To appear in the *Journal of Applied Statistics*
Vol. 00, No. 00, Month 20XX, 1–25

Statistical models for short and long term forecasts of snow depth

Hugo Lewi Hammer
Department of Computer Science
Oslo and Akershus University College of Applied Sciences
Norway
Email: hugo.hammer@hioa.no

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

Forecasting of future snow depths is useful for many applications like road safety, winter sport activities, avalanche risk assessment and hydrology. Motivated by the lack of statistical forecasts models for snow depth, in this paper we present a set of models to fill this gap. First, we present a model to do short term forecasts when we assume that reliable weather forecasts of air temperature and precipitation are available. The covariates are included nonlinearly into the model following basic physical principles of snowfall, snow aging and melting. Due to the large set of observations with snow depth equal to zero, we use a zero-inflated gamma regression model, which is commonly used to similar applications like precipitation. We also do long term forecasts of snow depth and much further than traditional weather forecasts for temperature and precipitation. The long-term forecasts are based on fitting models to historic time series of precipitation, temperature and snow depth. We fit the models to data from six locations in Norway with different climatic and vegetation properties. Forecasting five days into the future, the results showed that, given reliable weather forecasts of temperature and precipitation, the forecast errors in absolute value was between 3 and 7 cm for different locations in Norway. Forecasting three weeks into the future, the forecast errors were between 7 and 16 cm.

Keywords: forecasting; meteorological data; snow depth; time series; zero-inflated gamma model

Classification codes: 62M10; 62M20; 62P12

Index to information contained in this guide

1. Introduction

The amount of snow, or snow depth, is important to many applications like road safety [9], risk assessments [1, 2], winter sport activities [7, 21], hydrology [8] and climate change research [3, 6, 13, 18, 24].

The modeling of snow depth is typically divided into three parts: snow accumulation (snowfall), snow aging and melting. The physics behind the different parts is quite complicated and depends on many factors. For snowfall, a common rule is the 10:1 rule stating that the density of the arriving snow is one tenth of the density of water. Following this rule, 10 mm of precipitation results in 10 cm of snow. In reality, the relation is more complicated. The density of snowfall is related to the ice-crystal structure by virtue of the relative proportion of the occupied volume of crystal composed of air. Snow density is regulated by in-cloud processes that affect the shape and size of growing ice crystals, subcloud processes that modify the ice crystal as it falls, and ground-level compaction due to prevailing weather conditions and snowpack metamorphism. Understanding how these processes affect snow density is difficult because direct observations of cloud microphysical processes, thermodynamic profiles, and surface measurements are often unavailable.

Roebber et al. [16] builds a neural network to classify snow density of snowfall to three classes heavy (1:1 < ratio < 9:1), average (9:1 < ratio < 15:1), and light (ratio > 15:1) where ratio refers to the density of water compared to the density of the arriving snow. The authors use the predictors solar radiation, temperature, humidity, precipitation and wind. The method only classify to correct density class in 60% of the cases which emphasize the difficulty in forecasting snow density and snow depth. Snow aging and melting of snow and resulting changes in snow depth depends on many factors like temperature, precipitation, solar radiation and the age and density of the snow pack. Compared to other factors, precipitation and temperature are the main drivers of changes in snow depth and most snow models are only based on these factors [3, 11].

Weather forecast services routinely forecast quantities like temperature, precipitation, wind and air pressure, but very rarely snow depth or other snow related quantities. A couple of exceptions are the `snow-forecast.com` [21] which forecasts snow depths for skiing resorts around the world and the `senorge.no` [19] which forecasts future snow depth in Norway.

The small number of weather services forecasting snow depth or other snow related quantities are in a big contrast to the amount logged snow depth data available around the world. In the US and Canada daily snow depth are logged for at least 8000 locations [3] and in Norway for over 1100 locations [17]. Motivated by the lack of models to forecast future snow depths and the large amount of historic snow depth data available, in this paper we present models to fill this gap. The contributions of this paper are three fold:

- Numerical snow depth models like Brown et al. [3] and *See Norway* [19] only use logged snow depths indirectly to calibrate the parameters of the models. In this paper we fit the statistical model to all available snow log data and calibrate *all* the parameters of the models. This may therefore result in better forecasts.
- We build statistical forecasts models. As described above, forecasting of snow depth is challenging and it is important to give uncertainty estimates on the forecasts. In addition, when making decisions it is oftn not the mean value that is of main interest but other quantities of the prediction distribution. E.g. a skiing resort may decide to open for the season if the probability that the snow depth the upcoming weekend will be below 30 cm is less then 10%.
- Available show depth models typically only predict snow depth when reliable forecasts of temperature and precipitation are available. In this paper we also present models to predict further into the future.

The paper is organized as follows: In Sections 2 and 3 we describe the statistical models for short and long-term forecasts of snow depth. The models are analyzed on real data in Section 5 and the paper ends with some closing remarks in Section 6.

2. Short term forecasting of snow depth

Temperature and precipitation are the main drivers of changes in accumulated snow depth. Better forecasts of future snow depth can therefore be achieved by including weather forecasts of temperature and precipitation in the forecast model. Such a model will be presented in this section.

Let D_t denote the current snow depth at a specific time of day at day $t \in \{1, 2, \ldots, n\}$ where n is the number of days with observations. Further let R_t and T_t denote the total precipitation and average air temperature for the last 24 hours.

Snow depth will for a large portion of days be zero (no snow), and else always larger than zero. Precipitation has the same properties and is typically modeled by a zero-inflated gamma model [14, 20, 22]. The model results in a good fit also to the snow depth data and can be formulated as follows

$$f(D_t) = P(D_t = 0)\,\mathrm{I}(D_t = 0) + P(D_t > 0) g(D_t)\,\mathrm{I}(D_t > 0) \tag{1}$$

where $g(D_t)$ is the gamma density and $\mathrm{I}(x)$ return 1 if x is true and 0 else.

We model $g(D_t)$ as a generalized linear model, and the expectation of the gamma density is linked to covariates as follows. We follow the typical assumption of changes in snow depth by dividing in snowfall (increase) and snow aging/snow melt (decrease) [3, 11, 12]. First we find a model for snowfall. If the temperature is sufficiently cold, precipitation will arrive as snow, and if the temperature is sufficiently warm, the precipitation will arrive as rain. For temperatures around zero °C, precipitation will arrive as a mixture of snow and rain. Several functions are suggested to model this, see e.g. Kienzle [10], Wen et al. [23]. We use the inverse logit function which fits well to our data and is also in accordance with observations from earlier research, see e.g. Figure 6 in Kienzle [10]. More specifically we assume that the expected depth of snow from R_t mm of precipitation at temperature T_t °C is given by

$$R_t\,\beta_0\,\mathrm{logit}^{-1}(\beta_1 + \beta_2 T_t) \tag{2}$$

where

$$\mathrm{logit}^{-1}(x) = \frac{e^x}{1 + e^x} \tag{3}$$

The parameter β_0 refers to the snow water density ratio and is usually set to 10 [16]. We will estimate the parameter from snow depth observation. We expect that β_2 will be negative such that higher temperatures result less snow (lower snow depth).

The snowpack tend to sink with time (compaction), but less if it is very cold. Further, for temperatures around and above zero °C the snow will additionally melt (transformed to water). The aging/melting is going faster if it is raining on the snowpack and is typically modeled with the index method (see e.g. Scherrer et al. [18]) which simply is an interaction between the temperature and the amount of rain, $(\beta_4 + \beta_5 R_t)T_t$. To ensure nonnegative snow depths we insert the index method in an inverse logit function. This means that we assume that the expected portion of the snow (depth) that disappears is given by

$$D_{t-1}\mathrm{logit}^{-1}(\beta_3 + (\beta_4 + \beta_5 R_t)T_t) \tag{4}$$

Adding the two parts together, given that $D_t > 0$ we assume that the expected snow depth at time t, is given by

$$E(D_t \,|\, D_t > 0) = e^{\mu} + \underbrace{R_t\,\beta_0\,\mathrm{logit}^{-1}(\beta_1 + \beta_2 T_t)}_{\text{snowfall}} + \underbrace{D_{t-1}\,\mathrm{logit}^{-1}(\beta_3 + (\beta_4 + \beta_5\,R_t)T_t)}_{\text{melting/aging}} \tag{5}$$

Since the expectation in (5) always is larger than zero, we simply use the identity link function when linking the expectation to the gamma distribution. The intercept e^{μ} typically becomes very small (see Table 2).

In gamma regression the far most common is to assume that the shape parameter, k, is constant such that the coefficient of variation becomes constant

$$\frac{SD(D_t)}{E(D_t)} = 1/\sqrt{k} = \text{const} \tag{6}$$

where $SD(D_t)$ denotes the standard deviation of D_t. Under this parameterization the standard deviation of D_t increases with the expected value of D_t. This is not a natural assumption for the modeling of snow depth and turns out to give a very poor fit to the data. If the temperature is below zero °C and it is no precipitation, the changes in snow depth is small, and D_t can be predicted with high precision from D_{t-1} even when D_{t-1} is high. A more natural assumption is that the variance depends on the expected *change* in snow depth. The larger expected changes, the larger uncertainty. To avoid the possibility that the variance is equal to zero, we combined this with the assumption of constant variance resulting in the following model for the variance

$$\mathrm{Var}(D_t|D_t > 0) = \sigma_1^2 + \sigma_2^2(E(D_t|D_t > 0) - D_{t-1})^2 \tag{7}$$

Given the expectation and variance, the shape, k, and scale, θ, parameter of the gamma distribution can be computed in the usual way

$$\begin{aligned} k &= \frac{E(D_t \mid D_t > 0)^2}{\mathrm{Var}(D_t \mid D_t > 0)} \\ \theta &= \frac{\mathrm{Var}(D_t \mid D_t > 0)}{E(D_t \mid D_t > 0)} \end{aligned} \tag{8}$$

An alternative to the model formulation above is to fit the data with a double generalized linear model, see e.g. the `dglm` package [5] in R [15]. We have not tried this.

Next we turn to $P(D_t = 0)$ in (1) which is modeled by logistic regression. One may estimate this probability using the covariates R_t, T_t and D_{t-1}. Instead we use $E(D_t \mid D_t > 0)$ as the only covariate

$$P(D_t = 0) = \mathrm{logit}^{-1}(\beta_6 + \beta_7 E(D_t \mid D_t > 0)) \tag{9}$$

which turns out to perform well. The intuitive is that higher expected snow depth, given by (5), results in lower probability of no snow.

Finally, given the covariates R_t, T_t, D_{t-1} and $R_{t'}, T_{t'}, D_{t'-1}$ for $t' \neq t$, we assume that D_t and $D_{t'}$ are independent. This makes it straight forward to put up the likelihood function for the snow depth observations. Due to the nonlinearities to the covariates and the coupling between the logistic and gamma part of the model, to the best of our knowledge there exists no statistical packages in R [15] or elsewhere to estimate the parameters of the model. Instead we estimate the parameters by implementing a steepest descent optimization algorithm and find the parameters that optimize the likelihood function.

Given the estimated parameters of the model, forecasting of future snow depths can be computed using Monte Carlo simulations. We assume that a reliable weather forecast of T_t and R_t is available for the next few days. The model above can then be used to "track" the probability distribution of snow depths into the future as follows. Given the current snow depth D_t and weather forecasts of temperature and precipitation the next day (T_{t+1} and R_{t+1}), generate a large set of realizations from the distribution $f(D_{t+1})$ in (1). Next, for each sample from $f(D_{t+1})$ and given weather forecasts two days into the future (T_{t+2} and R_{t+2}), generate a sample from $f(D_{t+2})$ and so on.

We tried some extensions to the model as described below, but neither of them improved the model with respect to the Akaike information criterion (AIC).

- One may expect an unsymmetry for the snowfall inverse logit function and thus we considered the extension $R_t\,\beta_0\,\mathrm{logit}^{-1}(\beta_1 + \beta_2 T_t + \beta_8 T_t^2)$.

- For the melting part of the model one may argue that if D_{t-1} is large, not the whole snow pack will be exposed for the air temperature and snow melting may go slower. We therefore tested the extension of the melting part of the model $D_{t-1}\text{logit}^{-1}(\beta_3 + (\beta_4 + \beta_5 R_t)T_t + \beta_8 D_{t-1})$.
- We also conditioned on previous values of snow depth adding the term $\exp(\beta_8 + \beta_9 D_{t-2})$ to (5).

3. Long term forecasting of snow depth

Weather forecasts for temperature and precipitation are typically reliable for three to six days into the future. In this section, we build models to forecast snow depth further into the future than this timespan. We consider two different strategies

- Model 1: When reliable weather forecasts are not available, we use historical observations of precipitation and temperature to build statistical time series models for these variables. The forecasts from these models are further used as input to the model in the previous Section.
- Model 2: We build a time series model for the snow depth data directly.

The strength of model 1 is that the model gives us simultaneous forecasts of temperature, precipitation and snow depth trends.

3.1 *Model 1*

Let $s(t)$ denote the day during a season for observation time t. E.g. if t refers to December 31 for some year, $s(t) = 365$ for ordinary years and 366 for leap years. For long-term forecasts of temperature and precipitation, the seasonal trends will be important. We apply Fourier series, which are able to model complex seasonal patterns with only a few parameters

$$h_m(t) = a_0 + \sum_{k=1}^{m} a_k \sin\left(k\frac{2\pi}{366}s(t)\right) + b_k \cos\left(k\frac{2\pi}{366}s(t)\right) \tag{10}$$

Except for the seasonal trend, temperature data fits well to an autoregressive process. Thus, we model the temperature time series using an Autoregressive process with a seasonal trend given by (10)

$$T_t - h_{m_T}(t) = \sum_{j=1}^{p_T} \alpha_j (T_{t-j} - h_{m_T}(t-j)) + \epsilon_t \tag{11}$$

where $\epsilon_t \sim N(0, \sigma_T)$ where $N(\mu, \sigma)$ denote a normal distribution with expectation μ and standard deviation σ. Further we assume that $\epsilon_t, t = 1, 2, \ldots, n$ are independent. Standard packages in R Core Team [15] can be used to fit this model, but instead we found the parameters that maximized the likelihood function using a steepest descent optimization algorithm.

For precipitation we follow the model in Stern and Coe [22] with some exceptions that will be explained below. Because of the many days with no precipitation the zero-inflated gamma model in (1) is suitable also to model precipitation

$$f(R_t) = P(R_t = 0)\,\mathrm{I}(R_t = 0) + P(R_t > 0)g(R_t)\,\mathrm{I}(R_t > 0) \tag{12}$$

Similar to the model for snow depth and the model in Stern and Coe [22] we model $g(R_t)$ with a gamma regression model. We define the expectation as

$$E(R_t \,|\, R_t > 0) = \exp\left(h_{m_R}(t) + \sum_{j=1}^{q_R} \gamma_{Rj} R01_{t-j} + \sum_{j=1}^{s_R} \kappa_{Rj} T^j\right) \tag{13}$$

were $R01_t$ is defined such that $R01_t = 1$ if $R_t > 0$ and $R01_t = 0$ if $R_t = 0$. Stern and Coe [22] also considers interactions between $R01_{t-j}$ for different values of j. We achieve almost as good fit with respect to AIC by instead increasing the value of q_R. The difference in AIC using interactions or not were between 2 and 8 for the data series considered in this paper. For simplicity, we therefore omitted interactions, which made the model easier to interpret. In contrast to Stern and Coe [22] we also include the current temperature as a predictor, which results in a substantial improvement of the model. The standard approach of holding the coefficient of variation constant (equation (6)) resulted in a good fit to the precipitation data.

In Möller et al. [14], Sloughter et al. [20] it is suggested to model $R_t^{1/3}$ in stead of R_t with a gamma distribution. Based on goodness of fit analyzes we were not able to show that one of these alternatives resulted in a better fit then the other and decided to model R_t as gamma distributed as shown above.

Similar to the snow depth model, we model $P(R_t = 0)$ with logistic regression. For the snow depth data using $E(D_t \,|\, D_t > 0)$ as the only covariate performed well, but using only $E(R_t \,|\, R_t > 0)$ as a covariate turns out to perform poorly for the precipitation data. A better fit is achieved using the same covariates as above

$$P(R_t = 0) = \text{logit}^{-1}\left(h_{m_{R_0}}(t) + \sum_{j=1}^{q_{R_0}} \gamma_{R_0 j} R01_{t-j} + \sum_{j=1}^{s_{R_0}} \kappa_{R_0 j} T^j\right) \tag{14}$$

Also for this model, we used a steepest descent optimization algorithm to find the parameters that maximized the likelihood function.

Given forecasts of temperature and precipitation using the models above, the model in Section 2 can further be used to forecast snow depth.

3.2 *Model 2*

While model 1 in the previous section perform long term forecasts of snow depth by first doing long term forecasts of temperature and precipitation, in this section we instead model the snow depth time series directly. The model will be exactly as the model in Section 2 except that the covariates must be changed since precipitation and temperature is unknown. Therefore we change (5) with

$$E(D_t \,|\, D_t > 0) = \exp\left(h_{m_D}(t) + \sum_{j=1}^{q_D} \gamma_{Dj} D01_{t-j} + \sum_{j=1}^{s_D} \eta_{Dj} D_{t-j}\right) \tag{15}$$

were $D01_t$ is defined such that $D01_t = 1$ if $D_t > 0$ and $D01_t = 0$ if $D_t = 0$. It turned out that using both the covariates $D01_{t-j}$ and D_{t-j} resulted in a better fit than using only $D01_{t-j}$ or only D_{t-j}.

3.3 *Monte Carlo procedure*

For the first days in to the future when reliable weather forecast of temperature and precipitation is available, the Monte Carlo method described in Section 2 will be used. Let δ denote the number of days with reliable weather forecasts. Now suppose that we want to forecast $D_{t+\delta+1}$. After running the Monte Carlo method in Section 2 we have a set of realizations of the time series $D_{t+1}, \ldots, D_{t+\delta}$. Long-term forecasts can then be computed as follows

- Using model 2, forecasting of $D_{t+\delta+1}$ can be achieved by generating a realization from model 2 conditioned on each of the realizations of the time series $D_{t+1}, \ldots, D_{t+\delta}$.
- Using model 1, first a large set of realizations of $T_{t+\delta+1}$ and $R_{t+\delta+1}$ is generated conditioned on the weather forecasts $T_{t+1}, \ldots, T_{t+\delta}$ and $R_{t+1}, \ldots, R_{t+\delta}$. For each realization of $T_{t+\delta+1}$ and $R_{t+\delta+1}$, a realization of $D_{t+\delta+1}$ is generated using the model in Section 2.

The procedures above can be repeated for as long into the future that forecasts of snow depth is needed.

4. Model evaluation

In this section we describe approaches to evaluate the performance of the models in the previous Section.

4.1 *Goodness of fit*

Assume that X is a stochastic variable with a cumulative distribution function $F_X(x)$. It's a well-known fact that $F_X(X) \sim U[0,1]$ where $U[0,1]$ denote a uniform distribution on the $[0,1]$ interval. The procedure can be used for the different models presented above. For the model in Section 2 the cumulative distribution for D_t can be computed as

$$\begin{aligned} F_{D_t}(d) =& P(D_t \leq d) = \text{logit}^{-1}(\beta_6 + \beta_7 E(D_t \mid D_t > 0)) + \\ &+ \text{I}(d > 0)(1 - \text{logit}^{-1}(\beta_6 + \beta_7 E(D_t \mid D_t > 0))) \int_0^d g(D) dD \end{aligned} \tag{16}$$

where $\int_0^d g(D)dD$ is the cumulative gamma distribution. Let $d_1, d_2, \ldots, d_n$ denote the real observations of snow depth. If the model fits the data well, we expect the distribution of $F_{D_1}(d_1), F_{D_2}(d_2), \ldots, F_{D_n}(d_n)$ to be close to uniformly distributed. In the computation of $E(D_t \mid D_t > 0)$, the real observations of temperature, precipitation and snow depth from the previous day are used as input according to (5). The same procedure will be used also for the other models above.

4.2 *Comparison with a numerical snow depth model*

To evaluate the performance of the models in Sections 2 and 3, it is natural to compare to other snow depth forecast models. We've only found one model doing forecasts in Norway. The model is developed by The Norwegian Water Resources and Energy Directorate (NVE) in partnership with Norwegian Meteorological Institute and The Norwegian Mapping Authority and is the engine behind the web site `senorge.no` (See Norway). In the rest of the paper we denote the model the NVE model. For more details on the

Table 1. Properties of meteorological data.

Place	Altitude (m)	Timespan with observations	Vegetation zone
Oslo (Blindern)	97	June 11 1955 – June 11 2015	City/boreal forest
Drevsjø	678	January 1 1957 – October 20 2016	Boreal forest (taiga)
Geilo	810	September 1 1966 – November 30 2006	Alpine tundra
Værnes	14	January 1 1950 – October 20 2016	Marine
Tromsø	100	January 1 1955 – September 1 2015	Arctic marine
Kautokeino	345	January 1 1950 – December 31 1969 January 1 1997 – October 18 2016	Arctic tundra

Table 2. Estimated parameters of the model in Section 2.

Place	μ	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	σ_1^2	σ_2^2
Oslo	−6.92	0.96	0.88	−1.76	1.99	−0.30	−0.03	4.13	−1.97	0.63	1.79
Drevsjø	−5.87	0.81	0.74	−1.01	2.73	−0.11	−0.05	3.97	−1.16	1.28	3.24
Geilo	−5.88	0.72	1.74	−1.19	2.86	−0.25	−0.05	3.61	−0.75	1.04	2.83
Værnes	−5.23	2.29	−1.16	−0.38	2.11	−0.08	−0.10	3.02	−1.09	1.33	2.38
Tromsø	−4.23	0.89	2.02	−0.83	2.64	−0.16	−0.04	3.38	−0.64	0.98	8.56
Kautokeino	−4.98	1.11	0.78	−1.50	2.00	−0.36	−0.02	3.49	−1.34	1.24	1.06

model, please see [17]. The model is fairly advanced taking into consideration both solar radiation and vegetation type.

We compare the forecasts from our models with the NVE model. The NVE model is constructed to run for a whole winter season based on daily inputs of temperature and precipitation. Suppose that at time t, the observed snow depth is D_t and that the computed snow depth from the NVE model at times t and $t+\delta$ are D_t^* and $D_{t+\delta}^*$, respectively. We then forecast the snow depth at time $t+\delta$ given D_t by bias correcting the numerical NVE simulations, i.e. we forecast the snow depth at time $t+\delta$ with the value $\max\{D_t - D_t^* + D_{t+\delta}^*, 0\}$

5. Real data example

In this section, we forecast future snow depths using the models introduced in the previous sections. We downloaded average daily temperature, precipitation and snow depth from six locations in Norway from the web portal `eklima.met.no`. Properties of the six locations are shown in Table 1. The locations represent the main climate and vegetation zones in Norway. We have not included any locations from the southwestern part of Norway since it usually is very little snow there.

We now fit the model in Section 2 for each of the six locations using the time series of temperature, precipitation and snow depth. Table 2 shows the estimated parameters for the six locations. For the fitted models Figures 1, 2 and 3 show curves for $E(D_t)$ for different values of T_t, R_t and D_{t-1}. $E(D_t)$ is computed using the law of total expectation

$$\begin{aligned} E(D_t) &= E(D_t \,|\, D_t = 0)P(D_t = 0) + E(D_t \,|D_t > 0)P(D_t > 0) \\ &= E(D_t \,|D_t > 0)P(D_t > 0) \end{aligned} \tag{17}$$

where $E(D_t \,|D_t > 0)$ and $P(D_t > 0)$ is computed using (5) and (9), respectively.

Figure 1 snows expected snow depth from 10 mm precipitation for different temperatures. We set $D_{t-1} = 0$ so that we only look at the snowfall part of the model, recall (5). As expected the snow depth decreases rapidly around 0 °C and the curves are in accordance with earlier research on such curves [10, 23]. Due to differences in climate, the curves varies with location in Norway and such differences is also observed in earlier research [10, 23].

Figure 2 shows $E(D_t)$ for different temperatures when assuming that $R_t = 0$ mm and $D_{t-1} = 30$ cm. Since R_t is set to zero, this figure shows the melting/aging part of the model. Figure 3 shows $E(D_t)$ for different amounts of precipitation when $T_t = 5\,^\circ$C and

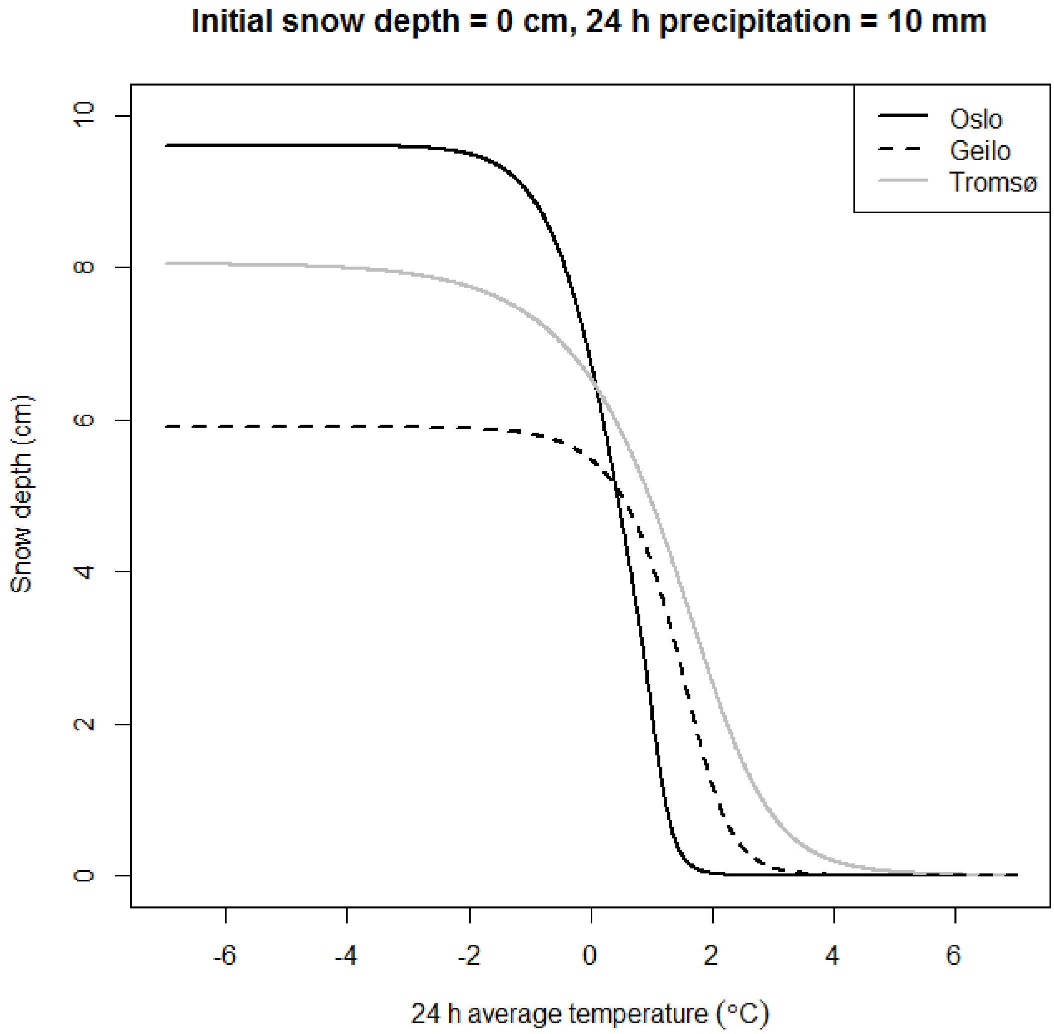

Figure 1. Relation between $E(D_t)$ from 10 mm precipitation for different temperatures.

$D_{t-1} = 30$ cm. This figure shows the effect of precipitation on the reduction of snow depth. We see that rain has a strong impact on the reduction of snow depth. We also observe that the curves seem to be quite linear which is in accordance with the index model [18]. The Figures $1-3$ show that the model captures the main effects of snow accumulation, aging and melting. It is also interesting to see how well the observations agreed with the curves shown in Figures 1 to 3. The results are shown in Figure 4. The black curves in the panels in the first column show $E(D_t)$ from Figure 1 for the locations Oslo, Geilo and Blindern. Further we selected observations satisfing $D_{t-1} = 0$ and 5 mm $< R_t <$ 15 mm and plotted $(R_t/10)D_t$ (we rescale to 10 mm precipitation) against T_t (the circles). Finally the gray curves show the average of the observations for some temperature intervals. The second column shows the same with respect to Figure 2 with the exception that D_{t-1} was set equal to 10 cm for Oslo to be able to compare with more observations. We selected observations satisfying 25 cm $< D_{t-1} <$ 35 cm (8 cm $< D_{t-1} <$ 12 cm for Oslo) and R_t = 0 mm and plotted $D_t - D_{t-1} + 30$ cm ($D_t - D_{t-1} + 10$ cm for Oslo) against T_t. The third column show the same with respect to Figure 3 with the expectation that we set D_{t-1} equal to 10 cm to be able to compare with more observations. We selected observations satisfying 7 cm $< D_{t-1} <$ 13 cm and

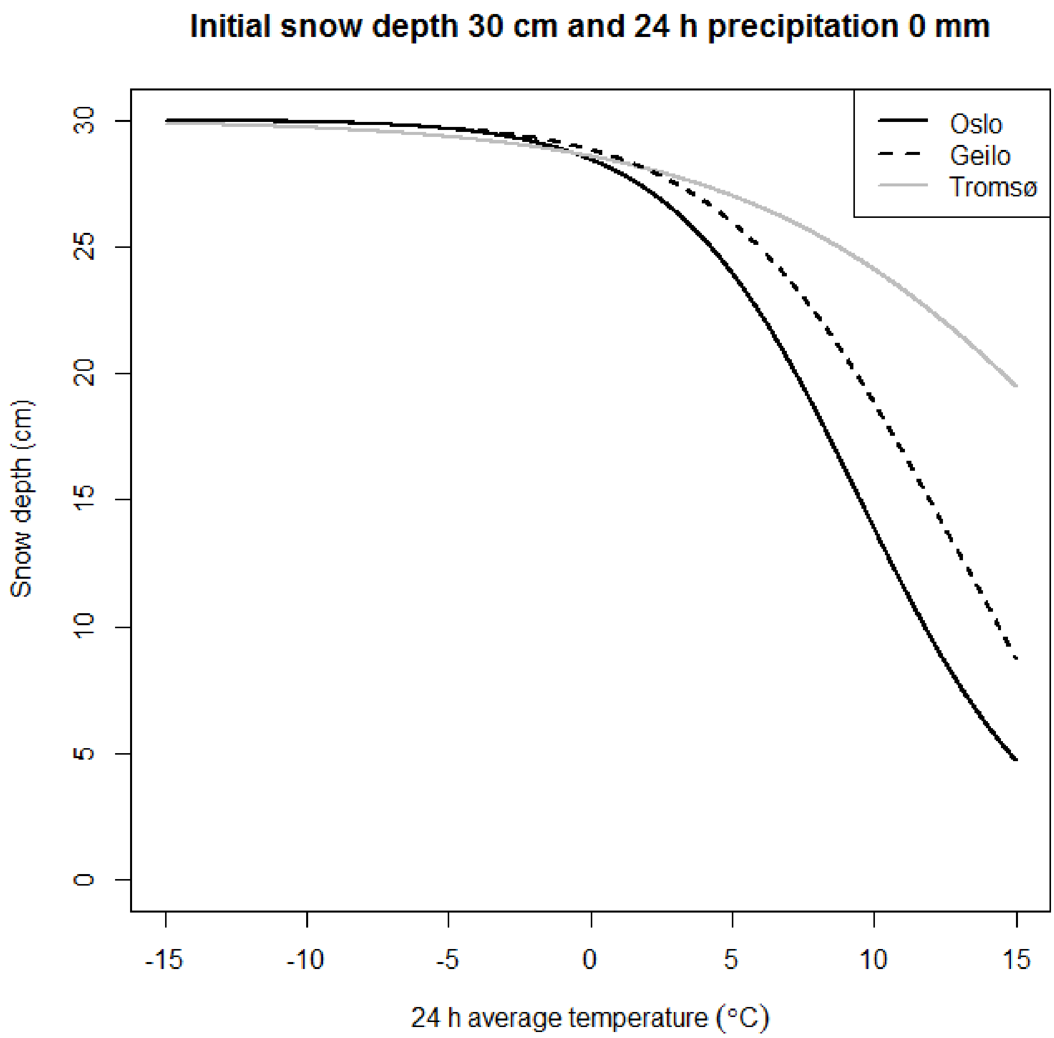

Figure 2. Relation between $E(D_t)$ and temperature when $D_{t-1} = 30$ cm and $R_t = 0$ mm, i.e. the melting/aging process.

$4° \text{ C} < T_t < 6° \text{ C}$ and plotted $D_t - D_{t-1} + 10$ mm against R_t.

For the panels in the first column it seem to be a good agreement between the average curves (gray) and the model expectations (black). Further we observe a large variation in snow depth for a given temperature.

For the panels in the second column, again we see a good agreement between the average curves and the model expectations, but the expectation curve seem to be a little above the observations for Tromsø.

For the third panel, comparison between the expectation curves and the observations was challenging since it was only a few samples satisfying a snow depth around 10 cm and rainfall with a temperature around 5 ° C.

Figure 5 shows the snow depth observations and residuals $(D_t - E(D_t))$ for the winter months December, January and February. We see that the residuals are almost symmetrically distributed around zero. For most of the observations, the residuals are close to zero.

To perform long therm forecasts of snow depth, we fitted the time series of temperature, precipitation and snow depth using the models in Section 3. In each of these models

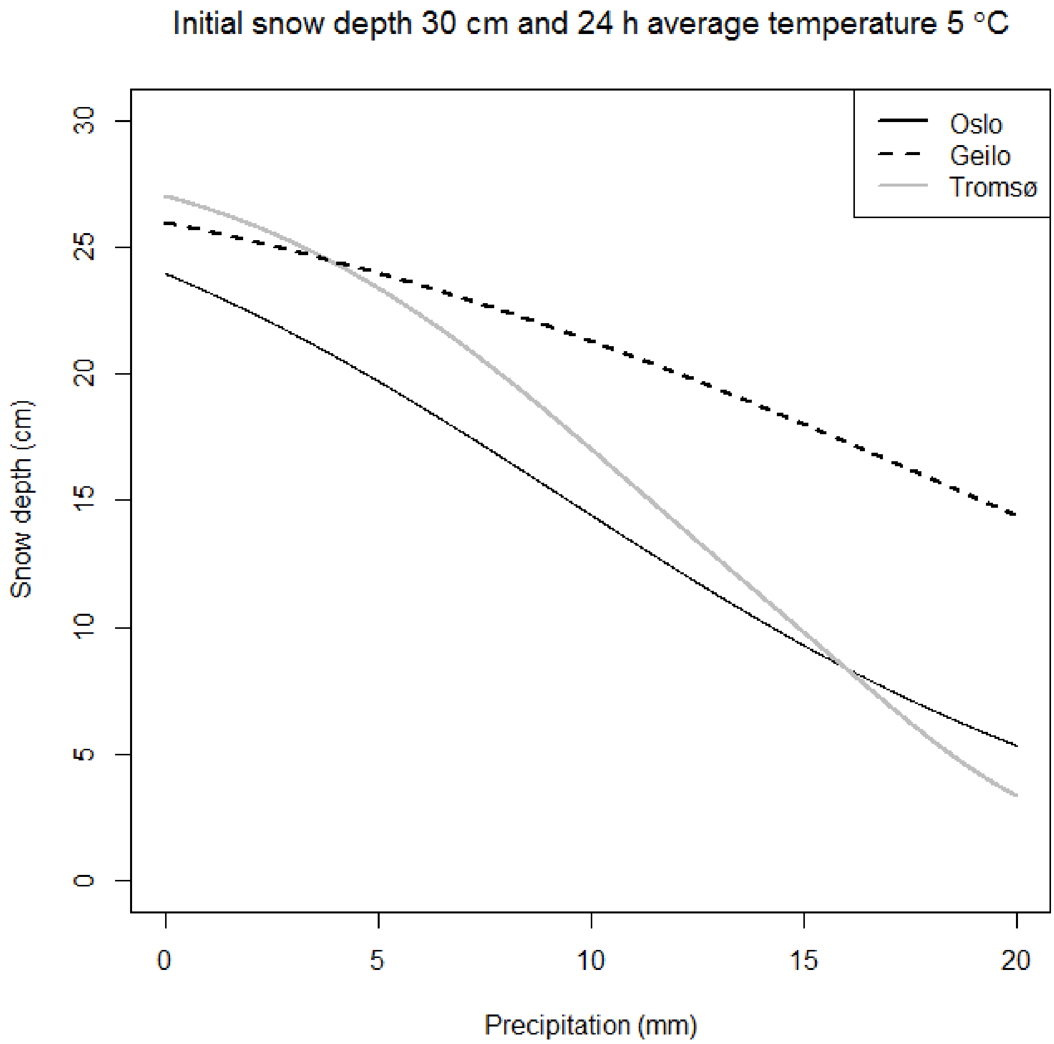

Figure 3. Relation between $E(D_t)$ and precipitation when $D_{t-1} = 30$ cm and $T_t = 5\,°\text{C}$, i.e. the melting/aging process.

Table 3. Number of parameter included in the precipitation, temperature and snow depth models.

	Temperature		Precipitation						Snow depth		
Place	m_T	p_T	m_R	q_R	s_R	m_{R_0}	q_{R_0}	s_{R_0}	m_D	q_D	s_D
Oslo	2	3	3	5	4	3	5	4	3	1	5
Geilo	2	3	2	3	3	3	5	3	3	1	3
Tromsø	2	4	3	6	3	2	6	3	3	1	6
Drevsjø	4	3	2	3	2	2	6	4	3	1	6
Værnes	2	3	3	6	3	3	5	4	3	1	2
Kautokeino	2	4	3	5	4	3	5	4	3	1	1

the number of covariates where chosen based on AIC in a forward stepwise regression procedure. E.g. for temperature, first m_T is increased to one, then p_T to one, then m_T to two and so on. The resulting number of parameters are shown in Table 3.

5.1 *Goodness of fit*

Figure 6 shows goodness of fit histograms of the models in Sections 2 and 3 using the approach described in Section 4.1. The upper row in Figure 6 shows from left to right

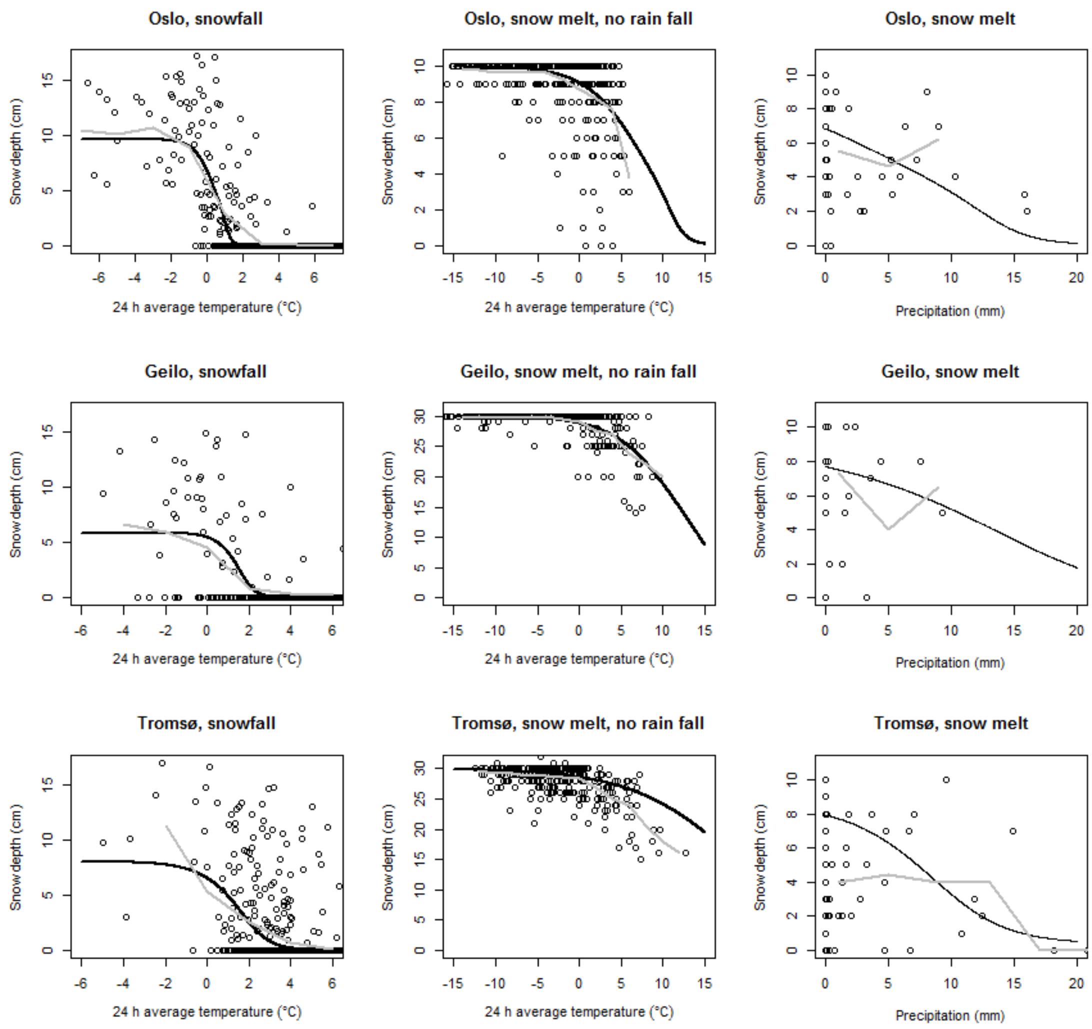

Figure 4. The figure show agreement between model expectations curves and snow depth observations.

goodness of fit histograms for temperature and precipitation, respectively, while the bottom row shows from left to right goodness of fit histograms for the model in Section 2 and model 2, respectively. The goodness of fit histograms for temperature and precipitation are based on observation for the whole year while for the histograms for the snow depth models are based on the winter months December, January and February. All the histograms are for Oslo, but the histograms for the other locations were similar. We see that the histograms for temperature and precipitation looks fairly uniformly distributed. The two models for snow depth (bottom row) show an overrepresentation of values around 0.5. When the temperature is below 0°C and it is no precipitation, the changes in snow depth is minimal and the model overestimates the variance in this cases resulting in too many values around 0.5 in the goodness of fit histograms. During winter, these weather conditions are of course very common. We can of course reduce the value of σ_1^2 in (7) to reduce the variance when the expected change in snow depth is small, but that will result in other negative consequences for the model.

To summarize the goodness of fit for the other locations, we use the reliability index measure, see e.g. Section 3 in Delle Monache et al. [4]. For a goodness of fit histogram, like the ones in Figure 6, we compute the reliability index as follows. Suppose that we have n values and we want measure the deviation from a perfect uniform distribution.

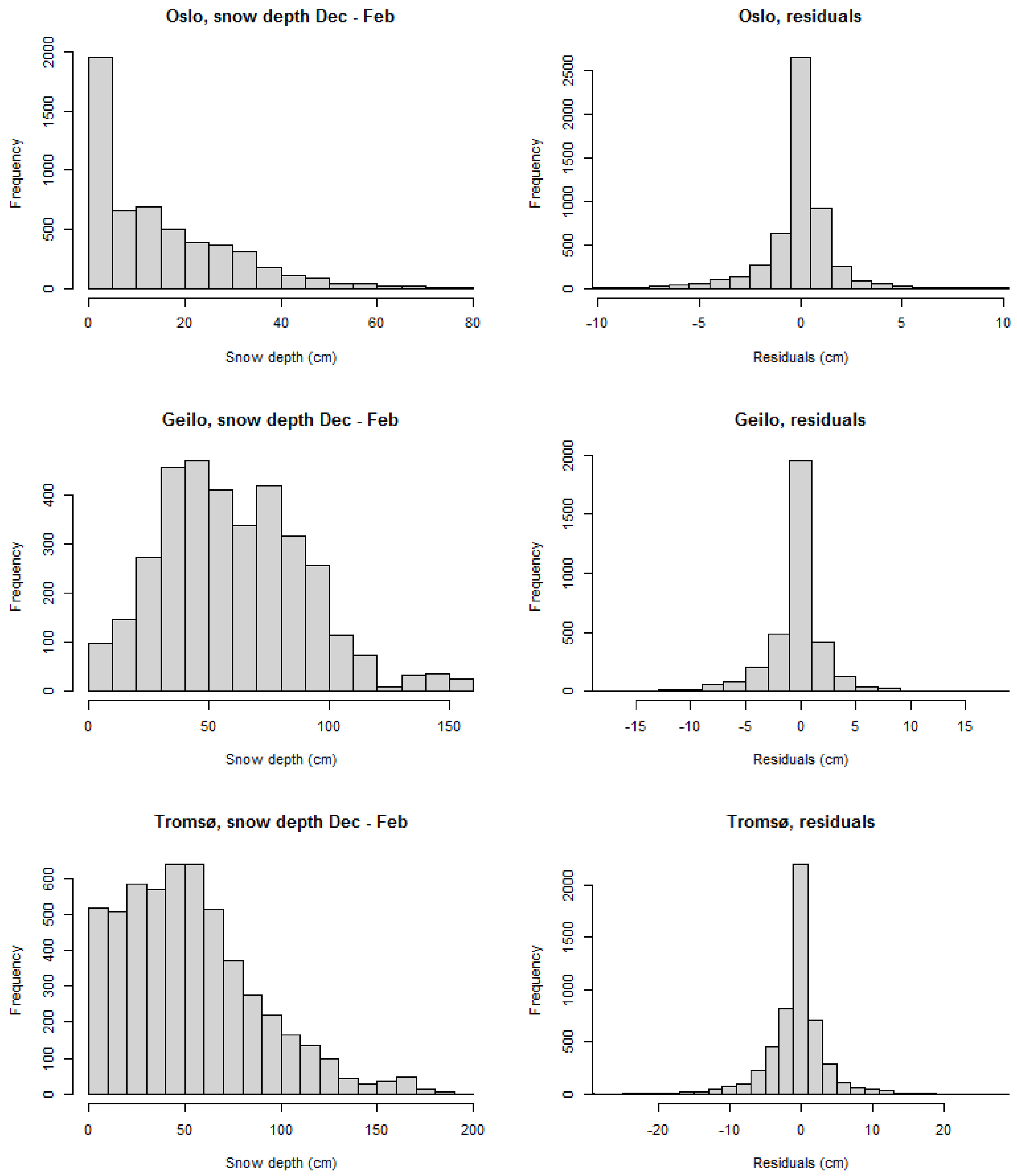

Figure 5. The panels in the left column show histograms of the snow depth observations and the right column the residuals $D_t - E(D_t)$.

Location	Snow depth	Temperature	Precipitation
Oslo	2.31	1.08	1.03
Geilo	2.61	1.52	0.64
Tromsø	3.52	0.69	1.01
Drevsjø	3.85	2.05	1.06
Værnes	2.69	1.14	0.49
Kautokeino	4.72	2.56	2.61

Table 4. Relibability index for the prediction models for each of the six different locations in Norway.

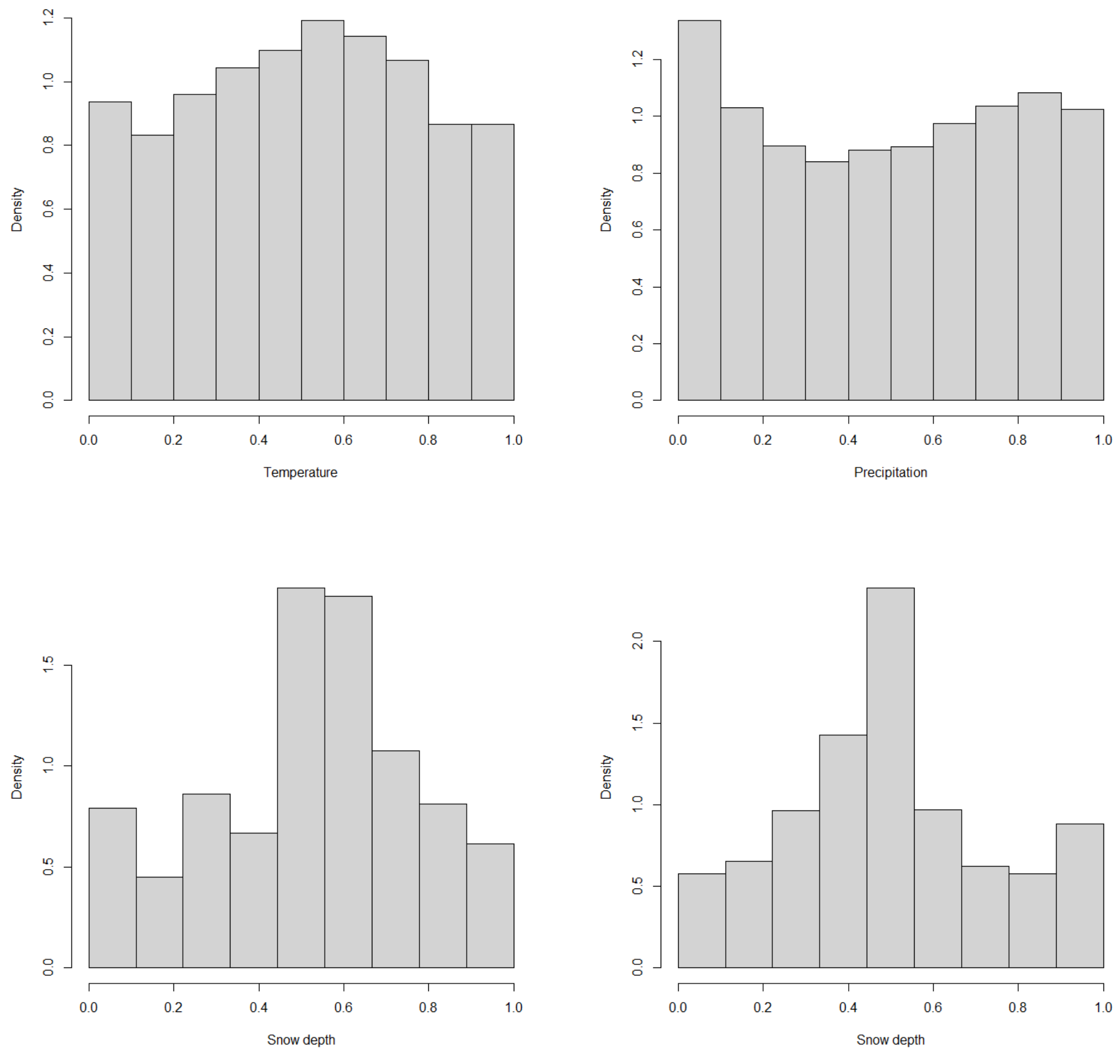

Figure 6. Goodness of fit histograms for Oslo. Upper row from left to right shows temperature and precipitation, respectively. The bottom row from left to right show goodness of fit histograms for the model in Section 2 and model 2, respectively.

We divide the $[0, 1]$ interval in B equally wide bins. Let n_b denote the number of the n values with a value in bin $b \in 1, 2, \ldots, B$. The reliability index is then computed as follows

$$\text{Reliability index } = \frac{1}{B}\sum_{b=1}^{B}\left|\frac{n_b}{n} - \frac{n}{B}\right| \times 100$$

The reliability index for the different prediction models for each of the six locations in Norway are shown in Table 4. The reliability indexes are computed using $B = 10$ bins. We see that the reliability index is low and fairly similar for all the models and for all the geographical locations.

5.2 *Forecast performance*

We now inspect the forecast performance of the snow depth models. Forecasts are performed using the Monte Carlo procedures described in Sections 2 and 3. In each time step we forecast using the mean value of the Monte Carlo samples. Forecast error is measured by the average difference in absolute value between the observed and forecasted snow depth. Forecasting is performed for the winter months December, January and February in a cross validation procedure. All data except for one year from July 1 to June 30 the next year is used to fit the models, and further used to forecast snow depths for December, January and February for the year of data not included in the model fitting. The procedure is repeated for each year. We consider three different cases where we assume that reliable weather forecasts of temperature and precipitation are available for zero, five and ten days into the future. For the days were we assume that reliable weather forecasts are available, the real observed values of temperature and precipitation are used, i.e. we assume “perfect” weather forecasts. For comparison we also consider a version of model 1 and 2, where all covariates except the periodic covariates are set to zero, i.e. all the variables p_T, q_R, s_R, q_{R_0}, s_{R_0}, q_D, s_D are set to zero which means that only m_T, m_R, m_{R_0} and m_D can be larger than zero. In addition, we assume that the number of days with reliable weather forecasts of precipitation and temperature are zero. In other words, these versions of model 1 and 2 do not take advantage of previous observations for the given season or weather forecasts and forecast only based on seasonal properties. We expect that the further we forecast into the future, the less will the usefulness of previous observations for the given season and weather forecasts be.

The results for Oslo, Geilo and Tromsø are shown in Figures 7 − 9. The gray and black curves show results for model 1 and model 2, respectively. The dashed, dotted and dash-dotted curves show forecasts when we assume that reliable weather forecasts of temperature and precipitation are available for zero, five and ten days, respectively. For the days when reliable weather forecasts are available, the model in Section 2 is used. The solid curves show forecast performance where only the periodic covariates are included as explained above. Finally, the solid line with circles show forecast error using the NVE model. The left vertical axis show average forecast error in absolute value, while the right vertical axis show the forecasting error normalized with the average snow depth for the months of December, January and February.

We see that model 2 perform better then model 1. Further we see that given reliable weather forecasts the forecast error is about half compared to not having reliable weather forecasts. We also observe that useful forecasts is possible long in to the future. Forecasting three weeks into the future and given five days of reliable weather forecasts the forecast error is a little over half of the forecast error using only the periodic covariates. In comparison, weather forecasts of temperature and precipitation are typically completely dominated by the seasonal trends after only a few days. Forecast error is lower for Oslo compared to Geilo and Tromsø , but relative to the average snow depth, forecast error for Oslo is higher than for Geilo and Tromsø. Finally we see that the dash dotted curves are below the curves for the NVE model meaning that the model in Section 2 forecasts with less error then the NVE model.

Figures 10 and 11 show forecasts for future snow depths in Oslo for a season with little and much snow, respectively. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution. For the days with reliable forecasts of temperature and precipitation, we forecast using the model in Section 2 and else we use model 2 in Section 3 since it performed better than model 1. Figure 12 −

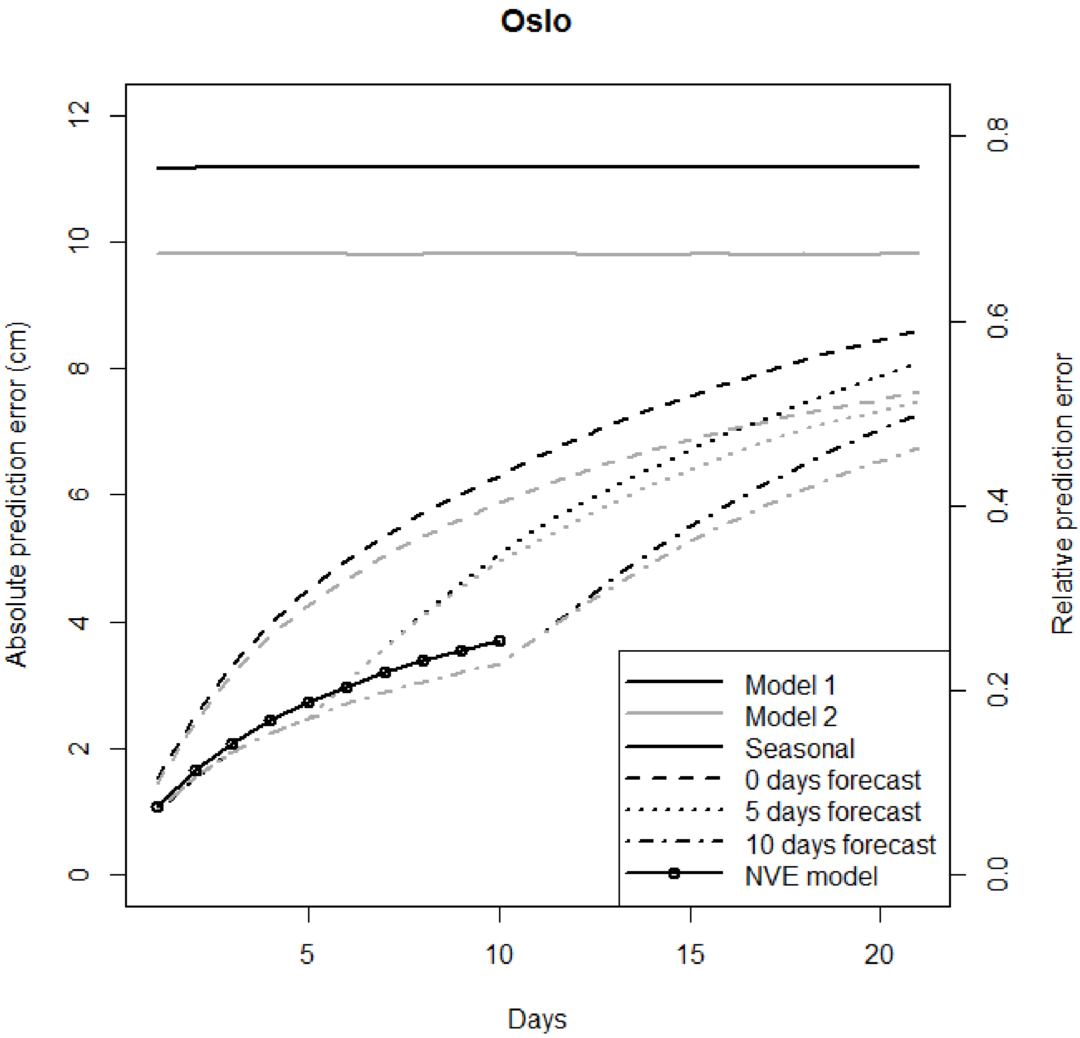

Figure 7. Forecast performance of snow depth for Oslo. The gray and black curves show results for model 1 and model 2, respectively. The dashed, dotted and dash-dotted curves show forecasts when we assume that reliable forecasts of temperature and precipitation are available for zero, five and ten days, respectively. The solid curves show forecast performance where only the periodic covariates are included. The left vertical axis show average forecast error in absolute value, while the right vertical axis show the forecast error normalized with the average snow depth. Finally, the solid line with circles show forecast error using the NVE model.

15 show the same for Geilo og Tromsø. We see that given reliable weather forecasts more precise snow depth forecasts can be achieved. The upper right panels show the only cases where we assume reliable weather forecasts for the whole forecast period and the forecasts thus are only based on the model in Section 2. We see that for this case, the model tracks the true snow depth very well. For the other panels we see that since we do not have weather forecasts for the whole forecast period, the tracking of snow depth lags behind, but still makes useful forecasts. E.g. comparing the bottom rows of Figures 10 and 11, 12 and 13 and 14 and 15, we see that the forecasts for three weeks into the future is quite different for seasons with little and much snow.

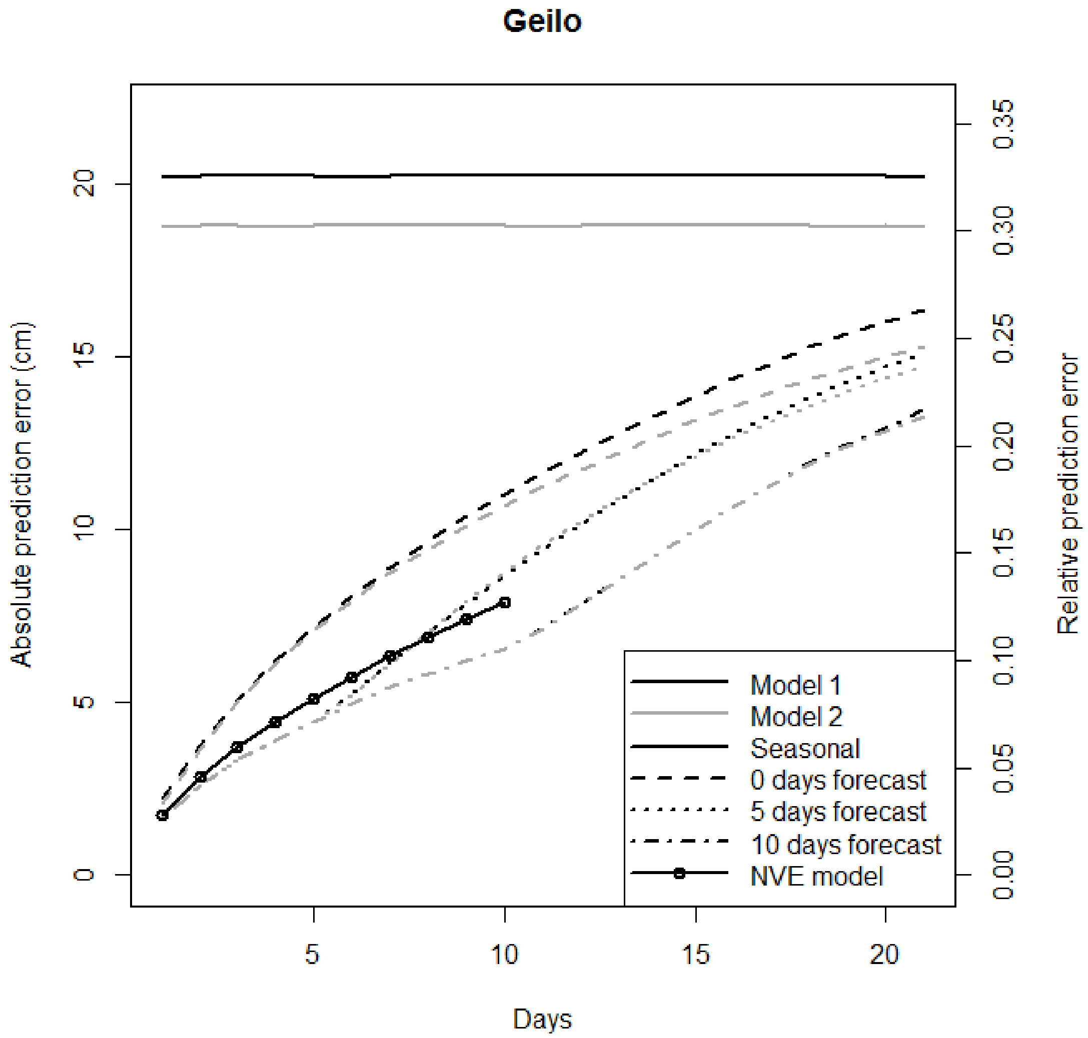

Figure 8. Forecast performance of snow depth for Geilo. The gray and black curves show results for model 1 and model 2, respectively. The dashed, dotted and dash-dotted curves show forecasts when we assume that reliable forecasts of temperature and precipitation are available for zero, five and ten days, respectively. The solid curves show forecast performance where only the periodic covariates are included. The left vertical axis show average forecast error in absolute value, while the right vertical axis show the forecast error normalized with the average snow depth. Finally, the solid line with circles show forecast error using the NVE model.

5.3 *Spatial robustness of model parameters*

The models in this paper have limited usefulness if they only can be used in locations where logged snow depth data are available. In this section we therefore investigate the potential of using parameters from one location to forecast snow depth at other location. The results are summarized in Table 5. As expected, the best predictions will be achieved using the parameters fitted at the given location, but that forecasts using parameters from other locations also do very well. Almost all of the forecasts outperform the NVE model.

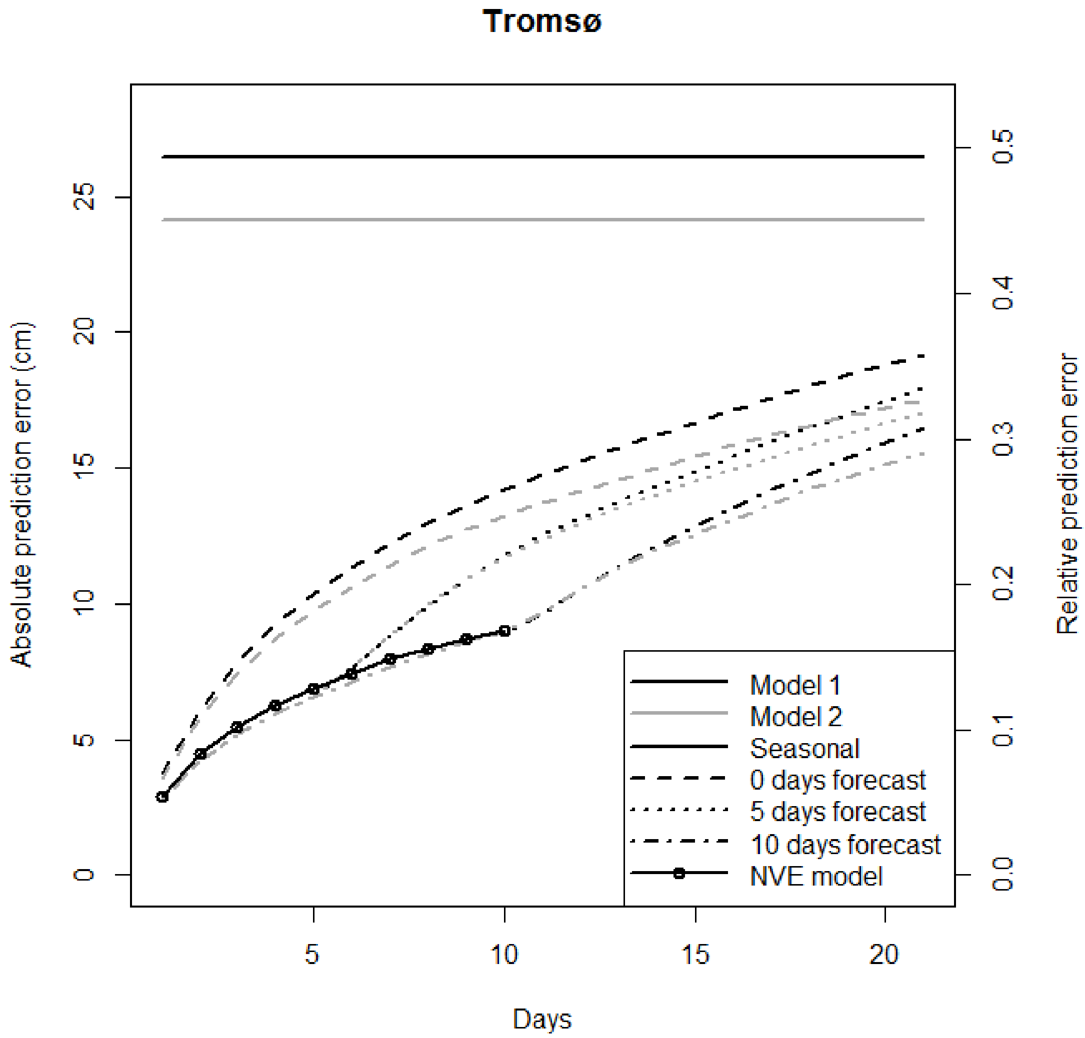

Figure 9. Forecast performance of snow depth for Tromsø. The gray and black curves show results for model 1 and model 2, respectively. The dashed, dotted and dash-dotted curves show forecasts when we assume that reliable forecasts of temperature and precipitation are available for zero, five and ten days, respectively. The solid curves show forecast performance where only the periodic covariates are included. The left vertical axis show average forecast error in absolute value, while the right vertical axis show the forecast error normalized with the average snow depth. Finally, the solid line with circles show forecast error using the NVE model.

Table 5. Forecasting one day into the future when reliable weather forecast of temperature and precipitation are available. The values in the first row show average forecast error (in absolute value) using the parameters estimated from snow depth observation in Oslo to forecast snow depth at other locations. The other rows show the same when the model parameters are estimated from snow depth observations from the other locations. The last row show forecast error using the NVE model.

	Oslo	Drevsjø	Geilo	Værnes	Tromsø	Kautokeino
Oslo	**0.438**	0.738	1.136	0.633	1.606	0.557
Drevsjø	0.476	**0.718**	1.181	0.683	1.633	0.673
Geilo	0.535	0.746	**1.128**	0.757	1.637	0.623
Værnes	0.624	1.046	1.841	**0.725**	2.349	1.042
Tromsø	0.571	0.760	1.188	0.785	**1.623**	0.676
Kautokeino	0.466	0.794	1.214	0.648	1.652	**0.571**
NVE model	1.086	1.295	1.764	1.572	2.906	1.008

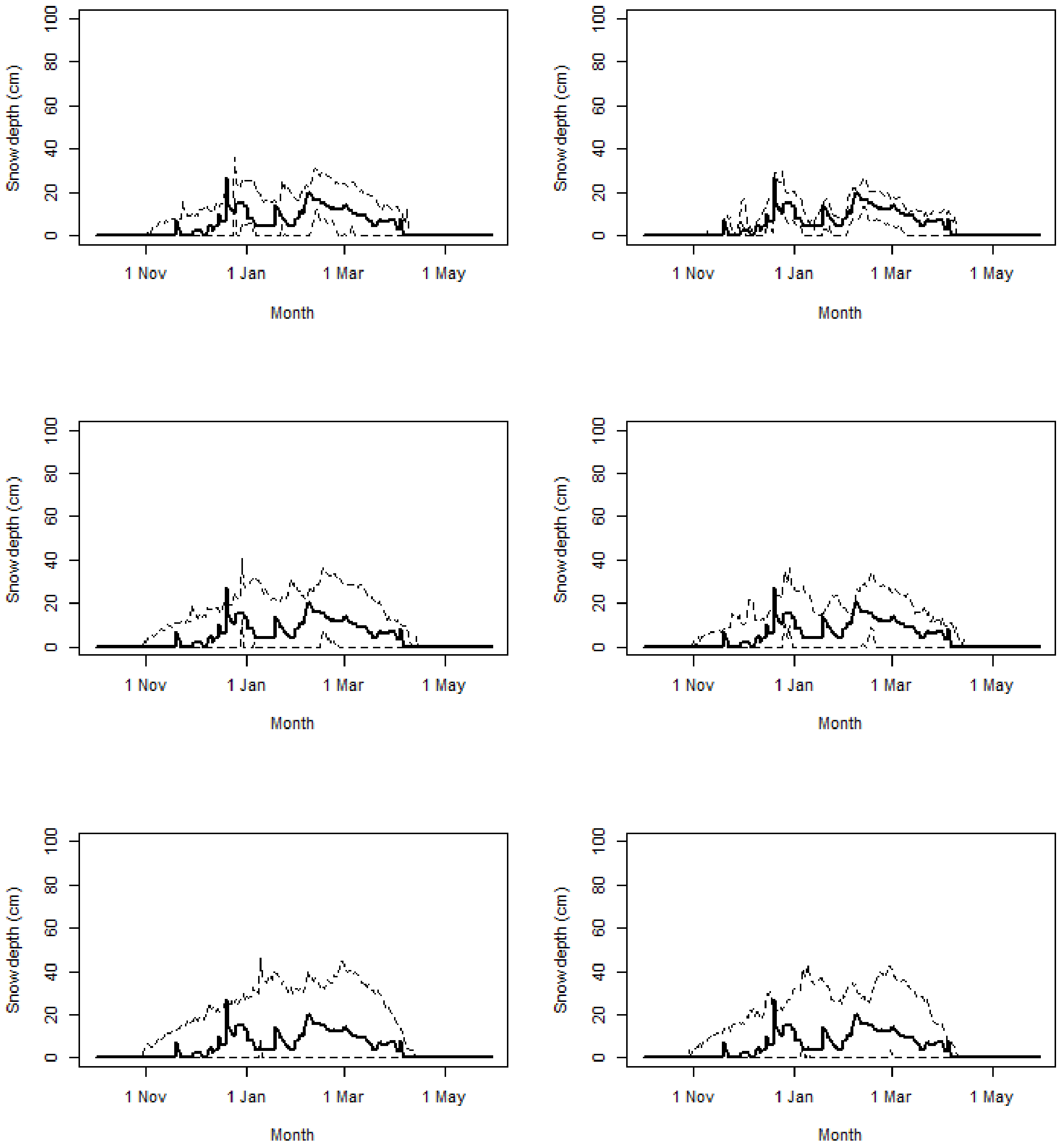

Figure 10. Forecasts of snow depth for Oslo using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

6. Closing remarks

This paper presents a first attempt to build statistical models for short and long term forecasts of snow depth. The results show that it is possible to do useful forecasts of snow depth long into the future. Further we found that model 2 (Section 3.2) perform better then model 1 (Section 3.2), but the advantage of model 1 compared to model 2 is that long term simultaneous scenarios of temperature, precipitation and snow depth is computed. This can be useful in for many applications. E.g. with respect to road safety the risk of slippery roads is especially high when the snow depth is above zero cm and at the same time the temperature is below zero °C. The results show that the model

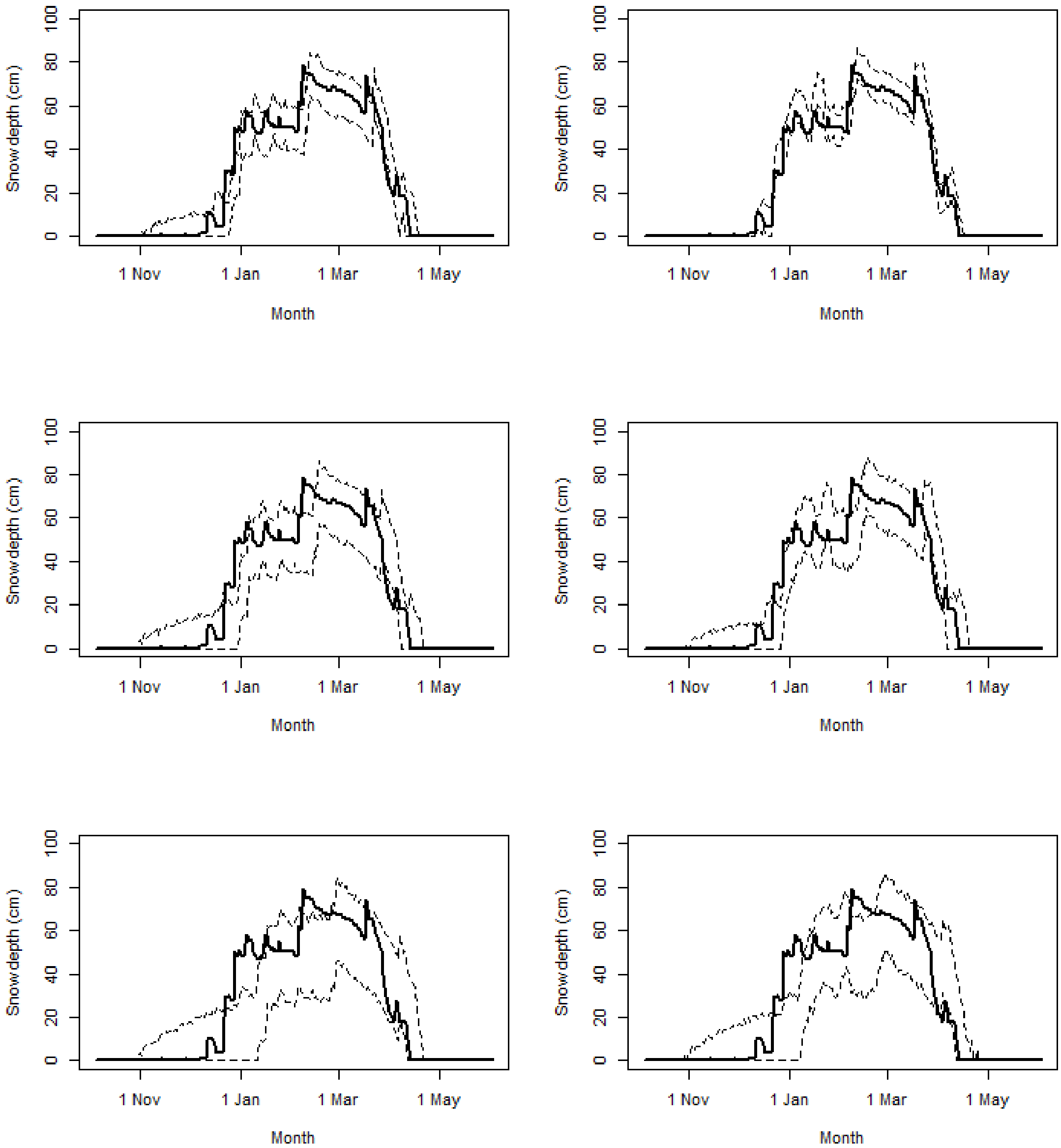

Figure 11. Forecasts of snow depth for Oslo using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

perform very well and better than the NVE mode.

Several extensions to the suggested models are possible. Including other covariates like solar radiation, humidity, wind and the age of the snowpack may improve the forecasts [12, 16]. Models that better separate sinking/aging from melting may be achieved by including the water content in the snow as a hidden layer in the model. Light snow tend to sink faster than denser snow and is not possible to separate in the model presented in this paper. The results in Section 5.3 show that the fitted parameters are robust with respect to location and documents the potential of extending the model to a spatio-temporal model.

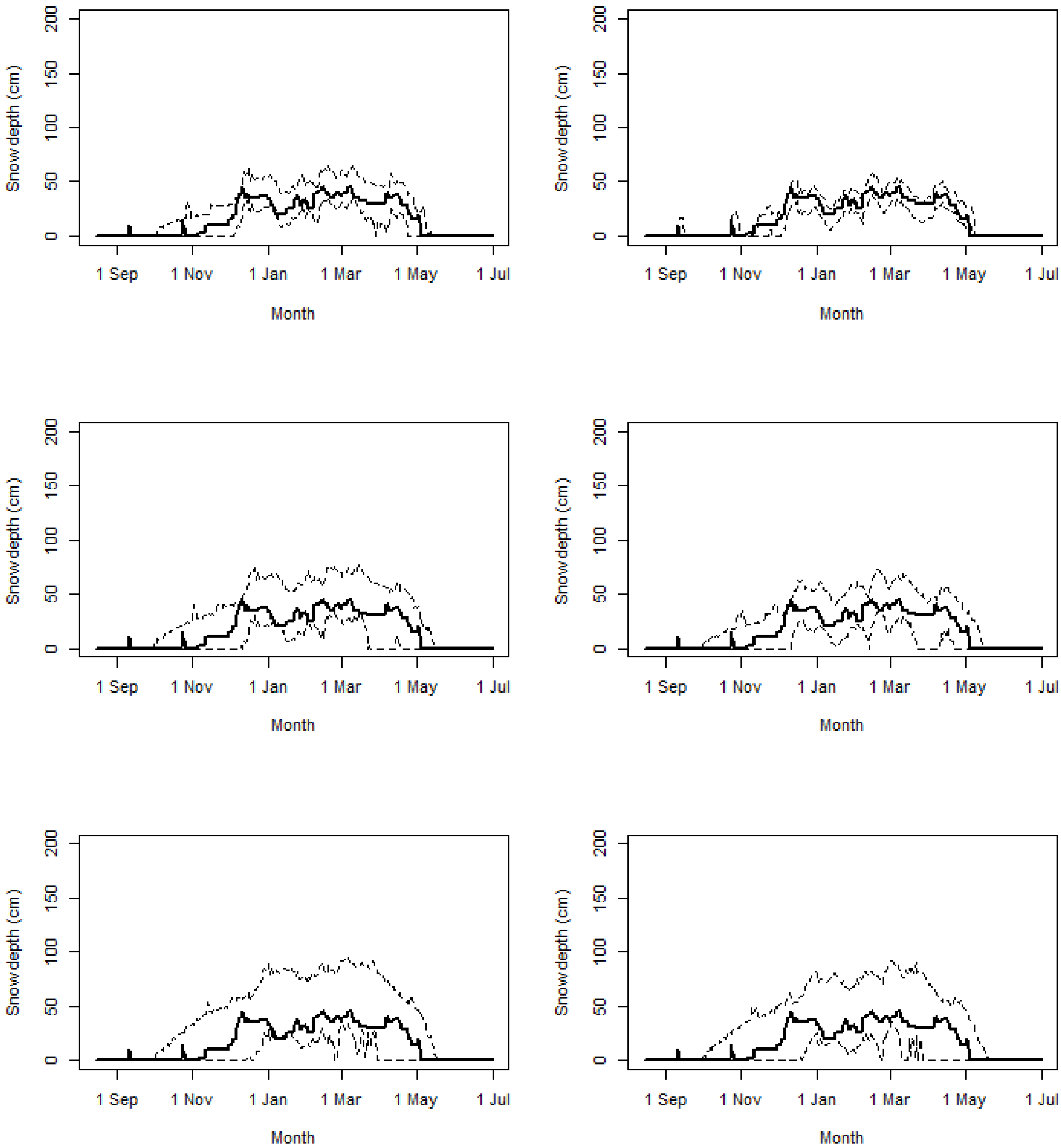

Figure 12. Forecasts of snow depth for Geilo using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

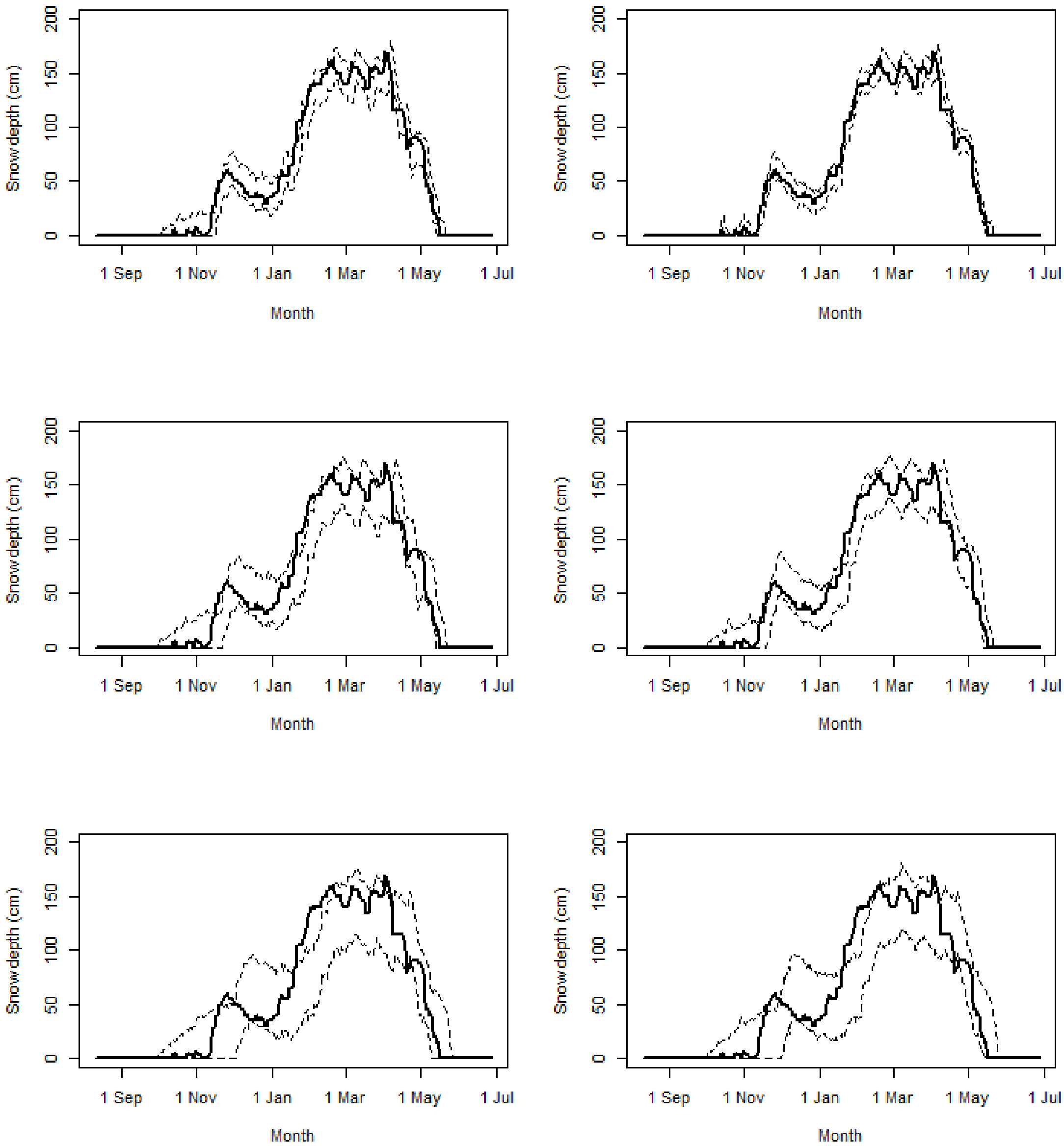

Figure 13. Forecasts of snow depth for Geilo using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

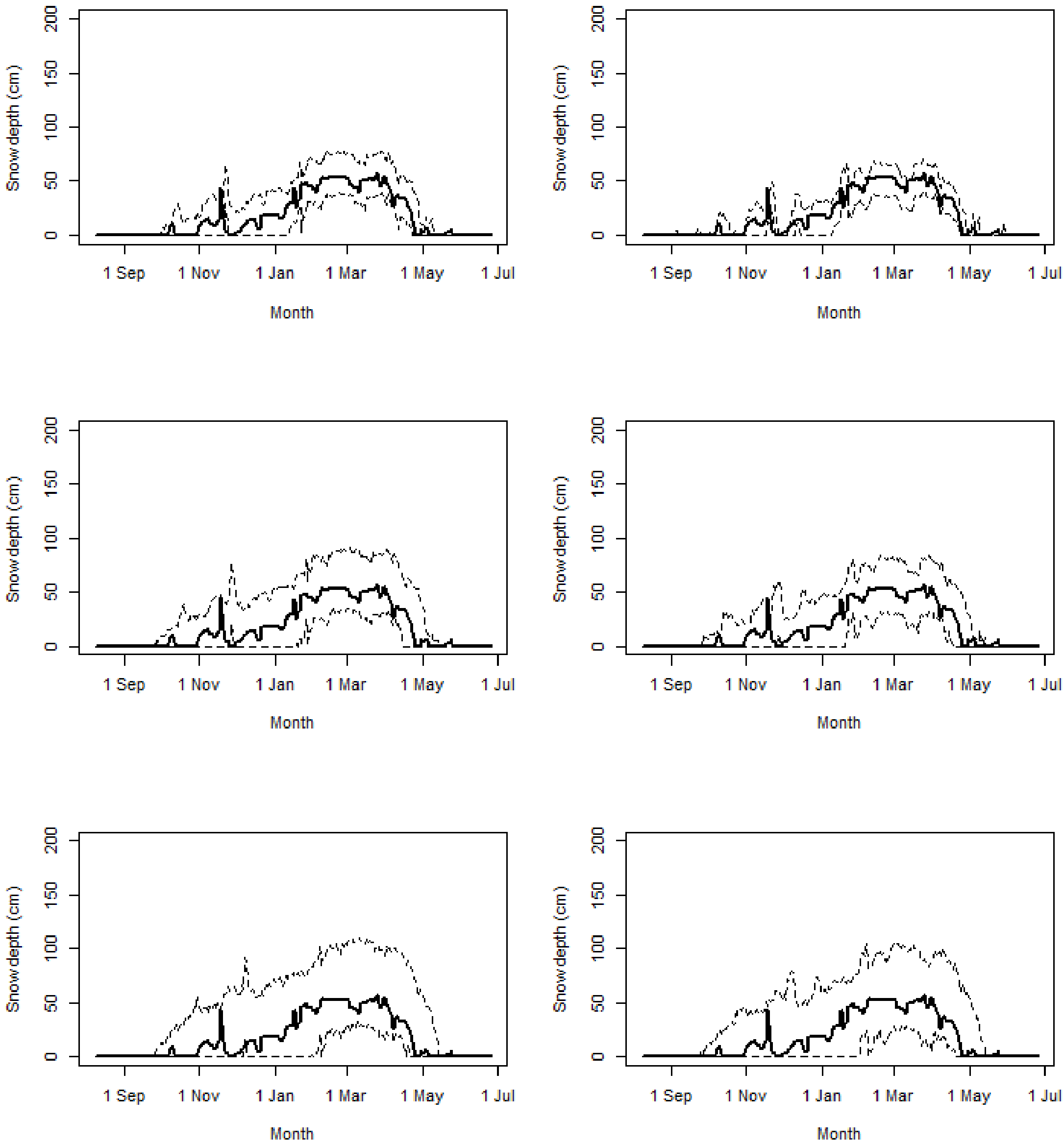

Figure 14. Forecasts of snow depth for Tromsø using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

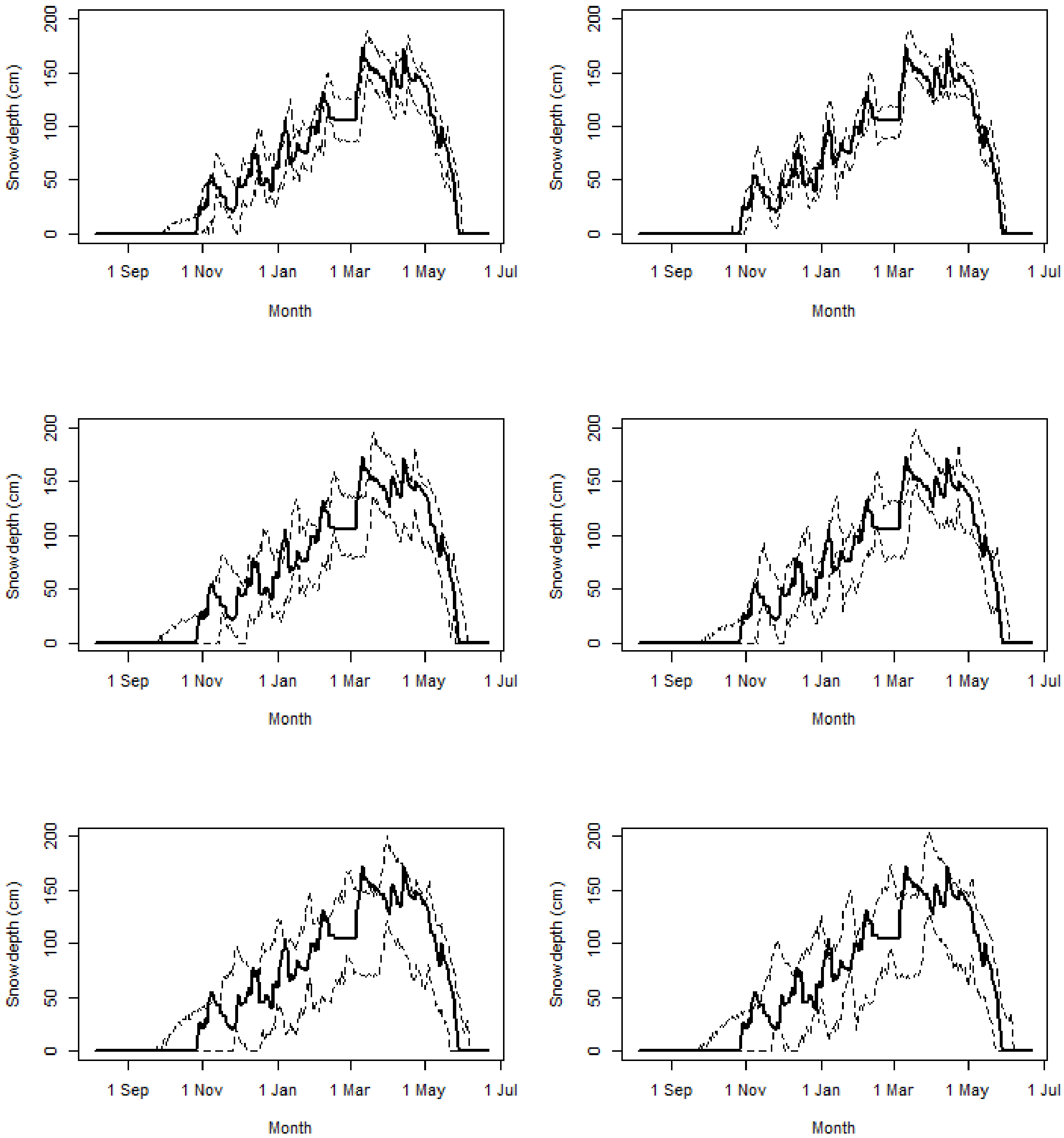

Figure 15. Forecasts of snow depth for Tromsø using model 2. The first, second and third row show forecasts for five, ten and three weeks into the future. In the left and the right column, we assume that zero and five days of reliable weather forecasts of temperature and precipitation are available, respectively. The solid curve shows the real snow depth data, while the dashed curves show 5% and 95% quantiles of the forecast distribution.

References

[1] Blanchet, J. and Davison, A. C. [2011]. Spatial modeling of extreme snow depth, *Annals of Applied Statistics* **5**(3): 1699–1725.

[2] Bocchiola, D., Medagliani, M. and Rosso, R. [2006]. Regional snow depth frequency curves for avalanche hazard mapping in central Italian Alps, *Cold Regions Science and Technology* **46**(3): 204–221.

[3] Brown, R. D., Brasnett, B. and Robinson, D. [2003]. Gridded North American monthly snow depth and snow water equivalent for GCM evaluation, *Atmosphere-Ocean* **41**(1): 1–14.

[4] Delle Monache, L., Hacker, J. P., Zhou, Y., Deng, X. and Stull, R. B. [2006]. Probabilistic aspects of meteorological and ozone regional ensemble forecasts, *Journal of Geophysical Research: Atmospheres* **111**(D24).

[5] Dunn, P. K. and Smyth, G. K. [2015]. *dglm: Double Generalized Linear Models*. R package version 1.8.2.
URL: *http://CRAN.R-project.org/package=dglm*

[6] Falarz, M. [2002]. Long-term variability in reconstructed and observed snow cover over the last 100 winter seasons in Cracow and Zakopane (southern Poland), *Climate Research* **19**(1): 247–256.

[7] *Fnugg* [n.d.]. `fnugg.no`. Accessed: 2016-11-14.

[8] Jonas, T., Marty, C. and Magnusson, J. [2009]. Estimating the snow water equivalent from snow depth measurements in the Swiss Alps, *Journal of Hydrology* **378**(1): 161–167.

[9] Juga, I., Nurmi, P. and Hippi, M. [2013]. Statistical modelling of wintertime road surface friction, *Meteorological Applications* **20**(3): 318–329.
URL: *http://dx.doi.org/10.1002/met.1285*

[10] Kienzle, S. W. [2008]. A new temperature based method to separate rain and snow, *Hydrological Processes* **22**(26): 5067–5085.
URL: *http://dx.doi.org/10.1002/hyp.7131*

[11] Kohler, J., Brandt, O., Johansson, M. and Callaghan, T. [2006]. A long-term Arctic snow depth record from Abisko, northern Sweden, 19132004, *Polar Research* **25**(2): 91–113.

[12] Kuusisto, E. [1984]. Snow accumulation and snowmelt in Finland, *Technical Report 55*, National Board of Waters, Finland, Helsinki.

[13] McCabe, G. J. and Wolock, D. M. [2010]. Long-term variability in Northern Hemisphere snow cover and associations with warmer winters, *Climatic Change* **99**(1): 141–153.

[14] Möller, A., Lenkoski, A. and Thorarinsdottir, T. L. [2013]. Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas, *Quarterly Journal of the Royal Meteorological Society* **139**: 982–991.

[15] R Core Team [2015]. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

[16] Roebber, P. J., Bruening, S. L., Schultz, D. M. and Jr., J. V. C. [2003]. Improving Snowfall Forecasting by Diagnosing Snow Density, *Weather and Forecasting* **18**(1): 264–287.

[17] Saloranta, T. [2012]. Simulating snow maps for norway: description and statistical evaluation of the senorge snow model, *The Cryosphere* **6**(6): 1323–1337.

[18] Scherrer, S. C., Wthrich, C., Croci-Maspoli, M., Weingartner, R. and Appenzeller, C. [2013]. Snow variability in the Swiss Alps 18642009, *International Journal of Climatology* **33**(15): 3162–3173.
URL: *http://dx.doi.org/10.1002/joc.3653*

[19] *See Norway* [n.d.]. `senorge.no`. Accessed: 2016-11-14.

[20] Sloughter, J. M., Raftery, A. E., Gneiting, T. and Fraley, C. [2007]. Probabilistic quantitative precipitation forecasting using Bayesian model averaging, *Monthly Weather Review* **135**(1): 3209–3220.

[21] *Snow Forecast* [n.d.]. `snow-forecast.com`. Accessed: 2015-11-09.

[22] Stern, R. D. and Coe, R. [1984]. A Model Fitting Analysis of Daily Rainfall Data, *Journal of the Royal Statistical Society. Series A (General)* **147**(1): 1–34.
URL: *http://www.jstor.org/stable/2981736*

[23] Wen, L., Nagabhatla, N., Lu, S. and Wang, S.-Y. [2013]. Impact of Rain Snow Threshold Temperature on Snow Depth Simulation in Land Surface and Regional Atmospheric Models, *Advances in Atmospheric Sciences* **30**(5): 1449–1460.

[24] Zhang, Y., Li, T. and Wang, B. [2004]. Decadal Change of the Spring Snow Depth over the Tibetan Plateau: The Associated Circulation and Influence on the East Asian Summer Monsoon, *Journal of Climate* **17**(1): 2780–2793.