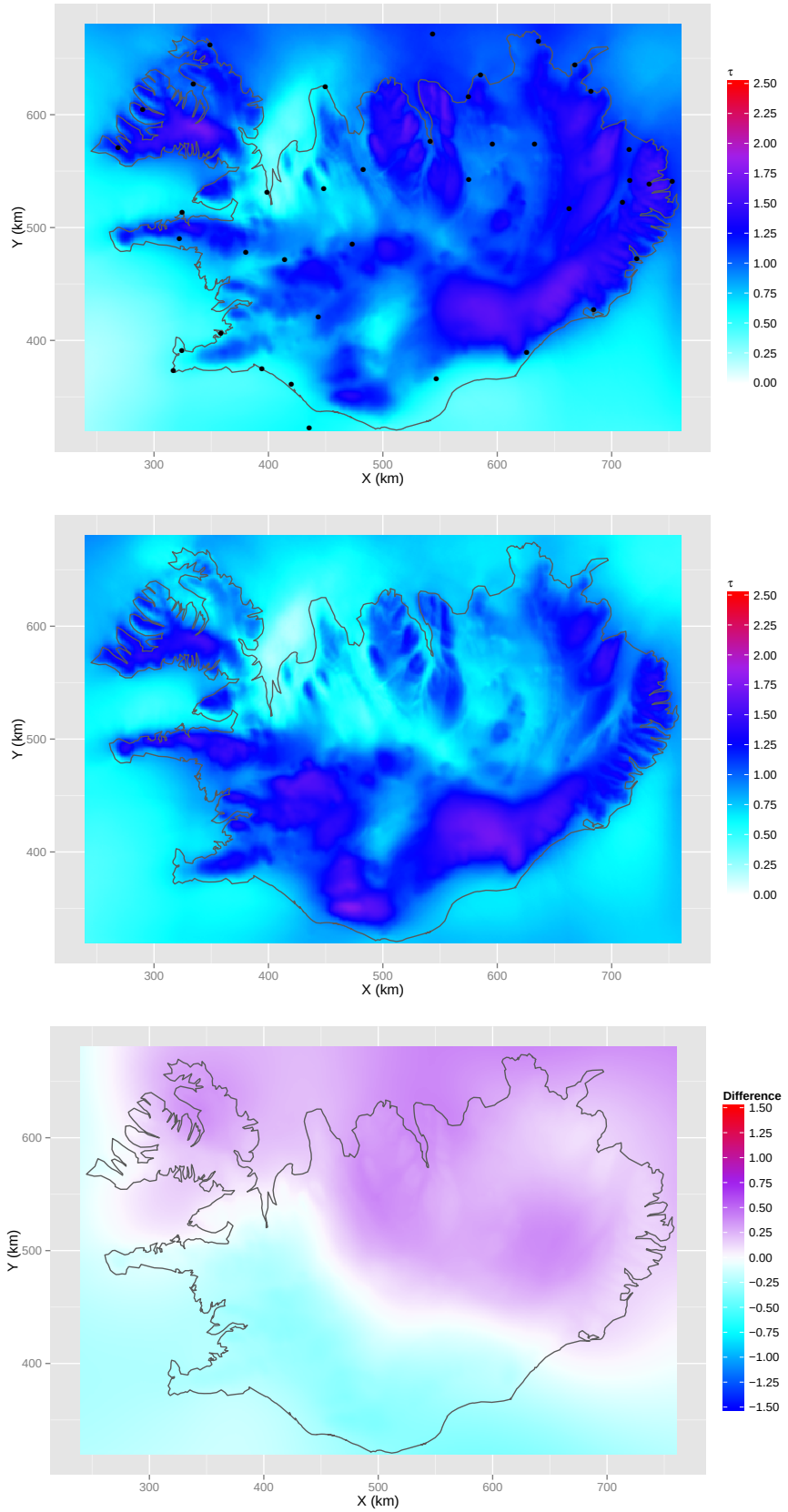

Figure 4.22: Spatial field for $\vec{\tau}$ in April. Posterior estimate, meteorological covariate and the difference between them.

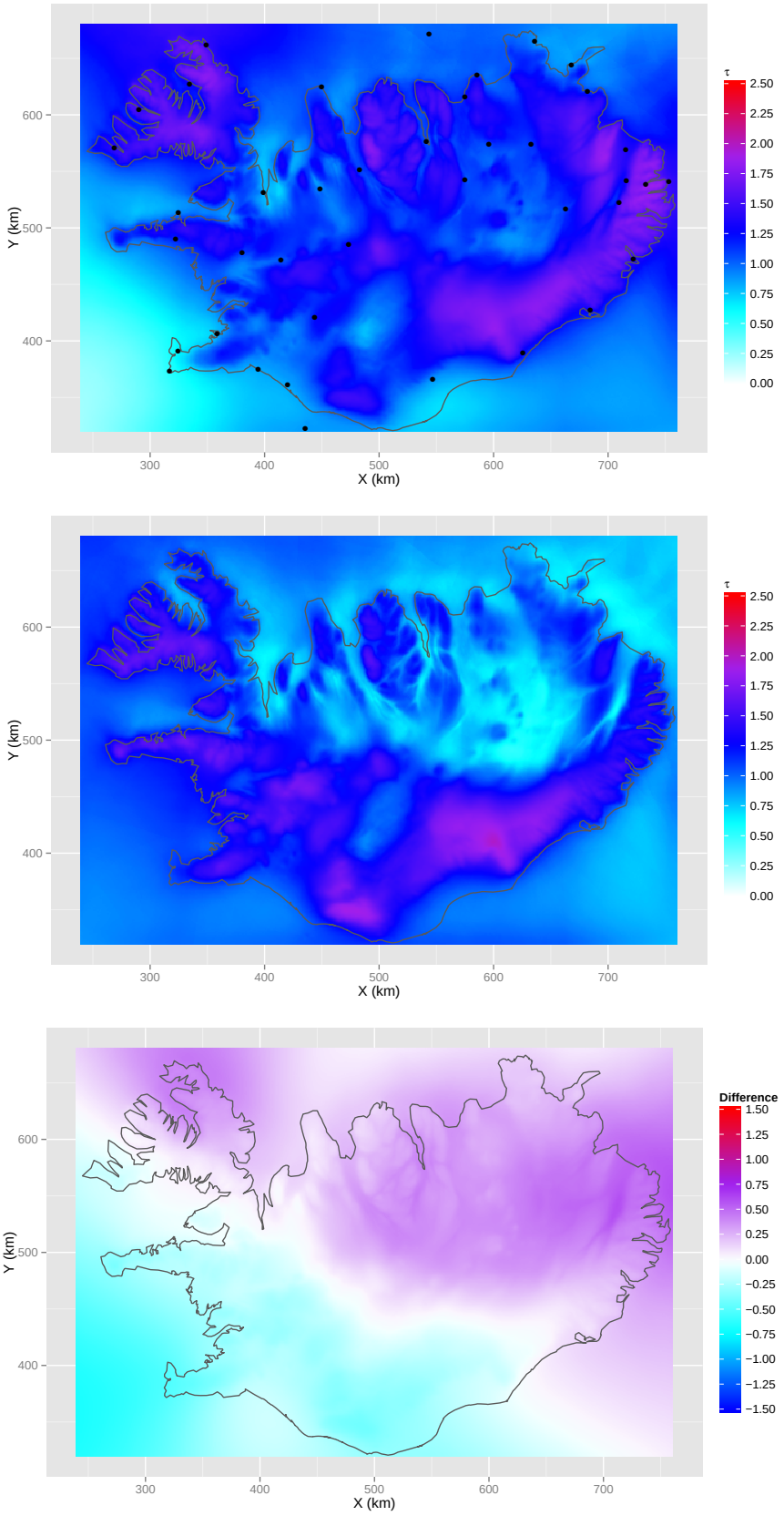

Figure 4.23: Spatial field for $\vec{\tau}$ in May. Posterior estimate, meteorological covariate and the difference between them.

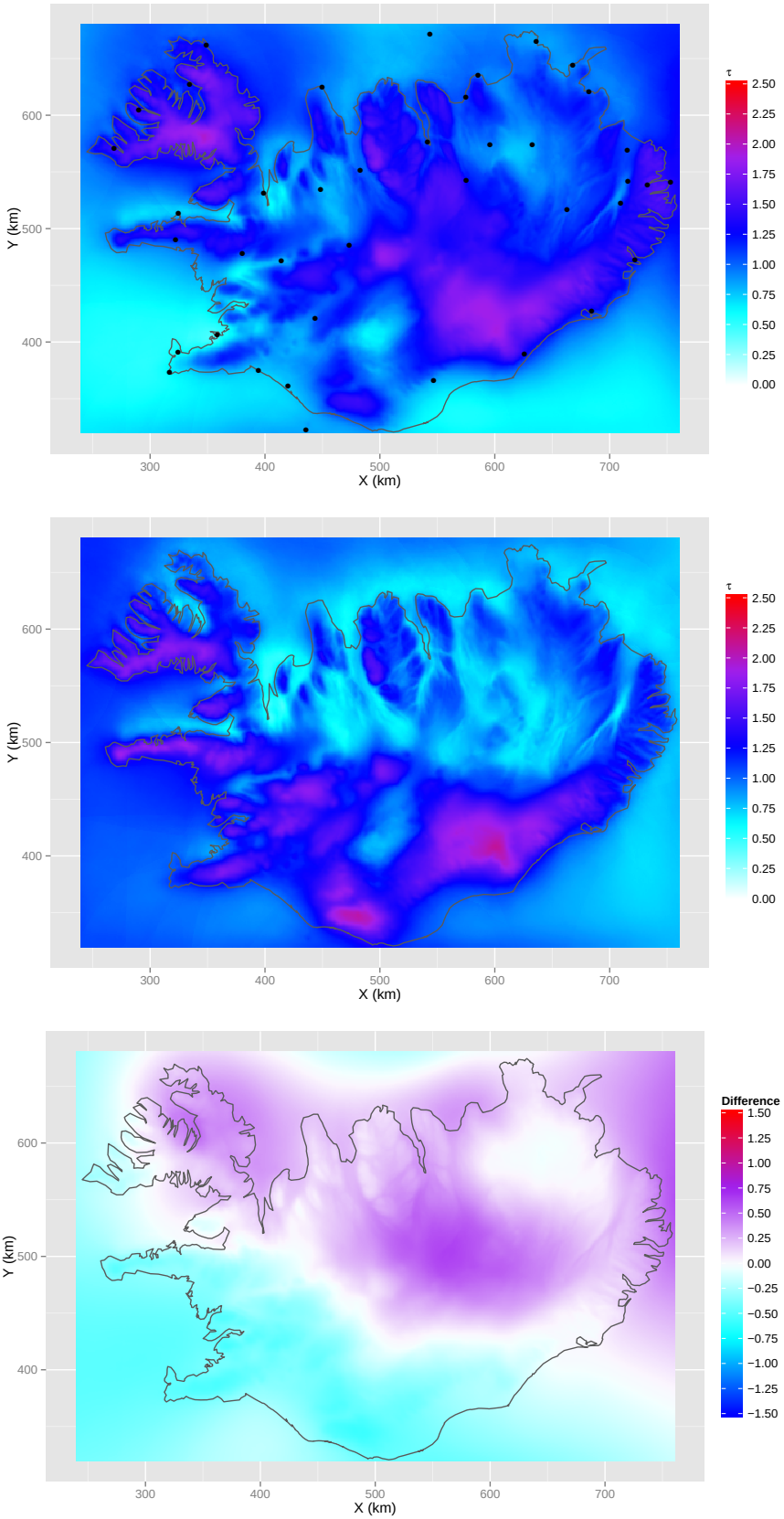

Figure 4.24: Spatial field for $\vec{\tau}$ in June. Posterior estimate, meteorological covariate and the difference between them.

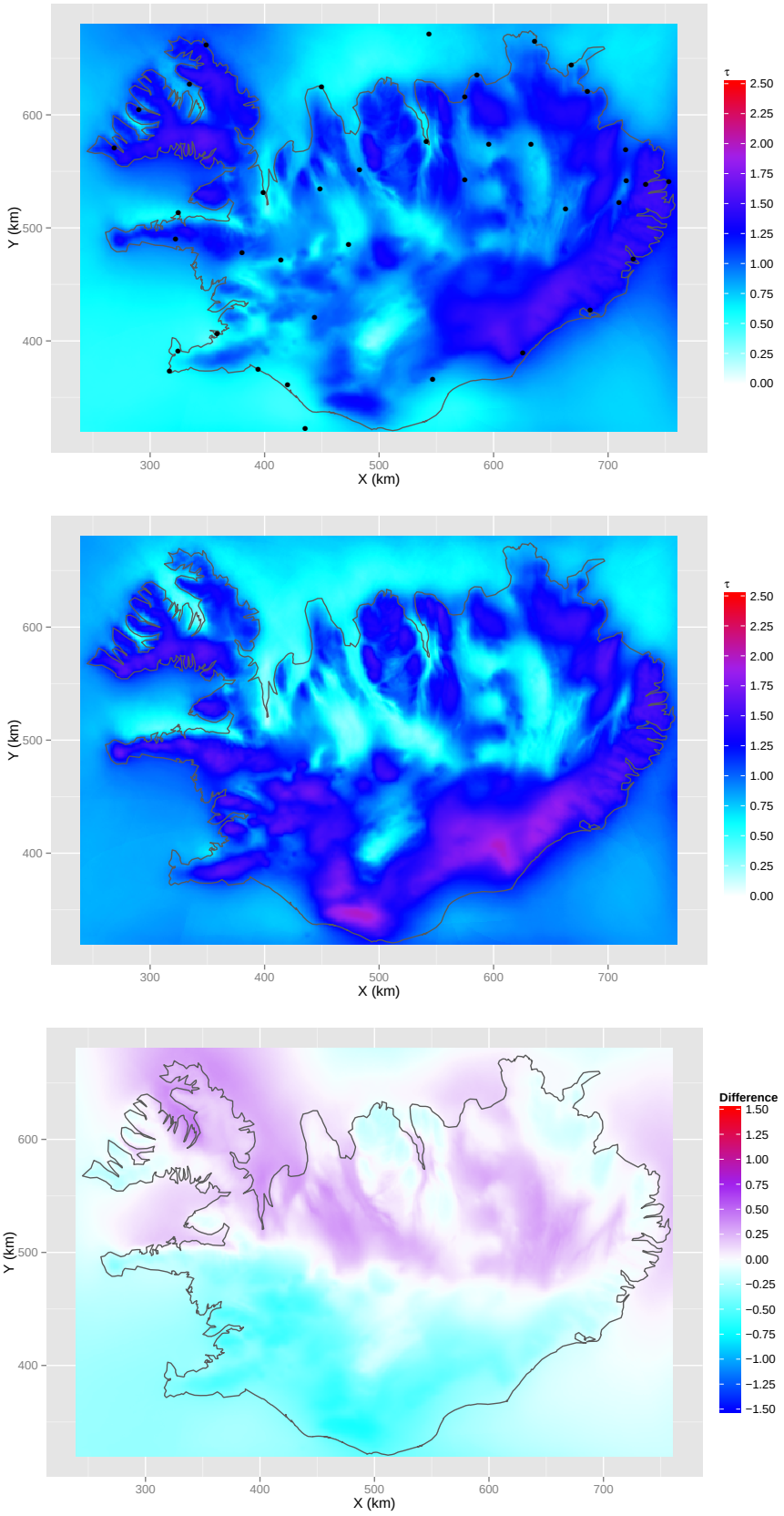

Figure 4.25: Spatial field for $\vec{\tau}$ in July. Posterior estimate, meteorological covariate and the difference between them.

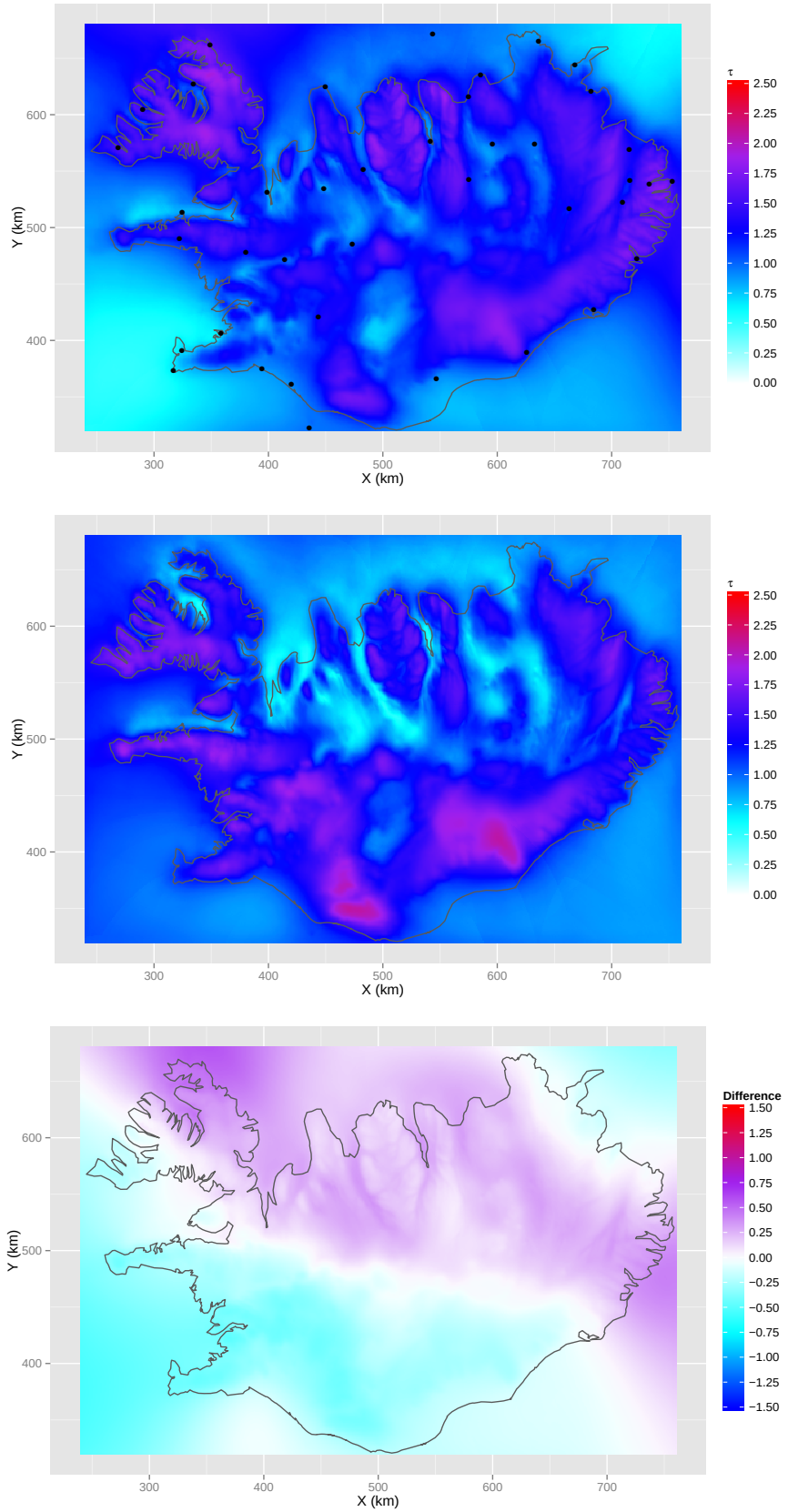

Figure 4.26: Spatial field for $\vec{\tau}$ in August. Posterior estimate, meteorological covariate and the difference between them.

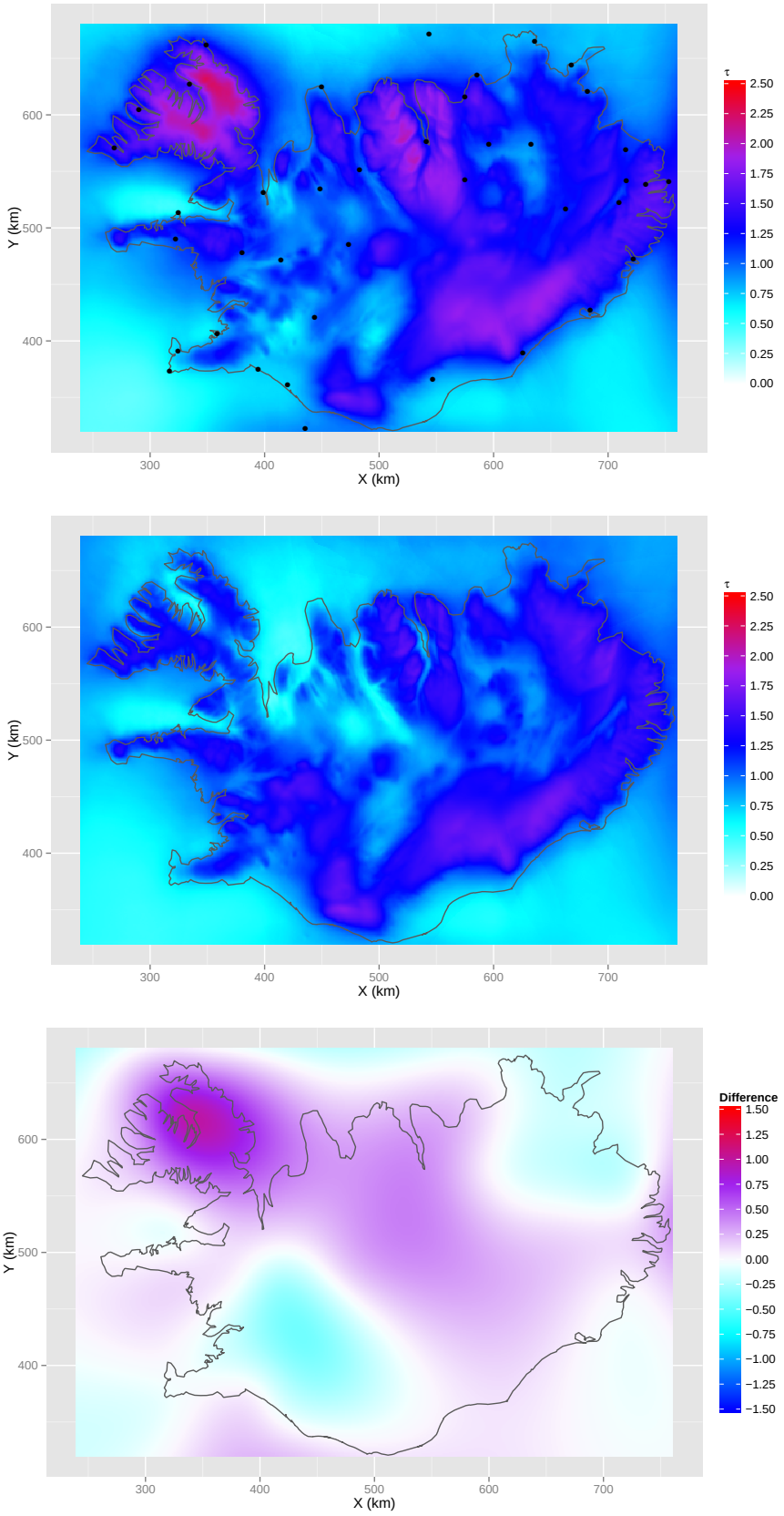

Figure 4.27: Spatial field for $\vec{\tau}$ in September. Posterior estimate, meteorological covariate and the difference between them.

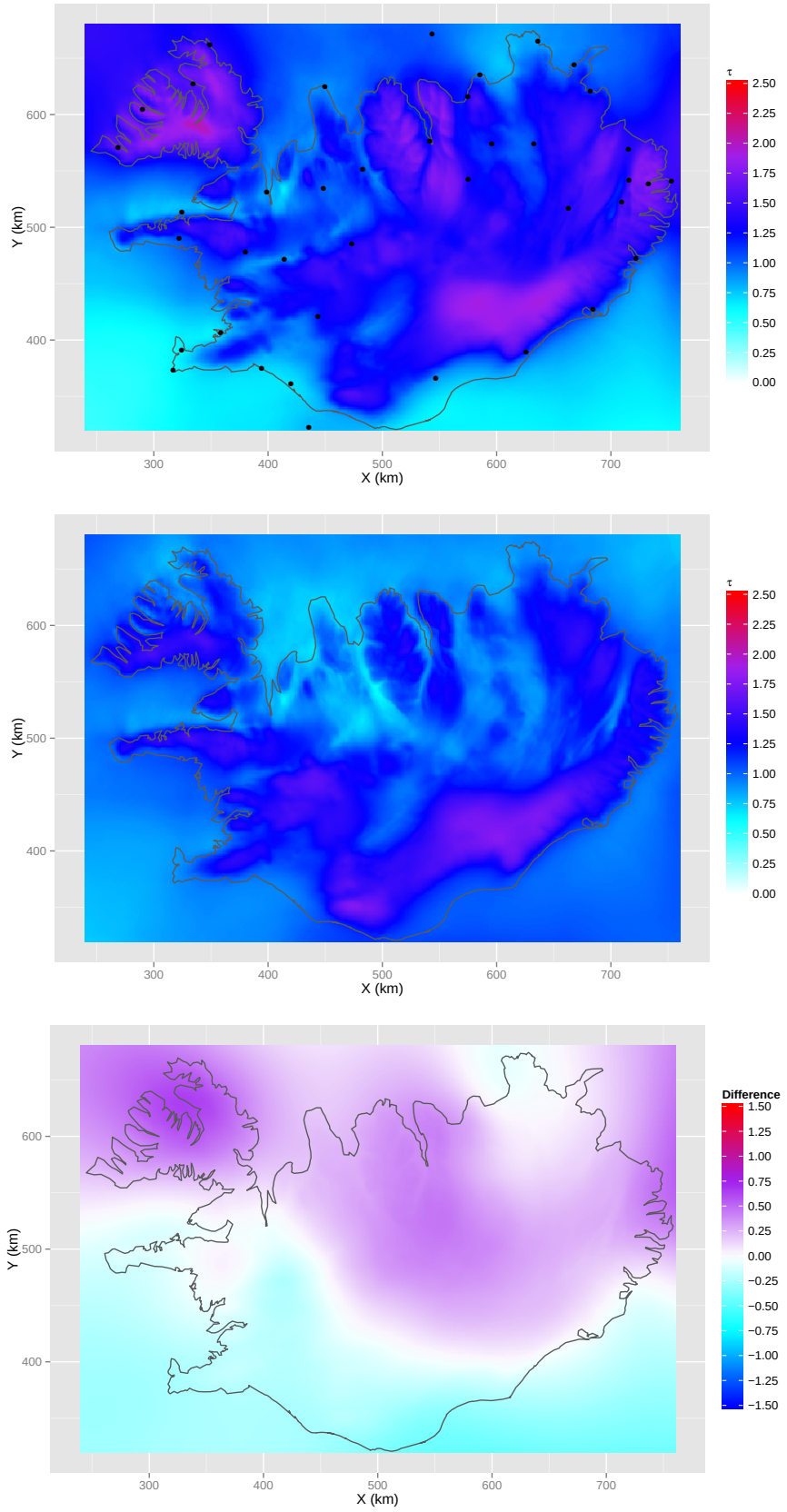

Figure 4.28: Spatial field for $\vec{\tau}$ in October. Posterior estimate, meteorological covariate and the difference between them.

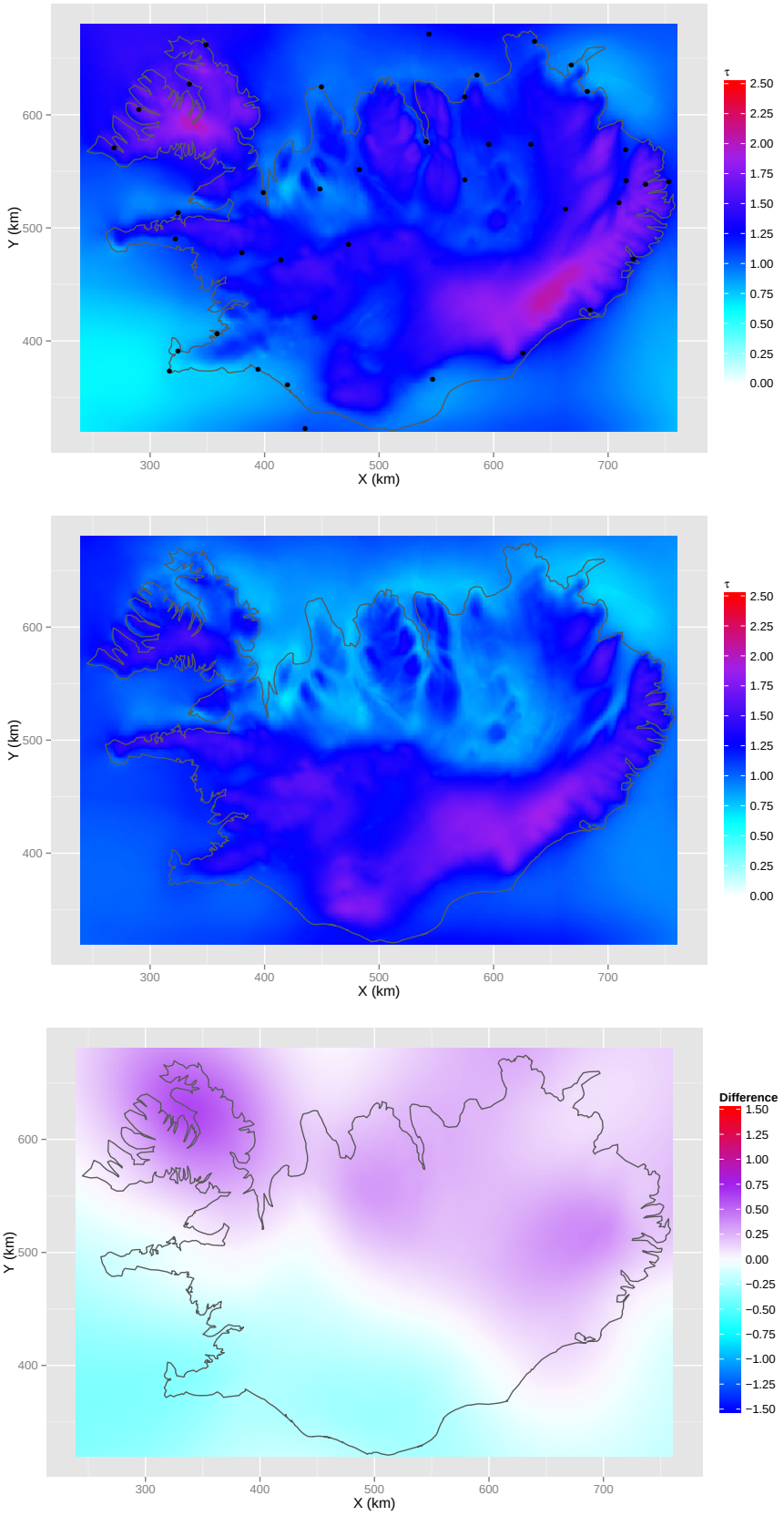

Figure 4.29: Spatial field for $\vec{\tau}$ in November. Posterior estimate, meteorological covariate and the difference between them.

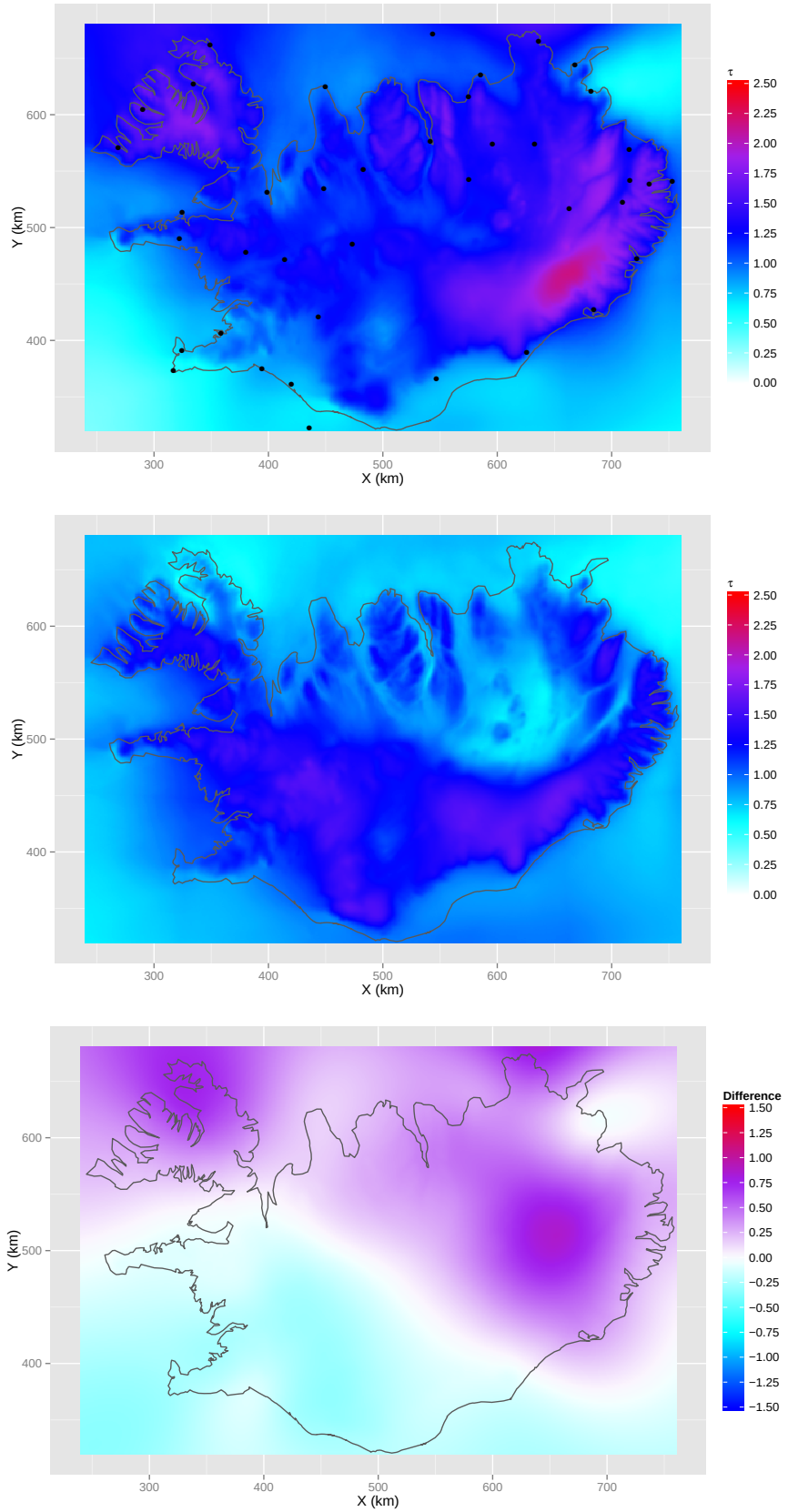

Figure 4.30: Spatial field for $\vec{\tau}$ in December. Posterior estimate, meteorological covariate and the difference between them.

4.2 Discussion

In this section we will discuss the model output, figures and tables. Since the results for all the months are quite similar, we will discuss the main model aspects for all the months together instead of discussing each month separately. First we shall discuss the spatial mean field, then the spatial field for $\vec{\boldsymbol{\tau}}$ and finally the temporal process $\vec{\boldsymbol{\gamma}}$ along with the parameters ψ and κ.

The spatial field for the mean monthly precipitation, $\vec{\boldsymbol{\alpha}}$, is controlled by three parameters, β, σ_α^2 and ϕ. β is the effect of the meteorological covariate and controls the mean of the field. σ_α^2 and ϕ control the amplitude and range of the spatial correlation, respectively, of the mean zero field $\vec{\boldsymbol{\varepsilon}}_\alpha$, where $\vec{\boldsymbol{\alpha}} = X\beta + \vec{\boldsymbol{\varepsilon}}_\alpha$. Firstly, it can be seen in Figures 4.7 to 4.18 that there is more precipitation during the winter than the summer. The values for the parameter β in the spatial mean field can be seen in Tables 4.4 and 4.5. The interpretation of the values of β is such that when it is larger than 1.0, the meteorological model is underestimating the mean monthly precipitation, and when β is smaller than 1.0 the model is overestimating the mean monthly precipitation. As can be seen in the tables the point estimate for β (the posterior median) is generally above the value 1.0. It is largest 1.08 in February, November and December but smallest 0.98 in August and 0.99 in May. In case of the winter months November to February the value 1.0 is not contained in the 95% posterior interval for β, and for March, April, September and October the value 1.0 is exactly at the 2.5% quantile. On the other hand, for the summer months, May to August the point estimate is very close to the value 1.0 (1.01 for both June and July) and the 95% posterior interval for all these months contain the value 1.0. This means that the output from the meteorological model describes the mean precipitation quite well for the months May to August. However, it underestimates the precipitation in the winter months, November to February and also gives a slight underestimation for the months in spring and autumn. The point estimate for the parameter σ_α^2 ranges from 1.2 in May to 2.64 in January. Its value is higher for the winter months than the summer months. This shows that not only is the mean monthly precipitation greater during the winter than the summer, but also the variability of the mean zero field, $\vec{\boldsymbol{\varepsilon}}_\alpha$, is greater as well. The estimate for the parameter ϕ is very similar for all the months. Its value ranges from 0.46 to 0.5. This value is quite low and maybe not what one might expect. It means that the range of the spatial correlation of the mean zero field, $\vec{\boldsymbol{\varepsilon}}_\alpha$, is very short. One might have expected that this range would turn out to be larger, since sites close in space should exhibit a similar error term. This small value can probably be explained by the quality of the covariate being used.

The spatial field for $\vec{\boldsymbol{\tau}}$ describes the variability of the monthly precipitation. It has the same structure as the field for $\vec{\boldsymbol{\alpha}}$, the mean monthly precipitation, and is controlled by the parameters η, σ_τ^2 and ρ. The parameter η in the spatial field

for $\vec{\tau}$ is the effect of the meteorological covariate. The same goes for η as it does for β, i.e. when it is larger than 1.0, the meteorological model is underestimating the variability of the monthly precipitation and when η is smaller than 1.0 it is overestimating the variability of the monthly precipitation. As can be seen in Tables 4.4 and 4.5 the point estimate for η (the posterior median) differs quite a lot from 1.0 with values ranging from 0.7 to 1.23. The fact that the range of values for η is larger than for β shows that the meteorological model does not give as good a result when describing the variability of monthly precipitation as it does in case of the mean. The posterior 95% intervals for η are wider than those for β. That is to be expected since the parameters $\vec{\tau}$, and therefore η, are more difficult to estimate than the parameters $\vec{\alpha}$ and β. The 95% posterior intervals that do not include the value 1.0 are those of January and May to August. In January the meteorological model is underestimating the variability and for the summer months May to August it is overestimating the variability. The estimates for the parameter σ_τ^2 are quite low. They range from 0.05 to 0.12 between the months. They are quite similar for all the months and there is no clear pattern to be seen e.g. between summer and winter. These low values indicate that the amplitude of the mean zero field $\vec{\varepsilon}_\tau$, where $\vec{\tau} = V\eta + \vec{\varepsilon}_\tau$, is relatively small. The final estimates of the parameter ρ can be found in Table 4.1. They differ somewhat between the months and range from around 140 km to around 309 km. Note that the Matérn smoothness is $\nu = 1.5$ and the spatial correlation at 3ρ is around 0.2. The values obtained for ρ are much higher than those for the parameter ϕ in the spatial field for $\vec{\alpha}$. There is of course some difference in how these estimates are obtained. ϕ was included in the Gibbs sampler but ρ was estimated separately based on numerical methods and the posterior estimates of other parameters. Nevertheless, we are reasonably confident in these estimates for ρ. These high values that are obtained for ρ indicate that the range of the spatial correlation is quite long in the mean zero field $\vec{\varepsilon}_\tau$. This means that sites close to one another are highly correlated when it comes to the error term in the $\vec{\tau}$ field for monthly precipitation.

The temporal process $\vec{\gamma}$ serves to raise or lower the spatial field $\vec{\alpha}$ within each month to best describe the mean of the monthly precipitation. As can be seen in Figures 4.5 and 4.6 it's not identically zero and therefore it appears to have clear purpose in describing the data. Figure 4.6 show the parameter $\vec{\gamma}$ for all the months taken together in correct temporal order. The figure shows no apparent temporal trend in monthly precipitation over the years 1958-2006. However, by looking at Figure 4.5 it can be seen that some months appear to be more stable when it comes to monthly precipitation than others. In April, June, July and September there is little fluctuation in the parameters $\vec{\gamma}$ meaning that these months are quite stable. This means that the spatial field for mean monthly precipitation, $\vec{\alpha}$, describes it quite well, in these months, with relatively small adjustments made over the years. Other months exhibit more fluctuation, especially February and October. In these months the monthly precipitation seems to vary greater over the years than in the more stable months.

The final estimates of the parameter ψ can be found in Table 4.1. They range from around 85 km. up to around 150 km. between months. The parameter ψ is the range parameter of the Matérn spatial field describing the correlation at the data level. This spatial field and the parameter $\vec{\boldsymbol{\gamma}}$ describe the deviation from the expected field $\vec{\boldsymbol{\alpha}}$. Note that, as with ρ, the Matérn smoothness is $\nu = 1.5$ and the spatial correlation of the data at 3ψ is around 0.2.

Table 4.1 also shows the final estimates of the parameter κ. They are quite similar between months with values ranging from 0.014 to 0.03. The purpose of the parameter κ is to account for measurement error. The values of the estimates indicate that the standard deviation of the measurement error is from around 5.2% to around 28.8% (calculated according to $100\sqrt{\kappa}e^{\tau_i}/\sqrt{E_{ii}}$) of the point estimate for the median precipitation at each site.

5 Conclusions

5.1 Modeling conclusions

In this section we will discuss the model being used in this project and reflect on its structure, application, and pros and cons.

The model used in this project is not a standard one and developed for the two data sources described here. It incorporates some things not widely used when modeling similar kinds of data, but which are however intuitive to use. The nonstandard parts of the model are the use of data level correlation, as well as measurement error, spatially varying variance and last but not least the use of an output from a meteorological model as a covariate. There were of course some simplification that had to be made. Firstly, not all of the parameters could be included in the Gibbs sampler. Secondly, it was decided to analyze each month separately, which makes the analysis simpler.

Data level correlation structure is a natural modeling aspect when it comes to meteorological data taken at sites located close to one another. With Iceland being relatively small, most weather systems passing by affect the whole country, therefore it is natural to assume that measurements of phenomena such as monthly precipitation are quite correlated over a large part of Iceland. Including a correlation parameter in the model does come with its problems though. The dimensionality of the data is quite large. In this project we examined 40 sites over 49 years (assuming no missing data) which adds up to 1960 data points for each month. Including a correlation parameter at the data level in the Gibbs sampler would have resulted in manipulations of rather large matrices in each iteration of the sampler. Even though these matrices are relatively sparse these manipulations are quite time consuming and it may be next to impossible to run the Gibbs sampler for long enough in a reasonable amount of time. This problem could of course be avoided by paralleling the sampler and running it on multiple processors but that task was to big a feat for this project. Therefore we opted to evaluate this parameter outside of the sampler according to the method described in Section 3.4. Even though we were not able to include this parameter in the sampler we feel that the model is better with it than without and we are quite confident in its estimates.

Including a measurement error is a natural modeling assumption. At first the measurement error parameter, κ, was not included in the model. In that setup the parameter $\vec{\tau}$ was ill-behaved and converged slowly to its target distribution. After introducing κ in the model things ran much more smoothly and $\vec{\tau}$ started to converge at a reasonable rate. It seems as in the absence of κ, $\vec{\tau}$ was trying to take on its role with little success.

The incorporation of site specific variance is natural when modeling precipitation. Modeling the variance with a spatial field with some correlation structure is even more natural since one would expect sites close to one another to exhibit similar characteristics when it comes precipitation. The problem with including such a field in the model is mainly the way it enters into the likelihood at the data level. This parameter does not have a known conjugate prior distribution and therefore a Metropolis Hastings step is needed for it in the sampler. This parameter is also multidimensional (40 dimensional in this project) and thus, in the Metropolis Hastings algorithm, one needs to update a vector and not just a univariate parameter. This makes it more difficult to find a usable proposal distribution, one which moves enough around the target mode and has a reasonable acceptance rate, and it is very probable that independent proposals will not work very well. The scheme that we used gave quite good results but it needed a separate algorithm to be used to estimate the covariance matrix of the target distribution. Also, since the Metropolis Hastings algorithm is used for this parameter, its sample tends to be highly correlated which means that it converges slower than other parameters, which results in the need to run the sampler with longer chains which can be time consuming. Also, one can run into trouble with the hyperparameters for the spatial variance field, like we did with the parameter ρ. The range parameter, ρ, has no known conjugate priors so it too needed to be sampled using a Metropolis Hastings algorithm. Including ρ in the sampler would have resulted in the need for some sort of block updating scheme since its draws would have been highly correlated with the draws of $\vec{\tau}$. This was outside the scope of this project and therefore we opted for other methods. Estimating ρ outside of the Gibbs sampler makes it harder to judge the accuracy of the estimate, but we are reasonably confident in its estimated values.

Using an output from an advanced meteorological model as a covariate in a statistical model has two benefits. Firstly, the covariates extracted from such an output are probably quite good, as turned out to be the case in this project. Meteorological models usually incorporate other covariates such as height above sea level, wind direction and topological factors that affect the weather and therefore there is less need to include such other covariates in the statistical model. Also since the meteorological model takes these factors into account one gets a good covariate, which incorporates the necessary factors, without having a to have detailed knowledge of meteorology.

The decision to analyze each month separately was taken quite early in the develop-

ment of this project. This simplified the problem at hand. Firstly the dimensionality of the data was reduced extensively and much fewer parameters were needed in the model, as there was no need to include a temporal correlation structure or a seasonal component. Despite this simplification, we feel that the analysis in this project is quite extensive and that the model performs well in describing the monthly precipitation and its various aspects.

Overall our conclusion is that the model presented in this project is quite detailed and accurate in describing the data at hand. It incorporates most of which one can imagine is needed in the analysis of monthly precipitation when it is analyzed separately for each month. However using such a model is not straight forward and results in various difficulties and problems to be solved. The time it takes to run the model is also quite long, one run of the Gibbs sampler for all the months takes around 13 hours. Thus, the setup of this model is not ideal if one needs fast computations.

5.2 Comparison with the meteorological model

In this section we will discuss the comparison between the output from our statistical model and the output from the meteorological model, which can be seen in Chapter 4.

The output from the meteorological model discussed in Section 2.3 gives a good description of the mean monthly precipitation. This can be seen in Figures 4.7 to 4.18. The estimated median precipitation fields seem to have a very similar shape as those for the meteorological covariate and generally the meteorological model is in fair agreement with the data. The values of the parameter β, discussed in Section 4.2, show that the output from the meteorological model describes the mean precipitation very well for the months May to August. However, it underestimates the precipitation in the winter months, November to February, and also gives a slight underestimation for the months in spring and autumn, though there is less discrepancy there than in the winter months. It can also be seen in Figures 4.7 to 4.18 that the underestimation of the model is greater at the more extreme parts of the field i.e. where there is more precipitation. Recall from Chapter 4 that the estimated values for ϕ were quite small. This result further indicates the accuracy of the meteorological model as an estimate for the mean monthly precipitation, as these small values can probably be explained by the fact that the meteorological covariate describes the $\vec{\boldsymbol{\alpha}}$ field very well, i.e. the field can be adequately estimated by the covariate corrected by the factor β and there is little need for further spatial correction.

The output from the meteorological model does not describe the variability of the

monthly precipitation as well as it describes the mean. This can be seen in Figures 4.19 to 4.30. When looking at the difference between the output from our statistical model and the meteorological model one can see that there is always some discrepancy and this generally is displayed in big chunks. For most of the months the meteorological model underestimates the variability in the North-East part of Iceland and overestimates it in the South-West corner. The values of the parameter η, discussed in Section 4.2 also show the inaccuracy of the meteorological models estimation of the variability. Both the range of the values of η, across the months, is greater than that of β and also the 95% posterior intervals are wider, although that can probably be attributed to the fact that η is more difficult to estimate than β. The values of η show that in January the meteorological model is underestimating the variability and for the summer months May to August it is overestimating the variability. Overall, the posterior estimates and intervals for η show that the meteorological model does not describe the variability in monthly precipitation as well as it does the mean. The values of the parameter ρ in the spatial field for $\vec{\boldsymbol{\tau}}$ are much higher than those of ϕ in the spatial field for $\vec{\boldsymbol{\alpha}}$. This is further indication of the inaccuracy of the meteorological covariate in describing the $\vec{\boldsymbol{\tau}}$ field. The reason for these large values of ρ is probably that the meteorological covariate being used is not adequately describing the $\vec{\boldsymbol{\tau}}$ field and thus the discrepancy is corrected with error field with high spatial correlation.

5.3 Future studies

In conjunction with previous sections it is natural to list a few extensions of the model to be considered for future studies.

What first comes to mind is to try to incorporate all the parameters in the Gibbs sampler, i.e. to include the range parameter ψ of the data level correlation, the measurement error parameter κ and also the range parameter ρ for the correlation of $\vec{\boldsymbol{\tau}}$. Including ψ and κ will result in manipulations of large matrices in each iteration of the sampler and therefore a way has to be found to do this rapidly to get a reasonable runtime. Including ρ will most likely result in some updating problems of ρ and $\vec{\boldsymbol{\tau}}$ and therefore some specialized updating scheme will be needed.

In this project we opted for using the same Box-Cox parameter λ to transform all the data to normality. Perhaps it gives better results to use different λ for different months and maybe even also for different sites. It could also be possible to include the parameter λ in the Gibbs sampler.

By looking at the figures for the spatial field for $\vec{\boldsymbol{\tau}}$ (see Figures 4.19 to 4.30), especially examining the difference between the output from our statistical model and

the output from the meteorological model, it seems that it might be interesting to include longitude and latitude, or other grid variables, as covariates, together with the meteorological covariate, for the parameter $\vec{\tau}$.

An obvious extension of the analysis in this project is to analyze all the months together in correct temporal order. That would result in a spatio-temporal model with much higher dimensionality, which on its own is problematic. For such an analysis one would need a seasonal component to adjust the mean of the temporal process. This seasonal component could be modeled with trigonometric functions or perhaps B-splines. A correlation structure in the temporal process would be needed as well. An analysis, of that kind, would result in many additional parameters for the model, needed to describe the temporal process. One would probably have to find a clever way to estimate all these parameters since it might be quite difficult to include them all in a Gibbs sampler.

Bibliography

[1] Philippe Crochet, Tómas Jóhannesson, Trausti Jónsson, Oddur Sigurðsson, Helgi Björnsson, Finnur Pálsson, and Idar Barstad. Estimating the spatial distribution of precipitation in iceland using a linear model of orographic precipitation. *Journal of Hydrometeorology*, 8(6):1285–1306, 2007.

[2] Christopher K Wikle, L Mark Berliner, and Noel Cressie. Hierarchical bayesian space-time models. *Environmental and Ecological Statistics*, 5(2):117–154, 1998.

[3] Pingping Xie and Phillip A Arkin. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *Journal of climate*, 9(4):840–858, 1996.

[4] Claudia Tebaldi and Bruno Sansó. Joint projections of temperature and precipitation change from multiple climate models: a hierarchical bayesian approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1):83–106, 2009.

[5] Pinhas Alpert and Haim Shafir. Meso-scale distribution of orographic precipitation: Numerical study and comparison with precipitation derived from radar measurements. *J. Appl. Meteor*, 28(10):1105–1117, 1989.

[6] Pingping Xie and Phillip A Arkin. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society*, 78(11):2539–2558, 1997.

[7] Christine L Goodale, John D Aber, and Scott V Ollinger. Mapping monthly precipitation, temperature, and solar radiation for ireland with polynomial regression and a digital elevation model. *Climate Research*, 10:35–49, 1998.

[8] Christopher Daly, Ronald P Neilson, and Donald L Phillips. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of applied meteorology*, 33(2):140–158, 1994.

[9] Daniel Cooley, Douglas Nychka, and Philippe Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.

[10] George EP Box and David R Cox. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 211–252, 1964.

[11] Thomas W Anderson and David A Darling. Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, 23(2):193–212, 1952.

[12] Michael A Stephens. Tests based on edf statistics. *Goodness-of-Fit Techniques, RB D'Agostino and MS Stephens, Eds. Marcel Dekker, New York and Basel*, 1986.

[13] Elies Kouider and Hanfeng Chen. Concavity of box-cox log-likelihood function. *Statistics & probability letters*, 25(2):171–175, 1995.

[14] Peter Guttorp and Tilmann Gneiting. Studies in the history of probability and statistics xlix on the matern correlation family. *Biometrika*, 93(4):989–995, 2006.

[15] Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. *Bayesian data analysis*. Chapman & Hall/CRC, 2004.

[16] Birgir Hrafnkelsson, Jeffrey S Morris, and Veerabhadran Baladandayuthapani. Spatial modeling of annual minimum and maximum temperatures in iceland. *Meteorology and Atmospheric Physics*, 116(1-2):43–61, 2012.