A physiographic approach to downscaling fractional snow cover data in mountainous regions

Reggie D. Walters [a], Katelyn A. Watson [a], Hans-Peter Marshall [a,b], James P. McNamara [a], Alejandro N. Flores [a,*]

[a] Department of Geosciences, Boise State University, 1910 University Drive, Boise, ID 83725, United States
[b] Center for Geophysical Investigation of the Shallow Subsurface (CGISS), United States

ABSTRACT

Accurate characterization of snow-covered area (SCA) and snow water equivalent (SWE) in complex terrain is needed to improve estimation of streamflow timing and volume, and is important for land surface modeling. Direct field observations of SWE, SCA and atmospheric forcing inputs for models of snow accumulation and ablation are typically sparsely sampled in space. Satellite imagery is, therefore, a critical tool for verification and confirmation of snow model estimates of SCA. The Landsat system provides snow-covered area estimates at a spatial resolution of 30 m with a 16-day return interval, while daily estimates of SCA and fractional SCA (f_{SCA}) are available at 500 m from the Moderate Resolution Imaging Spectroradiometer (MODIS). This study describes and tests a linear model to downscale MODIS MOD10A1 f_{SCA} (500 m) data to higher-resolution (30 m) spatially explicit binary SCA estimates. The algorithm operates on the assumption that two variables, potential insolation and elevation, control differential ablation of snow cover throughout spring melt at 30 m to 500 m scales. The model downscales daily 500 m f_{SCA} estimates from MODIS to provide daily SCA estimates at a spatial resolution of 30 m, using limited Landsat SCA for calibration and independent Landsat SCA estimates for validation. Downscaled SCA estimates demonstrate statistically significant improvement from randomly generated model ensembles, indicating that insolation and elevation are dominant factors controlling the snow cover distribution in the semi-arid, mountainous region in southwestern Idaho, USA where this study is performed. Validation is performed with Landsat data not used for calibration, and is also performed using Landsat 500 m aggregate f_{SCA} instead of MODIS f_{SCA} as an ideal case. Downscaled estimates show reasonable accuracy (test metric outperforms random ensembles at $p = 0.01$ significance level for multiple ranges of snow cover) with only one calibrated parameter.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

In complex terrain, the characterization of snow-covered area (SCA) and snow water equivalent (SWE) at a resolution less than 500 m could substantially improve estimation of streamflow timing and volume, as SCA can vary over length scales much less than the resolution of data from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Anderson, McNamara, Marshall, & Flores, 2014). Since more than one-sixth of the world's population depends on seasonal snowmelt for water resource supply (Barnett, Adam, & Lettenmaier, 2005), predicting the spatiotemporal evolution of snow processes is of great importance for conveying reliable hydrologic information. Accumulation and melting of snow occur variably, producing heterogeneity in snowpack disappearance that must be modeled with accuracy in order to estimate melt runoff for a catchment (Clark et al., 2011). However, the ability to predict these variable snow processes is limited in part because variability occurs at length scales of less than 100 m, while in situ observation networks have a resolution several orders of magnitude larger and are confined to a relatively narrow elevation range in flat terrain (Bales et al., 2006; Martinec & Rango, 1981). In addition, while field sampling can be performed at the necessary spatial resolution, it is time consuming and costly and therefore is typically limited to small spatial extents and coarse temporal resolution (Elder, Dozier, & Michaelsen, 1991). Thus, satellite remote sensing observations are often employed in conjunction with simulation models to improve the estimation of snowpack states and resultant hydrologic fluxes. For example, studies assimilating satellite-derived areal snow cover information into hydrologic models have demonstrated improvements to simulated streamflow and SWE (Clark et al., 2005; Rodell & Houser, 2004; Thirel, Salamon, Burek, & Kalas, 2011). In other studies, snowmelt depletion curves have been accurately constructed using similar SCA information in combination with energy balance melt modeling (Homan, Luce, McNamara, & Glenn, 2011). Retrospective analysis of SCA data combined with distributed temperature-index and energy balance snowmelt modeling

* Corresponding author at: 1910 University Drive, MS 1535, Boise, ID 83725-1535, USA. Tel.: +1 208 426 2903; fax: +1 426 4061.

E-mail addresses: reggiewalters@u.boisestate.edu (R.D. Walters), katelynwatson@u.boisestate.edu (K.A. Watson), hpmarshall@u.boisestate.edu (H.-P. Marshall), jmcnamar@boisestate.edu (J.P. McNamara), lejoflores@u.boisestate.edu (A.N. Flores).

has been used to reconstruct basin-wide SWE at the time of maximum accumulation, comparing favorably with results of intensive field campaigns (Cline, Bales, & Dozier, 1998; Durand, Molotch, & Margulis, 2008; Martinec & Rango, 1981; Molotch, 2009). This approach depends on distinct accumulation and ablation periods, however, and therefore may provide accurate peak SWE distribution only under certain conditions. Mid-winter rain and melt, as well as late spring snowfall can cause problems with this approach, as the technique assumes ablation dominates and albedo is constant between SCA estimates.

The space-borne Landsat remote sensing system is capable of retrieving snow-covered area and albedo data for hydrologic studies at the catchment scale (Dozier, 1989; Dozier & Marks, 1987; Rosenthal & Dozier, 1996). Similarly, the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument, aboard NASA Aqua and Terra satellites, can be used to map SCA at a much higher temporal but lower spatial resolution (Hall & Riggs, 2007; Hall, Riggs, & Salomonson, 1995; Justice et al., 1998; Painter et al., 2009; Riggs, Hall, & Salomonson, 1995; Salomonson & Appel, 2004, 2006). These products are highly valuable for their utility in updating and constraining distributed snow models (e.g., Clark et al., 2005; Luce, Tarboton, & Cooley, 1998, 1999; Thirel et al., 2011). However, each of these snow cover products has spatial or temporal limitations. For instance, Landsat has a relatively high spatial resolution of 30 m with a 16-day return interval under ideal, cloud-free conditions. Conversely, estimates from MODIS can be derived at a 500 m spatial resolution on a daily basis. The fractional snow-covered area (f_{SCA}) product derived from MODIS (Hall & Riggs, 2007; Riggs et al., 1995; Salomonson & Appel, 2004, 2006) provides a sub-grid approximation by estimating the percentage of each pixel that is snow-covered, but does not explicitly resolve SCA at sub-pixel scales. The MODIS f_{SCA} product has led to significant improvements in ablation modeling at coarse resolutions (e.g. Yatheendradas et al., 2012). Many model applications, however, require snow cover information at finer resolutions. For example, hydrologic models in mountainous complex terrain commonly adopt the 30 m resolution of readily available digital elevation models. At this scale, knowledge of the percentage of a 500 m pixel that is snow-covered is of value only in understanding basin-scale trends in snow cover. Whereas Landsat offers a spatially finer resolution product, the temporal resolution is not sufficient. Since snow cover often varies extensively within 500 m and over 2 weeks, it is desirable in many applications to have snow cover information at the spatial resolution of Landsat and the temporal resolution of MODIS.

The objective of this work is to develop and describe an efficient approach to downscaling melt-season fractional snow-covered area (f_{SCA}) data from MODIS (spatial resolution 500 m) to a higher-resolution (spatial resolution 30 m), yielding a spatially explicit SCA estimate at 30 m resolution. The derived high-resolution snow cover product is meant to be used to constrain future snow cover simulations. The proposed model is based on the hypothesis that the distribution of snow-covered area in a partially snow-covered region is non-random and can be predicted using terrain physiographic features (elevation, slope, and aspect). Further, since these terrain features are relatively constant year to year, snow distribution patterns are assumed to occur in similar patterns from year to year in agreement with observations (Sturm & Wagner, 2010). The algorithm we develop is based on physiographic characteristics that can be derived from ancillary data products, principally digital elevation models (DEMs). The algorithm assigns binary snow cover to a grid co-registered with a 30 m DEM that is used to derive normalized potential incoming solar radiation (insolation) and normalized relief within each 500 m MODIS pixel. The method preserves the predicted snow cover fraction at the 500 m scale. We calibrate and test the model using 13 Landsat images for a region in southwestern Idaho. The proposed approach assumes that potential insolation and elevation control the spatial distribution of snow cover at the sub 500 m scale. A similar approach could be developed for other factors controlling the distribution of snow cover in regions in which SCA is controlled by other processes.

In Section 2 we describe the satellite and terrain datasets used in this study. Section 3 outlines algorithm design, development, and procedures used for calibration and validation. Model results are presented in Section 4 and we provide a brief synopsis of the results, implications, and model limitations in Section 5.

2. Datasets

2.1. Remotely sensed snow cover: hillslope scale

The Landsat Program has provided multispectral, high-resolution data observations across the entire globe for over 40 years, offering a unique retrospective and near real-time data record for many applications. Multispectral band information from the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) instruments is often used for automated mapping of snow cover. The Normalized Difference Snow Index (NDSI) has been used in efforts to distinguish snow-covered pixels from other land surfaces, leveraging the large difference in reflectance of snow in the visible and shortwave infrared portions of the electromagnetic spectrum (Dozier, 1989). This ratio is described as

$$NDSI = \frac{R_{vis} - R_{swir}}{R_{vis} + R_{swir}} \qquad (1)$$

where R_{vis} represents reflectance in a visible band and R_{swir} is reflectance in a short-wave infrared band. These correspond to Landsat TM bands 2 and 5, respectively. The NDSI is an analog to the ubiquitously used Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) which utilizes similar principles to estimate vegetation properties. Other studies have exploited spectral mixture analyses to classify snow-covered and snow-free pixels, utilizing spectral libraries for pure end-member reflectance values and solving a set of linear combinations of their relative fractions for the observed reflectance in each pixel (e.g., Nolin, Dozier, & Mertes, 1993; Painter, Dozier, Roberts, Davis, & Green, 2003; Rosenthal & Dozier, 1996).

In this study, a series of 13 Landsat scenes over a mid-latitude, semi-arid region in southwestern Idaho (path/row 41/30) are compiled over a range of snowmelt season dates (January to May) between 2000 and 2011. We employ the NDSI to estimate binary snow coverage. These scenes serve as high-resolution calibration and validation data for the development of the downscaling routine. A combination of Landsat TM and ETM+ scenes are chosen in which cloud cover is minimal (i.e. <20% for whole scene) and qualitatively inspected such that any perceived cloud cover does not occur over the mountainous regions of interest (Table 1). Within the Landsat scenes, we chose subsets known to retain seasonal snow cover for calibration and validation regions and that are of interest for modeling exercises (Fig. 1). Subset (*a*) in Fig. 1 contains the Dry Creek Experimental Watershed (DCEW), a 27 km² watershed north of Boise

Table 1
Image calibration dates and sensors utilized with corresponding cloud coverage for the Landsat scenes used in this study. All scenes are path 41, row 30.

Scene #	Date	TM 5	ETM+ 7	Cloud cover %
1	02/03/2000	x		0
2	02/19/2000		x	0
3	03/01/2001		x	0
4	04/18/2001		x	4
5	05/04/2001		x	0
6	02/16/2002		x	2
7	03/04/2002		x	2
8	04/08/2003		x	6
9	05/10/2006	x		0
10	04/27/2007	x		0
11	03/12/2008	x		20
12	05/15/2008	x		13
13	02/01/2011	x		6
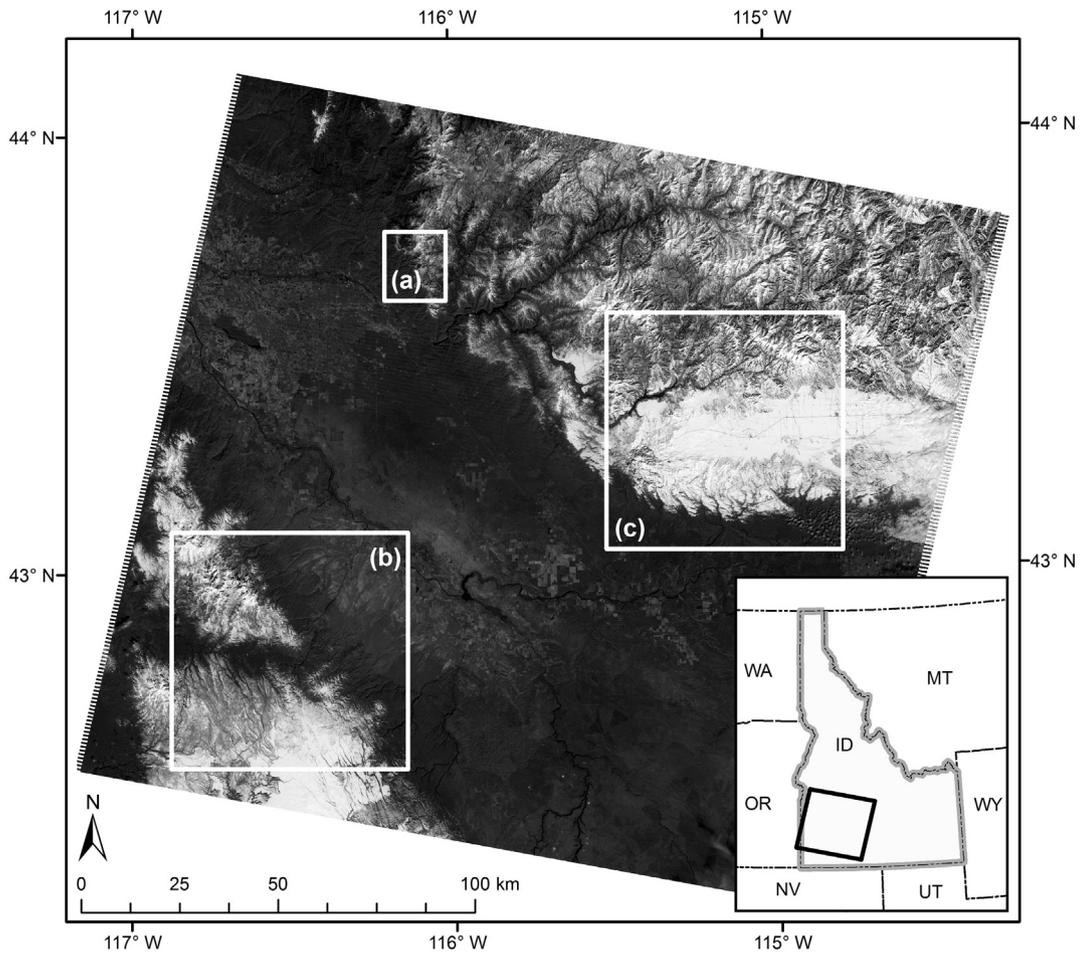
Fig. 1. Landsat band 2 image (path/row: 41/30) for 18 March 2010 with SW Idaho location inset. The model is calibrated over subset *a* and validated over subsets *b* and *c*.

(Graham, Barnard, Kavanagh, & Mcnamara, 2012; Kelleners, Chandler, McNamara, Gribb, & Seyfried, 2010; Kunkel, Flores, Smith, McNamara, & Benner, 2011; McNamara, Chandler, Seyfried, & Achet, 2005; Stieglitz et al., 2003; Tyler et al., 2008), while subset (*b*) contains Reynolds Creek Experimental Watershed (RCEW) and Critical Zone Observatory, a 239 km^2 watershed in the Owyhee mountains that is maintained by the USDA Agricultural Research Service (Flerchinger, Marks, Reba, Yu, & Seyfried, 2010; Johnson & Hanson, 1995; Marks, Link, Winstral, & Garen, 2001; Reba et al., 2011; Winstral & Marks, 2002). Methods for classification of Landsat snow-covered pixels are described in further detail in Section 3.

2.2. Remotely sensed snow cover: MODIS scale

MODIS MOD10A1 Level 3 Version 5 (Hall and Riggs, 2007; Riggs et al., 1995) data are obtained for downscaling. This product is a daily retrieval of f_{SCA} based on data in visible and short-wave infrared bands from the MODIS sensor onboard the NASA Terra satellite. The MOD10A1 product produces f_{SCA} estimates based on an empirical linear relationship between MODIS NDSI and pixel snow fraction, inferred from Landsat data using an NDSI threshold approach. A similar product is available from Aqua (MYD10A1), however, we choose to use the Terra version in this study because Terra f_{SCA} estimates performed better than Aqua with lower RMSE values found during a validation study; this is largely due to a pixel misregistration issue between NDSI bands in the Aqua MODIS instrument (Salomonson & Appel, 2006). Recent work has suggested that the MOD10A1 fractional snow cover potentially overestimates snow cover fraction in some regions, particularly in North America (Rittger, Painter, & Dozier, 2012). A more recent f_{SCA}

product, the MODIS Snow-Covered Area and Grain Size (MODSCAG) model, utilizes spectral mixture analysis with MODIS reflectance data to estimate f_{SCA} and has demonstrated promising results in validation (Painter et al., 2009; Rittger et al., 2012), but these data are not readily available to the general scientific community at this time. Since the purpose of this study is not to evaluate the skill of a particular remote sensing product, we use Terra MOD10A1 data for this project. Future work may explore the utility of the downscaling approach with other snow remote sensing products.

2.3. Terrain data

We use Shuttle Radar Topography Mission (SRTM) elevation data acquired from the Global Land Cover Facility. This dataset was acquired by the Space Shuttle Endeavour on mission STS-99 during February 2000 and provides nearly global coverage at 30 m resolution. The WRS-2 tile edition of the 30 m DEM, which is co-registered with the Landsat scene of interest was obtained, substantially easing comparisons between modeled and observed Landsat snow-covered/snow free pixels during the calibration and validation phases of the study. Local topographic slope (gradient in direction of steepest descent) and aspect (cardinal direction of steepest descent) are extracted using the topographic modeling tools in ENVI resulting in 30 m grids with identical spacing and extent to that of the DEM. A mountain slope solar radiation algorithm (Swift, 1976) is applied to these terrain data in MATLAB producing indices for daily, integrated potential incoming solar radiation (insolation) at 30 m resolution. Details concerning the radiation computation and treatment are discussed in the methodology section of this paper.

3. Downscaling methods

In this section we describe the downscaling model used to derive 30 m binary snow cover maps using 500 m f_{SCA} estimates from MODIS. The developed algorithm makes several important assumptions about drivers of variable snow cover in the study region, which are enumerated here:

- (1) Two variables, potential insolation and elevation, dominantly control the differential ablation of snow cover throughout spring melt at 30 m to 500 m scales (Anderson et al., 2014).
- (2) Snow will disappear from pixels prone to increased solar radiation exposure before those that are more obscured.
- (3) Higher elevation pixels will retain a deeper snow pack due to orographic effects during initial accumulation and temperature lapse rates during melt onset. This effect should be preserved during ablation, as lower elevation pixels melt before higher ones.

The degree to which these predictor variables affect the snow cover distribution is a question that is addressed by this study and is likely a function of the spatial resolution at which the f_{SCA} observation occurs (e.g. 500 m in the case of individual MOD10A1 grid cells).

This section is organized as follows: first, the calculation of the solar radiation index is described. Next, the derivation of the 30 m binary snow-covered maps from the Landsat scenes shown in Table 1 is discussed. Reprojection and subsetting of the MODIS data for the region is presented. The downscaling routine is then described, followed by the parameter calibration and validation methods.

3.1. Terrain influence: normalized solar radiation index and elevation

To calculate solar radiation, we employ a simple algorithm requiring Julian date, latitude, slope inclination and aspect as inputs to estimate daily potential integrated irradiation at each DEM pixel (Swift, 1976). We extract the terrain slope and aspect data from a 30 m DEM. The radiation model does not account for atmospheric attenuation due to changes in optical depth or aerosol presence, making only a series of trigonometric adjustments to the extraterrestrial solar constant. We are interested in the relative comparison of insolation values over complex terrain within each MODIS pixel and thus do not attempt to adjust for these effects as they would be nearly constant at the 500 m scale. In order to achieve a relative value for solar radiation, we normalize the potential insolation for each MODIS pixel to that of a corresponding horizontal plane at the same location,

$$f_{sl} = \frac{R_{slope}}{R_{horiz}} \tag{2}$$

where R_{slope} and R_{horiz} are pixel-scale potential insolation on the local slope and horizontal plane, respectively, and f_{sl} is the normalized term referred to as the slope factor. We compute f_{sl} for every Julian date over each subset domain depicted in Fig. 1. For the computed time-series of f_{sl} over each subset region (Fig. 1), the minimum and maximum slope factors over the entire season, f_{sl}^{max} and f_{sl}^{min}, are stored for each domain.

A normalized slope factor is next computed within each MODIS pixel. Rather than normalizing to the extreme values within each MODIS pixel, we normalize to the maximum and minimum calculated slope factors across the domain for the entire season,

$$f_{norm} = \frac{f_{sl}}{f_{sl}^{max} - f_{sl}^{min}} \tag{3}$$

where f_{sl}^{max} is the maximum slope factor for the whole season, and f_{sl}^{min} is the respective minimum. Since there is a 30 m pixel in every domain that is sufficiently steep and North-facing that there is zero potential direct irradiance at least one day of the year, the minimum observed slope factor, f_{sl}^{min}, for each domain is zero. Thus, Eq. (3) reduces to

$$f_{norm} = \frac{f_{sl}}{f_{sl}^{max}} \tag{4}$$

Because the slope factor is normalized to the value for a horizontal surface, f_{sl}^{max} is largest in the winter when sun angles are small. f_{norm} therefore has a large range in winter, and a much smaller range in late spring, capturing the dynamic range of solar declination angles throughout the year and the distribution of slopes and aspects in a given region of complex terrain. This serves to capture the intra-seasonal variations in solar declination and, therefore, in insolation as a driver of ablation. The range of values of f_{norm} within each MODIS pixel will thus be narrower than [0,1] on any given date except for the time and location(s) containing the maximum seasonal slope factor, f_{sl}^{max}.

Elevation for each 30 m pixel is normalized to the maximum range of values observed in a given 500 m MODIS grid cell throughout the domain of interest. This process standardizes the elevation variable to represent the degree of ruggedness or flatness within a MODIS cell. For the MODIS cell(s) containing this maximum range, the normalized values will span the [0,1] interval. All other 500 m cells will have a narrower interval. Elevation is normalized in a reverse fashion such that the minimum observed elevation has a greater value and the maximum elevation pixel takes on a lower normalized value:

$$z_{norm} = \frac{z_{max} - z}{z_{max}^R} \tag{5}$$

where z_{norm} is the variable name assigned to the normalized elevation grid, z_{max} represents the maximum elevation within the MODIS cell, and z_{max}^R represents the maximum 500 m window range throughout the domain.

3.2. Landsat binary grid processing

An empirical method is used to identify snow-covered pixels within Landsat subset grids. We employ the NDSI with a threshold criteria test similar to the SNOMAP algorithm (Hall et al., 1995). NDSI is computed via Eq. (1) with R_{vis} and R_{swir} corresponding to Landsat bands 2 and 5, respectively. Pixels are classified in binary manner according to the following criteria:

$$s_o = \begin{cases} snow, & NDSI \geq 0.4 \\ no\ snow, & otherwise \end{cases} \tag{6}$$

where s_o represents the classified state of the observed pixel. Dozier (1989) suggests other criteria for shadowed areas and cloud discrimination, but these are neglected here; we rely on (1) the NDSI to have minimal influence of viewing geometry over complex terrain and (2) our qualitative manual cloud cover selection criteria.

3.3. MODIS grid processing

We subset MODIS MOD10A1 data (Riggs et al., 1995) over the study region and re-project from its native Sinusoidal grid into a UTM coordinate system using the MODIS Reprojection Tool (MRT) from the NASA Land Processes Distributed Active Archive Center (LP-DAAC). For simplicity, we remap the 500 m MODIS f_{SCA} product onto a 30 m resolution grid that is co-registered with the SRTM DEM. It is on this geospatial template that the binary snow-covered classification is assigned. We use a nearest-neighbor scheme to register the MODIS f_{SCA} grid to the resolution and position of the DEM. The 30 m resolution grids, which include the remapped MODIS f_{SCA}, Landsat SCA, slope factor (f_{sl}), and elevation (z), are cropped to the MODIS grid boundaries, eliminating a small number of pixels at the grid margins. Because the boundaries of

the reprojected MODIS grid do not perfectly align with the boundaries of the SRTM DEM and Landsat multispectral images, there is an approximate geolocation offset of 9.96 m between the grids. It should be noted that this offset is well below the MODIS geo-location uncertainty of approximately 50 m at nadir (Wolfe & Nishihama, 2009). Results depicting differences between MODIS f_{SCA} and the f_{SCA} at 500 m computed from Landsat SCA show approximately zero mean and normally distributed disparities, suggesting that the grid offset does not produce systematic errors that would influence the calibration of the downscaling routine.

Although a MODIS snow cover grid is produced for each day, some regions and days are frequently corrupted by cloud cover and poor sensor viewing geometry. To combat this issue, temporal smoothing splines can be employed to estimate snow cover fraction during periods with data gaps (e.g., Dozier and Frew, 2009). In this study, we execute preliminary downscaling on a MODIS f_{SCA} image that temporally coincides with 30 m data from a clear Landsat overpass (18 March, 2010) which is used for preliminary validation. Since both Landsat and MODIS obtained cloud-free observations on this same date, no temporal smoothing is necessary.

3.4. Downscaling routine

We achieve a standardized grid at each MODIS pixel by combining the normalized insolation and elevation with a weighted average, using a constant weight across the domain, to produce a "terrain score":

$$T_s = w \cdot X_1 + (1-w) \cdot X_2 \tag{7}$$

where the matrix T_s represents a standardized grid representing the terrain score for which the constant weights w and $1-w$ represent the respective contributions from insolation (X_1) and elevation (X_2) to the snow cover distribution. The weight w determines the relative importance of insolation and elevation. Since the range of solar insolation is largest in winter, this parameter dominates the spatial distribution in the early season, and by the end of spring when solar angles are large and insolation is less spatially variable, elevation becomes more important. Within each MODIS pixel, the empirical cumulative distribution function (CDF) of the terrain score can be constructed. We then use

the retrieved MOD10A1 f_{SCA} value, which varies from zero to one, to invert the empirical CDF of T_s to determine a critical value of T_s that we take as the threshold for the occurrence of snow cover within the MODIS pixel. This critical value is designated as variable T_k. Specifically, snow cover is assigned to pixels in T_s with values below T_k. The remaining pixels, representing a fraction $1 - f_{SCA}$ within each MODIS pixel, are assumed to be snow-free. This approach preserves the original snow fraction observed by the individual MODIS pixel and thus preserves f_{SCA} at the basin-scale. The snow-mapping model is of the form:

$$s_m = \begin{cases} \text{snow,} & \int_0^{T_k} F(T_s) \\ \text{no snow,} & \int_{T_k}^{T_m} F(T_s) \end{cases} \tag{8}$$

where s_m is the classified state of the 30 m modeled pixel and T_m is the maximum value occurring in T_s. The overall schematic approach is illustrated in Fig. 2. Numerically, this algorithm is straightforward and computationally inexpensive since the only variable requiring solution is w.

This approach reflects our assumptions about the drivers of ablation in a couple of important ways. First, by normalizing the elevation pixels in reverse order, we assign greater values to pixels of lower elevation. Within each MODIS pixel boundary, the 30 m pixels with the lowest elevations will tend to be associated with higher values of T_s, and will therefore tend to be assigned a snow-free status before those pixels with higher elevations. In contrast to the reverse-normalizing process used with elevation, slope factors are normalized such that those pixels with the highest values of f_{sl} within a MODIS pixel also have the highest values of f_{norm}. Thus, those pixels with the highest values of f_{sl} (prone to relatively more solar radiation) will correspondingly have higher values of f_{norm}, which will tend to increase T_s. All else being equal, these pixels will tend to be classified as snow-free more frequently. Finally, by combining z_{norm} and f_{norm} in an efficient linear combination model, the approach requires only one parameter (w) to fit (since T_k is defined by MODIS f_{SCA} and w). This parameter is informed by Landsat data and represents the relative contributions of potential solar radiation and elevation within each MODIS pixel to the persistence of snow cover. The
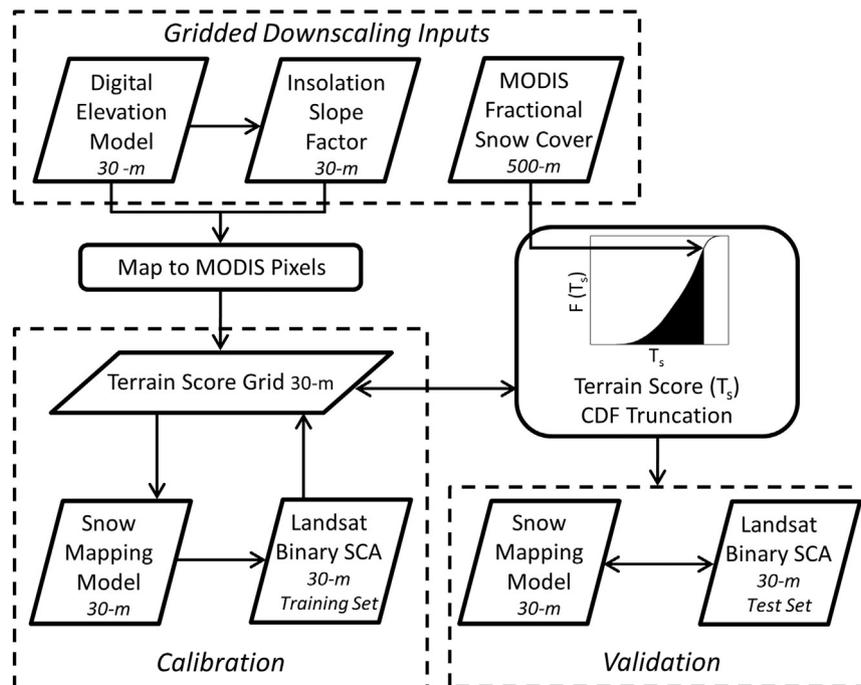
Fig. 2. Flow chart schematic outlining the fractional snow cover downscaling routine.

iterative calibration and validation methods are outlined in the next section.

3.5. Parameter calibration

As described, the downscaling model requires only an input value for the scalar weight w in Eq. (8) in addition to the z_{norm} and f_{norm} maps. We use an iterative approach to calibrate this parameter that optimizes the downscaled binary snow cover maps relative to the Landsat binary snow cover predictions. To test the transferability of parameters we use only region a in Fig. 1 to calibrate w.

During the calibration phase, the 500 m resolution f_{SCA} is computed directly from Landsat binary snow cover maps using an iterative, moving-window approach. This eliminates the impact of potential errors in the MODIS retrieval of f_{SCA} from influencing the calibration of the algorithm. The calibration approach iteratively moves a 500 m square window over three different co-registered 30 m resolution grids for region a: (1) Landsat SCA, (2) normalized elevation, and (3) normalized potential insolation. For robustness, we move the window across every possible 500 m window over the subset domain, rather than just those windows falling on the MODIS grid footprint. Within each window, the fractional snow cover is derived as the fraction of snow-covered pixels predicted with the critical NDSI threshold,

$$f_{SCA}^{Landsat} = \frac{n_s}{N} \qquad (9)$$

where $f_{SCA}^{Landsat}$ is the fractional snow cover inferred from the Landsat scene within the 500 m window, n_s represents the number of pixels where snow cover is observed, and N represents the total number of pixels in the window. By default, the downscaling model performs perfectly when f_{SCA} equals zero or one, regardless of the value of w. Therefore, we analyze only those windows with fractional snow cover within the range $0.1 \leq f_{SCA}^{Landsat} \leq 0.9$. This prevents misleading statistical analysis of the model results.

For windows satisfying the given range, we iterate through candidate values, w, on [0,1] at intervals of 0.01. For each iteration, binary classification performance metrics are used to construct a confusion matrix. True positive (TP), false positive (FP), true negative (TN), and false negative (FN) occurrences are obtained by comparing the Landsat derived 30 m binary snow cover map to each of the realizations for different values of w on a pixel-by-pixel basis. Where the model predicts snow when Landsat NDSI suggests no snow, a false positive (i.e. type I error) occurs. Conversely, false negatives (type II errors) occur where Landsat NDSI suggests snow cover but where the model predicts snow-free conditions. The occurrence of these four possible outcomes allows us to calculate precision (p) and recall (r). Precision equals the proportion of modeled snow-covered pixels that are correctly identified as snow,

$$p = \frac{TP}{TP + FP}. \qquad (10)$$

Recall equals the proportion of observed snow-covered pixels that are correctly modeled (Powers, 2011),

$$r = \frac{TP}{TP + FN}. \qquad (11)$$

This analysis closely follows the approach of Rittger et al. (2012) in which the accuracy of MODIS snow cover products is evaluated. Accordingly, the harmonic mean of p and r is the so-called F-measure,

$$F = 2 \cdot \frac{p \cdot r}{p + r} \qquad (12)$$

which provides a robust statistical balance for such a binary test. We store the F-measure for each candidate w for every analyzed window.

Taking the mean of all windows for each value of w produces a vector of F-measure as a function of w, from which the maximum F-measure and associated w can be retrieved. We are thus calibrating the value of w to maximize the spatially averaged F-measure. This analysis is repeated for all available Landsat scenes used in calibration (Fig. 3).

3.6. Parameter validation

Since we drive the downscaling routine with f_{SCA} information, the outcome of the performance metrics are subject to particular constraints. For instance, when f_{SCA} input to the downscaling model is constrained to the value of f_{SCA} inferred from Landsat, $f_{SCA}^{Landsat}$, the confusion matrix between observation and model is symmetric, meaning the number of false positives and false negatives are equal. This occurs because every time a false positive arises, we miss an opportunity to correctly classify a snow-covered pixel, thus generating a false negative elsewhere in the grid. This leads to equality in precision and recall and it can be demonstrated that the F-measure simplifies to p or r when the f_{SCA} value input to the model and the f_{SCA} value to which the model is being calibrated are identical. Thus, the y-axis in Fig. 3 could also be labeled as precision or recall since, during calibration, our model always receives as input the value of $f_{SCA}^{Landsat}$. Calibration could, therefore, be simplified by simply maximizing precision. However, during validation the map of f_{SCA} values input to the downscaling model comes from the MOD10A1 product and is not necessarily equal to the map of f_{SCA} that would have been derived from Landsat at the same spatial scale. In validation, therefore, it is necessary to perform assessment via the F-measure since p and r will often not be equal.

It is also important to determine a minimum acceptable value of the F-measure that indicates the model is outperforming a random assignment of binary snow cover for a given MOD10A1 f_{SCA}. For any given Landsat-derived binary snow cover map, a corresponding map of randomly assigned snow cover with identical f_{SCA} will produce an F-measure that is on the order of the f_{SCA} value. Therefore, for a given f_{SCA} value, an F-measure for the downscaling model that is outside the range of variability of the corresponding F-measures of an ensemble of randomly produced SCA within a MODIS pixel would indicate that the model is outperforming a random assignment of SCA within the pixel. We estimate the range of variability through a Monte Carlo simulation for a range of f_{SCA} possibilities from 0.1 to 0.9. For each possible f_{SCA},
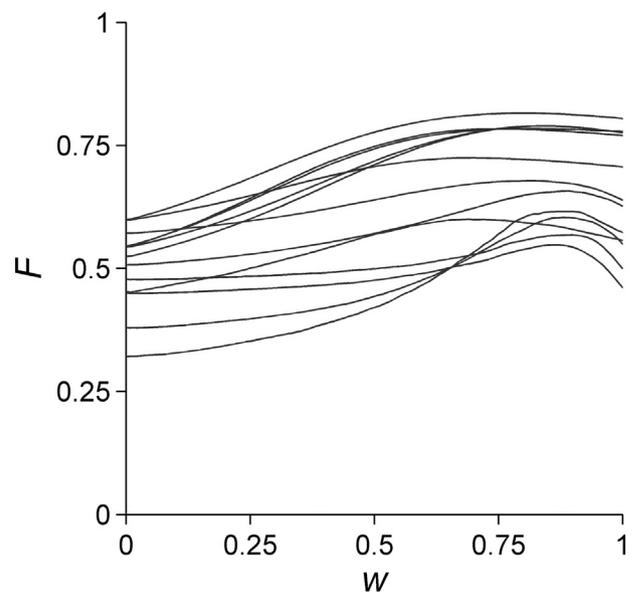
Fig. 3. Mean F-measure by parameter w for subset a for 13 temporally separated Landsat scenes. For each date, a calibration optimum is chosen that corresponds to the maximum value taken on by F.

we randomly produce a synthetic true binary snow cover map. For each synthetic true map, we generate an ensemble of 10,000 random binary snow cover maps having an f_{SCA} value that is within a fixed percentage, δ, of the snow cover fraction of the synthetic truth. This approach allows for discrepancies between the retrieved f_{SCA} (i.e., that which would be input to the downscaling model) and the true binary snow cover (i.e., that which would be inferred from Landsat). We calculate the F-measure for all random snow cover maps and every f_{SCA} value. For a given f_{SCA}, we compute the expected value of F and its variance, denoted μF_{random} and σF_{random}, respectively. The variability in F-measure tends to decrease as f_{SCA} increases because pixels randomly assigned as snow cover are more likely to be correct than at lower values of f_{SCA}. F-measure ensembles tend to be distributed normally across each value of f_{SCA} with a variance that decreases with increasing f_{SCA}.

4. Results

4.1. Calibration results

To obtain a calibrated parameter value, w, used in the downscaling model, we examine the performance of the F-measure across snow cover grids for 13 different Landsat scenes in the manner described above. A vector of F-measure statistics is generated for each candidate w and 500 m window. The mean F-measure across all 500 m windows and for each candidate w is computed. Note that, as described in Section 3.6, the F-measure is identical to precision and recall in this characteristic case. The w value that maximizes the magnitude of F (p or r) is chosen as the optimum weight parameter for each Landsat scene.

For 13 different dates of Landsat acquisition, the calibrated parameter w varies within a relatively narrow range from 0.83 to 0.94. The distribution of optimal w values has a mean, $w\mu$, of 0.9069 and standard deviation, $w\sigma$, of 0.0364. The maximum, domain-averaged F-measure for all calibrated Landsat scenes ranges from 0.5403 on 15 May 2008 to 0.8163 on 16 February 2002. The f_{SCA} across the whole calibration domain within the calibration scenes ranges from less than 1% on 27 April 2007 to approximately 77% on 16 February 2002. Fig. 4 depicts these resulting parameter optima by date, along with their corresponding values of F and domain-wide f_{SCA} values. Note that the F-measure maxima tend to be highest when domain snow cover fraction is also relatively high.

Parameter w represents the weight assigned to the normalized slope factor f_{norm} while its complement, $(1 - w)$, is that assigned to normalized elevation z_{norm}. Thus, the calibration seems to suggest that spatial variability in insolation is the dominant predictor variable used because the weights appear to be relatively consistent over time (Fig. 4). The weights remain constant because the normalized insolation includes

the physics describing how insolation variability increases as the spring progresses. If insolation is not normalized in this way, the optimal weight will decrease during the season as insolation becomes less variable and the spatial distribution of f_{SCA} becomes more controlled, presumably, by elevation. By normalizing to the maximum relative insolation over the season, we are including this effect in the physics of the model and allowing the approach to use only one calibrated parameter. Otherwise one or more additional empirical parameters would be necessary to account for this temporal change. It should be reiterated, however, that values for f_{norm} vary with Julian date and are normalized to seasonal maxima whereas z_{norm} values are static and normalized to each grid window. As a result, insolation (as captured by f_{norm}) exhibits a larger dynamic range in the winter and early spring, and dominates where snow-covered pixels are assigned in the downscaling routine. However, as the season progresses elevation (as captured by z_{norm}) becomes increasingly important in predicting snow cover.

4.2. Validation results

Preliminary validation is performed using data from the MOD10A1 Version 005 f_{SCA} product (Hall, Riggs, & Salomonson, 2006) as input to the downscaling model and Landsat SCA to facilitate statistical analyses. We perform validation over regions b and c in the same Landsat scene (path/row 41/30) used for calibration (Fig. 1). The MOD10A1 and Landsat observations are acquired for 18 March 2010, a date not used in model calibration. This date is selected because both Landsat and Terra satellites passed over the region coincidentally during this date, obtaining a clear view of the domain. We operate the validation in similar fashion to the calibration procedures. Every MODIS pixel is downscaled to 30 m spacing, following the methods described in Section 3.4. We use the mean of calibration w values, $w\mu$, to generate downscaled estimates. For all MODIS pixels, we select for validation only those which satisfy $(0.1 \leq f_{SCA}^{Landsat} \leq 0.9)$ and $(0.1 \leq f_{SCA}^{MODIS} \leq 0.9)$. These criteria are used for two reasons:

1) By default, the model performs quite well when f_{SCA} is near zero or one so we choose to disregard the performance at these ranges.
2) We are interested in examining the results when both MOD10A1 and Landsat have an f_{SCA} not near 0 or 1 at the 500 m scale, since we are not evaluating the performance of MOD10A1, but rather the performance of the downscaling model when both products suggest the evidence of partial snow cover.

At windows compliant with the above criteria, the difference in f_{SCA} between the two products, Δf_{SCA}, is computed,

$$\Delta f_{SCA} = f_{SCA}^{Landsat} - f_{SCA}^{MODIS} \qquad (13)$$
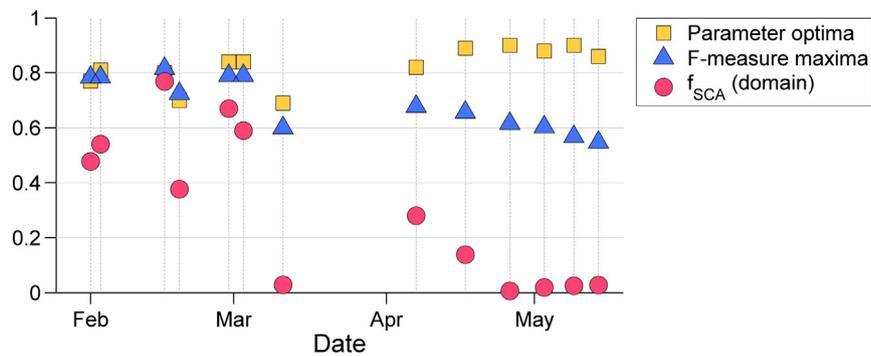
Fig. 4. Results from calibration phase. Parameter optima (yellow squares) are chosen as the corresponding weight which maximizes the average F-measure for each date. The F maxima are shown as blue triangles. Snow cover fractions for the whole domain are plotted as red circles. Note that the x-axis dates are relative, meaning they are not all from the same snow season. Also note that the y-axis is intentionally un-labeled as all three plotted datasets are unitless and fall on the same interval [0,1]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
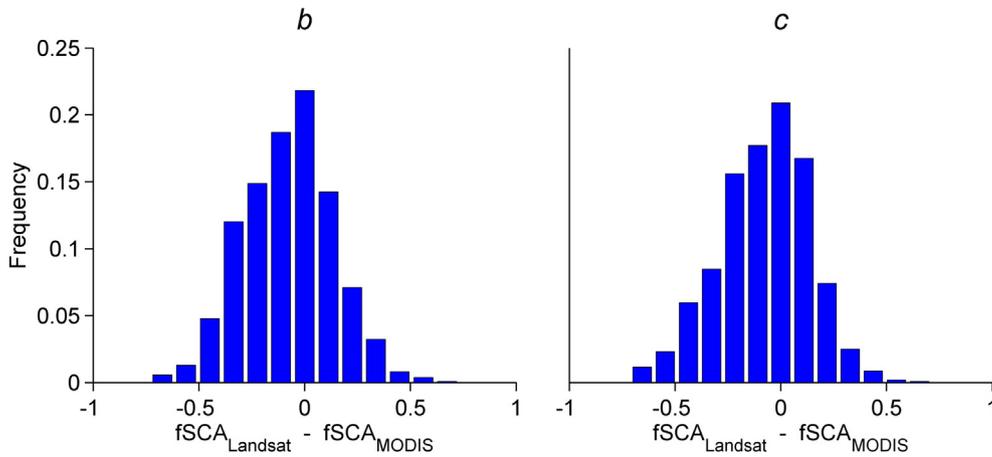
Fig. 5. Relative density histograms displaying the differences in snow cover fraction between Landsat and MODIS over the two validation subsets on 18 March 2010. Subsets b and c are shown at left and right, respectively. Note that these differences do not include occurrences where both satellites observe 0% or 100% snow cover.

Fig. 5 illustrates the histogram of f_{SCA} disparity between MODIS and Landsat across each validation subset.

We further restrict validation to those MODIS pixels satisfying ($|\Delta f_{SCA}| \leq 0.1$) to examine those observations where MODIS agrees with the observed Landsat snow fraction to within 10%. We choose this tolerance value to coincide with the choice of δ in Section 3.6. In subset regions b and c, 34.98% and 36.71% of windows fall, respectively, within this tolerance. Note that these percentages exclude occurrences where MODIS and Landsat observed completely snow-free or snow-covered conditions. We then compare the F-measures calculated between downscaled MODIS and observed Landsat binary snow cover with the scores computed via random model generation from

Section 3.6. If, for a given f_{SCA}, the corresponding F-measure within that MODIS pixel exceeds $\mu F_{random} + \sigma F_{random}$, we suggest that the model is performing agreeably relative to a random assignment of snow cover locations. Fig. 6 depicts the results of validation, displaying downscaling model F-measure for given values of f_{SCA}. Since we set constraints on Δf_{SCA} to be within 10%, we plot the measures against the snow fractions for both the downscaled MOD10A1 f_{SCA} and Landsat f_{SCA} computed over the MOD10A1 footprint. For a point plotting above $\mu F_{random} + \sigma F_{random}$, we are approximately 68% confident that the model is performing better than a random one. Similarly, for points plotting above $\mu F_{random} + 2 \cdot \sigma F_{random}$, we are 95% confident the model is outperforming a random one. Table 2 summarizes the fraction
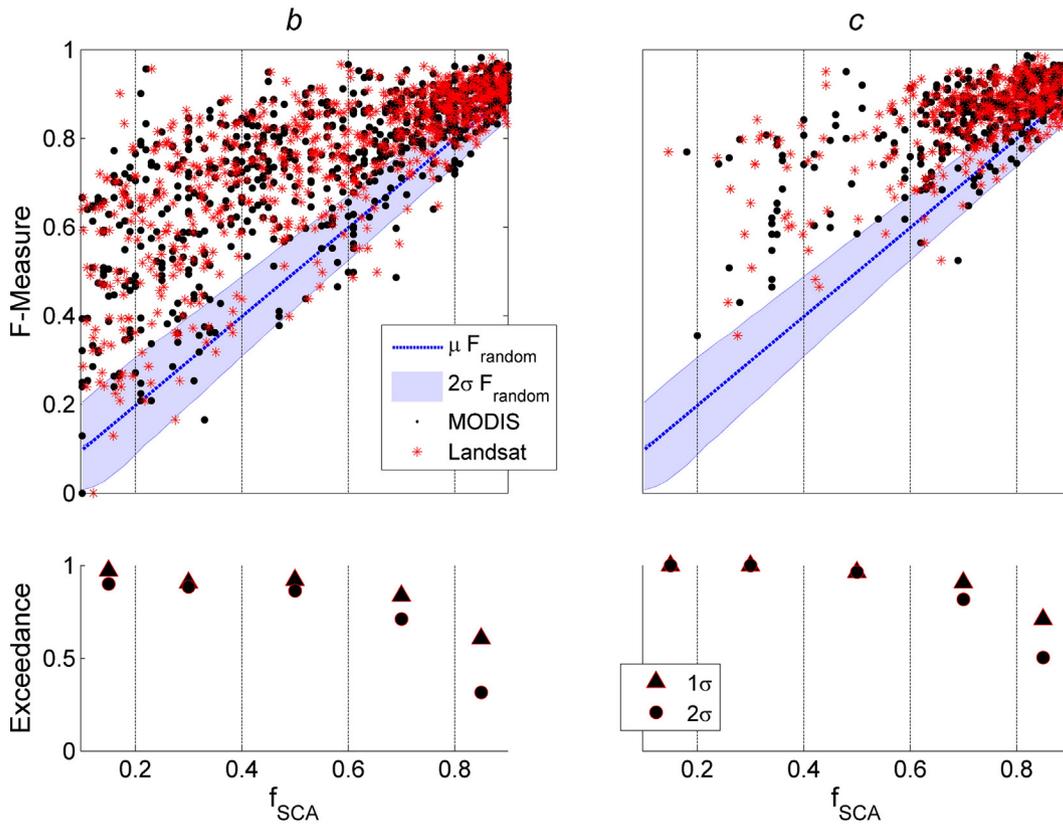
Fig. 6. Top row: Modeled F-measure (downscaling model versus Landsat observation) as a function of f_{SCA} (MODIS and Landsat) for windows where MODIS and Landsat f_{SCA} agree to within 10% ($|\Delta f_{SCA}| < 0.1$). Mean F-measure from random model ensembles ($\delta = 0.1$) is also plotted against f_{SCA} with error bars of $\pm 2\sigma$. Validation subset b and c are shown at left and right, respectively. Bottom row: Fraction of F-measures (from MODIS) that exceed random model ensemble ranges within f_{SCA} bins. Exceedance fractions are shown for 1σ and 2 σ levels of random variation.

Table 2
Fraction of downscaled windows ($|\Delta f_{SCA}| < 0.1$) in each subset region with F-measures exceeding the normal range of variability in those of random model ensembles ($\delta = 0.1$). Exceedance is shown for Landsat and MODIS.

f_{SCA}	Variability	Exceedance fraction	
Platform	Range	Subset b	Subset c
Landsat	$\mu F_{random} + \sigma F_{random}$	0.8122	0.8058
	$\mu F_{random} + 2 \cdot \sigma F_{random}$	0.7017	0.6325
MODIS	$\mu F_{random} + \sigma F_{random}$	0.7776	0.8320
	$\mu F_{random} + 2 \cdot \sigma F_{random}$	0.6823	0.6719

Table 3
Fraction of downscaled windows with ideal snow fraction ($\Delta f_{SCA} = 0$) in each subset region with F-measures exceeding the normal range of variability in those of random model ensembles ($\delta = 0.0$).

Standard	Exceedance fraction	
Deviations	Subset b	Subset c
1σ	0.8491	0.9011
2σ	0.8101	0.8723

of downscaled scenes that satisfy these criteria. To demonstrate statistical significance of the modeled F-measures, we perform a Wilcoxon rank-sum test on the values obtained from the downscaling model versus those from the random model assignments. This test is chosen because 1) we are not able to assume normal distributions for the modeled F-measures and 2) the distributions of modeled and random scores are not necessarily the same size. The rank-sum test is performed for each f_{SCA} bin indicated by Fig. 6 for a total of ten tests. For all but one bin, the test indicates that the F-measures from the downscaling model are statistically higher than random ones at the $p = 0.01$ significance level. For the 10–20% f_{SCA} bin in region c, there is simply not enough

data to perform a meaningful test of the hypothesis that the modeled F-measures are high and significant.

From a qualitative standpoint, the model more accurately downscales in subset validation region b. However, it is possible that this is a result of less agreement between Landsat and MODIS over region c as depicted by the sheer fewer number of such points in this region (Fig. 6c). Many occurrences of f_{SCA} between 0.8 and 0.9 are observed in both subset regions and, on the whole, most downscaling model F-measures plot above the μF_{random} lines in Fig. 6, indicating that the model is, at the very least, moving snow-covered pixels in the correct direction. However, there are points, especially at lower snow cover fractions, where the model performance exceeds randomness, yet still suffers a qualitatively low F-measure, meaning there are many type I and/or type II errors still occurring, albeit less than if the model were not tuned to terrain indices. Overall, the mean F-measure is higher over subset region c than for region b. However, in region c f_{SCA} varies mostly between 0.65 and 0.90, where F-measures must be correspondingly high to outperform a random assignment of snow cover.

4.3. Ideal simulation results

We also execute a validation simulation under circumstances of "ideal" fractional snow cover observations. That is, MODIS observations of f_{SCA} are substituted for the corresponding f_{SCA} values derived from Landsat over the MODIS footprint. In this fashion, the model is given maximum opportunity to downscale f_{SCA} windows containing the observed, "true" amount of snow-covered area since $\Delta f_{SCA} = 0$ everywhere. We examine the results of the downscaling model using the same methods outlined in Section 4.2. An updated ensemble of random models is generated with $\delta = 0$ to represent perfect snow cover observations. From this, we produce a new distribution of F-measures that is narrower than one resulting from a non-zero choice for δ. Table 3 provides the fraction of windows satisfying the criteria for outperforming
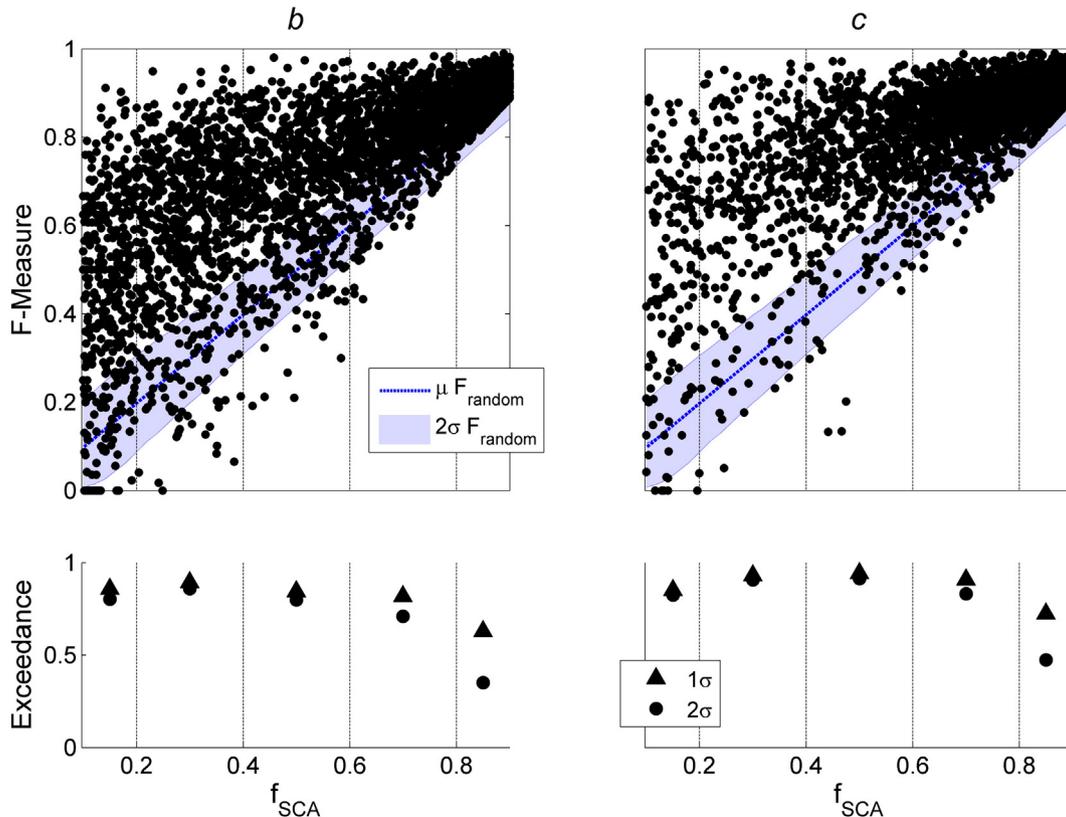
Fig. 7. Modeled F-measure as a function of "ideal" f_{SCA} for synthetic windows where MODIS is adjusted to equal Landsat. Mean F-measure from random model ensembles ($\delta = 0.0$) is also plotted against f_{SCA} with error bars of $\pm 2\sigma$. Validation subsets b and c are shown at left and right, respectively.

Table 4

Mean *F*-measures across regions *b* and *c* when downscaling MODIS and ideal (Landsat aggregate) f_{SCA} grids.

f_{SCA}	Mean *F*-measure (F_μ)	
Platform	Subset *b*	Subset *c*
MODIS	0.7313	0.8372
Ideal	0.7232	0.8182

random snow cover assignments outlined in the previous section. Fig. 7 depicts qualitative results. In these ideal circumstances, the downscaling model appears to perform better across subset region *c* than *b*. It should be noted that region *c* has many windows with Landsat-observed f_{SCA} between 0.9 and 1.0 (n = 10,115) in comparison to region *b* (n = 5,111). Thus, these would be excluded from the evaluation leading to less windows for validation in region *c* (n = 2,992) than in region *b* (n = 3,333). Of the windows in the evaluation range, region *c* has a higher mean *F*-measure across its domain than *b*. The mean *F* values for MODIS and ideal (Landsat) snow cover fraction evaluation are given in Table 4.

Wilcoxon rank-sum tests are also performed for the ideal simulations. For all f_{SCA} bins in regions *b* and *c*, the modeled *F*-measures are significantly higher than those obtained from random ensemble assignments (p = 0.01).

A *K*-fold cross-validation routine is also performed across subset region *a* for each of the Landsat observation dates (K = 13) used in initial model calibration. For each leave-one-out iteration, the mean of the remaining 12 optimized calibration parameters (*w*) is used to drive the downscaling model along with the ideal synthesized snow fraction values of $f_{SCA}^{Landsat}$. Downscaled results are compared with the original Landsat SCA observation for each date and the *F*-measure exceedance fractions are calculated. Table 5 depicts the cross-validation performance results. Fig. 8 provides a visual example of downscaled results at a small spatial scale to qualitatively depict the model's skill at predicting snow cover at the hillslope scale when accurate f_{SCA} data are available (i.e. MODIS and Landsat snow fractions are similar). At a larger glance, Fig. 9 illustrates example downscaled results for the entire subset region *a*.

5. Discussion

The developed downscaling model performs with significant skill for all subsets, requiring calibration of only one empirical parameter. Thus, we produce snow cover maps at the temporal resolution of MODIS and the spatial resolution of Landsat, a tool that is not currently available to hydrologic modelers. However, locations exist in *F*-measure plots within, or below, a range of *F*-measures that can be achieved through purely random assignment of snow-covered locations. Even those downscaled 500 m pixels that exceed $\mu F_{random} + 2 \cdot \sigma F_{random}$ do not always reveal a dramatic improvement from random. The fraction of pixels with *F*-measures exceeding this range was, however, promising, especially across subset region *c*. This region also exhibited the highest mean *F*-measure, performing even better with input MOD10A1 f_{SCA} than with the ideal snow fraction.

Ancillary information not currently included in the downscaling model may provide a source of improvement to the downscaling approach. Further analysis is necessary in order to gain insight into physiographic characteristics of terrain within MODIS windows where the model is failing (or succeeding) consistently. For example, since we do not account for a forest canopy, we may be assigning snow-free states to areas where a forest canopy impedes a great deal of insolation and obscures existing snow cover. Forest canopy is also a potential reason

Table 5

K-fold cross-validation results from subset region *a*. Snow fractions are given for the whole domain and the fractions of *F*-measures exceeding the normal range of variability from random model assignments are shown.

Date	f_{SCA}	Exceedance fractions	
		1σ	2σ
02/03/2000	58.3	0.8101	0.7597
02/19/2000	44.5	0.7876	0.7375
03/01/2001	64.7	0.9823	0.9764
04/18/2001	10.4	0.9558	0.9381
05/04/2001	2.3	0.9617	0.9469
02/16/2002	73.3	0.9646	0.9587
03/04/2002	61.3	0.9558	0.9440
04/08/2003	20.4	0.9381	0.9056
05/10/2006	3.0	0.9292	0.9027
04/27/2007	1.4	0.9322	0.9056
03/12/2008	3.4	0.9410	0.9086
05/15/2008	3.4	0.9292	0.8997
02/01/2011	52.5	0.9263	0.9027

for the f_{SCA} offset between MODIS and Landsat. For example, Raleigh et al. (2013) found that f_{SCA} derived from the MODSCAG model consistently under-predicted the observed snow fraction (quantified via temperature sensor networks) in a study plot with high canopy fraction (i.e. 79%), even when the remotely sensed f_{SCA} value was statistically corrected for the canopy cover.

Another parameter worthy of consideration is wind redistribution. With knowledge of time-series wind vectors or seasonal tendencies, terrain-based indices can be generated in order to predict areas prone to drifting and general redistribution during accumulation and settling of snow fall. Winstral, Elder, and Davis (2002) showed that a snow depth regression tree model based on elevation, solar radiation, and slope was substantially improved by the addition of a wind redistribution predictor variable. Using this parameter with a physically based, distributed snow model, improvements were made to the simulated distribution of snow cover, snow drifts and melt within RCEW (Winstral & Marks, 2002). RCEW is located within validation subset region *b*, which will facilitate future efforts to include such information into the downscaling model. In addition, in RCEW prevailing winds come from the southwest and resultant drifts tend to form on north facing slopes. Therefore, our model may be getting the right answer for the wrong reason, since the insolation component of the model causes similar spatial patterns of snow cover (Winstral & Marks, 2002). Anderson et al. (2014), however, demonstrated that in the DCEW (subregion a) differential ablation throughout the snow-covered season has stronger impact on the spatial variability of SWE than does wind redistribution.

Future work may also employ spectral mixture analysis snow cover products (Painter et al., 2009; Rosenthal & Dozier, 1996) in order to perform similar experiments using products that have been suggested to have a higher accuracy (particularly during the ablation phase with which we are most concerned) in reproducing the snow cover distribution. In applying the model developed here to satellite retrievals using spectral mixing models, care must be taken to ensure that the 30 m resolution binary snow cover map is produced using techniques similar to those used to estimate the 500 m resolution fSCA (i.e. Painter et al., 2003).

The linear method used here to combine the physiographic index variables (insolation and elevation) is a relatively simplistic way to combine the two effects captured by the predictor variables, making a speculative assumption that the variables' relative contributions to snow cover are dependent on a single-parameter, linear model. It is worth noting that more complex approaches exist for multivariate data characterization. A method to be explored in future work is the use of copulas for blending multiple variable indices. Copulas are

Fig. 8. Hillslope-scale downscaling visualization for two arbitrary subsets within greater subset region *a* for 18 March 2010. Each subset represents a 1500 m window (9 MODIS cells) and shows the Landsat-derived snow cover observation, downscaled Landsat aggregate f_{SCA}, downscaled MODIS f_{SCA} and performance metrics for each 1500 m window.
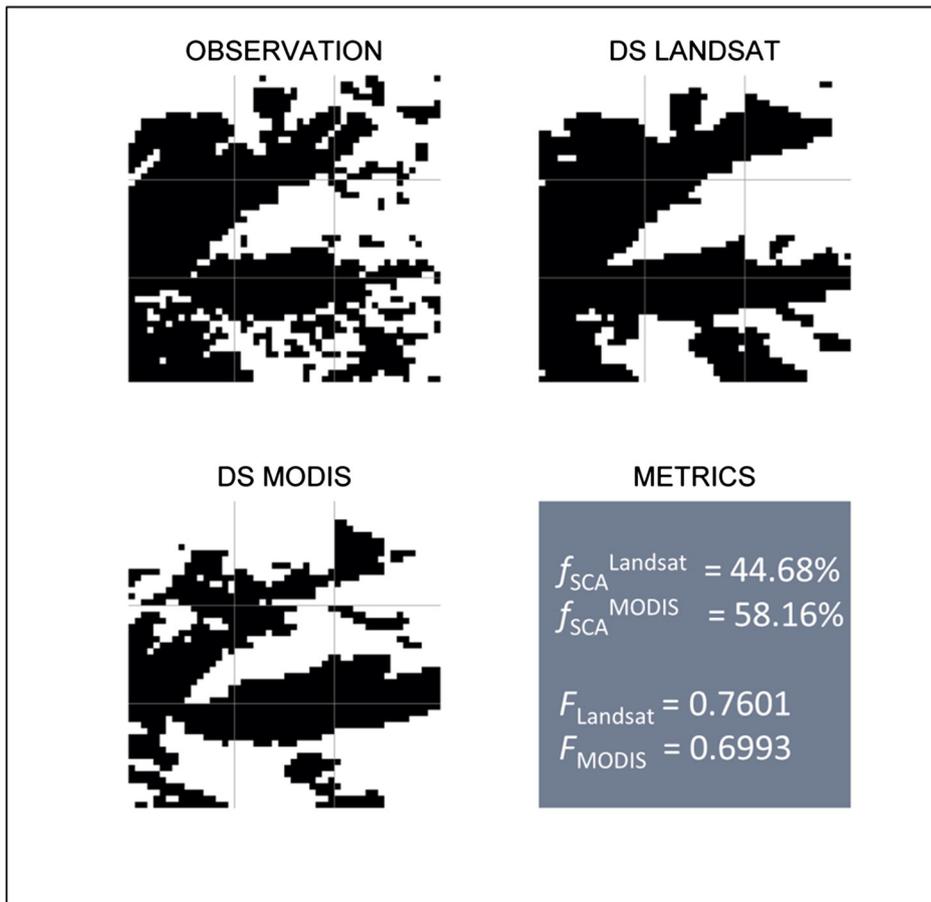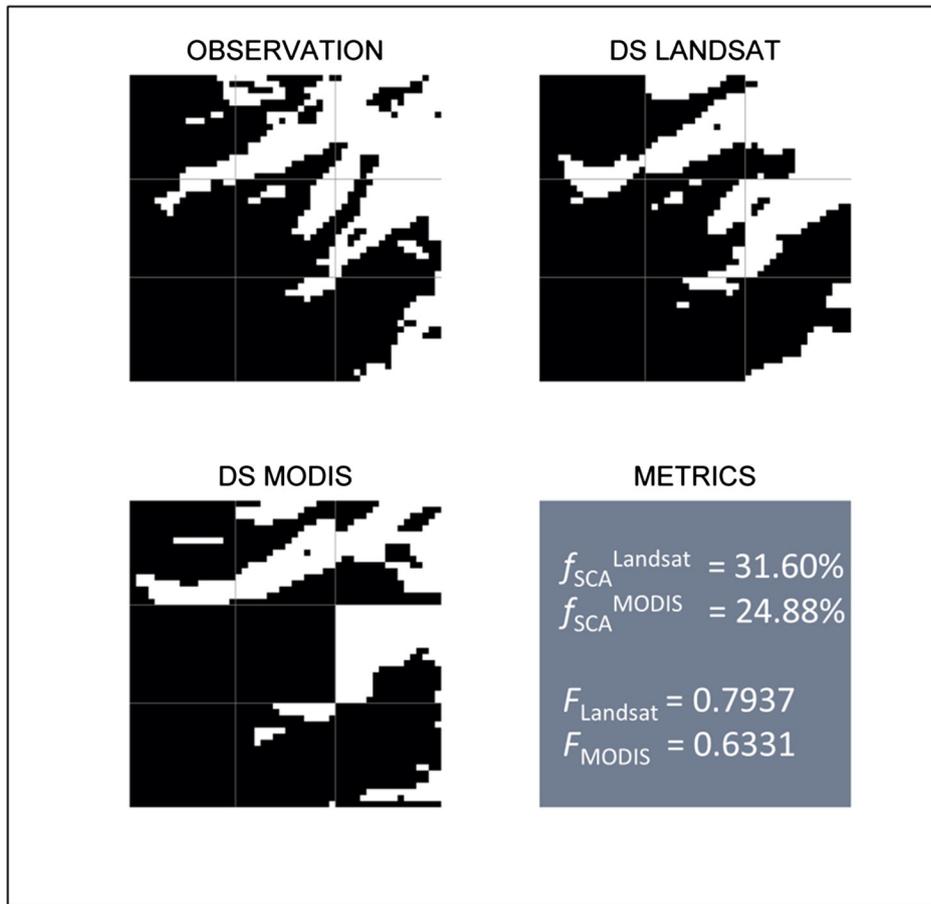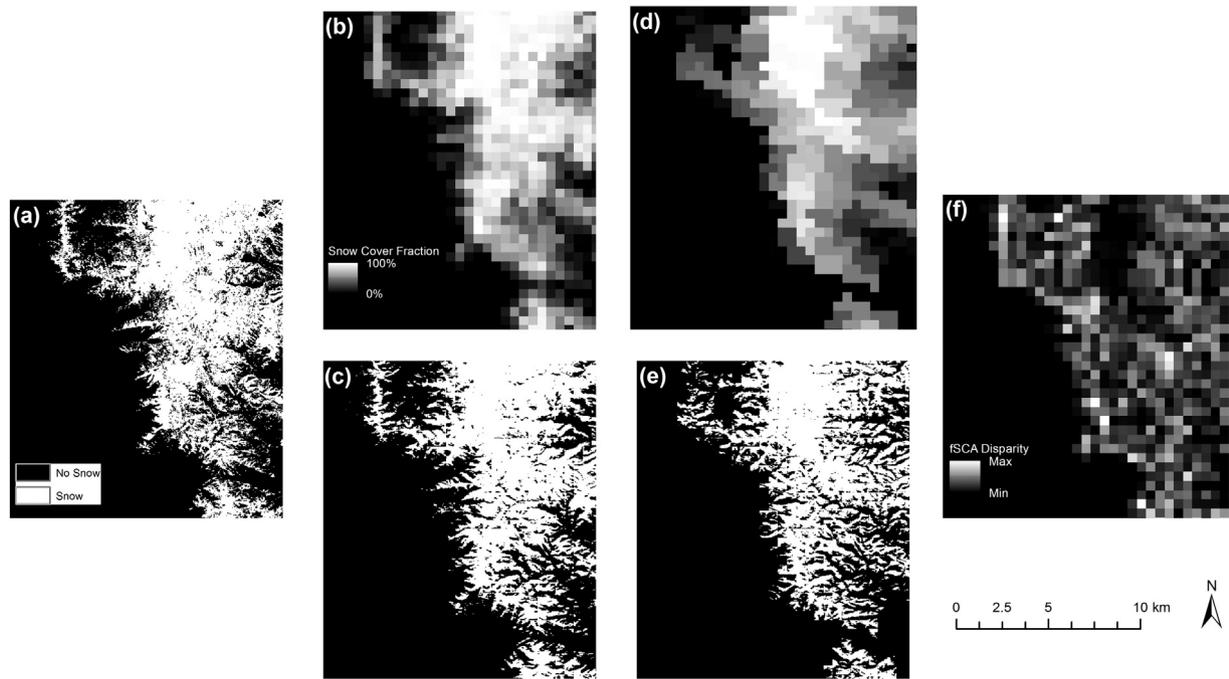
Fig. 9. Downscaling visualization over region *a* (18 March 2010) displaying (a) Landsat observed binary snow cover, (b) "ideal" f_{SCA} as 500 m aggregate from Landsat, (c) downscaled f_{SCA} from previous, (d) MOD10A1 500 m f_{SCA}, (e) downscaled f_{SCA} from previous, and (f) absolute difference grid between Landsat and MODIS, normalized to unity.

efficient tools used to describe the dependence structure of joint multivariate random variable distributions through examination of the underlying univariate marginal distributions (Schwarzlander, 2011). This method is frequently applied in financial risk analysis and has been used increasingly in the field of hydrology in the past several years (Schölzel & Friederichs, 2008). Such an approach may offer additional insight into the relationship between physiographic parameters and snow cover.

6. Conclusion

We describe here a model to downscale fractional snow-covered area (f_{SCA}) data from MODIS to a higher-resolution, spatially explicit binary grid based on physiographic indices derived primarily from digital elevation data. Parameterizations for elevation and insolation are combined to yield a terrain score that is used to assign binary snow cover on a 30 m grid with information about fractional snow cover on a 500 m grid. Using Landsat binary snow cover maps derived using an NDSI threshold, the model is calibrated to 13 Landsat binary snow cover scenes for one domain subset and validated against two independent domain subsets (each in the same scene, approximately 100 km away) for a date not used in calibration. Cross-validation is also performed on each of the 13 scenes across the original calibration subset region. Calibration results depict a general stability in weight parameters over varying times of the snow season and degrees of snow cover. Results from validation show that model's skill, as assessed via the *F*-measure, tends to exceed that of ensembles of randomly generated snow cover maps over a relatively large range of f_{SCA}. The statistical significance of these findings is confirmed via Wilcoxon rank-sum testing on the *F*-measures in binned ranges of f_{SCA}. Perhaps unsurprisingly, the model performs best in the region over which it is calibrated (as suggested by the cross-validation results) though it exhibits promising results in the independent validation regions as well. The developed model could assist hydrologic modeling by providing daily, high-resolution constraints on the distribution of snow in mountainous areas.

References

Anderson, B. T., McNamara, J. P., Marshall, H. -P., & Flores, A. N. (2014). Insights into the physical processes controlling correlations between snow distribution and terrain properties. *Water Resources Research*, *50*, http://dx.doi.org/10.1002/2013WR013714.

Bales, R. C., Molotch, N.P., Painter, T. H., Dettinger, M.D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*, *42*(8), W03537, http://dx.doi.org/10.1029/2005WR004387.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*, 303–309, http://dx.doi.org/10.1038/nature04141.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., et al. (2011). Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, *47*(7), http://dx.doi.org/10.1029/2011WR010745.

Clark, M. P., Slater, A. G., Barrett, A. P., Hay, L. E., McCabe, G. J., Rajagopalan, B., et al. (2005). Assimilation of snow covered area information into hydrologic and land-surface models. *Advances in Water Resources*, *29*(8), 1209–1221, http://dx.doi.org/10.1016/j.advwatres.2005.10.001.

Cline, D., Bales, R., & Dozier, J. (1998). Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. *Water Resources Research*, *34*(5), 1275–1285.

Dozier, J. (1989). Spectral signature of alpine snow cover from the Landsat Thematic Mapper. *Remote Sensing of Environment*, *28*, 9–22.

Dozier, J., & Frew, J. (2009). Computational provenance in hydrologic science: A snow mapping example. *Philosophical Transactions of the Royal Society A*, *367*, 1021–1033, http://dx.doi.org/10.1098/rsta.2008.0187.

Dozier, J., & Marks, D. (1987). Snow mapping and classification from Landsat Thematic Mapper data. *Annals of Glaciology*, *9*, 97–103.

Durand, M., Molotch, N.P., & Margulis, S. A. (2008). Merging complementary remote sensing datasets in the context of snow water equivalent reconstruction. *Remote Sensing of Environment*, *112*(3), 1212–1225, http://dx.doi.org/10.1016/j.rse.2007.08.010.

Elder, K., Dozier, J., & Michaelsen, J. (1991). Snow accumulation and distribution in an Alpine Watershed. *Water Resources Research*, *27*(7), 1541–1552, http://dx.doi.org/10.1029/91WR00506.

Flerchinger, G. N., Marks, D., Reba, M. L., Yu, Q., & Seyfried, M. S. (2010). Surface fluxes and water balance of spatially varying vegetation within a small mountainous headwater

catchment. *Hydrology and Earth System Sciences*, *14*(6), 965–978, http://dx.doi.org/10.5194/hess-14-965-2010.

Graham, C. B., Barnard, H. R., Kavanagh, K. L., & Mcnamara, J. P. (2012). Catchment scale controls the temporal connection of transpiration and diel fluctuations in streamflow. *Hydrological Processes*, http://dx.doi.org/10.1002/hyp.9334.

Hall, D. K., & Riggs, G. A. (2007). Accuracy assessment of the MODIS snow products. *Hydrological Processes*, *21*(12), 1534–1547, http://dx.doi.org/10.1002/hyp.6715.

Hall, D. K., Riggs, G. A., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote sensing of Environment*, *54*(2), 127–140.

Hall, D. K., Riggs, G. A., & Salomonson, V. V. (2006). *MODIS/Terra Snow Cover Daily L3 Global 500 m Grid, Version 5.* January 2006-March 2011. Boulder, Colorado USA: National Snow and Ice Data Center.

Homan, J. W., Luce, C. H., McNamara, J. P., & Glenn, N. F. (2011). Improvement of distributed snowmelt energy balance modeling with MODIS-based NDSI-derived fractional snow-covered area data. *Hydrological Processes*, *25*(4), 650–660, http://dx.doi.org/10.1002/hyp.7857.

Johnson, G. L., & Hanson, C. L. (1995). Topographic and atmospheric influences on precipitation variability over a mountainous watershed. *Journal of Applied Meteorology*, *34*(1), 68–87.

Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1228–1249, http://dx.doi.org/10.1109/36.701075.

Kelleners, T. J., Chandler, D.G., McNamara, J. P., Gribb, M. M., & Seyfried, M. S. (2010). Modeling runoff generation in a small snow-dominated mountainous catchment. *Vadose Zone Journal*, *9*(3), 517–527, http://dx.doi.org/10.2136/vzj2009.0033.

Kunkel, M. L., Flores, A. N., Smith, T. J., McNamara, J. P., & Benner, S. G. (2011). A simplified approach for estimating soil carbon and nitrogen stocks in semi-arid complex terrain. *Geoderma*, *165*(1), 1–11, http://dx.doi.org/10.1016/j.geoderma.2011.06.011.

Luce, C. H., Tarboton, D.G., & Cooley, K. R. (1998). The influence of the spatial distribution of snow on basin-averaged snowmelt. *Hydrological Processes*, *12*(10–11), 1671–1683.

Luce, C. H., Tarboton, D.G., & Cooley, K. R. (1999). Subgrid parameterization of snow distribution for an energy and mass balance snow cover model. *Hydrological Processes*, *13*(12–13), 1921–1933.

Marks, D., Link, T., Winstral, A., & Garen, D. (2001). Simulating snowmelt processes during rain-on-snow over a semi-arid mountain basin. *Annals of Glaciology*, *32*, 195–202.

Martinec, J., & Rango, A. (1981). Areal distribution of snow water equivalent evaluated by snow cover monitoring. *Water Resources Research*, *17*(5), 1480–1488, http://dx.doi.org/10.1002/hyp.7206.

McNamara, J. P., Chandler, D., Seyfried, M., & Achet, S. (2005). Soil moisture states, lateral flow, and streamflow generation in a semi-arid, snowmelt-driven catchment. *Hydrological Processes*, *19*(20), 4023–4038, http://dx.doi.org/10.1002/hyp.5869.

Molotch, N.P. (2009). Reconstructing snow water equivalent in the Rio Grande headwaters using remotely sensed snow cover data and a spatially distributed snowmelt model. *Hydrological Processes*, *23*(7), 1076–1089, http://dx.doi.org/10.1002/hyp.7206.

Nolin, A. W., Dozier, J., & Mertes, L. A. K. (1993). Mapping alpine snow using a spectral mixture modeling technique. *Annals of Glaciology*, *17*, 121–124.

Painter, T. H., Dozier, J., Roberts, D. A., Davis, R. E., & Green, R. O. (2003). Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data. *Remote Sensing of Environment*, *85*(1), 64–77, http://dx.doi.org/10.1016/S0034-4257(02)00187-6.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868–879, http://dx.doi.org/10.1016/j.rse.2009.01.001.

Powers, D.M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Raleigh, M. S., Rittger, K., Moore, C. E., Henn, B., Lutz, J. A., & Lundquist, J.D. (2013). Ground-based testing of MODIS fractional snow cover in subalpine meadows and

forests of the Sierra Nevada. *Remote Sensing of Environment*, *128*, 44–57, http://dx.doi.org/10.1016/j.rse.2012.09.016.

Reba, M. L., Marks, D., Seyfried, M., Winstral, A., Kumar, M., & Flerchinger, G. (2011). A long-term data set for hydrologic modeling in a snow-dominated mountain catchment. *Water Resources Research*, *47*(7), W07702, http://dx.doi.org/10.1029/2010WR010030.

Riggs, G. A., Hall, D. K., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment*, *54*, 127–140.

Rittger, K., Painter, T. H., & Dozier, J. (2012). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, *51*, 367–380, http://dx.doi.org/10.1016/j.advwatres.2012.03.002.

Rodell, M., & Houser, P. R. (2004). Updating a land surface model with MODIS-derived snow cover. *Journal of Hydrometeorology*, *5*(6), 1064–1075.

Rosenthal, W., & Dozier, J. (1996). Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. *Water Resources Research*, *32*(1), 115–130.

Salomonson, V. V., & Appel, I. (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment*, *89*, 351–360, http://dx.doi.org/10.1016/j.rse.2003.10.016.

Salomonson, V. V., & Appel, I. (2006). Development of the Aqua MODIS NDSI fractional snow cover algorithm and validation results. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(7), 1747–1756, http://dx.doi.org/10.1109/TGRS.2006.876029.

Schölzel, C., & Friederichs, P. (2008). Nonlinear Processes in Geophysics Multivariate non-normally distributed random variables in climate research — Introduction to the copula approach. *Nonlinear Processes in Geophysics*, *15*, 761–772.

Schwarzlander, H. (2011). *Probability concepts and theory for engineers.* Chichester, West Sussex, UK: John Wiley & Sons, Ltd.

Stieglitz, M., Shaman, J., Mcnamara, J., Engel, V., Shanley, J., & Kling, G. W. (2003). An approach to understanding hydrologic connectivity on the hillslope and the implications for nutrient transport. *Global Biogeochemical Cycles*, *17*(4), 1105, http://dx.doi.org/10.1029/2003GB002041.

Sturm, M., & Wagner, A.M. (2010). Using repeated patterns in snow distribution modeling: An Arctic example. *Water Resources Research*, *46*(12), W12549, http://dx.doi.org/10.1029/2010WR009434.

Swift, L. (1976). Algorithm for solar radiation on mountain slopes. *Water Resources Research*, *12*(1), 108–112.

Thirel, G., Salamon, P., Burek, P., & Kalas, M. (2011). Assimilation of MODIS snow cover area data in a distributed hydrological model. *Hydrology and Earth System Sciences Discussions*, *8*(1), 1329–1364, http://dx.doi.org/10.5194/hessd-8-1329-2011.

Tucker, C. J. (1979). Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment*, *8*(2), 127–150.

Tyler, S. W., Burak, S. A., Namara, J. P.M. C., Lamontagne, A., Selker, J. S., & Dozier, J. (2008). Spatially distributed temperatures at the base of two mountain snowpacks measured with fiber-optic sensors. *Journal of Glaciology*, *54*(187), 673–679.

Winstral, A., Elder, K., & Davis, R. E. (2002). Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology*, *3*, 524–538.

Winstral, A., & Marks, D. (2002). Simulating wind fields and snow redistribution using terrain-based parameters to model snow accumulation and melt over a semi-arid mountain catchment. *Hydrological Processes*, *16*(18), 3585–3603, http://dx.doi.org/10.1002/hyp.1238.

Wolfe, R. E., & Nishihama, M. (August 24, 2009). Trends in MODIS geolocation error analysis. *Proc. SPIE 7452. Earth Observing Systems XIV*, 74520L, http://dx.doi.org/10.1117/12.826598.

Yatheendradas, S., Peters Lidard, C. D., Koren, V., Cosgrove, B. A., De Goncalves, L. G. G., Smith, M., et al. (2012). Distributed assimilation of satellite-based snow extent for improving simulated streamflow in mountainous, dense forests: An example over the DMIP2 western basins. *Water Resources Research*, *48*, W09557, http://dx.doi.org/10.1029/2011WR011347.