A tool for downscaling weather data from large-grid reanalysis products to finer spatial scales for distributed hydrological applications

2 authors:

Avirup Sen Gupta

AIR Worldwide

8 PUBLICATIONS   **17** CITATIONS

SEE PROFILE
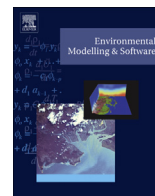
David G. Tarboton

Utah State University

223 PUBLICATIONS   **6,317** CITATIONS

SEE PROFILE

A tool for downscaling weather data from large-grid reanalysis products to finer spatial scales for distributed hydrological applications

Avirup Sen Gupta [a], [*], David G. Tarboton [b]

[a] Research and Modeling, AIR Worldwide, Boston, MA 02116, USA
[b] Department of Civil and Environmental Engineering, Utah State University, Logan, UT, USA

ARTICLE INFO

ABSTRACT

A downscaling tool was developed to provide sub-daily high spatial resolution surfaces of weather variables for distributed hydrologic modeling from NASA Modern Era Retrospective-Analysis for Research and Applications reanalysis products. The tool uses spatial interpolation and physically based relationships between the weather variables and elevation to provide inputs at the scale of a gridded hydrologic model, typically smaller (~100 m) than the scale of weather reanalysis data (~20–200 km). Nash-Sutcliffe efficiency (NSE) measures greater than 0.70 were obtained for direct tests of downscaled daily temperature and monthly precipitation at 173 SNOTEL sites. In an integrated test driving the Utah Energy Balance (UEB) snowmelt model, 80% of these sites gave NSE > 0.6 for snow water equivalent. These findings motivate use of this tool in data sparse regions where ground based observations are not available and downscaled global reanalysis products may be the only option for model inputs.

© 2016 Elsevier Ltd. All rights reserved.

Software availability

1. Introduction

High resolution weather data are increasingly used in distributed hydrologic modeling studies to simulate hydrological responses in heterogeneous areas. The outcomes of these studies are critical for water resources management decisions related to agricultural water supply, ecosystem services and hydropower production. While computer models in hydrology vary widely in purpose, complexity and spatial-temporal scale, physically based distributed models require as input continuous and complete time-varying weather data at each grid point or model element (Jeffrey et al., 2001). Moreover, physically based energy balance models often require incoming radiation fluxes and wind speed, which are not measured at all weather stations, especially in developing countries. Globally available climate reanalysis data provides an option for obtaining hydrologic model inputs where surface observations are limited or not available. However climate reanalysis data is often at a scale that is much coarser, typically 20–200 km, than the grid scale of physically based distributed hydrologic models, 100 m or less, derived from digital elevation models and scaled to represent topographic variability. There is thus a need for tools to produce inputs at the scale of hydrologic models from climate reanalysis data.

* Corresponding author. AIR Worldwide, 131 Dartmouth Street, Boston, MA 02116, USA.
E-mail address: avirup.sengupta@aggiemail.usu.edu (A. Sen Gupta).

In this study, we developed a spatial downscaling tool for generating 3-hourly grid surfaces of weather data over a complex terrain using reanalysis and satellite based precipitation data. The tool was developed to address the problem of obtaining sufficiently accurate input data to apply the Utah Energy Balance Snowmelt Model (UEB) to the melting of glaciers in the Himalaya region (Brown et al., 2014; Sen Gupta, 2014; Sen Gupta et al., 2015; Sen Gupta and Tarboton, 2013). The tool was designed to take inputs from large-grid reanalysis products such as NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011), and NOAA's Rainfall Estimation (RFE2) (Bajracharya et al., 2014, 2015; Shrestha et al., 2013; Xie and Arkin, 1996; Xie et al., 2002) products. UEB is an energy and mass balance snowmelt model designed for distributed application over a watershed at a grid scale fine enough to quantify topographic and vegetation variability including the variability in elevation, slope and aspect that are important for radiation inputs (Luce and Tarboton, 2010; Mahat and Tarboton, 2012; Tarboton and Luce, 1996). Typically the scale is chosen based on the scale of a digital elevation model (30–100 m), a scale we refer to as the hydrological scale. UEB requires inputs of precipitation, air temperature, downwelling long and short wave radiation, air humidity and wind speed at the scale and elevation of its grid cells. UEB inputs include slope and aspect and it adjusts internally for the effect of these on radiation, but adjustments due to processes in the atmosphere above the surface are not modeled within UEB and should be accounted for in inputs provided to UEB. These include adjustments in temperature and humidity due to lapse rates and the differences in elevation between the observation sites or nominal elevation of a meteorological model or reanalysis input. Precipitation, radiation and wind are also elevation dependent. This paper addresses the hydrometeorological downscaling required to adjust inputs from the scale of a meteorological model or reanalysis to the scale of the hydrological grid. Hydrometeorological downscaling is not unique to the application of UEB. It is required for any fine scale (DEM scale) hydrological model. It is distinct from the statistical or dynamical downscaling (e.g. Weather Research and Forecasting Model, nested within a General Circulation Model) used to go from climate model scale (~0.5–2°) to regional model scale (~2–100 km) (Benestad, 2004; Fowler et al., 2007; Wilby et al., 2002; Xu, 1999). Existing approaches for hydrometeorological downscaling include MTCLIM (Hungerford et al., 1989), Integrated Runoff Model Bultot (IRMB, Gellens et al., 2000), Daymet (Thornton et al., 2012), MicroMet (Liston and Elder, 2006) and TopoSCALE (Fiddes and Gruber, 2014).

Sparse meteorological data in the Himalayan region motivated developing a methodology for driving UEB using downscaled globally available reanalysis data. However there was insufficient data there to evaluate and validate the downscaling approaches described here. Instead, the methodology was evaluated at sites in the Western US where there is more data available. Precipitation and temperature were directly compared at 173 SNOTEL sites in Utah, Nevada, Idaho and California. Radiation and wind downscaling, data for which is less widely available, was tested using data from the Utah State University Doc Daniel site (NRCS, 2014) in the Logan River watershed from October 2009 to June 2010 where we had access to additional detailed field observations. The downscaled data were also used in an integrated test to drive the UEB snowmelt model to simulate the spatial and temporal variability of Snow Water Equivalent (SWE) at these SNOTEL sites. This tests the integral effect of downscaled inputs in the context of the UEB model. Results do depend on the sensitivity of the model to inputs, and thus this test may not reveal discrepancies in inputs to which the model is less sensitive, but ultimately it is the performance of the model that we are interested in, so discrepancies in

these variables are less important in this context.

The code for the tool we developed, called MERRA Spatial Downscaling for Hydrology (MSDH), is open source and available in a public bitbucket repository (https://bitbucket.org/AvirupSenGupta/msdh.usu/). In developing the tool described here we drew upon ideas in prior work (Fiddes and Gruber, 2014; Hungerford et al., 1989; Liston and Elder, 2006; Thornton et al., 2012), but new code was developed and made open source as we needed a tool that can produce hydrologic model inputs from globally available climate reanalysis data, and that can be freely distributed and is easy to use. Micromet (Liston and Elder, 2006) incorporates much, but not all of the physics we wanted, but operates on point data and the code for MicroMet is only available for a fee. The recent Fiddes and Gruber (2014) article has elements in common with our approach, but does not report on code availability and appeared after we had substantially developed MSDH. The contributions of this paper include the physically based hydrometeorological downscaling methodology, open source R code implementation and graphical user interface software that embeds direct access to MERRA and RFE2 data sources used as input.

In this paper we next give background on reanalysis data available from climate models (Section 2.1), notably the NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) model used in our study and review current hydrometeorological approaches for the generation of gridded data from point observations that provide the foundation for our approach (Section 2.2). Section 2.3 describes the UEB snow and glacier melt model. We then describe the hydrometeorological downscaling methodology (Section 3) and software implementation of the downscaling tool (Section 4). We then describe the data from the western US (Sections 5.1–5.3) and results (Section 5.4) from evaluation of the methodology. Sections 6 and 7 give discussions and conclusions respectively. Technical model details are given in appendix A.

2. Background

2.1. Literature review on climate reanalysis data

Climate reanalysis datasets are commonly used to complement a limited observational record. Climate reanalysis data is produced by re-analyzing historic observations using a climate model that has unchanging parameters and equations based on known physics. They assimilate measurements of different atmospheric variables (temperature, pressure, precipitation etc.) from many sources to produce spatially complete, gridded meteorological variables at a continental or global scale (Kucera et al., 2013; Rienecker et al., 2011). Most reanalysis data are also temporally complete during the satellite era (1979 to present) and are typically generated at a resolution (hourly, 3-hourly and 6-hourly) sufficient to capture the diurnal variability (Rienecker et al., 2011). There are a number of reanalysis datasets available including from European Centre for Medium-Range Weather (Dee et al., 2011), NOAA/NCEP (Kanamitsu et al., 2002), Japanese 55-year Reanalysis (Ebita et al., 2011) and NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011). These datasets have proven to be valuable research tools in meteorology, climatology, and ecology (Rienecker et al., 2011) and an important source for obtaining forcing variables to drive hydrological models in data scarce regions such as the Himalayas in South Asia (Xie et al., 2007) and the Blue Nile Basin in Africa (Dile and Srinivasan, 2014). However, reanalysis precipitation and surface fluxes contain uncertainty because of model biases in long term climatology and limitations in reproducing the diurnal cycle. A recent study by Kishore et al. (2013) shows that the mean difference between the

seasonal precipitation from various reanalysis datasets in the Western Himalayas can be as high as 86% from the observed value. This study also shows that the performance of reanalysis precipitation substantially varies over different seasons and regions in India. Thus, the accuracy of the reanalysis data must be taken into account before using them in hydrologic applications. This need motivated us to evaluate the accuracy of downscaled meteorological data in the context of it being used to drive an energy balance snow melt model.

This work was done as part of a NASA applications project (Brown et al., 2014; Sen Gupta, 2014; Sen Gupta et al., 2015) whose goal was to evaluate and apply NASA technology in the developing Himalayan region. This dictated the use of MERRA and RFE2 (Southern Asia Daily Rainfall Estimate) products as primary data sources for the downscaling and hydrologic modeling.

MERRA is a near-real-time global climate reanalysis product developed by NASA's Global Modeling and Assimilation Office providing data covering the satellite era (1979 to present). MERRA is derived from the Goddard Earth Observing System version 5 (GEOS-5), NASA general circulation model (Rienecker et al., 2011; Suarez et al., 2008) and National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) analysis (Wu et al., 2002). Hourly temperature, wind speed, and relative humidity are available at a spatial resolution of 2/3° longitude by 1/2° latitude, and 3-hourly incoming shortwave and longwave radiation are available at a coarser resolution of 1.0° by 1.25° (Lucchesi, 2012). Assimilation of satellite precipitation and in-situ information reduces the uncertainty in climate variable fields in MERRA and makes the data more useful for a variety of applications including flood and drought studies (Kucera et al., 2013).

Given shortcomings in MERRA precipitation fields, RFE2 is an alternative source for precipitation data. The merits of precipitation data from these two datasets (i.e. RFE2 and MERRA) are discussed by Shrestha et al. (2008) and Reichle et al. (2011), respectively. RFE2 was favored in our Himalayan application due to its adoption by our regional collaborators (Shrestha et al., 2013). RFE2 is a NOAA high resolution (0.1° × 0.1°) daily observation-based precipitation product over South Asia (Bajracharya et al., 2014, 2015; Shrestha et al., 2013; Xie and Arkin, 1996; Xie et al., 2002). Rainfall Estimation (RFE2) daily total precipitation estimates are constructed using four observational input data sources: approximately 280 GTS stations, geostationary infrared cloud top temperature fields, polar orbiting satellite precipitation estimate data from SSM/I, and AMSU-B microwave sensors (Xie et al., 2002). Near real-time daily rainfall estimations are available for the Southern Asian domain (70°–110° East; 5°–35° North) at a spatial resolution of 0.1° by 0.1° beginning on May 01, 2001.

2.2. Literature review on hydrometeorological downscaling

The Parameter-elevation Regressions on Independent Slopes Model (PRISM) is a widely used approach to produce high-resolution climate data in North America. PRISM generates gridded estimates of annual, monthly, and event-based climatic variables such as maximum and minimum temperature, precipitation, and humidity using observational data at point locations, DEM, other spatial data, and local information (Daly et al., 1994, 1997; 2000, 2008). Variables at a target site are calculated by using linear regression, with regression weighting factors estimated based on elevation, terrain aspect, coastal proximity, and vertical air mass layering (Hunter and Meentemeyer, 2005). The spatial scale of PRISM outputs can be as fine as 800 m, reducing the adjustments required for application at the hydrological scale (~100 m). PRISM was not considered for this study, due to its products being limited to the US. We nevertheless anticipate that

the methodology developed in this paper for MERRA and RFE2 data could be used with PRISM data to produce model inputs at the hydrologic scale within the US.

Physically based hydrometeorological downscaling techniques such as MTCLIM (Hungerford et al., 1989), DAYMET (Thornton et al., 2012), MicroMet (Liston and Elder, 2006) and TopoSCALE (Fiddes and Gruber, 2014) distribute point-measured information over a modeling domain or downscale from either regional or global information to a distributed local modeling domain. MTCLIM provides algorithms for extrapolating meteorological forcing variables such as daily air temperature, precipitation, solar radiation, and relative humidity at a location of interest by using point measurements at weather stations (Zimmermann and Roberts, 2001). This approach constructs climate data at any elevation by adjusting the observed data collected at lower elevation climate stations. Meteorological variables are adjusted for elevation difference between the weather station and target site, slope, aspect, east-west orientation and leaf area index (LAI). The main objective of developing MTCLIM was to provide inputs to an ecological model for simulating plant growth in mountainous regions where observed data is sparse. DAYMET extends MTCLIM algorithms to produce gridded daily meteorological variables by interpolating observations at multiple sites across larger regions (Thornton et al., 1997, 2012; Zimmermann and Roberts, 2001). MicroMet is a quasi-physically based spatial and temporal downscaling model capable of producing high-resolution (30–1000 m) climate data over a wide range of landscapes (Liston and Elder, 2006). Using ground-based observations of air temperature, precipitation, relative humidity, wind speed, and direction within or near the area of interest, MicroMet calculates high-resolution gridded air temperature, precipitation, pressure, relative humidity, wind speed and direction, and shortwave and longwave radiation. Spatial interpolations use the Barnes objective analysis scheme (Barnes, 1964) and adjustments are made for elevation, topography, and cloudiness (Liston and Elder, 2006). The TopoScale model (Fiddes and Gruber, 2014) does not use point observations as input. Instead it takes input from ERA-Interim gridded data (Dee et al., 2011) using interpolation of pressure level data according to a high-resolution DEM elevation. The physical concepts in TopoScale are quite similar to those of MicroMet, and those that we implemented here, but an important idea introduced with TopoScale is the quantification of lapse rates from information at higher levels in the reanalysis data, rather than relying on climatological averages or values estimated from ground stations. This allows lapse rate adjustments to be time varying based on re-analyzed atmospheric conditions at each time step.

2.3. Utah Energy Balance snow and glacier melt model

The Utah Energy Balance model is a spatially distributed model that uses energy balance formulations to simulate the snowmelt and SWE over a watershed, driven by gridded weather inputs (Luce and Tarboton, 2010; Mahat and Tarboton, 2012; Tarboton et al., 1995; Tarboton and Luce, 1996; You, 2004). UEB is physically-based and tracks point energy and mass balances to model snow accumulation and melt. UEB has four state variables: surface snow water equivalent, W_S (m); surface snow and substrate energy content, U_S (kJ m^{-2} hr^{-1}); the dimensionless age of the snow surface η; and the snow water equivalent of canopy intercepted snow, W_C, (m). The model is driven by time-varying air temperature, precipitation, wind speed, relative humidity, and incoming shortwave and longwave radiation at time steps sufficient to resolve the diurnal cycle. Sen Gupta et al. (2015) provides a detailed description of the distributed version of UEB.

3. Downscaling methodology

MSDH was developed to generate 3-hourly grid surfaces of temperature, precipitation, relative humidity, wind speed, and shortwave and longwave radiation over a complex terrain watershed using MERRA and RFE2 reanalysis inputs and a high-resolution digital elevation model (DEM) of the target area or watershed. The choice of DEM resolution is left to the user based on the watershed area, source of the DEM, availability of computer disk space, resource constraints, and use of the data. Our choice of a 3-hourly time step was largely influenced by the need for the input variables in a physically based energy balance snowmelt model to quantify the diurnal cycle. This is a common requirement in the computation of surface energy balance so we anticipate that this approach has broad applicability. The model is capable of producing spatially distributed weather data without requiring any ground-based observations, which makes it suitable for use in data scarce watersheds. However, when observed data is available it can be used to derive location specific precipitation and/or lapse rate adjustment coefficients and bias correction factors that improve the quality of the downscaled data. Post processing bias correction adjustments can also be applied to other variables when there is data available to support such adjustments. While developing the tool, we considered the following criteria.

(1) Given the target application in data scarce remote locations, often in developing countries, the tool should be based on a free and open source software solution.
(2) The tool should have an easy-to-use graphical user interface to hide internal codes and file-folder complexity and to provide an intuitive visual environment.
(3) The data should be stored in a standard file format that can be accessed by readily available software tools.
(4) The computational complexity should be limited so that the software tool can be used on a personal computer (PC)

The MERRA variables used in this study are listed in Table 1 and can be accessed and downloaded from NASA's Goddard Earth Science Data and Information Services Center website. RFE2 data are available in gridded binary format via NOAA's National Centers for Environmental Protection (NCEP) ftp website (ftp://ftp.cpc.ncep.noaa.gov/fews/S.Asia/).

To start, MSDH automatically downloads the coarse scale MERRA and RFE2 input data for the range of dates and spatial bounding box specified by a user. Next, MSDH interpolates this data to the finer scale of the hydrologic grid. Bilinear interpolation is used and coordinate transformations are done at this step. This includes interpolation of the geo-potential height that is the reference elevation for re-analysis data. Then the difference in elevation between the DEM and interpolated geo-potential height is used to adjust each of the variables being modeled. For temperature, a lapse rate is calculated based on the MERRA surface temperature and the two nearest elevations above the MERRA surface elevation at each time step and grid point. This is used to adjust MERRA surface temperature to the elevation of the DEM. There is also an option for a user to input the lapse rate to be used, for example from nearby station data.

For humidity, MERRA specific humidity is used to calculate the dew point temperature, which is then adjusted for DEM elevations using a monthly vapor pressure coefficient and parameters in the saturation vapor pressure function for ice, relying on the relatively linear relationship between dew point temperature and elevation We then evaluate actual vapor pressure from air temperature and saturated vapor pressure from dew point temperature (Liston and Elder, 2006). Relative humidity is quantified as the ratio of these two quantities.

Horizontal wind speed magnitude was obtained from eastward and northward wind components from MERRA and was interpolated bilinearly and projected to the DEM grid resolution. Then, the effect of slope, aspect and curvature on wind speed was accounted for following Liston and Sturm (1998).

For solar radiation, a pressure based atmospheric attenuation coefficient was calculated for each time step and used to adjust MERRA incoming solar-radiation to the grid DEM elevation. Incoming longwave radiation was estimated based on downscaled air temperature, cloud cover and atmospheric emissivity.

Precipitation is adjusted, following Liston and Elder (2006) using the following non-linear relationship between elevation and precipitation

$$PRCP_{DEM} = PRCP_{MERRA}\left[\frac{1 + \kappa_p(Z_{DEM} - Z_{MERRA})}{1 - \kappa_p(Z_{DEM} - Z_{MERRA})}\right] \qquad (1)$$

where $PRCP_{MERRA}$ is the MERRA or RFE2 reanalysis precipitation interpolated at DEM cell location, and κ_p is a coefficient that quantifies how precipitation varies with elevation. Liston and Elder (2006) provided a table (Table A1, Appendix A) that gives globally averaged monthly κ_p values that we use as defaults. We also provide the capability for users to input values fit for their location from observations.

Appendix A gives full details of the downscaling methodology.

4. Software implementation

4.1. Implementing downscaling algorithms in R

R is a statistical software and scripting language initially developed for statistical analysis such as hypothesis testing, time

Table 1
Input MERRA variables used for downscaling.

MERRA Variable	Description	Spatial resolution (longitude × latitude)	Temporal resolution
t2m	Temperature at 2 m above the ground (K)	0.67° × 0.5°	hourly
v2m	Northward wind at 2 m above the ground (m s^{-1})	0.67° × 0.5°	hourly
u2m	Eastward wind at 2 m above the ground (m s^{-1})	0.67° × 0.5°	hourly
ps	Time averaged surface pressure (Pa)	0.67° × 0.5°	hourly
qv2m	Specific humidity at 2 m above the ground (kg kg^{-1})	0.67° × 0.5°	hourly
swgdwn	Surface downward shortwave flux (W m^{-2})	1.25° × 1.0°	3-hourly
t850	Temperature at 850 hPa (K)	0.67° × 0.5°	hourly
t500	Temperature at 500 hPa (K)	0.67° × 0.5°	hourly
t250	Temperature at 250 hPa (K)	0.67° × 0.5°	hourly
h850	Elevation at 850 hPa (m)	0.67° × 0.5°	hourly
h500	Elevation at 500 hPa (m)	0.67° × 0.5°	hourly
h250	Elevation at 250 hPa (m)	0.67° × 0.5°	hourly

series analysis and plotting, and linear and nonlinear modeling (Carslaw and Ropkins, 2012). R is also extensively used in environmental data analysis, visualization, and modeling. Open source, highly optimized coding functionality, extensibility, and simplicity contributed significantly to the large popularity of R. Users can extend its functionality by writing R packages, collections of well-structured reusable functions and data. These packages can be distributed to the entire R user group through a single web repository (Horsburgh and Reeder, 2014; Pinheiro et al., 2011). In this study, we used several existing R packages such as utils, ncdf (Pierce, 2011), rgdal (Keitt et al., 2011), and raster (Hijmans et al., 2013). We also used NetCDF Operators (NCO) (Zender, 2008) and Climate Data Operators (CDO) (Schulzweida et al., 2006) tools for efficient manipulation of netCDF files. NCO and CDO are both collections of operators for statistical and arithmetic processes, subsetting, interpolation, extrapolation, and transformation of geospatial time series data stored in netCDF files. The windows version of NCO and CDO program executables are called from R using the system() function.

First, a R function was developed to download MERRA and RFE2 files for the variables listed in Table 1 for a specified spatial and temporal extent using the binary file transfer method provided in the function download.file() from utils package. Next, for each netCDF file, all the MERRA and RFE2 variables are aggregated into three hourly time steps. Hourly MERRA data, such as temperature, is averaged over a three-hourly time step using NCO's ncra command. Then, daily RFE2 precipitation is uniformly distributed into three-hourly time steps using CDO's arithmetic process capability on netCDF datasets (Schulzweida et al., 2006).

A TIFF or image file of the DEM is read into R using rgdal's readGDAL() function and converted into a RasterLayer object. A RasterLayer object is single layer of raster data described by a set of parameters, such as number of columns and rows, spatial resolution, the coordinates of its spatial extent, and map projection. The DEM RasterLayer represents the domain and modeling grid that is the target for the downscaling. Then MERRA and RFE2 variables such as temperature and precipitation are read from netCDF files for each time step as a two-dimensional array. Using latitude and longitude bounding box information, the array is projected into another RasterLayer, then the netCDF RasterLayer is projected to the DEM RasterLayer using the projectRaster() function from raster package. This function of the raster package bilinearly interpolates the values of the netCDF RasterLayer to the extent and resolution of the DEM and transforms its projection to the DEM's projection (coordinate reference system, CRS). MERRA Geo-potential height in netCDF files are converted to a MERRA height RasterLayer with the resolution and spatial extent of the DEM. The conversion of multiple two-dimensional data objects to a uniform RasterLayer eases the implementation of the topographical adjustment algorithms described in Appendix A. Once the adjustment algorithms are implemented, the final RasterLayer of each output variable is converted into a two-dimensional matrix in R and appended onto a designated netCDF file that holds the downscaled result.

4.2. Output data storage in netCDF

The input and output gridded data used in MSDH are stored in netCDF files. NetCDF is a binary, multidimensional format commonly used by the oceanographic and atmospheric scientific communities for storing and managing scientific data. NetCDF4 (Rew et al., 2006) is a machine-independent format that allows direct access, shared access, visualization, and appending of new data to portable binary files. The output netCDF files of MSDH are always three-dimensional: (a) X (m), (b) Y (m) and (c) time (hours). Since the weather variables are produced at the surface, altitude is

not a required dimension. The performance of reading the data from the files depends on the ordering of dimensions within the file and the programming language used to read the data. In MSDH we provide an option to the users to choose the order of the dimensions in the file.

Each of the six weather variables is associated with six attributes, such as short name, long name, unit, a numeric value to represent the missing data, and a plausible range of values. All six variables are stored in the same netCDF file with a data array for each variable corresponding to the same set of dimension vectors. A large volume of data might be generated if the program is run for multiple years or at a very high spatial resolution or combination of these two. To avoid storing a large volume of data in a single netCDF file, a separate file is created for each month. The temporal sequence of the data between multiple files is maintained by incrementing the time dimension from "time of origin" or start time. The units of time dimension stores the start time in each file.

4.3. MSDH graphical user interface

Using R packages is a relatively straightforward task for experienced users, but it can be challenging, with a steep learning curve, for beginners with no prior programming experience. We, therefore, developed a GUI in order to create a visual environment for the users to enter inputs and execute the R functions. The GUI was coded in R using the RGtk package and the R script runs from a C# wrapper program. Thus, the MSDH GUI hides the R code from the user and enables data downloading and downscaling tasks while eliminating the complexity of creating or editing codes, files, and folders.

The MSDH has three main tabs: (1) data download, (2) coefficient calculations and (3) data downscale. The "data download" tab (Fig. 1) provides an option for the users to download data for the variables listed in Table 1 using R's utils package. Precipitation can be downloaded from either RFE2 or MERRA. MERRA data is available globally, while RFE2 covers only the South Asian region, but with better resolution.

The "coefficients calculations" tab performs the optional task of calculating time varying lapse rate and precipitation adjustment coefficient using observational data from the site network within the target domain or a watershed.

The "data downscale" tab performs the four-step downscaling methodology described in Appendix A. The user only needs to specify a DEM of the target spatial domain (in image/TIFF format). The user is provided with a capability to choose the source of the precipitation adjustment factor from a set of options, such as (1) default specified by Liston and Elder (2006), (2) calculated from the "Coefficients calculations" and (3) user input.

5. Evaluation using western US SNOTEL sites

5.1. SNOTEL data sources

To test the downscaling methodology, software implementation, and test that the GUI functioned as intended, MSDH was run for one water year starting from October 2009 to September 2010 over the an area between 36.15° and 43.23° N latitude and 108.90°−121.92° W longitude (Fig. 2) at 120 m resolution. 173 U.S. Department of Agriculture snowpack telemetry (SNOTEL) sites are located within the study area (Appendix B). The elevation of the sites ranges from 1777 m to 3816 m, with an average elevation of 2537 m. Daily historical minimum, mean, and maximum temperature; daily precipitation; snow depth; and SWE data available at these sites was used to test the downscaling. Along with daily temperature and precipitation, hourly temperature, precipitation,
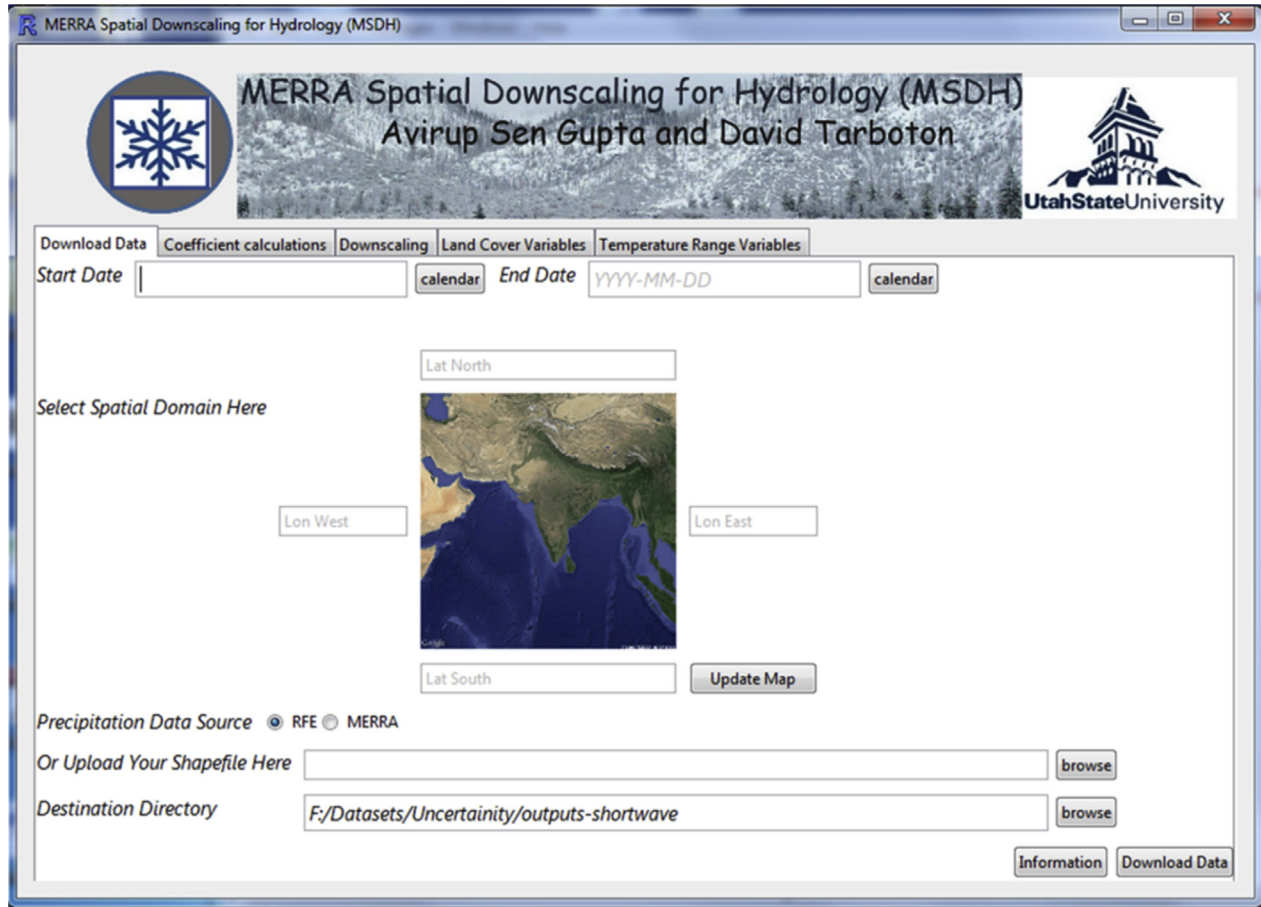
Fig. 1. Graphical user interface for MERRA spatial downscaling for hydrology (MSDH).

wind speed, relative humidity, and incoming shortwave radiation were available at the USU Doc Daniel site (Appendix B) from October 2009 to June 2010 from a separate study by Mahat and Tarboton (2012, 2013) and Mahat et al. (2013). This data was used to compare the downscaled relative humidity, solar radiation and wind speed data, and to conduct a sensitivity analysis.

The DEM of the Western United States was divided into a number of small regions to prevent MSDH output files from becoming too large. Fig. 3 shows the downscaling steps for the Logan River watershed in Utah where six SNOTEL sites are located. MERRA temperature data was downloaded for the contiguous United States (Fig. 3 (a)) and the four grid cells spanning the Logan River watershed (Fig. 3 (b)) were used in bilinear interpolation to obtain gridded temperature at the scale of the DEM (Fig. 3 (c)). This involved using R's raster library projection transformation capability to transform the data into the DEM's Universal Transverse Mercator (UTM) projection system and clip it to the extent of the DEM. This raster layer contains bilinearly resampled temperature data, while its spatial domain, resolution, and number of rows and columns are exactly the same as the DEM. Next, temperature was adjusted using the lapse rate and the difference between MERRA elevation and DEM elevation using the methodology described in Section 3 and Appendix A. This procedure was repeated for all time steps and grid cells. Other variables, such as incoming shortwave radiation and wind speed, were also downscaled to the DEM spatial scale using the physically based methodology described in Section 3. Precipitation was adjusted using equation (1) and bias corrected using equation (A18) using the average of SNOTEL stations within each MERRA grid cell.

Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE) and bias (BIAS) were used to compare the downscaled variables with observations. These are defined as follows:

$$NSE = 1 - \frac{\sum_{t=1}^{n} (Obs_t - Sim_t)^2}{\sum_{t=1}^{n} (Obs_t - Obs_{mean})^2} \qquad (2)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (Obs_t - Sim_t)^2}{n}} \qquad (3)$$

$$BIAS = \frac{1}{n} \sum_{t=1}^{n} (Sim_t - Obs_t) \qquad (4)$$

where Obs_t and Sim_t are observed and simulated values at any time step t, Obs_{mean} is the mean of observed values and n is the number of observations. NSE is a dimensionless metric quantifying error relative to variability, while RMSE and BIAS have the units of the quantity being evaluated and is representative of the scale of the error. NSE ranges from 1 for observations equal to simulations to 0 if simulations are no better than just picking the mean and may extend into negative values for even worse performance. Guidance on the interpretation of NSE is variable, but it is common practice to interpret the ranges <0.5 as poor, 0.5 to 0.65 satisfactory, 0.65 to 0.75 good, and >0.75 as very good (e.g. Kalra and Ahmad, 2012; Moriasi et al., 2007).
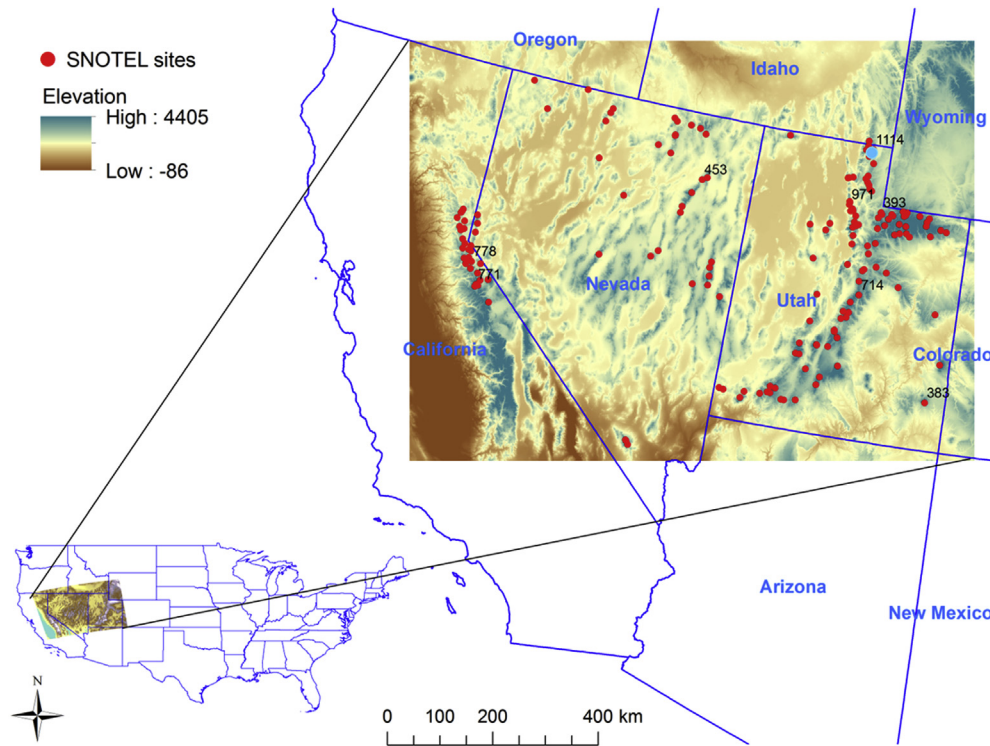
Fig. 2. Locations of the SNOTEL sites used in this study. Blue lines indicate state boundaries and red dots symbolize the SNOTEL sites. Utah State University Doc Daniel site is shown as a blue dot and 8 sites that are reported in Fig. 7 and Table 4 are shown by their station ID number. The Digital Elevation Model (DEM) from the National Elevation Model dataset shows the variability in terrain surface elevation (meters)4.2. Evaluation of Downscaling.

5.2. Detailed evaluation at USU Doc Daniel site

All five downscaled variables were compared with observations at the USU Doc Daniel site for the period October 2009 to June 2010 (Table 2, Fig. 4).

Fig. 4 shows that the downscaled data captures the seasonal pattern of low temperatures in December and high temperatures in June quite successfully. The very good NSE of 0.87 reflects this and demonstrates the model's capability to successfully reproduce observed temperature. Both downscaled incoming shortwave radiation and relative humidity capture the seasonal cycle of the observed data reasonably well; however, they fail to reproduce some short term changes and appear to fluctuate at smaller amplitude than the observations at short time scales for some months. This is reflected in their somewhat lower NSE (Table 2). Nevertheless, the NSE values obtained indicate the method's capability to reproduce these two variables at a "satisfactory" level. Compared to these variables, wind speed and precipitation perform rather poorly (i.e., precipitation NSE = 0.28 and wind speed NSE = 0.16). The wind discrepancies likely reflect the challenge in representing local (DEM grid scale) wind variability from regional information, while precipitation discrepancies originate both in the driving MERRA data and downscaling. Although 96% of precipitation events were simulated successfully by MERRA, it produces a considerable number of non-observed rainfall events with low magnitudes and fails to simulate the magnitude of observed rainfall events (Fig. 4, bottom right panel). Less intense precipitation events are often overestimated, and moderately heavy events are underestimated.

Note that in an earlier implementation version (Sen Gupta, 2014) we used linear regression between the mean monthly temperature gauge data and gauge elevation to calculate the monthly lapse rate. Precipitation adjustments used the same adjustment factor everywhere based on a single fitting of gauge precipitation and elevation. Here we have extended the methodology to use atmospheric profile information from MERRA at each time step for calculation the lapse rate. The monthly precipitation adjustment factor was calculated for each gauge as a nonlinear function of observed precipitation at a gauging stations and their respective elevation (compared to the MERRA grid elevation) at the grid cell in which the site is located. Quantitatively, the NSE for temperature remained about the same, while the precipitation NSE improved indicating the merit in this approach.

5.3. Broad evaluation across SNOTEL sites

Downscaling of daily maximum, minimum and mean air temperature (T_{max}, T_{min}, T_{mean}), and daily and monthly precipitation was evaluated at SNOTEL sites for water year 2010. Fig. 5 gives scatter plots of observed data at SNOTEL sites and downscaled data at DEM grid cells where those sites are located. Table 3 shows NSE, RMSE and BIAS between the observed and downscaled data for all the sites. Table 3 also reports these statistics between the observed and bilinearly interpolated MERRA data, without elevation adjustments. The difference reflects the value added due to application of the downscaling using high resolution topography data. In each variable, the downscaled data performed better which illustrates the added value of the downscaling approach over simple bilinear interpolation. Both daily T_{max} and T_{mean} show NSE of about 0.85. MSDH downscaling methods improve the daily T_{max} simulation by NSE of 0.63 and RMSE by 5.79. For Daily T_{min} and T_{mean} the improvements from the downscaling are relatively small. Downscaled T_{min} shows slightly lower NSE (0.74) compared to daily T_{max} and T_{mean}, indicating slightly lower performance in reproducing daily minimum temperature. Monthly aggregated downscaled precipitation also performs well against the observed SNOTEL
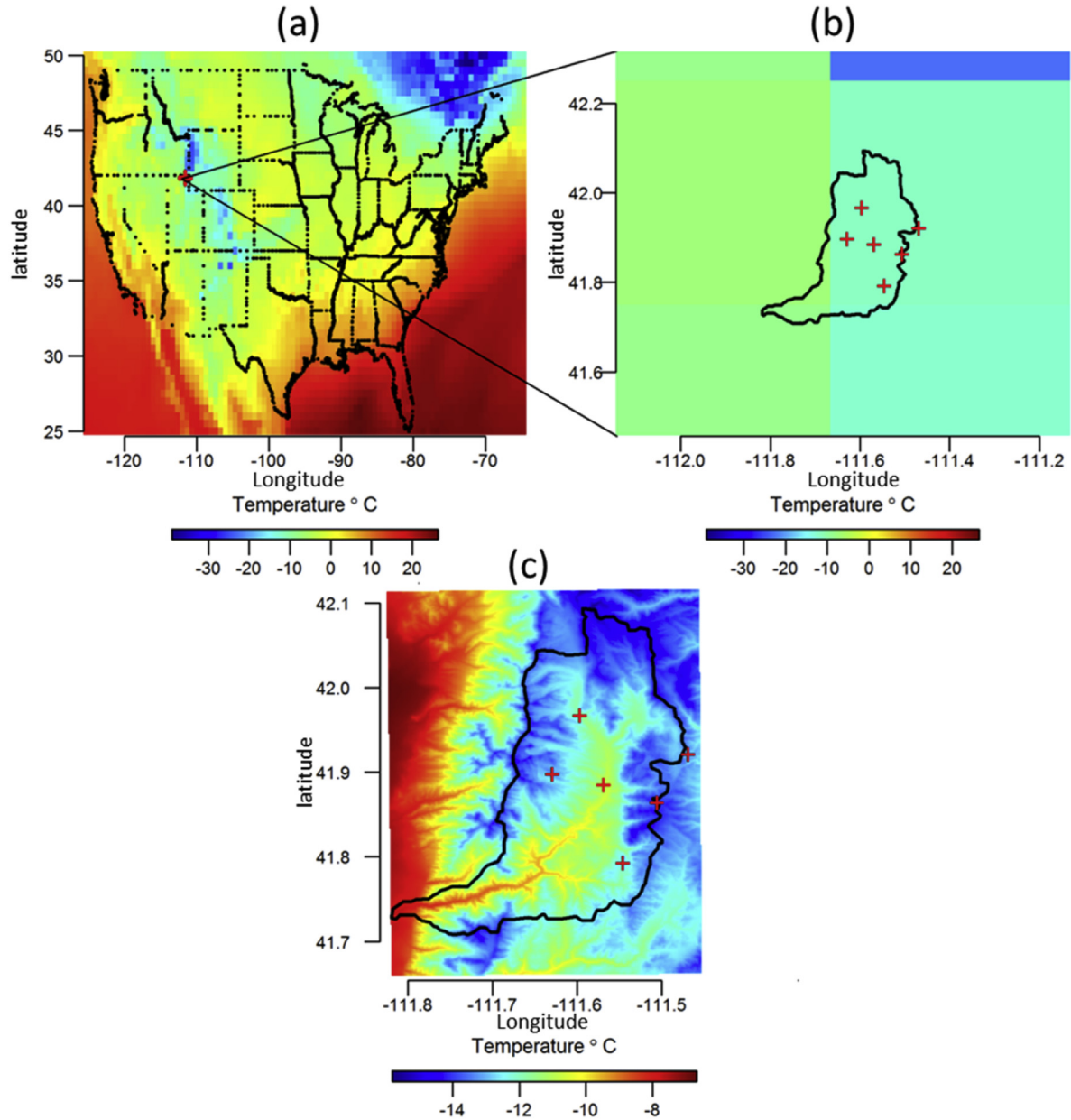
Fig. 3. Downscaled MERRA temperature (°C) for the Logan River watershed 18:00 UTC on Dec 24, 2009 (a) temperature reported in MERRA for Contiguous USA; (b) MERRA grid cells spanning Logan River watershed and surrounding areas and (c) downscaled temperature at DEM grid resolution.

measurements with NSE of 0.72 and RMSE of 23.83 mm. However, at daily time steps, precipitation simulation incorporates moderate to high uncertainty, especially during the late winter and early spring season. The relatively low NSE value (0.44) for daily precipitation (Table 3) indicates high uncertainty in precipitation downscaling at shorter time step. However, the performance improvement is significant when compared with the bilinearly interpolated MERRA precipitation, both at daily and monthly scale (Table 3).

Fig. 6 further complements Fig. 5 and Table 3 for model evaluation. The Taylor diagram provides a visual representation of multiple statistics quantifying the correspondence between the observed and modeled variables. Taylor diagrams in Fig. 6 (a–e) summarize model performance at each site for daily T_{mean}, T_{min}, T_{max}, precipitation and monthly precipitation, respectively. Each

point (red dot) on the polar style graph represents three statistics: normalized standard deviation, normalized centered root mean square error (CRMS, equation (5)) and correlation between the observed and modeled values at a SNOTEL station. The linear distance from the origin (0, 0) to any given point is the ratio of the modeled standard deviation to the observed standard deviation. The azimuthal position of a point shows the correlation coefficient between the observation and modeled data. The distance from each point to the point (1, 0) on the x-axis is the normalized CRMS for that point. The normalization is done with respect to the standard deviation of the observations. Together, these statistics are an easy and powerful way to depict the overall correspondence (how close to point 1 on x-axis), correlation and reproduction of variability (nearness to arc at radius 1 from the origin) for any given variable.

$$CRMS = \sqrt{\frac{\sum_{t=1}^{n} ((Obs_t - Obs_{mean}) - (Sim_t - Sim_{mean}))^2}{n}}$$

(5)

For T_{mean}, T_{min}, T_{max}, correlation is usually higher than 0.9, normalized CRMS is lower than 0.6 and modeled standard deviation is little higher than (up to 1.5 times) the observed standard deviation at most sites. The model's performance for daily and monthly precipitation varies widely from one site to another. A majority of sites show correlation of 0.6 or higher with the observations. However, normalized standard deviation less than 1.0 indicates under-dispersion of downscaled precipitation (both daily and monthly scale) compared to observed values for a majority of SNOTEL sites. Also, high normalized CRMS indicates lower precision of the model for precipitation.

5.4. Integrated evaluation using UEB snowmelt model

The Utah Energy Balance (UEB) snowmelt model was run using the downscaled data at 153 of the SNOTEL sites where SWE data was available to evaluate the simulation of snow accumulation and melt variability. The NSE and other statistics were evaluated and reported for the sites with best and worst NSE as well as sites ranked at 10th, 25th, 40th, 60th, 75th and 90th percentiles over the range of NSE obtained (Fig. 7, Table 4). Relative difference in peak SWE, P_{Diff}, (equation (6)), peak day difference, Δt (equation (7)), and volume ratio, V_{ratio} (equation (8)), were also evaluated.

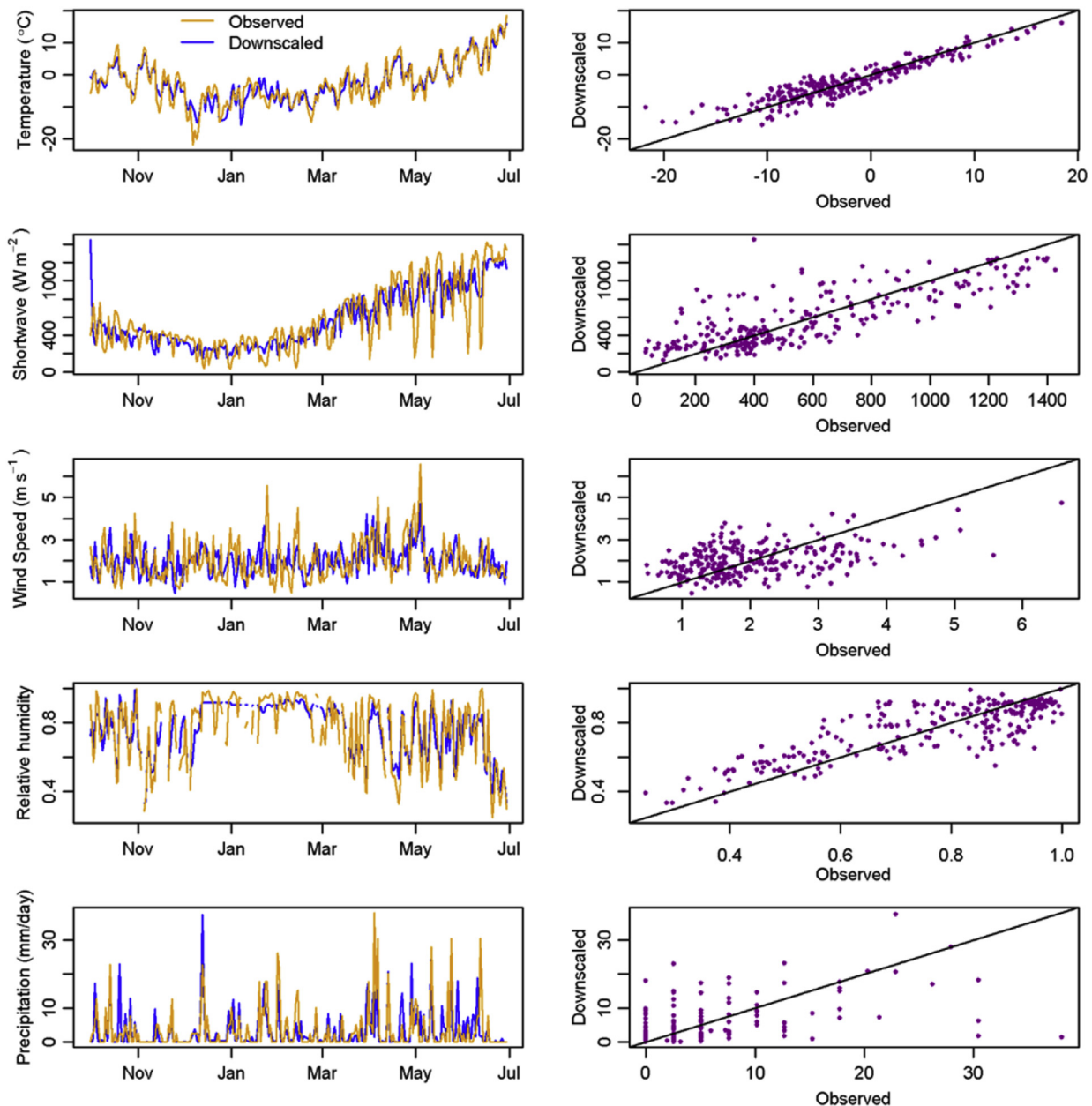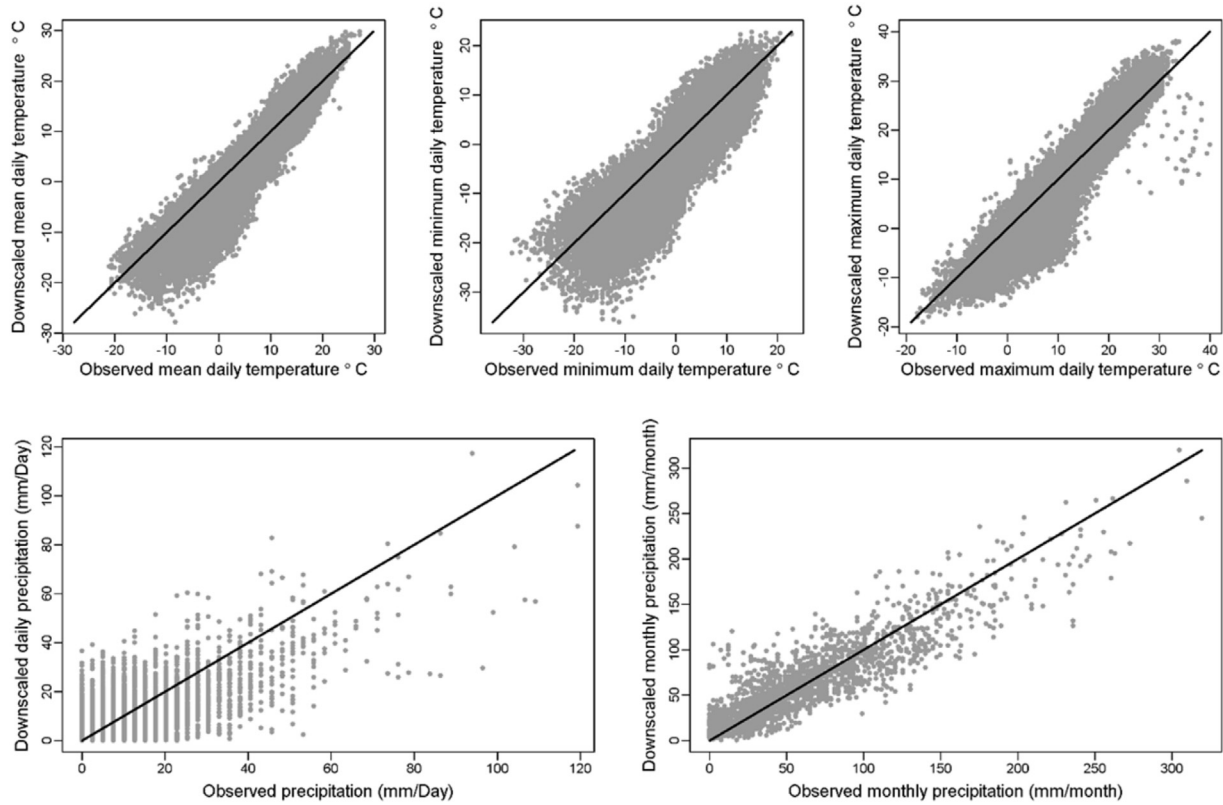$$P_{Diff} = \frac{[Max(SWE_{obs}) - Max(SWE_{sim})]}{Max(SWE_{obs})}$$

(6)

Fig. 4. Comparison of downscaled daily mean temperature, incoming shortwave radiation, wind speed, relative humidity, and precipitation with respect to measured data at the USU Doc Daniel SNOTEL site. A time series plot (left) and scatter plot (right) of observed and downscaled data are shown for each variable.

Fig. 5. Comparison of the downscaled data (y-axis) for daily mean, minimum and maximum temperature, daily and monthly precipitation with observed data (x-axis) at 173 SNOTEL sites for water year 2010 (Oct 01, 2009—Sep 30, 2010). The straight line at 45° indicates complete agreement between the observed and simulated data.

where $Max(SWE_{sim})$ is the modeled peak SWE, and $Max(SWE_{obs})$ is the observed peak SWE.

$$\Delta t = t(Max(SWE_{obs})) - t(Max(SWE_{sim})) \qquad (7)$$

where $t(Max(SWE_{obs}))$ is the date of the observed peak SWE and $t(Max(SWE_{sim}))$ is the date of the modeled peak SWE. $\Delta t = 0$ indicates peak SWE occurred on the same day for the observed and modeled SWE while a positive Δt indicates peak modeled SWE occurred earlier than the observed peak SWE and a negative Δt indicates peak modeled SWE occurred later than the observed peak SWE.

$$V_{ratio} = \frac{\sum_{t=1}^{n} Mod_t}{\sum_{t=1}^{n} Obs_t} \qquad (8)$$

where Mod_t and Obs_t are modeled and observed daily SWE respectively. The model SWE is produced at 3 h time steps, and the value at noon was used for comparison with daily observations. V_{ratio} compares the integral under the modeled and observed SWE curve; a value greater than 1 indicates that the model generally overestimates the SWE while a value less than 1 indicates the model generally underestimates the SWE over a snow season.

In Table 4, the absolute value of peak difference (P_{Diff}) and peak day difference decreases (Δt) from the top to bottom as the NSE values increase, while volume ratio (V_{ratio}) is more than double at Spratt Creek, but it is close to 1 for the best site, Parrish Creek.

At about 61% of sites integral volume and SWE peaks are underestimated reflecting either over prediction of melt or under prediction of snow precipitation. The model predicts the peak date very well (±5 days) at 48% of the sites. Only about 22% of the sites have a discrepancy more than 20 days in SWE peak date. Nearly 80%

of the sites have NSE higher than 0.6 and volume discrepancy less than 35%.

At the USU Doc Daniel SNOTEL site where we had observed temperature, precipitation, relative humidity, wind speed, shortwave and longwave radiation we ran the model driven only by observations. Then, for each input in turn we replaced the observed data by the downscaled data to study the error introduced by each input separately (Fig. 8).

The SWE simulation using the observed data matches the seasonal accumulation and ablation pattern reasonably (Fig. 8 a) with small underestimation of the SWE during the accumulation period and overestimation of the SWE during the melting season. The seasonal pattern and RMSE did not change significantly when observed temperature, wind speed and relative humidity data were replaced by downscaled data (Fig. 8 b, d and e). However, the performance decreased when precipitation and solar radiation downscaled data were used for the simulation (Fig. 8 c and f). For precipitation this is not surprising as precipitation statistics are poorly reproduced (Fig. 4, Table 2) a problem common with reanalysis data (Kucera et al., 2013) and indicating the sensitivity of UEB to precipitation inputs. For solar radiation this finding is a bit surprising as it was reasonably well downscaled (Table 2, Fig. 4). Nevertheless close examination reveals that there is a small positive bias in modeled solar radiation (Table 2). This is more so in the early season and manifests in the snow accumulation starting late and being underestimated. This reflects the sensitivity of UEB simulations to solar radiation inputs and underscores the importance of reducing solar radiation errors wherever possible.

We examined the sensitivity of the model to percentage changes in the two variables (shortwave radiation and precipitation) where the effects of downscaling are largest. In these sensitivity runs we
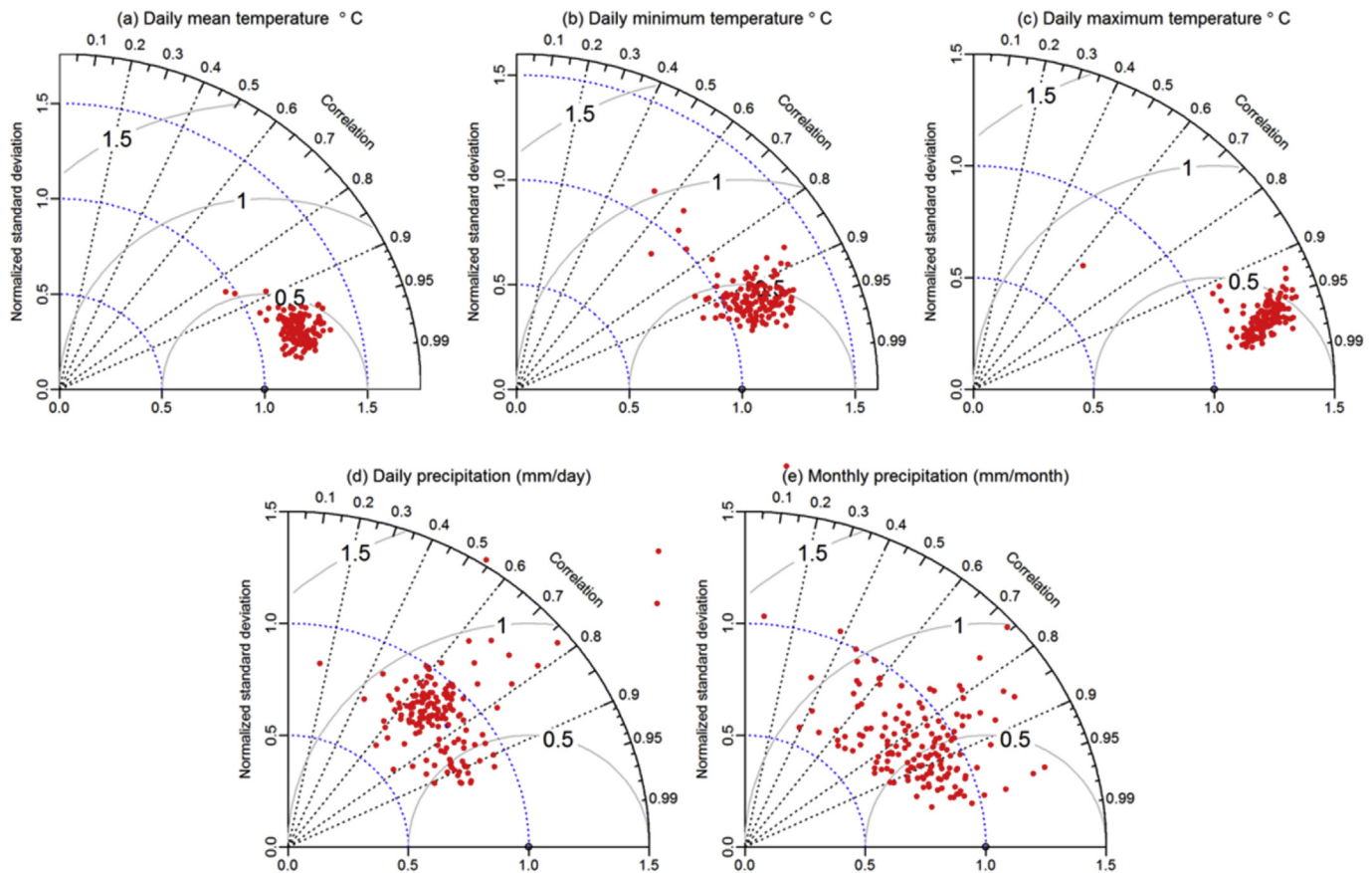
Fig. 6. Taylor diagram statistical comparison between the observations and downscaled estimates of daily mean, minimum and maximum temperature (T_{mean}, T_{min}, T_{max}) and precipitation and monthly precipitation (mm/month) at 173 SNOTEL sites.

increased each variable by 10% from the observed values. For shortwave radiation this increased the melt and thus reduced the SWE below observations increasing the RMSE by 0.034 m. This is consistent with the sensitivity to downscaled solar radiation (Fig. 8 f). For precipitation a 10% increase, increased the SWE generally and reduced RMSE by 0.007 m, indicating the smaller general sensitivity to precipitation than solar radiation.

6. Discussion

While developing MSDH, we recognized a number of limitations in downscaling methodologies and input data. As described by Liston and Elder (2006), this is a one-way approach where the vertical feedback between the near-land surface and atmosphere is ignored. While surface conditions such as presence of the canopy, soil moisture, and proximity to the water can have substantial impact on the local climate, MSDH adjusts the variables based on elevation alone.

In general, the interpolation of GCM outputs increases the spatial precision of the data, though often at a cost of increased uncertainty (Skelly and Henderson-Sellers, 1996). However, here the hydrometeorological downscaling approach has been designed to add value by adjusting to the local terrain and correcting some bias. Nevertheless the uncertainty of the original data may persist or increase when downscaling is performed. To evaluate this we compared both direct bilinearly interpolated and downscaled data against observations and found that downscaled temperature and precipitation data were always closer to observations than direct bilinear interpolations, evidencing the value of this

hydrometeorological downscaling.

Rienecker et al. (2011) explained many limitations of MERRA data including: (1) poor performance in capturing the diurnal temperature pattern by underestimating daily maximum and overestimating daily minimum temperature, (2) deviation of 3 °C or more from observations for daily temperature estimates, (3) short heavy precipitation events often simulated as precipitation drizzles and, (4) low solar radiation during daytime precipitation events often over estimated. These inaccuracies in MERRA are directly translated into the downscaled data and are responsible for some of the discrepancies found in Section 4.2.

The hydrometeorological downscaling evaluated here does depend on lapse rates for temperature and humidity as well as physical relationships of other variables (such as solar radiation) with elevation. Uncertainty in these input lapse rates may be a further source of error in the results. For temperature, we followed the approach of Fiddes et al. (2014) in using values higher up in the re-analyzed atmosphere profile to obtain a lapse rate specific for each time step and downscaled location. This is an advance over the more common approach of using climatological lapse rates. There is also an opportunity to do this for humidity or dew points; however some complexities are involved due to humidity being bounded by saturation. Our implementation thus retains a climatological dew point lapse adjustment following Liston and Elder (2006). Improvement of this approach using re-analyzed atmospheric profile information is left open for future research.

From this study it was apparent that reproduction of precipitation with a reasonable accuracy at a daily scale, or even at a monthly scale, was a challenge as manifested by the NSE values of
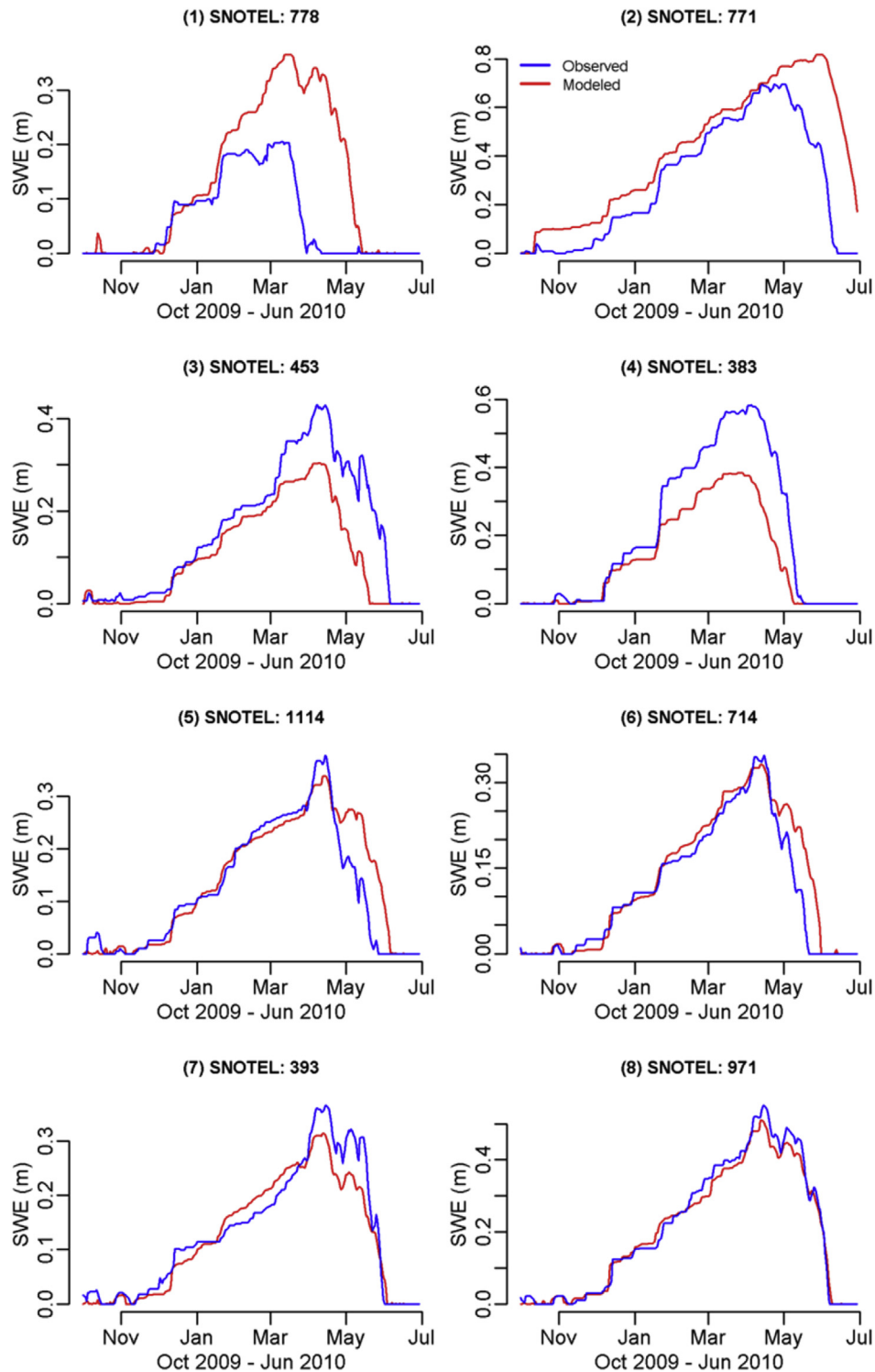
Fig. 7. Comparison between observed SWE and Utah Energy Balance (UEB) simulated SWE for water year 2010 (Oct 01, 2009–Sep 30, 2010) at eight selected SNOTEL sites.

0.44 at daily and 0.72 at monthly scale. We also noted a small, but meaningful positive bias in solar radiation. When the only down-scaled variable used as input to the model was solar radiation (Fig. 8 f) the start of snow accumulation is delayed and overall there is an under simulation of accumulation. From this we infer that even though the NSE for incoming solar radiation is relatively good overall (i.e., NSE = 0.65), and that discrepancies in Fig. 4 are hard to discern, the cumulative discrepancy in downscaled incoming solar

radiation results in erroneous melting too early and hence under simulation of the peak snow water equivalent. Interestingly, sensitivity analysis of downscaled variables at USU Doc Daniel SNOTEL site revealed that despite discrepancies in precipitation (i.e., NSE = 0.28), better reconstruction of snow water equivalent was obtained with downscaled precipitation inputs and other observed inputs (Fig. 8 c) than for solar radiation. UEB appears to be capable of producing good reconstruction of seasonal-scale SWE as

Table 2
Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE) and bias (BIAS) metrics at USU Doc Daniel site.

Variable name	NSE	RMSE	BIAS
Temperature (°C)	0.87	2.44	0.07
Shortwave radiation (Wm^{-2})	0.65	209	17.07
Wind speed (m/s)	0.16	0.85	−0.06
Relative Humidity	0.64	0.12	−0.02
Precipitation (mm/day)	0.28	5.23	−0.171

long as the aggregated precipitation matches with the observation during the accumulation season, regardless of the precise timing. This indicates a need to examine ways to improve incoming solar radiation downscaling in addition to precipitation downscaling.

Despite all the limitations, Nash-Sutcliffe efficiency (NSE) measures were greater than 0.70 for direct comparison of downscaled daily temperature and monthly precipitation at 173 SNOTEL sites. In an integrated test driving the Utah Energy Balance (UEB) snowmelt model 80% of these sites gave NSE > 0.6 for snow water equivalent. These findings motivate use of this tool in data sparse regions where ground based observations are not available and downscaled global reanalysis products may be the only option for model inputs.

Computational performance is another consideration to evaluate associated with hydrometeorological downscaling. The MSDH program's run time varied significantly depending on the number of rows and columns in the DEM raster file, as the process that takes the majority of runtime is interpolation of variables from MERRA to DEM resolution. The Logan River watershed used here consisted of 420 × 254 grid cells. Constructing data for six variables at 3-hourly time steps for a single month on this grid takes about an hour on a common commodity workstation (Dell Optiplex 780, with Intel Q9650 processor @ 3.0 GhZ and 8 GB RAM).

7. Conclusions

We have developed spatial hydrometeorological downscaling methods that adapt approaches from the MicroMet, DAYMET and MTCLIM to address the problem of downscaling climate reanalysis data for application on a fine resolution (30–100 m) grid over a watershed. This has general application in distributed hydrologic modeling, and was evaluated here for the generation of inputs to the Utah Energy Balance (UEB) snowmelt model. Variables downscaled include: temperature, precipitation, wind speed, relative humidity, shortwave and longwave radiation. The model produces 3-hourly, high resolution, gridded weather data for input to a spatially distributed hydrologic model. NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) climate products and Southern Asia Daily Rainfall estimate (RFE2) data are the major inputs to the program. In the first step of a two-step downscaling approach, we bilinearly interpolate RFE2 or MERRA reanalysis data to a high resolution digital elevation model (DEM) grid. In the second step, we make topographic adjustments using well-established relationships of elevation, slope, aspect, curvature, and cloudiness with the selected variables. The methods developed here are not limited to MERRA, and could be extended to any GCM, reanalysis, or regional climate model output or forecast that produced the same input quantities used by MSDH (Table 1), although use of other product inputs should be supported by further evaluation, as we have presented here for MERRA.

Development of MSDH was necessary for constructing topographically adjusted high resolution meteorological data to drive hydrological models in data scarce regions. Reanalysis data such as MERRA were developed to analyze the earth system at global or continental scales, whereas hydrological decision making for water availability and flood forecasting, for example, are studied at the watershed level. MSDH can be used as a tool to bridge the gap between the spatial scales of data and used in these two scientific domains. MSDH is capable of producing data at any grid resolution specified in an input DEM. The example application of the system produced the gridded surface of six variables at 120 m resolution and 3-hourly time steps for the Logan River watershed for 1 year starting on October 1, 2009. The data was then used to drive the Utah Energy Balance (UEB) snowmelt model to simulate one year of snow accumulation and melt. Daily temperature, shortwave radiation, relative humidity and monthly precipitation and UEB simulated SWE showed reasonably good agreement with the observations, indicating MSDH's capability to making estimates of good quality high resolution climate data using very limited

Table 3
Comparison between the bilinearly interpolated MERRA and downscaled daily mean, minimum and maximum temperature and daily and monthly precipitation at the NRCS SNOTEL sites. Nash-Sutcliffe Efficiency (NSE), BIAS and RMSE are used as performance evaluation statistics for the comparison.

Statistical criteria	Daily mean temperature (°C)		Daily minimum temperature (°C)		Daily maximum temperature (°C)		Daily precipitation (mm/day)		Monthly precipitation (mm/month)	
	Bilinearly interpolated	Downscaled	Bilinearly interpolated	Downscaled	Bilinearly interpolated	Downscaled	Bilinearly interpolated	Downscaled	Bilinearly interpolated	Downscaled
NSE	0.83	0.84	0.71	0.74	0.23	0.86	<0	0.44	0.11	0.72
RMSE	4.12	3.98	4.90	4.72	9.99	4.20	4.11	3.73	41.54	23.83
BIAS	−2.92	−1.18	2.94	−1.86	−9.96	−0.52	−0.81	0.008	−24.41	0.21

Table 4
Site metadata and performance statistics (NSE, R-squared, percent bias, peak difference, peak day difference and volume difference) of observed and modeled SWE at eight selected sites.

Site	Site ID number	State	Elevation (m)	NSE	RMSE (m)	Peak Difference, P_{Diff}	Peak day difference, Δt	Volume ratio, V_{ratio}
Spratt Creek	778	CA	1864	<0	0.12	−0.79	−7	2.24
Sonora Pass	771	CA	2690	0.35	0.20	−0.18	−38	1.45
Dorsey Basin	453	NV	2469	0.65	0.08	0.30	−3	0.67
Camp Jackson	383	UT	2733	0.72	0.11	0.34	10	0.64
Garden City Summit	1114	UT	2348	0.82	0.04	0.11	0	1.09
Red Pine Ridge	714	UT	2746	0.88	0.03	0.05	3	1.13
Chalk Creek #2	393	UT	2487	0.92	0.03	0.12	2	0.94
Parrish Creek	971	UT	2359	0.98	0.02	0.08	2	0.96
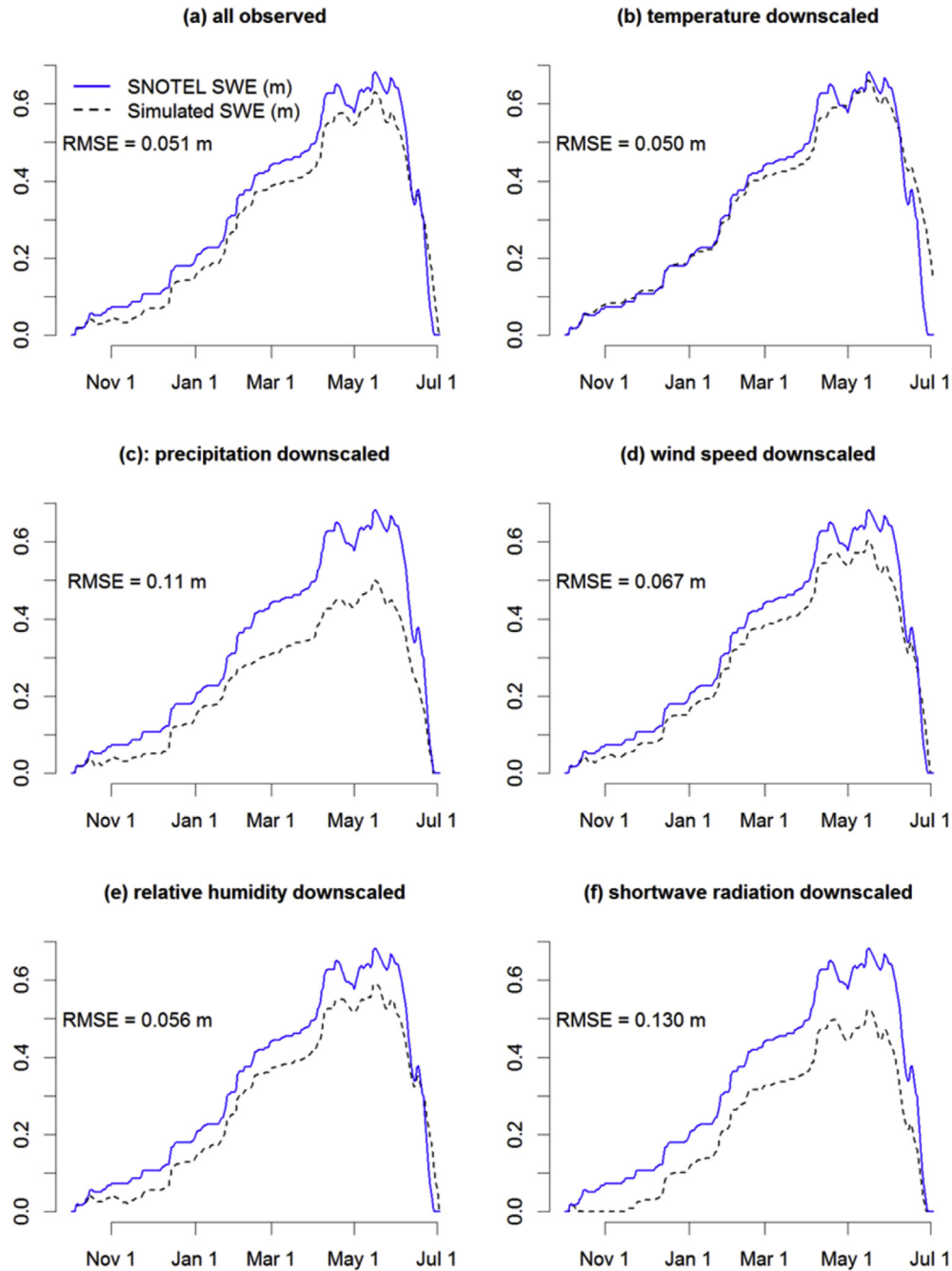
Fig. 8. Comparison between the observed and UEB simulated snow water equivalent (SWE) at the USU Doc Daniel SNOTEL site using (a) observed temperature, precipitation, wind speed, relative humidity and shortwave radiation, (b) downscaled temperature with observed data of other variables, (c) downscaled precipitation with observed data of other variables, (d) downscaled wind speed with observed data of other variables, (e) downscaled relative humidity with observed data of other variables, (f) downscaled shortwave radiation with observed data of other variables.

observational data.

This study showed that it is possible to obtain the input variables required to drive the UEB model entirely from climate reanalysis data extending its applicability to data scarce regions of the world. The discrepancies that result due to errors in the reanalysis data and downscaling model were quantified for a location in the US where there is detailed data available. Comparison between SNOTEL observations and the Utah Energy Balance Snowmelt Model-simulated snow water equivalent indicates the degree (i.e., mean Nash-Sutcliffe efficiency = 0.67) to which this method is effective. Sources of discrepancies, in terms of precipitation, and

solar radiation uncertainty were identified and motivate opportunities for future research to reduce uncertainty and improve simulations. These discrepancies need to be factored into the use of simulations driven by downscaled results for hydrological modeling and analysis.

The tool was developed using open source, freely available scripting language and programs. The R code is publically available in bitbucket (https://bitbucket.org/AvirupSenGupta/msdh.usu) so that the user community outside the initial development team can participate in future improvements of the software by integrating new approaches and analysis techniques. The program has a

graphical user interface (GUI) to make it accessible to users unfamiliar with R. Downscaled data is saved in CF-convention compatible three dimensional self-describing netCDF format, which makes the data portable across operating systems and accessible and displayable in a number of freely available software tools such as ncdump, ncBrowse, and Integrated Data Viewer (IDV; http://www.unidata.ucar.edu/downloads/idv).

The application demonstrated in this paper was successfully run on a PC with the Windows operating system. This is particularly advantageous for developing countries where students, engineers, or even researchers may not have access to the latest model high performance computing systems. Presently, MDSH has only been tested on Windows-based systems. The availability of R and all other required programs, such as NCO and CDO in UNIX/Linux operating systems suggests that the program could be ported to UNIX/Linux based computers with little code modification.

Acknowledgements

Appendix A. Downscaling methodology used in MSDH

Variables listed in Table 1 correspond to the elevations that are specified by geopotential height in MERRA's NASA general circulation model (Rienecker et al., 2011). Geopotential height is reported at the same spatial resolution with the corresponding variable and is constant over time. MSDH downscaling techniques follow a four-step procedure: (1) perform temporal averaging of MERRA hourly temperature, precipitation, eastward and northward wind speed, specific humidity, and pressure in 3 h blocks, (2) project MERRA data to the spatial projection of the DEM, (3) distribute the MERRA elevations and meteorological variables from MERRA resolution to DEM resolution using bilinear interpolation and (4) use known relationships between climate variables with elevation, slope, aspect, curvature and cloudiness to parameterize the effect of topography. RFE2 precipitation is reported as total daily values; thus, to obtain 3-hourly precipitation, we distribute the daily precipitation equally, assuming uniform precipitation throughout the day. In the third step, bilinear interpolation at any point on the DEM grid uses four surrounding MERRA grid cells to apply linear interpolation. The values of the interpolated surface at any grid cell at DEM resolution always remains within the minimum and maximum range of surrounding MERRA grid points, resulting in smoother high resolution MERRA data. In the following sections, where we describe the procedures implemented to adjust the selected variables, bilinearly interpolated high resolution MERRA data are subscripted as "MERRA" and physically (e.g., topographical) adjusted climate variables at DEM resolution are subscripted as "DEM".

Temperature

We calculated the temperature at DEM resolution using MERRA elevation and pressure obtained at 850, 500, and 250 hPa pressure levels at each time step with the following equation.

$$T_{DEM} = T_{MERRA} - \Gamma(z_{DEM} - z_{MERRA}) \tag{A1}$$

where T_{DEM} is topographically adjusted temperature at DEM resolution, T_{MERRA} is the interpolated MERRA temperature at DEM resolution, z_{DEM} is DEM elevation, z_{MERRA} is the elevation from MERRA geopotential height interpolated to DEM resolution and Γ is

the lapse rate calculated based on the MERRA surface temperature and the two nearest elevations above the MERRA surface elevation.

Shortwave radiation

First, we evaluate top of the atmosphere solar radiation (SW_{top}) for the 3 h interval based on solar constant (S^*), and the zenith angle (Z) of the sun, which is a function of latitude, date, and time (Dingman, 2002). A single value was assumed for the whole domain based on a central latitude and longitude.

$$SW_{top} = S^* \cos(Z) \tag{A2}$$

We then evaluate attenuation of solar radiation as the ratio of MERRA shortwave radiation (SW_{MERRA}) to the top of the atmosphere solar radiation (SW_{top}), expressed as a transmission factor, TF_{MERRA}.

$$TF_{MERRA} = \frac{SW_{MERRA}}{SW_{top}} \tag{A3}$$

We parameterize the attenuation of solar radiation using Beer's atmospheric transmission law assuming that the optical thickness above a point is based on the atmospheric pressure.

$$SW(P) = SW_{top}e^{-k \cdot P} \tag{A4}$$

where k is the atmospheric attenuation coefficient, P atmospheric pressure and SW shortwave radiation at a height with atmospheric pressure P. The following standard atmospheric pressure versus elevation function is used to relate pressure to elevation:

$$P(z) = P_o \left(\frac{T_o + z\lambda}{T_o} \right)^{-\frac{g}{R\lambda}} \tag{A5}$$

where P_o is standard sea level pressure (101,325 Pa), T_o is standard sea level temperature (288.15 K), g is earth gravitational acceleration (9.81 m s^{-2}), R is the gas constant for dry air (287.04 J kg^{-1} K^{-1}) and λ the lapse rate calculated by MSDH or provided by the user. The atmospheric attenuation coefficient is determined by solving equation (A6) for k at MERRA elevation and using the transmission factor evaluated in equation (A3).

$$k = \frac{-\log(TF_{MERRA})}{P(z_{MERRA})} \tag{A6}$$

Then equation (A4) is used with $P(z_{DEM})$ to obtain downscaled shortwave radiation.

Relative humidity

MERRA specific humidity is used to calculate actual vapor pressure at MERRA elevations that are specified by geo-potential height.

$$e_{MERRA} = \frac{q_{MERRA} * P_{MERRA}}{(0.622 + q_{MERRA})} \tag{A7}$$

where q_{MERRA} is interpolated MERRA specific humidity at DEM resolution, P_{MERRA} is interpolated MERRA pressure at DEM resolution and e_{MERRA} is actual air vapor pressure at DEM resolution. This is then used to evaluate dew point temperature at MERRA elevation ($T_{d-MERRA}$).

$$T_{d-MERRA} = \frac{c \ln\left[\frac{e_{MERRA}}{a}\right]}{b - \ln\left[\frac{e_{MERRA}}{a}\right]} \tag{A8}$$

where for ice/snow, a = 611.21 Pa, b = 22.452 and c = 272.55 °C. Dew point is then adjusted for DEM elevation using a monthly vapor pressure coefficient λ (m^{-1}) provided by Liston and Elder, Table 1, (2006).

$$T_{d-DEM} = T_{d-MERRA} + (z_{DEM} - z_{MERRA})\lambda\frac{c}{b} \tag{A9}$$

where $T_{d-MERRA}$ and T_{d-DEM} are dew point temperature at MERRA elevation and DEM grid elevation, respectively. The following saturation vapor pressure and temperature function:

$$e_s(T) = a\exp\left(\frac{b\,T}{c+T}\right) \tag{A10}$$

is then used to evaluate relative humidity as the ratio of actual and saturated air vapor pressure from dew point and air temperatures at DEM elevation.

$$RH_{DEM} = \frac{e_s(T_{d-DEM})}{e_s(T_{DEM})} \tag{A11}$$

Wind speed

MERRA eastward, E-W (U_{MERRA}), and northward, N-S (V_{MERRA}) wind components are combined using Pythagoras' equation (A12) to obtain the horizontal wind speed magnitude.

$$W_{MERRA} = \sqrt{\left(U_{MERRA}^2 + V_{MERRA}^2\right)} \tag{A12}$$

Wind direction, terrain slope and terrain aspect are calculated using equations (A13)–(A15) (Liston and Sturm, 1998).

$$\theta = \frac{3\pi}{2} - \tan^{-1}\left(\frac{V_{MERRA}}{U_{MERRA}}\right) \tag{A13}$$

$$\beta = \tan^{-1}\sqrt{\left[\left(\frac{\Delta z_x}{\Delta x}\right)^2 + \left(\frac{\Delta z_y}{\Delta y}\right)^2\right]} \tag{A14}$$

$$\gamma = \frac{3\pi}{2} - \tan^{-1}\frac{\left(\frac{\Delta z}{\Delta y}\right)}{\left(\frac{\Delta z}{\Delta x}\right)} \tag{A15}$$

Both slope and aspect are computed using the "Four nearest" method where Δz_x and Δz_y are the elevation differences between the two nearest cells of the target cell in horizontal and vertical directions, respectively.

Equation (A16) parameterizes the effect of the terrain slope and curvature on the MERRA wind speed (W_{MERRA}) (Liston and Elder, 2006).

$$W_{DEM} = W_{MERRA}(1 + \gamma_s\Omega_s + \gamma_c\Omega_c) \tag{A16}$$

where Ω_c (equation (A17)) and Ω_s (equation (A18)) are the curvature and slope in the direction of the wind, respectively.

$$\Omega_c = \frac{1}{4}\left[\frac{z - 0.5(z_w + z_e)}{2\eta} + \frac{z - 0.5(z_s + z_n)}{2\eta}\right.$$
$$\left. + \frac{z - 0.5(z_{sw} + z_{ne})}{2\sqrt{2}\eta} + \frac{z - 0.5(z_{nw} + z_{se})}{2\sqrt{2}\eta}\right] \tag{A17}$$

$$\Omega_s = \beta\cos(\theta - \gamma) \tag{A18}$$

Here $z_e, z_w, z_n, z_s, z_{sw}, z_{ne}, z_{nw}, z_{se}$ are the elevations at eight possible neighboring cells in the east, west, north and south, south-west, north-east, north-west and south-east direction from the target cell and η is the distance between the center of two neighboring cells. Note that the denominator in A17 includes η, not η^2 as would be the case for a conventional Laplacian finite difference approximation of curvature. The use of η only follows Liston and Elder (2006) and we interpret this to be Laplacian curvature scaled by cell size so that it is dimensionless when used in wind speed adjustments (equation (A16)). In equation (A16), γ_c and γ_s are weight factors that adjust wind magnitude based on curvature and slope respectively. Liston and Elder (2006) suggested that the valid range of γ_c and γ_s is between 0 and 1 such that $\gamma_c + \gamma_s = 1.0$. In MSDH, we approximated both of these quantities as 0.5 assuming equal weight for slope and curvature adjustments.

Precipitation

After the reanalysis precipitation is interpolated over the domain distributed at DEM spatial resolution, topographical adjustments are made using equation (1). Default monthly precipitation adjustment factor from Table 1 of Liston and Elder (2006) is encoded (Table A1), though users may also provide their own precipitation adjustment factors based on local data. To correct for bias in precipitation inputs users may also use the tool to compute a bias adjustment coefficient based on nearby station values. B_c is the ratio of the observed data at a precipitation measuring site and the downscaled data at the grid cell in which the site is located.

$$B_c = \frac{PRCP_o}{PRCP_d} \tag{A19}$$

where $PRCP_o$ and $PRCP_d$ are mean annual observed precipitation (mm) and downscaled precipitation (mm), respectively. If multiple sites are located in or near the target spatial domain, B_c is calculated for each site, and an average value is taken. Downscaled data is corrected by multiplying by the bias coefficient (B_c).

Table A1
Monthly adjustment factor for each month from Table 1 of Liston and Elder (2006).

Month	Precipitation adjustment factor, κ_p (km^{-1})
January	0.35
February	0.35
March	0.35
April	0.3
May	0.25
June	0.2
July	0.2
August	0.2
September	0.2
October	0.25
November	0.3
December	0.35

Longwave radiation

We estimated incoming longwave radiation based on downscaled air temperature following the methods of Liston and Elder (2006). First we evaluate the elevation at 700 hPa using linear interpolation of MERRA pressure and elevation information. Then air and dew point temperatures are evaluated at this elevation using (A1) and (A9), and relative humidity is evaluated at this elevation using (A11).

Implied cloud fraction σ_c and then emissivity ε is parameterized by Walcek (1994) using equation (A20) and by Iziomon et al. (2003) using equation (A21), respectively.

$$\sigma_c = 0.832\, exp\left(\frac{RH_{700} - 100}{41.6}\right) \tag{A20}$$

$$\varepsilon = \kappa_\varepsilon\left(1 + Z_s\sigma_c^2\right)\left(1 - X_s\, exp\left(\frac{-Y_s e_{DEM}}{T_{DEM}}\right)\right) \tag{A21}$$

where e_{DEM} is the atmospheric vapor pressure at DEM resolution and κ_ε is 1.08 (Liston and Elder, 2006). X_S, Y_S and Z_S are coefficients that vary depending on elevation. At elevations below 200 m, X_S, Y_S

and Z_S are 0.35, 0.1 K Pa^{-1} and 0.224, respectively. X_S, Y_S and Z_S are 0.51, 0.13 K Pa^{-1} and 1.1, respectively, at elevations above 3000 m. These coefficients vary linearly between these values for elevations from 200 to 3000 m. We then calculate incoming longwave radiation using the Stefan-Boltzmann equation.

$$Q_{li-DEM} = \varepsilon\, \sigma (T_{DEM})^4 \tag{A22}$$

where σ is the Stefan-Boltzmann constant $(5.670{,}373 \times 10^{-8}$ kg s^{-3} K$^{-4})$.

The downscaling parameterizations detailed above have been drawn from the literature or developed in this study based on physical principles for downscaling from the relatively coarse grid scale of MERRA variables at the MERRA geopotential height to the elevation associated with the fine scale grid used by a distributed hydrologic model.

Appendix B. NRCS SNOTEL sites in the study area

Site name	Site number	State	Latitude	Longitude	Elevation (m)
Ben Lomond Trail	333	UT	41.38	−111.92	1777
Sheldon	750	NV	41.9	−119.44	1786
Lamance Creek	569	NV	41.52	−117.63	1829
Lost Creek Resv	1118	UT	41.22	−111.36	1854
Little Grassy	583	UT	37.49	−113.85	1859
Spratt Creek	778	CA	38.67	−119.82	1864
Taylor Canyon	811	NV	41.23	−116.03	1890
Fallen Leaf	473	CA	38.93	−120.05	1901
Tony Grove Rs	1113	UT	41.89	−111.57	1930
Independence Creek	540	CA	39.49	−120.28	1968
Disaster Peak	445	NV	41.97	−118.19	1981
Truckee #2	834	CA	39.3	−120.18	1984
Little Bear	582	UT	41.41	−111.83	1995
Golconda	1195	NV	40.88	−117.59	2010
Ward Creek #3	848	CA	39.14	−120.22	2028
Laurel Draw	573	NV	41.78	−116.03	2041
Big Bend	336	NV	41.76	−115.69	2042
Louis Meadow	972	UT	40.83	−111.76	2042
Gutz Peak	1065	UT	37.5	−113.94	2061
Farmington Lower	1054	UT	40.99	−111.82	2066
Tahoe City Cross	809	CA	39.17	−120.15	2072
Css Lab	428	CA	39.33	−120.37	2089
Buckskin Lower	373	NV	41.75	−117.53	2108
Fawn Creek	476	NV	41.82	−116.1	2134
Independence Camp	539	CA	39.45	−120.29	2135
Dry Fork	906	UT	40.57	−112.17	2162
Seventysix Creek	746	NV	41.74	−115.47	2164
Leavitt Meadows	575	CA	38.3	−119.55	2194
Draw Creek	454	NV	41.66	−115.32	2195
Kilfoil Creek	1145	UT	41.25	−111.41	2201
Hardscrabble	896	UT	40.87	−111.72	2210
Jack Creek Upper	548	NV	41.55	−116.01	2210
Klondike Narrows	1115	UT	41.97	−111.6	2210
Lewis Peak	1006	NV	40.36	−116.86	2256
Vernon Creek	844	UT	39.94	−112.41	2256
Temple Fork	1013	UT	41.79	−111.55	2257
Long Valley Jct	593	UT	37.49	−112.51	2275
Bird Creek	1155	NV	39.46	−114.65	2286
Parley's Summit	684	UT	40.76	−111.63	2286
Smith & Morehouse	763	UT	40.79	−111.12	2316
Summit Lake	1194	NV	41.49	−119	2319
Clear Creek #2	400	UT	39.89	−111.25	2334
Echo Peak	463	CA	38.85	−120.08	2338
Rubicon #2	724	CA	39	−120.13	2344
Lamoille #3	570	NV	40.65	−115.38	2347
Toe Jam	1136	NV	41.32	−116.34	2347
Garden City Summit	1114	UT	41.92	−111.47	2348

(continued)

Site name	Site number	State	Latitude	Longitude	Elevation (m)
Poison Flat	697	CA	38.51	−119.63	2358
Parrish Creek	971	UT	40.93	−111.81	2359
Kalamazoo	1150	NV	39.56	−114.63	2360
Cascade Mountain	1039	UT	40.28	−111.61	2370
Hagan's Meadow	508	CA	38.85	−119.94	2370
Harris Flat	514	UT	37.49	−112.58	2377
Vaccaro Springs	1137	NV	39.45	−115.98	2388
Oak Creek	1146	UT	39.35	−112.19	2393
Marlette Lake	615	NV	39.16	−119.9	2402
Rock Creek	720	UT	40.55	−110.69	2405
Hole-in-mountain	527	NV	40.94	−115.1	2408
Rainbow Canyon	1110	NV	36.25	−115.63	2414
Bug Lake	374	UT	41.68	−111.42	2423
Gooseberry R.s.	495	UT	38.8	−111.68	2423
Beaver Dams	329	UT	39.14	−111.56	2435
Ben Lomond Peak	332	UT	41.38	−111.94	2438
Burts-miller Ranch	1135	UT	40.98	−11085	2438
Currant Creek	432	UT	40.36	−111.09	2438
Farmington	474	UT	40.97	−111.81	2438
Green Mountain	503	NV	40.38	−115.53	2438
Long Flat	592	UT	37.51	−113.4	2438
Forestdale Creek	1049	CA	38.68	−119.96	2444
Squaw Valley G.c.	784	CA	39.19	−120.26	2447
Diamond Peak	443	NV	39.56	−115.84	2448
Daniels-strawberry	435	UT	40.3	−111.26	2450
Blue Lakes	356	CA	38.61	−119.92	2456
Payson R.s.	686	UT	39.93	−111.63	2459
Franklin Basin	484	ID	42.05	−111.6	2464
Dorsey Basin	453	NV	40.89	−115.2	2469
Strawberry Divide	795	UT	40.16	−111.21	2476
Burnside Lake	1051	CA	38.72	−119.89	2478
Timpanogos Divide	820	UT	40.43	−111.62	2481
Chalk Creek #2	393	UT	40.89	−111.07	2487
Horse Ridge	533	UT	41.31	−111.45	2487
Lookout Peak	596	UT	40.84	−111.71	2499
Lightning Ridge	1056	UT	41.36	−111.49	2504
Mining Fork	631	UT	40.49	−112.61	2506
Big Meadow	340	NV	39.46	−119.94	2514
USU Doc Daniel	1098	UT	41.86	−111.51	2521
Beaver Divide	330	UT	40.61	−111.1	2524
East Willow Creek	461	UT	39.31	−109.53	2530
Monitor Pass	633	CA	38.67	−119.61	2533
Gardner Peak	1066	UT	37.4	−113.46	2537
Dry Bread Pond	455	UT	41.41	−111.54	2545
Independence Lake	541	CA	39.43	−120.31	2546
Carson Pass	1067	CA	38.69	−119.99	2546
Gooseberry Upper	1184	UT	38.79	−111.69	2560
Jacks Peak	549	NV	41.53	−116.01	2566
Lakefork #3	1116	UT	40.55	−110.35	2580
Tony Grove Lake	823	UT	41.9	−111.63	2583
Corral Canyon	417	NV	40.28	−115.53	2591
Granite Peak	498	NV	41.67	−117.57	2604
Horse Meadow	1050	CA	38.84	−119.89	2608
Heavenly Valley	518	CA	38.92	−119.92	2616
Lee Canyon	1112	NV	36.31	−115.68	2629
White River #1	864	UT	39.96	−110.99	2634
Big Creek Sum	337	NV	39.29	−117.11	2650
Merchant Valley	621	UT	38.3	−112.44	2653
King's Cabin	559	UT	40.72	−109.54	2659
Mammoth-cottonwood	612	UT	39.68	−111.32	2660
Timberline	1097	UT	39.68	−110.43	2663
Brighton	366	UT	40.6	−111.58	2667
Ebbetts Pass	462	CA	38.55	−119.8	2672
Bear River Rs	992	UT	40.89	−110.83	2675
Pine Creek	694	UT	38.88	−112.25	2679
Fish Lake Utah	1149	UT	38.5	−111.77	2682
Mt Rose Ski Area	652	NV	39.32	−119.89	2683
Sonora Pass	771	CA	38.31	−119.6	2690
Blacks Fork Jct	1162	UT	40.96	−110.58	2704
Agua Canyon	907	UT	37.52	−112.27	2713
Rocky Basin-settleme	723	UT	40.44	−112.24	2713
Clear Creek #1	399	UT	39.87	−111.28	2715
Monte Cristo	634	UT	41.47	−111.5	2731
Mill-d North	628	UT	40.66	−111.64	2733

(continued)

Site name	Site number	State	Latitude	Longitude	Elevation (m)
Camp Jackson	383	UT	37.81	−109.49	2733
Bristlecone Trail	1111	NV	36.32	−115.7	2737
Chalk Creek #1	392	UT	40.85	−111.05	2741
George Creek	1151	UT	41.92	−113.41	2745
Red Pine Ridge	714	UT	39.45	−111.27	2746
Berry Creek	334	NV	39.32	−114.62	2774
Squaw Springs	1156	UT	38.5	−112.01	2775
Kimberly Mine	557	UT	38.48	−112.39	2783
Hickerson Park	522	UT	40.91	−109.96	2787
Hole-in-rock	528	UT	40.92	−110.19	2789
Lily Lake	579	UT	40.86	−110.8	2791
Indian Canyon	543	UT	39.89	−110.75	2797
Dill's Camp	444	UT	39.05	−111.47	2799
Ward Mountain	849	NV	39.13	−114.96	2804
Webster Flat	853	UT	37.58	−112.9	2805
Kolob	561	UT	37.53	−113.05	2806
Hayden Fork	517	UT	40.8	−110.88	2808
Thaynes Canyon	814	UT	40.62	−111.53	2813
Lobdell Lake	587	CA	38.44	−119.37	2814
Summit Meadow	1052	CA	38.4	−119.54	2839
Ef Blacks Fork Gs	1163	UT	40.88	−110.54	2853
Buck Flat	371	UT	39.13	−111.44	2874
Virginia Lakes Ridge	846	CA	38.07	−119.23	2879
Black Flat-u.m. Ck	348	UT	38.68	−111.6	2884
Mosby Mtn.	643	UT	40.61	−109.89	2899
Trout Creek	833	UT	40.74	−109.67	2901
Hewinta	521	UT	40.95	−110.48	2901
Lasal Mountain	572	UT	38.48	−109.27	2914
Castle Valley	390	UT	37.66	−112.74	2920
Pickle Keg	691	UT	39.01	−111.58	2926
Leavitt Lake	574	CA	38.28	−119.61	2931
Snowbird	766	UT	40.56	−111.66	2938
Widtsoe #3	865	UT	37.84	−111.88	2938
Farnsworth Lake	475	UT	38.77	−111.68	2951
Jones Corral	1099	UT	38.07	−112.17	2971
Donkey Reservoir	452	UT	38.21	−111.48	2987
Midway Valley	626	UT	37.57	−112.84	2987
Box Creek	364	UT	38.51	−112.02	2996
Seeley Creek	742	UT	39.31	−111.43	3021
Brian Head	1154	UT	37.68	−112.86	3039
Trial Lake	828	UT	40.68	−110.95	3046
Clayton Springs	983	UT	37.97	−111.83	3063
Wheeler Peak	1147	NV	39.01	−114.31	3085
Steel Creek Park	790	UT	40.91	−110.5	3109
Spirit Lk	1117	UT	40.84	−110.01	3120
Big Flat	339	UT	38.3	−112.36	3154
Lakefork #1	566	UT	40.6	−110.43	3174
Cave Mountain	1152	NV	39.16	−114.61	3226
Chepeta	396	UT	40.77	−110.01	3228
Brown Duck	368	UT	40.58	−110.59	3231
Five Points Lake	481	UT	40.72	−110.47	3335
Lakefork Basin	567	UT	40.74	−110.62	3342

References

Bajracharya, S.R., Palash, W., Shrestha, M.S., Khadgi, V.R., Duo, C., Das, P.J., Dorji, C., 2015. Systematic evaluation of satellite-based rainfall products over the Brahmaputra basin for hydrological applications. Adv. Meteorology 2015 17.

Bajracharya, S.R., Shrestha, M.S., Shrestha, A.B., 2014. Assessment of high-resolution satellite rainfall estimation products in a streamflow model for flood prediction in the Bagmati basin, Nepal. J. Flood Risk Manag. http://dx.doi.org/10.1111/jfr3.12133.

Barnes, S.L., 1964. A technique for maximizing details in numerical weather map analysis. J. Appl. Meteorology 3 (4), 396–409.

Benestad, R., 2004. Empirical-statistical downscaling in climate modeling. Eos, Trans. Am. Geophys. Union 85 (42), 417–422.

Brown, M., Racoviteanu, A., Tarboton, D., Sen Gupta, A., Nigro, J., Policelli, F., Habib, S., Tokay, M., Shrestha, M., Bajracharya, S., Hummel, P., Gray, M., Duda, P., Zaitchik, B., Mahat, V., Artan, G., Tokar, S., 2014. An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas. J. Hydrology 519, 1859–1869.

Carslaw, D.C., Ropkins, K., 2012. openair—An R package for air quality data analysis. Environ. Model. Softw. 27, 52–61.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 28 (15), 2031–2064.

Daly, C., Neilson, R.P., Phillips, D.L., 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorology 33, 140–150.

Daly, C., Taylor, G., Gibson, W., 1997. The PRISM approach to mapping precipitation and temperature. In: Proceedings, 10th Conference on Applied Climatology, American Meteorology Society, pp. 10–12.

Daly, C., Taylor, G., Gibson, W., Parzybok, T., Johnson, G., Pasteris, P., 2000. High-quality spatial climate data sets for the United States and beyond. Trans. ASAE-American Soc. Agric. Eng. 43 (6), 1957–1962.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorological Soc. 137 (656), 553–597.

Dile, Yihun Taddele, Srinivasan, Raghavan, 2014. Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin. J. Am. Water Resour. Assoc. (JAWRA) 50 (5), 1226–1241. http://

dx.doi.org/10.1111/jawr.12182.

Dingman, S.L., 2002. Physical Hydrology, second ed. Prentice Hall.

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., 2011. The Japanese 55-year Reanalysis "JRA-55": an interim report. Sola 7, 149–152.

Fiddes, J., Gruber, S., 2014. TopoSCALE v.1.0: downscaling gridded climate data in complex terrain. Geosci. Model Dev. 7 (1), 387–405.

Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27 (12), 1547–1578.

Gellens, D., Barbieux, K., Schädler, B., Roulin, E., Aschwanden, H., Gellens-Meulenberghs, F., 2000. Snow cover modelling as a tool for climate change assessment in a Swiss alpine catchment. Nord. Hydrol. 31 (2), 73–88.

Hijmans, Robert J., Etten, J.v., Mattiuzzi, M., Summer, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B., Shortridge, A., 2013. Raster: Geographic Data Analysis and Modeling, 2.1-66. http://cran.r-project.org/web/packages/raster/.

Horsburgh, J.S., Reeder, S.L., 2014. Data visualization and analysis within a Hydrologic Information System: integrating with the R statistical computing environment. Environ. Model. Softw. 52, 51–61.

Hungerford, R.D., Nemani, R.R., Running, S.W., 1989. MTCLIM: A Mountain Microclimate Simulation Model, Forest Service Research Paper INT-414. U.S. Dept. of Agriculture, Washington, D. C.

Hunter, R.D., Meentemeyer, R.K., 2005. Climatologically aided mapping of daily precipitation and temperature. J. Appl. Meteorology 44 (10).

Iziomon, M.G., Mayer, H., Matzarakis, A., 2003. Downward atmospheric longwave irradiance under clear and cloudy skies: measurement and parameterization. J. Atmos. Solar-Terrestrial Phys. 65 (10), 1107–1116.

Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Model. Softw. 16 (4), 309–330.

Kalra, A., Ahmad, S., 2012. Estimating annual precipitation for the Colorado River Basin using oceanic-atmospheric oscillations. Water Resour. Res. 48 (6).

Kanamitsu, M., Kumar, A., Juang, H.-M.H., Schemm, J.-K., Wang, W., Yang, F., Hong, S.-Y., Peng, P., Chen, W., Moorthi, S., 2002. NCEP dynamical seasonal forecast system 2000. Bull. Am. Meteorological Soc. 83 (7), 1019–1037.

Keitt, T.H., Bivand, R., Pebesma, E., Rowlingson, B., 2011. Rgdal: Bindings for the Geospatial Data Abstraction Library. R Package Version 0.7-1. http://CRAN.r-project.org/package=rgdal.

Kucera, P.A., Ebert, E.E., Turk, F.J., Levizzani, V., Kirschbaum, D., Tapiador, F.J., Loew, A., Borsche, M., 2013. Precipitation from space: advancing Earth system science. Bull. Am. Meteorological Soc. 94 (3), 365–375.

Liston, G.E., Elder, K., 2006. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). J. Hydrometeorol. 7 (2), 217–234.

Liston, G.E., Sturm, M., 1998. A snow-Transport model for complex terrain. J. Glaciol. 44 (148), 498–516.

Lucchesi, R., 2012. File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3), 82 pp.

Luce, C.H., Tarboton, D.G., 2010. Evaluation of alternative formulae for calculation of surface temperature in snowmelt models using frequency analysis of temperature observations. Hydrol. Earth Syst. Sci. 14 (3), 535–543.

Mahat, V., Tarboton, D.G., 2012. Canopy radiation transmission for an energy balance snowmelt model. Water Resour. Res. 48, W01534.

Mahat, V., Tarboton, D.G., Molotch, N.P., 2013. Testing above- and below-canopy representations of turbulent fluxes in an energy balance snowmelt model. Water Resour. Res. 49 (2), 1107–1122. http://dx.doi.org/10.1002/wrcr.20073.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Natural Resource Conservation Service NRCS, 2014. USU Doc Daniel SNOTEL Site. http://www.wcc.nrcs.usda.gov/nwcc/site?sitenum=1098&state=ut. Last accessed on Oct 2, 2014.

Pierce, D., 2011. Ncdf: Interface to Unidata NetCDF Data Files. http://CRAN. R-project. org/package= ncdf, R package version 1(6).

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., 2011. R Development Core Team. 2010. Nlme: Linear and Nonlinear Mixed Effects Models. R Package Version 3.1–97. R Foundation for Statistical Computing, Vienna.

Reichle, R.H., Koster, R.D., De Lannoy, G.J., Forman, B.A., Liu, Q., Mahanama, S.P., Touré, A., 2011. Assessment and enhancement of MERRA land surface hydrology estimates. J. Clim. 24 (24), 6322–6338.

Rew, R., Hartnett, E., Caron, J., 2006. NetCDF-4: software implementing an enhanced data model for the geosciences. In: 22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanograph, and Hydrology.

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. J. Clim. 24 (14), 3624–3648.

Schulzweida, U., Kornblueh, L., Quast, R., 2006. CDO User's Guide. Climate Data Operators, Version 1(6).

Sen Gupta, A., 2014. Improving the Physical Processes and Model Integration Functionality of an Energy Balance Model for Snow and Glacier Melt. Civil and Environmental Engineering. Utah State University, Logan, UT, p. 196.

Sen Gupta, A., Tarboton, D., Hummel, P., Brown, M., Habib, S., 2015. Integration of an energy balance snowmelt model into an open source modeling framework. Environ. Model. Softw. 68, 205–218.

Sen Gupta, A., Tarboton, D.G., 2013. Using the Utah energy balance snow melt model to quantify snow and glacier melt in the Himalayan region. In: 81st Annual Western Snow Conference: Jackson Hole, Wyoming. http://www.westernsnowconference.org/sites/westernsnowconference.org/PDFs/2013SenGupta.pdf.

Shrestha, M., Artan, G., Bajracharya, S., Sharma, R., 2008. Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin. J. Flood Risk Manag. 1 (2), 89–99.

Shrestha, M., Rajbhandari, R., Bajracharya, S., 2013. Validation of NOAA CPC_RFE Satellite-based Rainfall Estimates in the Central Himalayas. ICIMOD Working Paper 2013/5: ICIMOD, Kathmandu. http://lib.icimod.org/record/28331.

Skelly, W.C., Henderson-Sellers, A., 1996. Grid box or grid point: what type of data do GCMs deliver to climate impacts researchers? Int. J. Climatol. 16 (10), 1079–1086.

Suarez, M.J., Rienecker, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H., Gu, W., Sienkiewicz, M., Koster, R., Gelaro, R., 2008. The GEOS-5 Data Assimilation System-documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0.

Tarboton, D.G., Chowdhury, T.G., Jackson, T.H., 1995. A spatially distributed energy balance snowmelt model. In: Tonnessen, K.A., Williams, M.W., Tranter, M. (Eds.), Biogeochemistry of Seasonally Snow-covered Catchments (Proceedings of a Boulder Symposium, July 1995). IAHS Publ. no. 228, Wallingford, pp. 141–155.

Tarboton, D.G., Luce, C.H., 1996. Utah Energy Balance Snow Accumulation and Melt Model (UEB). Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, Utah, p. 64.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., Cook, R., 2012. DAYMET: Daily Surface Weather on a 1 Km Grid for North America. 1980–2008. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge. T, N. doi 10.

Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. J. Hydrology 190 (3), 214–251.

Walcek, C.J., 1994. Cloud cover and its relationship to relative humidity during a springtime midlatitude cyclone. Mon. Weather Rev. 122 (6), 1021–1035.

Wilby, R.L., Dawson, C.W., Barrow, E.M., 2002. SDSM—a decision support tool for the assessment of regional climate change impacts. Environ. Model. Softw. 17 (2), 145–157.

Wu, W.-S., Purser, R.J., Parrish, D.F., 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Weather Rev. 130 (12), 2905–2916.

Xie, A., Ren, J., Qin, X., Kang, S., 2007. Reliability of NCEP/NCAR reanalysis data in the Himalayas/Tibetan plateau. J. Geogr. Sci. 17 (4), 421–430.

Xie, P., Arkin, P.A., 1996. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. J. Clim. 9 (4), 840–858.

Xie, P., Yarosh, Y., Love, T., Janowiak, J., Arkin, P.A., 2002. A real-time daily precipitation analysis over South Asia. In: 16th Conference on Hydrology, American Meteorological Society: Orlando, FL.

Xu, C.-y., 1999. From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Prog. Phys. Geogr. 23 (2), 229–249.

You, J., 2004. Snow Hydrology: the Parameterization of Subgrid Processes within a Physically Based Snow Energy and Mass Balance Model. Civil and Environmental Engineering. Utah State University, Logan.

Zender, C.S., 2008. Analysis of self-describing gridded geoscience data with netCDF Operators (NCO). Environ. Model. Softw. 23 (10), 1338–1342.

Zimmermann, N.E., Roberts, D.W., 2001. Final Report of the MLP Climate and Biophysical Mapping Project. Swiss Federal Research Institute WSL/Utah State University, Birmensdorf, Switzerland/Logan, USA.