NTNU – Trondheim
Norwegian University of
Science and Technology

Predicting Snow Density

Åshild Færevåg

Master of Science in Physics and Mathematics
Submission date: Januar 2013
Supervisor: Ingelin Steinsland, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

Snow density is an important measure in hydrological applications. It is used to convert snow depth to the snow water equivalent (SWE). A model developed by Sturm et al. (2010) predicts the snow density by using snow depth, the snow age and a snow class defined by the location. In this work the model is extended to include seasonal weather variables and variables concerning the location. The model is tested and fitted for 4040 Norwegian snow depth and densities measurements in the period 1998 – 2011. A Bayesian modeling framework is chosen. To do inference a Markov Chain Monte Carlo method with Gibbs sampler is used, and cross-validation is used for model evaluation. The final model improved the snow density predictions for the Norwegian data compared to the model of Sturm et al. (2010). In addition year specific measurements are performed in different areas, and included in the model by using random effects. The associated reduction in the prediction error is computed, indicating a significant improvement by utilizing information of annual snow measurements.

Sammendrag

Snøtetthet er et viktig mål i hydrologiske sammenhenger. Snøtettehet og snødybde kan brukes for å finne snøens vannekvivalent (SVE). En metode utviklet av Sturm et al. (2010) estimerer snøtettheten ved hjelp av snødybde, hvilken dag det er på året, og en snøklasse definert av målingens lokasjon. I oppgaven er denne modellen utvidet til å inneholde sesongavhengige værvariable og variable knyttet til lokasjon. Modellen er tilpasset og testet på 4040 norske snødybde- og tetthetsmålinger fra perioden 1998 – 2011. Bayesiansk statistikk og Markov Chain Monte Carlo simulering med Gibbs Sampler er brukt for å tilpasse og finne modellparametre, og kryssvalidering blir brukt for å evaluere modellen. Den endelige modellen forbedret prediksjonene av snøtettheten for norske målinger sammenlignet med modellen til Sturm et al. (2010). Årsspesifikke målinger er lagt til i modellen ved å bruke tilfeldige effekter. Den assosierte reduksjonen i prediksjonsfeilen er regnet ut og viser en signifikant forbedring ved bruk av informasjon fra årets målinger i modellen.

Contents

Preface

This Master's thesis is the final part of my Master of Science Degree in Applied Physics and Mathematics at the Norwegian University of Science and Technology. This thesis, TMA4905 Statistics, gives 30 credits and is a part of the Industrial Mathematics programme. This thesis was carried out in 20 weeks and was completed January 2013.

Measurements of the snow water equivalent (SWE) provide an estimate of how much water there is on the ground that can potentially run off into rivers and streams. One of the most important properties of snow is the density - the relationship between snow depth and its water equivalent. To determine the snow water equivalent, the density needs to be known as well as the snow depth. To measure the snow density is more time consuming than measuring the snow depth. The main focus of this thesis is the snow bulk density, and how it relates to snow depth, age and other covariates concerning the location and the climate.

Organizing and preparing the relevant data have been an important part of the work. The software programs MATLAB, WinBUGS, OpenBUGS and Microsoft Excel were used for analysis, simulation and creating map displays.

I would like to thank everybody who has contributed to this thesis. I am grateful to 'the snowmen' from Statkraft, Oddbjørn Bruland and Knut Sand, for discussions and ideas, and for presenting snow data and meteorological data. My thanks go to Brian Taras for WinBUGS code and Glen Liston for permission to use data files and background materials. Above all I would like to thank my supervisor, Professor Ingelin Steinsland, for our weekly meetings with discussions, a lot of good advice, motivating me and giving me helpful comments. Thanks for making this period so interesting and exciting!

Åshild Færevåg

Trondheim, January 30th 2013

Chapter 1

Introduction

1.1 Motivation and objectives

Density is an important physical property of snow and establishes the relationship between the snow depth and the water content in the snow. A large proportion of the precipitation in Norway falls in the form of snow. When the snow melts, it makes a major contribution to the water in the rivers and in reservoirs. Keeping track of the snow distribution is essential in connection with risk of flooding, avalanche warning, climate research and hydropower production. Therefore it is important to have good and reliable estimates of the amount of snow through the winter. In the hydropower production, where snow represents an energy storage, exact estimates of the snow reservoir are important for planning the power generation.

In hydrological applications, the amount of snow is characterized by its snow water equivalent (SWE). The SWE can be determined from the snow depth and the snow bulk density. Measuring the SWE have traditionally been done manually by taking the weight of a cylindrical snow sample. The disadvantage of manual methods is that it is time consuming and costly. Other methods to determine the amount of snow are use of mobile snow radar and satellite pictures. The snow radar measures the snow depth, and the satellite images give information of areas of the earth covered with snow. Measuring the snow depth is much cheaper than measuring the snow density.

Since SWE can be calculated from the snow depth and snow density, estimating the snow density can be done equally as estimating the SWE. The snow density varies with the shape and crystal size, and is influenced by physical processes in the atmosphere during and after snowfall. This can make it difficult to determine the exact snow density. Different models have been developed in hydrology that can be used to calculate the snow density from the snow depth. One rule of thumb is that 10 cm of snow melts to 1 cm of water, but this is a very inaccurate approximation (Duesken and Judson, 1997). Sturm, Tara, and Liston (2010) came up with a model that esti-

mates the local snow bulk density with respect to snow depth, snow age and snow class (Sturm et al., 1995). Their study was based on snow data from the United States, Canada and Switzerland.

The method presented in Sturm et al. (2010) is of great interest also in Norway. But a problem with the model is that it does not consider the seasonal variation in the snow density in terms of climate. In Norway there are big differences in local climate and topography, and therefore a large variation in the snow density can be observed. An alternative is to include the effect of the weather directly in the model. In this work the model of Sturm et al. (2010) is extended to include seasonal weather variables, together with other explanatory variables. By using $n = 4040$ snow measurements given by Statkraft, the model is fitted to the Norwegian data by using Bayesian analysis, and model parameters are estimated using a Gibbs sampling Markov Chain Monte Carlo method. We compare our model to the model in Sturm et al. (2010).

The snow measurement data used in this study are carried out in different areas in Norway in the period $1998 - 2011$. The snow density conditions for each of these areas may vary between areas and from year to year. We can assess the variation among the different areas and years by including random effects that are estimated from year and area specific measurements. When using year-area random effects, an important question is how many year specific measurements in each area are needed to gain reduction in the prediction error.

The aims of this thesis are (i) to describe and test a method that predicts the snow bulk density in conjunction with snow depth, climate and location, (ii) develop the model by means of weather data, (iii) to fit and test different models by using Norwegian data, (iv) include random year and area effect to see if performing multiple area and year specific measurements provide better predictions, and (v) look at the associated reduction in the error by collection of year specific snow density measurements.

In Chapter 2 some background information of snow processes and the model of Sturm et al. (2010) are given. The study region and data are described in Chapter 3, together with some exploratory analysis. Chapter 4 gives an overview of the fundamentals in Bayesian analysis that are used for modeling. In Chapter 5 the snow model is outlined, followed by Chapter 6 describing the model evaluation criteria. In Chapter 7 different models are fitted to the data, model parameters are estimated and finally the models are tested to find the model that provides the most reliable predictions. The model is extended to include random effects and tested against uncertainty reduction for a different number of annual measurements. In Chapter 8 we look closer at the area Ulla-Førre. The thesis ends with discussion and conclusion in Chapter 9.

Chapter 2

Background theory

2.1 Snow water equivalent and bulk density

The snow water equivalent (SWE) is a measure of the amount of water contained within a snowpack, for instance the height of water you get when you melt the snow. SWE is a function of the snow depth (h_s) and the bulk density (ρ_b), defined by

$$SWE = h_s \frac{\rho_b}{\rho_w}, \tag{2.1}$$

where ρ_w is the density of water, 1 g/cm^3. The bulk density ρ_b of snow is mass per unit volume, an indicator of how compact the snow is. One commonly used conversion from snowfall to water content is that 10 cm of snow melt to 1 cm of water (ten-to-one ratio), i.e. a snow density of 0.1. This is an inaccurate approximation. The density of pure ice is 0.917, and the density of snow can vary from 0.01 to as high as 0.7 g/cm^3. For snow on the ground, the density is normally between 0.15 and 0.5 g/cm^3 (Killingtveit and Sælthun, 1995).

2.2 Snow processes

Snow density is influenced by the atmospheric conditions during crystal formation and descent, and conditions experienced when and after landing on the surface. Snow depth is the accumulation of new and old snow measured on the ground. The density increases with age and compaction under the weight of new snow.

The conditions of temperature, wind and humidity determine the form of the snow crystals. The wind breaks down the crystal structure in the snow, and consequently the snow will be more packed and the density will increase. After snowfall, wind action will increase the density when there is minus degrees. Low density usually requires light winds. If the temperature during snowfall is well below freezing

point, the snow has typically a rather low density, while warm temperature favors high density (Duesken and Judson, 1997).

2.3 Sturm's model

A model that estimates the snow bulk density has been developed by Sturm et al. (2010), based on data from the United States, Canada and Switzerland. It takes snow depth, snow age and snow class as input variables.

The snow class is found by a classification system for seasonal snow cover proposed in Sturm et al. (1995). It has six classes, where each class is defined in terms of physical characteristics of the snow and the snow layers. The classes are also derived by using three different climate variables given by the weather stations at the different location - wind, precipitation and air temperature, in a binary classification system. The snow class distribution in Scandinavia is shown in Figure 2.1. In Norway tundra and maritime snow are dominant.

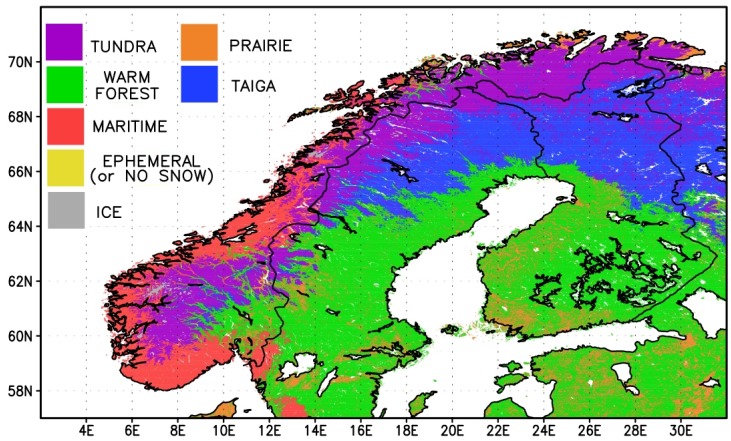


Figure 2.1: Snow class distribution in Norway. Source: Glen Liston (2011).

In the model of Sturm et al. (2010) bulk density is a function of snow depth (h_s), the day of year (DOY), and the snow class parameters k_1, k_2, ρ_0 and $\rho_{\max}$,

$$\rho_{h_i,DOY_i} = (\rho_{\max} - \rho_0)[1 - \exp(-k_1 \cdot h_i - k_2 \cdot \mathrm{DOY}_i)] + \rho_0, \tag{2.2}$$

where $0 < \rho_0 \leq \rho_{\max} < 1$. k_1 and k_2 are the densification parameters for depth and DOY, and $\rho_{\max}$ is maximum bulk density and and ρ_0 is the initial density of the individual snow layer. The snow class parameters are given in Table 2.1. The snow

season begins early in October. DOY represents the effect of snow aging and the number of days in the winter season, and runs from 1 October (-92) to 30th of June (+181).

Table 2.1: Model parameters by snow class (Sturm et al., 2010).

Snow class	$\rho_{\max}$	ρ_0	k_1	k_2
Tundra	0.3630	0.2425	0.0029	0.0049
Maritime	0.5979	0.2578	0.0010	0.0038
Prairie	0.5940	0.2332	0.0016	0.0031
Alpine	0.5975	0.2237	0.0012	0.0038
Taiga	0.2170	0.2170	0.0000	0.0000

Chapter 3

Data

3.1 Study region

In this study, snow depth and snow density observations from 244 locations within 17 different areas in Norway are available. In addition, meteorological data from representative weather stations are used.

The climate in Norway is characterized by the influence of the Gulf Stream and elevation differences. Norway is located between latitudes 57° and 71° N, and longitudes 4° and 32° E, with almost one-third of the country situated north of the Arctic Circle (Figure 3.1). Norway shares the same latitudes as Alaska and Greenland, but because of the location in the westerlies and a position on the east side of a great ocean with a warmer temperatures, Norway has a different climate in relation to these areas.

In the interior of Norway the temperature is determined by the solar radiation with warm summers and cold winters, while on the coast the sea temperature does not change that much, and there are relative mild winters and cool summers. The lowest temperatures can be found on Finmarksvidda, and the coldest temperature recorded is -51.4°C. The highest temperature measured is 36.5 °C in Nesbyen in the interior of Norway. The yearly mean temperatures varies from −3.1 °C in Finmarksvidda in the north to 7.7 °C on Karmøy on the west coast (met.no).

The west coast has some of the areas in Europe with most precipitation. The low pressure often comes towards the west coast, bringing mild and humid air from the ocean. In some places, up to 200 days of measurable precipitations are registered with more than 3000 mm of precipitation a year. The east of Norway is protected from the mountains, so there is less precipitation, normally around 1000 mm a year. Areas in the north are also protected by mountains, and have less precipitations, around 400 mm a year (met.no).

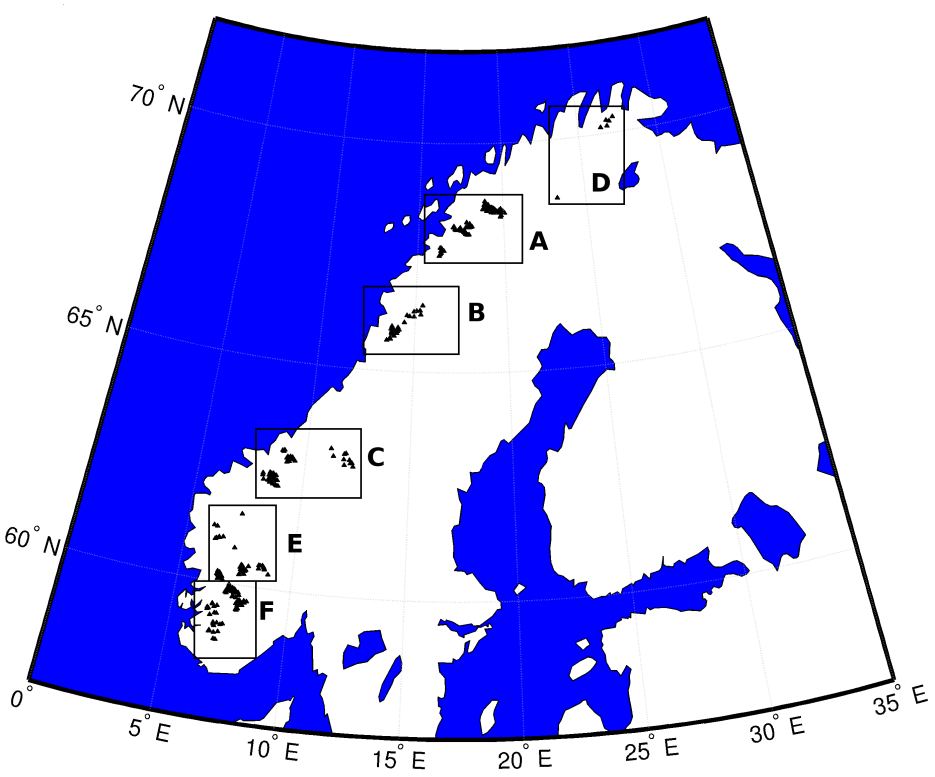


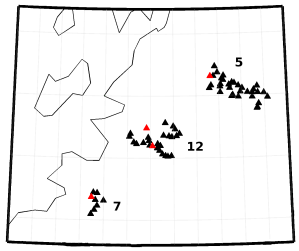
Figure 3.1: Locations in the study region. In Figure 3.2 each of the region A-F showed in this map are displayed in a separate map.

The location in the westerlies brings westerly and southwesterly winds over the country. The wind speed may reach storm on the coast, especially in the autumn and winter.

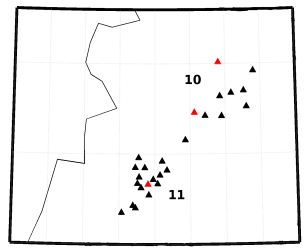
3.2 Snow data

Statkraft performs snow depth and density measurements in different areas in Norway. In this study we have used $n = 4040$ of these observations, from the period $1998-2011$. Snow data from 17 areas are used, and each area consists of several locations. These location compositions are set by Statkraft and displayed in Table 3.1 and specified in Figure 3.2. All locations are displayed in Figure 3.1.

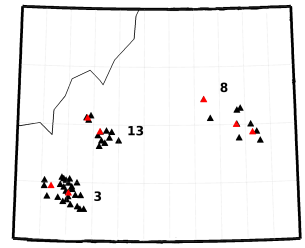
Snow depth and density data are carried out from November to May, $1-7$ times a year for each measurements area. The observations are based on manually snow



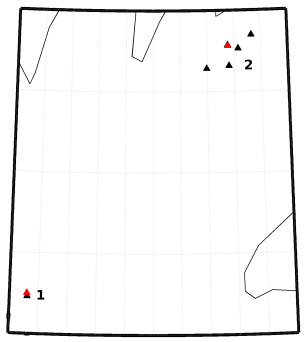
(a) 5: Innset, 7: Kobbelv, 12: Skjomen.



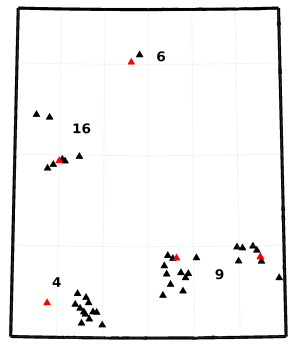
(b) 10: Rana, 11:Røssåga.



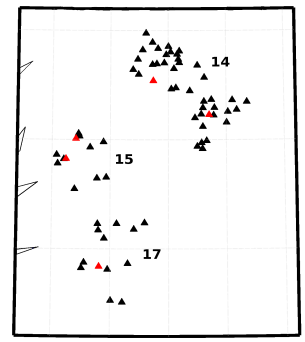
(c) 3: Aura/Grytten, 8: Nea-Nidelv, 13: Svorka/Trollheim.



(d) 1: Adamselv, 2: Alta.



(e) 4: Tysso-Folgefonn, 6: Jostedal,9: Nore, 16: Vik/Høyanger.



(f) 14: Tokke, 15: Ulla-Førre, 17: Sira-Kvina.

Figure 3.2: Map of the locations (black) and representative weather stations (red) used in this study. Figure (A)-(F) are the regions labeled in Figure 3.1.

measurements. The snow density is measured by taking the weight of a cylindrical snow sample tube. The density is obtained by dividing the weight by the volume of the tube. The snow depth is found by using a yardstick that measures the depth of accumulated snow.

3.2.1 Exploratory analysis

Table 3.1 displays the different measurement areas along with number of locations in each area, number of snow measurements used and mean characteristics of the snow data for each area.

There is variability in snowfall, temperature and wind between locations and between years. Because of differences in local climate and topography, a variation in the snow density in Norway can be observed both among areas and from year to year. To illustrate the variability, box and whiskers plots of the snow density in different years and in different locations are used. In these plots, the central mark indicates

Table 3.1: Information of the measurements areas used in this work: Number of locations, mean depth, snow density and snow water equivalent (SWE), and total number of measurements in each measurement area (n). Each area is given an id number that is used when refering to a specific area.

Id	Area	No. of locations	Depth (cm)	Density (g/cm^3)	SWE (mm)	n
1	Adamselv	5	74.1	0.345	26.4	42
2	Alta	1	51.7	0.266	13.4	18
3	Aura/Grytten	24/3	119.3	0.349	43.2	439
4	Tysso-Folgefonn	12	219.5	0.390	87.7	168
5	Innset	32	90.5	0.282	27.1	556
6	Jostedalen	1	190.3	0.327	64.1	39
7	Kobbelv	7	162.5	0.414	68.8	188
8	Nea-Nidelv	10	108.8	0.338	36.6	144
9	Nore	19	81.8	0.295	25.2	306
10	Rana	9	107.4	0.315	35.3	254
11	Røssåga	15	103.3	0.353	37.7	273
12	Skjomen	24	123.6	0.340	43.4	324
13	Svorka/Trollheim	3/9	126.5	0.376	50.3	243
14	Tokke	40	99.9	0.306	31.9	535
15	Ulla-Førre	10	138.6	0.398	57.6	243
16	Vik/Høyanger	5/2	184.8	0.384	75.4	191
17	Sira-Kvina	13	82.3	0.339	28.8	77
Total		244	118.2	0.337	42.5	4040

the median, while the edges of the box are the 25th and 75th percentiles. The outliers extend to the most extreme data. In Figure 3.3, densities from Ulla-Førre in the period 1983 – 2012 are used to show differences in the mean densities from one year to another. In Figure 3.4 densities in the different areas from the period used in this work are displayed.

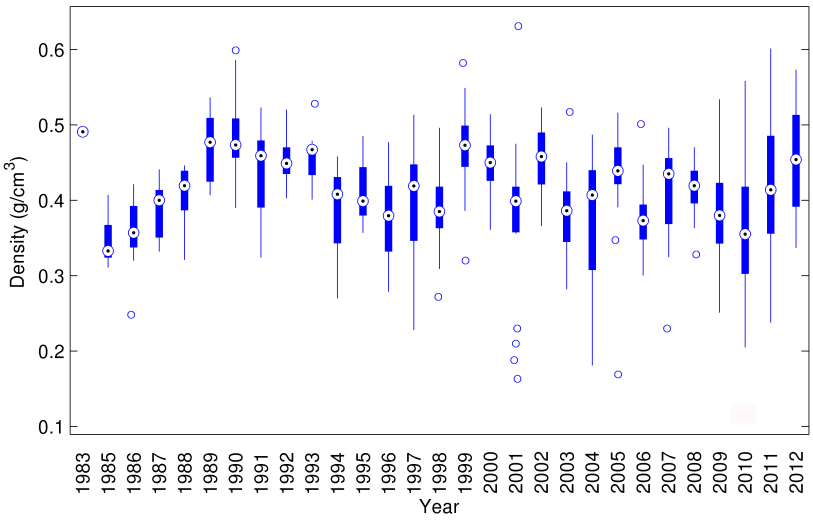


Figure 3.3: Box and whiskers plot: Observed snow density in the area Ulla-Førre in the period 1983 – 2012.

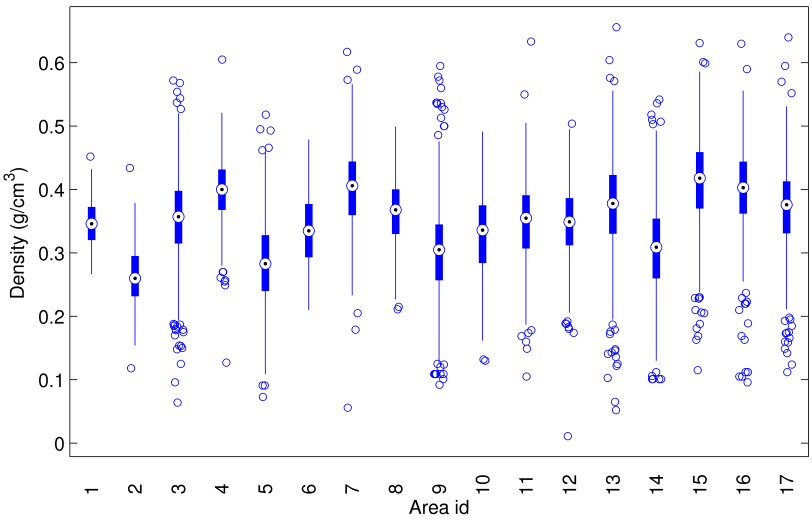


Figure 3.4: Box and whiskers plot: Observed density in the different areas in the period 1998 – 2011 (see Table 3.1 for definition of area id).

3.3 Meteorological data

Meteorological data from weather stations close to the locations need to be available. The weather stations used in this work are listed in Table 3.2, and their position relative to the locations are shown in Figure 3.2. Wind, precipitation and temperature data are available by Statkraft's own weather stations and from some operated by The Norwegian Meteorological Institute (DNMI). The weather observations used are registered hourly from 26 weather stations.

3.3.1 Correction of weather data

Temperature data for the different locations are taken from weather stations nearby. However, there are differences in the elevation between the locations where the measurement is made and at its representative weather station. The temperature normally decreases with an increase in elevation. Therefore, the temperature is corrected in relation to altitude by

$$T = T_0 + a(h - h_0), \tag{3.1}$$

where T is the air temperature to be found for the location at height h. T_0 and h_0 are the air temperature and height of the representative weather station. a is the rate of temperature change in the given region. In the troposphere $a = -6.5^{\circ}$C /km (Jacobson, 2005).

All of Statkraft's weather stations have wind measurements at the same height as the precipitation measurement, normally 2 meter above ground level. The DNMI weather stations have wind measurements 10 meters above the ground. The wind speed is a function of the logarithms to the altitude, so the wind speed for these stations needs to be corrected to a 2 meter level. A formula used by Statkraft is applied to these time series.

All precipitation measurements have error sources. Wind speed is the most important environmental factor which contributes to the underestimation of actual precipitation, and especially for snow precipitation (Killingtveit and Sælthun, 1995). Because of wind, the amount of precipitation measured in a gauge is less than actual precipitation reaching the ground. Thus the precipitation time series that are used are wind corrected by Statkraft.

3.3.2 Exploratory analysis

As mentioned in Section 3.1, different factors affect weather in Norway. Table 3.2 is an example of weather data in the period January-March 2011. In this period we can observe a variability in temperature, wind and precipitation. These climate variables may have a great impact on the accumulation and melting of snow, and consequently we can have a great variability in the amount of snow and snow density.

Table 3.2: In this table the 26 weather stations that are used in this study are listed. As an example, the characteristics of temperature, wind and precipitation data in the period January-March 2011 are shown.

Weather station	Area id	Temperature (°C)			Wind speed (m/s)		Precipitation (mm)
		mean	max	min	mean	max	total
Adam-Muora	1	-7.2	4.9	-25.3	5.2	18.7	112.9
Alta-Aidijavrre	2	-13.5	1.7	-36.2	1.6	7.0	2.3
Aura-Aursjøen	3	-5.9	4.3	-21.4	4.0	14.3	337.0
Aura-Eikesdal	3	0.8	13.9	-11.5	1.3	4.8	400.5
Folg-Botnane	4	-4.0	5.2	-16.4	2.8	18.0	923.6
Inns-Innset	5	-9.1	5.6	-26.1	2.1	5.4	349.5
Jost-Vigdøla	6	-4.7	4.9	-20.0	1.4	7.9	580.2
Kobb-Reinoksvatn	7	-5.5	6.0	-15.8	5.8	19.4	600.5
Sylsjøen	8	-6.9	5.1	-25.7	5.7	17.7	391.6
Nesjøen	8	-6.3	5.1	-25.1	2.9	14.4	527.5
Hersjøen	8	-3.7	7.1	-25.7	0.7	5.1	453.5
Nore-Lappstein	9	-8.8	1.2	-27.9	5.0	18.7	178.2
Nore-Pålsbu	9	-7.5	5.8	-32.9	0.9	4.8	81.77
Rana-Tverrvatn	10	-5.9	4.6	-30.0	2.8	11.7	564.3
Lang-Bjøllånes	10	-6.7	5.0	-26.8	0.8	5.9	632.2
Roes-Tustervatn	11	-5.6	4.8	-22.0	1.5	9.2	535.0
Skjo-Elvegård	12	-2.9	8.2	-19.6	1.3	4.7	206.7
Skjo-Kjørisdal	12	-8.2	3.3	-20.6	4.4	13.8	344.0
Svor-Solåsvatn	13	-4.3	8.7	-28.7	1.5	8.4	805.8
Tokk-Vinje	14	-3.9	9.9	-23.1	1.4	10.6	200.2
Tokk-Vågsli	14	-6.3	7.3	-27.0	1.7	9.6	430.6
Troll-Gråsjø	13	-2.6	11.4	-19.7	2.1	6.9	536.5
Ulla-Lauastøl	15	-0.7	13.9	-18.8	2.6	11.5	1021.0
Ulla-Osali	15	-2.6	4.9	-11.9	5.0	24.7	932.6
Vikf-Hestvollan	16	-5.3	2.7	-19.8	4.8	15.4	689.6
Sirdal-Sinnes (dnmi)	17	-3.7	6.0	-26.8	2.4	10.9	525.5

3.4 Construction of explanatory variables

Different factors influence the snow density. In this section we construct explanatory variables we believe might affect the snow density. These include seasonal weather variables as well as variables that relate to the location. In order to compute some of these variables we first need to explain a model that estimates the snow-melt process, the process of runoff caused by snow melting.

Snow accumulation and melt are important elements in the hydrological cycle in Arctic areas. In this application we use a snow-melt model to find a period for which the snow-melt does not exceed the snow accumulation.

3.4.1 Snow-melt: Degree-day model

This section is based on Chapter 3 in Killingtveit and Sælthun (1995).

The snow cover is exposed to thermodynamical processes. When these processes produce a positive energy supply to the snow cover, the snow melts. There are many factors that influence the melting of snow, like temperature, absorption of solar radiation, rain, wind, humidity and heat input from the ground (Duesken and Judson, 1997).

The degree-day model is an alternative to the more complex 'Energy balance' model, which is considered the most accurate for snow-melt computations (Killingtveit and Sælthun, 1995). A *degree day* is a measure that expresses a combination of temperature and time, and can be defined in different ways.

By using the air temperature above a given threshold temperature, the snow-melt can be computed by using a degree-day factor that assumes a correlation between average daily temperature and daily snow-melt. The degree-day model is expressed as

$$\begin{aligned} \mathrm{SM} &= C_X \cdot (T_a - T_s) && \text{if } T_a > T_s \\ \mathrm{SM} &= 0 && \text{if } T_a < T_s \end{aligned}$$

where SM is the amount of snow (mm/day) that is melted per day. This model assumes that the snow is melting as a function of the air temperature.

Table 3.3: Notation in snow accumulation and snow-melt model.

Parameter		Units
T_a	Mean daily air temperature	°C
T_s	Threshold air temperature for melting (usually 0°C)	°C
T_p	Threshold temperature rain-snow	°C
C_x	Degree-day factor	mm/(°C·Day)

The two others parameters are the degree-day factor C_X and the threshold air temperature for snow-melts, T_s. T_a is the mean air temperature in the location being investigated.

The HBV hydrology model is a mathematical model of the hydrological processes much used in Scandinavia (Killingtveit and Sælthun, 1995). Different areas operate with different degree-day factors and threshold temperatures. The parameters in Table 3.3 are estimated in the HBV-model for each area. To achieve the best model performance, these estimated parameters from Statskraft's operative HBV-model are used in this work. In the locations used in this study, the degree-day factors vary within a range 2.3 to 7, T_s from -0.9 to 0.8 and T_p between -0.5 and 0.6.

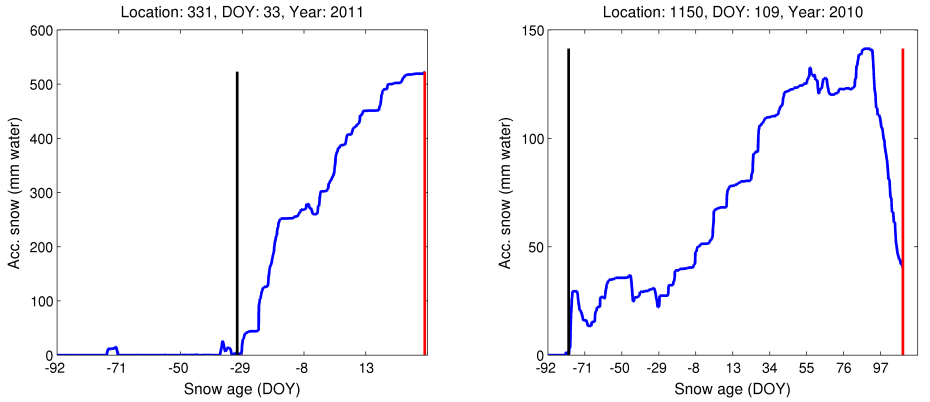
The accumulated snow (ACC) is the sum of precipitation (P), in mm/day, that comes as snow, minus the amount of snow that has melted (SM):

$$\begin{aligned} \mathrm{ACC}_i &= \max\{0, P_i + \mathrm{ACC}_{i-1} - \mathrm{SM}_i\} \quad \text{if } T_a > T_p \\ \mathrm{ACC}_i &= \max\{0, \mathrm{ACC}_{i-1} - \mathrm{SM}_i\} \quad \text{if } T_a > T_p \end{aligned}$$

where the ACC_i represents the accumulated snow at time i. T_p is the threshold temperature between rain and snow. In this study hourly measurements are used, so the degree-day factor needs to be adjusted to a degree-hour factor by dividing by 24 hours/day. Consequently, the snow-melt and precipitation measures in the model are recorded in mm/hour.

This snow-melt model is used in order to find which day in the snow season, before the day when the density measurement is performed (DOY), when the estimated snow-melt does not exceed the snow accumulation. This day is referred to as A_0. Day of year (DOY) was defined in Section 2.3 and can have values in the range -92 to +181. Hence, A_0 can take values between -92 and DOY, where -92 represents 1st of October. The use of A_0 will be further discussed in Section 3.4.2.

Figure 3.5 shows examples of the snow-melt and accumulation process in four different locations for different measurement dates, with different values of DOY. In Figure 3.5d we can see an estimated accumulation followed by an estimated snow-melt that exceeds the snow accumulation. A_0 in this plot is -18, and the accumulation period runs from $A_0 = 18$ to DOY$= 61$.



(a) Sira-Kvina (b) Adamselv

(c) Høyanger (d) Kobbelv

Figure 3.5: Examples of the snow-melt process in four different locations. The red line indicate the day of year (DOY) of the observation, the black line shows the accumulation start day (A_0), and the blue line the estimated accumulated snow.

3.4.2 Explanatory variables

Snow density at the time of observation is a product of the initial density of snow fall and further densification due to the weather and compaction of overlying snow. From the snow hits the ground, it begins a process of metamorphism, a process associated with recrystallization. The densification of snow is an irreversible process and the density increases with time and pressure.

Different explanatory variables (covariates) x_i are used in relationship to estimate the snow density. Three variables describe the snow density by its location and age, and six variables describe the climate during the snow season for the specific area. Here $T_t(i)$ is the air temperature (°C), $P_t(i)$ the amount of precipitation (mm) and $W_t(i)$ is the wind speed (m/s) registered at the weather stations at time t. The subscription i, means that the weather observation belongs to snow density measurement number i.

Below all accumulation of weather data are summarized hourly from the accumulation start date $A_0(i)$ to the day of measurement i, DOY(i). The accumulation starts from the time t corresponding $t = A_0$. The accumulation start date (A_0) is found by using the degree-day model in Section 3.4.1. If the snow-melt exceeds the snow accumulation, there will be no snow left on the ground. If this happens at some point before the snow depth measurement, there is no point in observing the weather's influence on the snow in this period, since the snow has nevertheless disappeared. A_0 is the last day after the beginning of the snow season (after 1st October) and before the measurement day, when this happens. In other words, A_0 is the day that defines the start of accumulation of snow towards the day the measurement is made (DOY).

All accumulated sums are scaled by a common number H. Here, H is chosen to be 24 for all locations and years, so the accumulation of for example hourly recorded wind speeds can be interpreted as accumulation of mean daily wind speeds. The index t represents the time in hours. The covariates are summed up in Table 3.4.

Snow depth, snow age and elevation: The snow density increases with age and higher compaction (snow depth). $x_1(i)$ represents the observed snow depth, which is the accumulated sum of new and old snow. $x_2(i)$ the day of year measurement i is carried out (DOY_i), representing the snow age. DOY starts at -92, the 1st of October and runs to +181, 30th of June. If DOY is 1, it means that the snow depth and density is measured 1st of January. $x_3(i)$ is the elevation, or the height, in meter above sea level (MASL) of the location of measurement i.

Plus degrees: In order to include the influence of high temperatures on the snow density, x_4 is the accumulated sum of plus degrees,

$$x_4(i) = \frac{1}{H} \sum_{t=A_0(i)}^{\text{DOY(i)}} T_t(i) \cdot \mathbb{1}\{T_t > 0\}. \tag{3.2}$$

The function $\mathbb{1}\{T_t > 0\}$ is the indicator function of the event $T_t > 0$ and has value 1 if $T_t > 0$, and 0 else.

Wind: Wind speed less than 2 m/s is here assumed to give no effect on the snow density. The variable x_5 is a measure of the accumulated sum of wind velocities above 2 m/s while temperature is below freezing point,

$$x_5(i) = \frac{1}{H} \sum_{t=A_0(i)}^{\text{DOY(i)}} W_t(i) \cdot \mathbb{1}\{T_t < 0\} \mathbb{1}\{W_t > 2\}. \tag{3.3}$$

Snowfall and wind: x_6 is the amount of accumulated precipitation falling as snow when there is wind,

$$x_6(i) = \frac{1}{H} \sum_{t=A_0(i)}^{\text{DOY(i)}} P_t(i) \cdot \mathbb{1}\{T_t < 0\} \mathbb{1}\{W_t > 0\}. \tag{3.4}$$

Precipitation type: It can snow at +2°C, and it can rain at -2°C. Therefore the density of the snow during precipitation are classified in three categories: x_7 is ratio of the total precipitation that comes when there is light snow, x_8 mixed snow and rain and x_9 when it is raining.

Light snow: $x_7(i) = \frac{1}{P_{\text{tot}}(i)} \sum_{t=A_0(i)}^{\text{DOY(i)}} P_t(i) \cdot \mathbb{1}\{T_t < -2\}$

Mixed snow/rain: $x_8(i) = \frac{1}{P_{\text{tot}}(i)} \sum_{t=A_0(i)}^{\text{DOY(i)}} P_t(i) \cdot \mathbb{1}\{-2 < T_t < 2\}$

Rain: $x_9(i) = \frac{1}{P_{\text{tot}}(i)} \sum_{t=A_0(i)}^{\text{DOY(i)}} P_t(i) \cdot \mathbb{1}\{T_t > 2\}$

$P_{tot}(i)$ is the total precipitation at the location of snow density observation i.

Table 3.4: Summary of the 9 explanatory variables constructed for the model.

Covariate	Name	Restrictions
x_1	Snow depth	-
x_2	Snow age (DOY)	-
x_3	Elevation (MASL)	-
x_4	Temperature	After and during snowfall/rain
x_5	Wind	During snowfall
x_6	Snow and wind	Snowfall and wind
x_7	Precipitation	Ratio as snow
x_8	Precipitation	Ratio as mixed snow/rain
x_9	Precipitation	Ration as rain

3.4.3 Correlation between covariates and snow density

Figure 3.6 shows the snow depth and SWE as a function of snow depth. We can observe a strong correlation between snow depth and SWE and also a tendency of increasing snow density with increasing snow depth.

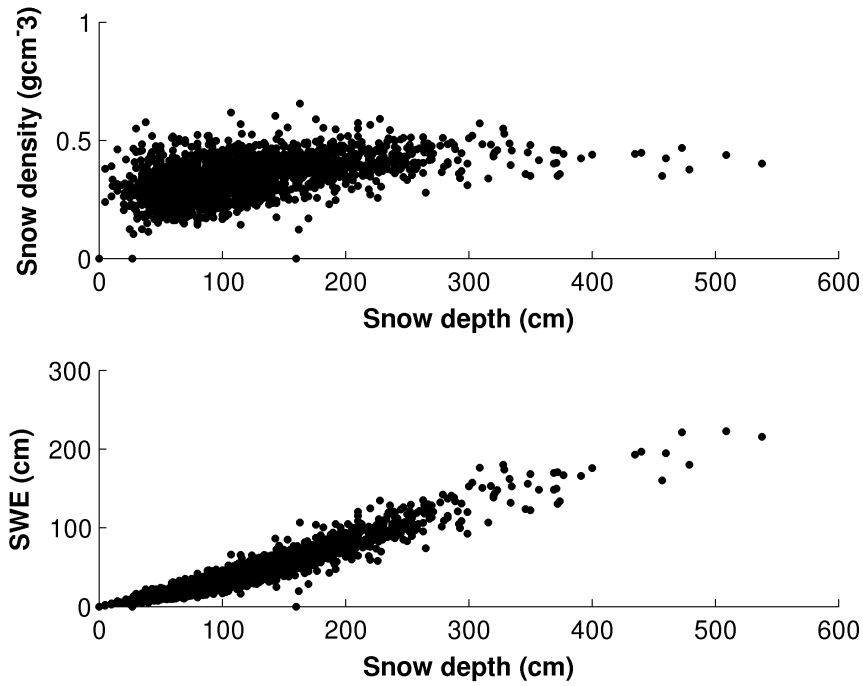


Figure 3.6: Measured depth versus measured density and SWE.

To measure the correlation between the density and the covariates, Pearson's coefficient of correlation is applied. Pearson's correlation coefficient between two variables is defined as the covariance of the two variables divided by the product of their standard deviations,

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y},$$

and can have a value between -1 and 1. The larger the value of the coefficient, the stronger the relationship is between the two variables, and the more likely it is possible to predict one variable by knowing the other one.

Table 3.5 contains all the correlations between the covariates and the density. This table shows that the correlation between depth and density is relatively high, meaning that high snow depth measured tends to be paired with relatively high snow

	Density	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1	0.48	1								
x_2	0.39	0.25	1							
x_3	0.20	0.39	-0.00	1						
x_4	0.16	-0.05	0.40	-0.30	1					
x_5	0.24	0.21	0.42	0.32	-0.06	1				
x_6	0.29	0.20	0.23	0.34	-0.02	0.53	1			
x_7	-0.05	0.08	0.09	0.34	-0.53	0.15	0.09	1		
x_8	0.07	-0.04	-0.09	-0.17	0.34	-0.02	0.08	-0.75	1	
x_9	0.03	-0.08	0.03	-0.35	0.60	-0.19	-0.19	-0.69	0.33	1

Table 3.5: Correlation coeffisient $\rho(x_i, x_j)$ between density and the covariates, and between two covariates, x_i and x_j. x_1: snow depth, x_2: day of year, x_3: elevation, x_4: accumulated wind, x_5: accumulated plus degrees, x_6: accumulated snow when there is wind, x_7: ratio light snow, x_8: ratio mixed snow and rain, x_9: ratio rain.

densities. The relationship between DOY and snow density exhibits an increase of snow density with DOY with a correlation coefficient of 0.39. This indicates that the snow depth and DOY can be used to predict the snow density. The correlation between density and elevation (x_3), wind (x_5) and precipitation (x_6), are also implying that there is an association between the density and these variables.

We can observe high correlations between the explanatory variables. For example there is a strong positive correlation between plus degrees (x_4) and type of precipitation ($x_{7,8,9}$). A positive correlation between x_4 and x_9 means that an increase in accumulated plus degrees gives a higher ratio of precipitation as rain.

Except among the variables $x_{7,8,9}$, the relationships between variables might not be as obvious as they seem. For example, high correlation between plus degrees and ratio of precipitation as rain does not necessarily mean that it is normally raining when there are high temperatures. There may be warmer periods without any precipitation, and more rain in colder periods.

Chapter 4

Bayesian analysis

4.1 Introduction

In statistics we distinguish between frequentistic and Bayesian statistics. In the frequentist approach, conclusions are drawn solely based on information provided by the random sample. One assumes that the data comes from some probability distribution, and considers the parameters in the distribution to be constant, but unknown. The key difference between these two approaches is that in the Bayesian one, these parameters are viewed as random variables. This chapter about Bayesian analysis is based on Chapter 2, 4 and 5 in Gamerman and Lopes (2006).

Statistical inference can be formulated as the process where conclusions are drawn from data. Let θ be the unknown parameter of interest and x the observation related to θ. $f(x|\theta)$ is the likelihood function of θ, and describes the assumption that the observed data were generated by θ.

In Bayesian inference, prior beliefs of the parameter θ before observing the value of x is defined. This is expressed as a probability distribution $p(\theta)$, and is called the prior distribution of θ. By combining the prior knowledge on θ and the information of θ provided by the data through the likelihood, the posterior distribution $p(\theta|x)$ can be obtained. It describes the information of θ after observing x. When the posterior distribution is obtained, inference using this distribution can be made.

The foundation of Bayesian statistics is Bayes' theorem. From the definition of conditional probability the posterior distribution is expressed as

$$p(\theta|x) = \frac{f(x|\theta)p(\theta)}{f(x)}, \tag{4.1}$$

where $f(x) = \int f(x|\theta)p(\theta)d\theta$ is the marginal density of x, and can be regarded as a normalizing constant in expression 4.1. However $f(x)$ is often difficult to evaluate, and only for very simple models the posterior $p(\theta|x)$ can be derived analytically. One

way to evaluate $p(\theta|x)$ is to use Markov Chain Monte Carlo (MCMC) methods which construct a Markov chain that generates samples for $p(\theta|x)$. Samples of the posterior distribution can be summarized through quantities as posterior mean, median and variance. Many summarization quantities are provided by integration of the form

$$I = E[t(\theta)] = \int t(\theta)p(\theta|x)d\theta. \tag{4.2}$$

This expression provides the posterior mean for $t(\theta)$. $t(\theta)$ is a function and $p(\theta|x)$ a probability function. It can be used in Monte Carlo approximation, when sampling from, for example, $p(\theta|x)$.

4.2 Markov Chain Monte Carlo (MCMC)

MCMC is a class of methods for sampling from a probability distribution using Markov chains. The MCMC approach is based on two ideas; Monte Carlo sampling and the stationary distribution of Markov chains.

Let $\theta^{(t)}$ denote the value of a random variable at time t, and let the *state space* refer to the range of possible θ values. The random variable is a *Markov process* if the transition kernel between different values in the state space depends only on the random variable's current state, i.e: $p(\theta^{(t+1)}|\theta^{(t)},\theta^{(t-1)},...,\theta^{(1)}) = p(\theta^{(t+1)}|\theta^{(t)})$. A *Markov chain* refers to a sequence of random variables $\theta = (\theta_0,...,\theta_k)$ generated by a Markov process. MCMC methods sample dependent realizations, $\theta^{(1)},\theta^{(2)},...,\theta^{(t)}$, that are asymptotically from the posterior distribution.

Assume a Markov chain $\theta^{(n)}$ with a continuous state space, a transition kernel $p(\theta,\cdot)$ and an initial distribution. To generate a value of this chain, a value for $\theta^{(t+1)}$ is sampled from $p(\theta^{(t)},\cdot)$. This procedure is repeated until convergence and a stationary distribution is reached. The Markov transition kernel $p(\theta,\cdot)$ needs to be defined. The two main methods to do this are Gibbs sampling and Metropolis-Hastings algorithm. In this work only the Gibbs sampler, described in Section 4.3, is used.

If a sample $\theta^{(1)},\theta^{(2)},...,\theta^{(t)}$ from $p(\theta|x)$ is available, then the integral in Expression 4.2 can be estimated by *Monte Carlo integration*

$$\hat{I} = \frac{1}{t}\sum_{i=1}^{t} t(\theta^{(i)}).$$

Sums of independent samples converge to Gaussian, but here the variables are dependent. Convergence will not be as fast as for independent samples, but the Monte Carlo estimator converges almost surely to the correct value when $t \to \infty$ (Gamerman and Lopes, 2006, p.96).

4.3 The Gibbs sampler

The Gibbs sampler is an important subclass of MCMC methods. Gibbs sampling is a MCMC scheme that uses the full conditional distribution as the transition kernel.

Assume that the distribution of interest is $p(\theta|x)$, where $\theta = (\theta_1, \theta_2, ..., \theta_k)$, x is the observation, and that the full conditional distributions of θ_i,

$$p(\theta_i|x, \theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_k) = p(\theta_i|x, \theta_{-i}),$$

which condition on both the data and the parameters, are available. Gibbs sampling is applicable when the joint distribution is not known explicitly, or when it is difficult to sample from it directly. The idea is that it is easier to sample from full conditional distributions than it is to obtain the marginal by integrations of the joint density.

One variable at a time is resampled, conditioned on all the others. This means that a set of variables are initialized randomly. For each iteration through the loop, one variable is selected and resampled based on all the other variables. This is repeated for new variables until convergence.

Algorithm for simulating a Markov chain with Gibbs sampler	
1. Initiate:	$\theta^{(0)} = (\theta_1^{(0)}, ..., \theta_k^{(0)})$, $t = 0$
2. For t=1,...,T:	a) Draw $\theta_i^{(t+1)} \sim p(\theta_i|x, \theta_{-i}^{(t)})$ for $i = 1, ...k$
	b) Set $\theta_i^{(t+1)} = \theta_i^{(t+1)}$
3. Change counter, t=t+1, and return to step 2 until convergence is reached.	

The value of the ith variable is drawn from the distribution $p(\theta_i|x, \theta_{-i})$. Thus during the tth iteration of the sample, $\theta_i^{(t)}$ is drawn iteratively from the full conditional distributions $\theta_i^{(t)} \sim p(\theta_i|x, \theta_1^{(t)}, ..., \theta_{i-1}^{(t)}, \theta_{i+1}^{(t-1)}, ..., \theta_k^{(t-1)})$. After one iteration with k variables the sampler becomes:

$$\begin{array}{lcl} \theta_1^{(t+1)} & \sim & p(\theta_1|x, \theta_2^{(t)}, \theta_3^{(t)}, ..., \theta_k^{(t)}), \\ \theta_2^{(t+1)} & \sim & p(\theta_2|x, \theta_1^{(t+1)}, \theta_3^{(t)}, ..., \theta_k^{(t)}), \\ \vdots & \vdots & \vdots \\ \theta_k^{(t+1)} & \sim & p(\theta_k|x, \theta_1^{(t+1)}, \theta_2^{(t+1)}, ..., \theta_{k-1}^{(t+1)}). \end{array}$$

The vectors $\theta^{(0)}, \theta^{(1)}, ..., \theta^{(t)}$ represent the realization of the Markov chain.

The Gibbs algorithm randomly samples from the posterior distribution and requires an initial starting point for the parameters. Then, one at a time, a value for each parameter of interest is sampled, given values for the other parameters and data.

Achieving good convergence can be tricky. A problem can be to determine how many iterations needed to reach the stationary distribution. Initial samples are not valid because the Markov chain has not been stabilized. A good chain will have rapid

burn-in, and convergence is reached quickly starting from an arbitrary position. The samples before stationarity and convergence are discarded, and the burn-in period is the number of these initial samples.

Another issue in MCMC is mixing. *Mixing* is the rate at which a Markov chain advances towards and explores the target distribution, before or after reaching the stationary distribution. Mixing problems can come from high correlations between model parameters and weakly identified model specifications (Levin et al., 2008).

4.4 Posterior predictive distribution

Predictions of future observables are based on predictive distributions, which refer to the distribution of the data averaged over all possible parameter values. The prediction of a future data $\tilde{y}$ after observed data y is based on the likelihood of the future data averaged over the posterior distribution (Ntzoufras, 2009). The posterior predictive distribution is defined by

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta,$$

where $p(\tilde{y}|\theta)$ is the sampling distribution, $p(\theta|y)$ is the posterior distribution $p(\theta|y) \propto l(\theta;y)p(\theta)$ and $\tilde{y}$ is the replicated value of y. The posterior predictive distribution can be compared to the observed data in order to evaluate the model.

4.5 Software: BUGS

The software package BUGS (Bayesian Analysis Using Gibbs Sampling) is a Bayesian analysis software that uses the Gibbs sampler to fit statistical models (Spiegelhalter et al., 2003). In this thesis two versions of BUGS, WinBUGS and OpenBUGS, are used to fit the models, and to find the posterior distributions for the model parameters.

Chapter 5

Snow density model

5.1 Introduction

In this chapter the statistical ideas behind the model in Equation 2.2 are presented. We also introduce a mixed model with random effects, where year and area specific measurements are included directly in the model.

5.2 Model

The snow density is modeled following Sturm et al. (2010), by assuming a beta distribution

$$Y_i \sim Beta(\alpha, \beta),$$

where Y_i is the snow density for the i'th observation.

A random variable $y \in [0,1]$, follows a beta distribution with parameters (α, β) if its density is given by

$$f(y;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1},$$

where $\alpha, \beta > 0$ and $\Gamma(\cdot)$ is the gamma function. The expected value and variance of the beta-distributed variable Y are

$$\begin{aligned} \mathrm{E}(Y) &= \frac{\alpha}{\alpha+\beta} \\ \mathrm{Var}(Y) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned}$$

It is convenient to parameterize a beta distribution in terms of its expectation. Following Kass and Raftery (1995, p.786), the beta distribution is reparameterized by setting

$$\nu = \frac{\alpha}{\alpha+\beta} \quad \text{and} \quad \omega = \frac{1}{\alpha+\beta}$$

where α and β are the parameters for the beta distribution: $\alpha = \frac{\nu}{\omega}$ and $\beta = \frac{1-\nu}{\omega}$. ν is the mean and ω is a precision parameter. So the model is reparameterized according to $(\alpha, \beta) = (\nu/\omega, (1-\nu)/\omega)$, and the distribution of Y becomes

$$f(y;\nu,\omega) = \frac{\Gamma(1/\omega)}{\Gamma(\nu/\omega)\Gamma((1-\nu)/\omega)} y^{(\nu/\omega)-1}(1-y)^{((1-\nu)/\omega)-1},$$

with $0 < Y < 1$ and $\omega > 0$. We can write $Y \sim \text{beta}(\nu, \omega)$. Here, E($Y$)=$\nu$ and Var(Y)= $\nu(1-\nu)/(1+\omega^{-1})$.

Following Sturm et al. (2010), the expected value is modeled by

$$\text{E}(Y_i) = (\rho_{\max} - \rho_0)[1 - \exp(\sum_{p=1}^{P} -k_p x_{p,i})] + \rho_0, \tag{5.1}$$

where k_p are the P number of originally unknown model parameters to be estimated, and $x_{p,i}$ are the explanatory variable for observation i belonging to model parameter p constructed in Section 3.4.2. It is the variables $x_1, x_2, ..., x_9$ that we in Chapter 7.2 will attempt to use in the model in different compositions.

The form of the equation implies that the expected value is between $\rho_{\min}$ and $\rho_{\max}$. The sign in front of the model parameters is negative, meaning that if the model parameters are positive, an increase in the covariates will consequently lead to a higher density.

Further, following Sturm et al. (2010) and WinBUGS code (see Appendix A) obtained by personal communication, ω is set to

$$\omega = \frac{\exp(\beta_0 + \beta_1 \cdot x_1)}{1 + \exp(\beta_0 + \beta_1 \cdot x_1)}.$$

Also, following Sturm et al. (2010), a uniform distribution was used for all parameters, $k_1 \sim \text{Unif}(0.0001, 0.0045)$, $k_2 \sim \text{Unif}(0.0001, 0.0045)$, $\rho_0 \sim \text{Unif}(0.3, 0.6)$, $\rho_{\max} \sim \text{Unif}(0.1, 0.7)$, $\beta_0 \sim \text{Unif}(-10, 1)$ and $\beta_1 \sim \text{Unif}(-0.1, 0)$.

In this work, other covariates with a different range of values are used, so the priors we have used are more vague and wider than the prior used in Sturm et al. (2010). They were chosen sufficiently vague to permit convergence and avoid run failure and set to

$$\begin{aligned}
k_p &\sim \text{Uniform(0, 0.08)},\\
\rho_0 &\sim \text{Uniform(0.1, 0.5)},\\
\rho_{\max} &\sim \text{Uniform(0.3, 0.8)},\\
\beta_0 &\sim \text{Uniform(-10, 1)},\\
\beta_1 &\sim \text{Uniform(-0.1, 0)}.
\end{aligned}$$

The densification parameters k_p are chosen positive, since the explanatory variables x_p are assumed to make the snow more compact and increase the snow density.

The MCMC with Gibbs sampler is run with 15000 iterations were the first 5000 discarded as a burn-in period to draw closer to the stationary distribution. After the model has converged, samples from the conditional distributions are used to summarize the posterior distribution of the model parameters. In this analysis there is no convergence problem, and only one set of initial values is used, generated by WinBUGS. The model that is implemented in BUGS is shown in Appendix A.

5.3 Random area and year effects

5.3.1 Area and yearly variation of snow density

In models where we use a dataset where the observations are grouped, it can be interesting to introduce *random effects*. We assume that there exist unobserved *latent variables* for each group of data, and the random effect describes the variation among groups in the dataset (Madsen and Thyregod, 2011).

The snow measurements used in this study are carried out in different areas in Norway annually in the period 1998 – 2011. The snow density conditions for each of these areas and years can vary among each other. The variability can be caused by various effects that is not explained by the weather variables or the other covariates. We can try to assess this variation among different areas by using random effects.

5.3.2 Model with random effects

Our model with random effects can be written as:

$$\mathrm{E}(Y_{ijk}) = (\rho_{\max} - \rho_0)[1 - \exp((\sum_{p=1}^{P} -k_p x_{p,ijk}) - \epsilon_{jk})] + \rho_0, \tag{5.2}$$

where Y_{ijk} is the response variable for observation i in area j in year k and $x_{p,i}$ are the explanatory variables.

The j different areas are likely to have different overall response for each year k. The model accounts for this by including a term ϵ_{jk}. ϵ_{jk} is the random effect for area j in year k. A distributional assumption needs to be added and we have chosen to use a uniform distribution as the random effects prior,

$$\epsilon_{jk} \sim \text{Uniform(-0.1,0.1)}.$$

The uniform distribution expresses the magnitude of the variability among locations, allowing stochastic group differences in various levels.

The random effects are estimated from annual measurements, because the model needs some kind of input of the specific year's condition. Computationally, the random effects can be implemented in the model in different ways. We have used three methods:

Method a) Here the model parameters k_p are fixed effects (effects that are assumed to be constant), and of the same value as the parameters in Table 7.3, and the random effect ϵ_{jk} are estimated in BUGS for each area and each year from the year and area specific measurements. This model equals to the model in Section 5.2 if no manually measurements are performed.

Method b) First the model parameters k_p are estimated together with the random effects ϵ_{jk} from all data in the training set. After, model parameters ,k_p, are used as fixed effects, while new random effects are estimated from measurements of the specific year. In a way the model expects random effects to be added to the model.

Method c) In a fully Bayesian setting it is desirable that all model parameters are estimated in terms of probabilistic distributions, and all parameters should be updated except for the data for which the model is tested.

Chapter 6

Evaluation

6.1 Introduction

Markov Chain Monte Carlo (MCMC) simulations are applied for Bayesian estimation by simulating the posterior distribution of the parameters of the model by using the softwares WinBUGS and OpenBUGS. The model performance is based on their predictive ability. To evaluate the models the mean error (ME), the mean absolute error (MAE), the root mean squared error (RMSE) and the continuous ranked probability score (CRPS) are used.

6.2 Mean error (ME)

The mean error is calculated by

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n} y_i - \hat{y}_i, \tag{6.1}$$

where y_i is the ith observation, $\hat{y}_i$ is the prediction and n is the number of observations. A positive value of the mean error implies that the predictions tends to be underestimated, and the opposite if the mean error is negative.

6.3 Mean absolute error (MAE)

The mean absolute error (MAE) measures the average of the absolute differences between predictions and observations. It is defined by

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|. \tag{6.2}$$

It is a quantity used to measure how close predictions are to the true observations, without considering the direction of the errors. The MAE is a linear score which

means that all the individual differences are weighted equally in the average. The error is measured in the same unit as the original data and the perfects score is 0.

6.4 Root mean square error (RMSE)

Another way of quantifying the error is the root mean square error. RMSE measures the square root of the average squared difference between the predicted and observed values. It is defined by

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{6.3}$$

Since the errors are squared before they are averaged, the RMSE gives a relatively higher weight to large errors than to small errors. RMSE is preferred when large errors are not desirable. The error is measured in the same unit as the original data and the perfect score is 0.

6.5 Continuous ranked probability score (CRPS)

The continuous ranked probability score (CRPS) is described in Gneiting and Raftery (2007). It can be used to compare the full probabilistic distribution with the observation.

Let the variable of interest be denoted y. In this work the prediction, y, is the snow density. The prediction pdf system is given by $f(y)$ and the observation y_o. The CRPS measures the distance between the probabilistic prediction f and the observation y_o. Figure 6.1 illustrates the idea behind CRPS.

CRPS is measured by the integrated squared difference between the cumulative distribution functions (cdf) of the predictions ($F(y)$) and the observations ($F_o(y)$). Here the cumulative distributions are

$$\begin{aligned} F(y) &= \textstyle\int_{-\infty}^{y} f(x)dx \quad \text{and} \\ F_o(y) &= H(y - y_o), \end{aligned}$$

where $F(y)$ is the predicted probability that y_o will be smaller than x, and H is the Heaviside function,

$$H(y) = \begin{cases} 0, & \text{if } y \le 0 \\ 1, & \text{if } y \ge 0 \end{cases}.$$

Then the equation for calculating the CRPS is

$$\mathrm{CRPS}(F, y_o) = -\int_{-\infty}^{\infty} (F(y) - \mathbb{1}\{y \ge y_o\})^2 dy. \tag{6.4}$$

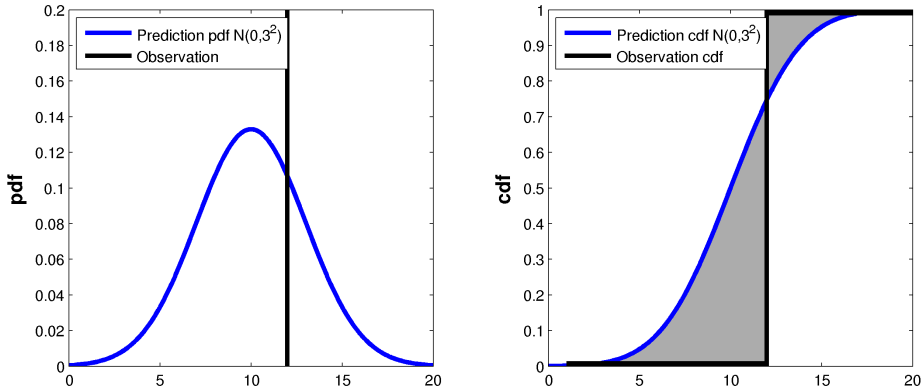


Figure 6.1: Illustration of the continuous ranked probability score. The CRPS is the integral of squared heights in the (shaded) region. The figure shades the absolute values not the squared heights of the CRPS. A normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 3$ is used in the illustration.

The CRPS can be decomposed into a reliability and a resolution part. Following Gneiting and Raftery (2007), CRPS can be written in closed form as

$$\text{CRPS}(F, y_o) = E|y_i - Y| - \frac{1}{2}E|Y - Y'|, \tag{6.5}$$

where F is the cdf of the predictive distribution for the model, y_i is the ith observed value, and Y and Y' are independent random variable from the distribution F. The expression in 6.5 can be approximated by and calculated from the MCMC samples using the estimates

$$\begin{aligned} \hat{E}|y_i - \tilde{y}_i| &= \tfrac{1}{N}\textstyle\sum_{n=1}^{N}|y_i - \tilde{y}_i^{(n)}| \\ \hat{E}|\tilde{y}_i - \tilde{y}_i'| &= \tfrac{1}{N^2}\textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}|\tilde{y}_i^{(n)} - \tilde{y}_j^{(n)}|, \end{aligned}$$

where $\tilde{y}_i$ and $\tilde{y}_j$ are independent replicated from the posterior predictive distribution $p(\tilde{y}|y)$. The CRPS can then be calculated by

$$\hat{\text{CRPS}} = \frac{1}{N}\sum_{n=1}^{N}|y_i - \tilde{y}_i^{(n)}| - \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|\tilde{y}_i^{(n)} - \tilde{y}_j^{(n)}|. \tag{6.6}$$

CRPS is negative oriented, so models with smaller CRPS are preferred. A perfect deterministic prediction has a CRPS of zero. The lowest possible score is achieved when $F = F_0$. CRPS is measured on the same scale as the observations. If the model is deterministic single-valued predictions, CRPS corresponds to the mean absolute error (Gneiting and Raftery, 2007).

The CRPS measures both sharpness and reliability of probabilistic predictions. Reliability says something about how many of the observations that are contained by the

predictive distribution. Sharpness refers to how wide the predictive distribution is, and CRPS rewards small spread (sharpness) if the prediction is accurate (Gneiting, Balabdaoui, and Raftery, 2007).

6.6 Weighted scores

In some locations there are more snow measurements than in other locations. If the mean score in each area are found based on all scores in the belonging locations, all areas contribute equally to the final score in MAE, RMSE and CRPS than other areas. The weighted score account for the difference in number of measurement in the locations.

The weighted score is the weighted combinations of all the scores in a set of data. For example, the weighted mean of a dataset $\{y_1, y_2,...,y_n\}$ with non-negative weights $\{w_1, w_2,...,w_n\}$ is

$$\bar{y} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}. \tag{6.7}$$

6.7 Evaluation schemes

6.7.1 Cross-validation in model selection

In order to get a more realistic prediction error, we would like to have a test dataset that is separated from our training dataset. A training set is the set of data that is used to estimate the model parameters, and a test dataset is the data used to assess the strength of the predictive model by using the evaluation criteria. In Table 6.1 a list of the training and test period is given for each area. Some areas have more than one period of training and test set. This is because one area can use multiple weather stations with different years of data, or because the various locations in one area do not have the same length in years of historical snow data. The training datasets in Table 6.1 are used to find model parameters both for the model with and without random effects.

6.7.2 Cross-validation in models with random effects

In the analysis of models with random effects, cross-validation is used. The data are partitioned k times into subsets, one training set and one test set. The training set is interpreted as the "measurements of the year". The data in the training set are randomly sampled from all the measurements of the specific year. The MCMC analysis is performed on the training subset, to estimate the random effect. The other test dataset is retained to evaluate the models by making predictions of the data. The average of the mean absolute values of the five subsets are used as prediction error.

Table 6.1: Years of snow and weather data in the training and test dataset.

Id	Area	Training years	Test years
1	Adamselv	2008-2010	2010-2011
2	Alta	2002-2007	2007-2011
3	Aura/Grytten	1998-2005	2006-2011
4	Tysso-Folgefonn	1999-2005	2006-2011
5	Innset	1998-2005	2006-2011
6	Jostedalen	1999-2005	2006-2011
7	Kobbelv	1998-2005	2006-2011
8	Nea-Nidelv	1998-2004	2006-2011
9	Nore	2001-2007/1998-2005	2007-2011/2006-2011
10	Rana	1998-2005	2006-2011
11	Røssåga	1998-2005	2006-2011
12	Skjomen	1998-2005	2006-2011
13	Svorka/Trollheim	1998-2005	2006-2011
14	Tokke	2005-2009	2009-2011
15	Ulla-Førre	1998-2005/2004-2008	2006-2011/2006-2011
16	Vik/Høyanger	1998-2005/2005-2009	2006-2012/2008-2011
17	Sira-Kvina	2009-2011	2011

Chapter 7

Results

7.1 Introduction

Based on about 4000 snow depth and density data, the model is tested and fitted for the Norwegian data. Bayesian analysis is used to estimate model parameters and random effects. We ran the Gibbs sampler for 15000 iterations. A burn-in of 5000 iterations was used for all MCMC runs, and model parameter estimates were based on a further 10000 iterations. First the model is found that predicts the snow density best without random effects. Afterwards, this model is tested with and without year and area specific measurements. The focus of the analysis is the snow density, since Equation 2.1 can be used to convert depth to SWE for all densities.

7.2 Model selection

In this section models with different explanatory variables (covariates) are tested. The models are tested for Norwegian snow depth data to see which model that provides reliable estimates of the bulk density. Equation 2.2 is applied to each observation using the different covariates, to estimate the bulk density.

The snow density is also calculated by Sturm's model in Equation 7.1, which uses the parameters in Table 2.1 defined by the snow classes. The snow classes for each of the 244 locations are found by the snow classification scheme described in Section 2.3.

All the models are compared between each other, and in addition versus Sturm's model, based on the the predictive distribution and the true measurement. To measure the model performance, the three criteria described in Section 6 are used.

10 different models where tested from the constructed explanatory variables. The models are listed in Table 7.1. We started with prior beliefs of what would affect the

Table 7.1: Different test models. x_1: snow depth, x_2: day of year, x_3: elevation, x_4: accumulated wind, x_5: accumulated plus degrees, x_6: accumulated snow when there is wind, x_7: ratio light snow, x_8: ratio mixed snow and rain, x_9: ratio rain.

Model	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
A	✓						✓	✓	✓
B	✓			✓	✓	✓			
C	✓			✓	✓		✓	✓	✓
D	✓			✓	✓				
E	✓		✓	✓	✓				
F	✓		✓	✓		✓			
G	✓	✓		✓	✓				
H	✓		✓	✓					
I	✓		✓	✓	✓	✓			
J	✓	✓	✓	✓	✓	✓	✓	✓	✓

snow density. Manually weight samples of the snow are potentially replaced by snow depth measurements. The snow density is highly correlated with the snow depth, so the snow depth is a part of every model. From one model to another, explanatory variables are removed or added, and a potential improvement is calculated by using the evaluation criteria MAE, RMSE and CRPS. Table 7.2 displays the result.

As a start, snow depth together with the ratio of the three precipitation ratio variables, x_7, x_8 and x_9, were selected (Model A). This was suggested by experts in Statkraft. Further, x_4 and x_5 were added (Model C), giving better results. By removing $x_{7,8,9}$ from Model C, we observe that the predictions are better without these variables (Model D).

By adding x_6 to Model D and E we get Model B and I, respectively, and we can see from Table 7.2 that x_6 provides poorer scores. This may be caused by the fact that the amount of snow is already represented through the snow depth x_1. Adding x_2 to Model D (Model G) and removing x_5 from Model E (Model H) give also worse predictions.

Model B, C and E provide quite similar results. Model B and C score best in the MAE, but Model E have a better score in CRPS. Since we are considering a predictive distribution, this model were chosen based on the lowest CRPS. The only difference between Model D and E is that the covariate elevation (x_3) is included in Model E. The wind is not corrected with respect to altitude, but one can assume that the elevation is already incorporated for example in the correction of air temperature due to altitude. But the CRPS implies a slightly better estimate.

In Figure 7.1 snow density estimates, Sturm's model and the posterior mean from model A, E, G are plotted together with the observed snow density. Sturm's model

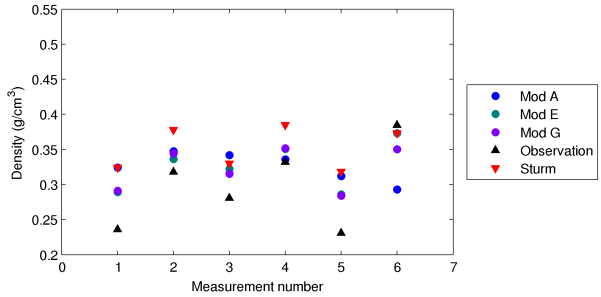
Table 7.2: Various evaluation criteria: Mean absolute error, root mean square error, continuous ranked probability score and weighted scores of MAE and RMSE. Different model are tested against measured densities using the test dataset. The framed value indicate the best score for the different evaluation criteria.

Model	MAE	Weighted MAE	RMSE	Weighted RMSE	CRPS
A	0.0508	0.0537	0.0607	0.0648	0.03776
B	0.0465	0.0491	0.0567	0.0602	0.07472
C	0.0460	0.0487	0.0556	0.0590	0.03440
D	0.0460	0.0486	0.0555	0.0589	0.03440
E	0.0462	0.0488	0.0558	0.0590	0.03418
F	0.0493	0.0520	0.0596	0.0632	0.07675
G	0.0467	0.0490	0.0561	0.0590	0.03451
H	0.0477	0.0504	0.0572	0.0608	0.03538
I	0.0465	0.0490	0.0567	0.0600	0.06931
J	0.1487	0.1497	0.1577	0.1593	0.08074
Sturm	0.0617	0.0636	0.0719	0.0743	0.06170

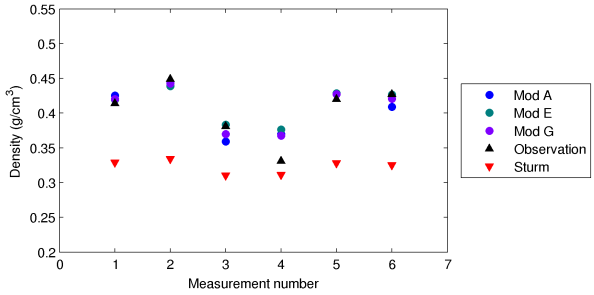
tends to underestimate the snow density for the tundra snow class. We can observe that these three other models perform quite similar results. In Figure 7.1b and 7.1c these three models predict the density well. In Figure 7.1a and 7.1d the estimated density are being overestimated and underestimated, respectively, but they still perform better than Sturm's model.

Evaluating the MAE and RMSE, most of the models in Table 7.2 provide better estimates than Sturm's model. This model must be assumed to be a deterministic model, hence the CRPS reduces to the MAE. Considering the CRPS, five of the models perform better than Sturm's model. Model J includes all the variables, and provides very poor results. This may be because the model has become overparameterized. Overparameterized models can cause overfitting to the training dataset, and consequently cause a poor fit to the test dataset.

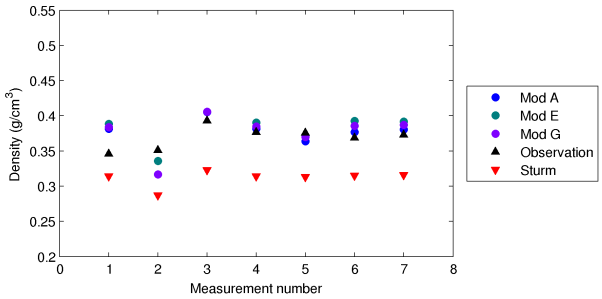
All combinations were not tested, but from Table 7.1 and Table 7.2, we can get a good idea of which explanatory variables that give an improvement or not. Other combinations of explanatory variables were tested, for example only including the ratio of heavy snow x_9. These models did not provide a better result and are not included in Table 7.2.



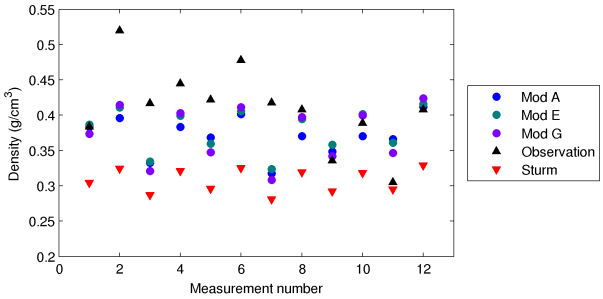
(a) Snow class: Maritime, location: Vinjerui, area: 14, period: 2009-2011.



(b) Snow class: Tundra, location: Illekleiv, area: 4, period: 2006-2011.



(c) Snow class: Tundra, location: Stordalen, area: 3, period: 2006-2011.



(d) Snow class: Tundra, location: Stearuvuggi, area: 7, period: 2006-2011.

Figure 7.1: Posterior mean estimates and observed density in four different locations for four different models without random effects.

7.3 Model 1: Prediction without year specific measurements

7.3.1 Model

Model E was selected by the CRPS criteria. This model is from now on referred to as 'Model 1'. The goal of the model is to predict the snow density in places where the density is not measured. The model takes snow depth (x_1), the elevation where the snow depth measurement is made (x_2), plus degrees (x_4), and wind speed (x_5) as input variables. Extending Sturm's model with new covariates, gives the following model

$$\mathrm{E}(Y_i) = (\rho_{\max} - \rho_0)[1 - \exp(-k_1 x_{1,i} - k_3 x_{3,i} - k_4 x_{4,i} - k_5 x_{5,i})] + \rho_0, \tag{7.1}$$

where Y_i is the estimated snow density for snow depth observation i.

A 5000 update burn-in followed by 10000 updates gave the parameter estimates in Table 7.3. The summary statistics, the empirical mean and standard deviation are displayed for each model parameter.

Table 7.3: Mean and standard deviation (std dev) of estimated model parameters and number of samples.

Variable	Mean	std dev	Sample size
ρ_0	0.1481	0.0148	10000
$\rho_{\max}$	0.4720	0.0128	10000
k_1	0.00503	5.48E-4	10000
k_3	0.00018	4.63E-5	10000
k_4	0.00477	6.82E-4	10000
k_5	0.00042	7.41E-5	10000

When using MCMC methods, it is important to make sure that the chain converges and reaches its stationary distribution. We need to make sure that we generate sufficient number of samples to provide a good estimate. Here we evaluate convergence based on the trace plot of the MCMC samples. Figure 7.2 shows that there is no convergence problems.

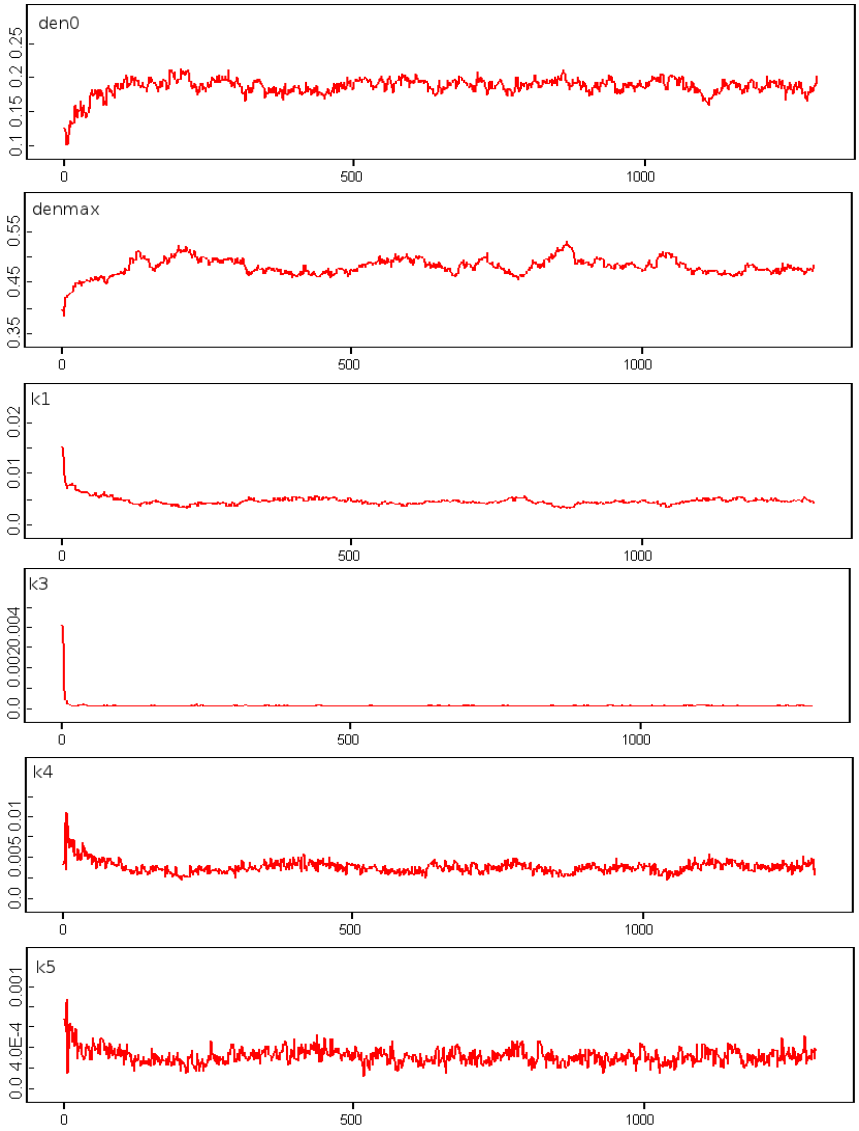


Figure 7.2: Trace plot of MCMC samples. The x-axis shows the number of iterations. The plot shows the 1500 first iterations.

7.3.2 Explanatory variables

For each of the snow density measurements, the values of the explanatory variables were calculated by using the equations in Section 3.4.2. Figure 7.3 shows the distribution of the explanatory variables. We observe that the SWE and snow depth have very similar distribution. The covariates vary within a different range of values, thus the model parameter values are adjusted to this value.

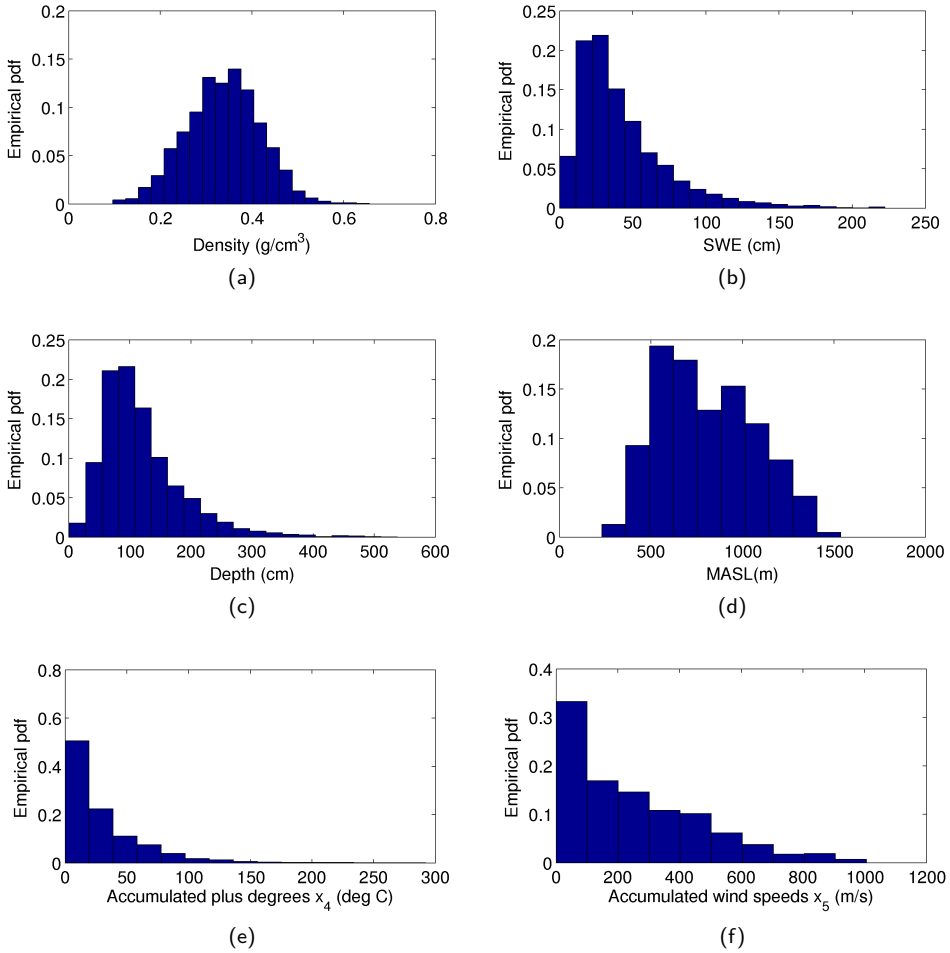


Figure 7.3: Histograms of the value of explanatory variables (both the training and test dataset).

7.3.3 Model properties

The model predicts the snow density by using the snow depth and other covariates as input variables. In Figure 7.4 different covariates are tested to see how the model estimates change within the value range of each covariate. The effects on density of one covariate are plotted, while the others are held fixed, using a representative value for each of these covariates. We can see that depth has the greatest impact on the snow density and elevation the least.

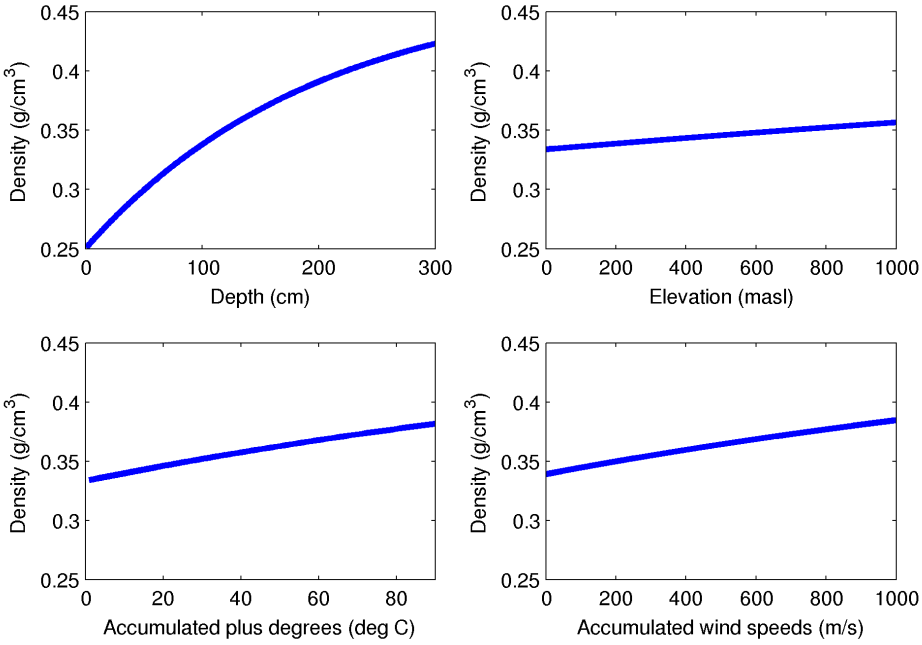


Figure 7.4: Estimated density by model with different values of x_1, x_3, x_4 and x_5. For the covariates not being evaluated, fixed values are used: Depth = 120, elevation = 800, accumulated plus degrees = 30, accumulated wind speeds = 250.

7.3.4 Error analysis Model 1

Additional error analyses of Model 1 are carried out by looking at the error in different years and in different locations. For simplicity only the mean error between measured density and posterior mean estimate of the density is considered.

Box and whiskers plot of the mean error for the different areas and in different years are illustrated in Figure 7.5 and in Figure 7.6. In these plots the central mark indicates the median, the edges of the box are the 25th and 75th percentiles. The outliers are plotted individually and extend to the most extreme data. In Figure 7.5 we can see that the error varies a lot. Area 5, 6 and 7 are some of the areas with the greatest prediction error, and while the mean error in area 13 is small, the error varies over a greater range than for example in area 1 and 2 with relative small mean errors. In Figure 7.6 the mean error in different years at Ulla-Førre is illustrated.

In Table 7.4 the mean error is found for different years in different areas. The bold numbers in this table are numbers that represent a positive/negative mean error, where the underlying errors the mean is calculated from, constitutes of 75% or more positive or negative errors of total errors, respectively. From Figure 7.6 and Table 7.4 we can conclude that the density in Ulla-Førre (area 15) tends to be underestimated. Table 7.4 together with Figure 7.5 may indicate that annual measurements and random effects can improve the model. This is discussed further in Chapter 7.4.

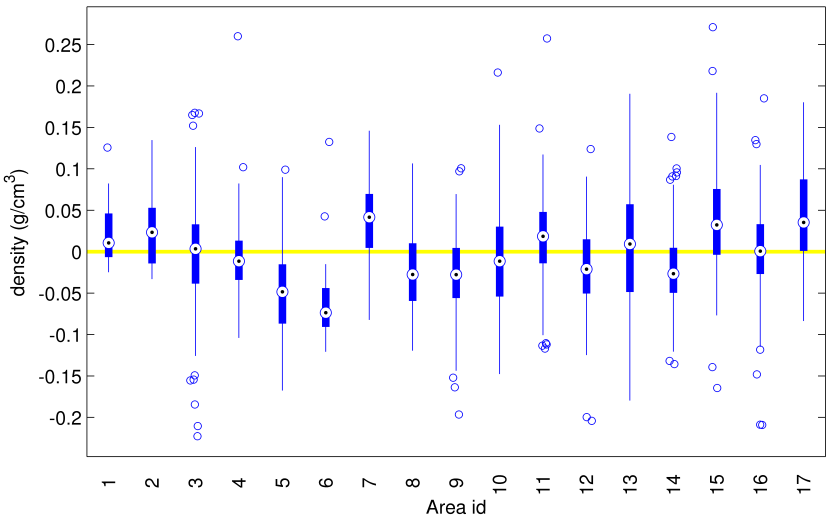


Figure 7.5: Box and whiskers plot: The mean error between observed and estimated snow density for the 17 different areas in the test period (see Table 3.1 for definition of area id).

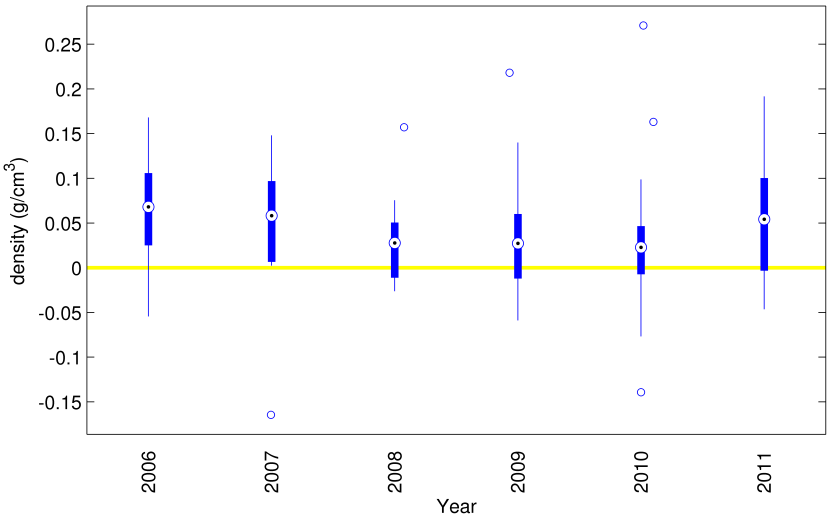


Figure 7.6: Box and whiskers plot: The mean error between observed and estimated snow density in the area Ulla-Førre in the period 2006 – 2011.

Table 7.4: Prediction mean error (ME) between observed and estimated snow density in the period 2006 – 2011 for all areas. Some of the areas do not have snow measurements from this period and can not be evaluated. The bold numbers are numbers where the underlying errors of the positive/negative ME constitute 75% or more positive or negative errors of total errors, respectively.

Area id \ Year	2006	2007	2008	2009	2010	2011
1	-	-	-	-	0.015	0.030
2	-	**0.046**	**-0.033**	0.022	**0.024**	0.062
3	**-0.045**	0.007	0.000	0.034	-0.019	-0.005
4	0.000	-0.016	**-0.005**	**-0.027**	**-0.018**	0.009
5	**-0.065**	-0.020	**-0.052**	**-0.049**	**-0.076**	-0.020
6	**-0.026**	**-0.050**	**-0.080**	**-0.087**	**-0.064**	**-0.047**
7	0.034	**0.064**	**0.065**	**0.052**	0.025	-0.003
8	**-0.056**	**-0.037**	**-0.025**	-0.014	0.004	-0.014
9	**-0.134**	0.018	-0.009	**-0.039**	**-0.041**	-0.028
10	-0.017	-0.015	-0.022	0.036	-0.013	-0.027
11	**0.020**	0.019	0.002	**0.054**	0.004	-0.001
12	-	0.021	**-0.029**	-0.017	**-0.037**	-0.026
13	0.010	**0.038**	-0.030	**0.072**	-0.033	-0.009
14	-	-	-	**-0.021**	-0.024	-0.023
15	**0.062**	**0.041**	0.027	0.035	0.026	0.052
16	-0.009	0.007	0.003	0.016	0.004	-0.020
17	-	-	-	-	-	**0.046**

7.4 Model 2: Predictions with year specific measurements

7.4.1 Model

To improve the predictions it is possible to collect annual measurements in every area. Here, the snow density is treated as different across areas and different among years following a random effects specification,

$$\mathrm{E}(Y_{ijk}) = (\rho_{\max} - \rho_0)[1 - \exp(-k_1 x_{1,ijk} - k_3 x_{3,ijk} - k_4 x_{4,ijk} - k_5 x_{5,ijk} - \epsilon_{jk})] + \rho_0 \tag{7.2}$$

where Y_{ijk} is the estimated snow density for snow measurement i in area j for year k. $i \in (1,...,n)$, $j \in (1,...,17)$ and $k \in (1998,...,2012)$, and ϵ_{jk} is the random effect of area j in year k.

Two different models with random effects are tested against the reduction in the prediction error by using snow density measurement information. This information is given by 5 and 20 snow density measurements in each area. These measurements are chosen randomly from all the location in the different areas and from the training dataset for the random effect. After randomly selecting a number of year specific snow densities for each area, these specific measurements are excluded from the evaluation and the remaining data are used as the test dataset.

In Chapter 5.3 we discussed three different implementations of the random effects. Here method a) and b) are modeled:

Model 2a) The parameters in Table 7.3 are fixed, and ϵ_{jk} are estimated by performing a MCMC simulation with 15000 iterations. These variables range from -0.1 to 0.1.

Model 2b) The model parameters and ϵ_{jk} are first estimated using the training dataset. The estimated fixed effects are displayed in Table 7.5. Thereafter new ϵ_{jk} are estimated in the same way as in Model 2a, by holding the pre-estimated k_p fixed, and new random effects are estimated using the specific training dataset of year specific measurements. In this way the model assumes that annual measurements are performed.

7.4.2 Error analysis Model 2

In Table 7.6 and Table 7.6 we show the mean absolute error for Model 2a and 2b and the associated uncertainty reduction compared to Model 1, discussed in Section 7.3 without the random effects. This model is tested on the same data as random effects Model 2a and 2b. The models are tested for all areas where 5 and 20 observations are available.

Table 7.5: Mean and standard deviation (std) of model parameters in Model 2a) and 2b).

Model		ρ_0	$\rho_{\max}$	k_1	k_3	k_4	k_5
2a)	Mean	0.1481	0.4720	0.00503	0.00018	0.00477	0.00042
	Std	0.0147	0.0128	5.48E-4	4.63E-5	6.82E-4	7.41E-5
2b)	Mean	0.2036	0.7966	0.00093	0.00011	0.00137	0.00015
	Std	0.0057	0.0033	5.76E-5	7.13E-6	1.02E-4	1.59E-5

The improvement from Model 1 to Model 2a and 2b with year specific measurements, is calculated as the relative change in percentage by $\frac{y-y_{\text{ref}}}{y_{\text{ref}}}$, where y is the mean absolute prediction error from one of the two models with random effects and y_{ref} is the model without random effects from Section 7.3 used as a reference model.

Table 7.6: Prediction mean absolute error (g/cm^3) and improvement (relative change in error) from Model 1 to Model 2a for available areas.

	MAE ($n = 5$)		MAE ($n = 20$)	
Year	Model 2a	Improvement %	Model 2a	Improvement %
1998	0.051	-3.0	0.035	-11.7
1999	0.070	-2.1	0.063	-10.9
2000	0.052	-1.2	0.037	-2.2
2001	0.056	-2.8	0.052	-10.8
2002	0.052	-2.4	0.050	-2.8
2003	0.054	-2.3	0.045	-5.3
2004	0.052	-0.4	0.051	-1.9
2006	0.055	-1.8	0.051	-9.2
2007	0.045	-0.7	0.043	-0.9
2008	0.044	-2.2	0.043	-6.7
2009	0.040	-3.9	0.033	-11.1
2010	0.051	-2.3	0.049	-7.2
2011	0.050	-0.6	0.051	-3.4

For Model 2a, the results are slightly improved by collecting $n = 5$ measurements. We can see that the results is consistently improved by the collection of more information through snow density measurements.

In Model 2b, the majority of the results for the different years are better with 20 measurements than for 5 measurements. The results are also much better than Model 2a. The two models are tested with the exact same dataset for the different years, and Model 2b performs better for all years when $n = 20$.

To illustrate the improvement of the model by performing annual measurements we look at snow density prediction with and without random effect in one specific year. Figure 7.7 displays the reduction in the error as a function of $n = \{0, 5, 10, 15, 25\}$ number of annual measurements in year 1998 for areas where data is available. To

Table 7.7: Prediction mean absolute error (g/cm^3) and improvement (relative change in error) from Model 1 to Model 2b for available areas.

	MAE ($n = 5$)		MAE ($n = 20$)	
Year	Model 2b	Improvement (%)	Model 2b	Improvement (%)
1998	0.046	-12.3	0.032	-19.5
1999	0.058	-23.8	0.048	-32.9
2000	0.051	-3.2	0.037	-2.7
2001	0.046	-20.1	0.039	-33.2
2002	0.047	-11.2	0.050	-2.9
2003	0.050	-10.3	0.044	-6.7
2004	0.052	-1.2	0.050	-4.3
2006	0.052	-8.0	0.042	-25.5
2007	0.044	-2.5	0.042	-4.1
2008	0.043	-3.6	0.036	-21.7
2009	0.034	-18.1	0.029	-22.0
2010	0.049	-6.6	0.041	-23.1
2011	0.051	+1.7	0.048	-9.7

evaluate on an equal basis, the random effects estimated for each area, year and n, are used on the exact same test dataset. This means that some of the random effects might have been estimated from some of these data, and will provide a better result than if proper cross validation was used. Still this plot can underpin the theory that Model 2b provides better results than Model 2a. It also supports the idea that multiple annual measurements give better result, but it is not necessarily true that the more the better. In Table 7.6 and 7.6 only 5 and 20 measurements in each area were tested. In the plot the error reduction stagnate after $n = 10-15$ measurements.

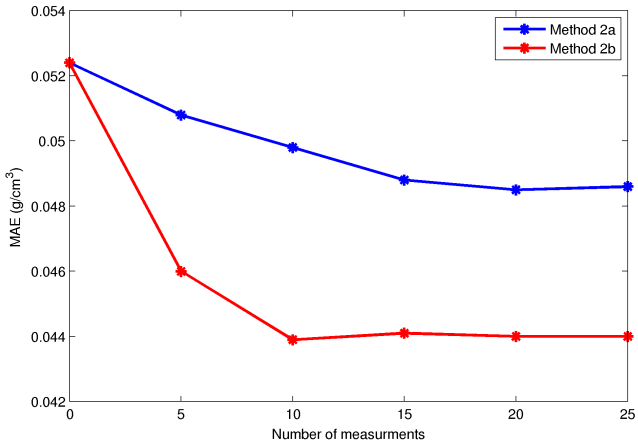


Figure 7.7: Error reduction in absolute error by collection of $n = \{0, 5, 10, 15, 20, 25\}$ number of year specific measurements. Data from year 1998.

7.4.3 Predictive distribution

A sample of 8 snow densities and the associated predictive distributions provided by models with and without year specific measurements are chosen. Posterior predictive densities of the snow density are shown in Figure 7.8. Densities are shown for Model 1 without random effects and Model 2b and 2c with random effects. We can observe that the models with random effects provide better predictions in most cases. Data from the area Ulla-Førre (area 15) are used to illustrate the predictive distributions, and we discuss this area and also Model 2c further in Chapter 8.

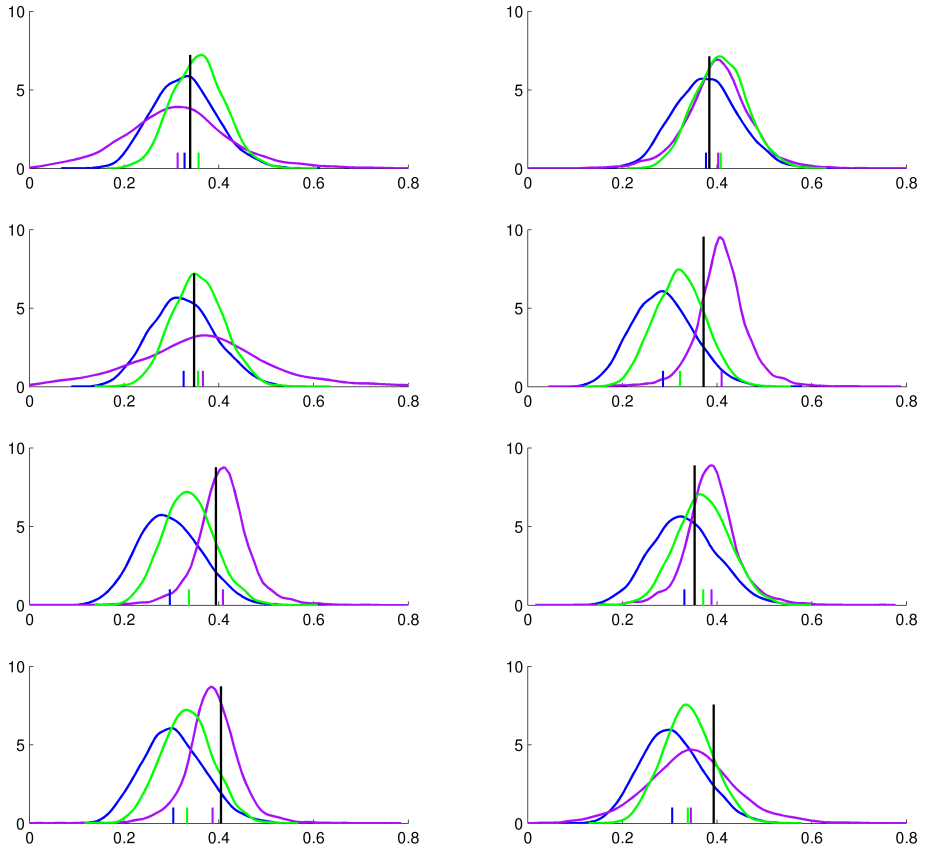


Figure 7.8: Predictive distribution of Model 1 (blue), Model 2b (purple) and Model 2c (green) for data from area 15 in year 2006. The number of year specific measurements is $n = 5$.. The black line indicates the observation and the other lines on the x-axis are the posterior mean of the three other models. y-axis: Probability density estimate of the samples, x-axis: Snow density (g/cm^3)

Chapter 8

Case study: Blåsjø

8.1 Introduction

In this chapter we look at snow measurements in one specific area in Norway - Blåsjø in Ulla-Førre (area 15). Ulla-Førre is Northern Europe's largest hydropower complex. Water for power generation comes from an area of about 2000 km^2. Blåsjø is the largest reservoir and consists of several mountain lakes 1000-1100 meters above sea level. They are governed into one large coherent reservoir. A large storage and pumping capacity makes Blåsjø an important part of the hydropower system in Norway (Statkraft).

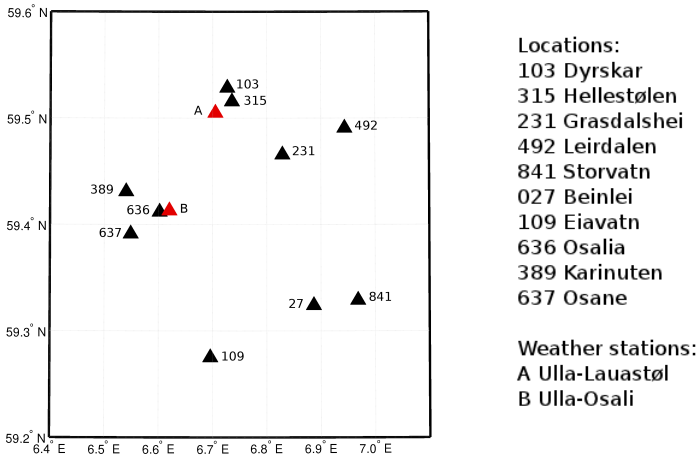


Figure 8.1: This map displays the two weather stations (red) and the locations (black) in this case study of the Ulla-Førre area.

Figure 3.3 in Section 3.2.1 shows the observed densities in the Ulla-Førre for different years. In this case study we want to apply the models with and without year specific measurements in this specific area. As Figure 7.6 and Table 7.4 indicated, utilizing information of annual measurements may be advantageous. The models are tested by collection of $n = \{0, 5, 10, 15, 20\}$ year specific measurements, that are used to estimate the random effects for each year in this area. We investigate the associated reduction in prediction error. The model with random effects are compared to the model without random effects.

8.2 Blåsjø

The data comprises $n = 141$ snow depth and density measurements from 10 locations in Ulla-Førre. Five locations close to Blåsjø are used: Grasdalshei, Leirdalen, Beinlei, Storvatn and Eiavatn. In order to get more data, five other locations in the Ulla-Førre area are also used: Hellestølen, Dyrskar, Osalia, Karinuten and Osane. Two weather stations are used in this case study, Ulla-Lauastøl and Ulla-Osali. The locations and belonging weather stations are presented in Table 8.3 and Figure 8.1.

In Table 8.2 the snow measurements from the training and test period, elevations and snow class for each location are presented. We can see that the snow depth varies over a much greater range than the snow density. By the snow classification system, 6 locations were classified as Tundra snow class, and 4 locations as Maritime snow class.

Table 8.1: Two weather stations are used for the locations in Ulla-Førre.

Name	Locations	Training dataset	Test dataset
Ulla-Lauastøl	231, 492, 315, 103	2004 – 2008	2008 – 2012
Ulla-Osali	27, 109, 841, 636, 389, 637	1998 – 2005	2006 – 2012

Table 8.2: Elevation, snow class and characteristics of density, depth and SWE data from the 10 locations in Ulla-Førre.

Location	Elevation (MASL)	Snow class	Density (cm)		Depth (g/cm^3)		SWE (cm)	
			Mean	Std	Mean	Std	Mean	Std
27	1080	Tundra	0.392	0.064	130.8	66.8	53.5	31.8
103	900	Maritime	0.418	0.092	166.4	65.6	72.6	37.2
109	1100	Tundra	0.419	0.066	165.9	61.4	71.1	33.2
231	1200	Tundra	0.414	0.061	157.2	43.9	66.2	23.9
315	650	Tundra	0.366	0.066	87.1	43.8	32.1	18.2
389	850	Maritime	0.413	0.064	153.0	74.6	65.5	37.3
492	1140	Tundra	0.403	0.051	239.8	78.4	99.3	42.2
636	750	Maritime	0.395	0.082	65.5	47.0	27.0	22.3
637	650	Maritime	0.384	0.094	69.7	37.4	27.5	17.3
841	1130	Tundra	0.422	0.047	193.4	73.4	83.3	35.8

8.3 Models

Different models are tested by collection of $n = \{0, 5, 10, 15, 20\}$ year specific measurements that are used to estimate the random effects for each year in this area. The model is compared to the model without any measurements. A brief summary of the models are given below:

Model 1: This is the base model from Section 7.3 which predicts the snow densities without any year specific measurements.

Model 2a: A random effect is added to model 1, estimated by using the measurements of the year (model described in Section 7.4). If $n = 0$ measurement is collected, the model equals to model 1.

Model 2b: This model assumes that random effects will be added to the model, and thus information through year specific measurements needs to be collected (model from Section 7.4). This model has other model parameters than model 1.

Model 2c: From a Bayesian view, it is desirable that the model parameters are estimated and not held fixed from training period. So in this model we use method c) described in Section 5.3. Except from the relevant test dataset, we are using all available information and the year specific measurements to estimate the model parameters.

8.4 Results

8.4.1 Analysis 1

Since Model 2b provided better results than 2a, Model 2b will be tested in this analysis. Model 2b requires annual measurements, so this model will be compared to Model 1 which equals to Model 2a with no annual measurements.

In this analysis we partition the data $k = 5$ times into training sets and test sets for each year, following the cross-validation procedure described in Section 6.7.2. The model predictions are compared with the observed snow densities. The average of the mean absolute values of the five subsets are shown in Table 8.3. The procedure is repeated for $n = \{5, 10, 15, 20\}$ number of year specific measurements in the period 2006 – 2011.

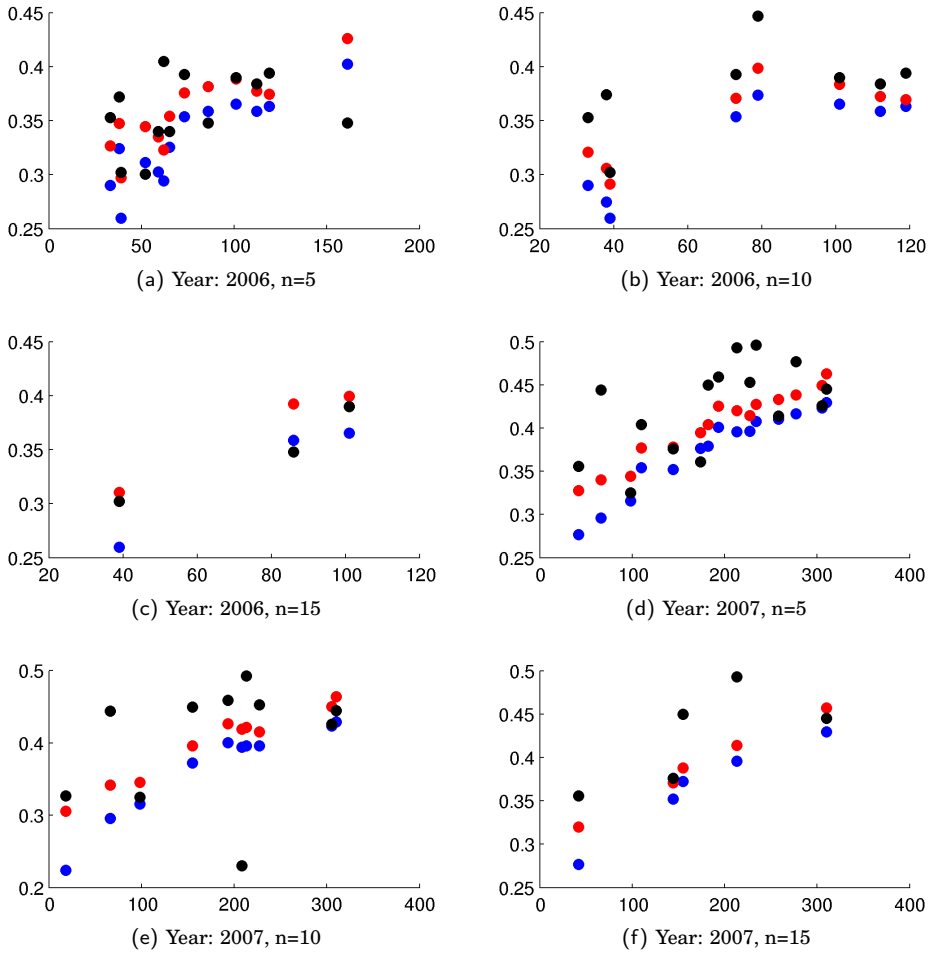
Table 8.3: Prediction mean absolute error (g/cm^3) and improvement (Imp.) from Model 1 to Model 2b for the data from Ulla-Førre.

Year	Total n	Mod 1	Mod 2b	Imp.	Mod 1	Mod 2b	Imp.
		$n = 0$	$n = 5$	%	$n = 0$	$n = 10$	%
2006	18	0.052	0.039	-24.2	0.054	0.037	-31.1
2007	20	0.058	0.047	-19.5	0.062	0.053	-14.8
2008	24	0.040	0.049	23.0	0.040	0.042	5.0
2009	26	0.050	0.043	-13.4	0.046	0.045	-3.1
2010	25	0.054	0.048	-11.4	0.053	0.052	-1.0
2011	27	0.064	0.059	-7.8	0.066	0.059	-11.3
		$n = 0$	$n = 15$	%	$n = 0$	$n = 20$	%
2006	18	0.029	0.027	-7.8	-	-	-
2007	20	0.067	0.053	-21.0	-	-	-
2008	24	0.040	0.040	0.5	0.041	0.035	-14.5
2009	26	0.044	0.038	-14.9	0.051	0.042	-17.9
2010	25	0.063	0.058	-8.5	0.047	0.042	-10.8
2011	27	0.065	0.057	-12.6	0.056	0.052	-7.5

The results presented in Table 8.3 show that Model 2b with year specific measurements provides almost consistently better results than Model 1 and 2b without year specific measurements. The only exception is year 2008, where Model 1 provides better estimates than the model with measurements. We notice that Model 2b in year 2008 is improving as the number of measurements increases. If we look at Table 7.4, we can see that the error in year 2008 in area 15 is relative small. If the model would have been perfect, the $n = 5$ collected measurements in the training dataset will influence the model in a disruptive way, and as n increases, the training set will approach the true model.

Another conclusion we can draw from this analysis is that an increase in the number of year specific measurements does not necessarily reduce the uncertainty. This can indicate that for example only $n = 5-10$ year specific of measurements are sufficient to establish the effects we want for the area in the specific year.

In Figure 8.2 some of the results of the mean posterior estimate of the snow density and the observations versus depth are presented. From this plot we can see that the effect of random effects potentially can change the estimated density up to 0.1 g/cm^3.



(a) Year: 2006, n=5

(b) Year: 2006, n=10

(c) Year: 2006, n=15

(d) Year: 2007, n=5

(e) Year: 2007, n=10

(f) Year: 2007, n=15

Figure 8.2: Observed and estimated density versus snow depth in Ulla-Førre in 2006 and 2007 for $\{n = 5, 10, 15\}$ measurements. Observation (black), Model 1 (blue), Model 2b (red). The x-axis is snow depth (cm) and on the y-axis the snow density (g/cm^3).

8.4.2 Analysis 2

In analysis 1 we observed that year specific measurements improved the model predictions, but not necessarily that the more measurements we make, the better the model predicts. In this analysis all the models are tested for $\{n = 0, 5, 10\}$. Here, also Model 2b is tested with $n = 0$ measurements, something that is missing in former analysis in this work.

Cross-validation for MCMC is time consuming. In this analysis the cross-validation procedure in Section 6.7.2 is performed, but only $k = 1$ training set is used to estimate the random effect, and one test dataset to evaluate the models. Consequently the analysis is not as robust as analysis 1, but we can get an impression of the strength of the different models. The results are presented in Table 8.4.

Table 8.4: The result of snow density and obtained by different model with and without random effects for each year in period 2006 – 2011. The models are evaluated by mean error (ME), mean absolute error (MAE) and continuous ranked probability score (CRPS) for $n = \{0, 5, 10\}$ annual observations.

Measure	Evaluation	N	2a ($n = 0$)	2b ($n = 0$)	2b ($n = N$)	2c ($n = N$)
Density (g/cm^3)	ME	5	0.034	0.029	0.012	0.021
		10	0.034	0.029	0.010	0.013
	MAE	5	0.050	0.048	0.043	0.044
		10	0.052	0.050	0.044	0.043
	CRPS	5	0.044	0.053	0.043	0.043
		10	0.043	0.042	0.035	0.041
SWE (cm)	ME	5	3.0	2.7	0.4	1.8
		10	3.2	2.9	-0.1	0.8
	MAE	5	5.9	6.1	5.6	5.6
		10	6.4	6.7	6.3	6.0

From Table 8.4 we can read that results with $n = 5$ and $n = 10$ are quite similar, but that models with random effects consistently leads to better predictions. It is noticeably that also Model 2b with $n = 0$ and Model 2a with $n = 0$ does not differ that much. This means that Model 2b actually predicts well, even if a random effects are assumed to be added. Model 2c was in advance assumed to give the best results, but gives very similar results as Model 2b with equal n. This may indicate that the training dataset in previous analysis is large and representative.

Chapter 9

Discussion and conclusion

In many applications the snow water equivalent needs to be measured. Measurements of the SWE provide an estimate of how much water that can potentially runoff into rivers and streams. For the long term inflow forecast it is important to have good and reliable estimates of the snow storage.

SWE can be determined by measurements of snow depth and density. Generally snow is not uniform in coverage, and several measurements are needed. The variability in snow depth are higher than for density, and the number of depth measurements are much greater than the number of density measurements. Snow depths are easier to measure than the snow density.

In this work, a method converting snow depth to SWE through estimating the snow bulk density has been discussed. However, the snow density is not always easy to estimate. A number of different processes influence the snow density. This can be climatic effects, topography and vegetation.

Sturm et al. (2010) came up with a model that uses snow depth, the age of the snow (DOY) and snow class found from geographical coordinates of the location as input variables. The snow density predicted from this model was in most cases underestimated for the snow class tundra. By using the snow classification dataset in Sturm et al. (1995) most of the 244 locations were classified as maritime (58) and tundra (175) snow class, and a few were classified as taiga snow. Apart from snow depth, the model of Sturm et al. (2010) does not include the variability in the density due to yearly variability in the climate. Using the snow classes to predict the snow density means that if depths measured on the same date in various years are equal, then the estimated density will be equal too.

Based on the previous work of Sturm et al. (2010), a model that directly includes the seasonal variability in the weather is developed. Our model is also an extension of the model of Sturm et al. (2010), using other explanatory variables and new model parameters found by Markov Chain Monte Carlo simulation with Gibbs sampler.

Several models consisting of different explanatory variables were tested against Norwegian snow data. The final model included the following predictor variables; (1) measured snow depth, (2) elevation (MASL) at the location, (3) temperature - accumulated plus degrees, and (4) wind - accumulated wind speeds above 2 m/s when the temperature is below freezing point. These variables need to be known for each location where snow density is to be predicted. We have used weather data from weather stations nearby the measurement locations.

The extended model includes the seasonal variability by the variables wind and temperature. The day of the year (DOY) is excluded from the model. This variable was relatively high correlated with the snow density, but gave worse predictions. This might be explainable since the effect of the age is already embedded in the measurement of snow depth and accumulation of weather variables. The elevation of the location did not influence the model much, but was chosen based on the CRPS score. An explanation can be that the effect of location altitude is already explained in the model through the correction of temperature due to altitude.

The accumulation period of the explanatory variables is defined from the last day after when the snow-melt does not exceed the snow accumulation to the day of year the measurements is performed. This is calculated by using the degree-day model. This model estimates the snow-melt based on the air temperature and takes location specific parameters as inputs.

Hence, various parameters and variables are involved in influencing the outcome of the model; Temperature, wind and precipitation data, parameters used in correcting these data series, the elevation of location and weather station, and of course the snow depth. In addition, parameters and data in the degree-day model, threshold temperatures for snow-melt and rain/snow and the degree-day factor are needed in each location. These parameters influence the number of weather data that are accumulated in the explanatory variables.

The model is easy to use as long as one have these input variables. All variables are subject to errors and there is a challenge to achieve the correct input data. Snow depth and elevation are almost exact variables, but there are more uncertainty associated with the weather data. The model requires that representative weather data are available, and hence the model is restricted by use of the data from weather stations nearby the locations. There can be a variation in topography between the location and the weather stations, and among locations using data from the same weather station. The model adjusts for this by correcting the temperature due to altitude of the different location. But still there might be a lot of difficulties in finding the correct precipitation, wind and temperature to each location. For example, the wind is not corrected due to altitude and some locations can be within a wind-protected area.

Even if the variability among areas and years is considered through the use of weather variables, it is conceivable that there is some other kind of variation in the properties of the snow densities that does not emerge through the explanatory variables. Thus, a random year-area effect is added to the model. This random effect is estimated by year specific measurements in different ways. The best predictions came from a model with random effects. This model used information from year and area specific measurements already in the beginning while estimating model parameters from the training dataset in the MCMC simulation. The use of year and area specific measurements assisted in bias removal and improved the model.

By increasing the collection of year specific measurements, it could in advance be assumed that the error in the estimated snow density would continue to decline. Analysis indicated that the model was improved, but that the result does not necessarily get better after more than $n = 10$ measurements.

The main focus of the evaluation and analysis in this work has been the bulk density and not the SWE. This is because it is the snow density that is dependent of the explanatory variables, and the SWE can be found from the density and snow depth. The snow density range is greater than the SWE range which has narrow limits. Consequently the SWE estimates derived from the measured snow depths fall close to the measured value. So even if there are relatively big differences between the measured and the estimated snow densities, the method gives better estimates for the SWE. If the error in snow density becomes large for small snow depths, the associated error in SWE will not be that much affected as it would have for large snow depths. As specified in the model, the uncertainty is larger for small snow depths. In future work different spatial models could be considered, and with other specifications of the variance.

The snow density prediction model developed in this study can be useful for estimating the snow density instead of manually measurements. As it takes wind speeds, temperature, elevation and snow depth as input variables, it is easy to use and applicable wherever weather stations are available and representative. This model is compared to the original model it is based on (Sturm et al., 2010), and gives more reliable predictions. Analyses show that information gained by collecting about 10 annual snow densities in each area can improve the predictions significantly.

Bibliography

P. Congdon. *Applied Bayesian Modelling*. 2003.

N. Duesken and A. Judson. *The Snow Booklet*. 2nd edition, 1997.

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

T. Gneiting, F. Balabdaoui, and A. E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 102:243–268, 2007.

T. Gneiting and A. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 2007.

M.Z. Jacobsen. *Fundamentals of Atmosphere Modeling*. 2nd edition.

R.E. Kass and A.E. Raftery. Bayes factor. *Journal of the American Statistical Association*, page 786, 1995.

M. Kery and M. Schaub. *Bayesian Population Analysis using WinBUGS*. Academic Press, 2011.

Å. Killingtveit and N. R. Sælthun. Hydrology. Norwegian Institute of Technology, Trondheim, 1995.

D. A. Levin, Y. Peres, and E.L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.

H. Madsen and P. Thyregod. *Introduction to General and Generalized Linear Models*. Chapman & Hall/CRC, 2011.

met.no. Værrekorder. *http://met.no/Klima/Klimastatistikk/Varrekorder/*, 2012.

I. Ntzoufras. *Bayesian Modelling Using WinBUGS*. Wiley, 2009.

D. J. Spiegelhalter, A. Thomas, N. G. Best, and D. Lunn. Winbugs user manual (version 1.4). 2003.

Statkraft. *Ulla-Førre (folder)*, 2012.

M. Sturm, J. Holmgren, and G. Liston. A seasonal snow cover classification system for local to global applications. 8:1261–1283, 1995.

M. Sturm, B. Tara, and G. Liston. Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, 11:1380–1394, 2010.

Appendix A

WinBUGS code

WinBUGS code for model and priors

```
#Likelihood
model
{
for(i in 1:n){
density[i]~dbeta(alpha[i],beta[i])
alpha[i]<-nu[i]/omega[i]
beta[i]<-(1-nu[i])/omega[i]
nu[i]<-(denmax-den0)*(1-exp(-k1*x1[i]-k3*x3[i]-k4*x4[i]-k5*x5[i]))+den0
omega[i]<-exp(beta0+beta1*depth[i])/(1+exp(beta0+beta1*depth[i]))
}

#Priors
denmax~dunif(0.3,0.8)
k1~dunif(0,0.08)
k3~dunif(0,0.08)
k4~dunif(0,0.08)
k5~dunif(0,0.08)
den0~dunif(0.1,0.5)
beta0~dunif(-10,1)
beta1~dunif(-0.1,0)
}
```

Table A.1: WinBUGS code for the model in Section 7.3