A Simple Data Assimilation System for Complex Snow Distributions (SnowAssim)

GLEN E. LISTON AND CHRISTOPHER A. HIEMSTRA

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

(Manuscript received 8 January 2007, in final form 14 August 2007)

ABSTRACT

A methodology for assimilating ground-based and remotely sensed snow data within a snow-evolution modeling system (SnowModel) is presented. The data assimilation scheme (SnowAssim) is consistent with optimal interpolation approaches in which the differences between the observed and modeled snow values are used to constrain modeled outputs. The calculated corrections are applied retroactively to create improved fields prior to the assimilated observations. Thus, one of the values of this scheme is the improved simulation of snow-related distributions throughout the entire snow season, even when observations are only available late in the accumulation and/or ablation periods. Because of this, the technique is particularly applicable to reanalysis applications. The methodology includes the ability to stratify the assimilation into regions where either the observations and/or model has unique error properties, such as the differences between forested and nonforested snow environments. The methodologies are introduced using synthetic data and a simple simulation domain. In addition, the model is applied over NASA’s Cold Land Processes Experiment (CLPX), Rabbit Ears Pass, Colorado, observation domain. Simulations using the data assimilation scheme were found to improve the modeled snow water equivalent (SWE) distributions, and simulated SWE displayed considerably more realistic spatial heterogeneity than that provided by the observations alone.

1. Introduction

The seasonal snow cover’s high albedo, low thermal conductivity, and considerable spatial and temporal variability play a key role in governing Earth’s global radiation balance (Walsh et al. 1985; Hall 1988; Karl et al. 1993); this balance is a primary driver of Earth’s atmospheric circulation system and associated climate. The problem of realistically representing seasonal snow in local, regional, and global atmospheric and terrestrial models is complex because of snow-related features that include considerable spatial variability at scales below those resolved by most models (e.g., Loth and Graf 1998; Pomeroy et al. 1998; Slater et al. 2001; Takata et al. 2003; Liston 2004). In light of the role snow cover plays in influencing land, atmospheric, hydrological, and ecosystem processes, it is essential that local, regional, and global models used to simulate these processes be capable of accurately describing seasonal snow evolution.

In contrast to many atmospheric applications in which the observation networks (e.g., synoptic) have been designed to capture existing spatial variability for the field of interest (e.g., air temperature), with snow the natural spatial variability is almost always greater than that represented by observations (Elder et al. 1991; Blöschl 1999; Sturm and Liston 2003; Liston 2004). This high spatial variability is generally the result of forest–canopy interactions with snowfall (e.g., Bonan 1991; Liston and Elder 2006a) and wind interactions among variable surface features, such as topographic ridges, bumps, and depressions, and vegetation characteristics (e.g., Hiemstra et al. 2002, 2006; Liston et al. 2002, 2007).

In recent years, considerable progress has been made developing models to simulate these detailed distributions and the associated physical processes (e.g., Pomeroy et al. 1998; Essery et al. 1999; Winstral and Marks 2002; Liston et al. 2007). Frequently, these models are able to simulate observed patterns, but the magnitudes are often deficient. This error can be the result of limitations in the model physics, errors in meteorological forcing (e.g., snow precipitation; Liston and Sturm 2004), and deficiencies in boundary conditions [e.g., relatively low-resolution topography and vegeta-
tion data; Liston and Sturm (1998); Hiemstra et al. (2006); Liston et al. (2007)]. These inadequacies can propagate the associated errors in many ways (Burrough and McDonnell 1998). For example, meteorological errors in air temperature and solar radiation can impact melt rates and other processes (e.g., Rivington et al. 2006), and seasonal snow simulations are affected through the additive processes of accumulation and ablation, such that errors early in the season can also exist late in the season. The presence of these model and data-forcing limitations suggest that merging a data assimilation scheme (SnowAssim) with the snow model to generate more accurate snow distributions is merited. In this process, data assimilation is used to constrain and enhance a model-simulated field (e.g., snow depth or water equivalent) using available observations that are irregularly (or regularly) distributed in time and space.

Data assimilation applications are common in the atmospheric sciences (see Daley 1991) and are becoming more frequent in land surface applications, such as those related to soil moisture (e.g., Houser et al. 1998; Reichle et al. 2002; Rodell et al. 2004). In contrast, snow-related assimilation applications are in their infancy. Initially, snow assimilation efforts followed a direct-insertion approach in which observations replaced the model-simulated values when and where the observations were available. For example, Liston et al. (1999) inserted snow water equivalent (SWE) data to identify snow-related deficiencies in a regional atmospheric model. Rodell and Houser (2004) used a similar approach in which the Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover data were incorporated into a global land surface hydrology model. This relatively simple approach can be improved by using an optimal, or statistical, interpolation scheme in which differences between modeled and observed values are distributed across the simulation domain and used to correct the model simulations (e.g., Rutherford 1972; Daley 1991). Brasnett (1999) used this method to distribute snow depths as part of the Canadian Meteorological Centre’s global analysis, and Brown et al. (2003) used it to grid snow depths and SWE over North America. At the next level of complexity are data assimilation schemes that use methods such as the various adaptations of Kalman filters (Gelb 1974; Evensen 1994), which account for relative observation and model-simulation uncertainty. Sun et al. (2004) used an extended Kalman filter to assimilate SWE observations for land surface model initial conditions. Slater and Clark (2006) used an ensemble Kalman filter to improve SWE estimates in Colorado based on snowpack telemetry (SNOTEL) and the National Weather Service’s (NWS) cooperative (COOP) observations, and Andreadis and Lettenmaier (2006) used an ensemble Kalman filter to assimilate snow cover extent data as part of SWE simulations in the Snake River basin. Snow data assimilation programs also exist within U.S. operational centers, such as those performed by the Air Force Weather Agency (AFWA), and the National Weather Service’s National Operational Hydrologic Remote Sensing Center’s (NOHRSC) Snow Data Assimilation System (SNODAS; Rodell and Houser 2004).

In an effort to constrain snow model outputs with observational data, we have developed a straightforward SnowAssim and merged it with the snow-evolution modeling system (SnowModel; Liston and Elder 2006a). SnowAssim’s primary application, and that presented here, is for retrospective applications. An example of this is the production of atmospheric and hydrologic reanalysis datasets, widely used in climate research (e.g., Kalnay et al. 1996; Fan et al. 2006; Mesinger et al. 2006). In these applications, observational datasets from before and after the analysis time are available to contribute to improved simulations of atmospheric and land surface processes (e.g., Cohn et al. 1994; Todling et al. 1998; Seo et al. 2003; Zhu et al. 2003).

The data assimilation scheme is composed of performing spatial interpolations of the model-simulated errors, defined as the differences between the observations and an initial model simulation of the observed variable, and using those difference fields to impose corrections to a second model simulation that produces an improved distribution of the simulated variable. This approach is consistent with Gandin’s (1963) optimal interpolation approach applied to an observed-minus-simulated difference field and with Rutherford’s (1972) data assimilation using simulated error fields. As such, it is considerably simpler than variational assimilation and fixed-interval Kalman smoothing methods, which are also appropriate for reanalysis applications (Fisher et al. 2005). In our application, the corrections are determined by comparing the observations with an initial model simulation and then applying the corrections retroactively as part of a second model simulation to create improved fields prior to the assimilated observation(s). As a result, snow distribution simulations are improved throughout the snow season, even when observations are only available late in the accumulation and/or ablation periods.

Our objective is to describe our data assimilation system and apply it to an example snow application. In this paper we follow Reichle et al. (2002) and Sun et al. (2004), by performing test simulations using synthetic
sets. MicroMet minimally requires screen-height air
and/or gridded atmospheric model, or analyses data-
tion model that uses meteorological station datasets
2006b). MicroMet is a data assimilation and interpola-
cal distribution model (MicroMet; Liston and Elder
(e.g., 1 m
1 km horizontal grid increment) meteorologi-
provided by a quasi-physically based, high-resolution
fields of topography and vegetation
surrounding landscape (primarily topography) to dis-
bounce the variables over any given landscape in
physically plausible and computationally efficient ways.
MicroMet performs two kinds of adjustments to the
meteorological data: 1) all available data, at a given
time, are spatially interpolated over the domain; and 2) 
physically based submodels are applied to each Mi-
CroMet variable to improve realism at any given point
in space and time. Station interpolations (horizontal) to 
a regular grid are done using a Barnes objective analy-
sis scheme (Barnes 1964, 1973; Koch et al. 1983). The
Barnes scheme applies a Gaussian distance-dependent
weighting function in which the weight that a station 
contributes to the overall value of the grid point de-
creases with increasing distance from the observation. 
Interpolation weights are objectively determined as a 
function of data spacing and distribution. At each time 
step, MicroMet generates distributions of air tempera-
ture, relative humidity, wind speed and direc-
tion, and precipitation data. The model uses known re-

terations where snow occurs. It is an aggregation of three submodels: EnBal (Liston 1995; Liston et al. 2000) cal-
culates surface energy exchanges, SnowPack (Liston 
and Hall 1995) simulates snow depth and water-
equivalent evolution, and SnowTran-3D (Liston and 
Sturm 1998; Liston et al. 2007) accounts for snow re-
distribution by wind.

SnowModel is designed to run on grid increments of
1–200 m and temporal increments of 10 min–1 day. It 
can be applied using much larger grid increments, if the
inherent loss in high-resolution (subgrid) information 
(Liston 2004) is acceptable. Simulated processes in-
clude the accumulation from snow precipitation; blow-
ing-snow redistribution and sublimation; interception,
unloading, and sublimation within forest canopies;
snow-density evolution; and snowpack ripening and 
melt. SnowModel includes first-order physics, required 
to simulate snow evolution within each of the global
snow classes (i.e., ice, tundra, taiga, alpine/mountain,
prairie, maritime, and ephemeral), defined by Sturm et 
al. (1995). An attractive feature of the distributed
SnowModel is that, for example, it can blow and drift 
snow in an alpine area of the simulation domain while
simultaneously melting snow in a valley below. Re-
quired SnowModel inputs include temporally varying 
fields of precipitation, wind speed and direction, air
temperature, and relative humidity, obtained from me-

terological stations and/or an atmospheric model lo-

cated within or near the simulation domain, and spa-
tially distributed fields of topography and vegetation

Meteorological forcings required by SnowModel are
provided by a quasi-physically based, high-resolution
(e.g., 1 m–1 km horizontal grid increment) meteorologi-
cal distribution model (MicroMet; Liston and Elder
2006b). MicroMet is a data assimilation and interpola-
tion model that uses meteorological station datasets
and/or gridded atmospheric model, or analyses data-
sets. MicroMet minimally requires screen-height air
temperature, relative humidity, wind speed and direc-
tion, and precipitation data. The model uses known re-
relationships between meteorological variables and the
surrounding landscape (primarily topography) to dis-
bounce those variables over any given landscape in
physically plausible and computationally efficient ways.

2. SnowModel and MicroMet summary

SnowModel (Liston and Elder 2006a) is a spatially
distributed snow-evolution modeling system, designed 
for application in all landscapes, climates, and condi-
tions where snow occurs. It is an aggregation of three
submodels: EnBal (Liston 1995; Liston et al. 2000) cal-
culates surface energy exchanges, SnowPack (Liston 
and Hall 1995) simulates snow depth and water-
equivalent evolution, and SnowTran-3D (Liston and 
Sturm 1998; Liston et al. 2007) accounts for snow re-
distribution by wind.

SnowModel is designed to run on grid increments of
1–200 m and temporal increments of 10 min–1 day. It 
can be applied using much larger grid increments, if the
inherent loss in high-resolution (subgrid) information 
(Liston 2004) is acceptable. Simulated processes in-
clude the accumulation from snow precipitation; blow-
ing-snow redistribution and sublimation; interception,
unloading, and sublimation within forest canopies;
snow-density evolution; and snowpack ripening and 
melt. SnowModel includes first-order physics, required 
to simulate snow evolution within each of the global
snow classes (i.e., ice, tundra, taiga, alpine/mountain,
prairie, maritime, and ephemeral), defined by Sturm et 
al. (1995). An attractive feature of the distributed
SnowModel is that, for example, it can blow and drift 
snow in an alpine area of the simulation domain while
simultaneously melting snow in a valley below. Re-
quired SnowModel inputs include temporally varying 
fields of precipitation, wind speed and direction, air
temperature, and relative humidity, obtained from me-

terological stations and/or an atmospheric model lo-

cated within or near the simulation domain, and spa-
tially distributed fields of topography and vegetation

Meteorological forcings required by SnowModel are
provided by a quasi-physically based, high-resolution
(e.g., 1 m–1 km horizontal grid increment) meteorologi-
cal distribution model (MicroMet; Liston and Elder
2006b). MicroMet is a data assimilation and interpola-
tion model that uses meteorological station datasets
and/or gridded atmospheric model, or analyses data-
sets. MicroMet minimally requires screen-height air
temperature, relative humidity, wind speed and direc-
tion, and precipitation data. The model uses known re-
relationships between meteorological variables and the
surrounding landscape (primarily topography) to dis-
bounce those variables over any given landscape in
physically plausible and computationally efficient ways.

OCTOBER 2008 LISTON AND HIEMSTRA 991
must be able to handle these relatively large computational domains given present-day computer resources.

The elemental snow processes are accumulation and ablation, and our assimilation scheme needs to account for these distinctly different phenomena. Consider, for example, a model-simulated seasonal snowpack evolution with well-defined accumulation and ablation seasons (Fig. 1). In addition, measured SWE at the end of the accumulation season and observed snow-free date are included. In this case, matching the model simulation to observations requires less modeled snow accumulation and faster modeled snowmelt.

Because precipitation \( P \) input is a dominant control on accumulation, we modify precipitation to obtain the required model constraint. Liston and Sturm (1998, 2002) used similar approaches to define the precipitation forcing required to reproduce observed arctic Alaska SWE distributions. The adjustment of precipitation is further justified for the snow-evolution case, because it is widely known that snow-precipitation measurements often contain significant errors (Yang et al. 1998, 2005; Liston and Sturm 2004). Because most snow-evolution models calculate a snowmelt rate, \( M \), this quantity will be modified to adjust modeled ablation.

Adjusting \( P \) to allow the model to more accurately simulate observed SWE (Fig. 1) requires the calculation of a precipitation adjustment factor, \( P_{\text{fact}} \). Thus,

\[
P_{\text{assim}} = P_{\text{fact}} P,
\]

where \( P_{\text{assim}} \) is the precipitation forcing required to simulate the observations. In the simplest terms,

\[
t_{-1} P_{\text{fact}} = \frac{\Delta \text{SWE}_{\text{obs}}}{\Delta \text{SWE}_{\text{mod}}} = \frac{\text{SWE}_{\text{obs}}^{t} - \text{SWE}_{\text{obs}}^{t-1}}{\text{SWE}_{\text{mod}}^{t} - \text{SWE}_{\text{mod}}^{t-1}}, \quad (2)
\]

where the subscripts obs and mod indicate observed and modeled SWE, respectively; \( t \) indicates the observation time; and the difference between \( t \) and \( t - 1 \) indicates the period between observations. If the denominator equals zero, the precipitation factor is set to unity. Similar precipitation corrections have been applied as part of other studies (Liston and Sturm 2002; Cherry et al. 2005).

Under the assumption that

\[
\sum_{t=1}^{f} P \approx \text{SWE}_{\text{mod}}^{t} - \text{SWE}_{\text{mod}}^{t-1},
\]

Eq. (2) becomes

\[
t_{-1} P_{\text{fact}} = 1 + \frac{(\text{SWE}_{\text{obs}}^{t} - \text{SWE}_{\text{mod}}^{t}) - (\text{SWE}_{\text{obs}}^{t-1} - \text{SWE}_{\text{mod}}^{t-1})}{\sum_{t=1}^{f} P},
\]

where if the denominator equals zero, the precipitation factor is set to unity. This precipitation adjustment factor is then substituted into Eq. (1), and the original precipitation \( P \) is modified to produce an adjusted precipitation used during the model assimilation cycle. As noted by Rutherford (1972), this approach is similar to Gandin (1963), in that differences between the observations and model simulations are used to determine required corrections.

Similarly, a melt correction factor, \( M_{\text{fact}} \), can be derived. In this case,

\[
\sum_{t=1}^{f} M = -(\text{SWE}_{\text{mod}}^{t} - \text{SWE}_{\text{mod}}^{t-1}), \quad (5)
\]

where \( M \) is the model simulated melt, and Eq. (4) becomes

\[
t_{-1} M_{\text{fact}} = 1 - \frac{(\text{SWE}_{\text{obs}}^{t} - \text{SWE}_{\text{mod}}^{t}) - (\text{SWE}_{\text{obs}}^{t-1} - \text{SWE}_{\text{mod}}^{t-1})}{\sum_{t=1}^{f} M},
\]

where if the denominator equals zero, the melt factor is set to unity. As with the precipitation adjustment, this is
used to adjust the original melt calculation over the observation interval following

\[ t^{-1}M_{\text{assim}} = t^{-1}M_{\text{fact}} t^{-1}M. \]  

Implementing this scheme in a snow-evolution model requires the ability to decide whether an accumulation or ablation correction will be performed. In SnowModel, this is done by calculating the relative contributions \( R \) of precipitation and melt during each observation interval using

\[
t^{-1}R_{\text{prec}} = \frac{\sum_{i=1}^{t} P}{\sum_{i=1}^{t} P + \sum_{i=1}^{t} M}, \quad \text{and} \quad t^{-1}R_{\text{melt}} = 1 - t^{-1}R_{\text{prec}}. \]

The greater of these defines whether precipitation or melt is corrected. During periods when there are no future observations, or when both the summed precipitation and melt are zero, the correction factor is defined to be unity. An additional requirement of our data assimilation system is that it be spatially distributed. This is described as part of our example simulations presented below.

4. Model simulations

a. Idealized examples

To illustrate the implementation and application of the above methodology, we performed MicroMet–SnowModel–SnowAssim simulations on an idealized test domain covering a 6 km \( \times \) 6 km area with a horizontal grid increment of 200 m (Fig. 2a). This domain was extracted from the CLPX Fraser, Colorado, mesocell study area (MSA) observation domain (Elder et al. 2008b) and was chosen to include both alpine (meadow) and subalpine (forest) vegetation types.

The 1 December 2002–15 May 2003 atmospheric forcing used to drive the model simulations (Fig. 3) was collected from the CLPX Fraser MSA and is typical of cold, windy, high-elevation sites in the Colorado Rockies. Over the spatial domain, we assumed there were three different observation sites representing commonly experienced snow data: a point or single model grid cell, a rectangular area representing a remote sensing pixel that would typically cover numerous model grid cells, and an arbitrarily shaped area resembling a remote sensing swath that would also include numerous model grid cells (Fig. 2b). Hypothetical SWE observations at these three sites were assigned and assumed to be measured 3 times during the winter (Table 1).

To illustrate the utility of the data assimilation procedure, we performed four different model simulations...
Table 2. A summary of the four example model simulations that were performed. See Fig. 2a for the vegetation distributions and Fig. 2b for the site locations and configurations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation 1</td>
<td>No data assimilation</td>
</tr>
<tr>
<td>Simulation 2</td>
<td>Assimilation using only site A</td>
</tr>
<tr>
<td>Simulation 3</td>
<td>Assimilation using site A, site B, and site C</td>
</tr>
<tr>
<td>Simulation 4</td>
<td>Assimilation using site A and site B over the alpine meadow and site C over the coniferous forest, then merging the results to cover the entire domain</td>
</tr>
</tbody>
</table>

Our implementation of the data assimilation scheme...
also provides the ability to stratify or divide observations to isolate data assimilation influences. For example, if the data assimilation process is being used to correct model deficiencies, it is reasonable to assume that model-simulated processes are different in the forest and alpine areas and that errors in the two regions will also be different. Likewise, observations in the forested and nonforested areas may also have dissimilar error characteristics.

A stratified approach was used to calculate a separate correction factor distribution for the alpine vegetation area, using site A and site B data, and another distribution for the forest, using site C data. These two correction factors were merged to cover the complete simulation domain (Fig. 7a). The resulting SWE distri-
bution on 15 May 2003 is shown in Fig. 7b. Because the forest area had only one observation, the correction factor distribution was uniform across the forested part of the domain. For the forested areas, this produced a spatially uniform increase in SWE over that in the control simulation (Fig. 4a). For the alpine region, the general south–north correction factor increase is less than that in simulation 3, and there is a corresponding reduction in the simulated SWE gradient.
b. CLPX example

Winter snow evolution was simulated using MicroMet–SnowModel–SnowAssim from 1 October 2002 to 1 April 2003 for the CLPX Rabbit Ears MSA in north-central Colorado (Elder et al. 2008b; Liston et al. 2008b). For this mountainous domain, this simulation period spans the snow accumulation season, from the initial snow accumulation in the fall to the beginning ablation in the spring. The spatial domain covered a 30 km × 30 km area centered over each 25 km × 25 km CLPX MSA; we defined these larger areas to ensure that the model simulations and outputs always included MSA-related observations. In the discussions that follow, MSA refers to the 30 km × 30 km simulation areas. Model simulations were performed using a 30-m horizontal grid increment (1001 grid cells in each of the x and y directions, or ~1 000 000 grid cells). Topographical data (30-m horizontal and 1-m vertical resolution) for the simulation domain were obtained from the U.S. Geological Survey’s National Elevation Dataset. Vegetation data of identical horizontal resolution were obtained from the 1992 National Land Cover Data (Vogelmann et al. 2001) and reclassified to match SnowModel’s defined vegetation classes (Liston and Elder 2006a). Meteorological data used in the model simulations take two primary forms: meteorological station (Elder et al. 2008a) and atmospheric analysis (Liston et al. 2008a). A complete description of the Rabbit Ears MSA forcing datasets can be found in Liston and Elder (2006b), Elder et al. (2008a), and Liston et al. (2008a).

The Rabbit Ears MSA includes the flat uplands of Rabbit Ears Pass across the Park Range of north-central Colorado (Fig. 8a). The uplands are flanked to the east and west by steep valleys and the relatively flat terrain of the river valleys and basins below. It has minimum, maximum, and mean elevations of 2025, 3330, and 2678 m, respectively. The upland portion of the MSA contains low-to-moderate topographic relief that extends north–south through the middle of the MSA. In the upland areas, forest cover consists of spruce-fir, lodgepole pine, and aspen (Populus tremuloides) interspersed with broad glades and grassy meadows (Fig. 8b). Vegetation in the lower elevations to the east and west consists of shrublands and pastures. The upland area receives heavy snowfall during the winter and spring, often developing the deepest snowpacks in Colorado (Doesken and Judson 1996).

The data assimilation incorporated two kinds of observational datasets: 1) ground-based snow depth and density observations that occurred over three 1 km × 1 km MSA subdomains called intensive study areas (ISAs; Elder et al. 2008b), and 2) airborne measurements of gamma radiation emitted from earth and converted to a measure of snowpack SWE (Cline et al. 2008). CLPX included two intensive observing periods.
TABLE 3. Area-averaged snow depth, density, and SWE for the CLPX ISAs during IOP3 (20–25 Feb 2003) and IOP4 (26–31 Mar 2003).

<table>
<thead>
<tr>
<th>ISA</th>
<th>IOP3</th>
<th>IOP4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rabbit Ears ISA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffalo Pass</td>
<td>275.3</td>
<td>322.5</td>
</tr>
<tr>
<td>Spring Creek</td>
<td>204.4</td>
<td>209.3</td>
</tr>
<tr>
<td>Walton Creek</td>
<td>197.1</td>
<td>203.1</td>
</tr>
<tr>
<td>Mean depth (cm)</td>
<td>258.7</td>
<td>340.3</td>
</tr>
<tr>
<td>Mean density (kg m(^{-3}))</td>
<td>51.9</td>
<td>67.2</td>
</tr>
<tr>
<td>Mean SWE (cm)</td>
<td>71.2</td>
<td>109.8</td>
</tr>
</tbody>
</table>

During winter 2002–03: IOP3 (19–25 February) and IOP4 (25 March–1 April).

CLPX ground-based snow depth and density observations from IOP3 and IOP4 were used to estimate area-averaged SWE for each of the three ISAs. All snow depth measurements were averaged over each ISA, including zero-depth values resulting from observations made at snow-free observation points. With 684 potential depth observations randomly distributed in one hundred 100 m × 100 m cells over each 1-km\(^2\) ISA (Elder et al. 2008b), this mean depth value is assumed to serve as an effective surrogate for mean ISA snow depth and includes the effective snow-covered area. Mean density was calculated using all available snow pit densities for each ISA and IOP. The mean SWE for each ISA, IOP3, and IOP4 was calculated as the product of mean depth and percent (snow/water) mean density (Table 3). Additional details of the ground-based snow observations can be found in Elder et al. (2008b) and Liston et al. (2008b).

The gamma SWE observations were derived from measurements of natural terrestrial gamma radiation emitted from potassium \(^{40}\text{K}\), uranium \(^{238}\text{U}\), and thorium \(^{232}\text{Th}\) radioisotopes in the upper 20 cm of soil. As part of the measurement program, the radiation was measured from an aircraft flying 150 m above the ground. Terrestrial gamma rays are attenuated by the atomic cross section of intervening mass (e.g., water in any phase, vegetation, and air) between the soil and the aircraft. The measurement of the attenuation (i.e., the difference between measurements over dry soil and moist soil or snow) has been found to be a reliable basis for soil moisture and SWE measurements (Jones and Carroll 1983). CLPX used the NWS’s Gamma Radiation Detection System (GAMMA) to measure mean areal SWE over the Rabbit Ears MSA. The CLPX GAMMA datasets used in model simulation included SWE observations collected in series of flight lines within the MSA (Fig. 8a). Initial background terrestrial gamma radiation measurements were collected under relatively dry soil conditions, with no snow cover present in September 2001. Coincident ground observations of gravimetric soil moisture were collected to estimate the mean soil moisture along each flight line. This was used to calibrate the background radiation measurements, accounting for the attenuation effects of existing soil moisture. Subsequent measurements of terrestrial gamma radiation over the flight lines during the experiment were used to determine the attenuation of the radiation signal solely due to the intervening water mass in the snow and soil. SWE was determined by estimating the proportion of attenuation resulting from soil moisture using IOP soil moisture observations. The Rabbit Ears GAMMA SWE data used in the model simulations, IOP3 and IOP4, are listed in Table 4.

The model simulation results for IOP4 are presented in Fig. 9; IOP3 results are not shown. The patterns are similar, but the SWE magnitudes are reduced because IOP3 was earlier in the snow accumulation season. The ISA and GAMMA SWE observations were interpolated to the simulation grid, using MicroMet’s Barnes spatial interpolation scheme (Fig. 9a). To perform this interpolation, average SWE for each ISA and GAMMA flight line was defined to fall in the center of the ISA flight line, and the resulting point values were interpolated over the domain. The interpolated ISA and GAMMA SWE observations (Fig. 9a) include a north–south band of higher values through the center of the domain, with relatively lower values to the east and west. The highest SWE values in this center band are located in the northern part of the domain. For reference, the Rabbit Ears, IOP4, and ISA and GAMMA observation areas are included in Figs. 9a and 9d.

During model simulations, SnowModel performed an initial simulation over the entire period. The results of this simulation were compared against concurrent observations. Under the assumption that the differences between the model and observations are largely the result of errors in the precipitation fields, a precipitation correction was calculated. This correction can be different for each interval between available observations. The precipitation correction factor distributions (Fig. 9b) were then used to scale the precipitation fields in the initial simulation, and this new precipitation forc-
ing was used for a secondary and final simulation (Fig. 9c). The precipitation correction factor distributions mimicked observed SWE distributions (Fig. 9b). This essentially provided different corrections for the eastern area, central highlands, and western valleys of the simulation domain. The final SWE distribution displays a relatively finescale structure related to orographic precipitation, low-elevation melt, wind redistribution (snow drifts above tree line), and patterns associated with the vegetation distribution. Intuitively, the general distribution patterns simulated by the model (Fig. 9c) are considerably more realistic than those defined by observations alone (Fig. 9a; see Liston references contained herein), and it generally reproduced the regional pattern found in the observations (Fig. 9a).

Using the observation mask defined in Fig. 8a, the average modeled SWE (Fig. 9c) corresponding to each ISA and GAMMA flight line was calculated for the initial (Fig. 10a) and second (Fig. 10b) model run outputs. Thus, Fig. 10 compares the modeled and observed SWE for each Rabbit Ears observation area, with and without the data assimilation. Both the square of the linear correlation coefficient $r^2$ and the root-mean-square error (rmse) have improved. Although not discussed herein, a precipitation preprocessing step was applied by Liston et al. (2008b) to make sure the CLPX MSA initial model simulations already closely fit the observed fields; without this step, the data assimilation procedure yields a much larger improvement (not shown). Differences between the simulated and observed values (modeled minus observed; Table 4) were also interpolated across the simulation domain using the Barnes scheme (Fig. 9d). The largest difference occurred in the northeast area of the domain and is influenced by SWE values associated with GAMMA line #311. The model simulation (Fig. 9c) included several deep drifts in the northeast corner that were underneath flight line #311. This yielded corresponding modeled SWE values that were nearly 3 times greater than the GAMMA observations (Table 4). Either the model overestimates snowdrift SWE, and/or the GAMMA system underestimates SWE in very deep drifts because the system saturates [e.g., Carroll and Carroll (1989) used GAMMA to measure 60 cm SWE in forests with errors of approximately 12%; other NOHRSC efforts indicate any gamma radiation emitted through more than 70–90 cm SWE is minimal (T. R. Carroll and D. W. Cline 2007, personal communications)].

5. Discussion

The modeling system and the associated model simulations represent a combined approach in which SnowModel defined the spatial SWE patterns, and the observations and the associated data assimilation system reconciled the SWE magnitudes. It is rare that snow observations are of sufficient resolution to capture the spatial variability present in the natural system. Available snow observations generally take two forms: ground-based observations and remotely sensed observations; examples of these (Figs. 2b and 8a) were included in our simulations. Ground observations usually provide point values (e.g., site A in Fig. 2b) and examples include the Natural Resources Conservation

<table>
<thead>
<tr>
<th>Observation ID</th>
<th>Easting (m)</th>
<th>Northing (m)</th>
<th>IOP3 observed SWE (cm)</th>
<th>IOP3 difference SWE (cm)</th>
<th>IOP4 observed SWE (cm)</th>
<th>IOP4 difference SWE (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffalo Pass ISA</td>
<td>357 576</td>
<td>4 488 391</td>
<td>71.2</td>
<td>0.8</td>
<td>109.8</td>
<td>-24.8</td>
</tr>
<tr>
<td>Spring Creek ISA</td>
<td>351 058</td>
<td>4 487 951</td>
<td>51.9</td>
<td>1.7</td>
<td>67.2</td>
<td>-4.1</td>
</tr>
<tr>
<td>Walton Creek ISA</td>
<td>360 145</td>
<td>4 473 579</td>
<td>50.4</td>
<td>-0.7</td>
<td>66.9</td>
<td>-6.6</td>
</tr>
<tr>
<td>303</td>
<td>345 190</td>
<td>4 470 221</td>
<td>-</td>
<td>-</td>
<td>16.3</td>
<td>-0.7</td>
</tr>
<tr>
<td>304</td>
<td>344 533</td>
<td>4 472 145</td>
<td>-</td>
<td>-</td>
<td>16.5</td>
<td>3.4</td>
</tr>
<tr>
<td>305</td>
<td>346 238</td>
<td>4 472 443</td>
<td>-</td>
<td>-</td>
<td>17.5</td>
<td>-4.9</td>
</tr>
<tr>
<td>306</td>
<td>353 939</td>
<td>4 467 181</td>
<td>-</td>
<td>-</td>
<td>32.0</td>
<td>7.5</td>
</tr>
<tr>
<td>308</td>
<td>353 755</td>
<td>4 485 596</td>
<td>-</td>
<td>-</td>
<td>34.0</td>
<td>28.5</td>
</tr>
<tr>
<td>309</td>
<td>351 275</td>
<td>4 490 443</td>
<td>-</td>
<td>-</td>
<td>30.5</td>
<td>19.0</td>
</tr>
<tr>
<td>310</td>
<td>362 298</td>
<td>4 488 299</td>
<td>-</td>
<td>-</td>
<td>32.5</td>
<td>28.9</td>
</tr>
<tr>
<td>311</td>
<td>366 324</td>
<td>4 488 870</td>
<td>-</td>
<td>-</td>
<td>21.8</td>
<td>38.0</td>
</tr>
<tr>
<td>312</td>
<td>354 381</td>
<td>4 475 079</td>
<td>-</td>
<td>-</td>
<td>27.7</td>
<td>23.7</td>
</tr>
<tr>
<td>313</td>
<td>361 927</td>
<td>4 470 916</td>
<td>-</td>
<td>-</td>
<td>51.3</td>
<td>-1.8</td>
</tr>
<tr>
<td>317</td>
<td>368 808</td>
<td>4 477 388</td>
<td>-</td>
<td>-</td>
<td>31.0</td>
<td>12.4</td>
</tr>
<tr>
<td>319</td>
<td>366 424</td>
<td>4 477 397</td>
<td>-</td>
<td>-</td>
<td>32.0</td>
<td>15.6</td>
</tr>
</tbody>
</table>
Service (NRCS)’s western SNOTEL automated snowpack data collection system (Carroll 1997) and transects such as snow courses measured by NRCS. Snow courses are permanent sites where manual measurements of snow depth and water equivalent are collected by trained observers. Measurements are usually taken monthly, starting at the beginning of the calendar year and continuing through the melt season, along courses that are approximately 300 m long and situated in small, wind-sheltered meadows.

Remote sensing observations usually cover specific areas (e.g., site B and site C in Fig. 2b and flight lines in Fig. 8a) and may coincide with numerous model grid cells, for example, the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) instrument on the NASA EOS Aqua satellite provides...
global passive microwave SWE measurements on 25-km Equal-Area Scalable Earth (EASE) grids (Kelly et al. 2004). In addition, MODIS on NASA’s Terra satellite is used to provide global snow-extent products on, for example, 500-m grids (Hall et al. 2006). Because MODIS is unable to discern snow concealed by clouds, temporal discontinuities exist within the datasets. As another snow remote sensing example, the NOHRSC runs an operational Airborne Gamma Radiation Snow Survey Program that makes winter airborne SWE measurements over a network of more than 1900 flight swaths in 25 states and seven Canadian provinces (NOHRSC 2004).

Because of the limited number of observations in our example simulations (and in most real-world snow-observation programs), a simple interpolation (Figs. 4b and 9a) of the observations frequently does not mimic real-world snow distributions; improved distributions may be provided by the model (cf. Liston et al. 2008b). As part of the example simulations presented herein, we have assumed that the model (Figs. 4a and 9c) represents naturally occurring spatial variability; but in reality, there are still governing processes that exist at smaller scales than the model’s 200-m, or even 30-m, grid increments (Elder et al. 1991; Blöschl 1999; Liston 2004). A methodology is needed by which model simulations can be corrected using available observations, while preserving spatial patterns produced by snow interactions with terrain, wind, and vegetation. In addition, because the observations are often temporally discontinuous and frequently offer only a brief snapshot of snow conditions, a methodology is required that can realistically improve the model simulation for time periods prior to the measurement(s). As highlighted by Figs. 5–7, and 9, the methodology introduced herein accomplishes these goals.

\subsection*{a. Applications}

In some applications, realistic snow distributions are only required at specific points in time. For example, successful snowmelt runoff calculations generally require knowledge of SWE distribution before net ablation occurs (not considering critical soil freeze–thaw conditions; Zhang 2005). In contrast, realistic SWE distributions throughout the winter are required for SWE distribution-dependent winter surface energy flux calculations. As an extreme example, consider the case in which only end-of-winter (i.e., before spring melt) SWE distribution observations are available but realistic SWE distributions for each day throughout the winter snow season are required for some land surface climate interaction application. In this case, the methodology presented herein would potentially produce the required winter SWE evolution. In addition, other available SWE data would be used to constrain the simulation during the winter accumulation and spring melt periods. Here we provided a real-world example, the MicroMet–SnowModel–SnowAssim application, in which the modeling system was used to simulate 30-m resolution, snow water equivalent distributions over the CLPX Rabbit Ears’ 30 km × 30 km study area in Colorado.

**Fig. 10.** Comparison of Rabbit Ears MSA, IOP4, and modeled and observed SWE over the ISA and GAMMA observation masks for the (a) initial model run without data assimilation and (b) second model run with data assimilation. Included are the square of the linear correlation coefficient \( r^2 \) and rmse.
This methodology has the benefit of being conceptually and computationally simple; one of our example applications had 10^6 grid cells. Observations are generally not error free, and the Barnes analysis objectively determines the degree to which correction fields fit the observed data. To some degree, this accounts for data outliers because the resulting surface is a relatively smooth representation of the data distribution. In our model implementation, the degree of smoothing can be overridden by modifying the influence radius used in the analysis. For conditions in which different simulation subdomains are dominated by distinctly different processes (e.g., forested and nonforested regions), the model can be configured to calculate separate correction distributions. In addition, to account for different levels of uncertainty and observation error between in situ and remotely sensed data, user-prescribed weights can be defined to control the relative influence of the various observations. These weights can take into account information such as knowledge of the data collection instruments and methods, quality control and quality assurance measures, and statistical quantities. The same weighting procedures can be used to control the relative influence of the observed and model-simulated fields on the final assimilated distributions.

The data assimilation methodology relies on the land surface model to account for the physical evolution of the snow cover. For example, in the simulations presented herein, SnowModel accounted for snow-evolution processes such as blowing-snow redistribution, sublimation, forest–canopy snow interception, snowmelt, and so on. In addition, the general method can be easily applied to other variables or models. For example, if the user is interested in soil moisture, the evaporation or transpiration flux could be used as the adjusted variable. The methodology can also be applied to much coarser-resolution modeling systems, such as those represented by global or regional atmospheric and hydrologic models. It can also be applied to models designed to represent subgrid processes and distributions within these coarse resolution models (e.g., Liston 2004).

b. Limitations

Depending on the modified variable, this methodology may or may not conserve mass and energy. In our example application, the simulations conserved mass under the assumption that it was acceptable to increase or decrease precipitation inputs. In contrast, our melt rate increases or decreases were performed without regard to whether that energy was actually available or what happened with any extra energy not used for melting. Also, in our particular example application, the Eq. (3) assumption requires that precipitation inputs exert a dominant control on the SWE distribution. Whereas this may be generally true, it is not always the case. For example, Liston and Sturm (2004) note that snow sublimation in katabatic flow regimes can be comparable to the precipitation inputs. In places or conditions where precipitation does not dominate the SWE distribution, another variable could be adjusted.

As with most data assimilation schemes, this method provides no simulation improvement without observations. In contrast to most assimilation schemes, it does provide a retrospective analysis that corrects the simulation prior to the availability of an observation. This analysis comes at the price of running the model twice over the same temporal domain, thereby doubling the computational time compared with the control simulation.

6. Conclusions

In light of the role that snow cover plays in modifying weather, climate, hydrologic, and ecologic features, recent land surface hydrology and atmospheric modeling efforts are working to improve their representations of seasonal snow evolution. In addition to these modeling efforts, snow-related observational datasets, such as those provided through remote sensing, are becoming increasingly available and with improved reliability. Data assimilation methodologies are needed to merge observations and models. We have presented a simple data assimilation scheme, suitable for reanalysis applications, that provides the ability to move model-simulated snow evolution closer to observational datasets. In addition, the implemented retrospective corrective scheme allows improved simulation during periods prior to assimilation of the observation. The resulting snow-evolution datasets are available for model simulations requiring improved realism in both space and time, throughout the snow season.

Although we have focused on snow-related variables, the same techniques could be applied to other variables as part of other applications. In addition, the general methodologies are adaptable for use in a wide range of terrestrial modeling systems, including those focusing on land surface hydrology and ecosystem processes.

Acknowledgments. The authors would like to thank Svetlana Berezovskaya and Kelly Elder for their thorough and insightful reviews of this paper. This work was supported by NASA Grants NAG5–11710, NNG04GP59G, and NNG04HK191; NOAA Grant NA17RJ1228; and National Science Foundation Grants OPP-0229973 and ARC-0629279.
REFERENCES


——, 1973: Mesoscale objective analysis using weighted series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK 73069, 60 pp. [NTIS COM-73-10781].


