The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

The Cryosphere Discussions
Open Access
EGU

How Much Cryosphere Model Complexity is Just Right? Exploration Using the Conceptual Cryrosphere Hydrology Framework

Thomas M. Mosier[1,2], David F. Hill[1,3], and Kendra V. Sharp[1,2]

[1]Water Resources Graduate Program, Oregon State University, Corvallis, Oregon, USA 97331
[2]Mechanical, Industrial, and Manufacturing Engineering, Oregon State University, Corvallis, Oregon, USA 97331
[3]Civil and Construction Engineering, Oregon State University, Corvallis, Oregon, USA 97331

Correspondence to: mosier.thomas@gmail.com

Abstract. Robustly understanding connections between the climate, cryosphere, and streamflow is necessary to make informed decisions regarding water resources in mountainous regions. Conceptual models, for instance variations on simple degree-index (SDI) or enhanced temperature index (ETI) formulations of cryosphere heat transfer and melt, are often applied in mountain environments. SDI and ETI models perform relatively well over short time periods or small geographic extents, but their performance is known to degrade over longer periods or across larger geographic regions, which makes them ill-suited for climate change impact studies. Our work seeks to quantitatively assess differences in performance of multiple existing and novel conceptual cryosphere model formulations across climatic conditions and geographic regions. To accomplish this, we present the Conceptual Cryosphere Hydrology Framework (CCHF), which is a set of coupled process modules, computational engine, and analysis tools, to assess the accuracy and precision of conceptual cryosphere models. We compare existing SDI and ETI formulations and novel conceptual cryosphere process modules that include net longwave radiation in the heat transfer formulation (termed the Longwave, Shortwave, Temperature - LST - formulation) and a cold content representation of the snowpack internal energy. We implement the CCHF for the Gulkana and Wolverine glaciated watersheds in Alaska, which have markedly different climates and contain longterm US Geological Survey benchmark glaciers. Overall, of the models tested, the LST heat representation with a cold content internal energy representation is the most accurate and precise model. The CCHF is open-source and we encourage interested parties to use it to identify further refinements to cryosphere model formulations.

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

1 Introduction

Many regions of the Earth's cryosphere are currently changing and are projected to continue changing significantly under a range of projected future climate scenarios (*Stocker*, 2013; *Mankin et al.*, 2015). Understanding the impacts of climate change on the cryosphere, and in turn on timing and magnitude of streamflow originating from the cryosphere, requires implementing a spatially and temporally robust cryosphere hydrology model. Two categories of cryosphere models are energy balance and conceptual. Our basic definition of the difference between these model types is that energy balance models represent each heat flux independently and base calculation of heat fluxes on all physically relevant parameters, while conceptual models lump multiple heat fluxes together and use parameterizations to reduce the number of input variables. In practice, these model types exist on a spectrum rather than as discrete categories.

Energy balance models are theoretically more robust than conceptual models because the relative balance of heat fluxes changes over time and between locations (*Kustas et al.*, 1994; *Sicart et al.*, 2008; *Huss et al.*, 2009). Yet, many cryosphere studies utilize conceptual models, such as simple degree index (SDI) and enhanced temperature index (ETI) structures (e.g. *Wang et al.*, 2004; *Hagg et al.*, 2007; *Jung and Chang*, 2011; *Lutz et al.*, 2014). In contrast, energy balance models require a variety of inputs that are even more sparsely measured in mountain environments and that vary significantly over short spatial and temporal scales (see discussion on the surface energy balance in *Sturm*, 2015). Wind speed, for example, linearly scales each convective heat flux (i.e. sensible and latent), and is therefore necessary for energy balance models. The spatial distribution of wind speed is, however, difficult to characterize even if there are measurements of wind speed at points within the region (*Marks et al.*, 1992; *Marks and Winstral*, 2001). Therefore, it warrants consideration of whether including variables such as wind speed improves or reduces model performance.

Conceptual models, which typically only require climate inputs of precipitation and temperature, perform well under many circumstances (*Ohmura*, 2001), with ETI models typically performing better than SDI models (*Hock*, 1999; *Pellicciotti et al.*, 2005). The reasons for relatively good performance of SDI and ETI models are that many heat fluxes are functions of temperature (e.g. sensible convection and longwave radiation), seasonality of air temperature and shortwave radiation are often correlated for a given location, and shortwave radiation fluxes tend to be significantly larger than convective heat fluxes during melt conditions (*Marks et al.*, 1992; *Cline*, 1997). While previous studies have compared ETI to energy balance models (e.g. *Hock*, 2005), the present study seeks to assess whether adding more physical basis to conceptual cryosphere structures (while holding the required inputs constant) enhances model accuracy and precision.

Identifying improvements to conceptual models requires a framework for comparing existing structures and systematically assessing novel representations. Other hydrology intercomparison projects have been carried out (e.g. *Smith et al.*, 2004, 2012; *Clark et al.*, 2008) and are being developed (e.g. *Clark et al.*, 2015a, b). The comparison carried out by *Smith et al.* (2004) and *Smith et al.* (2012) assess several hydrologic models, many of which are conceptual, but implement each model independently instead of through a common framework. *Clark et al.* (2008) creates a common modeling structure for implementing lumped variants of several hydrologic models for two basins in the USA, but does not model snowpack accumulation and melt. The Structure for Unifying Multiple Modeling Alternatives (SUMMA) is a much more generalizable intercomparison framework

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

by *Clark et al.* (2015a). In SUMMA, the relevant mass and energy conservation relationships are implemented, and the structure is designed to allow alternative spatial representations of each process relevant to the system. Thus, in theory SUMMA can be developed to implement any physically-based representation of a distributed hydrologic system. In practice, SUMMA development has not focused on the processes and conditions typical of data-sparse mountain regions (*Clark et al.*, 2015a, b).

Our present work seeks to assess and improve conceptual representations of snow and glacier processes. To achieve this, we have developed the Conceptual Cryosphere Hydrology Framework (CCHF). The CCHF modularizes alpine and glacier surface processes, and includes routines to optimize model parameters and analyze outputs. We assess four heat flux representations (including three existing and one novel formulation) and two mass and energy representations (one existing and one novel). We focus on these processes because their representation may depend significantly on climate regime (i.e. existing conceptual representations of these processes may not be robust under climate change). We conduct our assessment for Gulkana and Wolverine watersheds in Alaska (Fig. 1), for the period July 2000 through June 2010, and evaluate model performance using streamgage, glacier stake, and snow-covered area (SCA) observations. In addition to our present assessment, a valuable attribute of CCHF is that it is open-source and can easily be implemented for other regions, different climate forcing datasets, or novel process representations.

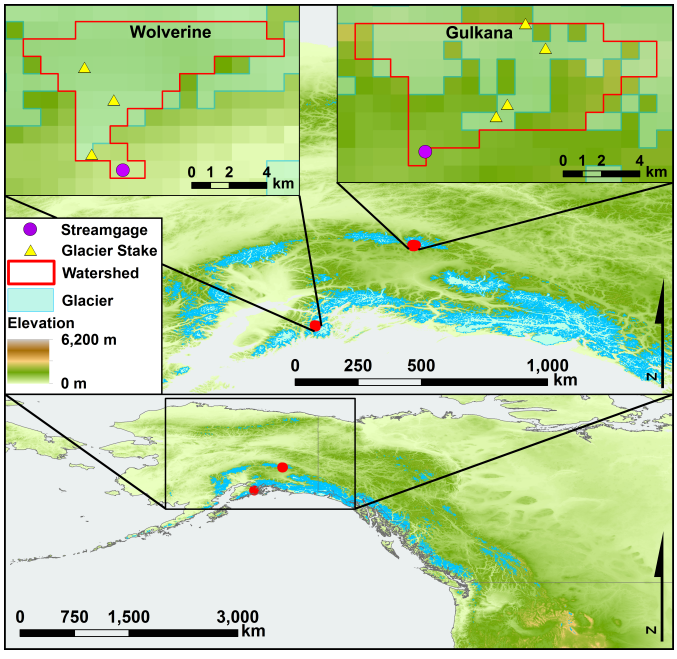

Figure 1. Geographic extent of the glaciated Gulkana and Wolverine watersheds. The digital elevation model is produced by WorldClim (*Hijmans*, 2011) and glacier coverage is from the Randolph Glacier Inventory version 5 (*Pfeffer et al.*, 2014).

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

1.1 Existing Conceptual Cryosphere Representations

1.1.1 Heat Flux to Melt

A basic yet still widely used conceptual cryosphere heat and melt formulation is the SDI structure (two well known examples are Hydrologiska Byråns Vattenbalansavdelning, presented in *Lindström et al.*, 1997, and the Snowmelt-Runoff Model, presented in *Martinec*, 1975). In SDI models, melt, M ($[M] = \mathrm{m}^{-1}\ \mathrm{s}^{-1}$, which throughout represents depth of melt per unit time per unit area; square brackets denote a variable's units), occurs when the surface air temperature at a grid cell, T_a ($[T_a] = {}^{\circ}\mathrm{C}$), is above a temperature threshold, T_0 ($[T_0] = {}^{\circ}\mathrm{C}$) (*Singh et al.*, 2000; *Wang et al.*, 2004). Otherwise, no melt occurs and incident precipitation during that time step adds to the snowpack. When melt occurs at a given grid cell, its magnitude is determined as

$$M = \begin{cases} f_m \left(T_a - T_0\right) & \text{if } T_a > T_0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where f_m is the degree-index factor ($[f_s] = \mathrm{m}^{-1}\ {}^{\circ}\mathrm{C}^{-1}\ \mathrm{s}^{-1}$).

ETI formulations are a category of conceptual cryosphere formulations that model heat transfer as a function of air temperature and shortwave radiation. Two ETI formulations, by *Hock* (1999) and *Pellicciotti et al.* (2005) respectively, are

$$M = \begin{cases} \left(a_i \left(1-\alpha\right) I + f_m\right) \left(T_a - T_0\right) & \text{if } T_a > T_0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$M = \begin{cases} a_i \left(1-\alpha\right) I + f_m \left(T_a - T_0\right) & \text{if } T_a > T_0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where a_i is a fitting parameter that scales melt from shortwave radiation (Eq. 2: $[a_i] = \mathrm{m}^{-1}\ \mathrm{J}^{-1}\ {}^{\circ}\mathrm{C}^{-1}$; Eq. 3: $[a_i] = \mathrm{m}^{-1}\ \mathrm{J}^{-1}$), α is the surface albedo (unitless), I is the measured or modeled shortwave radiation ($[I] = \mathrm{W\ m}^{-2}$). Throughout, the ETI heat formulations described in Eqs. 2 and 3 are referred to respectively as ETI(H) and ETI(P). As a note, within the CCHF, the units of fitting parameters in Eqs. 1-3 are altered such that the lefthand side of the equations have units of W per meter squared. ETI(P) is more physically representative than ETI(H) because temperature and shortwave radiation scale melt independently (*DeWalle and Rango*, 2008); however, *Pellicciotti et al.* (2005) only describes application of ETI(P) to individual points on a glacier, rather than as a distributed model. Conversely, ETI(H) is freely available as a spatially-distributed modeling package (*Hock*, 1999).

1.1.2 Internal Energy of Snow and Ice

Snow and ice melt in Eqs. 1-3 is proportional ti the heat transfer in the current time-step (modified using a step function), and therefore these formulations do not represent the internal energy of snow or ice in a physically meaningful way. A common concept in snow process modeling, however, is cold-content, CC ($[\mathrm{CC}] = \mathrm{m}$), which is defined as

$$\mathrm{CC} = \frac{c_i\, \rho_s}{L_f\, \rho_w}\, d_s \left(T_m - T_{s,i}\right) \tag{4}$$

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

where c_i is the specific heat of ice ($[c_i] = \text{J kg}^{-1}\ \text{K}^{-1}$), ρ_s is the density of the snow mass ($[\rho_s] = \text{kg m}^{-3}$), L_f is the latent heat of ice ($[L_f] = \text{J kg}^{-1}$), ρ_w is the density of water ($[\rho_w] = \text{kg m}^{-3}$), d_s is the depth of snow ($[d_s] = \text{m}$), T_m is the temperature at which melt occurs (i.e 0 °C; $[T_m] = °\text{C}$), and $T_{s,i}$ is the internal temperature of the snow mass (assuming the snow is isothermal or that $T_{s,i}$ is the mass-averaged temperature; $[T_{s,i}] = °\text{C}$). CC is simply a statement of the magnitude of sensible energy deficit of the snow mass in terms of the latent energy required to melt the same mass of water.

CC is a natural conceptual formulation of internal energy to combine with a conceptual heat flux representation (e.g. Eqs. 1-3) because both are formulated as depth of water equivalent. *Hock* (2005) states that the concepts of energy deficit (referring to *Van de Wal and Russell*, 1994) and CC, or 'negative melt' (referring to *Braun and Aellen*, 1990), have been incorporated into previous conceptual cryosphere models; however, the above references that *Hock* (2005) cites do not mention treatment of internal energy. Therefore, we are not aware of any previous conceptual cryosphere models that incorporate internal energy or 'negative melt' into their formulation. Many conceptual cryosphere model implementations are able to avoid including internal energy representations by focusing only on the melt season (e.g. *Kustas et al.*, 1994, *Hock*, 1999, *Pellicciotti et al.*, 2005, and *Sicart et al.*, 2008). Section 2.2 explains how CC is incorporated into the CCHF.

Even after the CC of a snowpack is satisfied (i.e. becomes zero), the snowpack must ripen before releasing water. Snowpack ripening refers to the process by which snow holds water interstitially until the interstitial forces are overcome by gravity and the melt water drains from the snowpack. While the percentage holding capacity of snow varies, a typical value is 3% (*DeWalle and Rango*, 2008). Some existing conceptual cryosphere models include snowpack liquid water holding capacity (e.g. *Kokkonen et al.*, 2006) and others do not (e.g. *Wang et al.*, 2004). Liquid water holding capacity is included in the CCHF and discussed in Section 2.2.

1.2 Climates of Gulkana and Wolverine Watersheds

While Gulkana and Wolverine watersheds are both within the Gulf of Alaska, their climates are markedly different (Fig. 2). With respect to air temperature, the upper end of the air temperature ranges (i.e. summer temperatures) are similar between Gulkana and Wolverine but Gulkana is significantly colder at the low end of the distribution (i.e. corresponding to colder winter temperatures at Gulkana). At all percentiles in the distribution, Wolverine's precipitation is greater than Gulkana's. The zeroth percentile non-exceedance probability of precipitation for Wolverine is greater than 150 mm per day while for Gulkana it is approximately 60 mm.

Based on these differences, Gulkana's climate can be categorized as more continental while Wolverine's is more maritime (*Armstrong and Armstrong*, 1987; *Ukraintseva*, 2009). These climate classifications are consistent with findings for the Alaska region presented in *Bieniek et al.* (2012) (using cluster analysis of station records) and (*Simpson et al.*, 2002) (based on Parameter-elevation Regressions on Independent Slopes Model interpolated climate surfaces; *Daly et al.*, 2002). Continental cryosphere climates tend to be colder and drier than maritime cryosphere climates. These climatic differences lead to snowpacks in continental climates being much more polythermal and generally less dense than snowpacks in maritime climates (*Armstrong and Armstrong*, 1987; *DeWalle and Rango*, 2008; *Ukraintseva*, 2009). Higher densities in maritime snowpacks are partially caused by atmospheric conditions during snowfall and also by repeat freezing and melting cycles, which are more

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

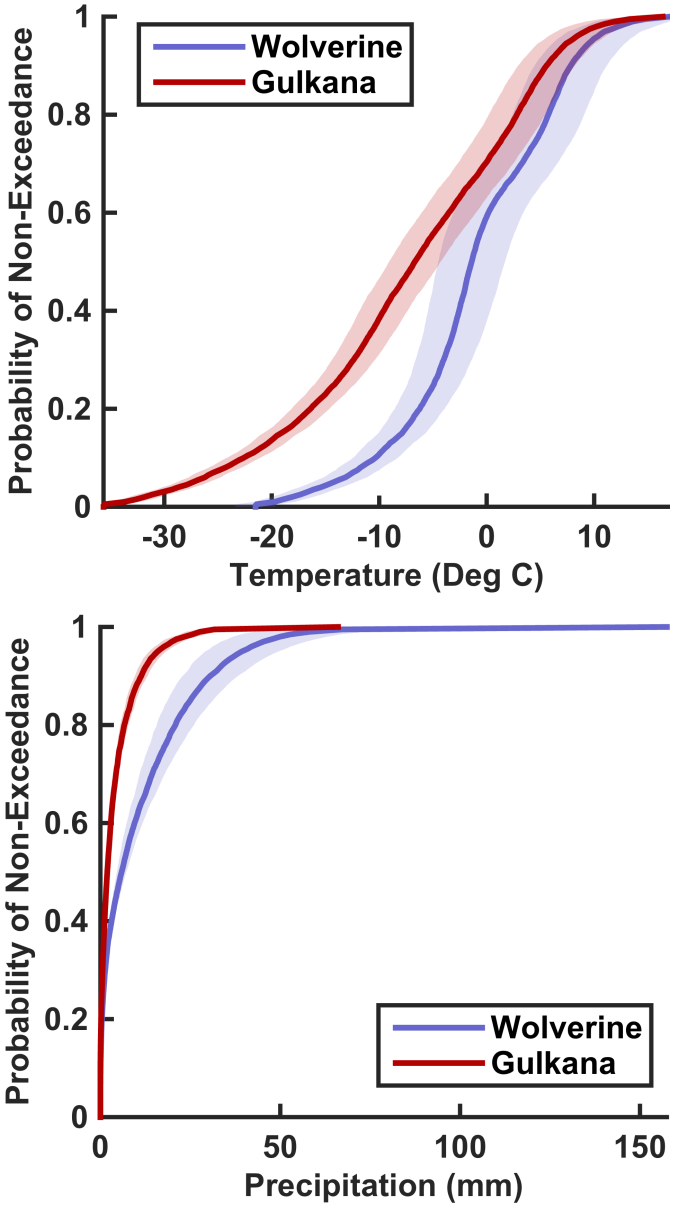

Figure 2. Cumulative probability distributions of daily precipitation and mean surface air temperature for the Gulkana and Wolverine watersheds (Fig. 1) from 1980-2009. The lines correspond to the spatially averaged precipitation and temperature values and the shaded region indicates one standard deviation in the spatial distribution of precipitation and temperature within the watershed. Climate representation is derived from the daily downscaled Climate System Forecast Reanalysis (*Saha et al.*, 2010) product discussed in Section 3.1.

common to maritime snowpacks. These differences between continental and maritime cryosphere climate affect the thermal

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

properties resulting snowpacks because the thermal conductivity of snow is primarily a function of density and water content (*Sturm et al.*, 1997; *DeWalle and Rango*, 2008).

2 The Conceptual Cryosphere Hydrology Framework (CCHF)

The CCHF models alpine and glacier systems using nine process modules (see Figure 3). While it is common for conceptual cryosphere models to combine heat and snow/ice melt into a single formulation, here they are separated into "cryosphere heat transfer" (referred to as the heat module for brevity; described in Section 2.1) and "cryosphere energy and mass" (referred to as the energy and mass module; described in Section 2.2) as this allows more versatility in the process representations. While we have programmed multiple representations for each process module in CCHF, we only vary these two modules. We choose this focus because the processes represented in the heat and energy and mass modules are the most sensitive to climatic differences (resulting from either interregional variability or long-term climatic changes).

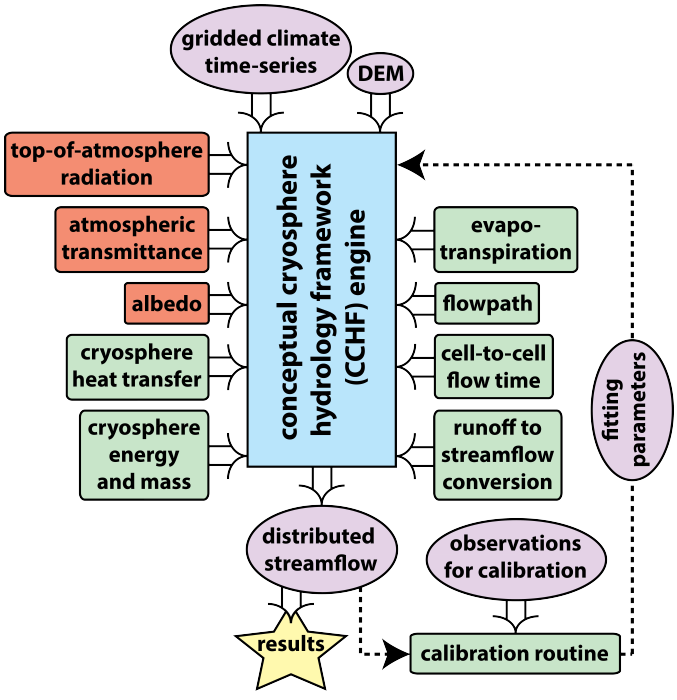

Figure 3. Categories of process modules included in the current version of the conceptual cryosphere hydrology framework (CCHF). Rounded blocks refer to process modules and ovals refer to data inputs and outputs. Blocks tinted green denote modules used for all models and blocks tinted orange denote modules only used for models that include shortwave radiation.

The process representations used for each of the seven fixed modules are provided in Table 1. The process representations for fixed modules (Table 1) are used because preliminary analysis indicated these modules work well for the conditions under which the CCHF is currently implemented. Section 2.4 discusses the modules involved in translating snow and ice melt and

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

rain into streamflow, which are represented by the four modules on the right hand side of the engine in Fig. 3. In addition to the nine process modules CCHF also includes a built-in optimization routine (Fig. 3), which we discuss in Section 2.5.

Table 1. Process representations for modules that are fixed in the present CCHF implementation. t_c, t_s, and c_a are fitting parameters (unitless), ΔT is maximum temperature minus minimum temperature ($[\Delta T] = \text{K}$), α_f is the albedo of fresh snow (unitless), α_o is the albedo of old snow (unitless), $\sum T_{max}$ is the cumulative sum of maximum temperature since last precipitation event ($[\sum T_{max}] = \text{K}$).

Module Name	Summary	Source
top-of-atmosphere radiation	function of latitude, slope, aspect, and day of year	*DeWalle and Rango*, 2008
atmospheric transmittance	$t_c \left(1 - \exp(-0.01\, t_s\, \Delta T^{2.4})\right)$	*DeWalle and Rango*, 2008
albedo	$\alpha_f - c_a \log\left(\sum T_{max}\right)$	*Pellicciotti et al.*, 2005
potential evapotranspiration	Eq. 16	*Allen et al.*, 1998; *Lu et al.*, 2005
flowpath	Sec. 2.4	manual
cell-to-cell flow time	Eq. 17	*Johnstone et al.*, 1949
runoff to streamflow conversion	Sec. 2.4	*Moore et al.*, 2012

2.1 Cryosphere Heat Transfer

Since the CCHF separates heat transfer from the representation of energy and mass, we use units of Watts per meter squared (W/m^2) for all heat module representations. The three existing heat modules we implement are SDI (Eq. 1), ETI(H) (Eq. 2), and ETI(P) (Eq. 3). Reformulating the lefthand side of each of these formulations to W/m^2 changes the units of the coefficients on the right hand side of the equations and the optimal range of fitting parameter values relative to the references on each of these formulations.

A novel conceptual heat representation implemented here is the Longwave, Shortwave, and Temperature (LST) formulation. The LST representation is similar to ETI(P) except that LST includes a longwave radiation balance term and LST does not have a fitting parameter to scale shortwave radiation. Net longwave radiation is calculated as

$$L_{\text{net}} = \sigma \left(\varepsilon_a T_{a,K}^4 - T_{s,K}^4\right) \tag{5}$$

where σ is the Stefan-Boltzmann constant ($\sigma = 5.67 \times 10^{-8}$ W m^{-2} K^{-4}), ε_a is the effective atmospheric emissivity (unitless), subscript K indicates the temperature is expressed in units of Kelvin rather than Celsius, and T_s is the snow surface temperature ($[T_s] = ^\circ\text{C}$). The LST formulation assumes ε_a to be 0.7 when there is no precipitation and 1 otherwise, which roughly translates to clear-sky and cloudy conditions (*Hock*, 2005; *Sedlar and Hock*, 2009). The LST heat formulation, H_{LST} ($[H_{\text{LST}}]$= W

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

m^{-2}), is thus

$$H_{\text{LST}} = (1-\alpha)\, I + f_m\, T_a + \sigma \left(\varepsilon_a\, T_{a,K}^4 - T_{s,K}^4\right) \tag{6}$$

where in this equation f_m has units of W °C^{-1} m^{-}2, which differs from f_m's units in Eqs. 1-3. Unlike the SDI, ETI(H), and ETI(P) formulations, the LST process representation requires modeling the temperature of snow and ice, which is discussed in Sec. 2.2.

2.2 Cryosphere Energy and Mass

We implement two process representations to relate the net heat flux to changes in the internal energy and mass of snow and ice (i.e. the 'cryosphere energy and mass' module). The first representation is a step function, denoted as Δ, which is implicit in the SDI and ETI formulations used in existing conceptual models (i.e. Eqs. 1-3). The other representation is based on the cold content (Eq. 4), and is denoted as CC. In both the Δ and CC formulations, heat is translated into melt potential, M_P ($[M_P] =$ m^{-1}; i.e. depth of melt per unit area), as

$$M_P = \frac{H\, \Delta t}{\rho_w\, L_f} \tag{7}$$

where H is the net heat flux (i.e. formulations described in Sec. 2.1; $[H] =$ W m^{-2}) and Δt is the model time step ($[\Delta t] =$ s).

In the Δ representation, all melt potential results in melt when the net heat flux is beyond a threshold value, which is described mathematically as

$$M = \begin{cases} M_P & \text{if } M_P > m_t \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where m_t is a fitting parameter ($[m_t] =$ m^{-1}). In the CC representation, CC ([CC] = m^{-1}) is calculated as

$$\text{CC}_i = \begin{cases} \text{CC}_{i-1} - c_c\, M_{P,i} & \text{if } M_{P,i} < 0 \\ \text{CC}_{i-1} - M_{P,i} & \text{otherwise} \end{cases} \tag{9}$$

where subscript i refers to the current time step, subscript "$i-1$" refers to the previous time step, and c_c is a unitless fitting parameter that provides a hysteresis in the accumulation and depletion of CC. Physically, the hysteresis relates to differences in the conduction of energy through the snowpack during accumulation (dry) versus ablation (wet) conditions. Melt, M, in the CC representation is calculated as

$$M = \begin{cases} M_P - CC & \text{if } M_P > CC \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where the negative bounds on CC is zero (i.e. CC is set to zero if it becomes negative). CC is always zero at grid locations without snowpack (i.e. CC is always zero for grid cells containing ice but without snow). While snowpack temperature cannot

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

be modeled in the Δ formulation, the CC formulation models the average internal snowpack temperature, $\bar{T}_{s,i}$, as

$$\bar{T}_{s,i} = -c_t \, \mathrm{CC} \tag{11}$$

where c_t is a fitting parameter to translate CC into internal temperature ($[c_t] = °\mathrm{C\ m^{-1}}$). Thus, the assumption is made in the CC representation that the snowpack is isothermal (i.e. $\bar{T}_{s,i} = T_{s,i}$). When liquid water is present in the snowpack, the snowpack temperature is set to 0 $°C$.The LST heat representation is the only formulation implemented here that uses Eq. 11. In the LST model, it is therefore assumed that the snowpack surface temperature equals the average internal temperature. This is a poor assumption for dry, cold snowpacks.

Both the Δ and CC energy and mass representations contain a maximum snowpack liquid content, MLC ($[\mathrm{MLC}] = \mathrm{m^{-1}}$, i.e. depth per area), represented by

$$\mathrm{MLC} = c_s \, \mathrm{SWE}_s \tag{12}$$

where c_s is a unitless fitting parameter and SWE_s is the snowpack's solid water content expressed as depth of liquid water equivalent ($[\mathrm{SWE}_s] = \mathrm{m^{-1}}$, which is depth per unit area). All liquid water in excess of the snow's liquid water holding capacity drains from the snowpack at the end of each time step. In the CC formulation, the snowpack's liquid water is refrozen when CC is greater than 0, and the CC is reduced in direct proportion to the depth of water frozen. Only water that drains from the snowpack is counted as a change in the snowpack's total water content.

2.3 Glacier Treatment

Fully representing ice processes requires accounting the the ice's internal energy. Conceptual models typically do not do this, instead assuming the the ice temperature is 0 °C when there is no snow present and the heat flux into the ice is positive (*Hock*, 1998, 1999). No energy and mass modules that include representation of ice's internal energy have been programmed in the current version of CCHF, although this is an area that should be developed in the future. Instead, CCHF assumes the glacier surface albedo is 0.4 and that the glacier's surface temperature is the minimum of the air temperature in the previous time step and 0 °C. The glacier's surface temperature only impacts net longwave heat flux calculated in the LST model.

Differences in energy required to melt ice versus snow are accounted for in the Δ and CC formulations by including a fitting parameter that scales glacier melt differently than snow melt. Thus, glacier melt is calculated as

$$M_{ice} = c_g \frac{H_{\mathrm{ice}}}{H} M_{P,ice} \tag{13}$$

where c_g is a unitless fitting parameter to scale glacier melt, H_{ice} is the heat flux for the ice surface properties, H is the equivalent heat flux for snow surface properties, and $M_{P,ice}$ is the melt potential remaining after snow melt is calculated ($[M_{P,ice}] = \mathrm{m^{-1}\ s^{-1}}$).

2.4 Melt to Streamflow

Once melt drains from the snowpack, CCHF models its transport through the landscape with the four modules on the right side of Fig. 3. Two general phases of this transport are (1) from melt to runoff and (2) through the stream network. To model phase

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

one, we implement a leaky groundwater bucket model as described in *Moore et al.* (2012). In the bucket representation, all liquid water released from snow, snlr ($[\text{snlr}] = \text{m}^{-1}$), released from ice, iclr ($[\text{iclr}] = \text{m}^{-1}$), and from rain ($[\text{rain}] = \text{m}^{-1}$) enters the groundwater bucket. The water in the bucket is described as the soil moisture, SM ($[\text{SM}] = \text{m}^{-1}$), with a total soil moisture capacity, SMC ($[\text{SMC}] = \text{m}^{-1}$; SMC is a fitting parameter). SM is updated during each time step as

$$\text{SM}_i = \begin{cases} \text{SM}_{i-1} + \text{rain}_i + \text{snlr}_i + \text{iclr}_i - r_{f,i} - \text{PET}_i & \text{if no snow or ice present} \\ \text{SM}_{i-1} + \text{rain}_i + \text{snlr}_i + \text{iclr}_i - r_{f,i} & \text{otherwise} \end{cases} \tag{14}$$

where subscript i refers to the value during the i^{th} time step, and PET is potential evapotranspiration ($[\text{PET}] = \text{m}^{-1}$; described below). Runoff in the bucket model is then calculated based on SM and SMC as

$$r_f = \begin{cases} A\left((d_r - 1)\,\text{SMC} + \text{SM}\right) & \text{if SM} \geq \text{SMC} \\ A\,d_r\,\text{SM} & \text{otherwise} \end{cases} \tag{15}$$

where r_f is the runoff ($[r_f] = \text{m}^3$), A is the grid cell's area ($[A] = \text{m}^2$), and d_r is the drain rate of soil moisture from the leaky groundwater bucket per day ($[d_r] = \text{days}^{-1}$; fitting parameter ranging from zero to one).

PET is calculated using a formulation of the Hamon Equation (*Hamon*, 1961) based on *Allen et al.* (1998) and *Lu et al.* (2005), in which

$$\text{PET} = 218.39 \times c_e \times N \left(\frac{6.108 \exp\left[\frac{17.27 T_a}{T_a + 273.15}\right]}{T_a + 273.15} \right) \tag{16}$$

where c_e is a unitless fitting parameter and N is the number of daylight hours ($[N] = \text{hrs}$). PET is set to zero when snow or ice is present.

Direction of flow between grid cells in the CCHF can be input by the user (in the form of an ESRI formatted flow direction grid) or calculated automatically using algorithms from the TopoToolbox (*Schwanghart and Kuhn*, 2010). The travel time through each cell (Δt; $[\Delta t] = \text{s}$) is then calculated using a formulation by *Johnstone et al.* (1949),

$$\Delta t = t_l\, 3600 \sqrt{\frac{\Delta l}{\sqrt{m}}} \tag{17}$$

where c_t is a unitless fitting parameter, Δl is the distance between the centroids of grid cells in the flowpath ($[\Delta l] = \text{m}$), and m is the slope between cells in the flowpath expressed as a ratio (unitless).

Two methods available in CCHF to route the streamflow are lumped and Muskingum. The primary difference between these two methods is that lumped method does not allow dispersion of the flow and the Muskingum method does (*Bedient et al.*, 2012). The results of both flow routing methods are comparable for simulating small watersheds at large time steps; however, dispersion becomes more important as the size of the watershed increases relative to the time step (*Bedient et al.*, 2012).

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

2.5 Optimization Routine

Several optimization routine options are built into the CCHF, including a genetic algorithm (*Wang*, 1991), Monte Carlo simulation (*Sawilowsky and Fahoome*, 2003), Particle Swarm Optimization (PSO; *Poli et al.*, 2007), and a hybrid algorithm developed by us. The hybrid algorithm works by implementing the Monte Carlo routine for the first several calibration generations (ranging from four to 15 generations based on the number of fitting parameters) and updating the fitting parameters in the following generations using PSO until two consecutive generations stagnate to the same fitness score (referred to here as an initial stagnation). Once an initial stagnation occurs, the hybrid algorithm uses a combination of Monte Carlo simulations and linear sensitivity analysis to alternately add diversity to the population (Monte Carlo simulation) and explore the local parameter space (linear sensitivity). If a new optimum is found, PSO is reinitiated. The process repeats until the same best fitness score is returned for 15 consecutive generations (referred to here as terminal stagnation). A population of 30 parameter sets is used in each generation of the optimization process.

Calibration in the CCHF can be conducted in one or multiple stages. If multiple stages are implemented, the user must determine which categories of fitting parameters to optimize in each stage. The two performance metrics we use are the Nash Sutcliffe Efficiency (NSE; *Nash and Sutcliffe*, 1970) score and Kling Gupta Efficiency (KGE; *Gupta et al.*, 2009) score. The NSE score is extensively used to assess hydrologic models and, as is shown in *Gupta et al.* (2009), can be decomposed into the form

$$\mathrm{NSE} = 2\alpha r - \alpha^2 - \beta^2 \tag{18}$$

where α is the ratio of the model standard deviation to the observed standard deviation, r is the linear correlation coefficient, and β is the bias normalized by the observed standard deviation. From Eq. 18 it is evident that the NSE score unequally weights standard deviation, correlation coefficient, and bias. Therefore, *Gupta et al.* (2009) propose using the KGE score, defined as

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{19}$$

Thus, the KGE metric equally weights errors in the correlation, standard deviation, and the non-dimensional bias.

A perfect KGE or NSE score is one and the negative bound on the scores is negative infinity. An NSE score of zero indicates the mean of the observation time-series are as as good of a predictor as the modeled time-series. Note, though, that a KGE score of zero does not share this interpretation. The advantage of using the KGE for calibration instead of the NSE is that the NSE is sensitive to errors in extreme values and less sensitive to errors in the overall distribution relative to the KGE metric (*Legates and McCabe Jr*, 1999).

3 Methods for Model Comparison

The objectives of the present model assessment using CCHF are to (1) demonstrate that CCHF is a useful tool for developing novel conceptual cryosphere hydrology models and (2) identify differences in accuracy and precision of existing and novel conceptual cryosphere hydrology models. We achieve these objectives by applying the CCHF to Gulkana and Wolverine

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

glaciated watersheds in Alaska (Fig. 1) for seven model formulations. Both of these watersheds are longterm US Geological Survey (USGS) study sites and exist within different climatic regimes of Alaska (Sec. 1.2). Specifically, the different climatic regimes serve as a proxy to determine how robust each model formulation will likely be when used to simulate the cryosphere response to altered climate scenarios.

The model formulations we assess are outlined in Table 2 (the fitting parameters contained in each model are displayed in Tables 9 and 10, Appendices E and F). Section 3.1 describes the inputs we use to implement each model within CCHF, Sec. 3.2 describes the observation data we use to assess model performance, and Sec. 3.3 describes how we assess each model.

Table 2. Conceptual heat (Sec. 2.1) and energy and mass (Sec. 2.2) module combinations that we assess. In all cases the seven fixed module formulations are those presented in Table 1. Note that as discussed in Section 2.1, all units of heat are expressed as W/m^2, which differ from the units in many of the reference equations.

Model Label	Heat Eq. #	Energy and Mass Eq. #
SDI_{Δ}	1	8
SDI_{cc}	1	9
$ETI(H)_{\Delta}$	2	8
$ETI(H)_{cc}$	2	9
$ETI(P)_{\Delta}$	3	8
$ETI(P)_{cc}$	3	9
LST_{cc}	6	9-11

3.1 Model Inputs

CCHF presently supports inputs of precipitation and temperature (mean, minimum, and maximum) time-series, a digital elevation model (DEM; we use the WorldClim 30 arcsecond DEM, *Hijmans*, 2011), a glacier or ice cover raster (we use the Randolph Glacier Inventory version 5, *Pfeffer et al.*, 2014), and a flow direction raster (FDR; can also be generated automatically within CCHF using algorithms from the TopoToolbox, *Schwanghart and Kuhn*, 2010). We manually delineate the watersheds and create the FDR using a combination of ArcGIS algorithms and visual analysis.

The climate time-series inputs used here are derived from the Climate Forecast System Reanalysis (CFSR), which is produced by the National Centers for Environmental Prediction (NCEP) (*Saha et al.*, 2010). CFSR is an hourly climate product available from 1979-2010 as 0.312 degree grids. *Wang et al.* (2011) and *Lader* (2014) find that CFSR represents precipitation and temperature variability as well or better than other reanalysis products such as MERRA (*Rienecker et al.*, 2011) and ERA-Interim (*Dee et al.*, 2011) for the Alaska region. CFSR is also demonstrated to reproduce total volumes of water flux through the Gulf of Alaska better than estimates derived from other reanalysis products (*Beamer et al.*, 2015).

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

We temporally aggregate the hourly 0.312 degree CFSR product to daily time steps. Precipitation and mean temperature are used in all instances, while minimum and maximum temperature are only used in model formulations that include shortwave radiation. We input the temporally aggregated 0.312 degree CFSR product into MicroMet (*Liston and Elder*, 2006a). Micromet resamples the input time-series to the spatial grid of the DEM and applies temperature lapse rates and precipitation correction factors based on elevation distribution, which vary as a function of the day of the year, creating a 30 arcsecond daily time-series. These spatially downscaled and temporally aggregated time-series are then used as inputs to CCHF.

3.2 Observation Data

The three types of observation data we use to assess model performance are USGS streamgage measurements of flowrate (referred to as flow), Moderate Resolution Imagining Spectroradiometer (MODIS; *Hall et al.*, 2006) images of SCA (referred to as SCA), and USGS stake measurements of changes in snow and ice water equivalent (referred to as stake). Both Gulkana and Wolverine are long-term benchmark glaciers maintained by the USGS. Correspondingly, the USGS has been measuring glacier and snow accumulation and ablation at three locations on Gulkana glacier from 1974 through the present (with an additional stake from 1990 through the present) and three locations on Wolverine from 1965 through the present (*Van Beusekom et al.*, 2010). Glacier stakes measure changes in depth of snow and ice at one location between two observation dates. These changes in depth are then converted to changes in mass through measuring the density of snow and assuming a density for ice. Wolverine glacier is entirely free of surface debris, while a portion of the ablation zone of Gulkana glacier contains scattered debris cover, but we model both glaciers as debris-free.

There are streamgages located slightly below the termini of Gulkana and Wolverine glaciers. The streamgages were installed in 1966 and are maintained by the USGS. The temporal coverage of the streamgages is not continuous. For example, the streamgage at Gulkana is missing measurements from October 1978 through September 1989 and the streamgage at Wolverine does not have measurements from October 1978 through September 2000. The USGS also cautions that there may be significant errors in the flows, especially at high flows, because the streamgages are located in geomorphologically active channels.

Monthly estimates of SCA are obtained from the MODIS instrument aboard the National Aeronautics and Space Administration's (NASA) Terra and Aqua satellites. MODIS images have been processed to observe many properties, but the dataset used here is MOD10CM, which is a 0.05 degree spatial and monthly temporal resolution approximation for the percent SCA of each grid cell (*Hall et al.*, 2006). While daily and 8-day MODIS images are available (MOD10C1 and MOD10C2, respectively), these datasets are less useful than the monthly product for this region due to a large portion of the images being obstructed by clouds. Individual grid cells are even obstructed in the monthly images due to prevalence of cloud cover during a given month. To overcome this, we spatially average all MODIS SCA percentages over each watershed and only use the spatially-averaged monthly SCA values for which 30% or fewer of contributing MODIS grid cells are obstructed for the given time step.

3.3 Assessment Strategy

As is seen in Fig. 2 and Fig. 5 (Appendix A), climate varies more between the Gulkana and Wolverine watersheds than it varies between decades at either of the watersheds. Due to this, and because more observation data are available for 2000-2010 than

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

for other periods (see discussion in Sec. 3.2), we calibrate each of the seven models for Gulkana and Wolverine separately for July 2000 through June 2010 and then validate each model by applying it to the opposite watershed but the same time period. Using only the period 2000-2010 allows us to use flow, SCA, and stake measurements for both calibration and validation at each site. We view the ten year calibration and validation periods as necessary to assess long-term performance of the models. Often models are assessed for shorter periods (e.g. *Singh et al.*, 2000, *Liston and Mernild*, 2012, and *Hock and Holmgren*, 2005) but their is evidence that assessing models over short periods can lead to incorrect assessments (e.g. see *Razavi and Tolson*, 2013).

We calibrate each model in two stages. In the first stage parameters relating to cryosphere modules (see Fig. 3) are optimized by comparing modeled and measured SCA and stake changes. The calibrated cryosphere parameters are then set to their optimized values during the second stage of calibration and all remaining parameters (i.e. those related to runoff generation and routing processes; see Sec. 2.4) are optimized by comparing modeled and measured flow. The KGE score is used to assess the fitness of parameter sets during calibration. Since stake observations are available at multiple points, fitness scores during the first stage of calibration are calculated at each stake location and then averaged to produce a single stake score for each parameter set. This stake score is then averaged with the SCA score to determine the overall cryosphere fitness of the parameter set.

4 Results and Discussion

Each of the models from Table 2 is calibrated and validated using the procedure outlined in Sec. 3.3. The corresponding KGE scores for each model's calibration and validation runs are summarized in Fig. 4 (full results, including NSE scores, are provided in Tables 7 and 8 in Appendices C and D, respectively). The corresponding calibrated parameter values are provided in Tables 9 and 10 (Appendices E and F). While Fig. 4 efficiently summarizes the results, trends become more apparent when model performance is aggregated in specific ways. We therefore separate this section into several subsections to highlight interesting properties of the model intercomparison. Some of the findings are that (1) which energy and mass module representation performs best depends on the watershed (Sec. 4.1), (2) overall performance varies between the two watersheds (Sec. 4.2), and (3) which model is most accurate depends on the observation variable considered but the LST_{cc} model is the most robust between regions (Sec. 4.3). Section 4.4 then summarizes general observations.

We also want to note at the outset that the choice of metric is a significant determinant of perceived model performance. Throughout, we primarily report the KGE (Eq. 19), although Tables 7 and 8 (Appendices C and D) provide model performance for the NSE (Eq. 18) as well. As explained in Sec. 2.5, whether the NSE or KGE score is greater depends on the characteristic of the errors (i.e. if the errors are primarily in the correlation, distribution, or bias). To highlight this, Table 3 displays the types of stake measurement errors for the LST_{CC} model. This analysis shows that the linear correlation and standard deviation of stake observations are reproduced well and that the largest source of error for the stake observations is the bias. This explains the better NSE scores for stake measurements, because the NSE is less sensitive to errors in the bias if the correlation coefficient is

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

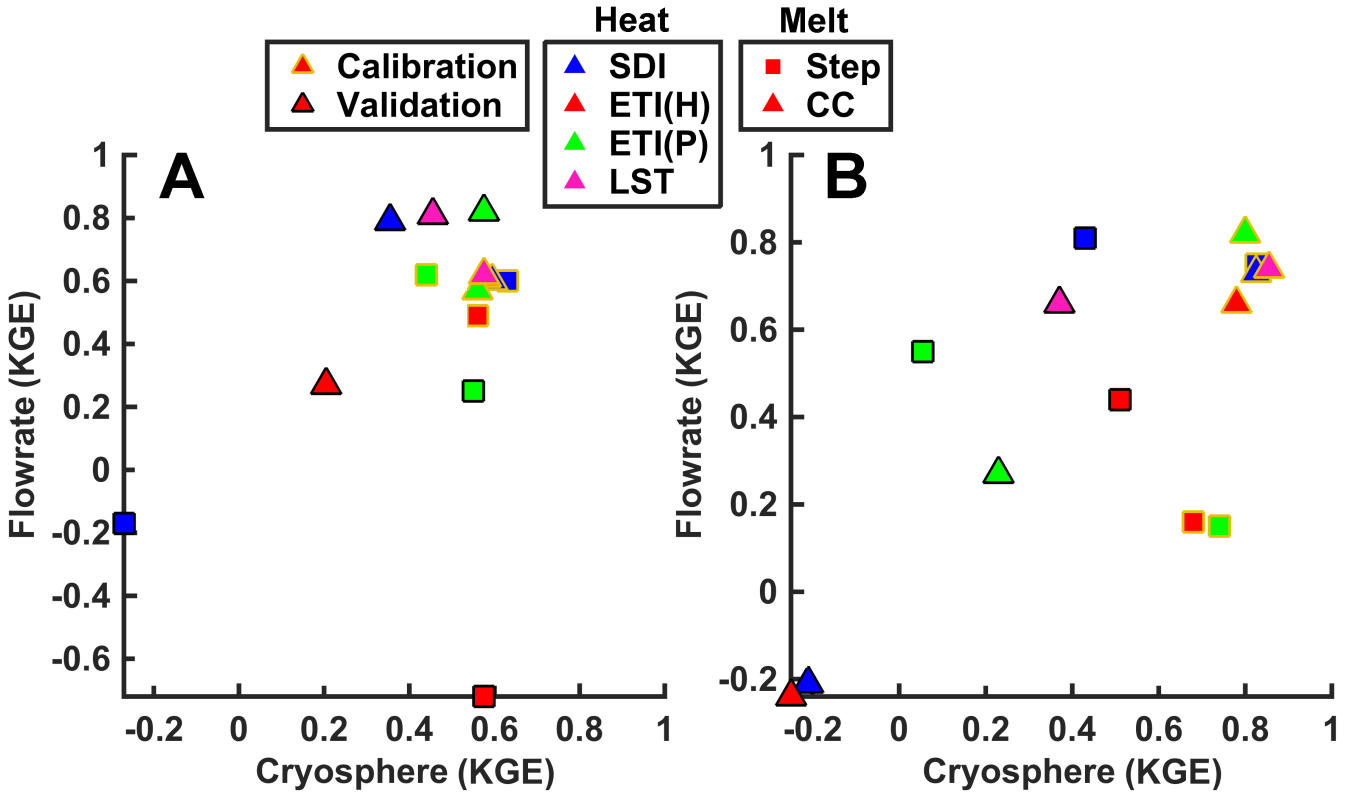

Figure 4. Modeled flow performance during validation as a function of combined cryosphere performance, described using the Kling Gupta Efficiency (KGE; Eq. 19). As described in the legends, the marker border indicates phase of assessment, face color indicates the heat module representation, and shape indicates the energy and mass module representation. In subplot A, models are calibrated for Gulkana and validated for Wolverine. In subplot B, models are calibrated for Wolverine and validated for Gulkana. The cryosphere KGE score is the average of the stake and snow-covered area KGE scores (described in Sec. 3.2).

strong and the standard deviation is reproduced well (Eq. 18). It also highlights that choice of error metric, and understanding the interpretation of each, is important in comparing results across studies.

Table 3. Types of error in reproduction of stake measurements by the LST_{CC} model during validation for Gulkana and Wolverine glaciers. The three error types correspond to the terms of the Kling Gupta Efficiency (Eq. 19), which are the error in the linear correlation coefficient (r), the ratio of the model standard deviation to the observed standard deviation (α), and the bias normalized by the observed standard deviation (β).

Error Term	Gulkana	Wolverine
$\lvert r-1\rvert$	0.12	0.06
$\lvert \alpha-1\rvert$	0.15	0.22
$\lvert \beta-1\rvert$	1.12	0.94

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Many other modeling studies implement similar conceptual models and reported "better" model performance. Better is placed in quotation marks because what constitutes better depends on the metric used and how the model is assessed. Commonly, glacier models are evaluated for only a few years (e.g. *Liston and Mernild*, 2012) or a few days (e.g. *Singh et al.*, 2000). It has been demonstrated by *Razavi and Tolson* (2013) that when calibrating and validating hydrologic models, use of short durations can lead to inaccurate assessments of model performance. In the context of identifying robust models for use in projecting the impacts of climate change, it is therefore necessary to ensure that the validation is sufficiently long and that validation is conducted for multiple climatic regimes.

4.1 Energy and Mass Module

Average performance of the two energy and mass module representations tends to vary by watershed (Table 4). On average the Δ representation performs better for Gulkana and the CC representation performs better for Wolverine. A difference between these watersheds is that Gulkana is drier and colder than Wolverine (Fig. 2 and Fig. 5 in Appendix A). Although it is beyond the present scope to establish precisely why the Δ performs better for Gulkana and CC performs better for Wolverine, it may be related to the way each module represents the timing of accumulation and melt.

The CC module assumes that the snowpack is isothermal, which impacts energy storage in all of the heat module representations and the net heat flux in the LST_{CC} model. The former point is caused by the thermal conductivity of snow, which cannot be accounted for under the isothermal assumption. The thermal conductivity is lower for less dense snowpacks, which means that less dense snowpacks will absorb energy from the surface less quickly than snowpacks with a higher density. Therefore, negative heat fluxes during the extremely cold winter months may impact the average CC of the snowpack less than as calculated by the CC energy and mass module representation, particularly for Gulkana because of its more continental cryosphere climate.

Table 4. Kling Gupta Efficiencies (KGE; Eq. 19) during validation by energy and mass module representation and watershed, averaged over heat module representation.

Model	Gulkana			Wolverine		
	Flow	SCA	Stake	Flow	SCA	Stake
Δ	0.60	0.94	-0.28	-0.21	0.73	-0.16
CC	0.12	0.37	-0.30	0.67	0.85	-0.06

One characteristic of an improved energy and mass module would be that it allows snowpack temperature to vary vertically through the snowpack. This would have two advantages over the CC representation: (1) the net surface heat fluxes could be better calculated across climatic regimes and (2) the effect of the snowpack's thermal conductivity could be accounted for in updating the snowpack's average internal energy. An approach to accomplish this would be to iteratively solve for the surface temperature by requiring that the surface energy balance be zero, as is done in *Liston and Elder* (2006b). Using such a method,

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

melt would occur based on the energy balance at the surface of the snowpack rather than as a function of the snowpack's average internal energy, as is the case in the CC representation.

4.2 Regional Differences

If validation results are aggregated over models within each region, it appears that on average the seven models reproduce streamflow better for Gulkana but that they reproduce SCA and stake observations better for Wolverine (Table 5). This is a surprising result given that snow and ice melt are significant contributors to streamflow. A few points should be noted though, which are that SCA performance does not necessarily mean the spatial distribution of snow within the watershed is modeled correctly and that averaging over models ignores intermodel differences in performance between models.

To better understand causes of interregional model performance, we also calculate the linear correlation coefficient, r value, between performance in each of the cryosphere observation types (i.e. SCA and stake) and flow, by region (Table 5). For Wolverine, the correlation between SCA and flow performance is 0.92, while for Gulkana it is 0.22; however, inspection of Fig. 6 (Appendix B) reveals that ETI(H)$_{\Delta}$ appears to be an outlier from the general trend because it has strong SCA performance but very poor streamflow performance. Removing this model, the correlation between SCA and flow for Gulkana increases to 0.71. Weaker correlations between stake and flow performance, as seen in Table 5, are expected given that glacier melt contributes a smaller portion of total annual flow than snowmelt.

Table 5. Regional differences during validation performance and correlation between the Kling Gupta Efficiency (KGE; Eq. 19) scores for each cryosphere observation type and flow. The top row is the regionally averaged KGE score. The number in parentheses is the linear correlation coefficient between the KGE scores of each the cryosphere observation type and flow. For Gulkana, the first correlation is for all models and the second excludes ETI(H)$_{\Delta}$, which appears to be an outlier (Fig. 6 in Appendix B).

Region	Flow	SCA	Stake
Gulkana	0.33	0.61	-0.29
		(0.22; 0.71)	(0.20; 0.45)
Wolverine	0.29	0.80	-0.10
		(0.92)	(0.46)

4.3 Accuracy and Precision of Performance

The primary goal of this work is to assess the robustness of conceptual cryosphere modules between regions and climatic conditions. Two aspects of robustness are overall accuracy (i.e. predictive skill for each model application) and precision (i.e. how the predictive skill varies between applications of the model across regions and climatic conditions). Table 6 provides the validation performance of each model combination averaged across regions and the difference in performance between each region. Under this analysis, the LST$_{CC}$ has the best predictive skill for flow, ETI(H)$_{\Delta}$ has the predictive skill for SCA, and ETI(P)$_{CC}$ has the predictive skill for stake measurements (Table 6).

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Differences in model performance between regions are indicative of how uniform the predictive skill is across regions and climates. In this regard, the LST_{cc} model has the most uniform predictive skill for streamflow and stake measurements, and the $ETI(H)_{\Delta}$ has slightly more uniform predictive skill for SCA (Table 6). Overall, the results indicate that the LST_{cc} model represents a significant improvement in precision of predictive skill relative to the other conceptual models assessed.

Given that the climate, cryosphere, and streamflow are physically linked, it would be desirable for one model to perform best for all observation types. Selecting a "best" overall model is subjective, depending on which metric, observation data, and time periods are used. The final column of Table 6 displays the simple mean of KGE scores for flow, SCA, and stake measurements; under this evaluation the LST_{cc} performs better than the other models (mean KGE of 0.52), with $ETI(P)_{cc}$ in second (mean KGE of 0.45), and $ETI(P)_{\Delta}$ in third (mean KGE of 0.33).

Table 6. Validation KGE scores for each model averaged over watersheds for each observation type and the average of all three observation types. Values in parentheses are absolute difference in validation KGE scores between regions. The complete and unaggregated validation results are provided in Appendix D, Table 8.

Model	Flow	SCA	Stake	Mean
SDI_{Δ}	0.32	0.74	-0.58	0.16
	(0.98)	(0.43)	(0.97)	(0.79)
SDI_{cc}	0.29	0.51	-0.37	0.14
	(1.00)	(0.66)	(0.47)	(0.71)
$ETI(H)_{\Delta}$	-0.14	0.93	0.16	0.32
	(1.16)	(0.05)	(0.18)	(0.46)
$ETI(H)_{cc}$	0.02	0.50	-0.54	-0.01
	(0.51)	(0.75)	(0.16)	(0.47)
$ETI(P)_{\Delta}$	0.40	0.85	-0.25	0.33
	(0.30)	(0.16)	(1.15)	(0.54)
$ETI(P)_{cc}$	0.55	0.60	0.21	0.45
	(0.55)	(0.46)	(0.23)	(0.41)
LST_{cc}	0.74	0.84	-0.01	0.52
	(0.15)	(0.07)	(0.10)	(0.11)

An important aspect of being robust is representing all fundamental processes well. In this regard, all of the models fail because none performs as well for stake measurements as for flow or SCA, as quantified by the KGE. There could be multiple explanations for the poor stake performance, including errors in the climate inputs, appropriateness of the model's spatial resolution, and lack of representation of the ice's internal energy. Our study does not identify which of these categories cause the poor stake performance. We do note, though, that future improvements to the energy and mass module should include representation of ice internal energy and allow polythermal snowpacks. Regardless of these possible future experiments, overall the LST_{cc} model appears to be more accurate and precise than the existing conceptual models that we implement.

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

4.4 General Remarks and Future Directions

As discussed in *Fujita and Ageta* (2000), glacier surface processes are not sufficient to describe glaciers; for example, because up to 20% of melt refreezes. Surface refreezing is allowed in the CC energy and mass module representation for snowpacks when the CC is positive, but not for glaciers. The primary differences in the conceptual models implemented here between snow and ice heat fluxes are that (1) ice has a different surface albedo and (2) ice includes a fitting parameter to allow heat fluxes to translate into melt at a different rate than for snow (see Section 2.3). None of the models contains a representation of the ice's internal energy. Even SnowModel, an energy balance cryosphere hydrology model, does not characterize the internal energy of ice (*Liston and Elder*, 2006b). Yet, conceptual and energy balance cryosphere hydrology models are often applied to assess the impacts of glacier melt on projected future streamflow (e.g. *Lutz et al.*, 2014). We therefore perceive this as a significant deficiency in current cryosphere hydrology models and an area where future work must be done.

None of the cryosphere hydrology models implemented here uses wind speed or humidity as inputs, but these variables are needed to explicitly represent the sensible and latent convective heat fluxes. The choice to exclude these inputs is made because they are less frequently observed for mountain environments, are difficult to characterize at high spatial resolutions, and are therefore not commonly used in conceptual models. For continental snow and glacier environments, such as Gulkana, sublimation is an appreciable source of energy and mass losses (*Ohmura*, 2001; *DeWalle and Rango*, 2008; *Sicart et al.*, 2008). Similarly, sensible convection is important during the times of the year when the air and snow/ice surface temperatures differ from one another the most (e.g. late summer during the glacier ablation season). Each of the conceptual models we implement accounts for sensible heat fluxes through a degree-index term, which is a significant abstraction from the actual process. In reality convective heat fluxes also depend on the wind speed and whether the atmosphere is in a stable, unstable, or neutral state (*DeWalle and Rango*, 2008). A potentially enlightening future experiment would be to develop a heat module representation that uses wind speed as an input and compare how uncertainties in characterizing wind speed impact representation of the heat balance relative to using a degree-index term. The same experiment could be conducted for humidity to directly account for sublimation.

5 Conclusions

Understanding how the cryosphere will respond to climatic changes has important water resources implications and requires implementing models that are robust across geographic domains and climatic conditions. Conceptual cryosphere hydrology models are often used to make these types of projections in mountain environments due to data paucity, yet existing conceptual models are known to be less robust than physically-based model structures. We have developed the CCHF (Fig. 3) in order to systematically intercompare the accuracy and precision of conceptual cryosphere models. We use the CCHF here to implement seven conceptual models (including existing and novel formulations; Table 2) for two glaciated watersheds in Alaska (Fig. 1). The CCHF enables us to interchange individual module representations, which provides more insight into the causes of differences in model performance compared to if we implemented standalone models. While no single model outperforms the

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

others for all categories of observations, the LST_{CC} model is the most accurate and precise between climatic conditions and geographic domains when model performance is averaged between flow, SCA, and stake observations.

Our model analysis provides some general insights and directions for future investigations. For example, we find that the Δ energy and mass module representation results in better KGE scores for models validated for Gulkana (the colder and drier of the two watersheds) and that the CC energy and mass module representation performs better under the same analysis for Wolverine (Table 4). Neither of these representations captures the impact of vertical temperature gradients within a snowpack on internal energy or heat fluxes across the boundary. Therefore, a future experiment should be to develop an energy and mass module representation that models these snowpack characteristics.

Assessed using the KGE score, all of the models better represent SCA than stake measurements (Table 6 and Fig. 6 in Appendix B). It is seen that the origin of this difference is larger bias errors in the reproduction of stake measurements (Table 4) than for the other observation types. One origin of these biases may be that the glacier's internal energy is not accounted for in any of the energy and mass module representations. This deficiency in cryosphere model formulation is consistent with other conceptual models (e.g. *Hock*, 1999, and *Pellicciotti et al.*, 2005), and even with energy balance models such as SnowModel (*Liston and Elder*, 2006b) that are used in data-sparse mountain environments. Thus, we believe an area of future work with CCHF will be to better represent glacier processes.

While the representations we assess are considered conceptual rather than physical, these model types exist on a spectrum rather than as discrete model types. We emphasize conceptual formulations in this work because this classification of processes representations is often used in data-sparse mountain regions. Of the representations we assess, we find that the more physically-based structures are more robust (i.e. LST is the most robust overall, ETI models are generally in the middle, and the SDI representation is the least robust). Therefore, interesting future work would be to incorporate more input data types, e.g. wind speed and vapor pressure, and assess the impacts of more physically-based heat flux representations on overall model accuracy and robustness. These experiments would be straightforward to implement within CCHF. Therefore, we expect CCHF will continue to be a useful tool to further our understanding of best practices for modeling the cryosphere in data-sparse mountain environments.

Acknowledgments

We would like to thank the Glumac Faculty Fellowship and the Oil Spill Recovery Institute for contributing funding to this project. We would also like to thank Anthony Arendt for his assistance in accessing the USGS stake measurements for Gulkana and Wolverine glaciers and Jordan Beamer for useful conversations about models and the Alaskan cryosphere.

Appendix A: Joint Distribution of Precipitation and Temperature

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

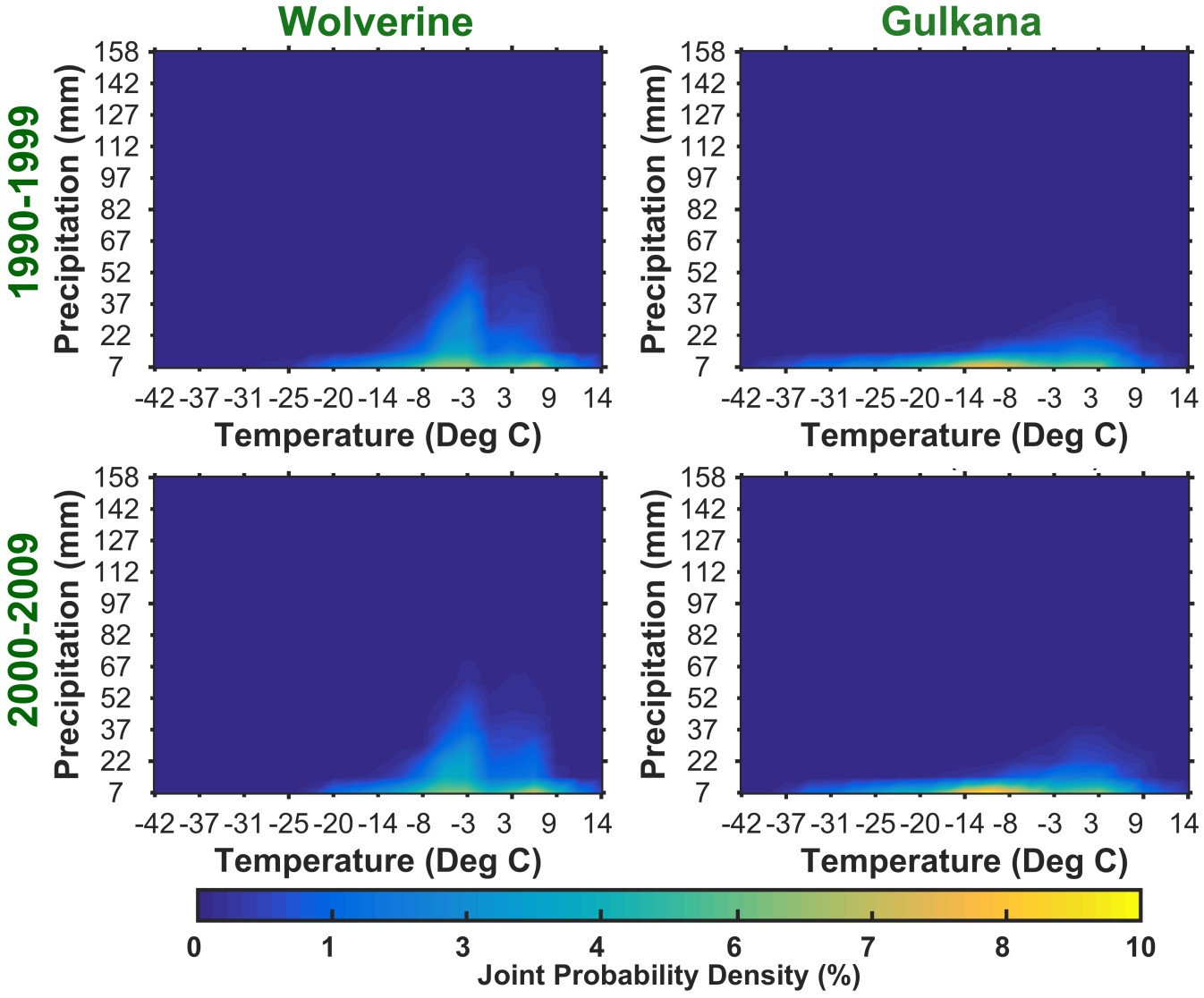

Figure 5. Joint distribution function of daily precipitation and mean surface air temperature for Gulkana and Wolverine watersheds (Fig. 1) from 1980-1989 and 1990-2009. The precipitation and temperature values are spatially averaged over each watershed. Climate representation is derived from the daily downscaled Climate System Forecast Reanalysis (*Saha et al.*, 2010) product discussed in Section 3.1.

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Appendix B: Validation Kling Gupta Efficiencies for Gulkana and Wolverine Watersheds

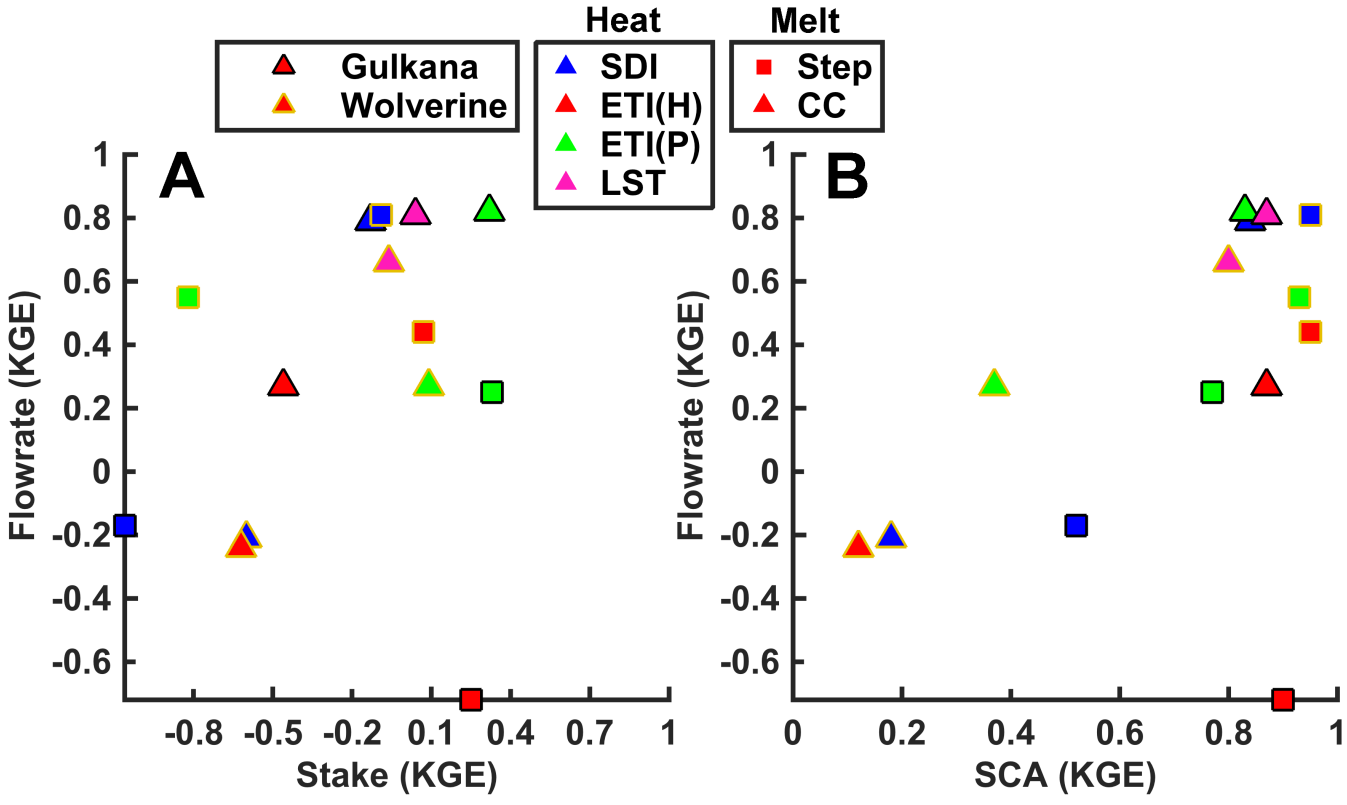

Figure 6. Modeled flow performance during validation as a function of SCA and stake performance, described using the Kling Gupta Efficiency (KGE; Eq. 19). The marker border indicates phase of assessment, face color indicates heat module representation, and shape indicates energy and mass module representation, as described in the legends. Subplot A displays KGE scores for stake measurements and flow. Subplot B displays KGE scores for SCA and flow.

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Appendix C: Calibration Performance Scores

Table 7. Model Kling Gupta Efficiencies (KGE; Eq. 19) and Nash Sutcliffe Efficiencies (NSE; Eq. 18) during calibration. The top metric is the KGE score and the bottom metric in parentheses is the NSE score. The KGE score is used to calibrate each model. All calibration runs are for July 2000 through June 2010.

Model	Gulkana			Wolverine		
	Flow	SCA	Stake	Flow	SCA	Stake
SDI_{Δ}	0.60	0.54	0.72	0.75	0.84	0.81
	(0.53)	(0.53)	(0.35)	(0.53)	(0.83)	(0.86)
SDI_{cc}	0.60	0.53	0.66	0.73	0.83	0.82
	(0.63)	(0.35)	(0.64)	(0.16)	(0.95)	(0.41)
$ETI(H)_{\Delta}$	0.49	0.90	0.22	0.16	0.95	0.41
	(-0.13)	(0.89)	(0.43)	(-1.22)	(0.90)	(0.78)
$ETI(H)_{cc}$	0.61	0.60	0.56	0.66	0.85	0.71
	(0.60)	(0.44)	(0.59)	(0.50)	(0.81)	(0.39)
$ETI(P)_{\Delta}$	0.62	0.94	-0.06	0.15	0.83	0.65
	(0.43)	(0.91)	(0.27)	(-1.31)	(0.86)	(0.69)
$ETI(P)_{cc}$	0.57	0.46	0.66	0.82	0.83	0.77
	(0.61)	(0.29)	(0.65)	(0.69)	(0.81)	(0.87)
LST_{cc}	0.62	0.58	0.57	0.74	0.86	0.85
	(0.68)	(0.43)	(0.61)	(0.70)	(0.83)	(0.85)

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Appendix D: Validation Performance Scores

Table 8. Model Kling Gupta Efficiencies (KGE; Eq. 19) and Nash Sutcliffe Efficiencies (NSE; Eq. 18) during validation. The columns labeled Gulkana are parameter sets that were calibrated for Wolverine and then validated for Gulkana, and vise versa for the Wolverine column. The top metric is the KGE score and the bottom metric in parentheses is the NSE score. The KGE score is used to calibrate each model. All validation runs are for July 2000 through June 2010.

Model	Gulkana			Wolverine		
	Flow	SCA	Stake	Flow	SCA	Stake
SDI_{Δ}	0.81	0.95	-0.09	-0.17	0.52	-1.06
	(0.60)	(0.90)	(0.37)	(0.01)	(0.46)	(-0.37)
SDI_{cc}	-0.21	0.18	-0.60	0.79	0.84	-0.13
	(-0.03)	(-0.03)	(-0.48)	(0.61)	(0.81)	(0.36)
$ETI(H)_{\Delta}$	0.44	0.95	0.07	-0.72	0.90	0.25
	(-0.23)	(0.92)	(0.48)	(-4.84)	(0.89)	(0.40)
$ETI(H)_{cc}$	-0.24	0.12	-0.62	0.27	0.87	-0.46
	(-0.06)	(-0.09)	(-0.57)	(-0.10)	(0.77)	(0.33)
$ETI(P)_{\Delta}$	0.55	0.93	-0.82	0.25	0.77	0.33
	(0.06)	(0.86)	(-0.88)	(-0.87)	(0.79)	(0.60)
$ETI(P)_{cc}$	0.57	0.46	0.66	0.82	0.83	0.32
	(0.44)	(0.19)	(0.26)	(0.66)	(0.82)	(0.54)
LST_{cc}	0.66	0.80	-0.06	0.81	0.87	0.04
	(0.70)	(0.67)	(0.38)	(0.61)	(0.83)	(0.57)

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Appendix E: Calibrated Model Parameters for Gulkana

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Table 9. All fitting parameters calibrated for Gulkana used in each of the models presented in Table 7.

	SDI_{Δ}	SDIcc	$ETI(H)_{\Delta}$	ETI(H)cc	$ETI(P)_{\Delta}$	ETI(P)cc	LST_{cc}
f_m (Eqs. 1-3 & 6)	41.02	35.83	139.12	173.39	23.48	17.39	28.56
a_i (Eqs. 2-3)	-	-	0.02	0.11	1.81	0.48	-
t_c (Table 1)	-	-	0.74	0.64	0.78	0.71	0.70
t_s (Table 1)	-	-	20.54	7.99	30.00	22.96	5.35
m_t (Eq. 8)	0.08	-	0.29	-	0.11	-	-
c_c (Eq. 9)	-	0.13	-	0.14	-	0.23	0.23
c_t (Eq. 11)	-	-	-	-	-	-	0.01
c_g (Eq. 13)	0.81	0.88	0.17	0.14	0.21	0.66	0.67
c_s (Eq. 12)	0.07	0.05	0.01	0.06	0.00	0.08	0.10
d_r (Eq. 15)	0.03	0.01	0.01	0.00	0.03	0.01	0.00
SHC (Eq. 15)	2.35	0.00	3.84	1.03	2.66	0.03	0.22
c_e (Eq. 16)	0.00	0.01	0.02	0.01	0.01	0.01	0.00
t_l (Eq. 17)	1.00	0.89	0.08	0.01	0.48	0.02	6.98

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Appendix F: Calibrated Model Parameters for Wolverine

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

Table 10. All fitting parameters calibrated for Wolverine used in each of the models presented in Table 7.

	SDI_Δ	SDIcc	$ETI(H)_\Delta$	ETI(H)cc	$ETI(P)_\Delta$	ETI(P)cc	LST_{cc}
f_m (Eqs. 1-3 & 6)	28.49	48.04	104.69	173.24	17.97	19.52	17.56
a_i (Eqs. 2-3)	-	-	0.00	0.29	4.19	0.51	-
t_c (Table 1)	-	-	0.83	0.71	0.85	0.74	0.73
t_s (Table 1)	-	-	17.79	19.63	19.49	24.02	12.04
m_t (Eq. 8)	0.02	-	0.14	-	0.20	-	-
c_c (Eq. 9)	-	0.22	-	0.32	-	0.26	0.39
c_t (Eq. 11)	-	-	-	-	-	-	0.09
c_g (Eq. 13)	0.65	0.36	0.13	0.08	0.24	0.83	0.57
c_s (Eq. 12)	0.03	0.10	0.01	0.00	0.05	0.10	0.04
d_r (Eq. 15)	0.01	0.03	0.02	0.03	0.01	0.03	0.00
SHC (Eq. 15)	0.25	0.01	0.20	0.11	3.19	0.26	0.86
c_e (Eq. 16)	0.01	0.01	2.72	0.05	0.01	0.00	0.00
t_l (Eq. 17)	0.13	0.12	9.51	0.14	0.73	0.23	0.56

The Cryosphere Discuss., doi:10.5194/tc-2016-17, 2016
Manuscript under review for journal The Cryosphere
Published: 24 February 2016

References

Allen, R. G., L. S. Pereira, D. Raes, M. Smith, and others (1998), Crop evapotranspiration - guidelines for computing crop water requirements - FAO irrigation and drainage paper 56, *FAO, Rome*, *300*(9), D05,109.

Armstrong, R. L., and B. R. Armstrong (1987), Snow and avalanche climates of the western United States: a comparison of maritime, intermountain and continental conditions, *IAHS Publication*, *162*, 281–94.

Beamer, J. J., D. F. Hill, A. Arendt, and G. Liston (2015), High-resolution modeling of coastal freshwater discharge and glacier mass balance in the Gulf of Alaska drainage, *Journal of Geophysical Research - Oceans (in preparation)*.

Bedient, P., W. Huber, and V. Baxter (2012), *Hydrology and Floodplain Analysis*, 5 edition ed., Prentice Hall, Upper Saddle River, NJ.

Bieniek, P. A., U. S. Bhatt, R. L. Thoman, H. Angeloff, J. Partain, J. Papineau, F. Fritsch, E. Holloway, J. E. Walsh, C. Daly, M. Shulski, G. Hufford, D. F. Hill, S. Calos, and R. Gens (2012), Climate divisions for Alaska based on objective methods, *Journal of Applied Meteorology and Climatology*, *51*(7), 1276–1289, doi:10.1175/JAMC-D-11-0168.1.

Braun, L. N., and M. Aellen (1990), Modelling discharge of glacierized basins assisted by direct measurements of glacier mass balance, *IAHS Publ*, *193*, 99–106.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resources Research*, *44*(12), doi:10.1029/2007WR006735.

Clark, M. P., B. Nijssen, J. D. Lundquist, D. Kavetski, D. E. Rupp, R. A. Woods, J. E. Freer, E. D. Gutmann, A. W. Wood, L. D. Brekke, J. R. Arnold, D. J. Gochis, and R. M. Rasmussen (2015a), A unified approach for process-based hydrologic modeling: 1. modeling concept, *Water Resources Research*, *51*(4), 2498–2514, doi:10.1002/2015WR017198.

Clark, M. P., B. Nijssen, J. D. Lundquist, D. Kavetski, D. E. Rupp, R. A. Woods, J. E. Freer, E. D. Gutmann, A. W. Wood, D. J. Gochis, R. M. Rasmussen, D. G. Tarboton, V. Mahat, G. N. Flerchinger, and D. G. Marks (2015b), A unified approach for process-based hydrologic modeling: 2. model implementation and case studies, *Water Resources Research*, *51*(4), 2515–2542, doi:10.1002/2015WR017200.

Cline, D. W. (1997), Snow surface energy exchanges and snowmelt at a continental, midlatitude Alpine site, *Water Resources Research*, *33*(4), 689–701.

Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris (2002), A knowledge-based approach to the statistical mapping of climate, *Climate Research*, *22*(2), 99–113.

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

DeWalle, D. R., and A. Rango (2008), *Principles of Snow Hydrology*, Cambridge University Press, Cambridge, United Kingdom.

Fujita, K., and Y. Ageta (2000), Effect of summer accumulation on glacier mass balance on the Tibetan Plateau revealed by mass-balance model, *Journal of Glaciology*, *46*(153), 244–252.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, *Journal of Hydrology*, *377*(1-2), 80–91, doi:10.1016/j.jhydrol.2009.08.003.

Hagg, W., L. N. Braun, M. Kuhn, and T. I. Nesgaard (2007), Modelling of hydrological response to climate change in glacierized Central Asian catchments, *Journal of Hydrology*, *332*(1), 40–53.

Hall, D. K., V. V. Salomonson, and G. A. Riggs (2006), MODIS/Terra snow cover monthly L3 Global 0.05Deg CMG, version 5. MOD10CM, *Tech. rep.*, NSIDC: National Snow and Ice Data Center, Boulder, Colorado USA, doi:10.5067/IPPLURB6RPCN.

Hamon, W. R. (1961), Estimating potential evapotranspiration, *Journal of the Hydraulics Division*, *87*(3), 107–120.

Hijmans, R. (2011), *GADM database of Global Administrative Areas, Version 2. University of Berkeley, CA, US, and the International Rice Research Institute, Los Baños, the Philippines*.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005), Very high resolution interpolated climate surfaces for global land areas, *International Journal of Climatology*, *25*(15), 1965–1978, doi:10.1002/joc.1276.

Hock, R. (1998), *Modelling of Glacier Melt and Discharge*, Zürcher Geographische Schriften 70, ETH Zürich, Germany.

Hock, R. (1999), A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, *Journal of Glaciology*, *45*(149), 101–111.

Hock, R. (2005), Glacier melt: a review of processes and their modelling, *Progress in physical geography*, *29*(3), 362–391.

Hock, R., and B. Holmgren (2005), A distributed surface energy-balance model for complex topography and its application to Storglaciaren, Sweden, *Journal of Glaciology*, *51*(172), 25–36.

Huss, M., M. Funk, and A. Ohmura (2009), Strong alpine glacier melt in the 1940s due to enhanced solar radiation, *Geophysical Research Letters*, *36*(23), L23,501, doi:10.1029/2009GL040789.

Johnstone, D., W. P. Cross, and others (1949), *Elements of applied hydrology*, Ronald Press Co., New York.

Jung, I., and H. Chang (2011), Assessment of future runoff trends under multiple climate change scenarios in the Willamette River Basin, Oregon, USA, *Hydrological Processes*, *25*(2), 258–277, doi:10.1002/hyp.7842.

Kokkonen, T., H. Koivusalo, T. Jakeman, and J. Norton (2006), Construction of a degree-day snow model in the light of the ten iterative steps in model development, in *Proceedings of the iEMSs Third Biennial Meeting: Summit on Environmental Modelling and Software. Environmental Modelling and Software Society, Burlington, USA*.

Kustas, W. P., A. Rango, and R. Uijlenhoet (1994), A simple energy budget algorithm for the Snowmelt Runoff Model, *Water Resources Research*, *30*(5), 1515–1527, doi:10.1029/94WR00152.

Lader, R. T. (2014), An evaluation of reanalysis products for Alaska to facilitate climate impact studies, Ph.D. thesis, University of Alaska Fairbanks, Fairbanks, Alaska.

Legates, D. R., and G. J. McCabe Jr (1999), Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, *35*(1), 233–241.

Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström (1997), Development and test of the distributed HBV-96 hydrological model, *Journal of Hydrology*, *201*(1-4), 272–288, doi:10.1016/S0022-1694(97)00041-3.

Liston, G. E., and K. Elder (2006a), A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), *Journal of Hydrometeorology*, *7*(2), 217–234.

Liston, G. E., and K. Elder (2006b), A distributed snow-evolution modeling system (SnowModel), *Journal of Hydrometeorology*, *7*(6), 1259–1276.

Liston, G. E., and S. H. Mernild (2012), Greenland freshwater runoff. part I: a runoff routing model for glaciated and nonglaciated landscapes (HydroFlow), *Journal of Climate*, *25*(17), 5997–6014, doi:10.1175/JCLI-D-11-00591.1.

Lu, J., G. Sun, S. G. McNulty, and D. M. Amatya (2005), A comparison of six potential evapotranspiration methods for regional use in the southeastern United States.

Lutz, A. F., W. W. Immerzeel, A. B. Shrestha, and M. F. P. Bierkens (2014), Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation, *Nature Climate Change*, *4*(7), 587–592, doi:10.1038/nclimate2237.

Mankin, J. S., D. Viviroli, D. Singh, A. Y. Hoekstra, and N. S. Diffenbaugh (2015), The potential for snow to supply human water demand in the present and future, *Environmental Research Letters*, *10*(11), 114,016, doi:10.1088/1748-9326/10/11/114016.

Marks, D., and A. Winstral (2001), Comparison of snow deposition, the snow cover energy balance, and snowmelt at two sites in a semiarid mountain basin, *Journal of Hydrometeorology*, *2*(3), 213–227.

Marks, D., J. Dozier, and R. E. Davis (1992), Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 1. meteorological measurements and monitoring, *Water Resources Research*, *28*(11), 3029–3042.

Martinec, J. (1975), Snowmelt-runoff model for stream flow forecasts, *Nordic hydrology*, *6*(3), 145–154.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An overview of the Global Historical Climatology Network - daily database, *Journal of Atmospheric and Oceanic Technology*, *29*(7), 897–910, doi:10.1175/JTECH-D-11-00103.1.

Moore, R. D., J. Trubilowicz, and J. Buttle (2012), Prediction of streamflow regime and annual runoff for ungauged basins using a distributed monthly water balance model, *JAWRA Journal of the American Water Resources Association*, *48*(1), 32–42, doi:10.1111/j.1752-1688.2011.00595.x.

Nash, J. V., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I: a discussion of principles, *Journal of hydrology*, *10*(3), 282–290.

Ohmura, A. (2001), Physical basis for the temperature-based melt-index method, *Journal of Applied Meteorology*, *40*(4), 753–761.

Pellicciotti, F., B. Brock, U. Strasser, P. Burlando, M. Funk, and J. Corripio (2005), An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland, *Journal of Glaciology*, *51*(175), 573–587.

Pfeffer, W. T., A. A. Arendt, A. Bliss, T. Bolch, J. G. Cogley, A. S. Gardner, J.-O. Hagen, R. Hock, G. Kaser, C. Kienholz, and others (2014), The Randolph Glacier Inventory: a globally complete inventory of glaciers, *Journal of Glaciology*, *60*(221), 537–552.

Poli, R., J. Kennedy, and T. Blackwell (2007), Particle swarm optimization, *Swarm Intelligence*, *1*(1), 33–57, doi:10.1007/s11721-007-0002-0.

Razavi, S., and B. A. Tolson (2013), An efficient framework for hydrologic model calibration on long data periods, *Water Resources Research*, *49*(12), 8418–8431, doi:10.1002/2012WR013442.

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

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, and others (2010), The NCEP climate forecast system reanalysis, *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057.

Sawilowsky, S. S., and G. Fahoome (2003), Statistics through Monte Carlo Simulation with FORTRAN, *Journal of Modern Applied Statistical Methods Inc., Michigan*.

Schwanghart, W., and N. J. Kuhn (2010), TopoToolbox: a set of Matlab functions for topographic analysis, *Environmental Modelling & Software*, *25*(6), 770–781, doi:10.1016/j.envsoft.2009.12.002.

Sedlar, J., and R. Hock (2009), Testing longwave radiation parameterizations under clear and overcast skies at Storglaciären, Sweden, *The Cryosphere*, *3*(1), 75–84.

Sicart, J. E., R. Hock, and D. Six (2008), Glacier melt, air temperature, and energy balance in different climates: the Bolivian Tropics, the French Alps, and northern Sweden, *Journal of Geophysical Research: Atmospheres*, *113*(D24), doi:10.1029/2008JD010406.

Simpson, J., G. Hufford, M. Fleming, J. Berg, and J. Ashton (2002), Long-term climate patterns in Alaskan surface temperature and precipitation and their biological consequences, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(5), 1164–1184, doi:10.1109/TGRS.2002.1010902.

Singh, P., N. Kumar, and M. Arora (2000), Degree-day factors for snow and ice for Dokriani Glacier, Garhwal Himalayas, *Journal of Hydrology*, *235*(1–2), 1–11, doi:10.1016/S0022-1694(00)00249-3.

Smith, M. B., D.-J. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong (2004), The distributed model intercomparison project (DMIP): motivation and experiment design, *Journal of Hydrology*, *298*(1), 4–26.

Smith, M. B., V. Koren, S. Reed, Z. Zhang, Y. Zhang, F. Moreda, Z. Cui, N. Mizukami, E. A. Anderson, and B. A. Cosgrove (2012), The distributed model intercomparison project phase 2: motivation and design of the Oklahoma experiments, *Journal of Hydrology*, *418*, 3–16.

Stocker, D. Q. (2013), Climate change 2013: the physical science basis, *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Summary for Policymakers, IPCC*.

Sturm, M. (2015), White water: fifty years of snow research in WRR and the outlook for the future, *Water Resources Research*, *51*(7), 4948–4965, doi:10.1002/2015WR017242.

Sturm, M., J. Holmgren, M. König, and K. Morris (1997), The thermal conductivity of seasonal snow, *Journal of Glaciology*, *43*(143), 26–41.

Ukraintseva, V. V. (2009), *Mountain weather and climate*, Taylor and Francis.

Van Beusekom, A. E., S. R. O'Neel, R. S. March, L. C. Sass, and L. H. Cox (2010), Re-analysis of Alaskan benchmark glacier mass-balance data using the index method, *US Geological Survey Scientific Investigations Report*, *5247*, 16.

Van de Wal, R. S. W., and A. J. Russell (1994), A comparison of energy balance calculations, measured ablation and meltwater runoff near Søndre Strømfjord, West Greenland, *Global and Planetary change*, *9*(1), 29–38.

Wang, J., M. Jin, D. Musgrave, and M. Ikeda (2004), A hydrological digital elevation model for freshwater discharge into the Gulf of Alaska, *Journal of Geophysical Research*, *109*, C07,009.

Wang, Q. J. (1991), The genetic algorithm and its application to calibrating conceptual rainfall-runoff models, *Water Resources Research*, *27*(9), 2467–2471, doi:10.1029/91WR01305.

Wang, W., P. Xie, S.-H. Yoo, Y. Xue, A. Kumar, and X. Wu (2011), An assessment of the surface climate in the NCEP Climate Forecast System Reanalysis, *Climate dynamics*, *37*(7-8), 1601–1620.