Hydrology and Earth System Sciences

A Bayesian spatial assimilation scheme for snow coverage observations in a gridded snow model

S. Kolberg[1]**, H. Rue**[2]**, and L. Gottschalk**[3]

[1]SINTEF Energy Research, Sem Sælands vei 11, 7465 Trondheim, Norway
[2]Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway
[3]Department of Geosciences, University of Oslo. P.O. Box 1047 Blindern, 0316 Oslo, Norway

Abstract. A method for assimilating remotely sensed snow covered area (SCA) into the snow subroutine of a grid distributed precipitation-runoff model (PRM) is presented. The PRM is assumed to simulate the snow state in each grid cell by a snow depletion curve (SDC), which relates that cell's SCA to its snow cover mass balance. The assimilation is based on Bayes' theorem, which requires a joint prior distribution of the SDC variables in all the grid cells. In this paper we propose a spatial model for this prior distribution, and include similarities and dependencies among the grid cells. Used to represent the PRM simulated snow cover state, our joint prior model regards two elevation gradients and a degree-day factor as global variables, rather than describing their effect separately for each cell. This transformation results in smooth normalised surfaces for the two related mass balance variables, supporting a strong inter-cell dependency in their joint prior model. The global features and spatial interdependency in the prior model cause each SCA observation to provide information for many grid cells. The spatial approach similarly facilitates the utilisation of observed discharge.

Assimilation of SCA data using the proposed spatial model is evaluated in a 2400 km² mountainous region in central Norway (61° N, 9° E), based on two Landsat 7 ETM+ images generalized to 1 km² resolution. An image acquired on 11 May, a week before the peak flood, removes 78% of the variance in the remaining snow storage. Even an image from 4 May, less than a week after the melt onset, reduces this variance by 53%. These results are largely improved compared to a cell-by-cell independent assimilation routine previously reported. Including observed discharge in the updating information improves the 4 May results, but has weak

effect on 11 May. Estimated elevation gradients are shown to be sensitive to informational deficits occurring at high altitude, where snowmelt has not started and the snow coverage is close to unity. Caution is therefore required when using early images.

1 Introduction

The utilization of water resources in mountainous regions is often in lack of precise measurements, due to harsh climate and remote situation. This also applies to the estimation of a seasonal snow pack. The potential of remote sensing to collect spatially distributed information of the snow storage is therefore of great interest for water supply or hydropower production planning, as well as flood forecasting. Satellite based measurements of snow water equivalent (SWE) is limited to coarse spatial scales, and with considerable restrictions on snow conditions. At 25 km resolution, passive microwave measurements from SSM/I are able to represent dry snow SWE fairly well in flat, low-lying regions, but show large deviations in mountainous areas (Brubaker et al., 2000). Research is currently progressing on the use of interferometric techniques on active microwave (radar) data, reporting an accuracy of 100 mm for 10×10 km averages, also restricted to dry snow (Engen et al., 2004). Further progress is expected, in particular with radar sensors planned for the near future.

In contrast, the snow covered area (SCA) can be retrieved with good accuracy at a range of scales. Both optical sensors and active microwave instruments have been used to map SCA, the latter with the advantage to measure independently of cloud coverage. Obviously, SWE and SCA are related, and using this relation to extract SWE information

Correspondence to: S. Kolberg
(sjur.kolberg@sintef.no)

from SCA images represents an alternative to direct SWE observation. A simple relationship based on a snow/no snow classification, can be combined with heuristic rules for how to adjust the SWE (Rodell and Houser, 2004), to validate the simulations of a hydrological model (Sheffield et al., 2003), or to support the interpolation of SWE from point measurements (Ranzi el al., 1999; Molotch et al., 2004). By relating observed snow cover patterns to process-governing terrain features, Blöschl et al. (1991) were able to link qualitatively the observation errors to specific process representations in the model, hence evaluating the model structure. Kirnbauer et al. (1994) also concluded that spatial observations of snow cover patterns were clearly able to reveal poor model assumptions, to which the measured streamflow time series was insensitive.

Multispectral techniques have proved effective in mapping SCA at sub-pixel scale, that is, the snow covered fraction of each satellite image pixel. Rosenthal and Dozier (1996) report a regression tree-based estimation of SCA from Landsat TM images, calculating fractional SCA within 30-m grid cells with an overall standard error of around 10%. Salomonsson and Appel (2004) extract fractional SCA from MODIS data at 500 m resolution using the NDSI index, also with a standard error of 10%. With a PRM simulation unit larger than the satellite image pixel size, sub-grid SCA may be determined even from a binary image classification. Thus, in alpine regions with deep, heterogeneous snow packs, fractional SCA can be observed at scales where the SWE-SCA relation is strong. In distributed precipitation-runoff models (PRMs), this relation is commonly represented by a snow depletion curve (SDC).

The SDC concept enables monitoring and updating of the snow cover mass balance with SCA observations made during the melt season. However, the SDC necessarily consists of several variables, whereas the SCA information is a single number. Thus, estimation of the SDC from SCA observations suffers from an information deficit, and the adjustment required to achieve the observed SCA is not uniquely defined. Technically, this problem can be overcome by a Bayesian approach, utilising prior knowledge (Kolberg and Gottschalk, 2005). Binley and Beven (1991) introduced this approach to general PRM parameter estimation, proposing the GLUE technique (General Likelihood Uncertainty Estimation). Relaxing the formal rigour of likelihood construction, GLUE is not a strict application of Bayes' theorem. A principal discussion of such relaxations for hydrological applications appears in Beven and Young (2003) and Gupta et al. (2003). Engeland and Gottschalk (2002) apply Bayes' theorem to the determination of parameters in a regional model, based on streamflow observed in several catchments within the region. Other methods for combined parameter and state estimation by data assimilation have been presented by Moradkhani et al. (2005) and Vrugt et al. (2003, 2005). Huang and Cressie (1996) applied a spatio-temporal model to snow water equivalent, and demonstrated the use

of a Kalman filter to interpolate between SNOTEL measurements.

Despite providing the formal apparatus for re-estimating multiple variables from a single observation, the Bayesian technique does not overcome the large information deficit in the current problem. Accordingly, the improvements of SDC precision using SCA observations reported by Kolberg and Gottschalk (2005) are moderate. The main aim of the current paper is to reduce the gap in information content between each grid cell's multi-variate SDC and single SCA observation. Recognising that SCA observations within a region depend on processes with similar response in many grid cells, we can reduce the effective dimensionality of the estimation problem by handling these processes at aggregated spatial level. Specifically, we first assume that the degree-day factor, as well as elevation gradients in precipitation and temperature are spatially constant, and confine the uncertainty caused by these processes to three discrete variables. Next, by tying two of the gridded variables together in spatial models, we reduce the grid cell uniqueness without formally reducing the number of variables. These transformations all take place in the joint prior distribution of the SDC variables; the assimilation method is identical to the one used by Kolberg and Gottschalk (2005).

2 Approach and method

2.1 The Snow Depletion Curve

The snow depletion curve (SDC) describes how snow covered area (SCA) reduces gradually through the melt season. Rango and Martinec (1997), Liston (1999, 2004), Luce et al, (1999), and Luce and Tarboton (2004) discuss the SDC concept in various formulations. In the current paper, the SDC is applied to each grid cell in a distributed PRM, giving the grid cell's proportional area of bare ground $y=1-$SCA, as a function of its accumulated melt depth. In this formulation, the SDC is also equivalent to the sub-grid cumulative distribution of point snow storage at the start of the melt season, provided that melt rate is sub-grid homogeneous (Liston, 1999). Once established at the end of a specific accumulation season, the SDC is not changed by the PRM; all temporal variability in the bare ground fraction y, the accumulated melt runoff Q, and the remaining snow storage SWE arise from changes in the accumulated melt depth λ, serving as argument to the SDC (Fig. 1).

Luce and Tarboton (2004) analyse nine years' empirical SDC for a small, wind-swept, low-vegetation mountainous catchment, and conclude that normalised SDCs are remarkably stable between years, and that the choice of a parametric model is less important than the estimation of a coefficient of variation. These results are both important for applying a SDC to every grid cell in a larger region, because even an extensive snow survey will only cover a small
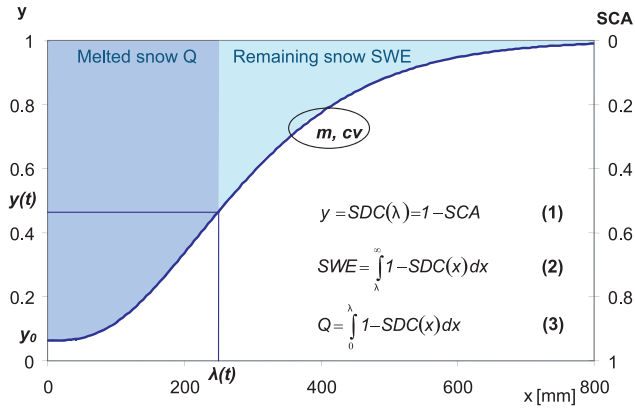
Fig. 1. Principle of the snow depletion curve, as it is applied in each grid cell individually. At the start of the melt season, the SDC represents the spatial heterogeneity of the point snow storage x within the grid cell. It is parameterized by the average storage m, the coefficient of variation cv, and the initial bare ground fraction y_0. During the snow melt season, these parameters are kept constant, and the SDC gives the fractional bare ground y as a function of the accumulated melt depth λ, which is assumed constant over the grid cell. During this period, the mass balance components Q and SWE are also functions of λ, and divide the initial snow pack m into accumulated snowmelt runoff and remaining snow storage, respectively.

fraction of the cells. Without specific information from the actual grid cell, it is necessary to use a parsimonious statistical model for the SDC. Several parametric models have been proposed, for instance 2-parameter Lognormal (Essery et al., 1999; Liston, 2004), 3-parameter Lognormal (Donald et al., 1995), 2-parameter Gamma (Skaugen et al., 2004), 3-parameter Beta (Brubaker and Menoes, 2001), or 3–5-parameter weighted combinations of Normal and/or Lognormal distributions (Bruland et al., 2001; Marchand and Killingtveit, 2004).

In this study, a 3-parameter mixed distribution is selected to describe the SDC in each grid cell. The parameter $y_0 = P(x=0)$, denotes the proportional area which stays snow-free though the whole winter. In the remaining, snow covered cell fraction, the distribution of point snow storage x follows a 2-parameter Gamma model, characterized by the mean value m and the coefficient of variation cv taken at the melt season onset. Using moments rather than the usual shape and scale parameters leads to prior distributions also specified in terms of moments, for which we are more likely to have prior information. Further, the shape and scale parameters, opposite m and cv, are shown to be spatially interdependent (Kolberg, 2001). With this SDC formulation, the bare-ground fraction y is given by the accumulated melt depth λ as:

$$y(t) = \text{SDC}(\lambda(t) \,|\, m, cv, y_0) = y_0 + (1-y_0) \cdot y_1(t) \quad (4)$$

$$y_1(t) = \int_0^{\lambda(t)} p(x; m, cv)\, dx = \gamma\left(\frac{1}{cv^2}, \frac{\lambda(t)}{m \cdot cv^2}\right) \quad (5)$$

Here, y_1 is the bare-ground area relative to the initially snow covered area $(1-y_0)$, $p()$ is the Gamma density and $\gamma(\cdot, \cdot)$ the cumulative Gamma distribution with shape cv^{-2} and scale $m \cdot cv^2$. It is worth noting that the main mass balance variables m and λ appear only as a ratio, and thus that the observable bare-ground fraction y is sensitive only to the relative magnitude of the two.

2.2 Assimilation by Bayesian updating

Bayes' theorem expresses how measured data D change a statistical distribution of a parameter vector $\boldsymbol{\theta}$, which may include any unknown or uncertain variable related to D. The prior knowledge about $\boldsymbol{\theta}$ is expressed as a joint distribution of parameters $p(\boldsymbol{\theta})$, and the relation between the parameters and the observed data is also given in terms of a statistical distribution $p(\boldsymbol{D}|\boldsymbol{\theta})$, expressing the probability (density) of the measured data given the parameters. Considered as a function in $\boldsymbol{\theta}$ with D as parameters, this is called the likelihood function, and its multiplication with the prior $p(\boldsymbol{\theta})$ gives the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{D})$:

$$p(\boldsymbol{\theta}|D) \propto (\boldsymbol{\theta}) \cdot (D|\boldsymbol{\theta}) \quad (6)$$

The proportionality symbol indicates the absence of the normalisation constant, which must be found by integrating the right hand side of Eq. (6). Properly normalised, the posterior distribution contains our knowledge of $\boldsymbol{\theta}$ conditioned on D, and is the basis for estimating elements of $\boldsymbol{\theta}$, typically by computing the expectation, mode, or median.

In the current application, D contains the satellite information, consisting of the y estimate and its associated standard error for each grid cell in the region. The vector $\boldsymbol{\theta}$ contains the variables describing the SDCs of all grid cells. The prior distribution $p(\boldsymbol{\theta})$ has a joint expectation $E[\boldsymbol{\theta}]$, typically consisting of the simulated values in the distributed PRM. Similarly, the expectation $E[\boldsymbol{\theta}|\boldsymbol{D}]$ of the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{D})$, represents updated estimates with which we can re-initialize the PRM. Figure 2 illustrates the procedure for a single grid cell. The prior variance $\text{Var}[\boldsymbol{\theta}]$ and posterior variance $\text{Var}[\boldsymbol{\theta}|\boldsymbol{D}]$ measure the uncertainty before and after updating, respectively, and from these we find the fractional reduction of variance, indicating the informational value of D. The assimilation is presented in an updating context, but also addresses variables normally considered as calibration constants rather than simulated states.

3 The proposed prior distribution

Figures 1 and 2 show how the SDC state, in a specific grid cell and at a given time, is defined by the four variables m, cv, y_0 and λ. For n grid cells and k images acquired through a

www.hydrol-earth-syst-sci.net/10/369/2006/

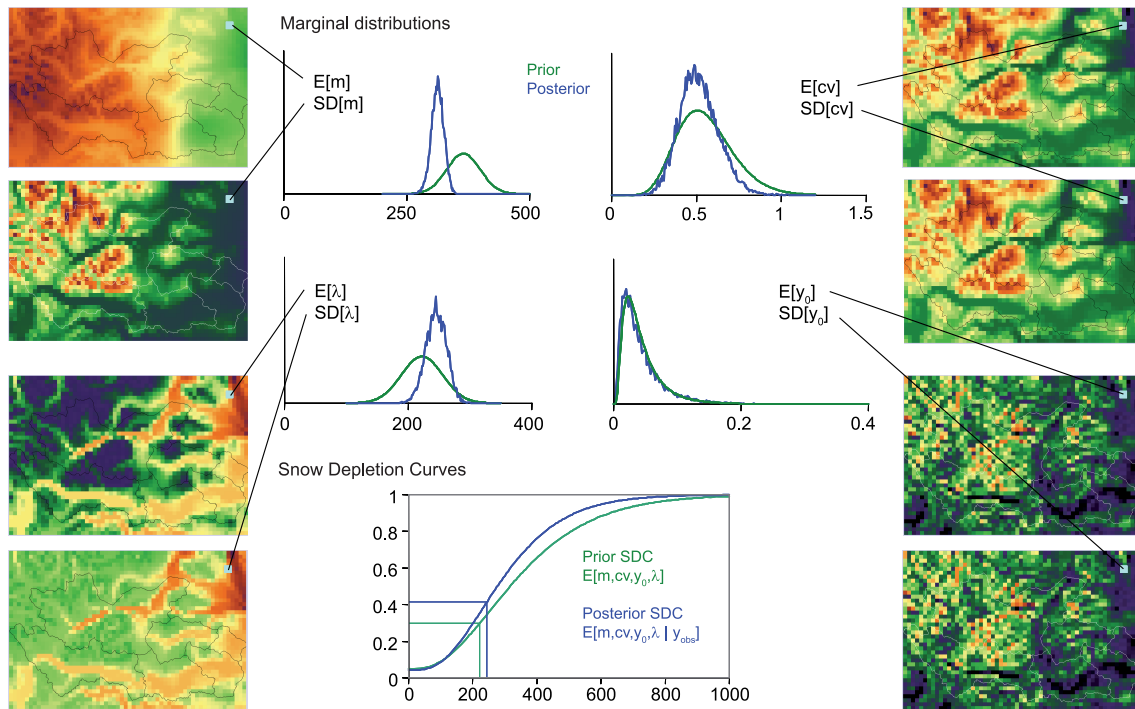Hydrol. Earth Syst. Sci., 10, 369–381, 2006

Fig. 2. The Bayesian assimilation illustrated for a specific grid cell. A priori, the joint distribution $P(m, cv, y_0, \lambda)$ is given by four 2-parameter marginal distributions (green pdf curves). The parameters are calculated by moments extracted from the corresponding maps. The 4-variate joint expectation specifies the prior SDC state. When assimilating satellite information, Bayes' theorem calculates the posterior $P(m, cv, y_0, \lambda | y_{obs})$, printed as blue pdf curves. The new joint expectation defines the updated SDC state. The four dimensions are now mutually dependent, so the marginal posteriors do not fully specify the joint posterior. The whole analysis is repeated for all grid cells in the region.

melt season, the Bayesian assimilation must thus estimate the joint distribution of $n \cdot (k+3)$ variables, since λ varies through time. The available information contains only $n \cdot k$ observations of y, one for each grid cell and image. Since k is usually small, the information deficit is considerable. To reduce this deficit, we relate the accumulated melt depth λ and the end-of-winter snow storage m to processes which act similarly in many grid cells. The resulting spatial interdependency adds conditional information, effectively reducing the size of the estimation problem. This section presents the transformations of λ and m into normalised spatial surfaces and global (spatially constant) variables; the spatial models applied to the normalised surfaces, and the joint prior model for the transformed SDC variables.

3.1 Identification of spatially constant variables

The first step to reduce the size of the estimation problem, is to identify processes which act similarly all over the region, and normalise the affected surfaces with respect to these. It is common practice in geostatistics to remove the effect of trends, periodicities, and other global features from a surface, prior to fitting a spatial dependency structure. In our case, the surfaces are simulated by a PRM, and we target the transfor-

mations at reversing the effect of global calculations usually applied by PRMs, rather than looking for an optimal transformation in terms of variance component separation. Most distributed PRMs apply global elevation gradients to precipitation and temperature, and use a degree-day approach to calculate snow melt from temperature.

For the average snow storage m calculated in the PRM, we estimate an elevation gradient g_m, and calculate the elevation-normalized surface of expected snow storage m^* by removing the effect of g_m. Formally, the transformation between m and m^* is:

$$m = m^* \, (1 + g_m)^{(h-h^*)} \tag{7}$$

Here, h is elevation, and the asterisk denotes a reference altitude. As is common for the precipitation lapse rate, the elevation gradient in m is expressed in relative terms, and thus the elevation normalization is multiplicative to ensure consistency and independence of the reference altitude. Regarding the accumulated melt depth, λ is first decomposed into a global degree-day factor Cx and a sum of positive temperatures z. The degree-day sum z is further decomposed into a global elevation gradient g_z and an elevation-normalized

degree-day sum z^*, which remains gridded. Formally, the transformations between λ and z^* are:

$$\lambda = Cx \cdot z; \quad z = \text{MAX}\left(0, z^* + g_z\left(h - h^*\right)\right) \qquad (8)$$

The elevation gradient g_z applies to the aggregated sum of positive degree-days, not to momentary temperature values, and consequently it varies from day to day. Its formulation as a linear trend is an approximation, not a result of its connection to the temperature lapse rate.

The variables y_0 and cv are not transformed. Over a region of n grid cells and for a single image, the prior distribution now consists of $4n+3$ variables; namely the four gridded maps of m^*, z^*, cv, and y_0, and the three global variables g_m, g_z, and Cx. The global variables are thus specified by distributions and updated just as the gridded maps. With k images, the dimensionality increases to $(3+k) \cdot n + 2 + k$, because z^* and g_z varies in time.

3.2 The prior distribution components

With the elevation gradients removed, the spatial surfaces of z^* and m^* naturally contain less variance than the corresponding surfaces of λ and m. More important, the transformed surfaces are also smoother than the original, because the variance components removed with the global effects resemble the spatial scale of the terrain, which is small compared to the spatial scale of the reference-altitude surfaces. This enables a further reduction in variance, by modelling the spatial dependency between neighbours.

A convenient class of spatial models for smooth surfaces are Gaussian Markov Random Fields (GMRF), see Rue and Held (2005). A GMRF have simplified expressions compared to the general Gaussian distribution for the full conditionals, which is the density of each z_i given the remaining z_j's for all $j \neq i$. In a GMRF, the full conditionals only depend on a few z_j's, in the neighbourhood of each i. This allows for fast computations of the full conditionals, which is important in the Markov chain Monte Carlo algorithm used for inference. In our case, the neighbourhood is set to contain the four directly adjacent cells. For each location i, we chose the conditional moments for z_i^* as

$$E\left[z_i^*|z_{\{j\}}^*\right] = E\left[z_i^*\right] + \frac{\sigma_i}{n_j}\sum_{j=1}^{n_j}\left(\frac{z_j^* - E\left[z_j^*\right]}{\sigma_j}\right);$$

$$\text{VAR}\left[z_i^*|z_{\{j\}}^*\right] = \frac{\sigma_i^2}{n_j} \qquad (9)$$

Here, index i denotes the local cell, $\{j\}$ denotes its n_j neighbours (with $n_j=4$ except at the region boundaries), and σ_i^2 denotes the marginal variance (before introducing spatial dependency) in grid cell i. Formally, this is a first order Intrinsic GMRF (see Rue and Held, 2005; Ch. 3), applied to standardised variables. Standardisation is necessary because both $E\left[z_i^*\right]$ and σ_i^2 vary over space.

Note that Eq. (9) encourage the standardised variables in each cell to be similar to the mean of the neighbours cells, as the conditional expectation is the mean of the neighbours. Even in the simplified GMRF, serial dependency causes all cells to be correlated. Based on the transformations given in Eq. (8), the prior distribution components governing λ_i are given by:

$$Cx \sim \text{Gamma}\left(\frac{E^2[Cx]}{\text{VAR}[Cx]}, \frac{E[Cx]}{\text{VAR}[Cx]}\right), \qquad (10)$$

$$g_z \sim N\left(E[g_z], \text{VAR}[g_z]\right) \qquad (11)$$

$$z_i^*|z_{\{j\}}^* \sim N\left(E\left[z_i^*\right] + \frac{\sigma_i}{n_j}\sum_j\left(\frac{z_j^* - E\left[z_j^*\right]}{\sigma_j}\right), \frac{\sigma_i^2}{n_j}\right) \qquad (12)$$

For m^*, with transformation given in Eq. (7), the prior distribution model and conditional moments are defined similarly as for z^*, omitting Eq. (10) (the temperature index). In prose, both elevation gradients are given Normal priors (Eq. 11), and the temperature index is given a Gamma prior (Eq. 10). The sub-grid coefficient of variation cv and the initial bare ground fraction y_0 are given independent prior distributions in each grid cell, with a Gamma density for cv, and a Lognormal density for y_0. Table 1 summarizes the models used for the different components of the prior distribution, and suggests the typical sources of information governing the estimation of prior moments.

The spatial connectivity in m^* and z^* aids in moving information between cells. An observed SCA close to 1 provides little or no mass balance information. However, a neighbouring cell with higher cv may exhibit a considerably lower SCA, making the observation more informative. With m^* and z^* bound by spatial dependence, this information transfers to the first cell, reducing its marginal variance. A special effect of this is that updating may take place also in grid cells with missing observations. Optical satellite images are notoriously sensitive to cloud coverage, and the value of an updating routine greatly increases with the usefulness of partially cloudy scenes.

4 The likelihood function and the observations

4.1 SCA-based likelihood

The second term on the right hand side of Bayes' theorem (Eq. 6) is the likelihood of the parameters, given the observations. At the core of this likelihood is the collection of grid cell SDCs, expressing the observable variable vector $\boldsymbol{y}$ as a function of the parameters. With y values constricted to $[0,1]$, a beta-distribution is selected to represent the likelihood in each cell. The expectation in this likelihood model is simply the SDC value provided by the actual parameter values, and the variance is the observation uncertainty. For

Table 1. The prior distribution components, and typical sources of information used in determination of prior moments. Besides the listed information sources, formerly assimilated SCA images may contribute information on all variables. The term "geostatistics" imprecisely means estimation from point measurements based on spatial covariance.

Par.	Variability	Model	Expectation	Variance
$m*$	Temporally constant Spatially variable	GMRF	PRM mass balance, or winter precip interpolation	Subjective experience, or geostatistics
g_m	Temporally constant Spatially constant	Normal	Subjective experience, or multi-gauge estimation	Subjective experience, or multi-gauge estimation
$z*$	Temporally variable Spatially variable	GMRF	Interpolation from temperature measurements	Geostatistics
g_z	Temporally variable Spatially constant	Normal	Subjective experience, or multi-gauge estimation	Subjective experience, or multi-gauge estimation
CX	Temporally constant Spatially constant	Gamma	PRM calibration	Subjective experience, or Monte Carlo based calibration approach
Cv	Temporally constant Spatially variable	Gamma	Snow surveys, historical or current-season.	Snow surveys, historical or current-season.
y_0	Temporally constant Spatially variable	Lognormal	Mid-winter satellite images, terrain dependencies	Satellite images, subjective experience

a single observed bare-ground fraction y^{obs} in grid cell i and the image acquired on date t, the likelihood is given as:

$$P\left(y_{it}^{\text{obs}}|m_i^*, g_m, z_{it}^*, g_{zt}, Cx, cv_i, y_{0i}\right)$$
$$= \frac{\Gamma\left(\varphi_{it} + \psi_{it}\right)}{\Gamma\left(\varphi_{it}\right)\Gamma\left(\psi_{it}\right)}\left(y_{it}^{\text{obs}}\right)^{\varphi_{it}-1}\left(1 - y_{it}^{\text{obs}}\right)^{\psi_{it}-1} \quad (13)$$

The parameters Φ_{it} and Ψ_{it} are given by the moments:

$$\varphi_{it} = E\left[y_{it}\right]\left(\frac{E\left[y_{it}\right]\left(1 - E\left[y_{it}\right]\right)}{Var\left[y_{it}^{\text{obs}}\right]} - 1\right);$$

$$\psi_{it} = \left(1 - E\left[y_{it}\right]\right)\left(\frac{E\left[y_{it}\right]\left(1 - E\left[y_{it}\right]\right)}{Var\left[y_{it}^{\text{obs}}\right]} - 1\right) \quad (14)$$

Recalling that the likelihood is conditioned on parameters, the first moment $E[y_{it}]$ is the y value provided by the SDC given the fixed parameters (combining Eqs. 5, 7, and 8):

$$E\left[y_{it}|m_i^*, g_m, z_{it}^*, g_{zt}, Cx, cv_i, y_{0i}\right]$$
$$= y_{0i} + (1 - y_{0i}) \cdot \gamma\left(\frac{1}{cv^2}, \frac{Cx \cdot MAX\left\{0, z_{it}^* + g_{zt}\Delta h\right\}}{cv^2 \cdot m_i^*(1 + g_m)^{\Delta h}}\right) \quad (15)$$

The second moment $Var[y_{it}]$ is the observation variance, as estimated from the satellite image analysis. This variance is spatially heterogeneous, depending on the uncertainty relating to classifying each particular pixel in the satellite image (Rosenthal and Dozier, 1996).

Since all the y values for a single cell over n observations in time are considered conditionally independent (that is, independent for fixed parameters), the multi-observation likelihood for a single cell is simply given by the product of the single-event likelihoods. Correspondingly, since the observations in different cells are also considered conditionally independent, the likelihood for the total set of observations is the product of all the single-cell likelihoods.

4.2 Discharge-based likelihood

The positive dependency between runoff-governing SDC parameters in neighbouring grid cells, makes calculated catchment outflow more responsive to their variation; increasing the validating power of runoff measurements. Engeland and Gottschalk (2002) have used measured discharge time series in Bayesian inference, and developed likelihood models including autoregressive terms. This paper takes a simpler approach, associating the sum of grid cell runoff to date with the sum of accumulated outlet discharge and the change in the internal water storage S_t. A simple linear tank model connects this storage to the observed daily discharge, and is assumed to be zero at the time of catchment melt onset.

Similarly to the SCA-based likelihood, the discharge-based likelihood expectation is given from the SDC parameter values, and the variance from the observation uncertainty. Assuming the tank content to be the major source of uncertainty in the estimated catchment runoff, the storage term is assigned the large relative standard deviation of 50%, compared to 5% for the measured accumulated discharge. The likelihood is given a Normal model, and with observed accumulated runoff Q_{acc} and storage S_t estimated from actual-day discharge Q_t, the likelihood of the sum $Q_{acc} + S_t$ is:

$$P\left(Q_{acc} + S_t|\bar{Q}_t\right) = \frac{1}{\sqrt{2\pi}\sigma_{qs}}\exp\left[-\frac{1}{\sigma_{qs}^2}\left(Q_{acc} + S_t - \bar{Q}_t\right)^2\right];$$
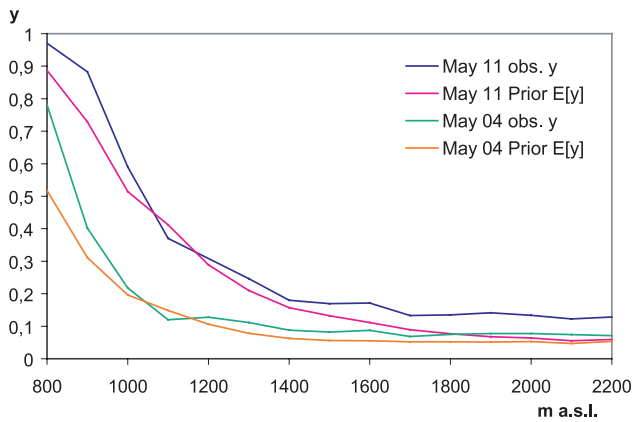
Fig. 3. Observed and a priori expected bare ground area y on 4 May and 11 May, averaged over elevation intervals of 100 m.

$$\bar{Q}_t = \frac{1}{n} \sum_{i=1}^{n} Q_i(\boldsymbol{\theta}_{i,t}) \qquad (16)$$

where $Q_i(\boldsymbol{\theta}_{i,t})$ is accumulated runoff at cell i and to date t, given the SDC parameters; and $\sigma_{qs}^2 = \mathrm{VAR}[Q_{acc} + S_t]$.

5 Site and data

The assimilation of observed snow covered area into spatially distributed snow depletion curves is evaluated over a 60×40 km rectangular region covering the Vinstra and Sjoa catchments in Jotunheimen, central Norway, at 61.4° N, 8.6° E. The elevation ranges from 710 m a.s.l. to 2240 m a.s.l measured at grid scale. Sparse forest and lakes each cover about 10% of the area, small glaciers are present at high altitudes. The accumulation season usually starts in November, and snowmelt in late April or May, depending on elevation.

From two Landsat 7 ETM+ images acquired during the 2000 melt season, gridded maps of fractional bare ground y and its standard error are estimated at 30 m scale, using the decision tree algorithm of Rosenthal and Dozier (1996). Minor adjustments are made to adapt the Landsat 5 TM-based routine to the Landsat 7 ETM+ sensor, which has a different sensitivity in some of the bands. Cloud and lake masks were applied at the same scale, discarding the corresponding 1 km cell if occupying more than 20% of its area.

Runoff observations exist in 6 catchments, which total area covering approximately half the total region. Five of these are nested, heavily regulated subcatchments constituting the Vinstra drainage area, with a total area of 744 km^2; whereas Sjoa is 473 km^2 and unregulated. Naturalized flow series from regulated catchments are known to be of reduced quality on daily and sub-catchment scale. Accumulated over the five subcatchments and several days, the Vinstra discharge data are assumed to have uncertainty similar to Sjoa's.

For the 2000 melt season, 9 updating experiments are run; with satellite data sets from 4 May and 11 May used sepa-
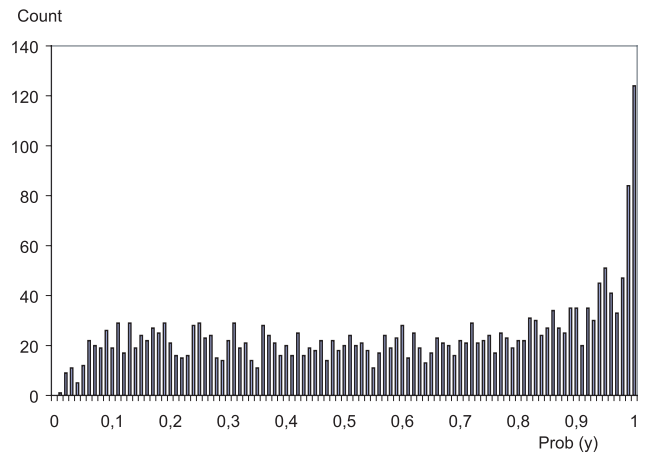

Fig. 4. Spatial histogram of a priori cumulative probability of the observation. A uniform histogram indicates that the prior distribution corresponds well to the observations.

rately and combined; and runoff information either not included, aggregated from the whole gauged part of the region, or from the Sjoa and Vinstra catchments separately. Simulated melt depths suggest a melting front at approximately 1500 m a.s.l. on 4 May and at approx. 1800 m a.s.l. on 11 May. Whether or not the discharge-based likelihood term is included, the posterior distribution is analytically intractable, and is sampled by Markov Chain Monte Carlo techniques using the Metropolis-Hastings algorithm (Chib and Greenberg, 1995). Due to the spatial dependency between cells, the algorithm must be implemented with the spatial loop inside the iteration, that is, each iteration traverses the whole map before the next iteration is executed.

6 Results

6.1 Direct comparison of prior and observed bare-ground fraction y

Before starting to analyse the results of the Bayesian updating, Figs. 3 and 4 both compare the observations to the simulated, a priori snow state. Figure 3 shows that on both dates, the observed bare-ground fraction y_{obs} is underestimated by prior $E[y]$ both at high and low elevation, but is approximately correct around 1100 m a.s.l. Figure 4 displays the spatial histogram of prior cumulative probabilities of the observed value on 11 May. A slight over-representation of cumulative probabilities above 0.8 reflects the tendency to underestimate y_{obs}, but the striking feature of Fig. 4 is the large frequency at the right tail. The prior distribution is obviously too restrictive towards the multi-dimensional tail corresponding to high y values. The tendency at the left tail is opposite, but weaker. Bearing in mind that the prior is subjectively specified with emphasis on marginal moments, the
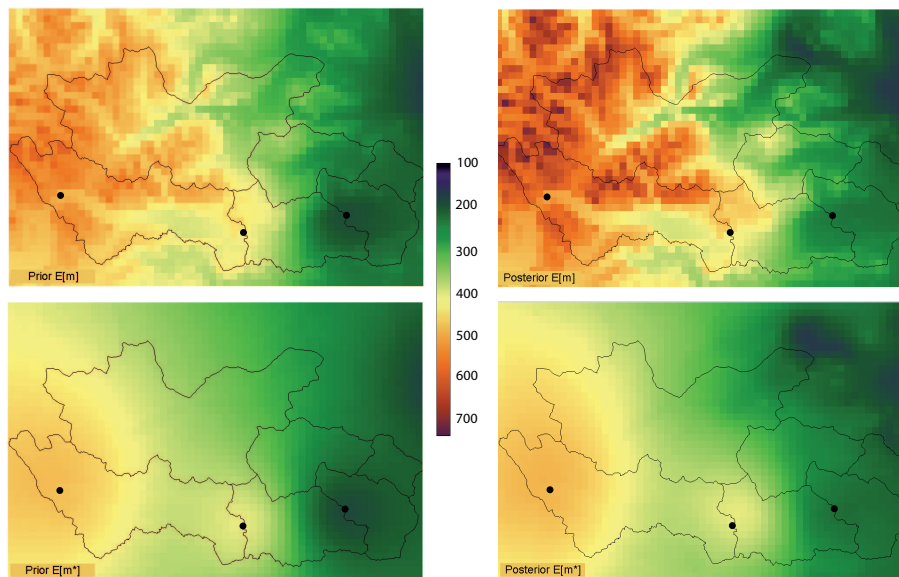
Fig. 5. Expected end of winter snow storage m and its reference-elevation component m^*, through transformation and updating. The elevation gradient is increased, and there are some changes to m^* in the eastern part of the region. Notice the smoothness of the m^* maps compared to the similar maps of m. Subcatchment boundaries are shown, with Sjoa (473 km^2) in the North, and the location of three precipitation gauges are indicated.
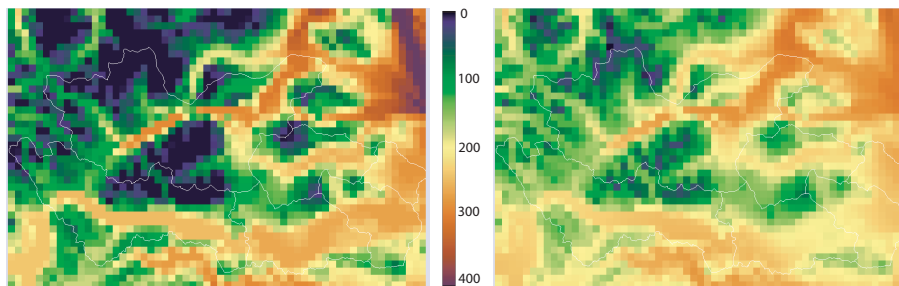

Fig. 6. Expected accumulated melt depth λ; a priori (left) and a posteriori (right). The elevation gradient is considerably less steep, with melt extending to greater elevation, and lower melt depths in the valley to the Northeast.

quantitative posterior estimates should be assessed with care in grid cells where the observation is unlikely high.

6.2 Changes to the end of winter snow storage $E[m]$

In the following, all graphical illustrations are based on the updating experiment using the 11 May image, and not using any runoff-based likelihood term. Figures 5a–d shows how the expected pre-melt snow storage $E[m]$ and its reference-elevation counterpart $E[m^*]$ are updated. Comparing the prior and posterior $E[m]$ maps (top row), it is evident that the posterior elevation gradient g_m is stronger than the prior. In addition, there are local changes to the reference-altitude $E[m^*]$ in the eastern part of the region (bottom row). Note the smoothness of the $E[m^*]$ maps compared to those of $E[m]$. The changes to $E[m]$ and $E[m^*]$ agree with the similar changes on 4 May with correlation coefficients of 0.87 and 0.81, respectively.

For the snow storage gradient g_m, the prior and posterior expectations and standard deviations are given in Table 2. The increased elevation dependency seen in Fig. 5 is confirmed by the 11 May g_m increasing from approx. 2.6%/100 m to more than 4%/100 m. Even more dramatic, the 4 May image results in a posterior $E[g_m]$ of more than three times the prior; far out on the prior distribution tail. The use of both images yields an intermediate $E[g_m]$, but closer to the moderate increase of 11 May. Including observed discharge in the likelihood produces a slight reduction in posterior $E[g_m]$, most notably for the 4 May image.

6.3 Changes to the accumulated melt depth $E[\lambda]$

Figure 6 shows the updating of $E[\lambda]$. Here, the negative elevation dependency is attenuated, with melt extending to a greater altitude in the posterior than in the prior case. Again, there are local changes as well, for instance in the south-
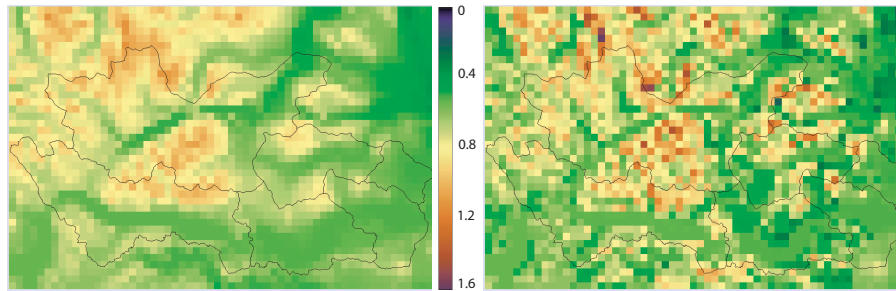
Fig. 7. Expected sub-grid SWE coefficient of variation *cv*; a priori (left) and a posteriori (right). The prior expectancy is linked to elevation, assuming that stronger winds at high altitude produce more re-distribution of snow. The posterior expectancy show more small-scale variability, but have approximately similar elevation dependency.

Table 2. Changes in expectation and standard deviation for the E[SWE] elevation gradient, resulting from different combinations of images and discharge data in the likelihood. The gradient is in % per 100 m.

	A priori	4 May posterior			11 May posterior			4 + 11 May posterior		
# discharge series:		0	1	2	0	1	2	0	1	2
$E[g_m]$	2.58	9.93	9.39	8.58	4.39	4.29	4.03	5.41	5.12	4.53
$sd[g_m]$	1.5	0.73	0.71	0.66	0.64	0.61	0.60	0.51	0.57	0.55

Table 3. Changes in expectation and standard deviation for Cx and gz, resulting from different combinations of images and discharge data. Two gauged subcatchments in the region are used separately or in sum. The gradient is in °C*d/100 m, Cx is the degree-day melt index. Observed gz between 850 and 1600 m a.s.l. is −7.7 degree-days/100 m on 4 May and −13.7 degree-days/100 m on 11 May, respectively.

	A priori	4 May posterior			11 May posterior			4 + 11 May posterior		
# discharge series:		0	1	2	0	1	2	0	1	2
$E[Cx]$	3.00	3.51	3.49	3.46	2.88	2.87	2.86	3.03	3.02	2.98
$sd[Cx]$	0.75	0.06	0.06	0.06	0.04	0.05	0.04	0.04	0.04	0.04
4 May $E[g_z]$	−9.32	−3.84	−3.90	−4.02				−4.10	−4.20	−4.23
4 May $sd[g_z]$	2.15	0.2	0.2	0.1				0.13	0.13	0.12
11 May $E[g_z]$	−12.12				−7.11	−7.15	−7.25	−6.11	−6.26	−6.40
11 May $sd[g_z]$	2.80				0.22	0.20	0.21	0.20	0.22	0.19

eastern part of the area, but the gradient driven changes dominate. The changes to the time variant $E[\lambda]$ and $E[z^*]$ on 11 May agree with the changes to these variables on 4 May with correlation coefficients of 0.85 and 0.89, respectively.

Table 3 shows the updating of the degree-day factor Cx and the degree-day sum gradient g_z, for the actual date. For $E[g_z]$, all experiments result in a large increase, i.e. an attenuation of the negative gradient, the 4 May image more than on 11 May. Observe that the two updates address different g_z, since the degree-day sum z varies in time. Using both images strengthens the 4 May $E[g_z]$, but further attenuates the 11 May $E[g_z]$, compared to the single-image posterior estimates. Adding discharge information has only marginal effect. For Cx, only moderate adjustments are made, and posterior $E[Cx]$ is well within the prior uncertainty.

Figure 7 shows the prior and posterior $E[cv]$, quantifying the sub-grid heterogeneity of snow storage and determining the shape of the Gamma model used for the SDC. The main effect of updating is an increase in the small-scale variability, evident as more scatter in the image. Also for this variable, there are instances of local neighbourhoods showing a common response. Since the *cv* prior has no spatial structure, this is either a direct result of the observations, or a result of posterior dependencies between *cv* and m^* or z^*, which do have spatial priors. For $E[cv]$, the correlation between 4 May and 11 May updates is 0.68.

The prior and posterior y_0 maps provide little valuable spatial information, and are omitted. $E[y_0]$ changes are generally small, and negative in the majority of the cells. However, the 11 May update cause an increase in $E[y_0]$ of more

Table 4. Fractional reduction of variance from prior to posterior state; map averages.

Fractional variance reduction [%] (Map averages)	4 May update # discharge series			11 May update # discharge series			4 + 11 May update # discharge series		
	0	1	2	0	1	2	0	1	2
cv	55.0	54.8	54.2	45.7	45.8	45.7	63.4	63.4	63.3
m	10.4	16.9	26.7	45.1	48.0	48.4	48.2	49.8	50.8
y_0	67.7	67.2	66.2	33.1	33.3	32.1	63.1	63.3	62.6
4 May λ	47.1	50.0	56.1				65.3	66.3	68.6
4 May y	73.4	72.8	72.9				75.5	75.5	76.0
4 May SWE	52.9	56.4	61.7				75.5	75.7	76.9
4 May Q	61.7	63.1	65.8				76.3	76.8	78.0
11 May λ				94.7	95.0	95.1	94.5	94.6	94.9
11 May y				82.6	83.0	83.0	88.0	87.8	88.0
11 May SWE				77.9	78.5	79.1	82.6	82.7	83.5
11 May Q				86.4	86.6	86.5	89.5	89.8	90.1

than 0.15 in approximately 50 grid cells, and the maximum posterior value is as high as 0.72. On 4 May only 11 pixels have $E[y_0]$ increased by more than 0.15, with 0.46 as the maximum posterior value. The correlation coefficient between the 4 May and the 11 May updates is 0.57.

6.4 Uncertainty reduction

A comparison of prior and posterior variance is useful to quantify the information content in the SCA observations, with respect to the SDC parameters and the derived snow states. Table 4 shows that the 11 May image provides more information on the mass balance governing m and λ. The earlier 4 May image provides more information on the variables cv and y_0, which mainly affect the sub-grid snow distribution. The two images used together generally reduce variance more than any of them alone, except for y_0 from the 4 May image and λ from the 11 May image. The inclusion of likelihood terms from discharge measurements provide little or no extra information for the 11 May and the 4 + 11 May updates, but significantly improves the mass balance precision on 4 May.

7 Discussion

Evaluating the relative variance reductions (Table 4), substantial improvement is achieved for all variables. Even with the 4 May image, less than a week after melt onset, the variance in remaining snow storage SWE is reduced by more than 50% on average. On 11 May the average variance reduction is more than 75%, i.e. the posterior variance is less than a quarter of the prior. The similar method used with independent priors in each grid cell (Kolberg and Gottschalk, 2005) reduced the prior variance in m on 4 May and 11 May by only 12% and 15%, respectively. Obviously, the spatial prior

largely compensates the information deficit apparent for each individual cell, at the cost of some spatial detail.

7.1 Gradient estimation difficulties

For the 4 May image, the posterior snow storage gradient $E[g_m]$ of nearly 10% per 100 m (Table 2) is beyond credibility. The poor variance reduction in snow storage m on 4 May (Table 4) is due to this large value, because g_m is in relative terms and thus increases the m uncertainty at high altitude. Also the posterior degree-day sum gradient g_z (Table 3) is extreme. Noticing that the observed changes in $E[g_m]$ and $E[g_z]$ affect $E[y]$ in opposite directions, the two gradients appear to attain unrealistic values while compensating each others effect on the posterior y. Recalling that the likelihood only relates y to the mass balance through the λ/m ratio (Eq. 5), changes in λ and m easily compensate. With elevation as a major source of variability in both λ and m, this compensation could transfer to the two gradients. In the posterior distribution, g_m and g_z are indeed positively dependent with an R^2 of 0.30 on 4 May, 0.40 on 11 May. This dependency is noticeable, but cannot alone explain the extreme values, in particular noticing the low posterior variance.

More than a problem of weak identification due to mutual dependency, the unlikely gradients on 4 May is a problem of extrapolation. Above the melting front, y is insensitive both to m and to any realistic change in λ, leaving the SCA observations merely non-informative. Figure 3 indicates that the elevation dependency of y is a priori underestimated below 1100 m a.s.l., and overestimated between 1100 m and 1400 m, while above 1400 m, y is close to y_0. With no restrictions from high-altitude information, g_m and g_z adapt freely to the situation below 1400 m a.s.l., and combine to mimic the observed gradient both below and above 1100 m. Even below 1400 m, there are plenty of cells contributing to the likelihood, ensuring the low posterior variance. In

Table 5. Prior and posterior mass balance of the Vinstra and Sjoa catchments, with the observed values estimated from the discharge series. Posterior columns identify the nine experiments defined by images used (4 May, 11 May, both) and number of discharge series used in the likelihood.

Mass balance [mm]			4 May posterior # discharge series			11 May posterior # discharge series			4 + 11 May posterior # discharge series		
	Prior	Obs	0	1	2	0	1	2	0	1	2
Vinstra											
$\overline{m}$	378	474	514	503	483	421	420	415	441	436	423
4 May $\overline{Q}$	55	95	106	104	100				99	98	95
11 May $\overline{Q}$	125	157				150	150	149	165	163	159
Sjoa											
$\overline{m}$	443	550	683	663	623	484	481	473	514	504	481
4 May $\overline{Q}$	30	35	99	97	89				85	85	79
11 May $\overline{Q}$	83	83				127	126	124	140	138	130

contrast, the poor two observations in the runoff likelihood only suffice to slight moderations in the posteriors which they strongly contradict.

7.2 Comparing the updates to measurements

In Table 5, the posterior end of winter snow storage $E[m]$ and runoff to date Q produced by the different updating experiments are averaged for the two measured catchments, and compared to their prior expectations as well as observed values. Average snow storage $\overline{m}$ is compared to the accumulated runoff at the end of the melt season, corrected for precipitation during the snow melt period. $\overline{m}$ is underestimated in the prior state, and increased by all updating experiments, in particular through a higher gradient. For the 4 May image, the high gradient discussed above yields an over-correction of $\overline{m}$; and in Sjoa, the posterior absolute error is even larger than the prior except when two observed runoff series are used.

Concerning accumulated runoff to date $\overline{Q}$, all the updating experiments result in increased estimates. For the Vinstra catchment, the posterior $E[\overline{Q}]$ is largely confirmed by the observed values, again with the changes on 4 May slightly too large. For Sjoa, however, observed $\overline{Q}$ on both dates corresponded well to the prior $E[\overline{Q}]$, whereas the 4 May update produces a three-fold increase in both runoff and melt depth. Less dramatic, but still too large, is the runoff increase produced by the 11 May update. These changes are mainly caused by the attenuation of the g_z gradient discussed above. The situation improves slightly by including runoff observation in the likelihood, for the 4 May case also by using both images.

A temperature station at 1600 m a.s.l. is not used in the prior estimation, and combined with one at 850 m a.s.l. it provides an independent evaluation of the posterior g_z estimates. Calculated gradients g_z are -7.73 and -13.67 degree-days/100 m on 4 May and 11 May, respec-
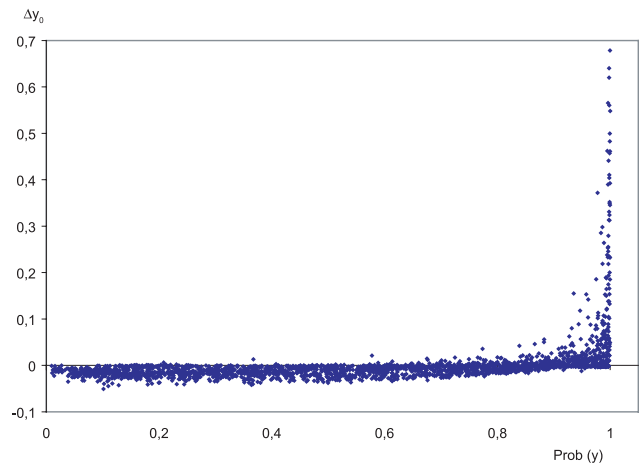

Fig. 8. Expectation changes for y_0 versus the prior cumulative probability of the observed y value. When y is observed at the right tail of its a priori distribution (unexpectedly low SCA), y_0 sometimes responds with an extreme increase. For observations appearing reasonably within the prior uncertainty, the resulting adjustment of y_0 is generally small. Thus, y_0 serves as a "safety valve" for unlikely high y observations.

tively. The prior distributions of g_z capture these observations within $+/-1$ standard deviation on both dates (Table 3), but the posteriors fail to do so, both by larger errors and by smaller standard deviations.

7.3 Small-scale variability and the role of y_0

Figure 8 shows that y_0 in some cases responds dramatically when unexpectedly high y values are observed. No other variable is changed this far out of normal range by such observations. The neighbourhood dependency applied to m and λ prevent these variables from being too hardly altered in isolated cells. With no such restrictions, the local variables cv

and y_0 are more likely to respond to small-scale variability, even when caused by local m or λ anomalies like a net transport loss of snow, or an increased melt due to sun exposure. The effect of cv on y is small for low λ values, whereas y_0 is directly connected, and has a heavily right skewed lognormal prior allowing it to respond. This may be the reason why y_0 is poorer defined by the use of both images than by the 4 May image only (Table 5); each of the images forces y_0 to compensate an unlikely situation. No single variable responds similarly to low y observations, which are more easily captured by moderate changes in all parameters.

Recognising that y_0 does not contribute to any relationship between snow coverage and melt depth, one might consider removing this parameter or fixing its value close to 0. For predictive purposes, observations beyond reasonable prior credibility could be simply disregarded. The effect is similar to a single-cell cloud in the satellite image, and the actual grid cell is only updated through neighbourhood dependency. For investigative purposes, the increase in observed y between the two dates at high altitude, with no elevation dependency above 1700 m a.s.l. (Fig. 3), suggests that neither the "initial" property of y_0, nor the degree-day model consequence of a distinct melting front, are well justified.

8 Conclusions

Identifying processes that operate similarly on many grid cells, is a powerful way to increase the informational value of remotely sensed snow coverage data. In this study, the isolation of two elevation gradients and a temperature melt index as spatially constant variables enabled a strong spatial dependence to be built into the prior distribution of the mean snow storage m and the accumulated melt depth λ. The transformed, spatial prior distribution produced largely better results than a non-spatial prior model previously reported.

The information content of a snow coverage observation depends on the situation it images. Early observations contain more information about the snow cover heterogeneity, in this case the sub-grid coefficient of variation cv and the premelt bare ground fraction y_0. In particular, observations prior to the melt onset are close to non-informative except for y_0. However, the spatial prior enables mass balance information gained in some grid cells, to be transferred to neighbouring or similar-altitude grid cells where the SCA observation was less informative.

The inability of a local SCA observation to contain mass balance information beyond relative terms, is to some extent reflected in the spatial data as compensation between the elevation gradients. In this case an early image leads to some unrealistic results, largely because the two gradients combine to adapt to a non-linear feature within a limited elevation interval. In the common situation with meteorological measurements concentrated at low altitude, an early image

lacking effective information at high-altitude is likely to produce estimates sensitive to extrapolation.

Combination of two images usually provides a greater reduction of variance than any of the images alone. The inclusion of observed runoff in the likelihood, however, generally supplies weak explanatory power, compared to the satellite image. This is due to the discharge being observed in only two catchments, whereas the SCA likelihood variance is greatly reduced by multiplying more than 2000 single-cell likelihood distributions. However, in the 4 May case, where the mass balance information of the image is smallest, the runoff observation is able to reduce some of the error in the gradient estimation, as well as the posterior variance in the mass balance variables.

Edited by: G. Pegram

References

Beven, K. J. and Young, P.: Comment on "Bayesian recursive parameter estimation for hydrologic models" by M. Thiemann, M. Trosset, H. Gupta, and S. Sorooshian, Water Resour. Res. 39, 5, doi:10.1029/2001WR001183. 2003.

Binley, A. M. and Beven, K. J.: Physically-based modelling of catchment hydrology: A likelihood approach to reducing predictive uncertainty, in: Computer modelling in the environmental sciences, edited by: Farmer, D. G. and Rycroft, M. J., Clarendon, Oxford, 1991.

Blöschl, G., Kirnbauer, R., and Gutknecht, D.: Distributed Snowmelt Simulations in an Alpine Catchment 1. Model Evaluation on the Basis of Snow Cover Patterns, Water Resour. Res., 27, 12, 3171–3179, 1991.

Brubaker, K. L., Jasinski, M., Chang, A. T., and Josberger, E.: Interpolating sparse surface measurements for calibration and validation of satellite-derived snow water equivalent in Russian Siberia, in: Remote sensing and Hydrology 2000, edited by: Owe, M., Brubaker, K., Ritchie, J., and Rango, A., IAHS Publ. No. 267, p. 93–98, 2001.

Brubaker, K. L. and Menoes, M.: A technique to estimate snow depletion curves from time-series data using the Beta distribution, Proceedings, 58th Eastern Snow Conference, Ottawa, Ontario, Canada, 2001.

Bruland, O., Sand, K., and Killingtveit, Å.: Snow distribution at a high arctic site at Svalbard, Nordic Hydrol., 32, 1, 1–12, 2001.

Chib, S. and Greenberg, E.: Understanding the Metropolis-Hastings algorithm, The American Statistician, 49, 4, 327–335, 1995.

Donald, J. R., Soulis, E. D., Kouwen, N., and Pietroniro, A.: A land cover-based snow cover representation for distributed hydrologic models, Water Resour. Res., 31, 4, 995–1009, 1995.

Engeland, K. and Gottschalk, L.: Bayesian Estimation of Parameters in a regional Hydrological Model, Hydrol. Earth Syst. Sci., 6, 5, 883–898, 2002.

Engen, G., Guneriussen, T., and Overrein, Ø.: Delta-K interferometric SAR technique for snow water equivalent (SWE) re-

trieval, IEEE Geoscience and remote sensing letters, 1, 2, 57–61, 2004.

Essery, R., Li, L., and Pomeroy, J.: A distributed model of blowing snow over complex terrain, Hydrological Processes, 13, 2423–2438, 1999.

Gupta, H. V., Thiemann, M., Trosset, M., and Sorooshian, S.: Reply to comment by K. Beven and P. Young on "Bayesian recursive parameter estimation for hydrologic models), Water Resour. Res., 39, 5, doi:10:1029/2002WR001405, 2003.

Huang, H.-C. and Cressie, N.: Spatio-temporal prediction of snow water equivalent using the Kalman filter, Computational Statistics and Data Analysis, 22, 159–175, 1996.

Kirnbauer, R., Blöschl, G., and Gutknecht, D.: Entering the era of distributed snow models, Nordic Hydrol., 25, 1–24, 1994.

Kolberg, S. A.: A snow routine for Bayesian updating, Proceedings, 13th Northern Research Basins international symposium and workshop, Saarisälke, Finland, August 19–24, p. 121–130, 2001.

Kolberg, S. A. and Gottschalk, L.: Updating of snow depletion curve with remote sensing data, Hydrological Processes, in press, 2005.

Liston, G. E.: Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for Atmospheric, Hydrologic, and Ecologic modelling, J. Appl. Meteorol., 38, 1474–1487, 1999.

Liston, G. E.: Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models., J. Climate, 17, 6, 1381–1397, 2004.

Luce, C. H., Tarboton, D. G., and Cooley, K. R.: Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model, Hydrological Processes, 13, 1921–1933, 1999.

Luce, C. H. and Tarboton, D. G.: The application of depletion curves for parameterization of subgrid variability of snow, Hydrological Processes, 18, 1409–1422, 2004.

Marchand, W-D., and Killingtveit, Å.: Statistical properties of spatial snowcover in mountainous catchments in Norway, Nordic Hydrol., 35-2, 101–117, 2004.

Molotch, N. P., Fassnacht, S. R., Bales, R. C., and Helfrich, S. R.: Estimating the distribution of snow water equivalent and snow extent beneath cloud cover in the Salt-Verde River Basin, Arizona, Hydrological Processes, 18, 1595–1611, 2004.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28, 135–147, 2005.

Rango, A. and Martinec, J.: Water storage in mountain basins from satellite snow cover monitoring, in: Remote sensing and Geographic Information Systems for design and operation of water resources systems, edited by: Baumgartner, M. F., Shultz, G., and Johnson, I., IAHS Publ. No 242, p. 83–91, 1997.

Ranzi, R., Grossi, G., and Bacchi, B.: Ten years of monitoring areal snowpack in the Southern Alps using NOAA-AVHRR imagery, ground measurements and hydrological data, Hydrological Processes, 13, 2079–2095, 1999.

Rodell, M. and Houser, P. R.: Updating a land surface model with MODIS-derived snow cover, J. Hydrometeorol., 5, 1064–1075, 2004.

Rosenthal, W. and Dozier, J.: Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper, Water Resour. Res., 32, 1, 115–130, 1996.

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

Salomonsson, V. V. and Appel, I.: Estimating fractional snow cover from MODIS using the normalized difference snow index, Remote sensing of environment, 89, 351–360, 2004.

Sheffield, J., Pan, M., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C, Robock, A., Lohmann, D., Cosgrove, B., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., Tarpley, J. D., and Ramsay, B. H.: Snow process modelling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent, J. Geophys. Res., 108(D22), 8849, doi:10.1029/2002JD003274, 2003.

Skaugen, T., Alfnes, E., Langsholt, E. G., and Udnæs, H.-C.: Time-variant snow distribution for use in hydrological models, Annals of Glaciology, 38, 180–186, 2004.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A shuffled Complex Evolution Metropolis algorithm for optimisation and uncertainty assessment of hydrologic model parameters, Water Resour. Res., 39, 8, doi:10.1029/2002WR001642, 2003.

Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, Water Resour. Res., 41, W01017, doi:10.1029/2004WR003059, 2005.