The Impact of Precipitation Type Discrimination on Hydrologic Simulation: Rain–Snow Partitioning Derived from HMT-West Radar-Detected Brightband Height versus Surface Temperature Data

NAOKI MIZUKAMI,* VICTOR KOREN, AND MICHAEL SMITH
Office of Hydrologic Development, NOAA/National Weather Service, Silver Spring, Maryland

DAVID KINGSMILL
University of Colorado Boulder, CIRES, and NOAA/Earth System Research Laboratory, Boulder, Colorado

ZIYA ZHANG, BRIAN COSGROVE, AND ZHENGTAO CUI
Office of Hydrologic Development, NOAA/National Weather Service, Silver Spring, Maryland

(Manuscript received 15 March 2012, in final form 17 January 2013)

ABSTRACT

Hourly surface precipitation type (Ptype) grids (a total of 408 h from 1 December 2005 through April 20, 2006) were generated by mapping the elevation of the radar-detected brightband height (BBH) to terrain elevation during the 2005/06 observation period of the western Hydrometeorology Testbed (HMT-West) in the North Fork American River basin. BBH Ptype grids were compared to those derived by the standard National Weather Service (NWS) temperature threshold method. In this method, a fixed threshold temperature separating rain and snow was applied to hourly 4-km gridded temperature data. The BBH Ptype grids agreed well (>90%) with the temperature threshold–based grids below an elevation of 1524 m. The agreement dropped to below 60% above this elevation, and BBH Ptype produced more rainfall than the temperature-based Ptype. Continuous hourly streamflow simulations were generated using spatially lumped and distributed hydrologic models with and without the BBH Ptype data from 1 October 2005 through 30 September 2006. Simple insertion of BBH Ptype data did not always improve streamflow simulations for the 11 events examined relative to corresponding simulations using temperature threshold–derived precipitation type, possibly because of the use of the models calibrated with the temperature-based Ptype. The simple sensitivity test indicated simulations of both peak flows from midwinter storms and spring snowmelt runoff are affected by errors in precipitation type estimates.

1. Introduction

In the California Sierra Nevada, approximately 70% of annual precipitation occurs in winter (Serreze et al. 1999). A well-known atmospheric condition referred to as an atmospheric river (i.e., a narrow band of concentrated water vapor transport in the atmosphere) in the Pacific causes extreme winter precipitation events along the West Coast and the windward side of the Sierra Nevada (B. L. Smith et al. 2010). A recent study showed up to 40% of seasonal snow accumulation in the Sierra Nevada originates from intense Pacific storms associated with atmospheric rivers (Guan et al. 2010). Such storms can also cause winter flooding if precipitation falls as rain. Because of a wide range of terrain elevation in the Sierra Nevada, knowledge of the elevation where precipitation transforms from snow to rain (hereafter referred to as RSELV) is crucial in determining the spatial extent of each type of precipitation within a basin during a given storm. Errors in estimating the RSELV for a given storm potentially have a large impact on the simulation of streamflow over a mountainous basin along the West Coast (White et al. 2002; Maurer and Mass 2006; Lundquist et al. 2008). For example, White et al. (2002) estimated

* Current affiliation: Research Application Laboratory, National Center for Atmospheric Research, Boulder, Colorado.

Corresponding author address: Naoki Mizukami, Research Application Laboratory, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.
E-mail: mizukami@ucar.edu

DOI: 10.1175/JHM-D-12-035.1

© 2013 American Meteorological Society
that a 600-m increase in the RSELV led to a threefold increase in streamflow volume given the topographic relief in the Sierra Nevada, indicating that errors in RSELV estimates can be a predominant source of uncertainty in streamflow prediction. RSELV errors also impact snow cover area estimates throughout the season. Shamir and Georgakakos (2006) detected errors in snow-covered areas derived from a distributed snow model compared to satellite images in lower elevations of the American River basin of the Sierra Nevada. These errors could result in runoff volume estimate errors during the spring snowmelt.

RSELV can be estimated based on freezing-level observations or near-surface air temperature measurements at climate stations and environmental temperature lapse rates with the use of a threshold air temperature (PXTEMP) that separates rain and snow. While traditionally used in hydrologic modeling, surface air temperatures are not considered to be the best indicator of precipitation type on the ground (Maurer and Mass 2006) and ignore atmospheric processes aloft as described by Minder et al. (2011). Snowfall can occur in a wide range of air temperatures depending on other atmospheric factors including atmospheric pressure (Dai 2008) and density and size of snow particles (Lundquist et al. 2008). Relationships between air temperature and the probability of snowfall (i.e., snow frequency curves) have been empirically developed for many regions of the world (Auer 1974; Rohrer 1989; L’hôte et al. 2005; Dai 2008; Kienzle 2008). For example, Yang et al. (1997) identified that the 2°C temperature corresponds to the 50% probability of snow from the curve developed by Auer (1974). Operational forecasters in the National Oceanic and Atmospheric Administration’s (NOAA) National Weather Service (NWS) River Forecast Centers (RFCs) employ a numerical weather prediction model to generate estimates of the RSELV (Lundquist et al. 2008; White et al. 2010). The RSELV is determined by “offsetting” the freezing level by prescribed heights. The typical offset recommended by CNRFC is 305 m below the RUC model freezing level. These vertical offsets typically correspond to a temperature difference of 1°–2°C.

In radar meteorology, the bright band (BB) is a layer of enhanced radar reflectivity where the beam encounters ice particles coated with liquid water. Radar-detected BB information has been found useful for characterizing the transition from snow to rain in the atmosphere (Martner et al. 1993; Ryzhkov and Zrnic 1998). For a hydrologic model application, Maurer and Mass (2006) applied radar-detected BB to distributed hydrologic modeling to evaluate the value of BB data for hydrologic simulations. They estimated the top of the BB, or the freezing level, and assumed a BB thickness of 300 m. Data from a NOAA Earth System Research Laboratory (ESRL) S-band precipitation profiling (S-PROF) radar were obtained for the period of 26 November through 22 December 2001, during which five precipitation–runoff events occurred in a 1440 km² basin in Oregon. Using a simple mapping of BB elevation to terrain, precipitation falling above the BB was automatically declared to be snow, while precipitation falling below the BB was typed as rain. Within the 300-m thick BB, the hydrologic model linearly interpolated a mixture of rain and snow for the 150-m grid cells of the Distributed Hydrology–Soil–Vegetation Model (DHSVM; Wigmosta et al. 1994). The authors found that the use of the BB data improved 3-h simulations of streamflow and snow water equivalent (SWE) compared to a base case in which temperature data were used to partition rain and snow.

More recently, White et al. (2002) developed a bright-band height (BBH) detection algorithm using Doppler wind profiling radars. Hereafter, this approach is referred to the “White algorithm.” As shown in Fig. 1, the BBH is the elevation of the maximum reflectivity and is called the radar-derived snow level (White et al. 2010). It should be noted that BBH is lower than the freezing level. Precipitation particles gain fall speed [Doppler vertical velocity (DVV)] as the phase changes to liquid within the BB. The White algorithm first analyzes vertical gradients of both DVV and reflectivity to determine the existence of a BB. Upon a detection of the BB, the algorithm then searches upward for the maximum radar reflectivity and declares this elevation to be the BBH, that is, the RSELV. The depth of the melting layer can induce uncertainty into the RSELV estimation from the White algorithm. The melting layer thickness can be over 500 m (White et al. 2002) depending on the atmospheric conditions such as humidity, wind, density, and size of snow particles (Lundquist et al. 2008). Nevertheless, BBH is considered the best remotely sensed proxy for the rain–snow line (Minder et al. 2011).

During the western Hydrometeorology Testbed (HMT-West) experiment operated by NOAA (Ralph et al.
2005), observations of hourly BBH were collected from a S-PROF radar deployed at Alta (39.20°N, 120.82°W, elevation 1085 m) within the North Fork American River basin (NFARB; see Fig. 2 for the location). Using data from the 2005/06 winter, White et al. (2010) compared freezing-level forecasts used at CNRFC with radar-derived observations of the freezing level. They identified the presence of biases in the forecast freezing levels. In addition, they found that the typical CNRFC rain–snow line offset of 305 m below the freezing level is 68–75 m lower than the snow level detected by the White algorithm. On the average, the radar-derived snow level was found to exist at the temperature at which falling precipitation has a 50% probability of reaching the ground as snow.

Profiling radars are being increasingly deployed for operational detection of RSELV. Three such radars were implemented to monitor snow level and other variables in response to the 2009 near failure of the Howard Hanson Dam in the state of Washington (White et al. 2012). The NOAA ESRL is developing low-cost, portable, S-band snow-level radars to be deployed in 10 key watersheds at the direction of the California Department of Water Resources (White et al. 2010; Ralph and Dettinger 2011).

The above discussion highlights the importance of accurate descriptions of RSELV and describes some of the advances and issues in detecting RSELV using radar. However, to the best of our knowledge, only one study (Maurer and Mass 2006) investigated the use of radar-derived RSELV data for hydrologic modeling. Thus, there is a growing gap between the ability to detect and analyze radar-derived RSELV and the actual use of the data for hydrologic modeling. To this end, the main goal of our paper is to assess the impacts of using the BBH data derived by the White algorithm on retrospective streamflow simulations. We pose the following question: can the use of radar-derived RSELV data lead to improved retrospective streamflow simulations compared to rain–snow partitioning based on surface temperature? The question is framed by 1) using a spatially uniform mapping of BBH-derived RSELV to terrain elevation to partition rain and snow and 2) direct replacement (hour by hour) of precipitation type grids determined by surface temperature with radar-derived RSELV grids.

This study is the first to use the BBH data derived from the White algorithm in multimonth continuous hourly streamflow simulations. Moreover, our study is the second to use actual radar-derived RSELV data for hydrologic simulation studies (the other being Maurer and Mass 2006). Our simulations span the period from 1 October 2005 to 30 September 2006, far longer than the 1-month period analyzed by Maurer and Mass (2006). In addition, we use two hydrologic models (spatially lumped and spatially distributed), both of which performed very well in the NFARB experiments of phase 2 of the Distributed Model Intercomparison Project (DMIP 2; Smith et al. 2012b, manuscript submitted to J. Hydrol.).

The analysis starts with a comparison of fractional snowfall areas and amounts computed from PXTEMP and BBH-based methods. Then the framework of DMIP 2 (Smith et al. 2012b, manuscript submitted to J. Hydrol.) was used to generate hourly retrospective streamflow simulations over the entire observation period of the 2005/06 HMT-West year. The hydrologic modeling intended to employ BBH-derived RSELV as input data in addition to precipitation and temperature as opposed to model parameter values.

The remainder of this paper is organized as follows. In section 2, the study basin and data are described. In section 3, Precipitation type (snow or rain; hereafter called Ptype) partitioning methods used in the NWS hydrologic models are described. Section 4 presents the

![Fig. 1. Schematic of mapping radar-detected BBH to topography and surface precipitation type.](image-url)
results of the analysis performed in this study: 1) characterization of BBH Ptype data by comparing with temperature threshold–derived Ptype data and 2) hydrologic modeling analyses using the BBH Ptype data. Finally, the paper provides a summary in section 5, followed by recommendations for future work in section 6.

2. Study basin and data

a. North Fork American River basin

The NFARB is a narrow, westward-facing, steep-sided basin, as shown in Fig. 2. The outlet of this basin is the U.S. Geological Survey (USGS) gauge at Lake Clementine, formed by the North Fork Dam. The drainage area at this point is 886 km². Precipitation amount is enhanced by orographic effects, with mean annual precipitation varying from 813 mm at Auburn (elevation 393 m) to 1651 mm at Blue Canyon (elevation 1676 m). The mean annual precipitation is 1530 mm, and the annual runoff is 850 mm (Lettenmaier and Gan 1990). Because of its large elevation range (260–2600 m), precipitation type can vary greatly within the basin. Streamflow is about two-thirds wintertime rainfall and snowmelt runoff and less than one-third springtime snowmelt runoff (Dettinger et al. 2004). The basin is
highly forested and varies from pine–oak woodlands to shrub rangeland to ponderosa pine and, finally, to subalpine forest as one moves up in elevation. Much of the forests are secondary growth because of the extensive timber harvesting to support the mining industry in the late 1800s (Jeton et al. 1996). Soils in the basin are predominately clay loams and coarse sandy loams. The geology of the basin includes metasedimentary rocks and granodiorite (Jeton et al. 1996).

b. Data

Our study leveraged datasets provided by the DMIP 2 project in the NFARB (M. B. Smith et al. 2010b; http://www.nws.noaa.gov/oh/hrl/dmip/2/data_link.html). Hourly gridded precipitation and temperature were developed for DMIP 2 over the period from October 1987 to September 2006, as described in the following sections. The grid used the Hydrologic Rainfall Analysis Projection (HRAP) described by Reed and Maidment (1999). The HRAP resolution is approximately 4 km in the horizontal length. Hourly streamflow observations from the USGS were also available at the outlet of the NFARB for the years 1987–2006.

1) GRIDDED QPE

The 4-km gridded quantitative precipitation estimates (QPE) from the DMIP 2 (M. B. Smith et al. 2010b; Smith et al. 2012b, manuscript submitted to J. Hydrol.) were generated solely from an interpolation of gauge observations. The QPE dataset was intended to be a gridded representation of the historical mean areal precipitation (MAP) time series used by the CNRFC for model calibration.

Two sources of precipitation observations were used. First, daily and hourly precipitation gauge observations from the NOAA Cooperative Observer Program (COOP) network were obtained through the National Climatic Data Center (NCDC). Second, we used observations from snowpack telemetry (SNOTEL) stations operated by the Natural Resources Conservation Service (NRCS; Serreze et al. 1999). Locations of both observation networks are shown in Fig. 2.

Stations having more than 10 years of reasonably complete data were selected for initial analysis. Forty-one stations in and near the NFARB were selected. Considerable effort was spent in the quality control (QC) of these gauge data. QC procedures included double mass analysis to locate and correct manmade inconsistencies, correcting observation times for daily stations, and correcting erroneous flags for missing and accumulated precipitation (M. B. Smith et al. 2003, 2010b, 2012b, manuscript submitted to J. Hydrol.). Interested readers are also referred to the DMIP 2 web documentation for specific examples of the QC procedures (http://www.nws.noaa.gov/oh/hrl/dmip/2/wb_precip.html).

Two major steps were involved in the QPE derivation process. First, serially complete hourly time series of precipitation spanning 1987–2006 were derived at all gauge locations. Missing data were estimated by an inverse distance weighted average of the observations from the closest station in each of four quadrants. Daily precipitation totals were time distributed by using a weighted average of the surrounding hourly stations when data corresponded to the observation time of the daily station.

The multisensor precipitation estimation algorithm (MPE; Seo 1998) was used to spatially interpolate the hourly point precipitation data to the 4-km grid. MPE uses climatological monthly mean precipitation from the Precipitation-Elevation Regression on Independent Slopes Model (PRISM; Daly et al. 1994) to adjust the interpolation of point precipitation to grids. We selected the 800-m resolution PRISM monthly climatological precipitation data derived for 1971–2000 (Daly et al. 2008) to be consistent with the river forecasting operations at CNRFC (Hartman 2010). The final DMIP 2 QPE data did not include any corrections for gauge undercatch. With the final QPE data, a Budyko curve (Budyko 1974) water balance analysis was performed using basins from a wide range of climatic regimes (M. B. Smith et al. 2010a). This analysis confirmed that the QPE data were reasonable for hydrologic modeling tests within DMIP 2.

2) GRIDDED TEMPERATURE DATA

Hourly temperature grids from the DMIP 2 were also used for our study. The data sources for the derivation of hourly gridded temperature fields are quality controlled observations at NCDC COOP and NRCS SNOTEL stations around the DMIP 2 basins. Only daily observations of maximum and minimum temperatures were used to derive temperature grids because spatial interpolation of the station data, as described below, utilized PRISM temperature grids, which supply monthly mean of daily temperature values. Temporal disaggregation was performed using the derived daily maximum and minimum temperature grids for each grid box and each day. This hourly analysis is based on an algorithm developed by Parton and Logan (1981), which accounts for dependency of diurnal temperature pattern on latitude and time of year. For each day this pattern is used to interpolate between sequential maximum and minimum values and minimum and maximum values to obtain hourly values while preserving the daily maximum and minimum values.
The daily maximum and minimum temperature grids for each day were generated based on an inverse distance approach. The interpolation procedure for daily maximum and minimum temperature analysis attempts to interpolate a measure of local temperature anomalies because these may be spatially less dependent on the complex terrain than observed local temperatures. This procedure uses PRISM gridded temperature data to define climatological mean values of daily maximum and minimum temperature at both gauge and grid point locations and then compute this anomaly measure. PRISM values are estimates of the climatological monthly mean maximum or minimum temperature for the period 1961–1990 at 2.5 arc second (~4 km) spatial resolution. This assures that the anomalies at both grid points and gauges are defined in a consistent way. It also eliminates the need for gauge normals to be defined for a consistent climatological period. As a result, gauges having relatively short historical records can be used in analysis, giving better spatial definition to daily spatial temperature variability.

For each gauge, the algorithm computed the anomaly defined herein as the difference between the gauge value for the given day and the corresponding monthly PRISM climatological mean value. The differences at the gauges were interpolated to HRAP grid points using an inverse distance interpolation with the distance exponent equal to one. Difference values for the nearest two gauges in each of four quadrants surrounding each HRAP grid point were used. For each HRAP grid point, the PRISM monthly average maximum or minimum temperature value obtained from the PRISM grid box containing each HRAP grid point was added to the analyzed difference value to obtain the HRAP grid point value of the analyzed maximum or minimum daily air temperature.

As a final check of the QPE and temperature grids for DMIP 2, these data were used as input for lumped and distributed hydrologic modeling tests. These tests revealed that the data were adequate to support hourly hydrologic model simulations called for in DMIP 2. At each time step and each grid point, Ptype was assigned to snow if terrain elevation is greater than BBH and to rain otherwise. Figure 2b presents the 4-km gridded elevation data used for NFARB for DMIP 2, while Fig. 2c shows an example of BBH Ptype data for 1 h on 18 December 2005.

It is well known that the BBH in the free atmosphere drops by several hundreds of meters as storms approach a mountain range from an upwind direction (Medina et al. 2005; Kingsmill et al. 2008; Lundquist et al. 2008; Thériault and Stewart 2008). Such BBH behavior is due to diabatic cooling of uprising air over the mountain slope, increasing frozen hydrometeor size and thus increasing the time required for snow particles to melt and the cooling of latent heat absorbed by snow particle melt (Minder et al. 2011). Minder and Kingsmill (2013) quantified the lowering of the BBH along a transect beginning in the Central Valley and proceeding up into the NFARB (see Fig. 2) based on the measurements at Sloughhouse (elevation 50 m MSL) and Alta during the seven largest storms in each of the winters of 2005/06, 2006/07, and 2007/08. Their study determined an average lowering of the BBH of 170 m from Sloughhouse to Alta. It is assumed that the elevation at which the BBH intersects with the terrain can be lower than the BBH measured at the Alta site. For practical purposes in our study, BBH measured at Alta was assigned to all the grid boxes. To account for the orographic effect on BBH spatial variability (i.e., BBH lower near the mountain slope), however, we created additional datasets by lowering the BBH measured at Alta by several distances (100 and 200 m offsets) to investigate the sensitivity of hydrologic simulations to precipitation typing.

The radar at Alta was not always able to detect the BBH even during precipitation events for a number of reasons. A total of 408 h of BBH observations were collected from 12 storm events, which produced precipitation for 1365 h in total. White et al. (2010) reported that approximately one-fourth of the precipitation that fell at the Alta S-PROF radar site was in the form of snow or a mixture of rain and snow during the 2005/06 study period. This indicates that for the hour when snowfall occurred at the Alta site, the BBH was either too low to be detected by the radar, or was nonexistent during the event. Consequently, time series of the BBH data were intermittent during the study period. For example, Fig. 3 shows hourly MAP for the lower- and upper-elevation zones of NFARB, which are delineated by the 1524-m contour (see section 3.2.1), during a span of 30 h starting at 0000 UTC on 1 December 2005. Only 14 h (total precipitation of 53.4 mm and 27.0 mm in the upper and
lower zone, respectively) of the 30 (total precipitation of 175 mm and 123.5 mm in the upper and lower zone, respectively) had an estimate of BBH and subsequently BBH Ptype data, and with the exception of the first 7 h, the observations were intermittent in time. On the other hand, Fig. 4 shows the availability of BBH Ptype data along with the MAP for both elevation zones during the large precipitation–runoff event of 28 December 2005 to 2 January 2006. In this event, the BBH Ptype data were more consistently available. For hydrologic modeling application, the other methods are required to determine the form of precipitation for the hours of missing BBH Ptype data, which will be discussed more in the next section.

Fig. 3. Hourly time series of MAP for NFARB (top) upper- and (bottom) lower-elevation zones from 1 to 4 Dec 2005. Hours when BBH Ptype data are available are shaded in gray. Upper and lower zones were defined as areas above and below the 1524-m contour line in the NFARB, respectively.

Fig. 4. As in Fig. 3, but for the storm event from 30 Dec 2005 through 2 Jan 2006.
3. Hydrologic models

Widely used and well-tested operational hydrologic models were used in our analyses. These include the SNOW-17 (Anderson 1973) and the Sacramento Soil Moisture Accounting models (SAC-SMA; Burnash et al. 1973). Both models are used by hydrologists at NWS RFCs in spatially lumped mode to generate operational river simulations and forecasts. Both SAC-SMA and SNOW-17 are also used within the distributed model being deployed by the NWS for operational forecasting (Koren et al. 2004; Smith et al. 2012a). The details of both models can be found in the above references. The following subsections focus on how Ptype is determined in spatially lumped and distributed SNOW-17 as well as how the new datasets were used in the hydrologic models.

a. Lumped model

The spatially lumped model typically treats the entire basin as a single entity with basin average meteorological input [MAP and mean areal temperature (MAT)] and model parameters. In mountainous areas, basins are typically disaggregated into two or more elevation zones to account for climatic variation with elevation. The NFARB is separated into upper and lower zones at the 1524-m contour line to capture the predominant form of precipitation. Time series of MAP and MAT were generated by spatially averaging the gridded precipitation and temperature data over each elevation zone for each hour.

The lumped application of SNOW-17 uses three methods to specify the form of precipitation at each time step in the following order of precedence. The first method is to explicitly specify the fraction of the precipitation that is in the form of snow for each precipitation data time interval. The second is to specify the rain–snow elevation (i.e., RSELV time series) in conjunction with an areal-elevation curve to determine which portion of the basin or elevation zone receives snow. For this option, RSELV can be estimated based on freezing-level data from numerical weather prediction models, radiosonde measurements, or BBH measurements. RSELV can also be estimated using air temperature. In either case, RSELV is computed using Eq. (1):

\[
RSELV_t = elev + (T_t - PXTEMP) \frac{100}{Lp},
\]

where \(T_t\) is the temperature in degrees Celsius at elev (meters), \(PXTEMP\) is a threshold temperature in degrees Celsius separating rain or snow as described before, \(Lp\) is the lapse rate during the precipitation period in degrees Celsius per 100 m, and \(RSELV_t\) is the elevation separating rain from snow in meters; \(T_t\) and elev are specified as pairs in Table 1 depending on the availability of input data. Finally, SNOW-17 uses \(PXTEMP\) to type precipitation as snow when temperatures are below \(PXTEMP\); otherwise, it is rain. In this last option, precipitation type is set to either snow or rain over the basin or elevation zone.

During a simulation, SNOW-17 can use a mix of these typing methods depending on data availability. Each time step, SNOW-17 checks first for a time series of fractional areas of snowfall (psfrac, percent) over a basin. If these data are not available for a time step, SNOW-17 uses a RSELV time series (if present) in conjunction with an areal-elevation curve to compute the areal percentage of precipitation that fell as snow versus rain. Finally, SNOW-17 will default to use the \(PXTEMP\) value in conjunction with the air temperature time series.

In our analysis, the psfrac values for the upper- and lower-elevation zones of the NFARB basin were computed using the hourly 4-km BBH Ptype grids. The psfrac value is the percent of the elevation zone in which the falling precipitation is typed as snow. There were 17 grid boxes in the upper zone and 31 grid boxes in the lower zone based on the 4-km DEM used in the DMIP 2 experiments (see Fig. 2c). Hours for which there were no BBH Ptype data were set to missing. For these hours of missing BBH Ptype data, SNOW-17 defaulted to use the RSELV time series generated using Eq. (1) (fixed \(Lp = 6.5^\circ C\)) and MAT data. The third option of using \(PXTEMP\) solely with the MAT time series was not used in this study.

b. Distributed model

The SNOW-17 and SAC-SMA models can be run in gridded mode in the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM; Koren et al. 2004). Figure 2b shows the 4-km grid used for distributed modeling in this study.

HL-RDHM follows a different order of precedence compared to the lumped model for determining the form of precipitation in each grid box. HL-RDHM first looks for the existence of hourly Ptype grids, which
come from hourly BBH Ptype data in this study. If available for a particular time step, these data are used to explicitly type the precipitation as rain or snow in each HRAP grid box. In those time steps in which BBH Ptype data were not available, the HL-RDHM defaults to use a PXTEMP value to determine the form of precipitation in each grid box. In the latter case, the DMIP 2 hourly temperature grids were used. Table 2 summarizes the priority and types of data used to classify precipitation as rain or snow in the lumped and distributed model tests.

4. Results and discussion

a. Comparison of BBH and PXTEMP Ptype data

We examined the agreement between the BBH Ptype grids and the precipitation type determined by using PXTEMP and the DMIP 2 gridded hourly air temperature. A PXTEMP value of 2°C was chosen for the comparison. The PXTEMP value at each time step was used with the hourly 4-km temperature grids to determine which grid boxes would be typed as rain or snow. Hereafter, we refer to this latter surface precipitation type as PXTEMP Ptype.

Figure 5 shows the fraction of the 408 total hours of BBH observations in which both methods indicated the same Ptype for each grid box. As expected, the most consistent agreement occurs in the lower elevations of the basin where both methods indicate rainfall. The agreement drops to approximately 50%–60% for the upper elevations of the basin indicating increasing uncertainty of Ptype estimates with elevation. We further looked into what kinds of difference in precipitation type detection occurred.

Two types of discrepancies were defined as follows. A Type 1 discrepancy is when the BBH method indicated rain while the PXTEMP Ptype indicated snow. A Type 2 discrepancy is when the BBH method indicated snow while the PXTEMP Ptype indicated rain. The percentages of Type 1 and Type 2 discrepancies for each grid box are shown in Figs. 6a and 6b, respectively. For the entire Ptype grid domain, Type 1 discrepancies are more prominent, indicating that the BBH-based method tends to produce more rainfall than 2°C PXTEMP Ptype.

Table 2. Precedence of methods used to type precipitation as rain or snow.

<table>
<thead>
<tr>
<th>Modeling approach</th>
<th>1st priority</th>
<th>2nd priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumped</td>
<td>Hourly psfrac time series derived from BBH Ptype grids</td>
<td>Hourly RSELV time series derived from MAT and 2°C PXTEMP [Eq. (1)]</td>
</tr>
<tr>
<td>Distributed</td>
<td>Hourly BBH Ptype grids</td>
<td>Hourly temperature grid with 2°C PXTEMP</td>
</tr>
</tbody>
</table>

However, examination of the NFARB domain reveals that BBH method produced more snow in higher parts but more rain in lower parts of the NFARB upper zone compared to the 2°C PXTEMP method. Considering the potentially increasing uncertainty of BBH with distance from the Alta radar site (Maurer and Mass 2006; Lundquist et al. 2008), we limit the analysis and discussion to the grid boxes within the NFARB basin. BBH Ptype generated with lowering offset (−100 and −200 m) did not greatly change this spatial pattern (not shown).

Distributions of the DMIP 2 air temperature values for snow and rain grid boxes determined by BBH data (including the offsets of −100 and −200 m) are shown in Fig. 7. The median air temperatures for snow grid boxes based on BBH Ptype data are 3.3°, 3.5°, and 3.6° for BBH with 0-, −100-, and −200-m offsets, respectively. On the other hand, the median air temperature for rain grid boxes is 7.0°C with little sensitivity to BBH offset. Past studies indicated that at a surface temperature below 0°C, 90% of precipitation occurs as snow, while over 90% of precipitation occurs as rain above a temperature of 3°C (United States Army Corps of Engineers 1956; United States Army Corps of Engineers 1988).
Compared to these past studies, our results show that 90% of snowfall occurred below approximately 5.5°C, which is much higher than expected. However, 90% of rainfall grids are over 3.8°C, which is closer to what the past studies suggested. The median temperature of the snow grid boxes is over 1.3°C higher than the PXTEMP of 2°C. These higher temperature snow grid boxes from BBH Ptype data occurred mostly at high elevations where more rain grid boxes are produced by the PXTEMP method as shown in Fig. 6. It is likely that there is uncertainty in the interpolated temperature values from the DMIP 2 temperature grids, particularly at higher elevations where SNOTEL sites are located. SNOTEL temperature measurements were found to be more erroneous than the other meteorological measurements (Serreze et al. 1999). Issues with SNOTEL temperature data were also described by Pepin et al. (2005), who examined long-term observations of high-elevation surface temperature over the western United States. There is also uncertainty in BBH Ptype grids because of the possible spatial variation of BBH. In addition to the orographic effect on BBH at its intersection with the upwind mountain slope (Kingsmill et al. 2008; Lundquist et al. 2008; Minder et al. 2011), the spatial resolution used for this study (i.e., 4 km) is possibly too coarse to represent the topography for this rugged basin. There is a large elevation range within a 4-km grid box (e.g., 250–1250 m based on 500-m DEM). The variability of elevation within each 4-km grid box is likely to cause errors in BBH Ptype mapping.

Next, hourly psfrac and fractional areas of rainfall (prfrac, percent) over the entire NFARB basin were computed from BBH Ptype and three PXTEMP Ptype grids. Figure 8 shows time series of the computed psfrac and prfrac values for the period of 0000 UTC 20 December 2005 to 2300 UTC 5 January 2006, which contains the largest precipitation–runoff event in the HMT-West 2006 observation period. For this analysis, we compared the BBH Ptype grids with no offset and −200-m offset to PXTEMP Ptype grids derived using three values of PXTEMP = 1°C, 2°C, and 3°C. Values of psfrac and prfrac were computed only during hours of nonzero precipitation. As shown in the Fig. 8, there is little correspondence in both psfrac and prfrac between the BBH Ptype and any PXTEMP Ptype regardless of different PXTEMP values and BBH offsets. In particular, for the precipitation event in the period from 0000 UTC 30 December 2005 to 2300 UTC 31 December 2005, BBH Ptype indicated that little snowfall occurred across NFARB, but PXTEMP Ptype produced snowfall over 20% of the basin for 1°C and up to 40% of the basin with 3°C. Correspondingly, BBH Ptype indicated psfrac values of nearly 100%, while PXTEMP Ptype indicated that smaller portions of the basin area received rainfall. On 31 December 2005, maximum daily air temperatures in the upper zone of the basin were in the range of 6°C–9°C, while minimum temperatures were below 0°C when PXTEMP Ptype produced snowfall. During the later event in the period from 1200 UTC 2 January to 0000 UTC 3 January 2006, however, BBH Ptype produced more snowfall area than the PXTEMP method. This series of precipitation events illustrates that PXTEMP Ptype produced more snowfall than BBH Ptype prior to a frontal passage (i.e., 0000 UTC 30 December to 2300 UTC 31 December 2005) when snow level is high compared to the post-cold-frontal passage (i.e., 1200 UTC 2 January to 0000 UTC 3 January 2006). However, PXTEMP Ptype
produced less snowfall area after the cold front passed and the snow level dropped. The BBH snow grid boxes with temperature higher than 2°C occurred over the high-elevation portions of the basin (see Fig. 6b) after the front passed the basin. It is noted that hourly temperature disaggregation using climatological diurnal temperature patterns is unlikely to reproduce rapid temperature drops associated with a specific frontal storm.

**FIG. 7.** Cumulative distribution of air temperature values at (left) snow and (right) rain grid boxes derived from the BBH Ptype with various offsets during the period from 0000 UTC 20 Dec 2005 to 2300 UTC 24 Apr 2006 over the NFARB.

**FIG. 8.** Hourly time series of fraction of (top) snowfall area and (bottom) rainfall area over NFARB from 0000 UTC 30 Dec 2005 through 2300 UTC 5 Jan 2006. The fractional area of each type of precipitation was computed based on hourly BBH Ptype grid and hourly air temperature grids with three PXTEMP values.
To illustrate how these fractional area differences impact the actual depths of snowfall and rainfall within the basin, mean areal snowfall and rainfall volumes were computed based on the BBH Ptype and the three PXTEMP Ptype grids by averaging precipitation depths of each grid box assigned to snowfall or rainfall, respectively. Figure 9 shows mean areal snowfall and rainfall volume over the same events as in Fig. 8. The difference in the mean basin-wide rainfall depth was below 2 mm during the prefrontal passage period, while a larger discrepancy in snowfall depth (over 2-mm difference between BBH with 0-m offset and 2°C PXTEMP) was seen during the same period. The better agreement of rainfall depth estimates is reasonable given that the large event in this period was predominately a rain event evidenced by the higher prfrac (Fig. 8). As shown in Figs. 8 and 9, lowering BBH did not change snow amounts very much, and the amounts are still far from those generated by PXTEMP Ptype.

As indicated by the above analysis, there is a large discrepancy in precipitation typing between the two methods. Moreover, BBH Ptype produced less snowfall than PXTEMP Ptype. Given this large difference, we compare both methods against point snow observations. Table 3 shows seasonal snow accumulation (ΣSWE) and its ratio to total precipitation over the entire water year (ΣSWE/ΣPRCP). Seasonal snow accumulation was computed by summing up positive increments for each time step over the season. This computation was performed using observed SWE and gauge precipitation measured at three snow observation sites [Blue Canyon (BLC), Huysink (HYS), and the Central Sierra Snow Laboratory (CSS) SNOTEL site; see Fig. 2 for their locations] and corresponding gridded SWE values computed by SNOW-17 with PXTEMP and BBH methods. The ΣSWE/ΣPRCP ratio is a better indicator of precipitation partitioning than snow accumulation alone because the latter can be contaminated by potential errors or differences of precipitation compared to the gauge values. Also, this quantity is not affected by snowmelt errors from SNOW-17.

As shown in Table 3, the ΣSWE/ΣPRCP is similar to observations for the Huysink and CSS sites for all BBH cases and the PXTEMP 1°C case, indicating precipitation is partitioned fairly well, particularly at these higher-elevation sites. It is noted that precipitation partitioning seems to be less sensitive to BBH offset. However, the skill of precipitation typing deteriorates for both PXTEMP and BBH methods at the lower-elevation Blue Canyon site. From this result, it can be hypothesized that errors in precipitation typing may be frequent at elevations where RSELV is often located. The elevation of the Blue Canyon site is 1609 m, very near the middle elevation in the basin. It is noted that the PXTEMP method generally produces more snow than BBH for all the locations, which is consistent with the results from the above analysis. The next section illustrates how this discrepancy impacts hydrologic simulations.
b. Impacts of BBH Ptype on streamflow simulations

We generated lumped and distributed model simulations with and without the BBH Ptype grids for the period spanning 1 October 2005 to 30 September 2006. Tests were performed in the framework of the DMIP 2 experiments (M. B. Smith et al. 2010a, 2012b, manuscript submitted to *J. Hydrol.*). Our tests used the DMIP 2 hourly 4-km QPE and temperature data as model forcings. Simulations were run over the entire year to capture both the snow accumulation and melt periods. As in the DMIP 2 experiments in the NFARB, simulations were generated in retrospective mode with no data assimilation. The period from 1 October 2003 to 30 September 2005 was used as a “warm up” period to let the model states equilibrate.

For the lumped model tests, the hourly 4-km DMIP 2 gridded QPE data were converted to hourly MAP values for the upper- and lower-elevation zones of NFARB (i.e., above and below 1524 m). SNOW-17 and SAC-SMA parameters in each elevation zone were calibrated for use in the DMIP 2 project over the period from 1 October 1988 through 30 September 1997. Likewise, parameters for the gridded SNOW-17 and SAC-SMA models were taken from HL-RDHM calibrations for DMIP 2 (M. B. Smith et al. 2010a, 2012b, manuscript submitted to *J. Hydrol.*). HL-RDHM was run on a 4-km grid, which corresponds to the DMIP 2 4-km gridded QPE and temperature data. As in Maurer and Mass (2006), no recalibration of the SNOW-17 or SAC-SMA parameters was performed with the BBH Ptype data.

1) OVERALL COMPARISON OF STREAMFLOW SIMULATIONS

Figure 10 compares distributed model simulations using PXTEMP and BBH. The top panel shows the hours when BBH observations are available, and the bottom panel shows observed flows along with the baseline PXTEMP 2°C. The simulation using BBH data produces more midwinter peak flows, which come from rain events (note that BBH produces more rainfall, particularly over the mid-elevation portion of the basin). This reduces snow accumulation, thus affecting snowmelt runoff as indicated by less flow than the PXTEMP simulation. Given the larger magnitude of flow during early season and percentage difference (up to 30%–40% for some events), precipitation typing has a significant impact on streamflow simulations from a flood forecasting perspective.

2) PERFORMANCE STATISTICS

To quantify how the BBH method impacts simulation performance compared with observations, we present entire-season statistics and event statistics as follows. Hourly goodness of fit statistics were computed for the period 1 October 2005 to 30 September 2006 to show the overall impacts of using the BBH grids (without lowering offsets). Table 4 presents the hourly %Bias and $r_{mod}$ statistics computed for the lumped and distributed models with and without BBH Ptype data. Lumped and distributed model simulations without using BBH Ptype data are labeled “baseline.” %Bias is a measure of total volume difference between two time series and is computed as

$$%\text{Bias} = \frac{\sum_{i=1}^{N} (S_i - O_i)}{\sum_{i=1}^{N} O_i} \times 100,$$

where $S_i$ is the simulated discharge for each time step $i$, $O_i$ is the observed value, and $N$ is the total number of values within the time period of analysis.

The quantity $r_{mod}$ is a modified form of the correlation coefficient and measures the agreement in hydrograph shape (McCuen and Snyder 1975). Recognizing the
tendency of the normal correlation coefficient to be overly influenced by outliers and to be insensitive to differences in the size of hydrographs, we select $r_{\text{mod}}$ to objectively compare hydrographs. In this statistic, the normal correlation coefficient is reduced by the ratio of the standard deviations of the observed and simulated hydrographs. The minimum standard deviation (numerator) and maximum standard deviation (denominator) are selected so as to derive an adjustment factor less than unity; $r_{\text{mod}}$ is computed using Eq. (3):

$$r_{\text{mod}} = r \times \frac{\min(\sigma_{\text{sim}}, \sigma_{\text{obs}})}{\max(\sigma_{\text{sim}}, \sigma_{\text{obs}})}, \quad (3)$$

where $r$ is the correlation coefficient, $\sigma$ is the standard deviation of the simulation (sim) or observed (obs). These measures were used in the analysis of the DMIP1 and DMIP2 results and are described in Smith et al. (2004).

For the lumped modeling case, use of the BBH Ptype data resulted in worse run period statistics than the simulation without BBH Ptype data. However, the values of $-4.4$ and 0.78 for %Bias and $r_{\text{mod}}$, respectively, are still acceptable. For example, %Bias values within $\pm 5\%$ are the suggested goal in NWS model calibration.

**TABLE 4. Hourly %Bias and $r_{\text{mod}}$ streamflow simulation statistics for the period of 1 October 2005 to 30 September 2006.**

<table>
<thead>
<tr>
<th>Model</th>
<th>Simulation</th>
<th>%Bias</th>
<th>$r_{\text{mod}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumped</td>
<td>Baseline</td>
<td>$-0.9$</td>
<td>0.89</td>
</tr>
<tr>
<td>Lumped</td>
<td>With BBH Ptype</td>
<td>$-4.4$</td>
<td>0.78</td>
</tr>
<tr>
<td>Distributed</td>
<td>Baseline</td>
<td>0.33</td>
<td>0.88</td>
</tr>
<tr>
<td>Distributed</td>
<td>With BBH Ptype</td>
<td>0.43</td>
<td>0.80</td>
</tr>
</tbody>
</table>
(Smith et al. 2003). As shown in Table 4, the %Bias and $r_{mod}$ were slightly degraded in the distributed modeling case. As indicated by the %Bias values, the simulations with the BBH Ptype data resulted in less volume generated by the lumped model and more volume generated by the distributed model compared to the baseline cases.

Specific precipitation–runoff events were also examined, as experience in DMIP 1 and DMIP 2 showed that overall run period statistics may mask the impacts of a particular method or model (Reed et al. 2004; Smith et al. 2012a). Table 5 presents the hourly %Bias and $r_{mod}$ statistics computed for 11 specific precipitation–runoff events within the run period. The simulations for these events were extracted from the continuous hourly simulations. No adjustments were made to initial conditions at the start of each event. Bold font numbers indicate cases in which the %Bias and $r_{mod}$ statistics were improved by using BBH Ptype data (%Bias closer to 0 and values of $r_{mod}$ closer to 1.0). All other cases show degradation from BBH Ptype data. The use of BBH Ptype data resulted in improvement of both statistics in three events for the lumped simulations and four events for the distributed model. More mixed improvement (i.e., only one statistical measure improved) occurred for the lumped event simulations. The improvements in %Bias for some events were offset by other cases, so that the overall %Bias using BBH Ptype data was worse than the lumped baseline.

Most of the improvements from using the BBH method occurred when PXTEMP simulations underestimated the observed streamflow since BBH Ptype generate more rainfall than PXTEMP method. This is illustrated in Fig. 11, which shows simulated streamflow hydrographs for two cases improved by BBH data: events 1 (top panel) and 9 (bottom panel). Also shown is a case in which the BBH method led to a deteriorated simulation: event 5 (middle panel). For both models, use of the BBH Ptype data leads to higher peak flows and earlier hydrograph rising limbs. The one exception is the lumped simulation of the 1–2 December 2005 event in which the lumped simulation with BBH Ptype data has less volume and leads to deterioration of the statistics.

It is certain that further investigation is needed to understand the mixed impacts seen in Table 5. One of the reasons that more consistent improvement was not seen with the BBH Ptype data could be that the hydrologic model is calibrated with PXTEMP value of 2°C. Therefore, it is not surprising that the model performance deteriorated with new data introduced into the model even if the new data might have more accuracy. This adverse effect of a new data source on hydrologic model performance was also illustrated by Yatheendradas et al. (2012), who showed that assimilation of satellite snow cover into a well-calibrated hydrologic model degraded streamflow simulations. Another potential reason may lie in the complex relationship between the radar-derived snow level and the elevation where snow actually accumulates on the ground (Lundquist et al. 2008). They noted that whether the snowfall contributes to accumulation or immediate melt depends on several physical parameters/processes such as snowfall rate, sublimation, atmospheric lapse rate, and ground heat flux. In addition, the retrospective streamflow simulation results for this period may be dominated by the large 31 December 2005 to 2 January 2006 precipitation event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event. This event has been identified as an atmospheric event.
of using BBH Ptype data for these events. Also, it seems that the lumped model behaved differently than the distributed model when BBH Ptype was used.

3) STREAMFLOW SENSITIVITY TO PXTEMP AND BBH OFFSET

To provide insight into how RSELV errors propagate to streamflow errors, Fig. 12 shows time series of percentage difference in simulated streamflows from the distributed model with the measured BBH and the BBH lowered by two offset distances (−100 and −200 m). Before snowmelt season, larger offsets (lowering measured BBH) produce less streamflow because lowering BBH, and thus snow level, produces more snowfall and, therefore, less immediate runoff. During the snowmelt period, however, because of more snowpack accumulation for lowered BBH simulation, greater runoff is generated. It is noted that the magnitudes of differences in streamflow are greater during the snow melt period (~15% for 100-m offset and ~25% for 200 m at maximum) than the accumulation period (~5% for 100-m offset and ~15% for 200 m at maximum), and the differences persist until the end of the season. This result implies larger sensitivity of streamflow simulation to snow-level determination than snowfall fractional area and amount presented in the previous sections (i.e., Figs. 8, 9), not only in the earlier season when more precipitation events are common, but also in the snowmelt season.

We also examined the sensitivity of streamflow simulations to PXTEMP because there exists uncertainty in the value of such thresholds. Figure 13 displays percentage differences in streamflow simulated with a PXTEMP of 2°C (a base simulation) and with two PXTEMP values (1° and 3°C). The simulated hydrographs show the expected trend in the discharge magnitude: as PXTEMP increases (decreases), more (less) of the falling precipitation is typed as snow and less as rain, leading to less (more) runoff in snow accumulation period and more (less) snowmelt runoff during spring time. It is noted that the hydrographs are more sensitive to the variation in PXTEMP values than to BBH offsets. However, this might be caused by BBH data intermittency over the season.
5. Summary and conclusions

During the HMT-West experiment, BBH measurements were performed continuously from 1 December 2005 through 19 April 2006 and generated 408 h of the BBH-derived Ptype. These hourly BBH data were mapped to surface topography to develop grids of surface precipitation type (rain or snow) at the 4-km resolution. We analyzed these data, focusing on using the data in spatially lumped and distributed modeling frameworks in order to understand differences in Ptype estimates between traditional and new methods and to examine impacts of the new approach on hydrologic modeling applications.

The comparison between both Ptype estimates shows that below 1524 m, the BBH Ptype grids compared favorably to gridded precipitation type derived using the PXTEMP value of 2°C from DMIP 2. This is reasonable because when BBH is detected at the Alta S-PROF site, the precipitation type at elevations lower than the site is certain to be rainfall. Above this elevation, the agreement in precipitation type was between 50% and 60%. This disagreement can be caused by both uncertainty of BBH method or PXTEMP method. We conclude that, overall, BBH Ptype produced more rainfall than PXTEMP in midelevations, possibly during a storm before a cold front passes the area, but BBH Ptype produced more snowfall than PXTEMP after the frontal passage in higher elevations. Since greater precipitation amounts occur before the frontal passage, BBH Ptype may lead to more rainfall and, therefore, more immediate runoff. Lowering measured BBH does not change this conclusion.

One of the objectives of this study was to evaluate the impact of the new data on streamflow simulations (without recalibration) compared to model simulations with PXTEMP data. The BBH Ptype grids were used as an additional input data stream for existing lumped and distributed hydrologic models to generate continuous, hourly, retrospective streamflow simulations. Use of new Ptype data affects simulations of not only peak flow immediately after individual storms that produce rainfall, but also of spring streamflow caused by snowmelt. In addition, the simulations are impacted by various BBH offsets and PXTEMP values. This highlights the importance of Ptype estimations in hydrologic simulation over the Sierra Nevada basin.

In our experiment, the new Ptype data did not improve the overall model performance, though the %Bias and $r_{mod}$ run-period statistics were still within the bounds suggested for NWS hydrologic model calibration (Smith et al. 2003). Looking at individual events, mixed performance was seen, with some event simulations being improved while others were degraded. Improvements were seen when the simulation with PXTEMP Ptype underestimates peak flow. While these mixed results are disappointing, they are not surprising. It is important to note that both lumped and distributed models were calibrated as part of the DMIP 2 experiment using a PXTEMP value of 2°C. As with Maurer and Mass (2006), no recalibration was performed with the BBH Ptype data because there is no long-term BBH data available at this site for calibration. From a broader viewpoint, our mixed results perhaps reflect the fact that new data sources and hydrologic modeling advances may not consistently provide improved simulations compared to an existing approach or benchmark. For example, our results here (testing a new source of data) parallel the results of DMIP 2 in the NFARB (testing new hydrologic models). In the NFARB DMIP 2 experiments, distributed models provided more accurate outlet hydrograph simulations than a lumped model for some precipitation events in an 18-yr study period, but not all (Smith et al. 2012b, manuscript submitted to J. Hydrol.). Moreover, DMIP 1 and 2 tested over 20 distributed models applied to basins in Oklahoma (Smith et al. 2012a; Reed et al. 2004). In those experiments, only a few distributed models were able to provide improved simulations compared to the lumped model benchmark.

6. Recommendations

This study did not provide a conclusive answer to the question, does uniform-height BBH data produce better hydrologic model results than traditional temperature-based methods? Nonetheless, given that our mixed results herein complement our previous experience with investigating new sources of data and models (e.g., DMIP 1 and 2), we recommend continued analysis and application of measured BBH data to hydrologic
modeling. Estimation of surface precipitation type from BBH data is based on a physically observable quantity and more objective method than temperature threshold approaches, which are more empirical. Use of PXTEMP in both lumped and distributed modeling introduces a few major sources of uncertainty in precipitation type estimates. Such uncertainty originates from hourly distributed temperature estimates and the PXTEMP value itself. It is likely that errors in the precipitation type estimates come from using a temporally and spatially constant PXTEMP value, which in truth could vary in space and time depending on spatiotemporal variation of the atmospheric conditions. While the spatial variation of BBH exists and poses challenges to distributing BBH over the area, improvement of the PXTEMP method will depend on qualities of both temperature and PXTEMP grids. Lumped model applications require a temperature lapse rate, which is typically fixed in both BBH and PXTEMP grids. Lumped model applications require a temperature lapse rate, which is typically fixed in an NWS RFC operational setting (i.e., 6.5°C). In complex terrain, seasonal and diurnal as well as geographic variations of lapse rate have been evident based on observations (Minder et al. 2010). Therefore, fixed lapse rates contribute errors to PXTEMP estimates in lumped model applications. Finally, calibration of the hydrologic models with BBH Ptype data could become possible as datasets grow in length, which is believed to improve the simulation results with the BBH data.

Validation of distributed Ptype from both BBH and PXTEMP with independent sources is challenging but necessary to understand how errors of Ptype estimates from both BBH and PXTEMP methods propagate error in streamflow simulations. Precipitation type observed with optical disdrometers could be used for direct comparison. Unfortunately, this dataset was not available during our 2005/06 season. Disdrometer observations were collected during later years of the HMT-West program and should be used to assess the precipitation type on the ground.

The spatial variability of BBH and subsequently derived Ptype grids should be investigated. BBH measurement at a point scale is discussed by White et al. (2002, 2010). However, one of the potential error sources for distributed BBH Ptype was neglecting the spatial variation of BBH. Atmospheric simulations with a mesoscale numerical weather model can provide insight into detailed physical mechanisms of BBH variability (Minder and Kingsmill 2013). Understanding the relationship between the space- and time-varying BBH drops and atmospheric and terrain factors will be beneficial for determining the elevation of the BBH–terrain intersection from the point BBH measurement. Correspondingly, spatially distributed BBH data derived at finer grid resolutions should be considered.

One limitation to our study was the detection of BBH located below the elevation where the radar is situated (in this study, Alta is located at elevation 1085 m). This was one of the causes of missing BBH Ptype data. As a result, our BBH Ptype data was interspersed with PXTEMP data because when the BBH Ptype data were missing, the SNOW-17 model defaulted to other methods of rain–snow partitioning. As an interim solution, it may be beneficial to interpolate missing BBH observations rather than defaulting to another precipitation typing method. A profiling radar located at a lower elevation might capture more low-elevation BBH events. However, the siting or use of additional radars must consider the drop in BBH in the direction from upwind to the mountain slopes.

The retrospective simulations reported herein did not allow for the adjustment of initial streamflow and snow cover conditions at the start of each precipitation–runoff event. It is recommended that hindcasting experiments be considered in which the model states prior to each event are adjusted so that the impacts of using BBH Ptype data may be more clearly identified for each event. Techniques for streamflow data assimilation with the lumped model (e.g., Seo et al. 2009) and distributed model (e.g., Lee et al. 2011) are progressing.

Anticipating that further analyses would lead to more consistent improvements, use of the BBH data (spatially constant or variable) for precipitation typing in real time hydrologic forecasting would be computationally feasible. Gridded data such as multisensor QPE have been used in NWS operational forecasting for well over a decade. Based on such experience, we believe the generation and use of gridded BBH Ptype data would be straightforward.

Acknowledgments. The authors thank Pedro Restrepo of the North Central River Forecast Center, John Halquist of the National Operational Hydrologic Remote Sensing Center (NOHRSC), Art Henkel and Rob Hartman of the NWS California–Nevada RFC, and Dan Gottas and one anonymous reviewer at the Earth System Research Laboratory for critical reviews that helped improve initial versions of this manuscript. We are grateful for the assistance of John Schaake in deriving the hourly temperature data. We thank the reviewers for their helpful comments that further improved the paper.

REFERENCES


