NTNU

Norwegian University of
Science and Technology

A Bayesian Model for Area and Point Predictions

A Case Study of Predictions of Annual
Precipitation and Runoff in the Voss Area.

Thea Julie Thømt Roksvåg

Master of Science in Physics and Mathematics
Submission date: June 2016
Supervisor: Ingelin Steinsland, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

In this work we perform predictions of annual precipitation and runoff by spatial interpolation. For this purpose, we utilise both point observations of precipitation and/or area observations of runoff from several years. We suggest a statistical model for annual precipitation and runoff consisting of two spatial terms: One spatial term that is common for all years which models the climatology in the area of interest, and one spatial term for year-to-year variation. The model is set up as a Bayesian hierarchical model of three levels, and we use informative priors based on information from the available observations. A stochastic partial differential equation (SPDE) approach to spatial modelling is used to make inference and predictions less computationally expensive. The model is implemented by using the R-package R-INLA, and we demonstrate how R-INLA can be used for making predictions and drawing inference from a model based on both point observations (e.g of precipitation) and area observations (e.g of runoff). The statistical model is tested through a case study of catchments located around Voss in Norway and through simulation studies. The main focus is on the predictive performance. In particular we explore how the predictive performance is affected by having a spatial varying climate effect in the model. We find that the spatial predictions of runoff and precipitation often are uncalibrated if the spatial differences in the observed annual precipitation are stable from one year to another. The consequence of this model property is that an observation design that produces accurate predictions one year, also will produce accurate predictions other years. Further we compare the predictive performance for annual runoff when using observation samples consisting of (1) only observations of runoff, (2) only observations of precipitation and (3) observations of both runoff and precipitation. The results from the simulation studies did not favour one of the observation types (runoff and/or precipitation), and both observation types can produce accurate predictions of annual runoff depending on the underlying climatology. For the real dataset we saw that observation samples of only runoff produced the most accurate predictions. Observation samples of only precipitation were not suitable for runoff predictions for the real dataset and led to large biases between the true observations and the predicted values.

Sammendrag

I denne oppgaven utfører vi prediksjoner av årsnedbør og årsavrenning ved bruk av romlig interpolasjon. Vi benytter vi oss av både punktobservasjoner av nedbør og arealobservasjoner av avrenning fra flere år. Vi foreslår en statistisk modell for årsnedbør og avrenning som består av to romlige effekter: En romlig effekt som er felles for alle år og som modellerer klimaet i området, og en romlig effekt som forklarer de årlige nedbørsvariasjonene. Modellen er satt opp som en Bayesiansk hierarkisk modell med tre nivåer og vi bruker informative priorfordelinger basert på informasjon fra datasettet. En stokastisk partiell differensialligning (SPDE) er brukt for å gjøre modellen mindre beregningskrevende. Modellen er implementert ved å bruke R-pakken R-INLA, og vi demonstrerer hvordan R-INLA kan brukes for å gjennomføre prediksjoner og inferens basert på både punktobservasjoner (av for eksempel nedbør) og arealobservasjoner (av for eksempel avrenning). Den statistiske modellen testes ved bruk av et ekte datasett fra Vossområdet og gjennom simuleringsstudier. Hovedfokuset i oppgaven er å vurdere modellens evne til å gjennomføre prediksjoner. Spesielt er vi interessert i å undersøke hvordan prediksjonene påvirkes av at vi har inkludert en romlig effekt for klima i modellen. Resultatene viser at de romlige prediksjonene av avrenning og nedbør ofte er ukalibrerte hvis den romlige variasjonen av årsnedbør er stabil fra ett år til et annet. Konsekvensen av dette er at et observasjonsdesign som produserer gode prediksjoner ett år, sannsynligvis vil produsere gode prediksjoner for andre år. Videre sammenligner vi avrenningsprediksjonene som oppnås ved å bruke observasjoner av (1) kun avrenning, (2) kun nedbør og (3) både avrenning og nedbør. Resultatene fra simuleringsstudiene favoriserer ikke noen av observasjonstypene (nedbør og/eller avrenning), og begge observasjonstypene kan føre til gode prediksjoner av årsavrenning avhengig av det underliggende klimaet. For det ekte datasettet fra Voss så vi at de beste avrenningsprediksjonene oppnås når man kun benytter seg av avrenningsobservasjoner. Å utføre avrenningsprediksjoner basert på kun nedbørsobservasjoner fungerte dårlig for de ekte dataene og førte store forskjeller mellom de ekte observasjonene og de predikerte verdiene.

Preface

This master thesis is the product of my last semester of five years of study at the Norwegian university of science and technology. The work was done during the spring 2016 at the Department of Mathematical Sciences and completes my Master of Science Degree in Applied Physics and Mathematics.

I would like to thank my supervisor Ingelin Steinsland for excellent guidance through the work with the master thesis and the preliminary project. I would also like to thank Kolbjørn Engeland at NVE for providing hydrological data and for helpfully answering questions regarding the dataset. Finally, I would like to thank my friends at FysMat for making my five years at NTNU great.

Thea Roksvåg
Trondheim, June 2016

Contents

1. Introduction

In Norway statistical models for hydrological forecasting are demanded. Approximately 99% of the electricity generation in Norway comes from hydro-power production (Statkraft, 2016), and accurate forecasts of precipitation and runoff with uncertainty estimates make it easier for the hydro-power companies to schedule the hydropower production during the year. Hydrological forecasting models also contribute to improved flood warnings.

For this purpose, it is important to account for the greatest runoff generation that can be expected in a catchment over a longer or shorter time period. In Benestad et al. (2012) it is shown that there is a close relationship between the mean amount of precipitation and precipitation extremes at a specific location. Further, it is reasonable to assume that the mean amount of runoff generated in a catchment during a specific time period also is closely related to greatest amount of runoff generation that can be expected within a time interval. If this is the case, historical data can help us to gain knowledge of runoff extremes. However, a problem in hydrology is that in many catchments there are no observations of runoff and/or precipitation. This makes it challenging to gain knowledge of runoff and precipitation amounts in these areas, and is known as "the problem of ungauged basins" (Blöschl et al., 2013).

Motivated by "the problem of ungauged basins" we here present a statistical model for annual precipitation and runoff, and demonstrate how both observations of precipitation and/or observations of runoff can be used for spatial predictions in a catchment where measurements of runoff and/or precipitation don't exist. By spatial predictions we mean that we perform predictions in space and not in time. It is common to perform spatial predictions of runoff within an ungauged basin by interpolation of rainfall data, i.e point observations of precipitation are interpolated to a continuous surface or area. Here, we explore if also runoff observations can contribute to improved spatial predictions.

The results presented in the report will be based on a case study of several catchments located in the western parts of Norway, more specific around Voss. We chose to focus the analysis on data from the Voss area because this area is known to be flood exposed. In 2014 the Voss area was hit by a severe flood which caused damage on infrastructure and property (NRK, 2014). Hydrological forecasting models are therefore of high relevance in the area of interest and can contribute to better area planning and preventing severe damage on property and infrastructure in the future.

The starting point of any statistical modelling is to make a realistic model for the process of interest which at the same time is simple enough to draw inference from with available data. Precipitation is a complicated process driven by humidity, temperature, the distance from the ocean and interaction with the topography, and simplifications must be made in the modelling.

In Norway the differences in annual precipitation are mainly caused by orographic precipitation which occurs when air rises on the windward side of a mountain. This results in large values of annual precipitation in the western parts of the country and a dryer climate in the eastern parts. Thus, the amount of annual precipitation depends on the location in space. It

is also reasonable to assume that locations that are close in space have more in common than locations far apart. In statistics, such dependency structures can be modelled by a spatial effect. In this work, we suggest a model for precipitation consisting of two spatial effects.

The first spatial effect models the climatic differences of annual precipitation in the area of interest and is referred to as the climatology of the study area. The climatology of an area is in general defined as the weather conditions, here annual precipitation, averaged over a time period. Thus, the climatic spatial effect used in our model will be common for all years. Further, we assume that it occurs spatial variations of precipitation that can not be explained by the climatology. The second spatial effect is therefore a year dependent effect that models spatial variations within a specific year. We also include a year specific intercept in the precipitation model that adjusts the annual level of precipitation

Annual runoff can be regarded as the amount of annual precipitation that flows over land as surface water instead of evaporating or being absorbed into ground water (NVE, 2002). In this work, the annual runoff is modelled as an integral of precipitation over the surface area of a catchment, minus the annual evaporation. As we see, runoff and precipitation are closely related. We can regard observations of precipitation as point observations, and runoff as area observations of the same underlying process.

In our analysis we use precipitation and runoff data from several years to make inference and predictions. In the same manner as in Ingebrigtsen et al. (2015), each year of data is regarded as an independent realization of the statistical model. This is called replicates, and it was shown in Ingebrigtsen et al. that the use of replicates made the parameter estimates more precise and reduced the estimation bias. The use of replicates is in particular important in our model because we have a spatial effect that is common for all years, i.e the climatology. Thus, the replicates are needed to gain knowledge of the climatology in the study area. In this work we use observations from 10 years, i.e 10 replicates to predict annual runoff and precipitation.

The model for annual precipitation and runoff is set up as a Bayesian hierarchical model of three levels. As a Bayesian approach is used, the model parameters must be assigned prior distributions, and we use informative priors based on information about the available observations.

Drawing inference from a Bayesian hierarchical model is traditionally done by using MCMC-methods (Gamerman and Lopes, 2006). However, MCMC-methods can be slow. In this work we will use an alternative to MCMC-methods which is integrated nested Laplace approximation (INLA) (Rue et al., 2009). INLA can be used for approximate inference and predictions if the underlying process, i.e the second level of the hierarchical model, has a Gaussian distribution. As our two spatial effects will be modelled by stationary Gaussian random fields (GRFs), this requirement is fulfilled.

The drawback of using GRFs to model the spatial effects is that models including GRFs often are slow to draw inference from. To ensure fast computations by using INLA we therefore make an additional approximation in the model. In Lindgren and Rue (2011) it is shown how a stochastic partial differential equation (SPDE) can be applied for expressing a GRF as

a Gaussian Markov random field (GMRF). There are computational benefits linked to GMRFs allowing us to make faster inference (Rue and Held, 2005), and we use the SPDE approach to approximate the two GRFs in our precipitation model.

The INLA methodology and a framework for implementing SPDE models are implemented in the R-package INLA, and the package is available at `www.r-inla.org`. One of the main contributions of this work is to demonstrate that the INLA framework is suitable for making predictions and inference with a SPDE model based on both point observations (e.g of precipitation) and area observations (e.g of runoff).

The statistical model for precipitation and runoff is evaluated by exploring its ability to make accurate predictions in the Voss area, i.e the focus is on the predictive performance of the model and not on parameter estimation. We leave out observations from our dataset and reproduce the observations by spatial predictions based on observations from other locations and catchments. We also test the model on simulated data. In our analysis we will focus on answering five research tasks (RT1 -RT5) which we present now.

In the first research task we do some experiments to explore if the spatial variations of annual precipitation around Voss are mainly explained by climatic differences or by annual spatial variability. We expect that the spatial effect that models the climatology is dominating over the annual spatial effect, and explore how this affects the spatial predictions. The first research task is formulated as follows:

RT1: How does the climatology of the study area affect the spatial predictions of runoff and precipitation?

The motivation for research task two is to show that predictions of runoff and precipitation are different from each other. To verify this assumption, we try to find an observation design that is suitable for runoff predictions in a specific catchment, but not suitable for predictions of precipitation in the middle of the catchment. The second research task is formulated as follows:

RT2: If we find an observation design suitable for runoff predictions in a specific catchment, will the same observation design be suitable for point predictions of precipitation in the interior of the catchment?

In the next three research tasks, we focus on runoff predictions. These are considered more important than predictions of precipitation as the work is motivated by "the problem of ungauged basins" and runoff extremes.

A large catchment that generates runoff to a large river, can always be divided into smaller catchments that generate runoff to smaller rivers. The dataset used in this work consists of catchments of this type, and we say that the catchments overlap. In a preliminary simulation study we saw that accurate predictions of annual runoff within a larger catchment can be obtained by using observations of annual runoff from a smaller overlapping catchment. We are interested in testing if this applies for real data as well. In research task 3, 4 and 5 we therefore perform spatial predictions of runoff by using observations of precipitation and/or observations of annual runoff from overlapping catchments. Research tasks 3-5 are formulated

as follows:

RT3: Explore how spatial predictions of annual runoff are affected by using an observation sample consisting of observations of runoff from an overlapping catchment.

RT4: Explore how spatial predictions of annual runoff are affected by using an observation sample consisting of observations of precipitation.

RT5: Explore how spatial predictions of annual runoff are affected by using an observation sample consisting of both observations of precipitation and observations of runoff from an overlapping catchment.

The different observation types used in RT3-RT5 are evaluated and compared. We emphasise that research task 5 is particularly interesting as we here combine observations of annual runoff and precipitation. We expect this observation type to produce the most accurate predictions, because in RT5 we include as much information as possible in the model.

The rest of the report is organized as follows: In Section 2, the dataset is introduced. In Section 3, we present important background theory needed to construct statistical models for precipitation and runoff. This includes introductions to GRFs, GMRFs, the SPDE approach to spatial modelling and INLA. Next, the model for annual precipitation and runoff is presented in Section 4, before the observation designs and the experiments are explained in Section 5 and 6 respectively. Finally, the results are analysed and discussed in Section 7 and 8. Code examples from INLA are included in the Appendix.

2. Data

In this section the dataset is introduced. We will not use all of the data available and will focus on data from the Voss area. However, we present the whole dataset because it illustrates the forces behind precipitation and runoff, and verifies the choice of model.

Runoff observations

Annual data for runoff from 1980 to 2013 were provided by the Norwegian Water Resources and Energy Directorate (NVE). The dataset consists of annual runoff observations from the 9 catchments displayed in Figure 2.1. The catchments are located in the western parts of Norway, more specific in Hordaland, Sogn og Fjordane and Oppland. The area of each catchment and mean values for annual runoff are displayed in Table 2.1. Here, the year 1980 starts at 1.September 1980 and ends 31.August 1981. This is called a hydrological year, and by this definition of a year me make sure that precipitation that came as snow has melted and returned to the hydrological cycle.

Data for measurement uncertainty for the runoff observations were also provided by NVE. Each observation has its own uncertainty, and the mean uncertainty for each catchment is displayed in Table 2.1. We note that the relative standard deviation is very low, but as the unit is $[10^8 m^3/year]$ we are still dealing with large quantities of water. The method used for estimating the annual runoff with corresponding observation uncertainty is described in Reitan and Petersen-Øverleir (2009).

Table 2.1: Area of the catchments $[km^2]$, mean value of observed runoff with corresponding standard deviation $[10^8 m^3/year]$, mean runoff per square meter [m/year] and years in which we have available data.

Catchment nr.	Name	Area	Mean runoff	Runoff/m^2	Mean sd	Data available
1	Målset	7.7	0.15	2.0	0.0018	1980 - 2013
2	Dyrdalsvatn	3.3	0.11	3.2	0.0019	1980-1995, 2001-2013
3	Røykenes	50.1	1.3	2.7	0.027	1980-2013
4	Fjellanger	12.8	0.316	2.4	0.0078	1995-2013
5	Haukåselva	7.4	0.13	1.8	0.0011	2007-2013
6	Svartavatn	72.4	2.2	3.1	0.040	1987-2013
7	Slondalsvatn	41.9	0.99	2.4	0.021	1983 - 2013
8	Kinne	511.4	10	2.0	0.078	1984-2013
9	Bulken	1092.0	22.4	2.1	0.33	1980-2013

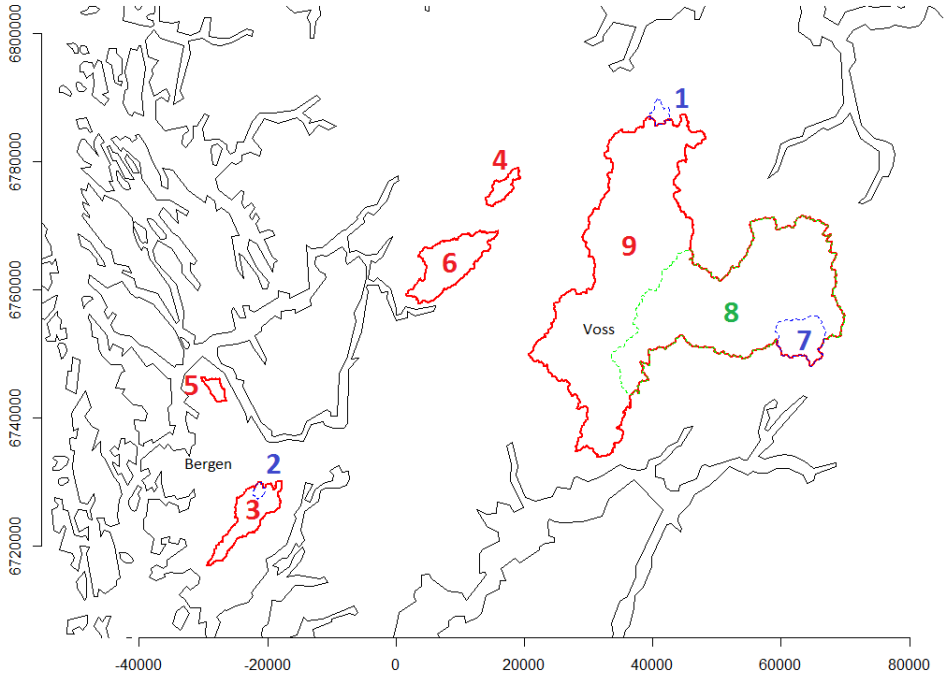


Figure 2.1: Nine catchments. The x- and y-axis show distances i meters.

Precipitation observations

Precipitation data for the same time period and area were downloaded from `www.eklima.no` which is a web site operated by the Norwegian Meteorological Institute. Only monthly observations were available, and the annual data are obtained by adding the monthly values. Data from 79 locations are available for years between 1980 and 2013, and mean values of precipitation in this time period are shown in Figure 2.2. Be aware that only some of the locations have annual data from the whole time period from 1980 to 2013. The 79 locations shown in Figure 2.2 have at least one measurement of annual precipitation between 1980 and 2013.

In Figure 2.2 we see that the highest measurements of annual precipitation are registered around a vertical line drawn from Kvamskogen to Takle. The annual values decrease as we move east or west from this line. West, close to the coast, the annual values of precipitation are high, but not as high as around Kvamskogen. East, around Reimegrend, moderate values are registered, and around Geilo we find low values of annual precipitation. This spatial pattern is apparent for all years. This can be seen in Figure 2.3 which shows the annual observations of precipitation from five locations with different climatic conditions: Bergen, Øvstedal, Brandset, Geilo and Voss. The ranking between the locations, from the highest amount of precipitation to the lowest is always Øvstedal, Bergen, Brandset, Voss and Geilo.

The spatial pattern in Figure 2.2 is caused by the topography of the area of interest. Humid oceanic winds from west hit the coast of Norway, and the mountains in the area force the air to elevate. This phenomenon is called orographic precipitation and results in large values

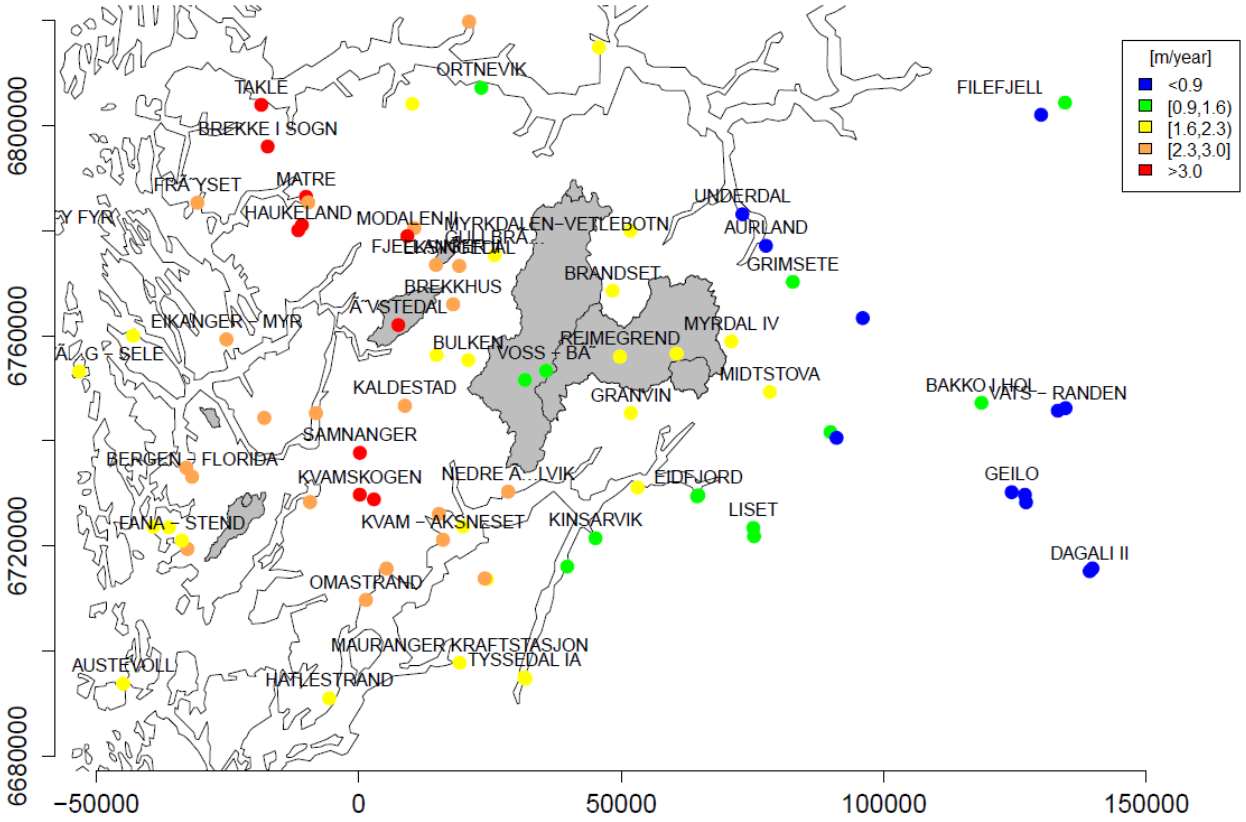


Figure 2.2: The mean annual precipitation from 1980 to 2013 at 79 locations. The x- and y-axis show distances i meters.

of annual precipitation on the windward side of the mountains, i.e in the western parts of Figure 2.2 (Ingebrigtsen et al., 2015). On the other side of the mountains we receive a dryer climate which can be observed as we move eastwards in Figure 2.2. The orographic precipitation defines the climatic differences in the area of interest, and a spatial effect common for all years seem to be suitable in the precipitation model.

In Figure 2.4 boxplots for annual precipitation for all locations are displayed for each year from 1980 to 2013. Recall that not all locations have observations from the whole time period. Thus, the 34 boxplots are not based on measurements from the same locations and might be misleading. However, the boxplots indicate that the median changes from year to year and suggest that the mean values of precipitation are year dependent. This is supported by what we observe in Figure 2.3: In 1988 we find larger values of precipitation for all locations except Geilo, while in 1990 we find lower values of precipitation everywhere. Thus, both Figures 2.3 and 2.4 suggest that some years in general are dryer than others and motivates for using a year dependent intercept in the precipitation model.

The boxplots also show that the variance within a year typically is larger for a year with a large median value (1988, 1989, 1991) than within a year with a low median value (1982,

2000, 2009). This suggests that there exist spatial variations that can not be explained by climatology, but by annual spatial variations. Based on this, we include a year dependent spatial effect in the precipitation model.

The measurement uncertainty for precipitation is unknown. According to Wolff et al. (2015) it is in general difficult to obtain accurate measurements of precipitation. Precipitation is typically measured by letting rain or snow fall into a bucket, but in windy areas biases in the measurements occur due to under-catch and turbulence. The rain or the snow simply blows away from the measurement bucket, and this is in particular a problem in the winter as snow is lighter than rain. We therefore assume that the measurement uncertainty for precipitation is large for our observations as they originate from a cold, mountainous, wind exposed area where a large proportion of the annual precipitation comes as snow.

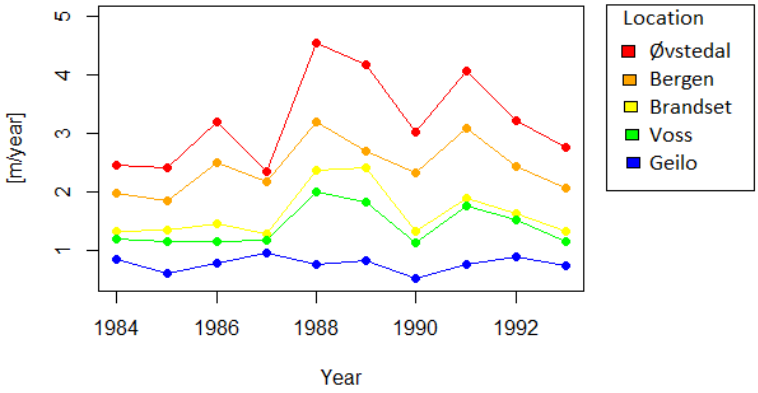


Figure 2.3: Observed annual precipitation from five different locations from 1984 to 1993.

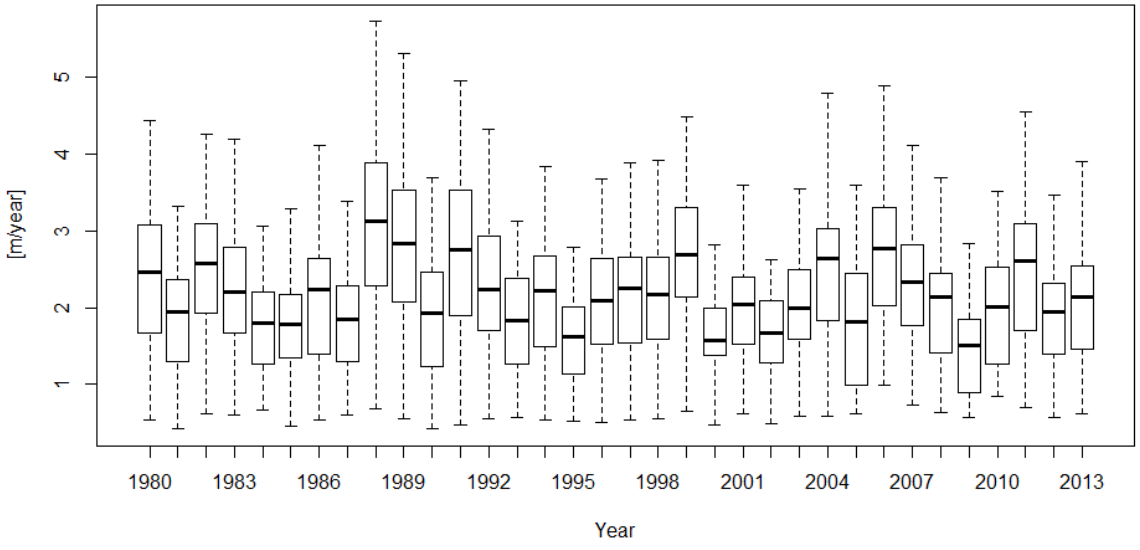


Figure 2.4: Boxplot of annual precipitation from 1980 to 2013 for all locations.

Evaporation data

Annual precipitation contributes to evaporation and runoff, and daily evaporation data for the nine catchments in Figure 2.1 were provided by NVE. The evaporation data are taken from the *MOD16 Global Terrestrial Evapotranspiration Data Set*, and the evaporation is derived by using satellite remote sensing data. See Mu et al. (2007a) and Mu et al. (2007b) for more information about the dataset.

The evaporation data received were daily evaporation values for 1 square kilometre for each of the catchments in Figure 2.1. Annual values for each catchment were obtained by adding the daily values and multiplying the result with the catchment area.

Evaporation data are available for all years from 1980 to 2013, and the mean values of annual evaporation for each catchment are displayed in Table 2.2. In Norway, the amount of evaporation is low because of the cold climate, and the mean amount of evaporation is 11% for our dataset. In Table 2.2 we also note that the correlation between the runoff data and the evaporation data is negative for the majority of the catchments: In years with high values of annual runoff, the evaporation is low. Evaporation in Norway is mainly driven by temperature as it is always humidity available. Typically, the temperature is lower for years with large annual values of precipitation and runoff, explaining the negative correlation between runoff and evaporation.

The estimation uncertainty for evaporation is unknown. According to Kolbjørn Engeland at NVE, the uncertainty is large, and it was suggested that the uncertainty was 20% of the provided evaporation value.

Table 2.2: Mean annual evaporation [$10^8 m^3$/year], evaporation per square meter [m/year] and the Pearson's correlation between evaporation and runoff for each catchment.

Catchment nr.	Mean evaporation	Evaporation/m^2	Corr(Runoff,Evaporation)
1	0.010	0.13	-0.48
2	0.015	0.45	0.11
3	0.30	0.60	0.34
4	0.026	0.20	-0.58
5	0.052	0.70	0.89
6	0.19	0.27	-0.23
7	0.061	0.15	-0.45
8	1.28	0.25	-0.37
9	3.1	0.28	-0.32

3. Background

In this section we present important background theory needed to construct a statistical model for annual precipitation and runoff. We start by an introduction to Gaussian random fields (GRFs) and Gaussian Markov random fields (GMRFs), and explain how these can be linked to each other through a Stochastic Partial Differential Equation (SPDE) formulation of the spatial field. Further, Latent Gaussian Models (LGMs) are presented, and we give an overview of how Integrated Nested Laplace approximation (INLA) can be used to make fast inference and predictions on such models.

3.1 Gaussian random fields

Tobler's first law of geography is the foundation of spatial statistics. It says that near things are more related than distant things (Tobler, 1970). This property is maintained by Gaussian random fields (GRFs) and is the motivation for using GRFs for modelling precipitation and runoff. In this section an introduction to Gaussian random fields is given. For details, we refer to Cressie (1993).

A Gaussian process is a process where the observations occur in a continuous domain, and where every finite collection of random variables from the process has a multivariate normal distribution. If the continuous domain is in space, the finite collection of random variables defines a Gaussian random field (GRF). For example, let $\mathcal{D} \subset \mathcal{R}^d$ be a spatial domain of interest where d typically is 2 or 3. The random field $\{x(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{D} \subset \mathcal{R}^d\}$ is a Gaussian random field if for each finite set of locations $\{\boldsymbol{s_1}, ..., \boldsymbol{s_n}\}$ in $\mathcal{D}$ it holds that

$$(x(\boldsymbol{s}_1), ..., x(\boldsymbol{s}_n)) \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where $\mathcal{N}_n$ is the n-variate normal distribution with mean $\boldsymbol{\mu} = (\mu(\boldsymbol{s}_1), ..., \mu(\boldsymbol{s}_n))$, covariance matrix $\boldsymbol{\Sigma}$ and $n \geq 1$. The covariance between the observations at locations i and j is given by element (i,j) of the covariance matrix, i.e $\Sigma_{ij} = Cov\{x(\boldsymbol{s_i}), x(\boldsymbol{s_j})\}$.

A random field is stationary if it is translation invariant. This means that $(x(\boldsymbol{s}_1), ..., x(\boldsymbol{s}_n))$ has the same distribution as $(x(\boldsymbol{s}_1 + \boldsymbol{t}), ..., x(\boldsymbol{s}_n + \boldsymbol{t}))$ where $\boldsymbol{t} = (t_1, ..., t_d)$. The field is also second-order stationary if the mean and the variance of the field do not depend on location, and if the correlation between two points only depends on the distance between them, i.e

$$\begin{aligned} \mathrm{E}\{x(\boldsymbol{s_i})\} &= \mu \\ \mathrm{Var}\{x(\boldsymbol{s_i})\} &= \sigma^2 \\ \mathrm{Corr}\{x(\boldsymbol{s_i}), x(\boldsymbol{s_j})\} &= \rho(\boldsymbol{s_j} - \boldsymbol{s_i}) \end{aligned}$$

for any i and j. If we add the requirement that $\mathrm{Corr}\{x(\boldsymbol{s_i}), x(\boldsymbol{s_j})\} = \rho(|\boldsymbol{s_j} - \boldsymbol{s_i}|)$ the field is isotropic. The covariance matrix of a GRF specifies the dependency structure of the field and is constructed from a covariance function. We will use a stationary and isotropic Matérn covariance function which is given by

$$Cov\{x(\boldsymbol{s_i}), x(\boldsymbol{s_j})\} = \frac{\sigma^2}{\Gamma(\lambda)2^{\lambda-1}}(\kappa||\boldsymbol{s}_i - \boldsymbol{s}_j||)^\lambda K_\lambda(\kappa||\boldsymbol{s}_i - \boldsymbol{s}_j||). \tag{3.1}$$

Here, σ^2 is the marginal variance of the spatial field, K_λ is the modified Bessel function of the second kind and order $\lambda > 0$, and $||\boldsymbol{s}_i - \boldsymbol{s}_j||$ is the Euclidean distance between the two locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$. The parameter κ is a scale parameter.

The range of the field tells us at which distance the correlation between two points in the field is approximately zero. It is shown empirically in Lindgren and Rue (2011) that the range of the Matérn covariance function is

$$\rho = \sqrt{8\lambda}/\kappa. \tag{3.2}$$

At this distance the spatial correlation is approximately 0.1.

3.2 Gaussian Markov random fields

Sometimes a spatial domain is discrete, for example if we consider pixels in an image. Gaussian Markov random fields (GMRFs) are useful in such situations and are defined as GRFs with a precision matrix $\boldsymbol{Q} = \boldsymbol{\Sigma^{-1}}$ characterized by a conditional independence structure. Let $\boldsymbol{x} = (x_1, ..., x_i, .., x_n)$ be a GMRF, and let $\boldsymbol{x_{-i}}$ be the vector $\boldsymbol{x}$ without element number i. The neighbourhood of element i in $\boldsymbol{x}$ is denoted $\boldsymbol{x_{ne(i)}}$, and can for example be the pixels in an image that share a border with element i. See Figure 3.1 for an example. For a GMRF the conditional distribution of x_i given $\boldsymbol{x_{-i}}$ can be expressed as the conditional distribution of x_i given its neighbours:

$$\pi(x_i|\boldsymbol{x_{-i}}) = \pi(x_i|\boldsymbol{x_{ne(i)}}).$$

This is the conditional independence structure of the GMRF, also known as the Markov property, and it is reflected in the precision matrix of the GMRF. For element $i \neq j$ it holds that

$$i \notin ne(j) \implies Q_{ij} = 0.$$

Thus, the precision matrix $\boldsymbol{Q}$ is sparse as it contains many zero elements.

Making inference and predictions with a spatial model require inversions and factorizations of the covariance matrix $\boldsymbol{\Sigma}$. The covariance matrix of a GRF is typically dense, and the computational cost for matrix operations on dense matrices has order $\mathrm{O}(n^3)$ where n is the dimension of the matrix. However, if the precision matrix $\boldsymbol{Q}$ has the conditional independence structure which is typical for a GMRF, the computational cost of matrix operations can be reduced to $O(n^{\frac{3}{2}})$. For this reason, GMRFs are easier to work with than GRFs allowing us to make faster inference and predictions.

We refer to Rue and Held (2005) for details on GMRFs and fast algorithms for sparse-matrix calculations.

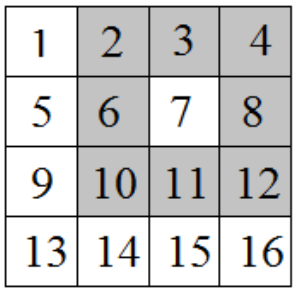


Figure 3.1: A discrete spatial field $\boldsymbol{x} = (x_1, ..., x_i, .., x_{16})$. The gray pixels are an example of a neighbourhood system of pixel 7 and is collected in vector $\boldsymbol{x_{ne(7)}} = (x_2, x_3, x_4, x_6, x_8, x_{10}, x_{11}, x_{12})$. If the discrete spatial field defines a GMRF, then $\pi(x_7|\boldsymbol{x}) = \pi(x_7|\boldsymbol{x_{ne(7)}})$.

3.3 The SPDE approach to spatial modelling

In Lindgren and Rue (2011) a new approach to spatial modelling is presented. In this paper it is shown that a GRF with a Matérn covariance matrix can be expressed as a solution to the following stochastic partial differential equation (SPDE):

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}}(\tau x(\boldsymbol{s})) = \mathcal{W}(\boldsymbol{s}). \tag{3.3}$$

Here, x(s) is a GRF, $\mathcal{W}(\boldsymbol{s})$ is spatial Gaussian white noise, α is a smoothness parameter, τ is a parameter controlling the variance of the GRF and $\kappa > 0$ is a scale parameter. The two dimensional Laplacian Δ is defined as $\sum_{i=1}^{d} \frac{\partial^2}{\partial x_i}$ where d is the dimension of our spatial domain $\mathcal{D}$.

The scale parameter κ in the SPDE (3.3) is the same κ as in Equation (3.1), and the parameter λ of the Matérn covariance function (3.1) is related to the SPDE through

$$\lambda = \alpha - \frac{d}{2}. \tag{3.4}$$

Further, the marginal variance σ^2 of Equation (3.1) is related to the SPDE through

$$\sigma^2 = \frac{\Gamma(\lambda)}{\Gamma(\alpha)(4\pi)^{\frac{d}{2}}\kappa^{2\lambda}\tau^2}. \tag{3.5}$$

An approximate solution of the SPDE (3.3) can be found by using a finite element method. This can be done by dividing the spatial domain $\mathcal{D}$ into a set of non-intersecting triangles. This leads to a triangulation mesh with m nodes and m basis functions ξ_k, where basis function ξ_k has value 1 at vertex k and value 0 at the other vertices. See Figure 3.2 for an example mesh .

The finite element method and the basis functions make it possible to express the GRF $x(s)$ through the basis function representation

$$x(\boldsymbol{s}) = \Sigma_{k=1}^{m} \xi_k(\boldsymbol{s}) w_k, \tag{3.6}$$

where w_k are weights with $\boldsymbol{w} = (w_1, ..., w_m) \sim N(0, \boldsymbol{Q}^{-1}(\tau, \kappa))$. The weights are chosen to approximate the solution $x(\boldsymbol{s})$ of the SPDE in the mesh nodes, and the basis functions transform the approximation of the GRF $x(\boldsymbol{s})$ from the mesh nodes to the locations of interest $\boldsymbol{s}$ through equation (3.6). Linear interpolation is used to determine the values in the interior of a triangle in the mesh.

The precision matrix of the weights $\boldsymbol{w}$ is defined as

$$\boldsymbol{Q}(\tau, \kappa) = \tau^2(\kappa^4\boldsymbol{C} + 2\kappa^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{C}^{-\mathbf{1}}\boldsymbol{G}),$$

where $C_{ii} = \int \xi_i(\boldsymbol{s})d\boldsymbol{s}$ and $C_{ij} = 0$ when $i \neq j$, and where $G_{ij} = \int \nabla\xi_i(\boldsymbol{s})\nabla\xi_j(\boldsymbol{s})d\boldsymbol{s}$ with $i = 1, ..., m$ and $j = 1, ..., m$. Often the precision matrix is parametrised as $\boldsymbol{Q}(\theta_\tau, \theta_\kappa)$ where $\theta_\tau = log(\tau)$ and $\theta_\kappa = log(\kappa)$, i.e:

$$\boldsymbol{Q}(\theta_\tau, \theta_\kappa) = \exp(2\theta_\tau)\{\exp(4\theta_\kappa)\boldsymbol{C} + 2\exp(2\theta_\kappa)\boldsymbol{G} + \boldsymbol{G}\boldsymbol{C}^{-\mathbf{1}}\boldsymbol{G}\}. \tag{3.7}$$

This parametrisation will be used throughout the text.

The precision matrix $\boldsymbol{Q}(\theta_\tau, \theta_\kappa)$ is sparse as opposed to the original precision matrix of the GRF $x(\boldsymbol{s})$. In fact $\boldsymbol{Q}(\theta_\tau, \theta_\kappa)$ defines a GMRF and is equipped with the conditional independence structure described in Section 3.2. In practice, we have now expressed the original GRF $x(\boldsymbol{s})$ as a GMRF through Equation (3.6). This allows us to reduce the computational cost of matrix operations. The computational cost is $\mathrm{O}(n^3)$ for matrix operations on a GRF evaluated at n locations, while the computational cost for the corresponding GMRF approximation is only $\mathrm{O}(m^{3/2})$ where m is the number of mesh nodes in the triangulation. Thus, we can make a coarser mesh to increase the computational speed of matrix inversions and factorizations. However, the accuracy of the GMRF approximation decreases with decreasing mesh density. In Figure 3.3 an example is shown giving an impression of how the GMRF approximation may look like for a specific triangulation mesh.

The SPDE approach has other advantages than computational benefits. It makes it possible to introduce non-stationary parameters in the covariance matrix of a GRF without thinking about the positive definiteness of the covariance matrix. This is because the SPDE approach allows us to regard the model parameters as parameters of the SPDE and not as building blocks of a covariance matrix. The SPDE approach is relatively new, and an overview of the current state of spatial modelling with systems of SPDEs can be found in Hu and Steinsland (2016). Interesting topics for further research are also suggested in this article.

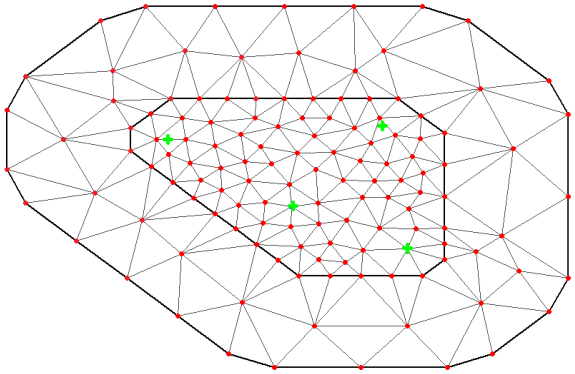


Figure 3.2: Example mesh with $m = 131$ nodes (red). We approximate the GRF $x(\boldsymbol{s})$ in the 131 mesh nodes by the GMRF $\boldsymbol{w}$. Further, we use the basis functions $\xi_k(\boldsymbol{s})$ to interpolate the spatial field from the 131 mesh nodes to the 4 locations of interest (green).

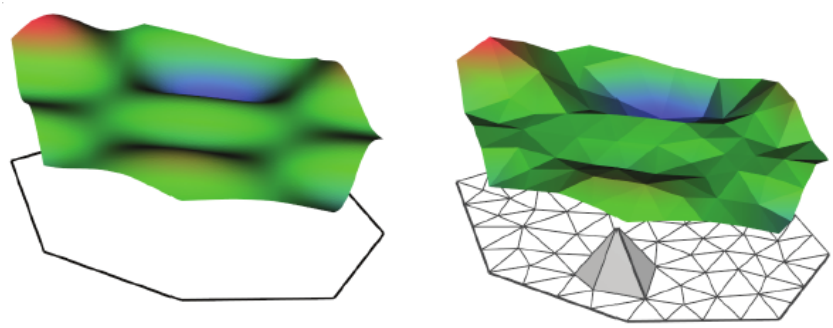


Figure 3.3: A smooth surface (left) can be approximated by using a basis function representation with piecewise linear basis functions defined on a triangulation mesh (right). The illustration is copied with permission from Hu and Steinsland (2016).

3.4 Bayesian hierarchical models and Latent Gaussian models

In this work the model for precipitation and runoff will be expressed as a hierarchical model of three levels. Hierarchical models are used to model complicated processes as a hierarchy of relatively simple statistical models (Hu and Steinsland, 2016). The first level of a hierarchical model is the observation model or observation likelihood $\pi(\boldsymbol{y}|\boldsymbol{\theta_1}, \boldsymbol{\psi}) = \prod_{i=1}^{n} \pi(y_i|\boldsymbol{\theta_1}, \boldsymbol{\psi})$ where $\boldsymbol{y}$ is a vector of n observations. The observation likelihood depends on a unobserved process $\boldsymbol{\psi}$ and some parameters $\boldsymbol{\theta_1}$.

The second stage of a hierarchical model is a model that describes the distribution of the unobserved process $\boldsymbol{\psi}$. This is called the process model or the latent model and is often a linear model of fixed and random effects. For example

$$\boldsymbol{\psi} = \boldsymbol{X\beta} + x(\boldsymbol{s}) \tag{3.8}$$

where $\boldsymbol{X\beta}$ is a fixed effect with design matrix $\boldsymbol{X}$ and coefficients $\boldsymbol{\beta}$, and where $x(\boldsymbol{s})$ is a GRF. The unobserved process $\boldsymbol{\psi}$ is then given a distribution $\pi(\boldsymbol{\psi}|\boldsymbol{\theta_2})$ where $\boldsymbol{\theta_2}$ is a vector of related parameters. For the above example, $\boldsymbol{\beta}$ and the parameters of the GRF $x(\boldsymbol{s})$ will be included

in $\boldsymbol{\theta_2}$.

If we use a Bayesian approach to statistics, the parameters $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \boldsymbol{\theta_2})$ must be assigned a prior distribution $\pi(\boldsymbol{\theta})$. Thus, we have a *Bayesian* hierarchical model where the third stage is the prior distribution of the model parameters $\pi(\boldsymbol{\theta})$. The parameters $\boldsymbol{\theta}$ are often called hyperparameters.

If the unobserved process $\boldsymbol{\psi}$ is Gaussian distributed given the parameters $\boldsymbol{\theta_2}$, we have a latent Gaussian model (LGM). Latent Gaussian models are a special case of Bayesian hierarchical models, and will be used for modelling precipitation and runoff in this work.

3.5 Integrated nested Laplace approximation (INLA)

Assume we have a Bayesian hierarchical model as described above. The objective of a Bayesian analysis is to gain knowledge about the unknown parameters $\boldsymbol{\theta}$ and the distribution of the unknown process $\boldsymbol{\psi}$ based on a set of observations $\boldsymbol{y}$. Thus, we need to calculate the posterior marginal distribution $\pi(\psi_i|\boldsymbol{y})$ of each element $i = 1, ..., m$ of the latent model $\boldsymbol{\psi}$ and the posterior marginal distribution $\pi(\theta_j|\boldsymbol{y})$ of each element $j = 1, ..., n$ of the hyperparameter vector $\boldsymbol{\theta}$. Traditionally, posterior marginals are computed by using Markov Chain Monte Carlo-methods (MCMC), but as m can be very large, typically larger than 100, MCMC-methods can be slow.

In Rue et al. (2009) an alternative to MCMC-methods is presented called Integrated Nested Laplace approximation (INLA). This is a methodology for doing approximate Bayesian inference on LGMs, and it enables fast computations of the posterior marginals. The approximations are very accurate and sometimes even long MCMC-runs are not able to detect the errors in the approximations.

The LGM must satisfy some requirements to ensure fast computations in INLA: (1) Sparse matrix algorithms are used to provide fast inference and predictions. For this reason, the latent Gaussian field $\boldsymbol{\psi}$ should be given a GMRF prior. (2) The number of hyperparameters $\boldsymbol{\theta}$ should not be too large as we need to perform numerical integration over the hyperparameter space.

We now give a short overview of how INLA works based on Chapter 4.7 in Blangiardo and Cameletti (2015). In INLA, the following integrals are used to compute the posterior marginals:

$$\pi(\psi_i|\boldsymbol{y}) = \int \pi(\psi_i, \boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta} = \int \pi(\psi_i|\boldsymbol{\theta}, \boldsymbol{y}) \pi(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta}, \tag{3.9}$$

$$\pi(\theta_j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta}_{-j}, \tag{3.10}$$

where $\boldsymbol{\theta_{-j}}$ denotes vector $\boldsymbol{\theta}$ without element j. To solve the integrals, we first need to compute the joint posterior distribution of the hyperparameters $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. In general, INLA utilises that

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{\psi}, \boldsymbol{\theta}|\boldsymbol{y})}{\pi(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y})} = \frac{\pi(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) \pi(\boldsymbol{\psi}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\pi(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y}) \pi(\boldsymbol{y})} \propto \frac{\pi(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) \pi(\boldsymbol{\psi}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\pi(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y})} \approx \left. \frac{\pi(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) \pi(\boldsymbol{\psi}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\tilde{\pi}(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y})} \right|_{\boldsymbol{\psi} = \boldsymbol{\psi}^*(\boldsymbol{\theta})} \tag{3.11}$$

where $\tilde{\pi}(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y})$ is the Gaussian approximation of $\pi(\boldsymbol{\psi}|\boldsymbol{\theta}, \boldsymbol{y})$ at its mode $\boldsymbol{\psi}^*(\boldsymbol{\theta})$ for a given $\boldsymbol{\theta}$.

Further, we need to compute $\pi(\psi_i|\boldsymbol{\theta}, \boldsymbol{y})$ in order to find the posterior marginal of the latent field (3.9). One of the options is to use a Gaussian approximation again

$$\begin{aligned} \pi(\psi_i|\boldsymbol{\theta}, \boldsymbol{y}) &= \frac{\pi((\psi_i, \boldsymbol{\psi_{-i}})|\boldsymbol{\theta}, \boldsymbol{y})}{\pi(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})} = \frac{\pi(\boldsymbol{\psi}, \boldsymbol{\theta}|\boldsymbol{y})}{\pi(\boldsymbol{\theta}|\boldsymbol{y})\pi(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})} \\ &\propto \frac{\pi(\boldsymbol{\psi}, \boldsymbol{\theta}|\boldsymbol{y})}{\pi(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})} \approx \frac{\pi(\boldsymbol{\psi}, \boldsymbol{\theta}|\boldsymbol{y})}{\tilde{\pi}(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})}\bigg|_{\boldsymbol{\psi_{-i}}=\boldsymbol{\psi^*_{-i}}(\boldsymbol{\psi_i}, \boldsymbol{\theta})}, \end{aligned} \tag{3.12}$$

where $\boldsymbol{\psi_{-i}}$ denotes the vector $\boldsymbol{\psi}$ without element i, and $\tilde{\pi}(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})$ is the Gaussian approximation of $\pi(\boldsymbol{\psi_{-i}}|\psi_i, \boldsymbol{\theta}, \boldsymbol{y})$ at its mode $\boldsymbol{\psi^*_{-i}}(\psi_i, \boldsymbol{\theta})$ for a given $\boldsymbol{\theta}$ and ψ_i.

When $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ and $\pi(\psi_i|\boldsymbol{\theta}, \boldsymbol{y})$ are approximated as outlined above, the next step in the INLA-methodology is to solve the integrals (3.9) and (3.10) numerically. If the observation likelihood of the LGM is Gaussian, the Gaussian approximations in Equations (3.11) and (3.12) are exact and can be omitted. This will be the case for our model for precipitation and runoff. In such cases, the accuracy of INLA is determined by the accuracy of the numerical integration scheme used for solving (3.9) and (3.10). We refer to Rue et al. (2009) for a description of the numerical methods used in INLA.

4. Statistical model for precipitation and runoff

The SPDE approach to spatial modelling is now used to construct a Latent Gaussian model for annual precipitation and runoff which is suitable for fast computations by using the INLA methodology.

4.1 Observation model for annual precipitation and runoff

Let $\mathcal{D} \subset \mathcal{R}^2$ be a spatial domain representing an area, e.g the whole area displayed in Figure 2.2, and let the true annual precipitation in year j at location $\boldsymbol{s} \in \mathcal{D}$ be denoted by $\eta_j(\boldsymbol{s})$ where $j = 1, ..., r$. We assume that the annual precipitation in year j at location $\boldsymbol{s_i}$ is observed with measurement uncertainty ϵ_{ij} where $i = 1, ..., n$. The observation model for the annual precipitation y_{ij} at location $\boldsymbol{s_i}$ can then be written as

$$y_{ij} = \eta_j(\boldsymbol{s_i}) + \epsilon_{ij}, \tag{4.1}$$

where $\boldsymbol{\epsilon_j} = (\epsilon_{1j}, .., \epsilon_{ij}, ..., \epsilon_{nj})^{\mathrm{T}}$ is assumed to be $\mathcal{N}(0, \tau_p^{-1}\mathrm{diag}(\boldsymbol{f_j})\mathbf{I})$, and where ϵ_{ij} is independent of $\eta_j(\boldsymbol{s}_i)$. The vector $\boldsymbol{f_j} = (f_{1j}, ..., f_{ij}, ..., f_{nj})$ is a vector of fixed scales allowing each observation to have its own variance. This is an important part of the model because the uncertainty typically is large for a large observation and smaller for a smaller observation.

Let $\mathcal{D}_l$ for $l = 1, ..., N$ be N catchments of interest located in our spatial domain $\mathcal{D}$. The annual runoff within catchment $\mathcal{D}_l$ in year j is denoted $\zeta_j(\mathcal{D}_l)$, while the annual evaporation in catchment l is denoted E_{lj}. As for precipitation, the runoff and evaporation within catchment $\mathcal{D}_l$ in year j are observed with measurement uncertainty. Recall that precipitation contributes to runoff and evaporation. Precipitation can thus be regarded as the sum of runoff and evaporation in a catchment. We use this quantity, the sum of runoff and evaporation, for constructing a second observation model.

Let e_{lj} denote the measurement uncertainty for the sum of annual runoff and evaporation in catchment $\mathcal{D}_l$, year j. The observed sum of runoff and evaporation within catchment $\mathcal{D}_l$ is denoted by z_{lj} and can be written as

$$z_{lj} = \zeta_j(\mathcal{D}_l) + E_{lj} + e_{lj}, \tag{4.2}$$

where $\boldsymbol{e_j} = (e_{1j}, ..., e_{lj}, ..., e_{Nj})^{\mathrm{T}}$ is assumed to be $\mathcal{N}(\mathbf{0}, \tau_r^{-1}\mathrm{diag}(\boldsymbol{v_j})\mathbf{I})$. As for the precipitation model, $\boldsymbol{v_j} = (v_{1j}, ..., v_{lj}, ..., v_{Nj})$ is a vector of fixed scales making it possible to have a unique variance for each observation z_{lj}.

4.2 Process model for annual precipitation and runoff

We assume that the true level of annual precipitation can be modelled by three components: A year specific intercept (β_j) and two stationary and isotropic GRFs, one that is common

for all years ($c(\boldsymbol{s})$) and one that is year specific ($x_j(\boldsymbol{s})$). The GRF $c(\boldsymbol{s})$ expresses the spatial variations that can be explained by climatic differences between locations, and we call this the climatology of the area of interest. The spatial field $x_j(\boldsymbol{s})$ expresses the annual spatial variations of precipitation and we call this the annual spatial variability. This leads to the following process model for the true annual precipitation in year j:

$$\eta_j(\boldsymbol{s}) = \beta_j + x_j(\boldsymbol{s}) + c(\boldsymbol{s}). \tag{4.3}$$

Further, it is assumed that $x_j(\boldsymbol{s})$ for $j = 1, ..., r$ represents r independent realizations of the spatial field. This is called replicates.

The model for precipitation in Equation (4.3) will be called Model 1 for further reference. We will also use a simpler model where the annual spatial variability is removed from the model. This will be referred to as Model 2:

$$\eta_j(\boldsymbol{s}) = \beta_j + c(\boldsymbol{s}). \tag{4.4}$$

Model 1 is used unless otherwise specified.

The annual precipitation contributes to evaporation and runoff. Thus, the true amount of runoff in catchment $\mathcal{D}_l$ of year j can be modelled as the integral of the precipitation over the area of $\mathcal{D}_l$ minus the evaporation E_{lj} in catchment l:

$$\zeta_j(\mathcal{D}_l) = \int_{\boldsymbol{s}\in\mathcal{D}_l} \eta_j(\boldsymbol{s}) d\boldsymbol{s} - E_{lj}. \tag{4.5}$$

We simplify the computations by discretising the catchment $\mathcal{D}_l$ into a regular grid denoted by $\mathcal{L}_{\mathcal{D}_l}$. For each grid node in $\mathcal{L}_{\mathcal{D}_l}$ it is assumed that the annual precipitation is either estimated or known. The true runoff in year j in $\mathcal{D}_l$ can now be approximated by summing the precipitation in the grid nodes inside the catchment of interest, and multiplying with an area element Δ_l. The area Δ_l is the area of catchment $\mathcal{D}_l$ divided by the number of grid nodes in $\mathcal{L}_{\mathcal{D}_l}$. Thus, we replace the integral with an sum, and get the following model for the annual runoff in year j within catchment $\mathcal{D}_l$:

$$\zeta_j(\mathcal{D}_l) = \Delta_l \sum_{\boldsymbol{s}_i\in\mathcal{L}_{\mathcal{D}_l}} \eta_j(\boldsymbol{s_i}) - E_{lj}. \tag{4.6}$$

Figure 4.1 shows how we discretise the catchments from Figure 2.1. For example, the discretisation of Catchment 6 (green) consists of 69 grid nodes, and Catchment 6 covers an area of 72.4 km^2 giving $\Delta_l = (72.4/69)$ km^2. The runoff within catchment 6 is then approximated by summing the annual precipitation in the 69 grid nodes and multiplying with Δ_l.

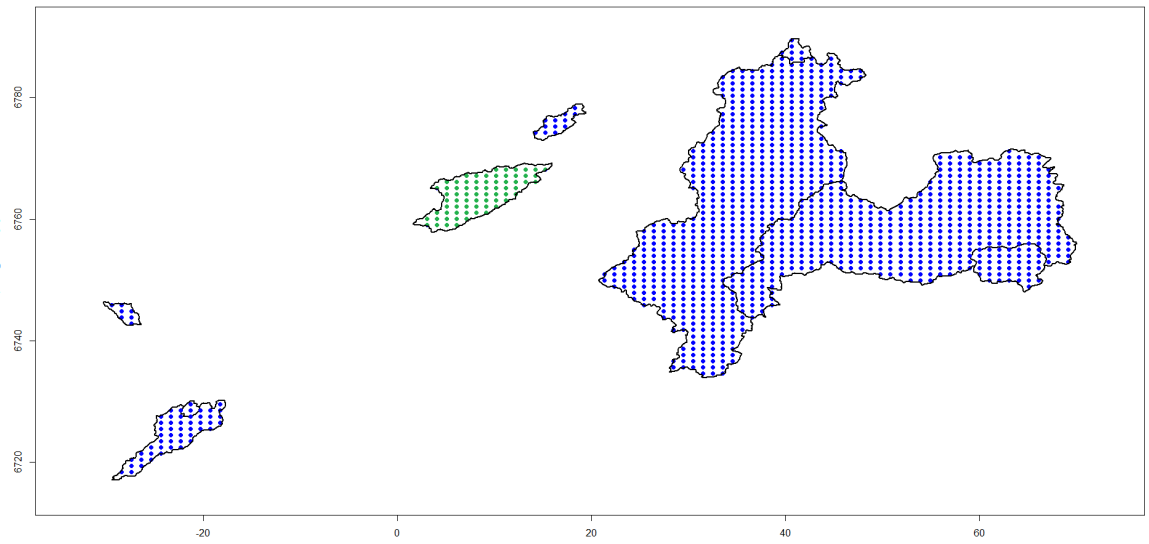


Figure 4.1: Regular grid used for computing or predicting the runoff within the catchments in Figure 2.1. Catchment 6 is marked with green.

4.3 The SPDE approach used on our model

We now use the SPDE approach introduced in Section 3.3 to express the GRFs $c(\boldsymbol{s})$ and $x_j(\boldsymbol{s})$ from Equation (4.3) as GMRFs to get a model that ensures fast computations in INLA. First, a triangulation mesh is constructed with basis functions ξ_k where $k = 1, ..., m$ and m is the total number of vertices in the triangulation mesh. The spatial variability $x_j(\boldsymbol{s})$ of year j can then be expressed as

$$x_j(\boldsymbol{s}) = \Sigma_{k=1}^{m} \xi_k(\boldsymbol{s}) w_{kj},$$

where w_{kj} is an approximation of $x_j(\boldsymbol{s})$ in the m mesh nodes. Note that we use the same triangulation mesh for all years. The basis functions ξ_k are therefore not indexed by year. The weights w_{kj} however differ between years as the spatial field is different for each j. The weights w_{kj} are given a normal distribution with precision matrix $\boldsymbol{Q}(\cdot, \cdot)$ given by (3.7):

$$\boldsymbol{w_j} = (w_{1j}, ..., w_{kj}, ..., w_{mj})^{\mathrm{T}} \sim \mathcal{N}(0, \boldsymbol{Q^{-1}}(\theta_{\tau,w}, \theta_{\kappa,w}))$$

with $\theta_{\tau,w} = log(\tau_w)$ and $\theta_{\kappa,w} = log(\kappa_w)$, and with range ρ_w and marginal variance σ_w^2. Thus, $\boldsymbol{w_j}$ defines a GMRF.

In the same manner, we approximate the climatology $c(\boldsymbol{s})$ by a finite basis function representation as well

$$c(\boldsymbol{s}) = \Sigma_{k=1}^{m} \xi_k(\boldsymbol{s}) u_k,$$

where u_k is the GRF $c(\boldsymbol{s})$ approximated in the m mesh nodes of the triangulation. The same triangulation mesh and basis functions are used as for the annual variability $x_j(\boldsymbol{s})$. The weights u_k define a new GMRF with precision matrix given by (3.7):

$$\boldsymbol{u} = (u_1, ..., u_k, ..., u_m)^{\mathrm{T}} \sim \mathcal{N}(0, \boldsymbol{Q^{-1}}(\theta_{\tau,u}, \theta_{\kappa,u}))$$

with $\theta_{\tau,u} = log(\tau_u)$ and $\theta_{\kappa,u} = log(\kappa_u)$, and with range ρ_u and marginal variance σ_u^2.

The process model for precipitation from Equation (4.3) can now be written as

$$\eta_j(\boldsymbol{s}) = \beta_j + \Sigma_{k=1}^{m} \xi_k(\boldsymbol{s})(w_{kj} + u_k). \tag{4.7}$$

We assume that we observe the process at n locations $\{\boldsymbol{s_1}, ..., \boldsymbol{s_i}, ..., \boldsymbol{s_n}\}$ in space, and the same n locations are observed for all replicates. The true precipitation at these locations in year j can then be written in vector notation as

$$\boldsymbol{\eta_j} = \mathbf{1}\beta_j + \boldsymbol{A_p}(\boldsymbol{w}_j + \boldsymbol{u}) = \boldsymbol{A_p}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u}), \tag{4.8}$$

with $\boldsymbol{\eta_j} = (\eta_j(\boldsymbol{s_1}), ..., \eta_j(\boldsymbol{s_i}), ..., \eta_j(\boldsymbol{s_n}))^{\mathrm{T}}$, and where $\boldsymbol{A_p}$ is a $n \times m$ matrix where element (i, k) is given by $A_{p(i,k)} = \xi_k(\boldsymbol{s}_i)$. The last equality in (4.8) is possible because the row sums of $\boldsymbol{A_p}$ are 1. This is because we use piece-wise linear basis functions that sum to 1 at each location (Rue and Lindgren, 2015).

By using the SPDE approach, the process model for runoff from Equation (4.6) can be written in vector form as well. Say, we observe the runoff within N of the catchments in Figure 2.1. The process model for runoff then becomes

$$\boldsymbol{\zeta_j} = \boldsymbol{A_\Sigma}\boldsymbol{\eta_j} - \boldsymbol{E_j} = \boldsymbol{A_\Sigma}(\boldsymbol{A_p}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u})) - \boldsymbol{E_j} = \boldsymbol{A_r}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u}) - \boldsymbol{E_j}, \tag{4.9}$$

where $\boldsymbol{\zeta_j} = (\zeta_j(\mathcal{D}_1), ..., \zeta_j(\mathcal{D}_l), ..., \zeta_j(\mathcal{D}_N))^{\mathrm{T}}$ and $\boldsymbol{E_j} = (E_{1j}, ...E_{lj}, ..., E_{Nj})^{\mathrm{T}}$. The variable $\boldsymbol{A_\Sigma}$ is a matrix that sums the precipitation in the correct grid nodes in Figure 4.1, and multiplies the resulting sum with the correct area element Δ_l.

4.4 Latent Gaussian model for annual precipitation and runoff suitable for INLA

The models (4.1) and (4.2) define the observation likelihoods and the first level of the LGM for precipitation and runoff. Because $\boldsymbol{\epsilon}_j$ and $\boldsymbol{e_j}$ are Gaussian with precisions τ_p and τ_r, the observation likelihoods are Gaussian given $\boldsymbol{w_j}$, $\boldsymbol{u}$, β_0, τ_p and τ_r i.e:

$$\boldsymbol{y_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_p, \beta_j) \sim \mathcal{N}(\boldsymbol{\eta_j}, \tau_p^{-1}\mathrm{diag}(\boldsymbol{f_j})\mathbf{I}) \tag{4.10}$$

$$\boldsymbol{z_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_r, \beta_j) \sim \mathcal{N}(\boldsymbol{\zeta_j} + \boldsymbol{E_j}, \tau_r^{-1}\mathrm{diag}(\boldsymbol{v_j})\mathbf{I}) \tag{4.11}$$

with $\boldsymbol{y_j} = (y_{1j}, ..., y_{ij}, ..., y_{nj})^{\mathrm{T}}$ and $\boldsymbol{z_j} = (z_{1j}, ..., z_{lj}, ..., z_{Nj})^{\mathrm{T}}$. The expected values of $\boldsymbol{y_j}$ and $\boldsymbol{z_j}$ are now given by:

$$\begin{aligned} E\{\boldsymbol{y_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_p, \beta_j)\} &= \boldsymbol{\eta_j} = \boldsymbol{A_p}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u}) \\ E\{\boldsymbol{z_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_r, \beta_j)\} &= \boldsymbol{\zeta_j} + \boldsymbol{E_j} = \boldsymbol{A_r}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u}), \end{aligned} \tag{4.12}$$

and we note that the two means have the same form. The only difference is the projection matrices $\boldsymbol{A_p}$ and $\boldsymbol{A_r}$. Having the means on the same form is convenient in the implementation

of the model (see Appendix B.4) because $\boldsymbol{y_j}$ and $\boldsymbol{z_j}$ now can be regarded as observations of the same process. The observations $\boldsymbol{y_j}$ can be seen as point observations of precipitation while the observations $\boldsymbol{z_j}$ can be seen as area observations of precipitation. This is the reason why the observed sum of runoff and evaporation is used to construct the observation likelihood instead of using only the observed runoff which is the quantity of interest.

The observation likelihoods (4.10) and (4.11) depend on the unobserved, latent components: $\boldsymbol{w_j}$ for $j = 1, .., r$, and $\boldsymbol{u}$. In the INLA framework, the intercept β_j is regarded as a part of the unobserved spatial field and not as a hyperparameter (Blangiardo and Cameletti, 2015). We gather the latent components in the vector $\boldsymbol{\psi} = \{\boldsymbol{w_1}, ..., \boldsymbol{w_j}, ..., \boldsymbol{w_r}, \beta_1, ..., \beta_j, ..., \beta_r, \boldsymbol{u}\}$. Because the SPDE approach is used, the latent components $\boldsymbol{w_j}$ and $\boldsymbol{u}$ are both GMRFs with precision matrices given by (3.7). Further, we give the intercepts β_j a Gaussian prior. This implies that the latent components all are Gaussian given the hyperparameters:

$$\boldsymbol{w_j}|(\theta_{\tau,w}, \theta_{\kappa,w}) \sim \mathcal{N}(0, \boldsymbol{Q^{-1}}(\theta_{\tau,w}, \theta_{\kappa,w})) \tag{4.13}$$

$$\boldsymbol{u}|(\theta_{\tau,u}, \theta_{\kappa,u}) \sim \mathcal{N}(0, \boldsymbol{Q^{-1}}(\theta_{\tau,u}, \theta_{\kappa,u})) \tag{4.14}$$

$$\beta_j \sim \mathcal{N}(\cdot, \cdot) \tag{4.15}$$

for j=1,..r. The latent components are also jointly Gaussian because $\boldsymbol{w_j}$, $\boldsymbol{u}$ and β_j all are Gaussian and independent of each other, i.e $\boldsymbol{\psi} \sim \mathcal{N}(\cdot, \cdot)$. This implies that we have a LGM as described in Section 3.4. Because $\boldsymbol{w_j}$ and $\boldsymbol{u}$ also are GMRFs, the model is suitable for fast computations in INLA.

The models for precipitation and runoff have 6 hyperparameters in total collected in the vector $\boldsymbol{\theta} = (\tau_p, \tau_r, \theta_{\kappa,w}, \theta_{\tau,w}, \theta_{\kappa,u}, \theta_{\tau,u})$. The joint distribution of the hyperparameters is

$$\pi(\boldsymbol{\theta}) = \pi(\tau_p)\pi(\tau_r)\pi(\theta_{\kappa,w})\pi(\theta_{\tau,w})\pi(\theta_{\kappa,u})\pi(\theta_{\tau,u})$$

where $\pi(\cdot)$ is the prior distribution of its argument. This is the third level of the LGM in our Bayesian model for annual precipitation and runoff.

4.5 Specification of parameter values and prior distributions

Before we start analysing the model, we need to specify the parameter values of the model and assign prior distributions to the model parameters.

Parameter values

Data for standard deviation for the observations of precipitation are not available, and as mentioned in Section 2 measurements of precipitation are known to be uncertain. We assume a priori that the standard deviation is 10% of the observed value which is a relatively large amount. Thus,

$$Var(y_{ij}) = (0.1y_{ij})^2.$$

The standard deviation of the evaporation data is also unknown, and we assume a priori that this is a even more uncertain quantity. We assume that the standard deviation of the evaporation data is 20% of the observed value. The variance of the runoff observations are given by NVE, and the variance of a runoff observation from catchment $\mathcal{D}_l$ in year j is denoted $\text{Var}(\zeta^*_{lj})$. Because we observe the sum of evaporation and runoff, i.e $z_{lj} = \zeta_j(\mathcal{D}_l) + E_{lj} + e_{lj}$, the variance of the observation z_{lj} is given by

$$Var(z_{lj}) = Var(\zeta^*_{lj}) + (0.2E_{lj})^2 + 2 \times (0.2E_{lj}) \times \sqrt{Var(\zeta^*_{lj})} \times Corr(\zeta^*, E),$$

where $Corr(\zeta^*, E)$ is the Pearson's correlation between all the observations of runoff and evaporation from all years. The correlations between annual runoff and evaporation for all catchments are shown in Table 2.2.

We choose $\alpha = 2$ as the smoothness parameter in Equation (3.3), because this choice is supported by INLA. As the analysis domain is in $\mathcal{R}^2$, we have $d = 2$ in Equation (3.4), leading to $\lambda = 1$. Thus, $\alpha = 2$, $d = 1$ and $\lambda = 1$ for both the SPDE representing $x_j(\boldsymbol{s})$ and the SPDE representing $c(\boldsymbol{s})$.

Prior distributions

All of the hyperparameters $\boldsymbol{\theta} = (\theta_{\kappa,w}, \theta_{\tau,w}, \theta_{\kappa,u}, \theta_{\tau,u}, \tau_p, \tau_r)$ and $(\beta_1, ..., \beta_r)$ are now assigned a prior distribution. We use informative priors based on the information given by the dataset.

The intercepts $(\beta_1, ..., \beta_r)$ which are a part of the latent component of the LGM, are all given the same Gaussian prior. The prior mean is set to 2 m/year because the observed annual precipitation typically lies around this value (see Figure 2.2). The prior variance is chosen to be 1^2. This gives the prior distribution $\mathcal{N}(2^2, 1^2)$ for β_j for $j = 1, ..., r$.

The hyperparameters $\theta_\kappa = log(\kappa)$ and $\theta_\tau = log(\tau)$ originating from the SPDE (3.3) have no direct physical interpretation which makes it challenging to choose prior distributions for these parameters. In Fuglstad et al. (2015) this problem is discussed, and practically and useful priors are constructed for the range and the variance for both stationary and non-stationary Matérn GRFs. In this work however, we choose an easier approach. The approach is taken from Ingebrigtsen et al. (2015), and we apply that the SPDE parameters τ and κ are linked to the range and the marginal variance of the spatial field through Equations (3.2) and (3.5). The relations are restated here

$$\rho = \sqrt{8}/\kappa \tag{4.16}$$

$$\sigma^2 = 1/\sqrt{4\pi}\tau\kappa, \tag{4.17}$$

where we use that $\lambda = 1$, $\alpha = 1$ and $d = 2$.

We first assign the following priors to the SPDE parameters: $\theta_\tau \sim \mathcal{N}(\mu_\tau, \sigma^2_\tau)$ and $\theta_\kappa \sim$

$\mathcal{N}(\mu_\kappa, \sigma_\kappa^2)$ where the prior means μ_τ and μ_κ, and variances σ_τ^2 and σ_κ^2 are unknown. Our task is now to find suitable values for these parameters. Using equations (4.16) and (4.17), and that $\theta_\tau = log(\tau)$ and $\theta_\kappa = log(\kappa)$, it can be shown that

$$\rho \sim log\mathcal{N}(log(\sqrt{8}) - \mu_\kappa, \sigma_\kappa^2)$$
$$\sigma \sim log\mathcal{N}(-log(\sqrt{4\pi}) - \mu_\tau - \mu_\kappa, \sigma_\tau^2 + \sigma_\kappa^2),$$

where log $\mathcal{N}(\cdot, \cdot)$ is the log-normal distribution. This implies that the p-quantiles of ρ and σ are

$$\rho(p) = \sqrt{8}\exp(-\mu_\kappa + \sigma_\kappa \Phi^{-1}(p)) \tag{4.18}$$
$$\sigma(p) = \frac{1}{\sqrt{4\pi}}\exp\left(-\mu_\tau - \mu_\kappa + \sqrt{\sigma_\tau^2 + \sigma_\kappa^2}\Phi^{-1}(p)\right), \tag{4.19}$$

where $\Phi^{-1}(p)$ is the cumulative distribution for the standard normal distribution and $0 < p < 1$. We can now choose two quantiles, for example p =0.5 and p =0.9 with corresponding $\rho(p)$ and $\sigma(p)$ and solve Equations (4.18) and (4.19) to find μ_τ, μ_κ, σ_τ^2 and σ_κ^2.

The above approach is now used to assign prior distributions to the SPDE parameters $\theta_{\kappa,u}$ and $\theta_{\tau,u}$ related to the climatology $c(\boldsymbol{s})$, and to $\theta_{\kappa,w}$ and $\theta_{\tau,w}$ related to the spatial annual variability $x_j(\boldsymbol{s})$. To get an idea of the magnitude of the spatial variations in the dataset, we compute the standard deviation of the mean values displayed in Figure 2.2. This is 0.85 m. It is hard to see from the dataset if this variation is mainly caused by a climatic effect or by annual variations. For this reason, $\theta_{\kappa,u}$ and $\theta_{\tau,u}$ are given the same priors as $\theta_{\kappa,w}$ and $\theta_{\tau,w}$ respectively.

The value 0.85 m is chosen as the 0.5-quantile for $\sigma(p)$ in Equation (4.19). The 0.9-quantile is chosen to be 2 m. The whole area in Figure 2.2 is of size approximately 200 km × 110 km. We choose the range 110 km as the 0.9-quantile in Equation (4.18) and one third of this, 33 km, as the 0.5-quantile. Inserting these four quantiles into Equations (4.18) and (4.19) and solving for μ_κ, μ_τ, σ_τ^2 and σ_κ^2, gives the following priors for the SPDE parameters

$$\theta_{\tau,w} \sim \mathcal{N}(1.41, 0.61^2) \tag{4.20}$$
$$\theta_{\kappa,w} \sim \mathcal{N}(-2.46, 0.94^2) \tag{4.21}$$

and

$$\theta_{\tau,u} \sim \mathcal{N}(1.41, 0.61^2) \tag{4.22}$$
$$\theta_{\kappa,u} \sim \mathcal{N}(-2.46, 0.94^2) \tag{4.23}$$

where $\theta_{\tau,w}$ and $\theta_{\kappa,w}$ are related to the annual variability $x_j(\boldsymbol{s})$, and $\theta_{\tau,u}$ and $\theta_{\kappa,u}$ are related to the climatology $c(\boldsymbol{s})$.

Finally, we have to assign prior distributions to the precisions τ_p and τ_r of the observations $\boldsymbol{y_j}$ and $\boldsymbol{z_j}$. We use the same prior distribution for τ_p and τ_r:

$$\tau_p \sim \text{gamma}(10, 10)$$
$$\tau_r \sim \text{gamma}(10, 10),$$

where the shape and rate parameters are both set to 10. This gives an expected prior mean of 1 and an expected prior variance of 0.1 for τ_p and τ_r.

The scales $\boldsymbol{f_j}$ from Equation (4.10) are then given by $(f_{1j}, ..., f_{ij}, ..., f_{nj})$ where

$$f_{ij} = \tau_p \cdot \text{var}(y_{ij}) = \text{var}(y_{ij}).$$

The precision τ_p is in the last equality replaced with its prior mean 1. Recall that y_{ij} is the observed precipitation at location i in year j and that the scales $\boldsymbol{f_j}$ allow each observation to have its own measurement uncertainty.

The scales $\boldsymbol{v_j}$ for runoff and evaporation are likewise given by $(v_{1j}, ..., v_{lj}, ..., v_{Nj})$ where

$$v_{lj} = \tau_r \cdot \text{var}(z_{lj}) = \text{var}(z_{lj}).$$

The precision τ_r is in the last equality replaced with its prior mean 1. Recall that z_{lj} is the sum of observed runoff and evaporation in Catchment l in year j and that the scales $\boldsymbol{v_j}$ allow each observation to have its own uncertainty.

5. Inference and evaluation

5.1 Implementation in R-INLA

We observe the annual precipitation at n chosen locations from Figure 2.2, and the runoff and evaporation within N catchments for r years. The observations are collected in the vector $\boldsymbol{y} = (\boldsymbol{y_1}, ..., \boldsymbol{y_r}, \boldsymbol{z_1}, ..., \boldsymbol{z_r})$. Thus, we have the observations $\boldsymbol{y}$, the latent Gaussian field $\boldsymbol{\psi} = (\boldsymbol{w_1}, ..., \boldsymbol{w_r}, \beta_1, ..., \beta_r, \boldsymbol{u})$ and the hyperparameters $\boldsymbol{\theta} = (\tau_r, \tau_p, \theta_{\kappa,w}, \theta_{\tau,w}, \theta_{\kappa,u}, \theta_{\tau,u})$. The observations are used to compute the posterior marginal distributions of the latent field and the hyperparameters, making predictions of precipitation at n^* locations and predictions of runoff within N^* catchments. The INLA methodology described in Section 3.5 is used for this purpose.

In Appendix B code examples are included showing how we make inference and predictions with our model by using R-INLA. We refer to Blangiardo and Cameletti (2015) and Krainski et al. (2016) for more examples and tutorials describing how R-INLA can be used for spatial modelling.

5.2 Parameter estimators and predictions

The posterior distribution is the probability distribution of a parameter of interest given the observations $\boldsymbol{y}$, but sometimes we are interested in point estimators instead of a distribution. The minimum mean square error (MMSE) estimator will be used as point estimator in this study (Kay, 1993). The MMSE estimator is defined as the mean of the posterior distribution of the parameter of interest, i.e the MMSE estimator $\hat{\beta}_1$ of parameter β_1 is given by

$$\hat{\beta}_1 = \int \beta_1 \pi(\beta_1|\boldsymbol{y}) d\beta_1 = E\{\beta_1|\boldsymbol{y}\}. \tag{5.1}$$

We can replace β_1 by a vector containing one or several of the parameters in $\boldsymbol{\psi}$ or $\boldsymbol{\theta}$ to find its MMSE estimator. Often we call $\hat{\beta}_1$ the posterior mean of β_1.

Further, the MMSE estimator is used for making predictions of annual precipitation and runoff. We use Equation (4.7) for true precipitation and replace the parameters by the corresponding MMSE estimators. The predicted annual precipitation in year j at location $\boldsymbol{s_i}$ is then given by

$$\hat{\eta}_j(\boldsymbol{s_i}) = \hat{\beta}_j + \Sigma_{k=1}^m \xi_k(\boldsymbol{s_i})(\hat{w}_{kj} + \hat{u}_k) = \hat{\beta}_j + \hat{c}(\boldsymbol{s_i}) + \hat{x}_j(\boldsymbol{s_i}) \tag{5.2}$$

The predicted amount of runoff in year j in the unobserved catchment $\mathcal{D}_l$ is likewise given by

$$\hat{\zeta}_j(\mathcal{D}_l) = \Delta_l \sum_{\boldsymbol{s}_i \in \mathcal{L}_{\mathcal{D}_l}} \hat{\eta}_j(\boldsymbol{s_i}) - E_{lj}. \tag{5.3}$$

Note that we need to compute the posterior mean of precipitation $\hat{\eta}_j(\boldsymbol{s_i})$ in all of the grid nodes of the discretisation of the catchment $\mathcal{L}_{\mathcal{D}_l}$ to make a runoff prediction. Also note that the annual evaporation E_{lj} is not estimated by the MMSE estimator. We regard E_{lj} as a fixed quantity which is given by NVE. However, its uncertainty is taken into account in the scales $\boldsymbol{v_j}$ as explained in Section 4.5.

5.3 Evaluation of the predictive performance

When analysing the LGM for precipitation and runoff we explore its ability to make accurate predictions in space. We leave out observations of runoff and precipitation from locations of interest, and use observations from other locations to reproduce the observations left out by using Equations (5.2) and (5.3).

To evaluate the predictive performance of (5.2) and (5.3), two criterion are used: The root-mean-square error (RMSE) and the continuous ranked probability score (CRPS).

The RMSE is defined as

$$\text{RMSE}_j = \sqrt{\frac{1}{n^*}\sum_{i=1}^{n^*}(\hat{\eta}_j(\boldsymbol{s_i}) - y_{ij})^2},$$

where y_{ij} and $\hat{\eta}_j(\boldsymbol{s_i})$ are the observed and the predicted precipitation at location i in year j. The variable n^* is the number of locations at which we predict precipitation and is the same for all years or replicates. Similarly, the RMSE for runoff is

$$\text{RMSE}_j = \sqrt{\frac{1}{N^*}\sum_{l=1}^{N^*}(\hat{\zeta}_j(\mathcal{D}_l) - (z_{lj} - E_{lj}))^2} = \sqrt{\frac{1}{N^*}\sum_{l=1}^{N^*}(\hat{\zeta}_j(\mathcal{D}_l) - \zeta^*_{lj})^2},$$

with $\hat{\zeta}_j(\mathcal{D}_l)$ and ζ^*_{lj} being the predicted and the observed runoff, and N^* being the number of catchments in which we predict the annual runoff.

Finally, we use the mean of all years r of either precipitation or runoff as a measurement of the predictive performance

$$\overline{\text{RMSE}} = \frac{1}{r}\sum_{j=1}^{r} RMSE_j,$$

and denote them as $\overline{\text{RMSE}}_{\text{precip}}$ and $\overline{\text{RMSE}}_{\text{runoff}}$ respectively.

The second criterion used to evaluate the predictive performance, is the CRPS which in general is defined as

$$\text{CRPS}(F, o) = \int_{-\infty}^{\infty}(F(u) - 1\{o \leq u\})^2 du,$$

where F is the predictive cumulative distribution function and o is the observed value of either y_{ij} for precipitation or ζ^*_{lj} for runoff (Gneiting and Raftery, 2007). As opposed to the RMSE, the CRPS criterion takes the whole posterior predictive distribution into account. The `R`-function `crps()` included in the package `verification` is used to compute the CRPS. In

this function it is assumed that the posterior distribution F is Gaussian with the posterior mean and the posterior standard deviation of precipitation or runoff as parameters.

The mean of the CRPS for all years of either precipitation or runoff is used as a measurement of the predictive performance

$$\overline{\text{CRPS}}_{\text{precip}} = \frac{1}{r}\sum_{j=1}^{r}\frac{1}{n^*}\sum_{i=1}^{n^*} CRPS(F_{ij}, y_{ij})$$

$$\overline{\text{CRPS}}_{\text{runoff}} = \frac{1}{r}\sum_{j=1}^{r}\frac{1}{N^*}\sum_{l=1}^{N^*} CRPS(F_{lj}, \zeta^*_{lj}).$$

A low value of $\overline{\text{RMSE}}$ or $\overline{\text{CRPS}}$ indicates an accurate prediction.

6. Observation designs and experimental set-up

As mentioned before, we will not use the whole dataset presented in Section 2 to test the statistical model for annual precipitation and runoff. We focus on observations from the area around Voss, and use them to answer the five research tasks presented in the introduction. In this section we present the observation designs and the experiments performed to answer the research tasks.

Observation design

Figure 6.1 shows all of the observations that will be used for spatial predictions. To answer research tasks 1-5 we use different combinations of the observations displayed here. For most experiments, we predict the runoff within Catchment 9 and/or the precipitation at Reimegrend by using Equations (5.2) and (5.3). The predictive performance is evaluated by the RMSE and the CRPS criterion.

We use observations from 1984 to 1993 and perform spatial predictions for the same time period, i.e we have 10 replicates. These 10 years are chosen because observations of annual runoff, evaporation and precipitation for most of the locations and catchments in Figure 6.1 are available for these years. Only one observation is missing and that is the annual precipitation at Reimegrend in 1989.

In addition to the observations shown in Figure 6.1, we construct three artificial catchments by using combinations of Catchments 7, 8 and 9. These are named Catchment (9-8), (9-7) and (8-7) and are shown in Figure 6.2. The runoff within Catchment (9-8) is for example computed by subtracting the observed runoff within Catchment 8 from the observed runoff within Catchment 9. The evaporation of Catchment (9-8) is likewise computed by subtracting the observed evaporation of Catchment 8 from the observed evaporation within Catchment 9.

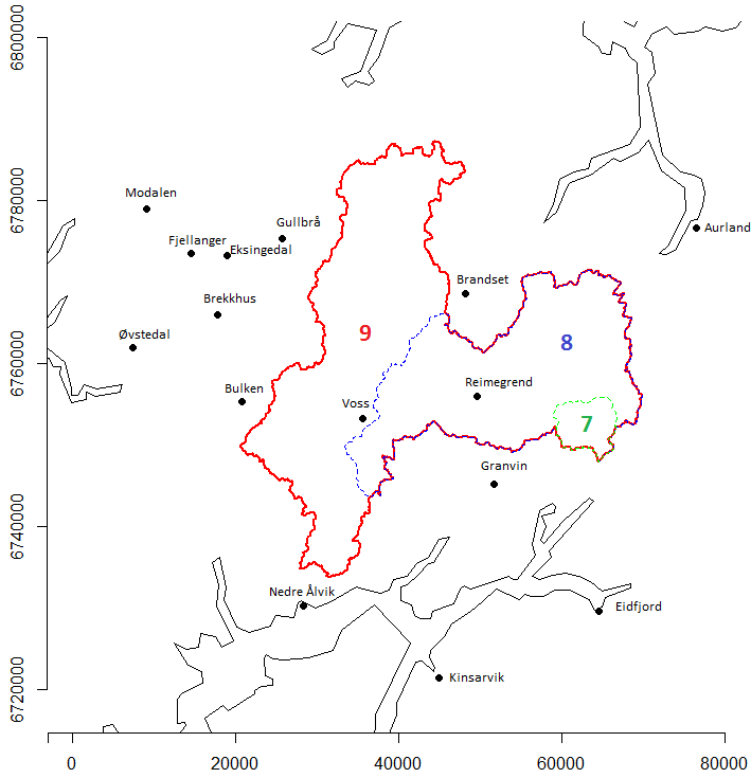
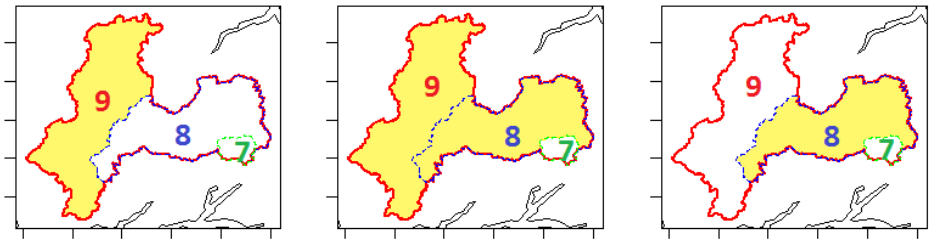


Figure 6.1: Observations used for inference and predictions. There are observations of precipitation from 15 locations, and observations of runoff from 3 catchments from 10 years.



(a) Catchment (9-8) (b) Catchment (9-7) (c) Catchment (8-7)

Figure 6.2: Combinations of Catchments 7, 8 and 9 are used to make three artificial catchments labelled (9-8), (9-7) and (8-7). The new catchments are marked in yellow.

RT1: The climatology

Research task 1 was to explore how the climatology of the study area affects the spatial predictions of annual runoff and precipitation. The properties of the climatology $c(\boldsymbol{s})$ are explored by predicting the runoff and the annual precipitation within Catchment 9 by using two different observation samples:

Observation sample 1: Annual precipitation from 12 locations. These are Granvin, Voss, Reimegrend, Bulken, Nedre Ålvik, Øvstedal, Brekkhus, Fjellanger, Eksingedal, Gullbrå, Aur-

land and Brandset.
Observation sample 2: Annual runoff and evaporation from Catchment 8.

We plot the posterior means of the climatology $c(\boldsymbol{s})$ and the annual part of the model $x_j(\boldsymbol{s})+\beta_j$ within Catchment 9, and consider the spatial patterns obtained from the predictions. These experiments are done by using the full precipitation model, i.e Model 1 from Equation (4.3) and we use real data. We also redo the predictions by using Model 2 from Equation (4.4), i.e we omit $x_j(\boldsymbol{s})$ from the model to see if the predictive performance changes.

To further explore the importance of the climatology, we perform a calibration test. A model is said to be calibrated if 20%, 40%, 60% and 95% of the true observations of precipitation or runoff lie inside 20%, 40%, 60% and 95% posterior prediction intervals. In the calibration test we predict the runoff within Catchment 8 and the precipitation at Reimegrend. Two different observation samples are used for this purpose:

Observation sample 3: Annual precipitation from all 14 locations in Figure 6.1 except Reimegrend.
Observation sample 4: Annual runoff and evaporation from Catchment 7.

For the calibration test we use simulated values of precipitation and runoff. We simulate 3 different climatologies $c(\boldsymbol{s})$. For each of the 3 climatologies we sample 150 replicates of $x_j(\boldsymbol{s})$ and β_j. The calibration test is then performed three times on the model, one time for each of the three climatologies. Ten and ten replicates are processed at the time.

The parameter values used for simulation are the posterior means obtained when making inference by using point and area observations originating from the same area. Thus, the parameter values used for simulation are realistic, and they are included in Appendix A.1.

RT2: Observation designs for points or areas

Research task 2 was: If we find an observation design suitable for runoff predictions in a specific catchment, will the same observation design be suitable for point predictions of precipitation in the interior of the catchment? To answer this question, we consider the results from the calibration test from RT1. We search for examples of an observation design that produces calibrated predictions of runoff, but uncalibrated predictions of precipitation, or opposite.

RT3: Observing the runoff within an overlapping catchment

Research task 3 was to explore how runoff predictions are affected by using observations of runoff from an overlapping catchment. To do this, we predict the runoff within Catchment 9 by using observations of runoff from one of the overlapping catchments in Figure 6.1 or 6.2, i.e Catchment 7, (8-7), 8, (9-8) or (9-7). Observations from only one of the overlapping catchments are used at the time, but we use observations from 10 years.

In this experiment we do the above experiment both by using real data and simulated datasets.

We make 9 simulated datasets, each with 10 replicates. Thus, we have 9 realizations of the climatology $c(\boldsymbol{s})$ and for each climatology there are 10 replicates of $x_j(\boldsymbol{s})$ and β_j just as for the real dataset.

In the presentation of the results, the predictive performance of runoff will be plotted against the *coverage percent* of the observed overlapping catchment. The coverage percent of Catchment 8 is for example 46.8% which means that the area of Catchment 8 covers 46.8 % of the area of Catchment 9. The coverage percent of Catchments 7, (8-7), (9-8) and (9-7) are 3.8 %, 43 %, 53.2 % and 96.2 %, respectively. See Table 2.1 for areas for all of the catchments.

RT4: Observing precipitation

Research task 4 was to explore how runoff predictions are affected by using observations of precipitation. To answer this, we predict the runoff within Catchment 9 by using observations of precipitation from 12 of the locations in Figure 6.1 from 10 years. These are: Reimegrend, Granvin, Brandset, Voss, Bulken, Nedre Ålvik, Brekkhus, Øvstedal, Fjellanger, Eksingedal, Gullbrå and Aurland.

We first perform the predictions by using the real dataset. Further, we redo the experiment by using the simulated datasets generated in the previous experiment (RT3). We compare the results from the real and the simulated datasets to see if they give similar results.

RT5: Observing precipitation and runoff

Research task 5 was to explore how runoff predictions are affected by using an observation sample consisting of both observations of annual runoff and precipitation. We predict the annual runoff within Catchment 9 as before, by using observations of precipitation from 11 locations: Granvin, Brandset, Voss, Bulken, Nedre Ålvik, Brekkhus, Øvstedal, Fjellanger, Eksingedal, Aurland and Gullbrå. The observations of precipitation are combined with observations of runoff from one of the overlapping catchments, i.e Catchment 7, (8-7), 8, (9-8) or (9-7). Observations from 10 years are used.

First, the experiment is performed by using the real dataset. We compare the predictive performance obtained when using only observations of precipitation, only observations of runoff and when combining the two observation types.

Further, we redo the experiment by using three of the ten simulated datasets from RT3 and RT4. This means that we redo the experiment for 3 simulated climatologies, each with 10 replicates of $x_j(\boldsymbol{s})$ and β_j.

For completeness of the results, we do the same experiment for predictions of precipitation at Reimegrend which is located in the middle of Catchment 9. These results are included in Appendix A.2. For RT5 we also include plots of the posterior distributions of the SPDE parameters $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$, $\theta_{\kappa,u}$, the precisions τ_p and τ_r and the intercepts for the different observation types when using real data. As the main focus of this report is on spatial

predictions, these plots are included in Appendix A.3.

Prior sensitivity

Finally, a simple analysis of the prior sensitivity of the parameters is done by trying 6 different priors for the parameters $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$. Recall that $\theta_{\tau,w}$ and $\theta_{\kappa,w}$ are related to the annual spatial variability $x_j(\boldsymbol{s})$, while $\theta_{\tau,u}$ and $\theta_{\kappa,u}$ are related to the climatology $c(\boldsymbol{s})$.

The 6 priors used are displayed in Table 6.1. Prior 1 is the prior used in the rest of the experiments, and was chosen by choosing 33 km and 110 km as the 0.5- and 0.9- quantiles for the range of the spatial fields, and 0.8 m/year and 2 m/year as the 0.5- and 0.9-quantiles for the spatial standard deviations in Equations (4.18) and (4.19). The five remaining priors are chosen by changing these quantiles, and the quantiles are displayed in Table 6.1.

We explore how the priors change the posterior distributions for all model parameters by making inference based on observations of runoff from Catchment 8 and observations of precipitation from the following 11 locations: Brandset, Aurland, Granvin, Voss, Nedre Ålvik, Bulken, Øvstedal, Brekkhus, Fjellanger, Eksingedal and Gullbrå.

The runoff within Catchment 9 and the precipitation at Reimegrend are also predicted to see if the resulting predictions are affected by the choice of prior.

Table 6.1: Six different prior distributions for the SPDE parameters.

	Prior 1	Prior 2	Prior 3	Prior 4	Prior 5	Prior 6
$\theta_{\tau,u}$	$\mathcal{N}(1.4, 0.61^2)$	$\mathcal{N}(0.91, 0.47^2)$	$\mathcal{N}(2.1, 0.61^2)$	$\mathcal{N}(2.3, 0.67^2)$	$\mathcal{N}(1.6, 0.47^2)$	$\mathcal{N}(2.3, 0.67^2)$
$\theta_{\kappa,u}$	$\mathcal{N}(-2.5, 0.94^2)$	$\mathcal{N}(-2.0, 0.86^2)$	$\mathcal{N}(-2.5, 0.94^2)$	$\mathcal{N}(-3.3, 0.25^2)$	$\mathcal{N}(-1.96, 0.86^2)$	$\mathcal{N}(-3.3, 0.25^2)$
$\theta_{\tau,w}$	$\mathcal{N}(1.4, 0.61^2)$	$\mathcal{N}(0.91, 0.47^2)$	$\mathcal{N}(2.1, 0.61^2)$	$\mathcal{N}(2.3, 0.67^2)$	$\mathcal{N}(2.3, 0.67^2)$	$\mathcal{N}(1.6, 0.47^2)$
$\theta_{\tau,w}$	$\mathcal{N}(-2.5, 0.94^2)$	$\mathcal{N}(-2.0, 0.86^2)$	$\mathcal{N}(-2.5, 0.94^2)$	$\mathcal{N}(-3.3, 0.25^2)$	$\mathcal{N}(-3.3, 0.25^2)$	$\mathcal{N}(-1.96, 0.86^2)$

Table 6.2: Quantiles used for assigning six different priors to the SPDE parameters as explained in Section 4.5. The unit of $\rho(\cdot)$ is km and the unit of $\sigma(\cdot)$ is m/year.

GRF	Quantiles	Prior 1	Prior 2	Prior 3	Prior 4	Prior 5	Prior 6
$c(\boldsymbol{s})$	$\rho(0.5)$	33	20	33	80	20	80
	$\rho(0.9)$	110	60	110	110	60	110
	$\sigma(0.5)$	0.8	0.8	0.4	0.8	0.4	0.8
	$\sigma(0.9)$	2	2	1	2	1	2
$x_j(\boldsymbol{s})$	$\rho(0.5)$	33	20	33	80	80	20
	$\rho(0.9)$	110	60	110	110	110	60
	$\sigma(0.5)$	0.8	0.8	0.4	0.8	0.8	0.4
	$\sigma(0.9)$	2	2	1	2	2	1

7. Analysis of the results

RT1: The climatology

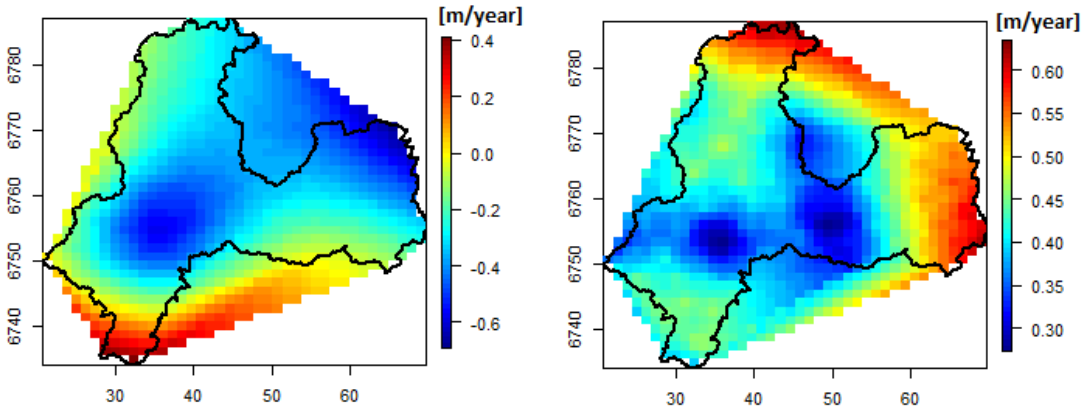
Research task 1 was to explore how the climatology of the study area affects the spatial predictions of runoff and precipitation. To answer research task 1 we first consider the posterior climatology within Catchment 9 for observation sample 1 where the observation sample consists of real observations of precipitation from 12 locations from 10 years. The climatology is shown in Figure 7.1. We compare this spatial pattern with the spatial pattern in Figure 7.2 which shows the posterior annual precipitation within Catchment 9 for 1984, 1988 and 1993. The spatial patterns in Figure 7.1 and 7.2 are very similar. Thus, the climatology $c(\boldsymbol{s})$ seems to be dominating over the annual spatial variability $x_j(\boldsymbol{s})$ in Catchment 9.

This is verified by Figure 7.3 where the posterior annual spatial variability for the same three years is shown. We see that the annual spatial variability is very low in 1993 and only determined by the annual intercept β_j. In 1984 we see more annual spatial variation, while in 1988 the annual spatial variations are quite large. The difference from the smallest to the largest value of the posterior annual variability in 1988 is approximately 0.6 m/year. The posterior marginal standard deviation of the GRFs $x_j(\boldsymbol{s})$ and $c(\boldsymbol{s})$ were $\hat{\sigma}_w = 0.19$ m/year and $\hat{\sigma}_c = 0.72$ m/year respectively. Thus, the climatology is the dominating GRF, but small annual spatial variations are also present within Catchment 9.

The results can be understood by considering the observation sample used (observation sample 1) which is displayed in Figure 7.4. The ranking between the locations, from the location with most precipitation to the location with least precipitation is the same from year to year with some exceptions. The lines are almost parallel as well. These tendencies will be explained by the climatology $c(\boldsymbol{s})$ and the lines show that the spatial variations of precipitation around Voss are quite stable. However, the plot also shows annual differences, but these are not that prominent. For example we see that in years with a lot of precipitation (1988, 1989) the difference between the smallest and the largest value in the observation sample is larger than in years with less precipitation (1984, 1993). The year 1988 is the most extreme year in the observation sample, and also the year in which the spatial annual variations were largest in Figure 7.3.

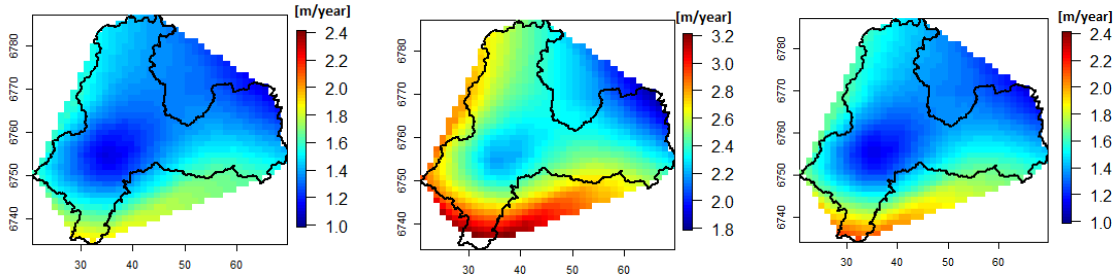
Further, we consider the posterior means for observation sample 2 which only consists of runoff observations from Catchment 8 from 10 years. Figure 7.5 shows the posterior climatology, Figure 7.6 shows the posterior annual precipitation for three years and Figure 7.7 shows the corresponding spatial annual variability added with the intercept. Again the annual spatial variability is low, close to non-existing. The difference from the previous observation design is that when the observation sample only consists of runoff observations from one catchment, the spatial variations explained by the climatology also are small as shown in Figure 7.5.

The images of the posterior means indicate that the climatology $c(\boldsymbol{s})$ dominates over the annual spatial variability. This is verified by considering the posterior prediction intervals in Figure 7.8. We see that the posterior quantiles are approximately equal whether $x_j(\boldsymbol{s})$ is



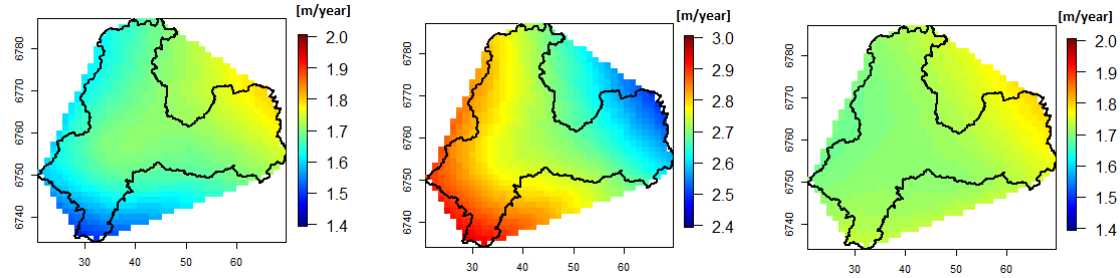
(a) Posterior climatology. (b) Posterior standard deviation.

Figure 7.1: Posterior climatology $\hat{c}(\boldsymbol{s})$ in Catchment 9 and its posterior standard deviation when the observation sample consists of observations of precipitation from 12 locations from 10 years.



(a) 1984 (b) 1988 (c) 1993

Figure 7.2: Posterior mean for precipitation $\hat{\eta}_j(\boldsymbol{s}) = \hat{c}(\boldsymbol{s}) + \hat{x}_j(\boldsymbol{s}) + \hat{\beta}_j$ in Catchment 9 when the observation sample consists of observations of precipitation from 12 locations from 10 years. Note that the scale is different in the plots, but the difference between the lowest and the largest value in the colorbar is 1.4 m/year.



(a) 1984 (b) 1988 (c) 1993

Figure 7.3: Posterior mean for the year dependent part of the precipitation model, $\hat{\beta}_j + \hat{x}_j(\boldsymbol{s})$, in Catchment 9 for observation sample 1. The observation sample consists of observations of precipitation from 12 locations from 10 years. Note that the scale is different in the plots, but the difference between the lowest and the largest value in the colorbar is 0.6 m/year.

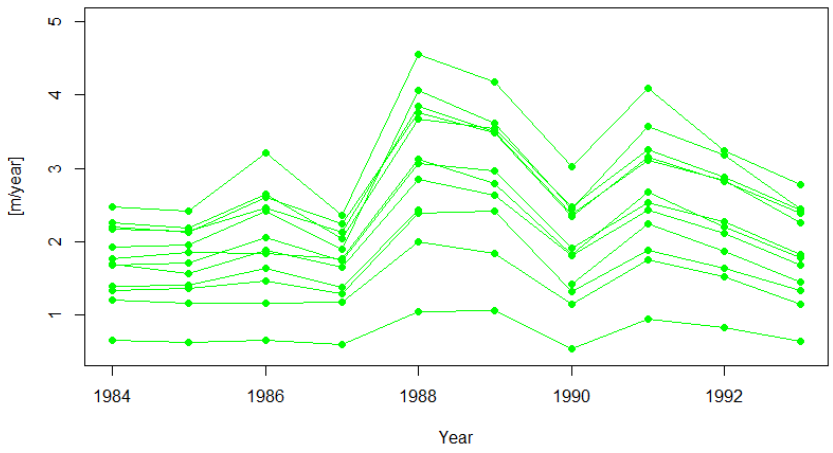
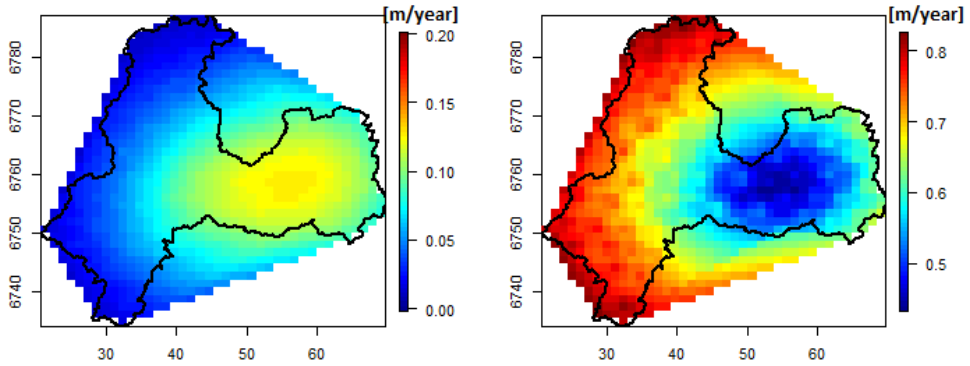
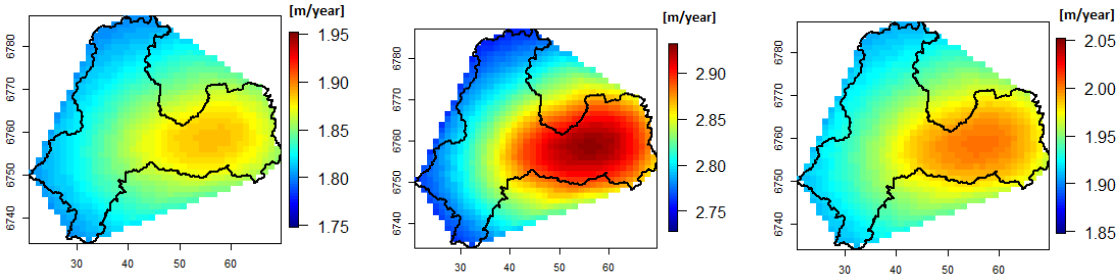


Figure 7.4: Observation sample 1: Observations of precipitation from 12 locations from 10 years. Each line corresponds to a location.



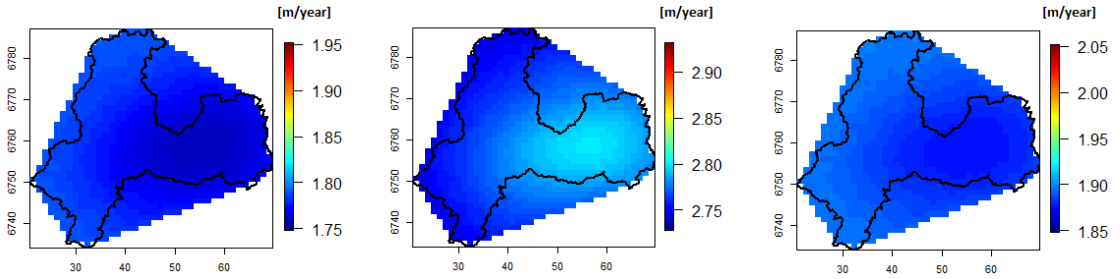
(a) Posterior climatology. (b) Posterior standard deviation.

Figure 7.5: Posterior climatology $\hat{c}(\boldsymbol{s})$ in Catchment 9 and its posterior standard deviation when the observation sample consists of runoff observations from Catchment 8 from 10 years.



(a) 1984 (b) 1988 (c) 1993

Figure 7.6: Posterior mean for precipitation $\hat{\eta}_j(\boldsymbol{s}) = \hat{c}(\boldsymbol{s}) + \hat{x}_j(\boldsymbol{s}) + \hat{\beta}_j$ in Catchment 9 when the observation sample consists of runoff observations from Catchment 8 from 10 years. Note that the scale is different in the three plots, but the difference between the smallest and the largest value in the colorbar is 0.2 m/year for all scalings.

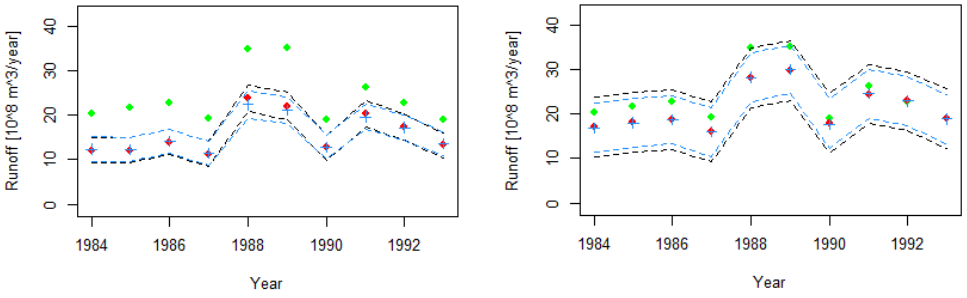


(a) 1984 (b) 1988 (c) 1993

Figure 7.7: Posterior mean for the year dependent part of the precipitation model, $\hat{\beta}_j + \hat{x}_j(\boldsymbol{s})$, in Catchment 9 for observation sample 2. The observation sample consists of runoff observations from Catchment 8 from 10 years. Note that the scale is different in the three plots, but the difference between the smallest and the largest value in the colorbar is 0.2 m/year for all scalings.

included in the model (Model 1) or not (Model 2). The largest differences between Model 1 and Model 2 are observed in Figure 7.8a in 1988 and 1989 which were the most extreme years in the observation sample.

The results from the experiments on the real dataset indicate that the spatial variations of precipitation in Voss are quite stable from one year to another. As the annual spatial variations are close to zero, the only annual variation of importance is the intercept β_j which adjusts the annual level of precipitation.



(a) Observation sample 1: Observations of precipitation from 12 locations.

(b) Observation sample 2: Runoff observations from Catchment 8.

Figure 7.8: 95% posterior prediction intervals for annual runoff within Catchment 9 when using Model 1 (black) and Model 2 (blue). Posterior mean for Model 1 (red), posterior mean for Model 2 (blue) and true observations (green). Real data are used.

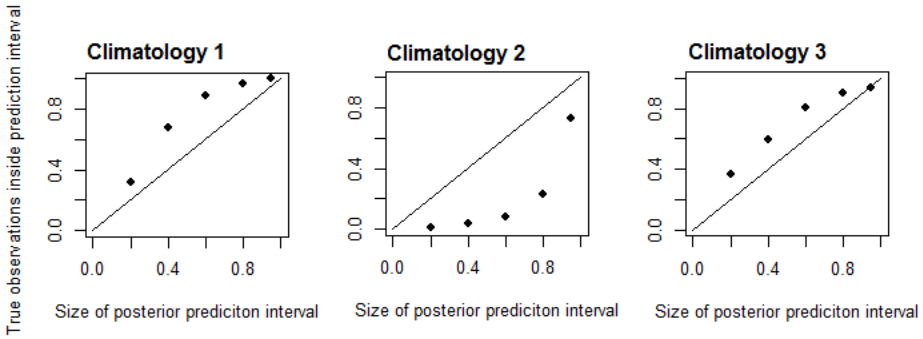
The importance of the climatology is further explored by using simulated values of precipitation and runoff, and the results from the calibration tests are displayed in Figures 7.9 and 7.10. The results show that the predictions of annual precipitation and runoff often are uncalibrated as long as the climatology is constant. This yields both when only using point observations (Figure 7.9) and when only using runoff observations from one small catchment (Figure 7.10).

For example for Climatology 2 in Figure 7.10a we see that the coverage of the prediction intervals of precipitation tends to be larger than expected, while for Climatology 1 the coverage of the prediction intervals in general is low with less than 10% of the true values within a 60% posterior prediction interval. In Figure 7.10b the model is close to calibrated for Climatology 1 and 2, but for Climatology 3 the coverage of the prediction intervals is in general poor.

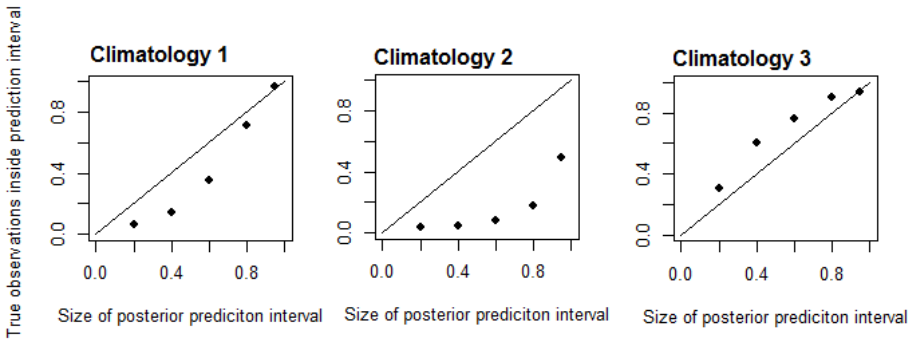
From a statistical point of view it is intuitive that the predictions are uncalibrated as long as the climatology is constant. Because the annual spatial variability $x_j(\boldsymbol{s})$ is close to zero, an accurate spatial prediction relies on an accurate prediction of $c(\boldsymbol{s})$, and as long as the climate is constant, we predict the same replicate of $c(\boldsymbol{s})$ again and again, from one year to the next. The predictions of runoff and precipitation would probably be calibrated if a new realization of the climatology $c(\boldsymbol{s})$ was drawn each time we performed a prediction.

The calibration tests emphasise that the underlying climatology has a large impact on the predictions of runoff and precipitation in areas like Voss where the spatial differences in precipitation are stable from one year to another. If an observation design produces biased predictions for one of the replicates, it is likely that poor predictions will be obtained for the remaining replicates as long as the same observation design is used. As long as the climatology is constant, we do the same systematic mistake in the predictions from one year to the next.

This model property is very important if we would like to use our model to produce flood warnings or hydrological forecasts in the Voss area. If we find a suitable observation design for past events based on available data, this observation design will also produce an accurate representation of the underlying climatology $c(\boldsymbol{s})$. As the climatology was shown to be the dominant component of the model, we therefore obtain a lot of information of future events.

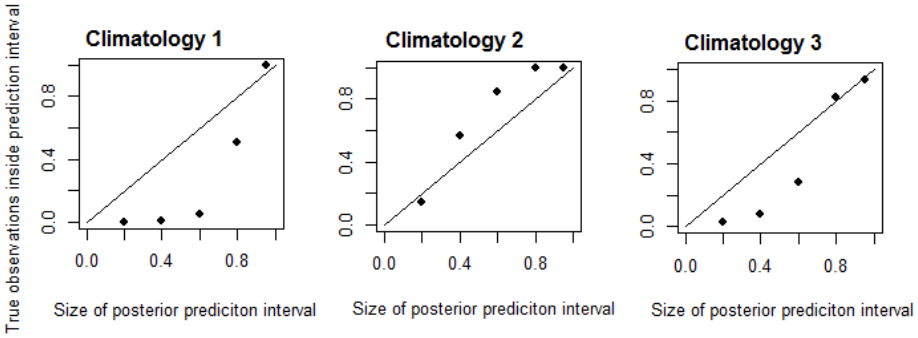


(a) Predictions of precipitation

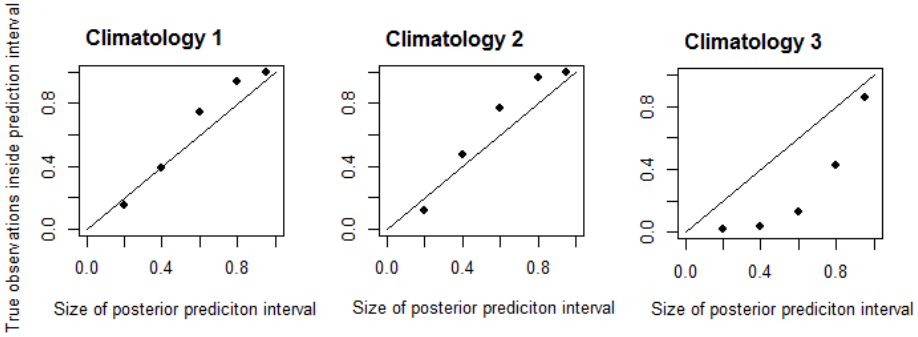


(b) Predictions of runoff

Figure 7.9: The amount of the true observations of precipitation that lies inside posterior prediction intervals of different sizes for the simulation study. Three simulated climatologies $c(\boldsymbol{s})$ are used and for each climatology there are 150 replicates of $x_j(\boldsymbol{s})$ and β_j. The observation sample consists of observations of precipitation from 14 locations.



(a) Predictions of precipitation



(b) Predictions of runoff

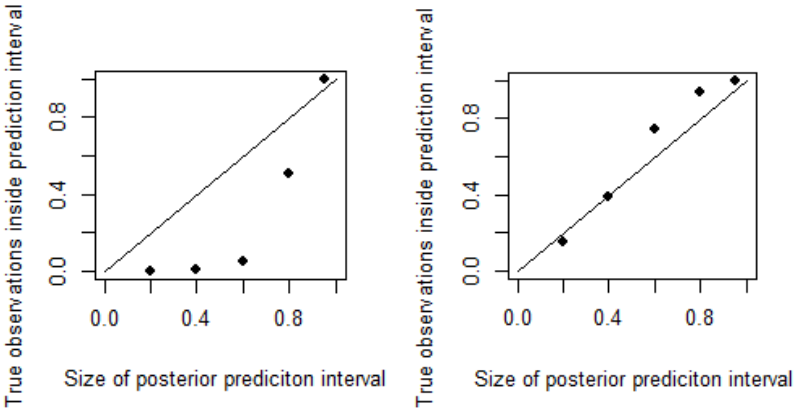
Figure 7.10: The amount of the true observations of precipitation that lies inside posterior prediction intervals of different sizes for the simulation study. Three simulated climatologies $c(\boldsymbol{s})$ are used and for each climatology there are 150 replicates of $x_j(\boldsymbol{s})$ and β_j. The observation sample consists of observations of runoff and evaporation from Catchment 7.

RT2: Observation designs for points or areas

The second research task was: If we find an observation design suitable for runoff predictions in a specific catchment, will the same observation design be suitable for point predictions of precipitation in the interior of the catchment? Already in the previous section we saw that the answer of research task 2 is no. The clearest proof is shown by comparing Figure 7.10a and 7.10b for Climatology 1, and we display the figures of interest again in Figure 7.11. Here, observations of runoff from Catchment 7 are used to predict the precipitation at Reimegrend and the runoff within Catchment 8. While this observation design provides calibrated predictions of the runoff within Catchment 8, the same observation design provides poor and biased predictions of annual precipitation at Reimegrend, although Reimegrend is located in the middle of Catchment 8.

To understand this, we consider Figure 7.1 where the simulated Climatology 1 is shown. We see that Catchment 7 is a catchment in which the annual precipitation is in general medium or low. Although there are areas with larger values of precipitation within Catchment 8, the annual precipitation within Catchment 7 is representative for the mean annual precipitation within Catchment 8. Thus, we obtain calibrated predictions of runoff for this observation design.

Further, we see that the annual precipitation at Reimegrend is very high compared to the annual precipitation within Catchment 7. Thus, we obtain uncalibrated and biased predictions of precipitation at Reimegrend for the same observation design.



(a) Predicted precipitation at Reimegrend.

(b) Predicted runoff in Catchment 8.

Figure 7.11: The amount of the true observations of precipitation that lies inside posterior prediction intervals of different sizes for the simulation study. Observations of runoff and evaporation from Catchment 7 are used to predict the runoff within Catchment 8 and the precipitation at Reimegrend.

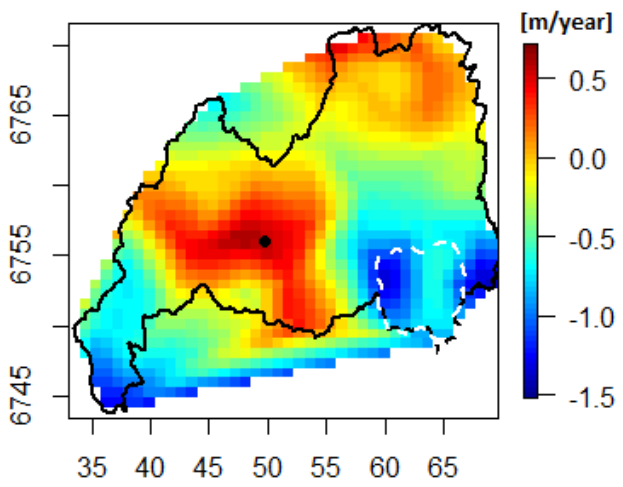


Figure 7.12: Simulated climatology (Climatology 1). Runoff observations from Catchment 7 (dashed) are used to predict precipitation at Reimegrend (black) and the runoff within Catchment 8 (black).

RT3: Observing the runoff within an overlapping catchment

Research task 3 was to explore how runoff predictions are affected by using an observation sample of runoff originating from an overlapping catchment. We first consider Figure 7.13 which shows the predictions obtained for annual runoff within Catchment 9 with corresponding 95 % posterior prediction intervals when using real data. The observation samples consist of runoff observations from one of the overlapping catchments from 10 years.

We note that the posterior prediction intervals become more and more narrow when the coverage percent of the overlapping observed catchment increases. This can be understood by considering Figure 7.14 which shows the distribution of the posterior standard deviation of precipitation within Catchment 9 for one of the replicates (1988).

In Figure 7.15 the RMSE and the CRPS for the real dataset are shown. The results are not as expected because we obtain a better prediction when using runoff observations from Catchment 7 than we get when using runoff observations from Catchment 8 or (8-7) according to the RMSE. However, using runoff observations from Catchment (9-7) results in very accurate predictions with RMSE and CRPS close to 0. This is as expected because Catchment (9-7) covers 96.2 % of Catchment 9, and the posterior standard deviation of precipitation becomes low in a larger area as seen in Figure 7.14. We saw in Figure 7.13 that this results in a narrow posterior prediction interval, where the true observations and the predicted ones are almost identical.

We now consider the results from the simulations to see if these correspond to the results from the real data. Figure 7.16 shows the CRPS for the simulated datasets when predicting the runoff within Catchment 9 by using observations from one of the overlapping catchments. The RMSE criterion shows the same tendency and is omitted. The plots in Figure 7.16 show that

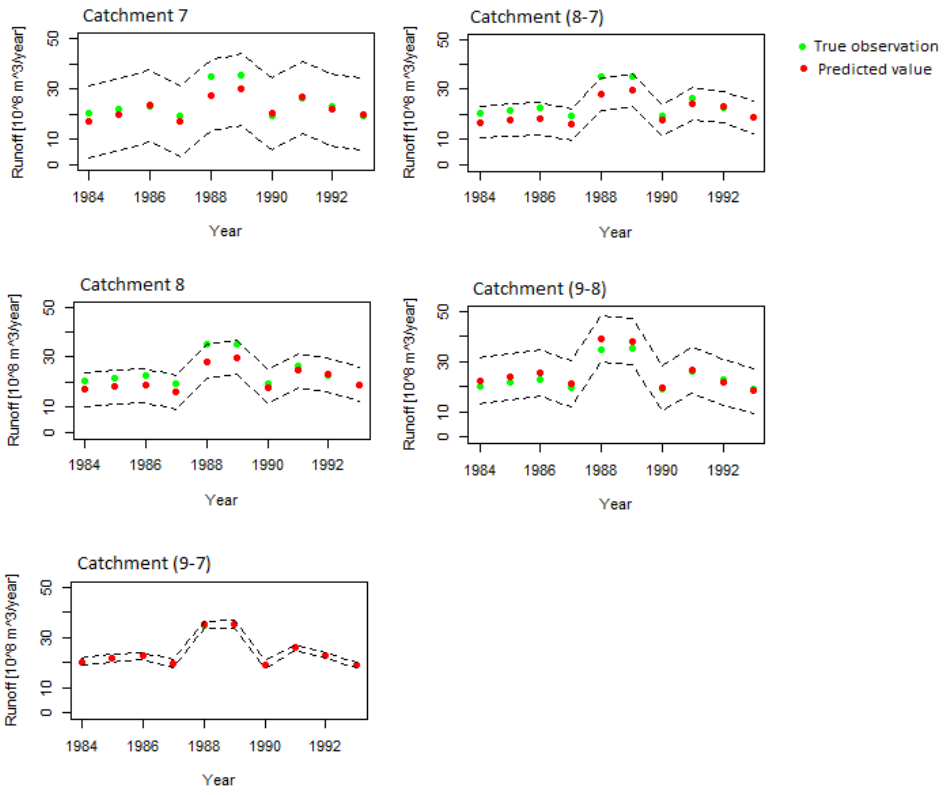
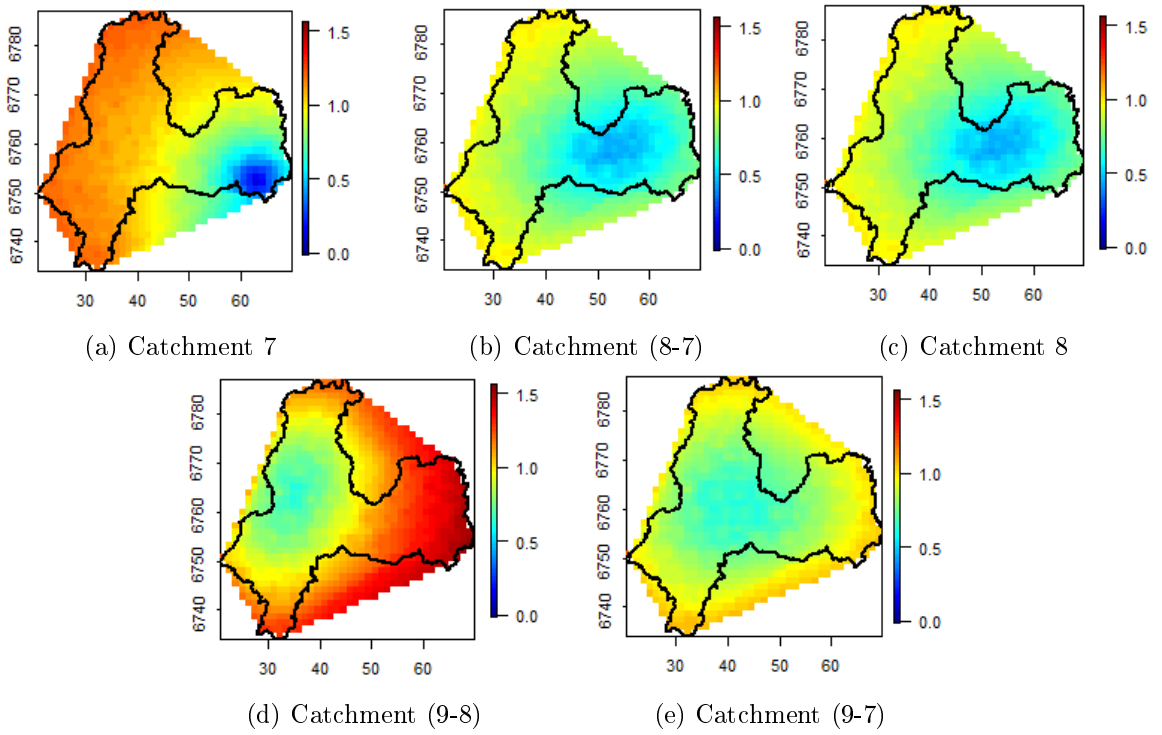


Figure 7.13: Posterior 95 % prediction interval for the runoff within Catchment 9 when the observation sample consists of runoff observations from an overlapping catchment. Real data from 10 years are used.

the predictions typically get more and more accurate as the coverage percent increases. Observing the runoff within Catchment (9-7) always gives a prediction that is significantly better than the prediction obtained when observing the runoff within Catchment 7, and for most of the simulations, runoff observations from Catchment (8-7), 8 or (9-8) provide lower CRPS values than runoff observations from Catchment 7. However, exceptions exist. In Simulation 8 in Figure 7.16 we see that runoff observations from Catchment 7 and (9-8) are approximately equally good. This simulation result corresponds to the results obtained when using real data.

In Figure 7.17 we display the 95% posterior prediction intervals for annual runoff within Catchment 9 when using observations of runoff from Catchment 7 for the simulated datasets. We chose to plot the prediction intervals when using observations from Catchment 7 because this catchment typically is the one giving the poorest predictions. In spite of this, Figure 7.17 shows that all of the true values lie within their 95% posterior prediction interval for all simulations. However we see that the predictions often are either over- or underestimated compared to the true value (Simulation 2, 3 and 8), but for some simulations the predictions are very accurate (Simulation 1, 4 and 7).

For some climatologies the annual precipitation within Catchment 7 is representative for the



(a) Catchment 7 (b) Catchment (8-7) (c) Catchment 8

(d) Catchment (9-8) (e) Catchment (9-7)

Figure 7.14: Posterior standard deviation for precipitation [m/year] in Catchment 9 in 1988 when the observation sample consists of only runoff observations from one of the overlapping catchments. Real data from 10 years are used.

annual precipitation within Catchment 9. This leads to accurate predictions as we saw for Simulation 1, 4 and 7 in Figure 7.17. For other climatologies the annual precipitation in Catchment 7 is not representative for the annual precipitation within Catchment 9. This leads to poor predictions as we saw for Simulation 2, 3 and 8 in Figure 7.17. Thus, it is difficult to know whether observations of runoff from a small overlapping catchment will provide good or poor predictions for a specific climatology. The results indicate that we can be lucky or unlucky with the observation sample. This is in particular important when the coverage percent of the overlapping catchment is small, because as the coverage percent increases we become more and more certain about the annual precipitation and runoff within the catchment of interest.

The objective of research task 3 was to explore how runoff predictions are affected by using observations of runoff from an overlapping catchment. Both the real dataset and the simulated ones showed that the true values of runoff within Catchment 9 typically are covered by the 95% posterior prediction intervals. Further, the experiments showed that the predicted value often is biased compared to the true value, but the bias decreases with increasing coverage percent. The results also showed that we can sometimes be lucky or unlucky with the observations, leading to unexpected results. This happened for the real dataset as the predictive performance was about equally good when using observations from Catchment 7 and 8, even though the coverage percentages were 3.8% and 46.8% respectively.

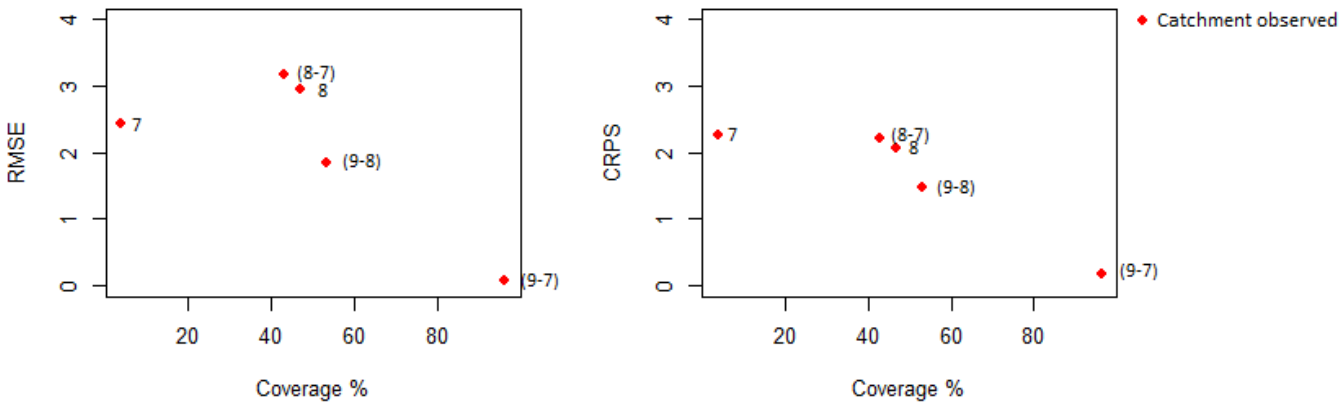


Figure 7.15: The predictive performance for the real dataset. The annual runoff within catchment 9 is predicted by using runoff observations from one of the overlapping catchments from 10 years.

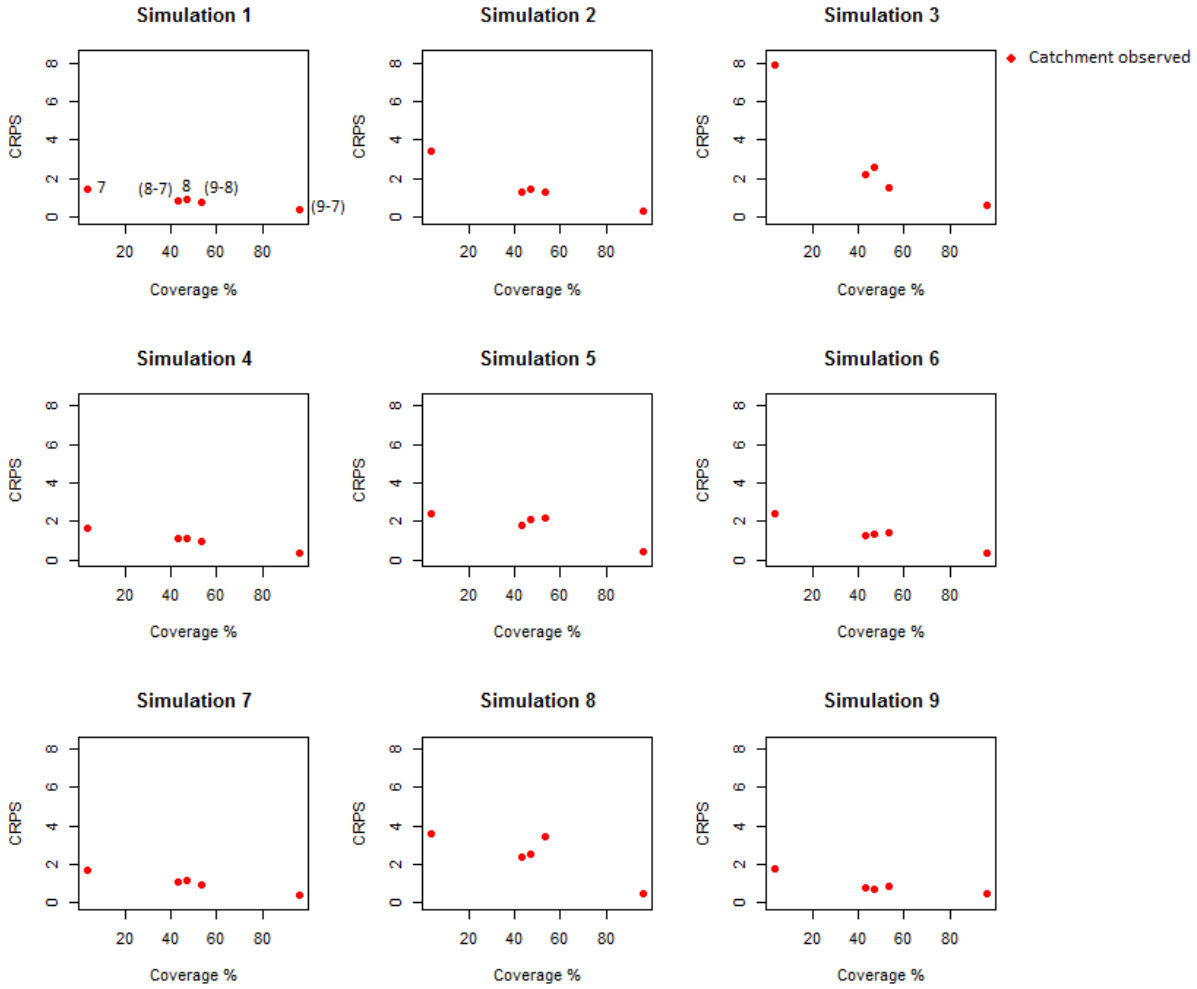


Figure 7.16: The predictive performance (CRPS) for the 9 simulated datasets. The runoff within Catchment 9 is predicted by using runoff observations from one of the overlapping catchments from 10 years.

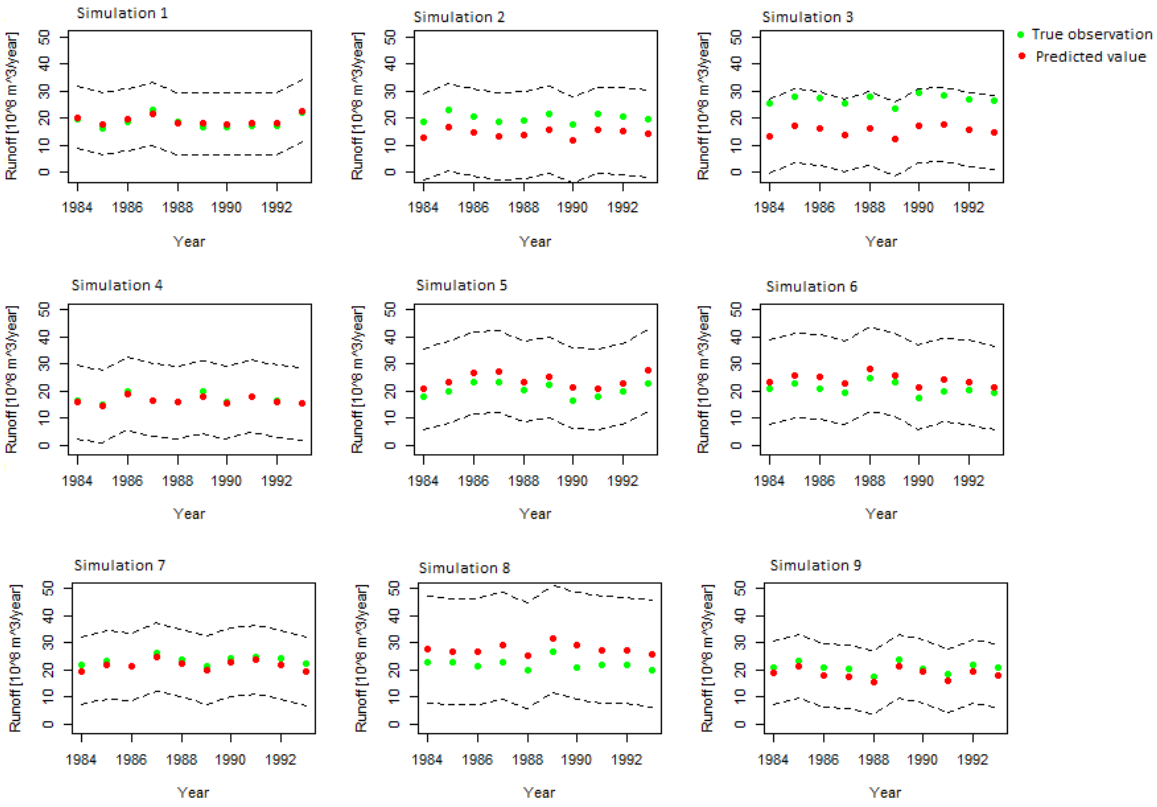


Figure 7.17: 95 % posterior prediction intervals for runoff within Catchment 9 when the observation sample consists of runoff observations from Catchment 7 from 10 years. The coverage percent of Catchment 7 is only 3.8 %. Simulated data are used.

RT4: Observing precipitation

Research task 4 was to explore how runoff predictions are affected by using an observation sample consisting of observations of precipitation. The predictive performance obtained when using observations of precipitation from 12 locations to predict the runoff within Catchment 9 were

$$\overline{\text{RMSE}}_{\text{runoff}} = 8.3$$
$$\overline{\text{CRPS}}_{\text{runoff}} = 7.4$$

for the real dataset. These values are more than twice as large than any of the RMSE or CRPS values obtained for the real dataset in Figure 7.15, and indicate that point observations are not suitable for runoff predictions within Catchment 9 for the climatology in Voss.

This is supported by Figure 7.18 which shows the corresponding 95% posterior prediction interval for the real dataset. The prediction interval shows that the predicted runoff within Catchment 9 is underestimated every year. The biases are large compared to the true observations, and the true observations are far outside the 95% posterior prediction interval.

The predicted runoff within Catchment 9 is most affected by the closest observations of precipitation. These are the observations from Gullbrå, Bulken, Reimegrend, Voss, Brandset, Granvin and Nedre Ålvik. These observations are displayed in Figure 7.19 along with the mean annual precipitation within Catchment 9. The figure shows that 6 of the locations provide values for annual precipitation that are lower than the mean annual precipitation within Catchment 9. We also see that the observations from the only two locations located in the interior of Catchment 9 are the ones giving the lowest annual values of precipitation. This explains why the point observations lead to an underestimation of the annual runoff within Catchment 9 for the real dataset.

Further, we consider the results from the simulation tests to see if other climatologies provide better results. Table 7.1 shows the predictive performance for the 10 simulated datasets. We

Table 7.1: Predictive performance when 12 point observations are used to predict the runoff within Catchment 9 for the simulated datasets.

Simulation	$\overline{\text{RMSE}}_{\text{runoff}}$	$\overline{\text{CRPS}}_{\text{runoff}}$
1	0.39	0.43
2	0.62	0.59
3	3.3	2.2
4	1.9	1.9
5	2.8	1.8
6	3.8	2.8
7	0.55	0.46
8	2.5	1.6
9	5.8	4.8

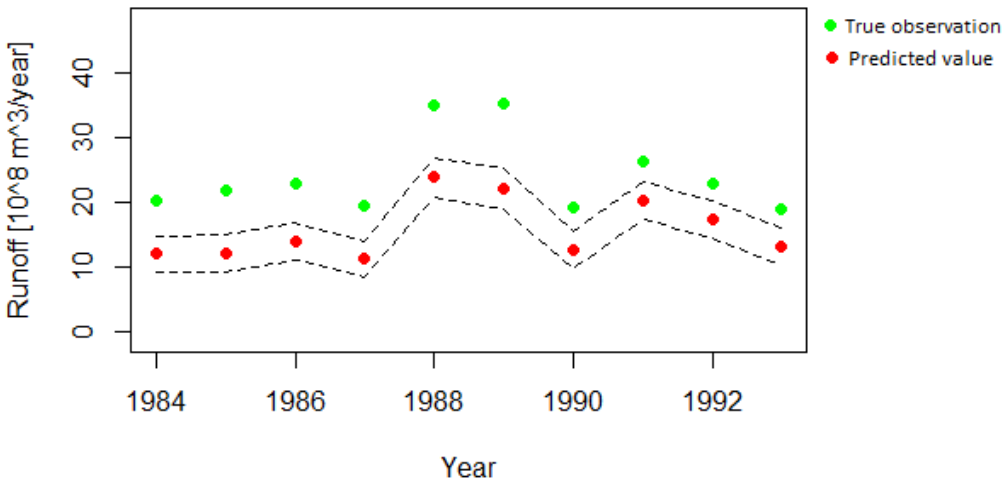


Figure 7.18: 95% prediction interval for real data when precipitation from 12 locations and 10 years are used to predict the runoff within Catchment 9.

see both low values (Simulation 1, 2, 7) and high values (Simulation 3, 6, 9) for $\overline{\text{RMSE}}_{\text{runoff}}$ and $\overline{\text{CRPS}}_{\text{runoff}}$.

In Figure 7.20 we show the mean annual precipitation in Catchment 9 (red) and the mean values of the observations from Gullbrå, Bulken, Reimegrend, Voss, Brandset, Granvin and Nedre Ålvik for the simulated datasets (yellow). These locations are chosen because they are located closest to Catchment 9. Comparing Figure 7.20 with the values in Table 7.1 we see that low RMSE and CRPS values typically correspond to closeness between the yellow and the red points in Figure 7.20. Thus, the mean values of the closest observations have a large impact on the resulting runoff prediction. We can both be unlucky or lucky with these mean values.

In Figure 7.21 the 95% posterior prediction intervals for the simulated datasets are shown. The posterior prediction intervals are quite narrow compared to the ones obtained when using observations of runoff from an overlapping catchment (Figure 7.13). Still, most of the true values are within the 95% posterior prediction intervals. Exceptions exists, e.g Simulation 6 and 9, but these predictions are less biased than the predictions obtained for the real dataset (Figure 7.18).

Research task 4 was to explore how runoff predictions are affected by using observations of precipitation. The simulation study shows that point observations of annual precipitation are suitable for runoff predictions for many climatologies. Under- and overestimations of runoff occur, and the biases are closely related to the mean values of the observations. Thus, we can be lucky or unlucky with the observation design.

The predictive performance obtained for the real dataset does not correspond to the simulation study because the predictions obtained for the real dataset were far more biased than any of the predictions from the simulation study.

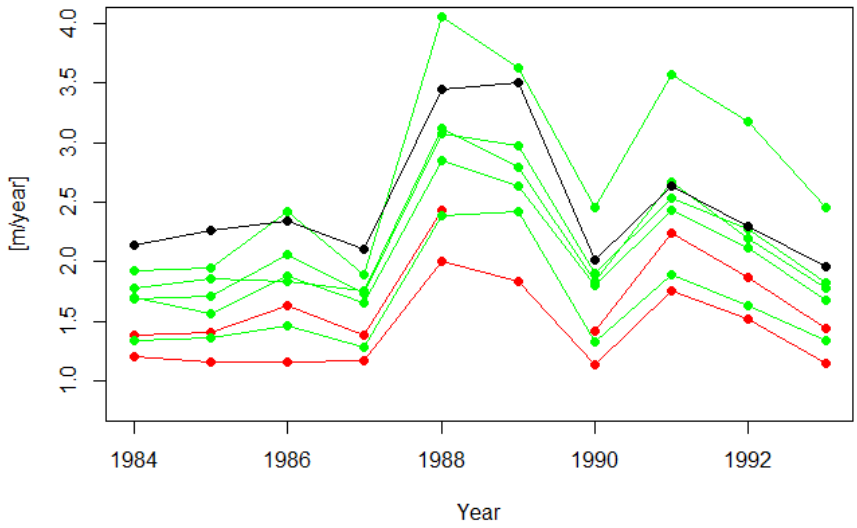


Figure 7.19: Black: The mean annual precipitation in each point in Catchment 9., i.e the (observed Runoff + evaporation)/(Catchment area) for Catchment 9. Green: Observations from Brandset, Nedre Ålvik, Gullbrå, Granvin and Bulken. Red: Reimegrend and Voss. These are inside Catchment 9 and are therefore marked in another color.

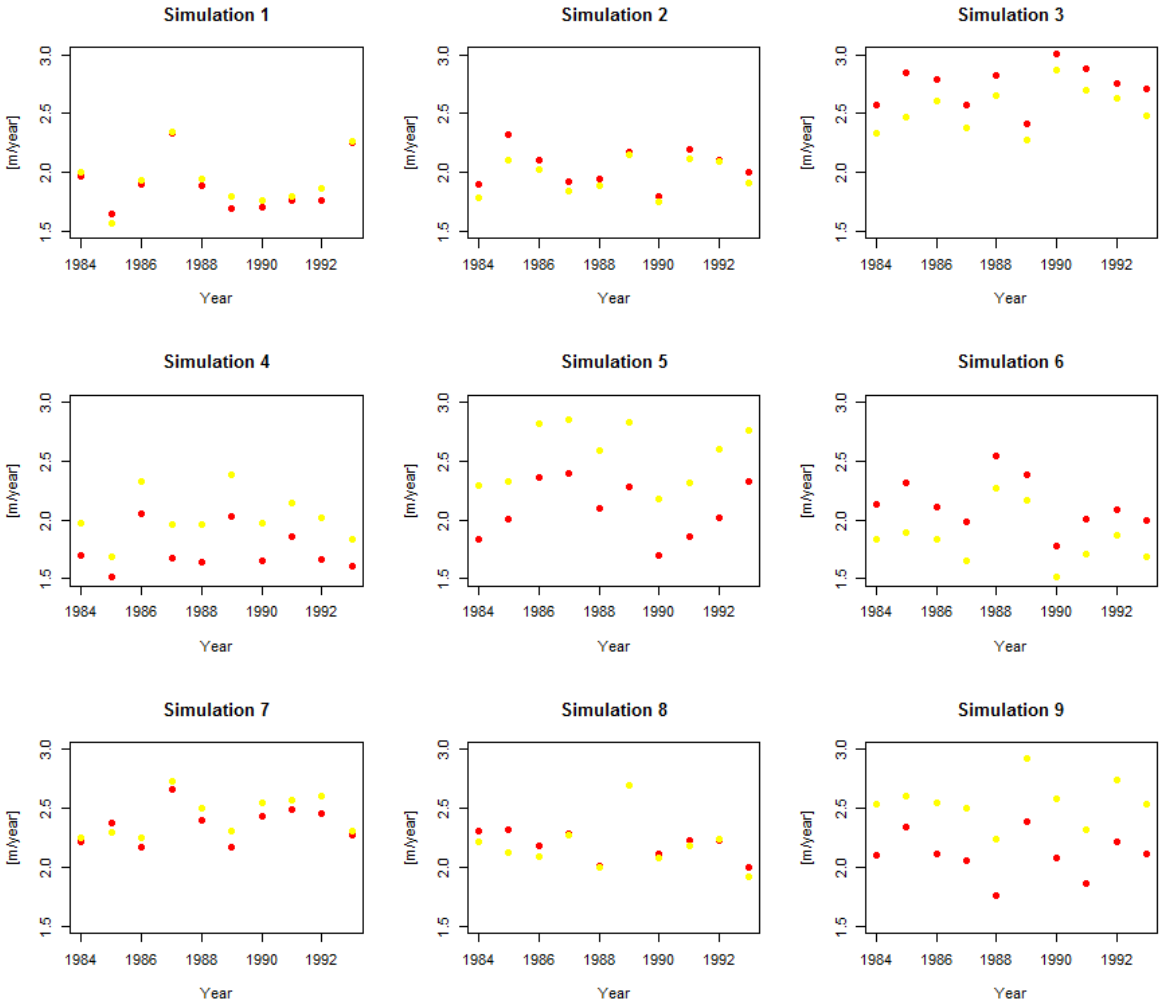


Figure 7.20: Red points: (Observed Runoff + evaporation)/(Catchment area) for Catchment 9. This is the mean annual precipitation in each point of Catchment 9. Yellow points: Mean values of annual precipitation for the 7 obervations located closest to Catchment 9. These are: Reimegrend, Brandset, Nedre Ålvik, Voss, Gullbrå, Granvin and Bulken.

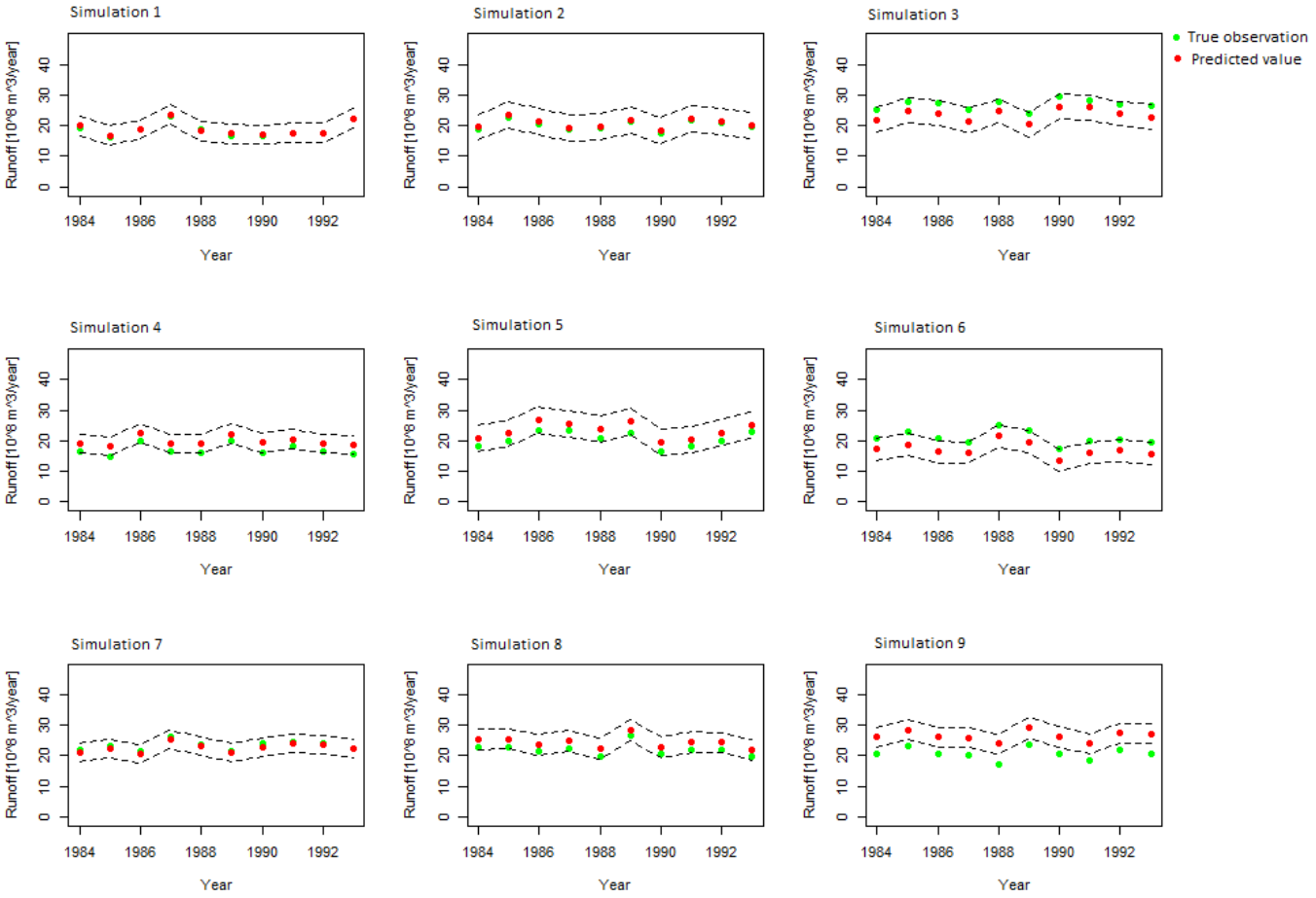


Figure 7.21: Posterior 95 % prediction interval for the simulated datasets when observations of precipitation at 12 locations are used to predict the runoff within Catchment 9.

RT5: Observing precipitation and runoff

Research task 5 was to explore how spatial predictions of runoff are affected by using an observation sample consisting of observations of both precipitation and observations of runoff from an overlapping catchment. Figure 7.22 shows the predictive performance when combining observations of runoff and precipitation for the real dataset. The observation samples consisting of only runoff observations perform better or approximately equally good as observation samples consisting of both runoff and precipitation for the real dataset: All of the red points in Figure 7.22 are below or at the same level as the corresponding blue point.

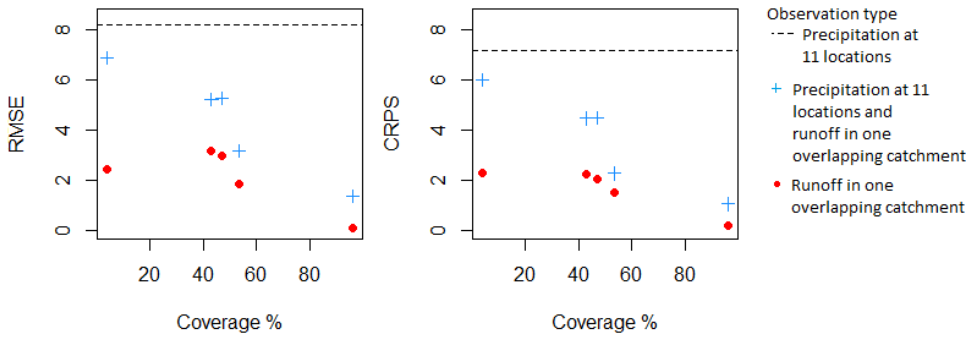


Figure 7.22: Predictive performance for runoff for the real dataset when the observation sample consists of runoff observations from an overlapping catchment.

Figure 7.23 shows posterior predictions intervals when the observation sample consists of annual runoff from one of the overlapping catchments for the real dataset. These prediction intervals are compared to the posterior prediction intervals obtained when using observations of runoff from an overlapping catchment and precipitation from 11 locations. When observations of precipitation are included in the observation sample, the posterior prediction intervals become narrow, and the majority of the true values of annual runoff within Catchment 9 fall outside the 95% prediction intervals. By adding point observations to an observation sample of only runoff observations, the resulting predictions both become more biased and the posterior standard deviation does not reflect the true uncertainty of the predictions. The conclusion from the real dataset is that combining observations of runoff and precipitation does not increase the predictive performance. From RT3 we know that point observations were not suitable for runoff predictions for the real dataset. This is probably the reason why an observation sample of only runoff observations performed better than an observation sample of both precipitation and runoff.

We now consider the results from the simulated datasets. In Figure 7.24 the predictive performance for runoff within Catchment 9 is shown for three simulations (with three different climatologies). For Simulation 1, the predictive performance is good when only using observations of precipitation (dashed line). For this dataset, the predictive performance is slightly increased when using both point and area observations.

For Simulation 8, the predictive performance is in most cases largest for an observation sample of only precipitation. An observation sample that combines observations of runoff and precip-

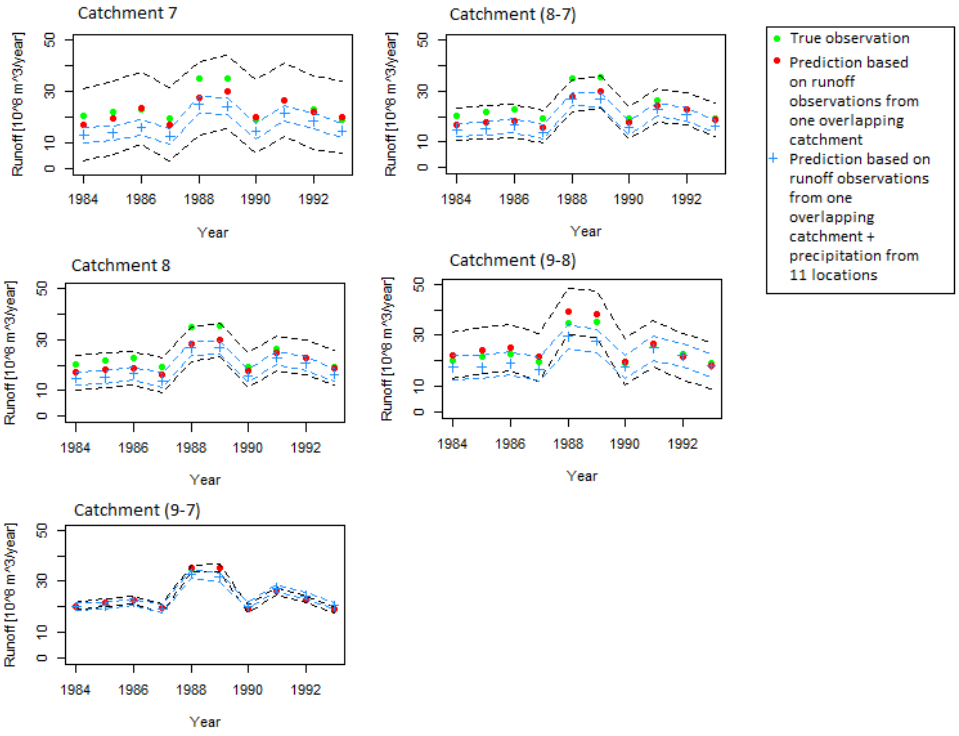


Figure 7.23: Posterior mean (red) with 95 % prediction intervals (black) for the runoff within Catchment 9 when the observation sample only consists of runoff observations from an overlapping catchment, and posterior mean with 95 % prediction intervals (blue) when the observation sample consists of runoff observations from an overlapping catchment and precipitation from 11 locations. Real data are used.

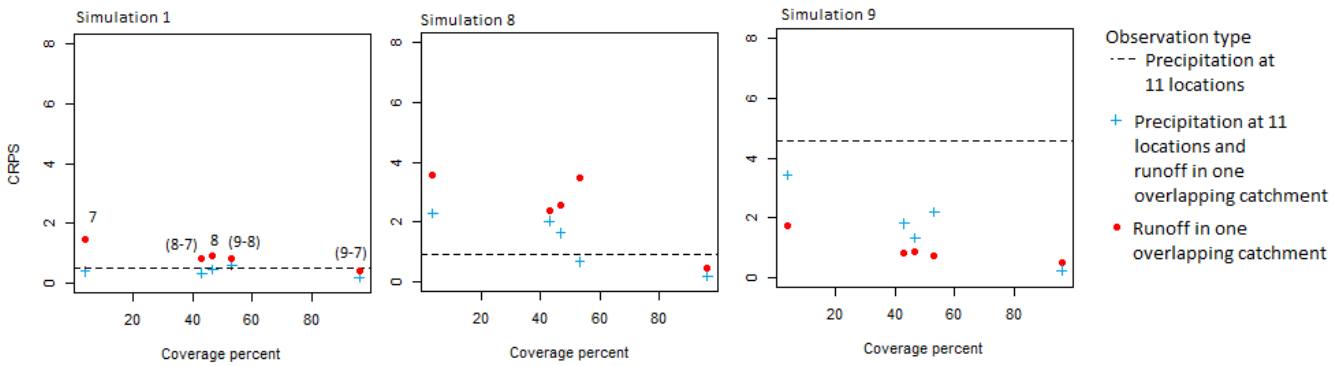


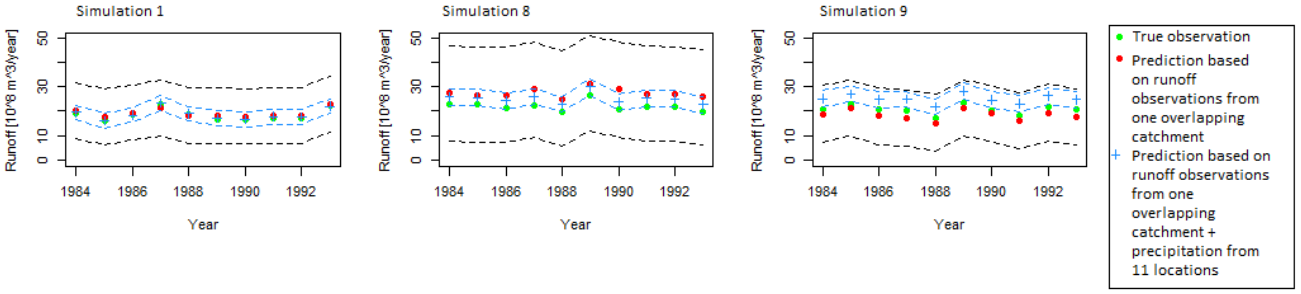
Figure 7.24: Predictive performance for annual runoff in Catchment 9 when the observation sample consists of runoff observations from an overlapping catchment and/or observations of precipitation from 11 locations. Simulated data are used.

itation performs better than an observation sample of only runoff.

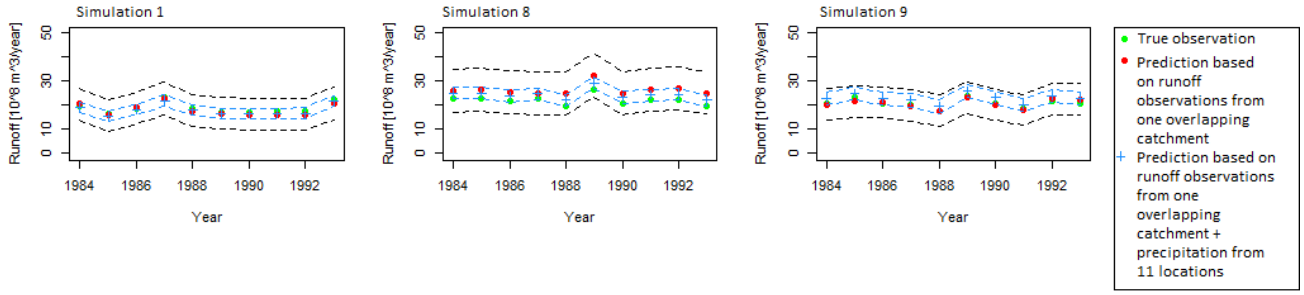
For Simulation 9, the predictive performance is largest when only using runoff observations. For this simulation, we see that the predictive performance is poor when only using observations of precipitation (dashed line). This is similar to what we obtained for the real dataset in Figure 7.22.

In Figure 7.25 we display some of the 95% confidence intervals for the runoff within Catchment 9 when using observations of runoff and/or precipitation for Simulation 1, 8 and 9. As for the real dataset, the prediction intervals become more narrow when observations of precipitation are included. For the simulated datasets, the true values are still covered by the prediction interval, except for Simulation 9 which was the dataset that provided the poorest predictions for the observation sample with both observations of runoff and precipitation. However, the biases for Simulation 9 are smaller than the biases observed for the real dataset when using a observation sample of both precipitation and runoff (Figure 7.23).

The conclusion from the experiments performed on the simulated datasets is the same as for the real dataset: From Figure 7.24 we see that including both observations of runoff and precipitation doesn't necessarily increase the predictive performance. In most cases an observation sample consisting of only runoff observations or only observations of precipitation performs better or equally good compared to an observation sample consisting of both observations of runoff and precipitation. However, the results from both the simulated datasets and the real dataset show that we never know which of the observation types that will produce the most accurate predictions for a specific climatology. Sometimes point observations perform better than runoff observations (Simulation 1) and sometimes it is the other way around (Simulation 9). This indicates that we in general will profit from combining observations of runoff and precipitation when making runoff predictions.



(a) Runoff observations from Catchment 7



(b) Runoff observations from Catchment 8

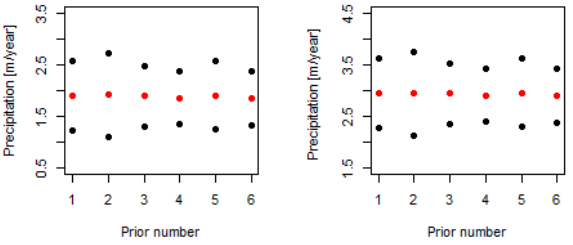
Figure 7.25: Posterior mean (red) with 95 % prediction intervals (black) for the runoff within Catchment 9 when the observation sample only consists of runoff observations from either Catchment 7 or 8, and posterior mean with 95 % prediction intervals (blue) when the observation sample consists of runoff observations from Catchment 7 or 8 and precipitation from 11 locations. Simulated data are used with 10 replicates.

Prior sensitivity

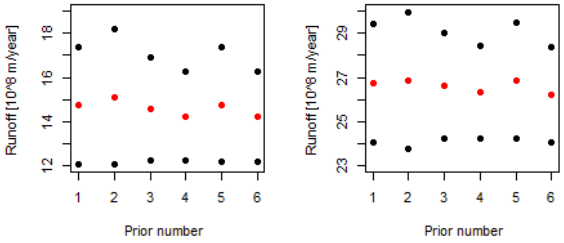
In Appendix A.4, the 0.025-, 0.5- and 0.975-quantiles for the posterior distribution of β_j are included when 6 different priors for the SPDE parameters are used. The parameter β_j is not sensitive to changes in the priors for $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$, and the posterior distributions are approximately equal for all priors.

Table 7.2 shows the posterior 0.025-, 0.5- and 0.975-quantiles for the remaining parameters when using prior 1 to 6 for the SPDE parameters. We see that $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$ are stable with respect to the choice of prior. The exception is the results from using prior 2 where we see posterior quantiles that are quite different from the other quantiles. The posterior means of $\theta_{\kappa,u}$ and $\theta_{\kappa,w}$ move towards the prior mean -2 which correspond to a lower range. Prior 2 also make an impact on the posterior quantiles of the precision τ_p, while τ_r is stable. Prior 2 is the prior where the strongest restrictions were set on the prior quantiles for the range for both $c(\boldsymbol{s})$ and $x_j(\boldsymbol{s})$, and the results show that this lead to an adjustment of the posterior distributions for several of the parameters.

Figure 7.26 shows posterior prediction intervals for runoff in Catchment 8 and posterior prediction intervals for precipitation at Reimegrend in two selected years: 1984 and 1988. Although prior 2 makes a large impact on the parameter values, this is not prominent in the resulting posterior mean of precipitation and runoff. However, the widths of the posterior prediction intervals vary more from prior to prior. For example in 1984 the width of the posterior prediction interval for runoff varies from $4.1 \cdot 10^8$ m^3/year for prior 4 to $6.1 \cdot 10^8$ m^3/year for prior 2. This is an increase of 50%. For precipitation the corresponding increase is from 1.0 m/year for prior 4 to 1.6 m/year for prior 2, i.e an increase of 60%.



(a) Precipitation 1984 (b) Precipitation 1988



(c) Runoff 1988 (d) Runoff 1988

Figure 7.26: Predictions of precipitation at Reimegrend and runoff in Catchment 9 in 1984 and 1988 (red) with 95% posterior prediction intervals (black). The observation sample consists of precipitation from 11 locations and runoff from Catchment 8.

Table 7.2: Quantiles for the posterior distributions of the precisions τ_p and τ_r and the SPDE parameters for 6 different priors. Recall that $\theta_{\tau,u}$ and $\theta_{\kappa,u}$ are related to the climatology $c(\boldsymbol{s})$ while $\theta_{\tau,w}$ and $\theta_{\kappa,w}$ are related to the annual variability $x_j(\boldsymbol{s})$.

	Quantiles	Prior 1	Prior 2	Prior 3	Prior 4	Prior 5	Prior 6
$\theta_{\tau,u}$	0.025	0.73	0.27	1.1	1.5	0.84	1.5
	0.5	1.5	1.1	1.7	1.9	1.5	1.9
	0.975	2.1	1.9	2.2	2.3	2.1	2.3
$\theta_{\kappa,u}$	0.025	-3.2	-2.9	-3.4	-3.7	-3.2	-3.7
	0.5	-2.4	-2.0	-2.7	-3.2	-2.4	-3.2
	0.975	-1.5	-1.0	-1.8	-2.8	-1.6	-2.8
$\theta_{\tau,w}$	0.025	2.8	0.56	3.0	3.3	3.3	2.3
	0.5	3.4	1.5	3.5	3.7	3.7	3.0
	0.975	3.9	2.4	4.1	4.1	4.1	3.6
$\theta_{\kappa,w}$	0.025	-3.6	-1.7	-3.7	-3.7	-3.7	-3.3
	0.5	-3.0	-0.92	-3.1	-3.3	-3.3	-2.6
	0.975	-2.3	-0.18	-2.5	-2.9	-2.9	-1.6
τ_p	0.025	1.5	1.1	1.5	1.5	1.4	1.5
	0.5	2.1	1.6	2.1	2.1	2.1	2.1
	0.975	3.0	2.4	3.0	2.9	2.9	3.0
τ_r	0.025	0.50	0.50	0.50	0.47	0.47	0.51
	0.50	0.95	0.96	0.95	0.92	0.92	0.97
	0.98	1.7	1.7	1.7	1.7	1.7	1.7

8. Discussion

In this work we have shown how we can construct a Bayesian statistical model for precipitation and runoff in R-INLA by using the SPDE approach to spatial modelling. The implementation supports both point observations of precipitation and area observations of runoff and evaporation, and the model was tested on catchments located around Voss both through real and simulated data. The focus of the experiments were on answering five research tasks to gain knowledge of the properties of the model.

The first research task was about exploring the importance of the climatology $c(\boldsymbol{s})$ which was modelled by a GRF common for all years. The experiments done by using real data from the Voss area showed that the climatology $c(\boldsymbol{s})$ was the dominating GRF of the model. The year dependent GRF $x_j(\boldsymbol{s})$ was neglected from the model without seeing significantly changes in the CRPS and RMSE. This indicates that the spatial variations in annual precipitation around Voss are quite stable from one year to another.

The consequence of this model property is that an accurate prediction of runoff or precipitation implies an accurate representation of the underlying climatology and the intercept β_j. In particular the climatology was shown to be important. The calibration tests based on simulated values suggested that as long as the underlying climatology is constant, the spatial predictions are not calibrated. Thus, an observation design that produces accurate predictions for one year will typically produce accurate predictions for other years.

Likewise will an observation design that produces poor spatial predictions for one year, produce poor predictions for other years. This is because we do the same systematic error in the predictions for a specific observation design as long as the spatial variations of precipitation in the area of interest are stable from year to year. Thus, if historical data show that a particular observation design produces poor predictions, we should try to change the observation design as we can't do anything about the climatology of an area in real life.

This conclusion applies for the area around Catchment 9 and probably for other areas where the climatic differences are similar from one year to another. In larger areas with more extreme weather, the annual spatial variability $x_j(\boldsymbol{s})$ might be of larger importance. The largest effect on the annual spatial variability $x_j(\boldsymbol{s})$ was seen in the posterior mean from 1988, and this was the most extreme year in our dataset.

In research task 2 we asked: If we find an observation design suitable for predicting runoff, will the same observation design be suitable for point predictions in the interior of the catchment? The calibration test based on simulated data showed that the answer on this question is no. This result illustrates that a runoff prediction and a prediction of precipitation are quite different from each other. When predicting the runoff within a catchment we first need to predict the annual precipitation in each point within the catchment of interest. To receive an accurate runoff prediction these predictions must be representative for the true climatic conditions within the catchment. When predicting the annual precipitation at a specific location, we only need to hit right at one point in space and not for many points.

In research task 3 we explored how runoff predictions are affected by using an observation sample consisting of runoff observations from an overlapping catchment. The real and the simulated datasets gave similar results and showed that the bias of the predictions decreases when the coverage percent of the observed catchment increases. The results from the simulations also showed that unexpected results can occur for some climatologies if we are unlucky or lucky with the observations.

In research task 4 we explored how runoff predictions are affected by using an observation sample consisting of observations of precipitation. For the real dataset, we saw that an observation sample of precipitation from 12 locations led to predictions that underestimated the true observations of runoff in Catchment 9. The prediction bias was large: The difference between the true and the predicted values were approximately twice as large as the width of the corresponding 95% prediction interval. This did not correspond to the simulation study. For some of the simulated climatologies we saw that point observations were suitable for runoff predictions, but for other climatologies biased predictions were obtained. However, the biases were not nearly as large as the biases obtained for the real dataset.

This result suggests that the correspondence between the observations of precipitation and the observations of runoff and evaporation was low for the real dataset. The low correspondence might be explained by the challenges related to measurements of precipitation in wind-exposed areas as reported in Wolff et al. (2015). In wind-exposed areas measurements of precipitation are often biased due to under-catch, and in RT4 we saw that almost all of the observations of annual precipitation were lower than what we would expect from the runoff observations from Catchment 9.

Another known problem in hydrology is that catchments of interest often are located at a higher elevation than the closest precipitation gauges. The reason for this is that it is challenging to build and maintain precipitation gauges in mountainous areas. As the precipitation gauges often are located at a lower elevation than the catchment of interest, it is common that the observations of precipitation are not representative for the true climatic conditions within the catchment of interest (Ingebrigtsen et al., 2013). This may be another explanation for the low correspondence between the observations of runoff and precipitation for the real dataset.

It is also possible that the predictive performance for runoff would have been better if more point observations within Catchment 9 were available. For our experiment, only 2 out of 12 observations of precipitation were located inside Catchment 9. In Engeland et al. (2016) it was shown that large differences in the posterior distribution of precipitation are obtained when one of the closest measuring stations are taken out of the dataset. This shows that missing data have a large impact on the resulting predictions. However, the simulated dataset gave more reasonable results than the real dataset, so the problem seems to lie in the data or in the precipitation model. The precipitation model can for example be improved by including a linear effect of elevation as precipitation is known to be affected by topography, or by replacing the stationary GRFs by non-stationary spatial fields as in Ingebrigtsen et al. (2013) and Ingebrigtsen et al. (2015).

In research task 5 we combined point observations of precipitation and area observations of runoff and used these for runoff predictions. For the real dataset we saw that the predictive

performance decreased when observations of precipitation and runoff were combined compared to an observation sample of only runoff observations. For most of the 95% posterior prediction intervals, the true observations fell outside the prediction interval when the observation sample consisted of both precipitation and runoff observations. The biases were quite large, and the poor predictive performance was caused by the point observations.

The work presented in this report was based on the idea that we should include as much information as possible in the statistical model to obtain accurate predictions. For most of the simulation experiments performed in research task 5, the predictive performance was better or equally good when using only one observation type (runoff or precipitation) compared to using both observation types (runoff and precipitation). Thus, including more information in the model, i.e more observations, did in most cases not lead to improved predictions, neither for the real dataset nor for the simulated ones.

However, the results from the simulation study did not favour one of the observation types (runoff or precipitation). For some simulations observation samples of only runoff performed better than observations samples of only precipitation. For other simulations, observations of precipitation produced the most accurate predictions. This indicates that we in average will profit from combining observations of runoff and precipitation, but we cannot be sure as this is based on results from only three simulated datasets and three climatologies.

A. Additional results

A.1 Parameters used for simulation

Some of the results in the main text are based on simulated values of precipitation and runoff. Table A.1 shows the parameter values used in the simulations. The values used for the SPDE parameters are the posterior means obtained when making inference based on an observation sample consisting of observations of precipitation from Aurland, Granvin, Reimegrend, Voss, Brandset, Nedre Ålvik, Gullbrå, Eksingedal, Fjellanger, Brekkhus, Bulken, Øvstedal and area observations from Catchments 1, 6, 8, 7 and 9 from 1984 to 1994 (see Figures 2.1 and 2.2). Thus, the SPDE parameter values chosen for simulations are realistic for our dataset. The year dependent intercepts β_j are sampled from a normal distribution with mean 2 and standard deviation 0.2.

The standard deviation of the observations of precipitation is set to 10% of the simulated value, and the standard deviation of the runoff observations is set to 3% of the simulated runoff value. The amount of evaporation is set to 10% of the simulated runoff value.

The prior distributions used in the simulation study, are the same as the ones presented in Section 4.5.

Table A.1: Parameter values used for simulation.

Parameter	Value
$\theta_{\tau,w}$	3.28
$\theta_{\kappa,w}$	-3.10
$\theta_{\tau,u}$	0.79
$\theta_{\kappa,u}$	-1.87
τ_p	1
τ_r	1

A.2 Predictions of precipitation

In the main text, the main focus was on runoff predictions. For completeness of the results, we give a brief overview of how predictions of precipitation are affected by observing runoff and/or precipitation. We observe the runoff within Catchment 7, (8-7), 8, (9-8) or (9-7) and/or precipitation at 11 locations for 10 years. The observations are used to predict the precipitation at Reimegrend. Thus, this is an equivalent experiment to the one performed for research task 5. The only difference is that we now predict precipitation and not runoff.

Figures A.1 and A.2 show the predictive performance obtained for predictions of precipitation at Reimegrend by using observations of precipitation and/or runoff. Reimegrend is located in the middle of Catchment (8-7), (9-8) and 8, and outside Catchments 7 and (9-8), but we don't see a clear correspondence between the predictive performance and the size and/or location

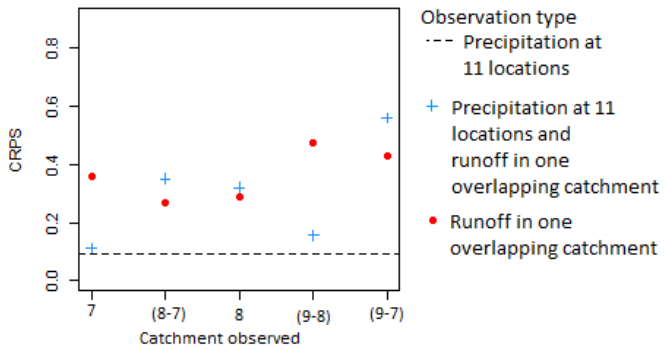


Figure A.1: Predictive performance for precipitation (CRPS) when the observation sample consists of runoff observations from an overlapping catchment and/or observations of precipitation from 11 locations. Real data are used.

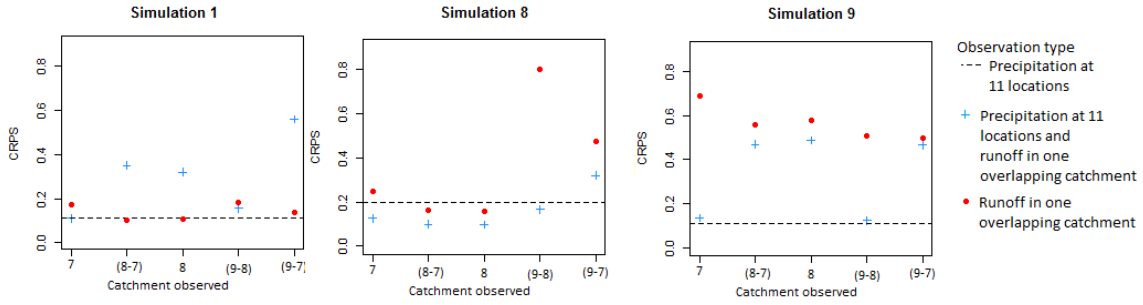


Figure A.2: Predictive performance for precipitation (CRPS) when the observation sample consists of runoff observations from an overlapping catchment and/or observations of precipitation from 11 locations. Simulated datasets are used.

of the catchment in which runoff is observed.

Further we notice that an observation sample consisting of observations of precipitation from 11 catchments (dashed line) provides a lower or approximately equally low CRPS compared to an observation sample consisting of only runoff observations (red points) both in Figure A.1 and A.2. This indicates that point observations from several locations are more suitable for making point predictions than area observations from one catchment.

However, we note that the CRPS values in general are close to zero for Simulation 1, showing that runoff observations from only one catchment can provide accurate point predictions of precipitation. This happens when the annual precipitation at Reimegrend is representative for the mean runoff within the observed catchment.

For Simulation 8 combining observations of runoff and precipitation gives the best predictions of precipitation at Reimegrend. For the other simulations, and for the real dataset, the predictive performance is largest when using an observations sample of only precipitation. Thus, combining the observation types did not lead to improved predictions for most of the experiments.

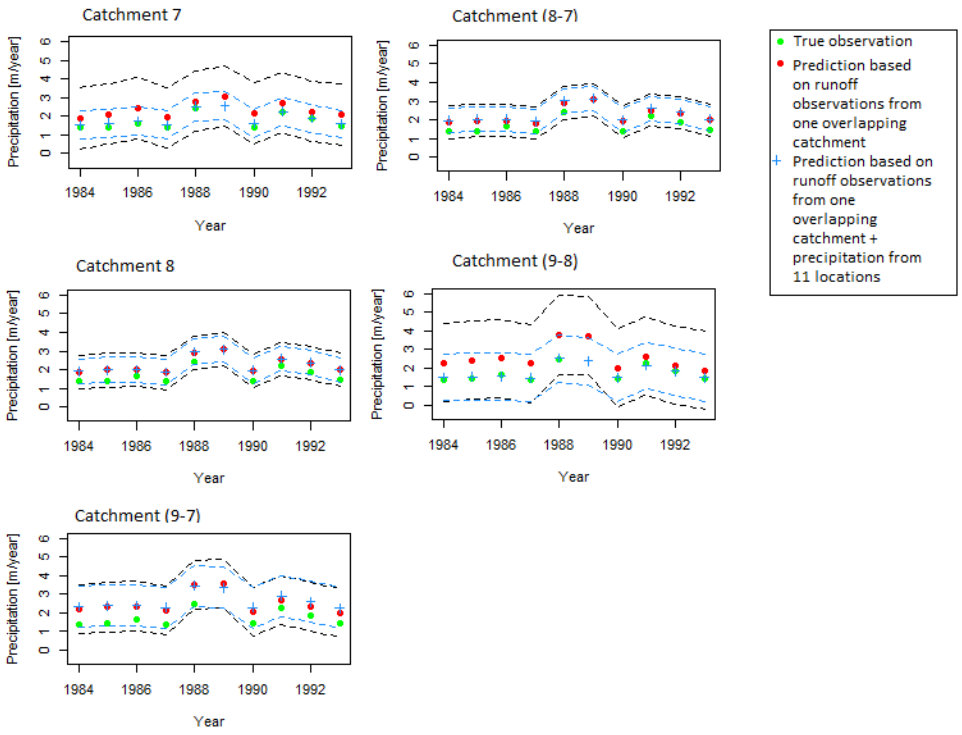
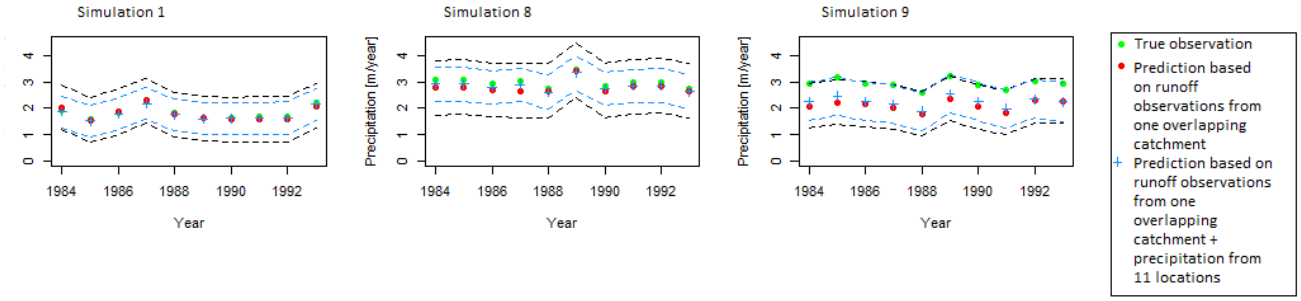
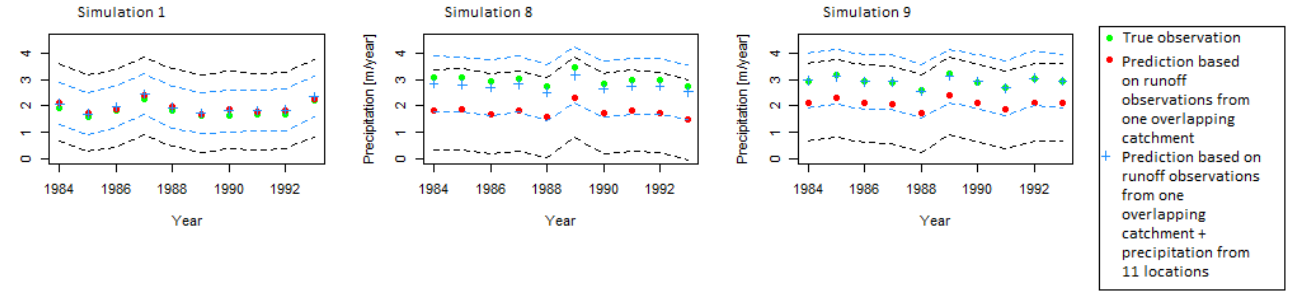


Figure A.3: Posterior mean (red) with 95 % prediction intervals (black) for the precipitation at Reimegrend when the observation sample only consists of runoff observations from an overlapping catchment, and posterior mean with 95 % prediction intervals (blue) when the observation sample consists of runoff observations from an overlapping catchment and precipitation at 11 locations. Real data are used.



(a) Runoff observations from Catchment 8



(b) Runoff observations from Catchment (9-8)

Figure A.4: Posterior mean (red) with 95 % prediction intervals (black) for the precipitation at Reimegrend when the observation sample only consists of runoff observations from an either Catchment 8 or (9-8), and posterior mean with 95 % prediction intervals (blue) when the observation sample consists of runoff observations from Catchment 8 or (9-8) and precipitation at 11 locations. Simulated datasets are used.

In Figure A.3 the posterior prediction intervals for the precipitation at Reimegrend for the real dataset are displayed. We see that the intervals are more narrow when the observation sample consists of runoff observations from either Catchment (8-7), 8 or (9-7). This is intuitive because the location of interest, Reimegrend, is located in the middle of these catchments and outside the remaining two catchments. In the middle of the observed catchment, the uncertainty is at its lowest as we saw in Figure 7.14 in the main text. Further, we observe that the prediction intervals become more narrow when point observations of precipitation are added to the observation sample. The true observations are still inside the 95% prediction interval.

In Figure A.4 we see some of the 95% posterior prediction intervals for the simulated datasets. The predicted values get closer to the true observation when point observations of precipitation are included in the observation sample. Especially when runoff is observed within Catchment (9-8). The location of interest, Reimegrend, is located outside this catchment and inside Catchment 8 so this is intuitive. Overall, the predictions are very accurate as long as point observations are included in the observation sample. We saw the same for the posterior prediction intervals for the real dataset.

A.3 Posterior distributions of the hyperparameters

The main focus of this work has been on spatial predictions. Here include plots of the posterior distributions of the hyperparameters based on some of the different observation samples from the main text. Real data from 10 years are used to compute the posterior distributions.

Figure A.5 shows posterior distributions of the SPDE parameters. We note that the posterior distributions obtained when the observation sample includes observations of precipitation in general are more narrow than the posterior distributions obtained when the observation sample only consists of runoff observations. Further, the posterior distributions originating from an observation sample of both runoff and precipitation (gray) have more in common with the posterior distributions originating from an observation sample of only point observations (blue) than an observation sample of only runoff observations (yellow, green, red). This indicates that the observations of precipitation dominate over the observations of runoff.

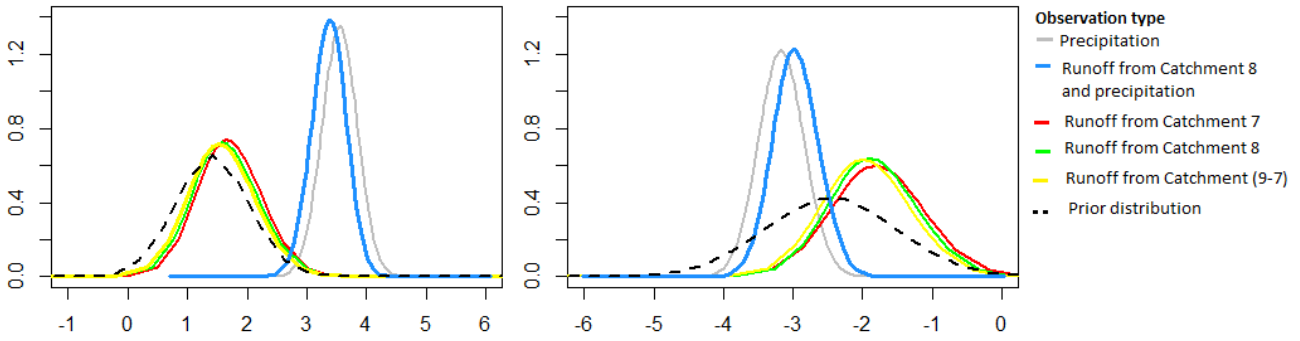
The posterior distributions in Figure A.5 shows that the parameter estimates obtained are very different from one observation sample to another. Especially for $\theta_{\tau,w}$ and $\theta_{\kappa,w}$ we see a large difference in the posterior mean when point observations are included in the observation sample or not.

Figure A.6 shows posterior distributions for the intercept for different observation samples. The standard deviation of the posterior distributions are approximately equal for all observation types, and the posterior means are also approximately equal.

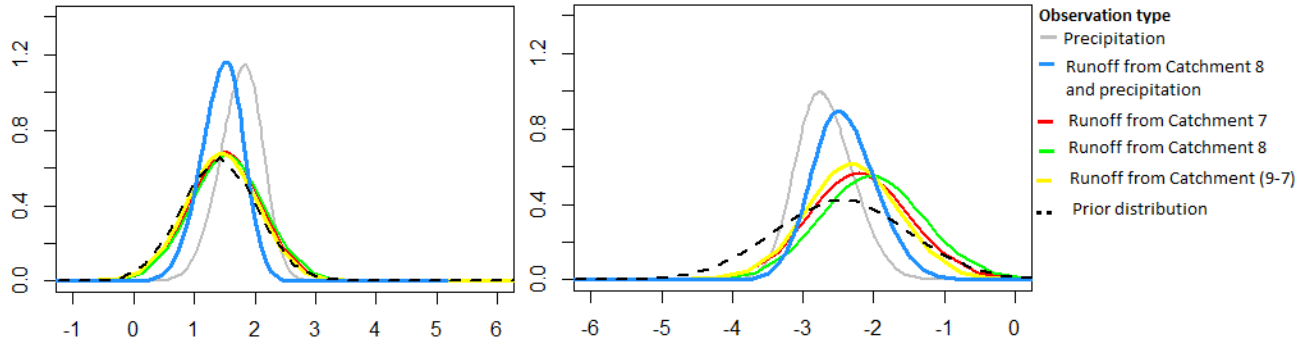
Figure A.7 shows posterior distributions for the precisions τ_p and τ_r. When no observations of precipitation are used, the posterior distribution of τ_p is equal to its prior distribution. Likweise, when no runoff observations are used, the posterior distribution of τ_r is equal to its prior distribution. The posterior distribution of τ_p indicates that the observations of precipitation are more certain than we assumed when the prior standard deviation of precipitation was set to 10% of the observed value.

The posterior distribution of τ_r is approximately equal to its prior distribution regardless of which catchment in which we observe runoff. This can be explained by the prior assumptions for the runoff uncertainty. The relative standard deviation for the runoff observations provided by NVE (Table 2.1) were in general low, so the prior assumptions were strict.

Analysing how the parameter estimates affect the spatial predictions has been outside the scope of this work.



(a) Posterior distributions of the parameters $\theta_{\tau,w}$ (left) and $\theta_{\kappa,w}$ (right) related to the spatial annual variability $x_j(\boldsymbol{s})$.



(b) Posterior distributions of the parameters $\theta_{\tau,u}$ (left) and $\theta_{\kappa,u}$ (right) related to the climatology $c(\boldsymbol{s})$.

Figure A.5: Posterior distributions of the SPDE parameters. The plots are based on real data from 10 years. The observations of precipitation are observed at the same 11 locations as in RT5. In addition we use observations of runoff from Catchment 7, 8 or (9-7).

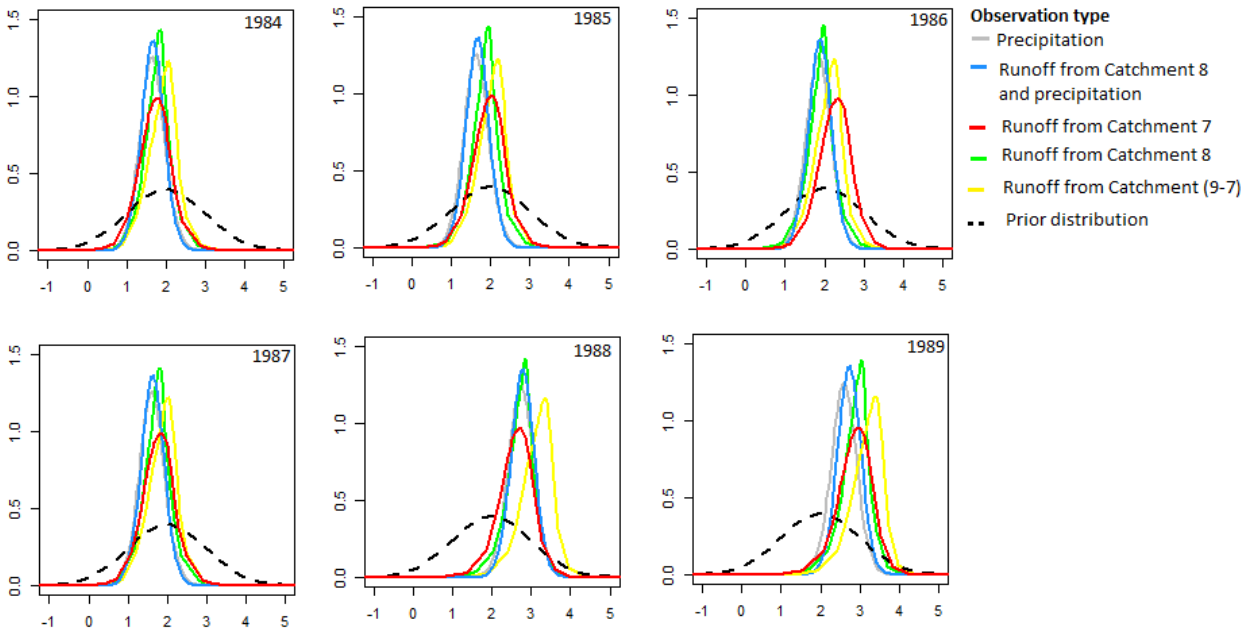


Figure A.6: Posterior distributions of β_j from 1984 to 1989. The distributions from 1990-1993 are omitted. The plots are based on real data from 10 years. The observations of precipitation are observed at the same 11 locations as in RT5. In addition we use observations of runoff from Catchment 7, 8 or (9-7).

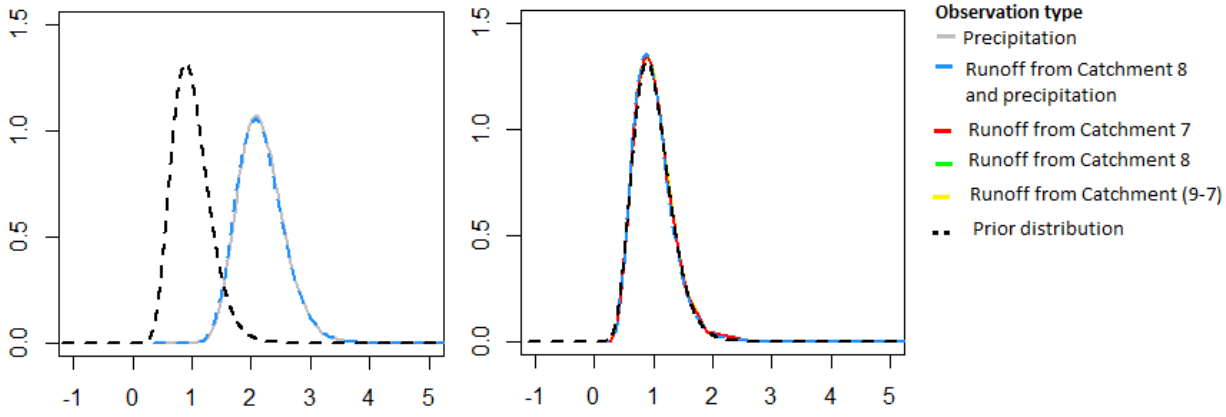


Figure A.7: Posterior distributions of precisions τ_p (left) and τ_r (right). The plots are based on real data from 10 years. The observations of precipitation are observed at the same 11 locations as in RT5. In addition we use observations of runoff from Catchment 7, 8 or (9-7).

A.4 Prior sensitivity for β_j

Table A.2 shows the posterior quantiles for the intercept β_j when 6 different priors for $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$ are used. We see that the parameter β_j is not sensitive to the choice of prior for the SPDE parameters.

Table A.2: Quantiles for the posterior distribution of β_j for 6 different priors.

	Quantiles	Prior1	Prior 2	Prior 3	Prior 4	Prior 5	Prior 6
β_1	0.025	1.1	1.2	1.0	0.92	1.0	0.98
	0.5	1.7	1.7	1.7	1.6	1.7	1.6
	0.975	2.3	2.1	2.3	2.4	2.3	2.3
β_2	0.025	1.1	1.3	1.0	0.92	1.0	0.99
	0.5	1.7	1.7	1.7	1.6	1.7	1.6
	0.975	2.3	2.2	2.3	2.4	2.2	2.3
β_3	0.025	1.3	1.4	1.2	1.1	1.2	1.2
	0.5	1.9	1.9	1.9	1.8	1.9	1.8
	0.975	2.5	2.3	2.5	2.6	2.5	2.5
β_4	0.025	1.0	1.2	0.97	0.88	0.99	0.94
	0.5	1.5	1.7	1.6	1.6	1.6	1.6
	0.975	2.2	2.1	2.3	2.3	2.3	2.2
β_5	0.025	2.2	2.3	2.1	2.0	2.1	2.1
	0.5	2.8	2.8	2.7	2.7	2.8	2.7
	0.975	3.4	3.2	3.4	3.4	3.4	3.4
β_6	0.025	2.1	2.4	2.0	1.9	2.1	2.1
	0.5	2.7	2.8	2.7	2.7	2.7	2.7
	0.975	3.3	3.3	3.4	3.4	3.3	3.3
β_7	0.025	1.2	1.3	1.1	1.0	1.2	1.1
	0.5	1.8	1.8.	1.8	1.8	1.8	1.8
	0.975	2.4	2.2	2.4	2.5	2.4	2.4
β_8	0.025	1.8	2.0	1.8	1.7	1.8	1.8
	0.5	2.4	2.4	2.4	2.4	2.4	2.4
	0.975	3.0	2.9	3.0	3.1	3.1	3.0
β_9	0.025	1.5	1.7	1.5	1.4	1.5	1.5
	0.5	2.2	2.2	2.2	2.1	2.2	2.1
	0.975	2.8	2.6	2.8	2.8	2.8	2.8
β_{10}	0.025	1.2	1.4	1.1	1.0	1.2	1.1
	0.5	1.8	1.8	1.8	1.8	1.8	1.7
	0.975	2.4	2.2	2.4	2.5	2.4	2.4

B. Implementation in R-INLA

Important aspects with the implementation in R-INLA is presented in this section. We will show a toy example where we are interested in making inference based on two observations of precipitation from two locations that are shown in Figure B.1, Brekkhus and Bulken. We also observe the runoff within Catchment 6.

Based on these observations, we estimate the runoff within Catchment 4 and the precipitation at Gullbrå.

We use 2 replicates, i.e $r = 2$, and use observations from 1995 and 1996.

For details regarding the functions used, we refer to (Krainski et al., 2016) and (Blangiardo and Cameletti, 2015).

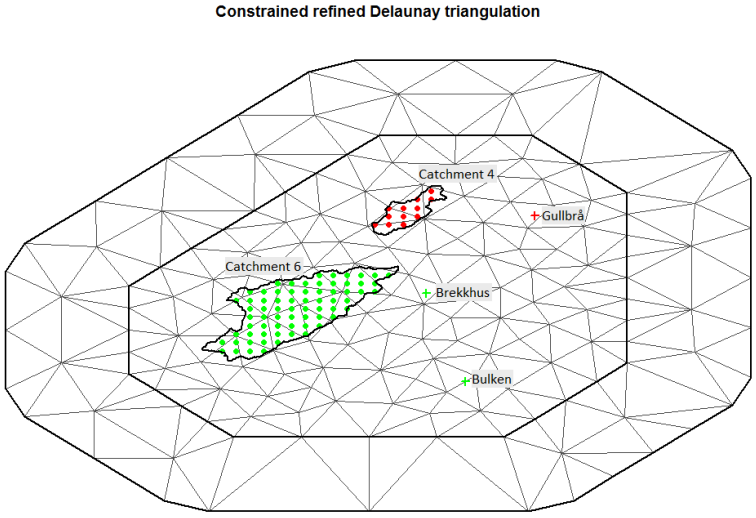


Figure B.1: Observations of precipitation from Brekkhus and Bulken (green) and runoff observations from Catchment 6 (green) are used to predict the precipitation at Gullbrå and the runoff within catchment 4 (red) in 1995 and 1996.

B.1 Data input in INLA

We gather the coordinates of the observed locations Bulken (20.8, 6755) and Brekkhus (17.96, 6766) in the matrix `point.obs.coords`, and the observed values of precipitation in the vector `y.obs`.

```
print(point.obs.coords)
        [,1]     [,2]
[1,] 20.83835 6755.341
[2,] 17.95662 6765.988
```

```
print(y.obs)
[1] 1.4742 1.8817 1.9509 2.5889
```

The first replicate from 1995 is located in the first two elements of the vector, i.e, `y.obs[1]` and `y.obs[2]` contain the precipitation in 1995 in Bulken and Brekkhus respectively, while `y.obs[3]` and `y.obs[4]` contain the precipitation in Bulken and Brekkhus in 1996.

We also need to store the coordinates of the 69 grid nodes of Catchment 6 from Figure B.1. These are stored in the matrix `catch6.coords`. In vector `z.obs` the sum of observed runoff and evaporation from 1995 and 1996 is stored:

```
print(catch6.coords)
          V1       V2
1   3.069280 6759.022
2   4.085079 6759.022
...  ...
...  ...
68 14.243063 6768.164
69 15.258861 6768.164

#The observed evaporation within Catchment 6 in 1995 and 1996.
print(evaporation6)
[1] 0.224 0.169

#The observed runoff within Catchment 6 in 1995 and 1996.
print(runoff6)
[1] 1.26 2.12

#z.obs is the sum of observed runoff and evaporation in 1995 and 1996.
z.obs=runoff6+evaporation6

print(z.obs)
[1] 1.484 2.289
```

Later, we will need the coordinates of the 12 grid nodes of Catchment 4, and the coordinates of Gullbrå (25.9,6775), the location at which we want to predict precipitation. We store the grid coordinates in the matrix `catch4.coords` and the coordinates of Gullbrå in the vector `point.pred.coords`:

```
print(catch4.coords)
         V1       V2
1  14.24306 6774.259
2  15.25886 6774.259
... ...
... ...
11 18.30626 6777.306
12 18.30626 6778.322
```

```
 print(point.pred.coords)
         [,1]     [,2]
[1,] 25.87003 6775.334
```

B.2 Using the SPDE approach to make the GMRFs $\boldsymbol{w_j}$ and $\boldsymbol{u}$

Recall that the responses of precipitation and (runoff+evaporation) are given by

$$E\{\boldsymbol{y_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_p, \beta_j)\} = \boldsymbol{\eta_j} = \boldsymbol{A_p}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u})$$
$$E\{\boldsymbol{z_j}|(\boldsymbol{w_j}, \boldsymbol{u}, \tau_r, \beta_j)\} = \boldsymbol{\zeta_j} + \boldsymbol{E_j} = \boldsymbol{A_r}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u})$$

with $\boldsymbol{w_j} \sim \mathcal{N}(0, \boldsymbol{Q}^{-1}(\theta_{\tau,w}, \theta_{\kappa,w}))$ and $\boldsymbol{u} \sim \mathcal{N}(0, \boldsymbol{Q}^{-1}(\theta_{\tau,u}, \theta_{\kappa,u}))$.

We use the coordinates of the grid nodes of both Catchment 4 and 6, and the coordinates of the points of interest to make the triangulation mesh shown in Figure B.1. Then, we assign prior means and variances for the SPDE-parameters $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$. The priors and the mesh are used for constructing two spde-objects `spde` and `spde.c`, one representing the annual spatial variability $\boldsymbol{w_j}$, and one representing the climatology $\boldsymbol{u}$:

```
#Specifying the properties of the mesh:
max.edge=c(0.2*25,0.5*35); offset=c(0.15,0.2)*45; cutoff=0.1*25

#We store all the coordinates that will be needed for
#predictions and inference in all_coords:
all_coords=rbind(catch6.coords,catch4.coords,point.obs.coords,point.pred.coords)

#Making the mesh:
mesh=inla.mesh.2d(loc.domain=all_coords,cutoff=cutoff,max.edge=max.edge,offset=offset)

#Specifying the properties of the year dependent GMRF w_j:
priors= list(sig_k=0.939465, mu_k= -2.456787, sig_t=0.6094179, mu_t=1.414418)
spde=inla.spde2.matern(mesh,B.tau=matrix(c(0,1,0),nrow=1,ncol=3),
      B.kappa=matrix(c(0,0,1),nrow=1,ncol=3),
     theta.prior.mean=c(priors$mu_t,priors$mu_k),
     theta.prior.prec=c(1/priors$sig_t^2,1/priors$sig_k^2))

#Specifying the properties of the climatology GMRF u:
priors.c=priors #We use the same priors as for the year dependent SPDE.
spde.c=inla.spde2.matern(mesh,B.tau=matrix(c(0,1,0),nrow=1,ncol=3),
      B.kappa=matrix(c(0,0,1),nrow=1,ncol=3),
     theta.prior.mean=c(priors.c$mu_t,priors.c$mu_k),
     theta.prior.prec=c(1/priors.c$sig_t^2,1/priors.c$sig_k^2))
```

The resulting mesh is displayed in Figure B.1 and has 177 mesh nodes, i.e $m = 177$. Each of these nodes are now given a index by using the function `inla.spde.make.index()`. We make one set of indices for $\boldsymbol{w_j}$ and one set for indices for $\boldsymbol{u}$.

In the function `inla.spde.make.index()` we also specify the number of replicates of the GMRFs $\boldsymbol{w_j}$ and $\boldsymbol{u}$. The object `spde` has 2 replicates because this object represents the annual spatial variability $\boldsymbol{w_j}$. The object `spde.c` has no replicates (or 1 replicate) as this represents the climatology $\boldsymbol{u}$ which is common for all years. This is how this is specified in INLA:

```
n.replic=2 #We have 2 replicates.

#Index for the annual SPDE-object. Two replicates:
s.index=inla.spde.make.index(name="field",n.spde=spde$n.spde,n.repl=n.replic)

#Index for the climatic SPDE-object. No replicates:
s.index.c=list()
s.index.c$field.c=rep(c(1:177),times=2)
s.index.c$field.c.repl=rep(1,length(s.index$field.repl))
s.index.c$field.c.group=rep(1,length(s.index$field.repl))
```

We now demonstrate the difference between `s.index` and `s.index.c`. The first vector of `s.index` and `s.index.c` is common. This vector contains the numbers from 1 to 177 repeated 2 times because $m = 177$ and $r = 2$.

```
print(s.index$field)
  [1]   1   2   ... 176 177 1 2 ... 176 177

  print(s.index.c$field.c)
  [1]   1   2   ... 176 177 1 2 ... 176 177
```

The second vector `group` is also common. This contains $m \times r$ ones, and is not relevant in this model.

In the last vector `repl` there is a difference between `s.index` and `s.index.c`. For the indices of the annual variability $\boldsymbol{w_j}$ we have that

```
s.index$field.repl=rep(c(1,2),each=177)
print(s.index$field.repl)
[1]   1 1 1 1 1 1 1 1 1 1 1 ... 1 1 2 2 2 ... 2 2
```

This vector specifies that the first $m = 177$ indices of `s.index$field` belong to the first replicate ($j = 1$) and that the next $m = 177$ indices of `s.index$field` belong to the second replicate ($j = 2$) of $\boldsymbol{w_j}$.

For the indices of the GMRF $\boldsymbol{u}$ representing the climatology, we have

```
print(s.index.c$field.c.repl)
[1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1....1 1
```

This is a vector of $m \times r$ ones. This way we mark that the GMRF $\boldsymbol{u}$ is common for all years and that we don't have replicates. All of the indices in `s.index.c$field.c` represent the same realization of the field.

B.3 Generating the projection matrices for inference

We now generate the projection matrix $\boldsymbol{A_p}$ from Equation (4.8) for precipitation which is straight forward with the function `inla.spde.make.A()`:

```
n.loc.obs=2 #We observe the precipitation at 2 locations, Bulken and Brekkhus.

#Projection matrix for precipitation
A.p=inla.spde.make.A(mesh,loc=as.matrix(point.obs.coords),
     index=rep(1:n.loc.obs,times=n.replic),
     repl=rep(1:n.replic,each=(n.loc.obs)))
```

This projection matrix projects the GMRFs $\boldsymbol{w_j}$ and $\boldsymbol{c}$ from the 177 mesh nodes to the locations at which we observe precipitation, Bulken and Brekkhus.

The projection matrix for runoff $\boldsymbol{A_r}$ from Equation (4.9) is more tricky. We first make the vector `group.obs`:

```
group.obs=rep(c(1,2),each=69)
```

Recall that the observed catchment, Catchment 6, consists of 69 grid nodes. Elements 1 to 69 of `group.obs` are 1's. This indicate that we want to sum the precipitation in all of the 69 grid nodes of replicate 1. Elements 70 through 138 of `group.obs` are 2's indicating that we want to sum the precipitation in all of the 69 grid nodes of replicate 2. We are now able to construct the projection matrix $\boldsymbol{A.r}$:

```
n.obsblock=69 #Number of grid nodes in Catchment 6.
Area6=72350713 #Area of Catchment 6 in meters.

#Projection matrix interpolating from mesh nodes to grid nodes in Catchment 6:
A.sum=inla.spde.make.A(mesh,loc=as.matrix(catch6.coords),
      index=rep(1:(n.obsblock),times=n.replic),repl=rep(1:n.replic,each=(n.obsblock)));

#Projection matrix for summing the precipitation in the grid nodes of Catchment 6:
A.r=inla.spde.make.block.A(A.sum,group.obs)

#We multiply A.r with an area element and scale the matrix
#to get the unit [m^3/year *10^{8}]
Delta6=(Area6/n.obsblock) * (10^-8)
A.r=A.r*Delta6
```

B.4 Drawing inference based on the observations

We further use the stack functionality in INLA. In `stack.p` we include all of the components of the response of precipitation. Precipitation is given by $\boldsymbol{\eta_j} = \boldsymbol{A_p}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u})$. Thus `stack.p`

includes the observations of precipitation and the projection matrix $\boldsymbol{A_p}$. In `stack.p` we also include `s.index`, `s.index.c` and `intercept` as these represent $\boldsymbol{w_j}$, $\boldsymbol{u}$ and β_j respectively. The sum of observed runoff and evaporation is given by $\boldsymbol{\zeta_j} + \boldsymbol{E_j} = \boldsymbol{A_r}(\mathbf{1}\beta_j + \boldsymbol{w}_j + \boldsymbol{u})$. Thus, `stack.r` includes the observations of runoff, the projection matrix $\boldsymbol{A_r}$, `s.index`, `s.index.c` and `intercept`:

```
#Year specific intercept is specified as a factor.
intercept=rep(1:n.replic,each=spde$n.spde)
#Here, spde$n is the number of mesh nodes, which is 177.

#Observation stacks:
#Precipitation:
stack.p=inla.stack(data=list(y=cbind(y.obs,NA)),
        effects=list(c(s.index,s.index.c,
        intercept=list(factor(intercept)))),A=list(A.p),tag="est.p")

#Runoff+evaporation
stack.r=inla.stack(data=list(y=cbind(NA,as.matrix(z.obs))),
        effects=list(c(s.index,s.index.c,
        intercept=list(factor(intercept)))),A=list(A.r),tag="est.r")

#We add the two observation stacks together
join.stack <- inla.stack(stack.p, stack.r)
```

Note that the data is added to the stacks as:

```
print(cbind(y.obs,NA))

      y.obs
[1,] 1.4742 NA
[2,] 1.8817 NA
[3,] 1.9509 NA
[4,] 2.5889 NA

cbind(NA,z.obs)
           z.obs
[1,] NA 1.486081
[2,] NA 2.289478
```

In INLA we need to specify the data with one column for each observation likelihood. In this case we have two observation types, precipitation and runoff, i.e we have two likelihoods. Thus, the data input need to have two columns. We add `NA` to the second column of `y.obs` specifying that `y.obs` comes from the first likelihood. Likewise, we add `NA` to the first column of `z.obs` specifying that `z.obs` comes from the second likelihood.

We specify the prior mean and precision for β_j, assign priors for the precisions τ_p and τ_r, and calculate the scales $\boldsymbol{f_j}$ and $\boldsymbol{v_j}$ allowing each observation to have its own uncertainty:

```
y.sd=(y.obs*0.1) #Standard deviation for y.obs. 10% of the observed value.
```

```
tau.p=1 #Prior mean for precision.
y.scale=(1/y.sd^2)*(tau.p^{-1}) #Scaled likelihood.

z.sd=c(0.0155,0.0112) #Standard deviation for z.obs. These are numbers given by NVE.
tau.r=1 #Prior mean for precision.
z.scale=(1/z.sd^2)*(tau.r^{-1}) #Scaled likelihood.

#Prior specification for tau.p:
precprior.p <- list(theta=list(param=c(shape=10,rate=10)))

#Prior specification for tau.r:
precprior.r=precprior.p

#Prior mean and precision for the intercepts:
prior.beta0=c(2,1/1^2)
```

The responses of precipitation and runoff from Equation (4.12), are specified in `formula`. Here, it is apparent why we chose to use (runoff+evaporation) as input: This way the responses of precipitation and runoff have the same form and can be specified through the same `formula` in INLA.

When the `formula` is specified, we can make inference by using the `inla()` function:

```
formula <- y~ -1 +intercept+f(field, model=spde,replicate=field.repl)
+f(field.c,model=spde.c)
#Note that we remove the default intercept by typing "-1" as we
#have specified the intercept manually in stack.p and stack.r.

#Making inference:
#The distributions of the two likelihoods are specified in the family-argument.
output=inla(formula,family=c("gaussian","gaussian"),
       data=inla.stack.data(join.stack),
       control.predictor=list(A=inla.stack.A(join.stack),compute=FALSE),
       control.family=list(list(hyper=precprior.p,prior="loggamma"),
       list(hyper=precprior.r,prior="loggamma")),scale=c(y.scale,z.scale),
       control.fixed=list(mean=prior.beta0[1],prec=prior.beta0[2]))
```

The object `output` consists the marginal posterior distributions of the SPDE parameters $\theta_{\tau,w}$, $\theta_{\kappa,w}$, $\theta_{\tau,u}$ and $\theta_{\kappa,u}$, the posterior marginals of the intercepts β_1 and β_2, and the posterior marginals of the precisions τ_p and τ_r.

B.5 Making predictions

We now show how we can predict the runoff within Catchment 4 and the precipitation at Gullbrå in 1995 and 1996. To make predictions, we need to make new projection matrices $\boldsymbol{A'_p}$ and $\boldsymbol{A'_r}$ that projects the GMRFs $\boldsymbol{w_j}$ and $\boldsymbol{c}$ from the m mesh nodes of the triangulation, to Gullbrå and to the grid nodes of Catchment 4. We use the same approach for making $\boldsymbol{A'_p}$ and

$\boldsymbol{A}_r'$ as we did for making $\boldsymbol{A}_p$ and $\boldsymbol{A}_r$. The only difference is that we now have new coordinates in the `loc` argument in `inla.spde.make.A()`.

```
n.loc.pred=1 #The number of locations
at which we want to estimate the annual precipitation.

#Projection matrix for precipitation
Apred.p=inla.spde.make.A(mesh,loc=as.matrix(point.pred.coords),index=rep(1:(n.loc.pred),
        times=n.replic),repl=rep(1:n.replic,each=(n.loc.pred)));

n.predblock=12 #The number of grid nodes in Catchment 4.
#Vector that indicates which grid nodes we should sum:
group.pred=rep(c(1,2),each=n.predblock)

#Projection matrix interpolating from mesh nodes to grid nodes in Catchment 4:
Apred.sum=inla.spde.make.A(mesh,loc=as.matrix(catch4.coords),
          index=rep(1:(n.predblock),times=n.replic),repl=rep(1:n.replic,each=(n.predblock)));

#Projection matrix for summing the precipitation in the grid nodes of Catchment 4:
Apred.r=inla.spde.make.block.A(Apred.sum,group.pred)

#We multiply Apred.r with an area element and scale the matrix
#to get the unit [m^3/year *10^{8}]
Area4=12828906 #Area of Catchment 4.
Delta4=(Area4/n.predblock)*10^{-8}
Apred.r=Apred.r*Delta4
```

We now make two new stacks, one for precipitation (`pred.response.y`) and one for runoff+evaporation (`pred.response.z`). The stacks don't contain data, only `NA`s. The `NA`s indicate that we want to predict the response. As before, the `data` input has two columns because we have two observation types.

After making the new stacks, we add the prediction stacks with the stack we used for making inference `join.stack`. Finally we use the `inla()` function to compute the predictions:

```
N=1 #The number of catchments in which we predict runoff.
#We are only interested in predicting the runoff within Catchment 4.

pred.response.y=inla.stack(data=list(y=matrix(NA,(n.loc.pred)*n.replic,2)),
               A=list(Apred.p),
              effects=list(c(s.index,s.index.c,intercept=list(factor(intercept)))),
              tag="pred.res.y")
pred.response.z=inla.stack(data=list(y=matrix(NA,N*n.replic,2)),
                  A=list(Apred.r),
                  effects=list(c(s.index,s.index.c,intercept=list(factor(intercept)))),
                  tag="pred.res.z")

#Add the prediction stacks and the inference stack.
```

```
join.stack.pred=inla.stack(join.stack,pred.response.y,pred.response.z)

#Make predictions:
output.pred=inla(formula,family=c("gaussian","gaussian"),
            data=inla.stack.data(join.stack.pred),
            control.predictor=list(A=inla.stack.A(join.stack.pred),compute=TRUE),
            scale=c(y.scale,z.scale,rep(1,dim(join.stack.pred
            $A)[1]-(n.loc.obs+N)*n.replic)),
            control.family=list(list(hyper=precprior.p),list(hyper=precprior.r)),
            control.fixed = list(expand.factor.strategy = "inla",mean=prior.beta0[1],
            prec=prior.beta0[2]))
```

We extract the results by using the `inla.stack.index()` function.

```
index.z.pred=inla.stack.index(join.stack.pred,tag="pred.res.z")$data
index.y.pred=inla.stack.index(join.stack.pred,tag="pred.res.y")$data

#Posterior mean and standard deviation for precipitation at Gullbrå:
pred=output.pred$summary.fitted.values[index.y.pred,"mean"]
sd=output.pred$summary.fitted.values[index.y.pred,"sd"]

#Posterior mean and standard deviation for (runoff+evaporation) within Catchment 4:
pred.z=output.pred$summary.fitted.values[index.z.pred,"mean"]
sd.z=output.pred$summary.fitted.values[index.z.pred,"sd"]

#The observed evaporation within Catchment 4:
print(evaporation4)
[1] 0.0323 0.0229

#We subtract the evaporation of Catchment 4 from the posterior mean of runoff:
pred.z=pred.z-evaporation4
```

Note that we need to subtract the evaporation of Catchment 4 from the posterior mean `pred.z`. Otherwise, the output would be (annual runoff + annual evaporation). We are mainly interested in annual runoff.

Finally we compare the predicted values with the "true", observed values. The results are displayed in Table B.1, and we see that the predictions are relatively close to the observed values.

Table B.1: The observed values of precipitation (y) at Gullbrå and the observed runoff (ζ^*) within Catchment 4 compared to the predicted values $\hat{\eta}$ and $\hat{\zeta}$ with posterior standard deviations.

	1995	1996	Unit
y	1.6	2.1	[m/year]
$\hat{\eta}$	1.8	2.6	[m/year]
Posterior sd	0.81	0.82	[m/year]
ζ^*	0.19	0.31	$[\cdot 10^8 \mathrm{m}^3/\mathrm{year}]$
$\hat{\zeta}$	0.22	0.34	$[\cdot 10^8 \mathrm{m}^3/\mathrm{year}]$
Posterior sd	0.088	0.089	$[\cdot 10^8 \mathrm{m}^3/\mathrm{year}]$

Bibliography

R.E Benestad, D. Nychka, and L. O. Mearns. Spatially and temporally consistent prediction of heavy precipitation from mean values. *Nature Climate change*, 2012.

M. Blangiardo and M. Cameletti. *Spatial and Spatio-temporal Bayesian Models with R-INLA*. Wiley, 1st edition, 2015.

G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije. *Runoff Prediction in Ungauged Basins*. Camebridge University press, 2013.

Noel Cressie. *Statistics for spatial data*. J. Wiley & Sons, 1993.

K. Engeland, I. Steinsland, S. Solvang Johansen, and A. Petersen-Øverleir. Effects of uncertainties in hydrological modelling. a case study of a mountainous catchment in southern norway. *Journal of Hydrology*, 2016.

Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Interpretable priors for hyperparameters for gaussian random fields, 2015.

D. Gamerman and H.F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC, 2006. ISBN 1584885874.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation, 2007.

X. Hu and I. Steinsland. Spatial modeling with system of stochastic partial differential equations. *WIREs Comput Stat*, 2016.

R. Ingebrigtsen, F. Lindgren, and I. Steinsland. Spatial models with explanatory variables in the dependence structure. *Spatial Statistics*, 2013.

Rikke Ingebrigtsen, Finn Lindgren, Ingelin Steinsland, and Sara Martino. Estimation of a non-stationary model for annual precipitation in southern Norway using replicates of the spatial field. *Spatial Statistics*, 2015.

Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

E. Krainski, F. Lindgren, D. Simpson, and H. Rue. The r-inla tutorial on spde models. `http://www.math.ntnu.no/inla/r-inla.org/tutorials/spde/spde-tutorial.pdf`, 2016.

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

Q. Mu, F.A Heinsch, M. Zhao, and S.W Running. Development of a Global Evapotranspiration Algorithm Based on MODIS and Global Meteorology Data. *Remote Sensing of Environment*, 2007a.

Q. Mu, M. Zhao, and S.W Running. Improvements to a MODIS Global Terrestrial Evapotranspiration Algorithm. *Remote Sensing of Environment*, 2007b.

NRK. Skadar for ein kvart milliard. http://www.nrk.no/sognogfjordane/skadar-for-ein-kvart-milliard-1.12536600, 2014. Accessed 26-May-2016.

NVE. Avrenningskart for norge. http://gis3.nve.no/metadata/tema/pdf/avrenningskart6190.pdf, 2002. Accessed 30-May-2016.

T. Reitan and A. Petersen-Øverleir. Bayesian methods for estimating multi- segment discharge rating curves. *Stoch. Environ Res Risk Assess*, 2009.

H. Rue and L. Held. *Gaussian Markov Random Fields: Theory and Applications*, volume 104 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 2005.

H. Rue and F. Lindgren. Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software, Volume VV, Issue II*, 2015.

H. Rue, S. Martino, and N. Chopin. *Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations*. Journal of the Royal Statistical Society, Series B. 2009.

Statkraft. Energikilder - vannkraft. http://www.statkraft.no/Energikilder/Vannkraft/, 2016. Accessed 24-May-2016.

W. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 1970.

M.A Wolff, A. Petersen-Øverleir, K. Ødemark, T. Reitan, and R. Brækkan. Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: results of a Norwegian field study. *Hydrology and Earth System Sciences*, 2015.