Evaluation and enhancement of permafrost modeling with the NASA Catchment Land Surface Model

Jing Tao[1], Rolf H. Reichle[2], Randal D. Koster[2], Barton A. Forman[3], Yuan Xue[3]

1 - Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland

2 - Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

3 - Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland

Correspondence to:

Dr. Jing Tao (JingTao@umd.edu)

Key Points

- Profile-average RMSE of simulated soil temperature versus in situ observations is reduced by using corrected local forcing and land cover
- Subsurface heat transport is mostly realistic; when not; it is improved via treatment of soil organic carbon-related thermal properties
- Mean bias and RMSE of climatological ALT between simulations and observations are significantly reduced with updated model version

Abstract

Besides soil hydrology and snow processes, the NASA Catchment Land Surface Model (CLSM) simulates soil temperature in six layers from the surface down to 13m depth. In this study, to examine CLSM's treatment of subsurface thermodynamics, a baseline simulation produced subsurface temperatures for 1980-2014 across Alaska at 9-km resolution. The results were evaluated using in situ observations from permafrost sites across Alaska. The baseline simulation was found to capture the broad features of inter- and intra-annual variations in soil temperature. Additional model experiments revealed that: (i) the representativeness of local meteorological forcing limits the model's ability to accurately reproduce soil temperature, and (ii) vegetation heterogeneity has a profound influence on subsurface thermodynamics via impacts on the snow physics and energy exchange at surface. Specifically, the profile-average RMSE for soil temperature was reduced from 2.96^{o}C to 2.10^{o}C at one site and from 2.38^{o}C to 2.25^{o}C at another by using local forcing and land cover, respectively. Moreover, accounting for the influence of soil organic carbon on the soil thermal properties in CLSM leads to further improvements in profile-average soil temperature RMSE, with reductions of 16% to 56% across the different study sites. The mean bias of climatological ALT is reduced by 36% to 89%, and the RMSE is reduced by 11% to 47%. Finally, results reveal that at some sites it may be essential to include a purely organic soil layer to obtain, in conjunction with vegetation and snow effects, a realistic "buffer zone" between the atmospheric forcing and soil thermal processes.

1. Introduction

Permafrost dynamics play a vital role in the water, energy and carbon cycles. Climate variability predominately controls the general patterns of permafrost occurrence and evolution at regional to global scales. At the local scale, many factors, including complex topography, soil type, vegetation and snow cover also strongly affect the thermal state of the subsurface. In situ permafrost measurement networks that provide near-surface and borehole temperature observations are critical for monitoring local permafrost conditions at the point scale [e.g., *Hinkel and Nelson* [2003], *Molders and Romanovsky* [2006], *Osterkamp and Romanovsky* [1999], *Romanovsky and Osterkamp* [1995, 1997], *Romanovsky et al.*[2010], *Shiklomanov et al.*[2010]]. However, in situ data are still too sparse in space and in time to allow their extensive use for monitoring permafrost at the regional scale, particularly in areas with a harsh environment and climate, such as Alaska.

Remote sensing techniques offer an alternative approach to monitoring the extent and distribution of permafrost at the regional scale. Specifically, remote sensing can detect (i) the surface expression of underground permafrost dynamics [*Farquharson et al.*, 2016; *Jones et al.*, 2011; *Panda et al.*, 2010], (ii) the freeze/thaw state based on microwave dielectric properties [*Frolking et al.*, 1999; *Kim et al.*, 2011; *Kimball et al.*, 2004; *Kimball et al.*, 2001; *Rautiainen et al.*, 2014; *Zhao et al.*, 2011] and (iii) the active layer thickness (ALT) based on measurements of surface subsidence [*Liu et al.*, 2012; *Liu et al.*, 2010]. The obvious drawback of remote sensing techniques, however, is that they cannot directly detect permafrost in the deep subsurface.

Other approaches for monitoring permafrost and/or the ALT include empirical, equilibrium and numerical modeling methods, as categorized in *Riseborough et al.* [2008]. Empirical methods estimate permafrost response to climate and environmental factors (e.g. soil properties, soil wetness, vegetation, etc.), such as geographically weighted regression methods [*Mishra and Riley*, 2014] and spatial analytic techniques based on the Stefan solution [*Nelson et al.*, 1997; *Shiklomanov and Nelson*, 2002; *Zhang et al.*, 2005], and usually require site-specific information to develop regression relationships. Equilibrium methods translate air temperature data into estimates of ground temperature and ALT [*Romanovsky and Osterkamp*, 1995; *Sazonova and Romanovsky*, 2003] and are typically suitable only for systems with limited complexity [*Jafarov et al.*, 2012].

Numerical modeling, in contrast, is not subject to the above limitations and can be an effective method to describe permafrost dynamics at regional to global scales with the unique advantage of being able to forecast the permafrost response to and feedback on climate change [*Jafarov et al.*, 2012]. However, numerical modeling requires realistic process parameterizations and accurate data to characterize the local topography, soil characteristics, land surface cover, and micro-climate [*Duguay et al.*, 2005]. With recent advances in the development of the necessary databases and improved model physics, numerical models, including Earth system models, have become increasingly useful for estimating permafrost [*Jafarov et al.*, 2012; *Riseborough et al.*, 2008]. For instance, numerical modeling studies have shown permafrost degradation in Alaska [*Jafarov et al.*, 2012; *Lawrence and Slater*, 2005]. However, more work is needed to quantify the skill of Earth system models to estimate permafrost conditions. Recent efforts to improve permafrost modeling have addressed using a deeper soil column [*Alexeev et al.*, 2007; *Lawrence* *et al.*, 2008], incorporating a surface organic layer [*Nicolsky et al.*, 2007], and accounting for the impact of soil organic carbon on the thermal and hydrologic properties of the soil [*Lawrence and Slater*, 2008]. In addition, models would benefit from an improved representation of the sub-grid variability of land surface properties such as vegetation properties and soil characteristics [*Riseborough et al.*, 2008].

In this paper, we systematically assess and improve the ability of a global land surface model (namely, the NASA Catchment Land Surface Model, or CLSM) to represent permafrost conditions in Alaska, extending through a more focused analysis the earlier and more limited evaluation of CLSM's permafrost performance included in *Stieglitz et al.* [2001]. Specifically, this work aims to (i) assess the performance of soil temperature profile estimates (and thus permafrost conditions) simulated by CLSM in Alaska, (ii) investigate the uncertainty associated with the meteorological forcing, land cover, and soil thermal parameter inputs, and (iii) improve the skill of CLSM for simulating permafrost dynamics.

2. Theoretical Background and Model Configuration

Permafrost is modeled here using CLSM [*Ducharne et al.*, 2000; *Koster et al.*, 2000], the land model component of the NASA Goddard Earth Observing System (GEOS-5) coupled Earth system model. Here, CLSM is used in an off-line (land-only) configuration. The CLSM subsurface heat transfer module uses six soil layers, each with its own prognostic heat content. For the land cover classes considered in this discussion, these six subsurface layers lie below a negligibly thin surface (skin) layer from which surface radiative and turbulent fluxes are computed. (As described by *Koster et al.* [2000], this surface layer in fact features three horizontally distinct temperatures tied to horizontally-varying hydrological regime.) The soil thickness for each subsurface layer increases with depth; the relevant depths are 0~0.1m, 0.1~0.3m, 0.3~0.7m, 0.7~1.4m, 1.4~3m, and 3~13m from top to bottom, respectively. Snow acts as a buffer that modulates the heat and water exchange between the overlying air and the underlying land surface and is simulated using a three-layer snow model that tracks the evolution of snow mass, snow depth, and snow heat content [*Stieglitz et al.*, 2001].

In the following, we outline the theoretical background of the soil heat transfer module in CLSM (section 2.1) and the current parameterization for soil thermal conductivity (section 2.2). Thereafter, we describe changes to the model parameterization that are designed to improve the simulation of permafrost (section 2.3). Finally, we discuss the model domain and ancillary forcing data (section 2.4).

2.1 Heat Transfer

Heat transfer in the subsurface is governed by the one-dimensional heat diffusion equation (Eq. 1):

$$C\frac{\partial T(z,t)}{\partial t}=\frac{\partial}{\partial z}\left(\lambda\frac{\partial T(z,t)}{\partial z}\right) \qquad \text{(Eq. 1)}$$

where C is the volumetric heat capacity ($Jm^{-3}K^{-1}$), which is equal to the sum of the specific heat capacities of the soil constituents (water, ice, soil minerals, organic matter, and air) multiplied by their respective volumetric fractions. The soil temperature at depth z and time t is denoted as

$T(z,t)$ (K), and λ is the soil thermal conductivity ($Wm^{-1}K^{-1}$), which also varies with depth and time. Using a finite-difference method, the heat diffusion equation (Eq. 1) can be discretized and approximately solved using

$$H(l,t+1) = H(l,t) + (F(l+1) - F(l))\Delta t \qquad \text{(Eq. 2)}$$

where $H(l,t)$ represents the heat content associated with soil layer l (J m^{-2}), with a zero reference value corresponding to a layer holding liquid water at exactly 0°C (so that "negative" heat contents imply the presence of ice and, potentially, subfreezing temperatures).

$H(l,t)$ is related to the temperature $T(l,t)$ and the fraction of ice in the layer, $f_{ice}(l,t)$, through consideration of the heat capacity, C, and the assumed amount of water, W, in the soil that can freeze or melt. The ice fraction is computed first:

$$f_{ice}(l,t) = 0. \quad \text{if} \quad H(l,t)/(L_s W) > 0.$$

$$f_{ice}(l,t) = 1. \quad \text{if} \quad H(l,t)/(L_s W) < -1. \qquad \text{(Eq. 3)}$$

$$f_{ice}(l,t) = -H(l,t)/(L_s W) \quad \text{otherwise.}$$

L_s here represents the latent heat of fusion. With the ice fraction known, we can compute $T(l,t)$, expressed here in degrees Celsius:

$$T(l,t) = H(l,t) / C \quad \text{if } f_{ice}(l,t) = 0$$

$$T(l,t) = (H(l,t) + L_s W) / C \quad \text{if } f_{ice}(l,t) = 1 \qquad \text{(Eq. 4)}$$

$$T(l,t) = 0 \quad \text{otherwise.}$$

The heat flux $F(l)$ due to heat diffusion along the temperature gradient between layer l-1 and l (Wm^{-2}), for use in (1), is expressed as

$$F(l) = K\frac{\Delta T}{\Delta z} = K\ \frac{T(l,t)\ - T(l-1,\mathrm{t})}{\mathrm{zc}(l) - \mathrm{zc}(l-1)} \qquad \text{(Eq. 5)}$$

where $K = \frac{[zb(l) - zc(l-1)]\lambda(l-1) + [zc(l) - zb(l)]\lambda(l)}{\mathrm{zc}(l) - \mathrm{zc}(l-1)}$ is the depth-weighted thermal conductivity (Wm^{-1}K^{-1}) between layers *l* and *l-1*, zb(l) represents the depth at the top of layer l, and zc(l) is the depth at the center of layer l.

Eq. 2 is solved using an explicit approach, that is, the soil temperatures at the current time step are determined from the heat contents (the model's prognostic variables) at the previous time step using (Eq. 3) and (Eq. 4) above. The heat flux at the uppermost soil boundary is equal to the ground heat flux, which is obtained by solving the surface energy-balance equation. A no-heat-flux boundary condition is applied at the lowest boundary (i.e., at ~13m depth). The key model parameters impacting the soil heat transfer is the thermal conductivity, which is further described in the next section.

2.2 Baseline Soil Thermal Conductivity Parameterizations

The soil thermal conductivity parameterization in CLSM is based on *Johansen* [1977] and *Farouki* [1981]. Specifically, the thermal conductivity λ of unsaturated soil is a weighted average of the saturated and dry thermal conductivities:

$$\lambda = K_e \lambda_{sat} + (1 - K_e)\lambda_{dry} \qquad \text{(Eq. 6)}$$

where K_e is the Kersten number, which is related to the degree of saturation of the soil layer [*Johansen*, 1977]. In CLSM, the soil water model component is only loosely coupled with the soil heat transfer component. The baseline CLSM version uses a constant saturation for the calculation of the thermal conductivity under unsaturated conditions, assuming that the soil water is always at 50% of saturation regardless of the modeled soil water conditions; that is, $K_e = 0.5$. Below the water table, fully saturated conditions are assumed. For the layer that contains the water table, the Kersten number is computed as $K_e = (\Delta z_1 * 0.5 + \Delta z_2)/(\Delta z_1 + \Delta z_2)$, where Δz_1 and Δz_2 are the partial layer thicknesses above and below the water table, respectively. In general, the computation of K_e is inconsistent with the modeled soil moisture conditions.

The thermal conductivity for dry soil, λ_{dry}, has the form

$$\lambda_{dry} = 0.039 \times n^{-2.2} \qquad \text{(Eq. 7)}$$

where n is the porosity, which is assumed to be 0.45 in the baseline CLSM version for the calculation of λ_{dry}. Thus, $\lambda_{dry} = 0.226$ $Wm^{-1}K^{-1}$ regardless of soil type. (Note that CLSM uses soil texture-dependent porosity values [*De Lannoy et al.*, 2014] for modeling soil moisture dynamics.) Finally, the thermal conductivity of saturated soil, λ_{sat}, is computed as

$$\lambda_{sat} = {\lambda_s}^{(1-n)} {\lambda_i}^{(n-w_u)} {\lambda_w}^{w_u} \qquad \text{(Eq. 8)}$$

where λ_w, λ_i and λ_s are the thermal conductivities for liquid water (0.57 $Wm^{-1}K^{-1}$), ice (2.2 $Wm^{-1}K^{-1}$), and soil solids (3 $Wm^{-1}K^{-1}$ in CLSM), respectively. The fractional volume of liquid water, w_u, is calculated as $w_u=n*(1.-f_{ice})$, where f_{ice} is the ice fraction.

2.3 Model Improvements

While the essential physical processes for soil heat transfer are considered in the baseline CLSM (section 2.2), three underlying assumptions potentially impair the model's ability to accurately simulate permafrost dynamics. The first assumption is the use of a constant soil water saturation of 0.5 for the calculation of the thermal conductivity under unsaturated conditions, which neglects the impact of soil water dynamics on the thermal processes. λ_{dry} and λ_s The second is the use of a constant soil water saturation of 0.5 for the calculation of the heat capacity, C. The third is the use of constant thermal conductivity values for λ_{dry} and λ_s regardless of soil mineral type and organic carbon content. Each of these issues was addressed in turn in the development of an improved treatment of subsurface heat transport.

To address the first issue, we modified CLSM to use the dynamically-varying modeled soil moisture estimates in the calculation of the thermal conductivity (Eq. 6). As a result, the updated CLSM now allows for more efficient heat transport when the soil is wetter. This modification of the code is employed in all of the simulations described in section 5.

Addressing the second issue with code modifications is not nearly as straightforward. As soon as heat capacity becomes a function of soil moisture content, energy balance calculations become significantly more complex, given that a proper energy balance requires that the energy attached to the dynamic water variable be transported with this water as it diffuses, drains, or is extracted for transpiration, all in addition to or in conjunction with energy transport through heat diffusion. Given the unusual water variables in CLSM – they are not strictly tied to soil layers, as in other LSMs, and in any case they are not coincident with the vertical temperature discretization – such energy-in-water accounting would quickly become intractable. In the face of these issues, we addressed the question of heat capacity instead with a series of five sensitivity experiments, assigning to a given experiment a non-dynamic specific heat capacity associated with one of five different water contents: w = 0., 0.25, 0.5, 0.75 and 1, where w is the soil's degree of saturation. The time series over multiple years of simulated subsurface temperatures at a representative site were found to be largely insensitive to the heat capacity employed, particularly for w ≥ 0.25 (see Figure S1 in the supplementary file). In light of this insensitivity, we retain the original assumption of w=0.5 for the calculation of the constant specific heat capacity, recognizing the potential for some error in very dry conditions (which are, in any case, relatively rare in permafrost areas).

To address the final issue above, we adopt a revised parameterization for the soil thermal properties that incorporates the impact of soil organic carbon based on *Lawrence and Slater* [2008]. In the revised parameterization, soil thermal properties are calculated as:

$$x = (1 - f_{sc})x_{mineral} + f_{sc}x_{sc} \qquad \text{(Eq. 9)}$$

where x represents a soil thermal property such as λ_s, λ_{dry}, the specific heat capacity of soil solid c_s, or the soil porosity that is used in heat transfer module. The corresponding thermal properties for mineral soil and soil carbon are denoted with $x_{mineral}$ and x_{sc}, respectively. The soil carbon fraction f_{sc} is described in more detail in section 3.2. To be consistent with *Lawrence and Slater* [2008], we further set the Kersten number to the degree of saturation ($Ke = S_r$) under frozen conditions and to K_e=log(S_r)+1 for thawed conditions (though we constrain it to lie between 0 and 1). This implies, however, that the soil porosities used for the soil thermal calculations (Eq. 7) differ from the porosities [*De Lannoy et al.*, 2014] used in the soil water module. The results with this revised CLSM version are discussed in section 5.3.

2.4 Model Domain and Ancillary Data

Although CLSM is typically used as a global model, we focus here on Alaska, where continuous, discontinuous, and sporadic permafrost conditions exist in areas ranging from the North Slope to the southern glacial, high-mountain region [*Duguay et al.*, 2005; *Zhang et al.*, 1999]. Alaska is a useful study area because suitable in situ observations are available for validation there (section 3.1). Figure 1a shows the model domain used here along with the elevation from the GEOS-5 modeling system [*Mahanama et al.*, 2015]. Figure 1b shows the 2-m air temperature climatology, calculated by averaging 35 years of data (1980-2014) from the Modern-Era Retrospective Analysis for Research and Applications-2 [MERRA-2; *Bosilovich et al.*, 2015] reanalysis. From north to south, the annual average air temperature ranges from about -10.8^{o}C to 6.4^{o}C. Figure 1c displays a map of permafrost extent in Alaska, showing four types of permafrost: continuous (90-100%), discontinuous (50- 90%), sporadic(10- 50%) and isolated patches (0 - 10%) [*Brown et al.*, 2002].

We conducted a baseline simulation at 9-km resolution for the entire domain from 1980 to 2014 using the baseline version of the CLSM. The model configuration within this system is similar to that used in the Soil Moisture Active Passive Level 4 Soil Moisture algorithm [*Reichle et al.*, 2016]. The model was forced with hourly surface meteorological forcing data from MERRA-2 [*Bosilovich et al.*, 2015; *Global Modeling and Assimilation Office (GMAO)*, 2015a, 2015b]. The precipitation forcing used here is essentially a rescaled version of the precipitation generated by the atmospheric general circulation model within the MERRA-2 system [*Reichle et al.*, 2017], with the (uncorrected) MERRA-2 precipitation rescaled to the long-term, seasonally varying climatology of the Global Precipitation Climatology Project version 2.2 (GPCP v2.2) product. (At latitudes south of 62.5°N, some information from the 0.5° degree, global Climate Prediction Center Unified gauge product is used as described in *Reichle et al.* [2017], but the impact of the gauge data is minimal for the high-latitude domain considered here.) The model was spun up, reaching a quasi-equilibrium, by looping 100 times through the one-year period from 01/01/2014 to 01/01/2015 and then once through the 35-year period from 01/01/1980 to 01/01/2015 period. Table 1 describes the land model parameters and boundary conditions used, including soil texture parameters, soil hydraulic parameters, soil depth, land cover, vegetation height, leaf area index (LAI), greenness fraction, and albedo [*Mahanama et al.*, 2015].

3. Datasets

3.1 In situ Permafrost Observations

To evaluate the simulation results and assess model performance, we used measurements from 51 active permafrost sites in Alaska [*Romanovsky et al.*, 2009] (http://permafrost.gi.alaska.edu/sites_map; see dots in Figure 1). Most of the permafrost sites are equipped with sensors that provide daily measurements of the soil temperature profile down to 0.5m~3m below the surface. The few sites that only have intermittent, deeper borehole observations down to 50m~60m are not used here. The in situ soil temperature observations were interpolated to the center of each CLSM layer using an Inverse Distance Weighting method. The aggregated daily soil temperature observations were then used for comparison with simulated, layer-based soil temperatures.

Problematic data records were screened out during a quality control review process. Simple cases include temperature values that were outside of the valid range as well as missing and null records. Moreover, we noticed some systematic errors. For instance, portions of some records exhibited an unnatural phase shift with respect to the corresponding multi-year climatology. It might be possible to use these records after correcting for the unnatural time shift, but in our work we simply excluded the affected measurements from the validation.

3.2 Soil Organic Carbon Database

We estimated vertical profiles of soil carbon fraction (f_{sc}) from two datasets that provide soil carbon content. The first dataset is the Global Gridded Surfaces of Selected Soil Characteristics product developed by the Global Soil Data Task Group of the International Geosphere-Biosphere Programme Data and Information System (IGBP-DIS) [*Carter and Scholes*, 2000; *Global Soil Data Task*, 2000; *Scholes et al.*, 1995]. The IGBP-DIS data cover the top 1.5m of the soil at 0.083° spatial resolution. The second dataset is the Northern Circumpolar Soil Carbon Database version 2 (NCSCD) [*Hugelius et al.*, 2013a; *Hugelius et al.*, 2014; *Hugelius et al.*, 2013b]. The NCSCD product is at finer resolution (0.012°) and covers the top 3m of soil providing data for the 0-0.3m, 0-1m, 1-2m and 2-3m depth ranges.

We interpolated the soil carbon content (kg m^{-2}) data to the 9-km model grid using the nearest neighbor method for both IGBP-DIS and NCSCD data. For the NCSCD data, simple aggregation of data for the 0~1m and 1~2m depth range was employed to obtain total carbon content in the top 2m. Next, we calculated the soil carbon density ρ_{sc} (kg m^{-3}). Following *Lawrence and Slater* [2008], we adopted the cumulative carbon storage profile for polar and boreal soils as identified in *Zinke et al.* [1986] to estimate vertical distribution (Vd) of soil carbon content. The soil carbon fraction for the l-th layer, $f_{sc}(l)$, was thus computed as $\rho_{sc}(l)/\rho_{sc,max}$, where ρ_{sc} is soil carbon density in the l-th layer calculated as $SCC \times Vd(l)/\Delta \mathrm{z}(l)$, SCC is the soil carbon content, and $\rho_{sc,max}$ is the maximum soil carbon density. The latter is set to the standard value for the bulk density of peat, 130kg m^{-3}[*Farouki*, 1981].

3.3 Weather Station Data

Weather station data were obtained from the Quality Controlled Local Climatological Data product, which provides hourly-to-monthly records and is available at the National Centers for Environmental Information (NCEI; http://www.ncdc.noaa.gov/orders/qclcd/). Specifically, we extracted measurements of dry bulb temperature, wet bulb temperature, dew point, relative humidity, wind speed, air pressure, and precipitation. Moreover, we downloaded and processed solar radiation measurements at weather stations from the National Solar Radiation Database at

NCEI (ftp://ftp.ncdc.noaa.gov/pub/data/nsrdb-solar/solar-only/). The weather station measurements were used to assess the MERRA-2 surface meteorological forcing data and to improve the forcing data by simple scaling methods (section 5.1).

Unfortunately, owing to the harsh environmental conditions, it is difficult to maintain weather stations in the high latitudes, particularly at high elevations, and this results in poor spatial and temporal coverage. In addition, due to the complex topography and micro-climates commonly found in Alaska, a particular weather station is often not representative of conditions within an associated 9-km grid cell. This is especially true for the interior of Alaska. Only one station, Deadhorse airport (Site ID: 70063727406), is co-located (within a distance of about 3.5 km) with a permafrost site (DH1) and could thereby be used in this study.

4. Assessment of Baseline Results

The baseline simulation was conducted using the original version of CLSM (section 2.2) for the period 1980 to 2014. Figure 2a illustrates the soil freeze/thaw variability in space and time using baseline simulated soil temperature at 8:30pm (local time) on the 16th day of every other month in 2014 as a typical example. The figure shows that for large regions the top three layers are frozen (indicated by the gray color) in late winter (February). The 4th and 5th layers continue to freeze into April whereas the top two layers are already starting to thaw in early spring. During the summer, the near-surface soil continues to thaw, and by August the top three layers are completely thawed while the 4th layer remains frozen in some parts of the North Slope. With the start of the cold season in October, the soil starts to re-freeze from the top down. Note that the

4th layer is much warmer compared to the upper layers during winter, and the re-freezing cycle in the 5th layer has an even greater time lag. The lagged freeze/thaw cycle in the different soil layers is also illustrated in Figure 2b, which shows, for each layer, the daily climatology of the frozen area in the domain. The shaded area indicates the inter-annual variability across the 35-year simulation period. The figure shows that the frozen area in the top three layers reaches zero around June. The 4th through 6th layers show much smaller seasonal variability compared with the upper layers, owing to the higher heat capacity in the deeper (thicker) layers. In the remainder of this section, we use the observations at the in situ permafrost sites (section 3.1; Figure 1a) to validate the simulated ALT (section 4.1) and soil temperature profiles (section 4.2).

4.1 Evaluation of Simulated Active Layer Thickness

Simulated ALT values were calculated for each year in the 35-year period based on (1) the model-simulated soil temperature profiles and (2) the ice content within the uppermost soil layer that is at least partially frozen. If the entire soil column remains thawed year-round, the simulated ALT is set to null (that is, permafrost-free). The spatial patterns of the 35-year minimum, mean, and maximum annual ALT in Alaska are shown in Figure 3a. Generally, the spatial permafrost distribution is consistent with the permafrost map shown in Figure 1c. Most of the continuous permafrost extent is captured by the model simulation, while some of the discontinuous and sporadic permafrost areas are not, perhaps due to model's coarse resolution. The spatial ALT pattern is also similar to that of previous studies [e.g. *Mishra and Riley*, 2014; *Sazonova and Romanovsky*, 2003] with relatively shallow ALT in the north and deeper values in the interior. Figure 3a also indicates that there is no permafrost in some southern areas of the domain (gray areas). This is consistent with the air temperature climatology (Figure 1b), which indicates annual average temperatures above -2°C. (Note that the effective annually-averaged temperature forcing is in fact slightly higher there given that the insulating properties of snow help shield the subsurface from cold winter air temperatures.) The permafrost-free areas may include patches of sporadic or isolated permafrost [*Zhang et al.*, 1999], but such patches are not resolved in the simulation owing to the relatively coarse (9-km) model resolution. Considering this, the permafrost-free area can be interpreted as indicative of having a low probability of permafrost, which is also consistent with the permafrost probability results reported by *Pastick et al.*[2015]. The temporal variations in the spatial mean air temperature and ALT (Figure 3b) are consistent for some years but show a lagged pattern (on the order of one year) for other years, depending on the magnitude of the temperature changes, which is reasonable. The figure suggests a decline in the regionally averaged ALT since 2010, but overall there is a slightly increasing trend in the regional ALT that is consistent with the increasing air temperature trend over the 35 years. The trend line of regional ALT has a positive slope suggesting an increasing rate about 0.4cm per year, and the warming rate for air temperature is about 0.02°C per year as shown in Figure 3b.

To validate the simulated ALT, multi-year average ALT values were calculated from the in situ soil temperature observations at the permafrost measurements sites. Figure 4 shows a scatter plot between the simulated and observed multi-year mean ALT values, along with the spatial distribution of the ALT values at the permafrost sites. The model clearly overestimates ALT at most sites compared to the observations, by an average of 0.36m. An outlier site IM1 has a deeper ALT in the observations (1.81m) than in the simulation (0.62m). Note that pixels that were permafrost-free in the simulation were excluded from the comparison. Thus, there are only

38 sites presented here. That is, among the 51 active permafrost sites, there are 13 sites for which the baseline simulation is permafrost free but observations show permafrost. It should be stressed that the model performances at these 13 sites are in fact the worst and that this is not reflected in the bias calculation. In the following, we carefully evaluate the modeled soil temperature results and then identify the key issues to address in our model simulations.

4.2 Evaluation of Simulated Soil Temperature Profiles

Daily estimates of the simulated soil temperature profiles were evaluated using observations from the permafrost sites (section 3.1). In addition to computing RMSE values for each layer, we also calculated a single, vertically-averaged RMSE value for each site with weights given by the layer thicknesses. This profile-average RMSE assigns more weight to the deeper (thicker) layers. The profile-average RMSE includes only layers for which measurements are available, which is rarely the case for the 6^{th} layer. This single statistic for each observation station permits a convenient, comprehensive assessment of the model's ability to capture subsurface heat transfer processes.

Generally, the baseline simulation results show fair performance at the regional scale (Figure 5a) with a spatially averaged RMSE of 3.48 ^{o}C (indicated by the horizontal red line in the figure). The performance varies from site to site with a minimum RMSE of 0.83 ^{o}C at COW and a maximum RMSE of 6.52 ^{o}C at S3-AWS. Sites within the same 9-km model grid cell (indicated by the background shading in Figure 5a) can exhibit large differences in performance. For instance, sites SL1, SL2, SL3, SL4 and UF1 are within a same model grid cell but have

RMSE values ranging from 2.29°C at SL3 to 4.49°C at SL4, demonstrating the large heterogeneity in local site conditions that cannot be captured by the model as applied here. Similarly, sites COF, COS, COT and COW have quite different RMSE values of 3.39°C, 4.00°C, 0.96°C and 0.83°C, respectively. The smallest RMSE at COW is attributed to the better simulation in the 2nd and 3rd layers compared to the other sites (Figure 5b). Note that most sites do not have RMSE values for the 5th and 6th layers due to lack of measurements.

The RMSE values of the 51 sites are mapped in Figure 5c. The figure suggests that, overall, the baseline simulation results show relatively better performance (blue and green colors) along or near the coastline and relatively worse performance in the interior of Alaska (yellow and red colors). This is possibly because the coastal areas generally have a less variable climate and, in the northern part of Alaska, less complex terrain than the interior. Coastal areas are thus better represented by the meteorological forcing data and the land model parameters from the GEOS-5 system. The greater heterogeneity in micro-climate, orographic effects, and landscape vegetation gradients in the interior region is less well described by the global-scale input data.

We selected 9 sites (as labeled in Figure 5c) for further investigation of these aspects, including a site that is close to the northern coast (DH1), three sites along the northern highway (FB1, SG2 and GL1), and five sites in the interior near Fairbanks (UF1, SL1, SL2, SL3 and SL4). The latter are located within the same 9-km model grid cell. The sites were selected primarily because of the availability of (1) soil temperature measurements in each soil layer, (2) long measurement records, and (3) local soil information. Geolocation and land surface information for the selected sites are provided in Table 2.

Our ultimate objective for investigating these 9 sites more closely is to improve the model's skill in reproducing the subsurface soil temperature profile. Specifically, DH1 is used to investigate the impact of errors in the MERRA-2 meteorological forcing data because there is a suitable weather station nearby (section 3.3). UF1 is used to study the influence of land cover type on permafrost simulation because its land cover is distinct from that of the other sites within the same 9-km model grid cell. For the remainder of the sites, including FB1, SG2, GL1, SL1, SL2, SL3 and SL4, soil survey information is available, permitting us to examine the impact on the model skill of using soil carbon information in the calculation of the soil thermal properties.

5. Towards Improving Permafrost Modeling

As mentioned in section 2.2, all of the experiments below, with the exception of the baseline experiment, use an updated model version that allows the simulated soil moisture dynamics to affect the thermal conductivity calculation (specifically, the Kersten number). Results obtained during the development of this version demonstrate that this facet of the model physics has only a marginal impact on modeled soil temperatures (not shown). We now evaluate the impact of three more important facets of the permafrost modeling problem: (1) the accuracy of the meteorological forcing (section 5.1), (2) the choice of land cover (section 5.2), and (3) the assigned soil thermal properties (sections 5.3 and 5.4).

In examining these three aspects, we essentially break down the heat transfer process into two vertical gradients [*Koven et al.*, 2013]. The first gradient (the "air to shallow soil" gradient) determines the heat transfer from the atmosphere to the shallow soil and is controlled in part by the meteorological forcing and land cover type. The second gradient (the "shallow to deep soil" gradient) is associated with heat transfer from shallow to deep soils and is controlled by the soil's thermal properties.

5.1 Meteorological Forcing

The evaluation of simulated 9-km grid cell-scale subsurface temperatures with point-scale in situ measurements is subject to scaling uncertainty. This is exacerbated by the coarse resolution of both the MERRA-2 meteorological forcing and the applied land surface parameters. Consider, for example, the five sites UF1 and SL1-4, as marked in Figure 5b. Although the UF1 and SL sites are within the same model grid cell (9-km) and thus use the same meteorological forcing in our simulations, the observed soil temperatures at these sites are markedly different – a result of some unresolved heterogeneity.

To assess the scaling problem, at least the part associated with meteorological forcing, we obtained local weather data from a weather station co-located with a permafrost site (site DH1; see section 3.3). We then filled the large temporal gaps in the station data using scaled MERRA-2 forcing fields – the original MERRA-2 variables at the grid cell containing the site were scaled with either multiplicative corrections (for specific humidity, wind speed, precipitation and solar radiation) or additive corrections (for air temperature and pressure) so that the climatological monthly means of the MERRA-2 data matched those of the station observations. We then forced the land model with the raw weather station data whenever they were available and with the scaled MERRA-2 data otherwise.

The multi-year mean seasonal cycles of the simulated subsurface soil temperatures obtained with the original MERRA-2 forcing and with the station-based forcing at DH1 are shown in Figure 6, along with observations. The figure shows that at this site, the original MERRA-2 forcing produces a reasonable simulation of subsurface temperature, capturing much of the observed seasonal cycle. The simulation results improve even further, though, when the station-based forcing fields are fed into the model (black line; see in particular the simulated-minus-observed differences shown in Figure 6b). With the original MERRA-2 forcing, the maximum errors appear in May to July due to a slightly earlier thawing time compared to observations. This problem is effectively alleviated in the simulation using the station-based forcing fields (black vs. gray in Figure 6b). The profile-average RMSE is 2.96°C for the daily soil temperature simulated using the original MERRA-2 forcing, and it reduces to 2.10°C when using the station-based forcing. As for the multi-year mean seasonal cycle, the profile-average RMSE is reduced by 60% (2.53°C vs. 0.95°C). This confirms that the forcing has a first order impact on the simulation of the subsurface temperatures. However, both simulations cannot pick up the zero curtains at the freeze up time around Nov. for the top three layers, which might be associated with some thermodynamic processes currently lacking in the model, such as the advection of heat upward or downward with the diffusion of moisture.

5.2 Land Cover

The land cover type chosen for a simulation can affect the energy (and water) partitioning at the land-atmosphere interface and can potentially have a strong impact on the transfer of heat between the air and the shallow soil. To examine this, we consider now the UF1 site near the University of Alaska, Fairbanks. When the land model is run globally (or across Alaska, as in Figure 2), the assigned vegetation class for this particular grid cell (and thus for our baseline UF1 simulation) is broadleaf deciduous tree. Site pictures and the site survey, however, indicate that the local land cover at UF1 is more like grassland (http://permafrost.gi.alaska.edu/site/uf1). Thus, we performed a new experiment at UF1 with grassland assigned as the surface type and with the associated vegetation height set to 0.6m (as standardly used in this model for grassland conditions). Aside from the aforementioned additional use of a moisture-dependent thermal conductivity, the experiment was otherwise identical to the baseline experiment.

The results from the two experiments are illustrated in Figure 7. The figure shows that modifying the land cover improves the simulation results at this site; the profile-average RMSE is reduced from 2.38^{o}C for the simulation (“Tree”) to 2.25^{o}C for the new experiment (“Grass”). The improvements are mainly seen in the 5th layer, which indirectly benefits from the better agreement between simulated snow depth for Grass and observations (see the top panel of Figure 7). The thicker snowpack generated in the “Grass” experiment acts as a stronger “thermal blanket” that slows down the release of energy from the ground during the cold season, which facilitates warmer, more accurate soil temperatures in the 5th soil layer. For example, the Grass simulation results show very good agreement with observations in the 5th layer in October of 2012, while the corresponding temperatures in the Tree experiment are about 3^{o}C colder. In May of 2013, the 5th layer temperatures simulated in the two experiments differ by up to 2.7°C, with solidly frozen soil in the Tree experiment and thawed soil (at 0.01°C) in the Grass experiment. Note that although the simulation of snow depth is more accurate in the Grass experiment, it is still underestimated in that experiment, and thus even this experiment shows earlier thawing compared to the observations. We expect, however, that further improvements could have been achieved by using local meteorological forcing fields (currently unavailable) in the simulations; as discussed in Section 5.1, simulations at DH1 demonstrated better thawing time with station-based forcing.

The change in the snowpack and the resulting changes in the subsurface temperatures in Figure 7 can be explained by the effect of vegetation height on the albedo of snow-covered areas. Because grassland is shorter than forest, less of its structure appears above the snow cover, resulting in a larger albedo for the snowpack; for forests in particular, modeled albedo in the presence of snow is significantly reduced by exposed tree branches and stems. Relative to forests, higher albedos over grassland for a given amount of snow lead to less melting and thus greater snow accumulation.

Overall, the results for UF1 illustrate the difficulty of using local, in situ measurements to evaluate model simulation results given that the large-scale parameter values assumed for the grid cell (here, values associated with forest cover) may be inconsistent with the local conditions at the measurement site. Although changing the assumed land cover to grassland led to significant improvements at UF1, subsurface temperatures there are still overestimated during summer and underestimated during winter, resulting in still-large inaccuracies in the simulated seasonal cycle. This may very well be due to inaccuracies in the MERRA-2-derived meteorological forcing. The weather station closest to this permafrost site is at the Fairbanks International Airport, about 5.5km away; the approach used above for DH1 to examine the impacts of meteorological forcing is thus not applicable here. Nevertheless, we will address in section 5.3 below how well the model works at UF1 under the assumption of a "perfect" air-to-shallow soil gradient (which would include an assumption of perfect meteorological forcing).

We now turn our attention to the other sites across Alaska. Inspection of site pictures suggests that most permafrost sites are found within grassy areas even when surrounding conditions are much different. For instance, the SL sites, which are installed in the forested area of Smith Lake near the University of Alaska, Fairbanks, are seen sitting amongst grassland patches within the forest (http://permafrost.gi.alaska.edu/site/sl4). This is reasonable given the logistics of installation and maintenance. Again, at UF1, assigning grassland rather than forest characteristics led to an improved simulation of subsurface temperatures; to see if this improvement is seen at other sites across Alaska as well, we repeated the experiment at these other sites. Figure 8a shows the profile-average RMSE from this new experiment ("Grass") minus that from the baseline simulation ("Baseline") at all of the sites. In the plot, negative values (blue colors) indicate improvement in model performance through the use of grassland parameters whereas positive values (orange and red colors) indicate degraded performance. While there is a mix of positive and negative differences, the spatial mean of the RMSE difference is negative (- 0.15^{0}C) indicating an overall improvement.

When considering the question of land cover impacts across the various in situ sites across Alaska, we should note that a comprehensive analysis of albedo effects on snow depth and of snow insulation effects on the simulation of permafrost is unfortunately limited by a lack of data, particularly snow depth and total albedo at the sites. (The availability of snow depth data at UF1 is one of the few exceptions.) Various ancillary products (e.g., albedo estimates from MODIS) may perhaps contribute information to a comprehensive study.

We now examine the consistency between improvements in simulating the aforementioned air-to-shallow soil temperature gradient and the shallow-to-deep soil temperature gradient. First, the temperature offset between the top soil layer and the overlying air, Ta0, was calculated at the monthly scale; this offset is taken to represent the temperature gradient from the air to the shallow soil. Similarly, the offset, T01, between the monthly temperatures in the 4^{th} layer (about 1 meter deep) and the top layer was computed to represent the shallow-to-deep soil gradient. We then computed the RMSE of the simulated Ta0 and T01 values against site observations for both the baseline and grassland experiments. Figure 8b shows the spatial distribution of the differences between the grassland and baseline experiments in the RMSE for Ta0, and Figure 8c shows the corresponding differences for the RMSE of T01. As before, negative values indicate improvements associated with the use of grassland parameters.

Theory suggests that improvements in Ta0 should translate to improvements in T01 – deep soil temperature variations are ultimately driven by variations in air temperature, and the deep soil cannot be simulated properly if the forcing from above is inaccurate. Similarly, degraded model performance along the air-to-shallow soil temperature gradient would presumably result in a degraded shallow-to-deep soil temperature gradient. This consistency is generally seen (for all but two sites) in Figure 8b and 8c – locations where Ta0 improves with the use of grassland conditions also show improvement in T01. The agreement supports the idea that the correct land cover type, which directly affects the shallow soil temperature, also eventually leads to improved heat transfer in the deeper soil.

5.3 Isolating Subsurface Heat Transport Processes

If the meteorological forcing and land surface parameterizations (including land cover) were perfect in our simulations, the simulation of subsurface temperatures might still be inaccurate due to a deficient parameterization of subsurface heat transport. To isolate these problems, we perform a series of experiments in which the top layer soil temperature is continually forced to agree with top layer soil temperature observations at a site (i.e., the simulated temperatures in the top layer are continually replaced with corresponding measured values). In the model, the top layer temperature is the sole boundary condition driving the evolution of the temperatures in the layers below. By prescribing the time variation of top layer temperature to observations, we effectively sidestep errors in meteorological forcing and surface parameters at a given site, allowing us to focus specifically on how well heat is transported in the subsurface.

The experiments in which the top layer temperature is prescribed are denoted “T1BC”, meaning that the top soil layer is effectively the upper boundary condition of the model. For these experiments, initial soil temperatures in the other soil layers were also prescribed to observations. The experiment was carried out at sites that have continuous long-period data records in at least the top four layers for at least three consecutive years: UF1, WD1, HV1, FB1, GL1, SG2, and SL1 through SL4. Due to similarity, results for some sites are not shown here; they can be found in the supplementary material.

The 5^{th} and 6^{th} layers required special treatment for the initialization because most sites do not provide corresponding measurements that deep. If the needed measurements were absent, these layers were initialized to values obtained from a fully spun-up T1BC simulation at that site. Note that this implies a potential source of error; spinning up the T1BC experiments over only a few recent years implies that the often warmer recent forcing temperatures (Figure 3b) are imprinted, perhaps unrealistically, on the 5^{th} and 6^{th} layers. This should be kept in mind when interpreting the T1BC results.

With a prescribed top layer temperature, the soil temperatures simulated in the layers below should be accurate if the heat transfer mechanism in the subsurface is adequately represented in the model. This is seen to be the case at UF1 as shown in the left panel of Figure 9. Other sites that show very good performance for the T1BC experiments include WD1 and HV1 (see Figures S2 and S3 in the supplementary file). Figure 9 indicates that the treatment of subsurface heat transport is not responsible for the errors in the UF1 simulation shown in Figure 7; these errors must be due to the meteorological forcing or to the treatment of the processes (including parameter values) that control the surface temperature itself. The model apparently represents well the physics of, for example, thermal conductivity and water/ice phase change in the subsurface at these sites (UF1, WD1 and HV1).

Other sites (FB1, GL1, SG2, and SL1-SL4), however, did not show the same success. As shown in right panel of Figure 9 for SL1 (and in supplementary Figures S4-S9 for the other sites), the T1BC results at these sites overestimate temperature in the warm period (June to September). Moreover, for all sites except for SL1, the summer overestimation eventually leads to an overestimation of temperature in the cold season (winter to early spring; see supplementary file). The SL1 site is in fact unusual in that its cold season subsurface temperatures in the T1BC experiment are greatly underestimated (Figure 9, right panel). For SL1, the problem is rectified in an additional experiment (T2BC) in which the temperatures of both the 1^{st} and 2^{nd} layer are prescribed to observations. With the 2^{nd} layer forced to be accurate as well, the simulated temperatures in the 3^{rd} through 5^{th} layers become realistic (black line in right panel of Figure 9; no observations are available for the 6^{th} layer.). From these results we conclude that for SL1, the treatment of subsurface heat transport in the model is adequate at and below the 3^{rd} layer, but that some aspect of the problem is poorly captured in the top and 2^{nd} layers. The sites FB1, GL1, SG2, SL2, SL3, and SL4 also appear to be deficient specifically in the top two layers, as these sites also show substantial improvement when the 1st and 2nd layers are prescribed to observations (see supplementary Figures S10-S15).

In summary, subsurface heat transfer appears accurate at a few sites but is deficient at several others, especially in the top and 2^{nd} layer. We address a possible reason in the next section.

5.4 Impacts of Organic Carbon

We hypothesize that the errors in the T1BC experiments seen in the right panel of Figure 9 for SL1 and in the supplementary material for several other sites relate to the treatment of organic carbon in the near-surface soil and its impacts on soil thermal conductivity. A rich, organic carbon content is associated with a small soil thermal conductivity, which would impede the insertion of energy into the soil during the warm season and the release of subsurface warmth to the atmosphere during the cold season. Site soil surveys indicate that all of the sites investigated in section 5.3 are organically rich, especially near the surface (Table 2). For instance, peat soil at FB1, SG2 and GL1 exists down to 15cm, 15cm and 55cm, respectively. Although there is no corresponding information available for SL2, SL3 and SL4, the soil survey indicates that at SL1, which is very close to SL2-SL4, peat soil is found down to a depth of 31cm.

Peat soil is poorly represented in the default model framework. Given the model assumptions regarding soil texture and organic carbon content, the peat soil information in the soil survey suggests that the thermal conductivites used in the default model are excessive, particularly near the surface. The improvement seen for SL1 in the T2BC experiment may even suggest the presence of a purely organic litter layer (e.g., decayed and undecayed leaves) at the site from which the observed top layer temperatures were measured.

As described in section 3.2, soil carbon fraction profiles were constructed from the IGBP-DIS and NCSCD soil data. Figure 10a illustrates the vertical profiles of soil carbon fraction at the seven sites examined here, including FB1, GL1, SG2, and SL1 through SL4. The profiles derived from the two different carbon datasets are nearly identical at the SL sites but differ significantly at the other sites, especially at SG2. Figure 10b shows the associated soil thermal properties at GL1. The impact of organic carbon content on the soil thermal properties (e.g., the thermal conductivities for soil solids λ_s and dry soil λ_{dry}, the specific heat capacity of the soil c_s, and the soil porosity) are illustrated by the differences between the original CLSM parameters and the new parameters derived from the soil organic carbon databases. With the new soil parameterization, λ_s and λ_{dry} are much smaller in the top two layers. Conversely, c_s and the porosity are much larger than the original CLSM values in the top two layers. In addition, for the new parameters the entire profile of λ_{dry} is much smaller than that of the original CLSM, whereas the porosity is much larger across all layers.

We incorporated the two different soil carbon fraction profiles into the CLSM using the soil parameterization scheme described in section 2.3. We then re-ran the T1BC experiment at FB1, GL1, SG2, and SL1-4. Results for GL1 and SL2 are shown in Figure 11. The subsurface temperatures obtained in the experiments using the organic carbon profiles (T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD) show an improved agreement with observations during warm periods (June through September) relative to the original T1BC experiment, especially for SL2. Results for sites FB1, SG2, SL3 and SL4 are similar; see supplementary Figures S16-S19. At GL1, for which the two sources of organic carbon profiles differ (see Figure 1110), use of the NCSCD information produces the more realistic subsurface temperatures, especially for the 3rd layer. This can be attributed to the larger carbon fraction in the 2nd and 3rd layers at GL1 for NCSCD, as highlighted in Figure 10.

Figure 12 summarizes the results obtained with the organic content profiles. Compared to the original T1BC results, the profile-average RMSE is reduced for T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD at all six of the study sites, with the better results often obtained with the NCSCD organic content data. The largest improvement in the profile-average RMSE is found at GL1 (about 56%) using NCSCD data. At individual soil layers, improvements are as high as 70% (Layer 3 at SL2, again using NCSCD data).

The behavior at site SL1 is anomalous and merits further discussion. As shown in Figure 12g, both T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD yielded larger profile-average RMSE values than T1BC (i.e., model results were degraded in an aggregate sense) despite considerable improvements during the warm period (see supplementary Figure S20) and a reduction of RMSE for the 2^{nd} and 3^{rd} layers. Nevertheless, both the T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD simulations still cannot capture the large contrast between the soil temperatures in the top and 2^{nd} layers. Furthermore, neither T1BC_OrgC_IGBP nor T1BC_OrgC_NCSCD correct the aforementioned underestimation problem at SL1 during the cold season. Moreover, when the T2BC experiment is performed (i.e., when both the top and 2^{nd} layer temperatures are prescribed to observations), the use of either the IGBP-DIS or NCSCD data still increases slightly the profile-average RMSE relative to the original T2BC experiment (Figure 12h). We can only speculate about this behavior. It is possible, for example, that relative to the cumulative carbon storage profile used to approximate the vertical distribution of carbon content at all sites, the soil carbon content at SL1 is more concentrated in the top two soil layers and much less so in the 3^{rd} and 4^{th} layers. Alternatively, the top two layers might be purely organic layers (a.k.a. litter layers) rather than the assumed composite of mineral soil and organic carbon; this particular explanation is consistent with our analysis in section 5.3.

Comparison of RMSEs for annual ALT from the different experiments reveals that simulated ALTs improve at six out of the seven test sites when soil carbon impacts are included, as shown in Figure 13 (green vs. cyan and magenta bars for simulations with MERRA-2 forcing, and blue vs. gray and black bars for simulations with prescribed top soil temperature). That is, by incorporating the thermal impacts of soil carbon into the model, simulated ALT is generally improved regardless of the quality of the forcing fields. In addition, despite the larger profile-average RMSE of soil temperature from T1BC compared to the two T1BC simulations incorporating organic carbon at SL1 as discussed above, the annual ALT at this site from baseline and T1BC simulations are significantly improved after incorporating soil carbon impacts. The only exception is SL3, which shows larger RMSE of annual ALT from T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD compared to T1BC. Nevertheless, all seven sites the simulations with MERRA-2 forcing (which is available everywhere and thus suitable for global simulations) demonstrate improved ALT by incorporating soil carbon impacts (cyan and magenta vs. green bars). One thing we should stress again is that for these sites a permafrost-free simulation is an error that cannot be quantified in terms of an RMSE of ALT; any simulation at these sites that has a meaningful ALT (e.g. M2_OrgC_IGBP and M2_OrgC_NCSCD at SLx sites) is a fundamental, if non-quantifiable, improvement over a permafrost-free simulation (e.g. Baseline simulation at SLx sites).

Figure 13, by the way, also shows that with the original carbon profile, the T1BC simulation tends to produce, as expected, more accurate ALT than the baseline simulation (dark blue versus green bars). We can only speculate on why the MERRA-2 versus T1BC ALT results are relatively mixed for the improved carbon cases (e.g., magenta versus black bars); perhaps it has to do with the aforementioned limitation regarding the spin-up of the 5th and 6th layers in the T1BC experiment.

Overall, the anomalous results at SL1 and SL3 aside, Figure 11, Figure 12 and 13 support our hypothesis regarding the importance of properly treating the impacts of organic carbon content on soil thermal properties and thereby on subsurface heat transfer – our simulations generally improve with a more careful treatment of organic carbon. The results indicate that the vertical profile of fractional organic matter within the soil composite should be specified realistically, as should the existence of any layers of organic matter sitting on top of the soil layers. A more realistic thermal “buffer zone” should indeed consider both snow and organic layers at some sites.

We now compare multi-year means of estimated ALT from the three simulations with MERRA-2 forcing (i.e., Baseline, M2_OrgC_IGBP and M2_OrgC_NCSCD) with the observed ALT at all sites across Alaska. The results are shown in Figure 14. Figure 14b shows that the RMSE of multi-year averaged ALT is reduced by 11% and 47% for the simulations using IGBP (0.49m vs. 0.55m) and NCSCD (0.29m vs. 0.55m) carbon data, respectively, compared to the baseline simulation. The overall bias values provided in Figure 14c reveal that the M2_OrgC_IGBP simulation still overestimates regional ALT but nevertheless shows a 36% improvement (0.23m vs. 0.36m) over the baseline, while the M2_OrgC_NCSCD simulation shows a very small negative bias (-0.04m, reduced by 89% compared to 0.36m in terms of absolute bias) in regional ALT, indicating a significant improvement.

6. Summary and Discussion

In this study we used the NASA Catchment land surface model to study permafrost conditions in Alaska. We first conducted a regional simulation using the current (baseline) model version and investigated the general pattern and evolution of the simulated permafrost dynamics across Alaska. The modeled ALT shows a large spatial and temporal variability that is consistent with the regional air temperature climatology (Figures 2, 3). However, the modeled ALT is overestimated by ~0.43m on average when compared against in situ observations from 38 permafrost measurement sites (Figure 4). The simulated soil temperature profiles have a spatially-averaged, profile-average RMSE of 3.48°C versus the in situ measurements (Figure 5).

Next, we investigated the soil temperature simulation errors along two vertical temperature gradients, the "air-to-shallow soil" gradient and the "shallow-to-deep soil" gradient. An accurate simulation of the first gradient is a prerequisite for the successful simulation of the subsurface temperature profile. Following this paradigm, we addressed two factors that affect the air-to-shallow soil gradient: (i) the quality of the forcing data and (ii) the land cover representation. Finally, we examined the performance of simulated subsurface heat transfer in isolation (i.e., we focused on the shallow-to-deep soil gradient) by prescribing the temperature in the surface soil layer.

In the context of our experiments, errors in the model forcing data have two potential sources: (i) inaccuracies in the GEOS-5 atmospheric modeling and assimilation system used to generate the forcing, and (ii) representativeness error, given the relatively coarse (0.5 degree) resolution of the GEOS-5 system and the point scale of the permafrost measurement sites. We addressed both error sources simultaneously by forcing the model at the DH1 site with measurements from a nearby meteorological station. The profile-average RMSE of simulated subsurface temperature at the DH1 site was thereby decreased from 2.96°C to 2.10°C, indicating that, as might be expected, meteorological forcing fields that better reflect the local conditions at a local site produce simulated soil temperature profiles that better agree with observations there.

Likewise, the model's land cover parameterization may be inaccurate, or the site-specific land cover conditions may not be representative of the grid-cell scale average conditions. In situ measurement sites are usually in more accessible, grassy areas (where snow can build up more easily), whereas larger-scale land cover in the areas studied is more typically forest or shrubs. Our results demonstrate that using grassland parameters rather than the default, grid-average land cover parameters produces soil temperature profiles that better agree with the observations. At the UF1 site, the profile-average RMSE in this experiment decreased from 2.38°C to 2.25°C.

Finally, we demonstrated that the baseline version of the CLSM can sometimes simulate subsurface thermal dynamics with high accuracy if the top layer temperature is simulated correctly – model simulations that prescribed the surface soil temperature (T1BC) showed success in simulating temperature in the subsurface at a number of sites (UF1, WD1 and HV1). However, at other sites, the T1BC results overestimated the soil temperature, especially during warm periods. For these other sites, the temperatures in both the top and 2nd layers needed to be prescribed to observations (the T2BC experiments) to produce accurate temperatures in the layers below. Overall, the T1BC and T2BC experiments suggest that, while CLSM's treatment of subsurface heat transport below the 2nd layer is accurate, at several sites the soil heat transfer properties in the top two layers of the baseline model are deficient.

This result led to an examination of the impacts of organic matter, which to date had not been properly considered in the CLSM representation of soil thermal processes. We conducted additional simulations that explicitly included the impact of soil carbon on soil thermal processes using the soil carbon parameterizations of *Lawrence and Slater* [2008]. These simulations utilized carbon data from two data sources (IGBP-DIS and NCSCD) and were run in the T1BC configuration, i.e., with top layer temperatures prescribed to observations. The results show that the more careful treatment of soil organic carbon led to greatly improved simulation results at sites with organic-rich soils. The profile-average RMSE for T1BC_OrgC_NCSCD was reduced by as much as 56% (at GL1) when compared to the original T1BC experiment, and indeed, an RMSE reduction was seen at all of the sites considered in this experiment except for SL1. At SL1, we speculate that the explicit modeling of a strictly organic layer (e.g., composed of leaf litter) may be needed to provide a more effective buffer zone between the air temperature and the deeper soil.

Simulations with the updated model version driven by MERRA-2 forcing also demonstrated improvements in ALT at the site scale, showing reduced RMSE of annual ALT compared to baseline results. At the regional scale (considering all sites across Alaska), our simulations show reduced RMSE of multi-year averaged ALT compared to the baseline results (by 47%) when NCSCD carbon information is used, along with a very small regional bias (-0.04m). Note that while our RMSE of ALT using NCSCD carbon information (0.29m) is somewhat higher than that found in a similar study by [*Jafarov et al.*, 2012] (0.08m), our model results (unlike theirs) did not benefit from calibration; also, our mean ALT bias (-0.04m) is very close to their value of -0.03m.

Overall, enhanced treatments of meteorological forcing, land cover type, and organic carbon-related soil thermal properties substantially improved CLSM's ability to simulate realistic subsurface temperatures. Progress toward an effective, large-scale representation of subsurface thermodynamics, however, was nevertheless hampered here by the local-scale character of the in situ measurements and, in any case, by the limited number of measurement sites. Looking ahead, it should be possible to continue model development on a regional, rather than local, scale using radar retrievals of ALT from the Airborne Microwave Observatory of Subcanopy and Subsurface (AirMOSS) instrument [*Chen et al.*, 2016].

Another issue that has not been addressed fully here but is worth investigating further is the impact of a purely organic layer on subsurface permafrost. Such an organic layer not only has unique thermal properties but also affects soil hydrologic processes by slowing down bare soil evaporation from the ground surface, reducing vegetation transpiration [*Luthin and Guymon*, 1974], altering downslope runoff pathways, and thus significantly affecting soil moisture underneath [*Hinzman et al.*, 1991], which can result in a dramatically different permafrost response. Some key parameters associated with an organic layer can possibly be characterized at the regional scale based on radar remote sensing, such as forthcoming organic layer thickness retrievals from the AirMOSS project (personal communication with with Mahta Moghaddam and Richard Chen). Once available, such radar retrievals should make it is possible for us to improve further the simulation of permafrost at the regional scale.

Acknowledgments

Funding for this work was provided by the NASA Interdisciplinary Science program. We are grateful for valuable discussions with Richard Chen, John Kimball, Mahta Moghaddam, and Yonghong Yi. We thank the two anonymous reviewers for their helpful comments. We acknowledge the University of Maryland supercomputing resources (http://www.it.umd.edu/hpcc) made available for conducting the research reported in this paper. Soil temperature observations used in this study are available from the Permafrost Laboratory at University of Alaska Fairbanks (http://permafrost.gi.alaska.edu/sites_map). The IGBP-DIS soil carbon data are available from the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (https://doi.org/10.3334/ORNLDAAC/569), and the NCSCD dataset is available from the Bolin Centre for Climate Research (http://bolin.su.se/data/ncscd/). The weather station data are available at the National Centers for Environmental Information (NCEI; http://www.ncdc.noaa.gov/orders/qclcd/). The baseline and revised simulation results are available at http://hdl.handle.net/1903/20168.

References

Alexeev, V. A., D. J. Nicolsky, V. E. Romanovsky, and D. M. Lawrence (2007), An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost, *Geophys Res Lett*, *34*(9), doi:10.1029/2007gl029536.

Baret, F., M. Weiss, R. Lacaze, F. Camacho, H. Makhmara, P. Pacholcyzk, and B. Smets (2013), GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production, *Remote Sensing of Environment*, *137*, 299-309, doi:10.1016/j.rse.2012.12.027.

Bosilovich, M. G., et al. (2015), MERRA-2: Initial Evaluation of the Climate*Rep.*, NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2015-104606, Vol. 43, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA.

Brown, J., O. Ferrians, J. A. Heginbottom, and E. Melnikov (2002), Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2. [Permafrost Extent], NSIDC: National Snow and Ice Data Center. Boulder, Colorado USA.

Camacho, F., J. Cemicharo, R. Lacaze, F. Baret, and M. Weiss (2013), GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products, *Remote Sensing of Environment*, *137*, 310-329, doi:10.1016/j.rse.2013.02.030.

Carter, A., and R. Scholes (2000), SoilData v2. 0: generating a global database of soil properties, *Environmentek CSIR*, *Pretoria*, *South Africa*.

Chen, R. H., A. Tabatabaeenejad, and M. Moghaddam (2016), A time-series active layer thickness retrieval algorithm using P-and L-band SAR observations, paper presented at Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International, IEEE.

De Lannoy, G. J. M., R. D. Koster, R. H. Reichle, S. P. P. Mahanama, and Q. Liu (2014), An updated treatment of soil texture and associated hydraulic properties in a global land modeling system, *J Adv Model Earth Sy*, *6*(4), 957-979, doi:10.1002/2014ms000330.

Dirmeyer, P., X. Gao, and T. Oki (2002), The second global soil wetness project (gswp-2), *International GEWEX Project Office Publication*, *37*, 75.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model 2. Parameter estimation and model demonstration, *Journal of Geophysical Research-Atmospheres*, *105*(D20), 24823-24838, doi:Doi 10.1029/2000jd900328.

Duguay, C. R., T. Zhang, D. W. Leverington, and V. E. Romanovsky (2005), Satellite remote sensing of permafrost and seasonally frozen ground, *Remote Sensing in Northern Hydrology: Measuring Environmental Change*, 91-118.

Farouki, O. T. (1981), The Thermal-Properties of Soils in Cold Regions, *Cold Reg Sci Technol*, *5*(1), 67-75, doi:Doi 10.1016/0165-232x(81)90041-0.

Farquharson, L., D. H. Mann, G. Grosse, B. M. Jones, and V. E. Romanovsky (2016), Spatial distribution of thermokarst terrain in Arctic Alaska, *Geomorphology*, *273*, 116-133.

Frolking, S., K. McDonald, J. Kimball, S. Running, and R. Zimmermann (1999), Using the space-borne NASA scatterometer (NSCAT) to determine the frozen and thawed seasons, *Journal of Geophysical Research: Atmospheres*.

Global Modeling and Assimilation Office (GMAO) (2015a), MERRA-2 inst1_2d_lfo_Nx: 2d,1-Hourly,Instantaneous,Single-Level,Assimilation,Land Surface Forcings V5.12.4, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed [03/13/2017] Greenbelt, MD, USA, doi:10.5067/RCMZA6TL70BG.

Global Modeling and Assimilation Office (GMAO) (2015b), MERRA-2 tavg1_2d_lfo_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Land Surface Forcings V5.12.4, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed [03/13/2017] Greenbelt, MD, USA, doi:10.5067/L0T5GEG1NYFA.

Global Soil Data Task (2000), Global gridded surfaces of selected soil characteristics (IGBP-DIS), *International Geosphere–Biosphere Programme—Data and Information Services. Available online [http://www.daac.ornl.gov/] from the ORNL Distributed Active Archive Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.*

Hinkel, K., and F. Nelson (2003), Spatial and temporal patterns of active layer thickness at Circumpolar Active Layer Monitoring (CALM) sites in northern Alaska, 1995–2000, *Journal of Geophysical Research: Atmospheres*, *108*(D2).

Hinzman, L. D., D. L. Kane, R. E. Gieck, and K. R. Everett (1991), Hydrologic and Thermal-Properties of the Active Layer in the Alaskan Arctic, *Cold Reg Sci Technol*, *19*(2), 95-110.

Hugelius, G., et al. (2013a), A new data set for estimating organic carbon storage to 3m depth in soils of the northern circumpolar permafrost region, *Earth Syst Sci Data*, *5*(2), 393-402, doi:10.5194/essd-5-393-2013.

Hugelius, G., et al. (2014), Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, *Biogeosciences*, *11*(23), 6573-6593, doi:10.5194/bg-11-6573-2014.

Hugelius, G., C. Tarnocai, G. Broll, J. G. Canadell, P. Kuhry, and D. K. Swanson (2013b), The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, *Earth Syst Sci Data*, *5*(1), 3-13, doi:10.5194/essd-5-3-2013.

Jafarov, E. E., S. S. Marchenko, and V. E. Romanovsky (2012), Numerical modeling of permafrost dynamics in Alaska using a high spatial resolution dataset, *Cryosphere*, *6*(3), 613-624, doi:10.5194/tc-6-613-2012.

Johansen, O. (1977), Thermal conductivity of soils*Rep.*, DTIC Document, No. CRREL-TL-637. COLD REGIONS RESEARCH AND ENGINEERING LAB HANOVER NH.

Jones, B. M., G. Grosse, C. D. Arp, M. C. Jones, K. M. W. Anthony, and V. E. Romanovsky (2011), Modern thermokarst lake dynamics in the continuous permafrost zone, northern Seward Peninsula, Alaska, *J Geophys Res-Biogeo*, *116*, doi:10.1029/2011jg001666.

Kim, Y., J. S. Kimball, K. C. McDonald, and J. Glassy (2011), Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing, *IEEE Transactions on Geoscience and Remote Sensing*, *49*(3), 949-960.

Kimball, J., K. McDonald, S. Frolking, and S. Running (2004), Radar remote sensing of the spring thaw transition across a boreal landscape, *Remote Sensing of Environment*, *89*(2), 163-175.

Kimball, J., K. McDonald, A. Keyser, S. Frolking, and S. Running (2001), Application of the NASA scatterometer (NSCAT) for determining the daily frozen and nonfrozen landscape of Alaska, *Remote Sensing of Environment*, *75*(1), 113-126.

Koster, R. D., and M. J. Suarez (1991), A simplified treatment of SiB's land surface albedo parameterization.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure, *Journal of Geophysical Research-Atmospheres*, *105*(D20), 24809-24822, doi:Doi 10.1029/2000jd900327.

Koven, C. D., W. J. Riley, and A. Stern (2013), Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models, *Journal of Climate*, *26*(6), 1877-1900, doi:10.1175/Jcli-D-12-00228.1.

Lawrence, D. M., and A. G. Slater (2005), A projection of severe near-surface permafrost degradation during the 21st century, *Geophys Res Lett*, *32*(24).

Lawrence, D. M., and A. G. Slater (2008), Incorporating organic soil into a global climate model, *Clim Dynam*, *30*(2-3), 145-160, doi:10.1007/s00382-007-0278-1.

Lawrence, D. M., A. G. Slater, V. E. Romanovsky, and D. J. Nicolsky (2008), Sensitivity of a model projection of near-surface permafrost degradation to soil column depth and representation of soil organic matter, *Journal of Geophysical Research-Earth Surface*, *113*(F2), doi:10.1029/2007jf000883.

Liu, L., K. Schaefer, T. Zhang, and J. Wahr (2012), Estimating 1992–2000 average active layer thickness on the Alaskan North Slope from remotely sensed surface subsidence, *Journal of Geophysical Research: Earth Surface*, *117*(F1).

Liu, L., T. Zhang, and J. Wahr (2010), InSAR measurements of surface deformation over permafrost on the North Slope of Alaska, *Journal of Geophysical Research: Earth Surface*, *115*(F3).

Luthin, J., and G. Guymon (1974), Soil moisture-vegetation-temperature relationships in central Alaska, *J. Hydrol.*, *23*(3-4), 233-246.

Mahanama, S. P., R. D. Koster, G. K. Walker, L. L. Takacs, R. H. Reichle, G. De Lannoy, Q. Liu, B. Zhao, and M. J. Suarez (2015), Land Boundary Conditions for the Goddard Earth Observing System Model Version 5 (GEOS-5) Climate Modeling System: Recent Updates and Data File Descriptions.

Mishra, U., and W. J. Riley (2014), Active-layer thickness across Alaska: comparing observation-based estimates with CMIP5 earth system model predictions, *Soil Sci Soc Am J*, *78*(3), 894-902.

Molders, N., and V. E. Romanovsky (2006), Long-term evaluation of the Hydro-Thermodynamic Soil-Vegetation Scheme's frozen ground/permafrost component using observations at Barrow, Alaska, *Journal of Geophysical Research-Atmospheres*, *111*(D4), doi:10.1029/2005jd005957.

Moody, E. G., M. D. King, C. B. Schaaf, and S. Platnick (2008), MODIS-Derived Spatially Complete Surface Albedo Products: Spatial and Temporal Pixel Distribution and Zonal Averages, *Journal of Applied Meteorology and Climatology*, *47*(11), 2879-2894.

Nelson, F., N. Shiklomanov, G. Mueller, K. Hinkel, D. Walker, and J. Bockheim (1997), Estimating active-layer thickness over a large region: Kuparuk River basin, Alaska, USA, *Arctic Alpine Res*, 367-378.

Nicolsky, D. J., V. E. Romanovsky, V. A. Alexeev, and D. M. Lawrence (2007), Improved modeling of permafrost dynamics in a GCM land-surface scheme, *Geophys Res Lett*, *34*(8).

Osterkamp, T. E., and V. E. Romanovsky (1999), Evidence for warming and thawing of discontinuous permafrost in Alaska, *Permafrost Periglac*, *10*(1), 17-37, doi:Doi 10.1002/(Sici)1099-1530(199901/03)10:1<17::Aid-Ppp303>3.0.Co;2-4.

Panda, S. K., A. Prakash, D. N. Solie, V. E. Romanovsky, and M. T. Jorgenson (2010), Remote Sensing and Field-based Mapping of Permafrost Distribution along the Alaska Highway Corridor, Interior Alaska, *Permafrost Periglac*, *21*(3), 271-281, doi:10.1002/ppp.686.

Pastick, N. J., M. T. Jorgenson, B. K. Wylie, S. J. Nield, K. D. Johnson, and A. O. Finley (2015), Distribution of near-surface permafrost in Alaska: Estimates of present and future conditions, *Remote Sensing of Environment*, *168*, 301-315.

Rautiainen, K., J. Lemmetyinen, M. Schwank, A. Kontu, C. B. Ménard, C. Maetzler, M. Drusch, A. Wiesmann, J. Ikonen, and J. Pulliainen (2014), Detection of soil freezing from L-band passive microwave observations, *Remote Sensing of Environment*, *147*, 206-218.

Reichle, R. H., G. J. M. De Lannoy, Q. Liu, J. V. Ardizzone, F. Chen, A. Colliander, A. Conaty, W. Crow, T. Jackson, J. Kimball, R. D. Koster, and E. B. Smith (2016), Soil Moisture Active Passive Mission L4_SM Data Product Assessment (Version 2 Validated Release)*Rep.*, NASA GMAO Office Note, No. 12 (Version 1.0), National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA.

Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka (2017), Land Surface Precipitation in MERRA-2, *Journal of Climate*, *30*(5), 1643-1664, doi:10.1175/jcli-d-16-0570.1.

Riseborough, D., N. Shiklomanov, B. Etzelmuller, S. Gruber, and S. Marchenko (2008), Recent advances in permafrost modelling, *Permafrost Periglac*, *19*(2), 137-156, doi:10.1002/ppp.615.

Romanovsky, V. E., A. L. Kholodov, W. L. Cable, L. Cohen, S. Panda, S. Marchenko, R. R. Muskett, and D. Nicolsky (2009), Network of Permafrost Observatories in North America and Russia. NSF Arctic Data Center, doi:10.18739/A2SH27.

Romanovsky, V. E., and T. E. Osterkamp (1995), Interannual variations of the thermal regime of the active layer and near-surface permafrost in northern Alaska, *Permafrost Periglac*, *6*(4), 313-335, doi:DOI 10.1002/ppp.3430060404.

Romanovsky, V. E., and T. E. Osterkamp (1997), Thawing of the active layer on the coastal plain of the Alaskan Arctic, *Permafrost Periglac*, *8*(1), 1-22.

Romanovsky, V. E., S. L. Smith, and H. H. Christiansen (2010), Permafrost Thermal State in the Polar Northern Hemisphere during the International Polar Year 2007-2009: a Synthesis, *Permafrost Periglac*, *21*(2), 106-116.

Sazonova, T., and V. Romanovsky (2003), A model for regional-scale estimation of temporal and spatial variability of active layer thickness and mean annual ground temperatures, *Permafrost Periglac*, *14*(2), 125-139.

Scholes, R., D. Skole, and J. Ingram (1995), A global database of soil properties: proposal for implementation*Rep.*, IGBP-DIS Working Paper. Report of the Global Soils Task Group, International Geosphere-Biosphere Programme - Data and Information System (IGBP-DIS). University of Paris, France.

Shiklomanov, N. I., and F. E. Nelson (2002), Active-layer mapping at regional scales: A 13-year spatial time series for the Kuparuk region, north-central Alaska, *Permafrost Periglac*, *13*(3), 219-230.

Shiklomanov, N. I., D. A. Streletskiy, F. E. Nelson, R. D. Hollister, V. E. Romanovsky, C. E. Tweedie, J. G. Bockheim, and J. Brown (2010), Decadal variations of active-layer thickness in moisture-controlled landscapes, Barrow, Alaska, *Journal of Geophysical Research: Biogeosciences*, *115*(G4).

Simard, M., N. Pinto, J. B. Fisher, and A. Baccini (2011), Mapping forest canopy height globally with spaceborne lidar, *Journal of Geophysical Research: Biogeosciences*, *116*(G4).

Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez (2001), The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales, *J. Hydrometeorol.*, *2*(3), 228-242.

Zhang, T., R. G. Barry, K. Knowles, J. Heginbottom, and J. Brown (1999), Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere, *Polar Geography*, *23*(2), 132-154.

Zhang, T., O. W. Frauenfeld, M. C. Serreze, A. Etringer, C. Oelke, J. McCreight, R. G. Barry, D. Gilichinsky, D. Yang, and H. Ye (2005), Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin, *Journal of Geophysical Research: Atmospheres*, *110*(D16).

Zhao, T., L. Zhang, L. Jiang, S. Zhao, L. Chai, and R. Jin (2011), A new soil freeze/thaw discriminant algorithm using AMSR-E passive microwave imagery, *Hydrol Process*, *25*(11), 1704-1716.

Zinke, P. J., A. G. Stangenberger, W. M. Post, W. R. Emanuel, and J. S. Olson (1986), Worldwide organic carbon and nitrogen data, *ONRL/CDIC-18, Carbon Dioxide Information Centre, Oak Ridge, Tenessee*.

List of Tables

List of Figures

Table 1 – Land model parameters and boundary conditions.

Land boundary conditions	Data source or generation method	Reference
Soil Depth	The Second Global Soil Wetness Project (GSWP-2).	[*Dirmeyer et al.*, 2002]
Soil parameters	Harmonized World Soil Data (HWSD-1.21) and the State Soil Geographic (STATSGO2) data set.	[*De Lannoy et al.*, 2014]
Land cover	USGS Global Land Cover Characteristics Data Base Version 2.0 (GLCCv2).	https://lta.cr.usgs.gov/glcc/
Vegetation height	The Geoscience Laser Altimeter System (GLAS) aboard ICESat (Ice, Cloud, and land Elevation Satellite).	[*Simard et al.*, 2011]
Leaf Area Index (LAI)	Moderate Resolution Imaging Spectroradiometer (MODIS) and GEOLAND2 LAI product.	[*Baret et al.*, 2013; *Camacho et al.*, 2013]
Greenness fraction	GSWP-2	[*Dirmeyer et al.*, 2002]
Albedo	Computed by a modified Simple Biosphere	[*Koster and Suarez*, 1991;

	Model (SiB) albedo parameterization scheme and (for the snow-free fraction) scaled by MODIS albedo climatology.	*Moody et al.*, 2008]

Table 2 – Permafrost sites used in Section 5.

Permafrost Sites	Latitude	Longitude	Local landcover*	Local soil information#	Purpose
DH1	70.1613°	-148.4653°	Landcover units include Graminoid-moss tundra and graminoid, prostrate-dwarf-shrub, moss tundra (wet and moist nonacidic).	15cm - Peat.	Examining Meteorological Forcing (section 5.1)
FB1	69.6739°	-148.7219°	Landcover units include Graminoid-moss tundra and graminoid, prostrate-dwarf-shrub, moss tundra (wet and moist nonacidic). This site is located on	15cm – Peat.	Examining upper boundary condition and soil organic carbon content (section 5.4)

			the inner coastal plain with river terraces.		
GL1	68.4774°	-149.5024°	Landcover units include Graminoid-moss tundra and graminoid, prostrate-dwarf-shrub, moss tundra (wet and moist nonacidic). Broad glaciated mountain valley.	80cm – Peat; 127cm - Silty loam; 199cm - Peat and silt mix; 278cm – silt.	Examining upper boundary condition and soil organic carbon content (section 5.4)
SG2	69.4283°	-148.7001°	Moist acidic tundra	15cm – Peat; 40cm - Silty loam.	Examining upper boundary condition and soil organic carbon content (section 5.4)
SL1	64.8694°	-147.8608°	Forest	31cm – Peat.	Examining upper boundary

					condition and soil organic carbon content (section 5.3 and 5.4)
SL2	64.8661°	-147.8568°	Forest	---	Examining upper boundary condition and soil organic carbon content (section 5.4)
SL3	64.8675°	-147.8588°	Forest	---	Examining upper boundary condition and soil organic carbon content (section 5.4)
SL4	64.8669°	-147.8584°	Forest	---	Examining upper boundary condition and soil organic carbon content (section

					5.4)
UF1	64.8529°	-147.8575°	Agricultural field	---	Examining land cover type and upper boundary condition (section 5.2 and 5.3)

* Information is from http://permafrost.gi.alaska.edu/sites_map.

Information is from personal communication with with Dr. Vladimir Romanovsky and Dr. Alexander Kholodov from University of Alaska Fairbanks.

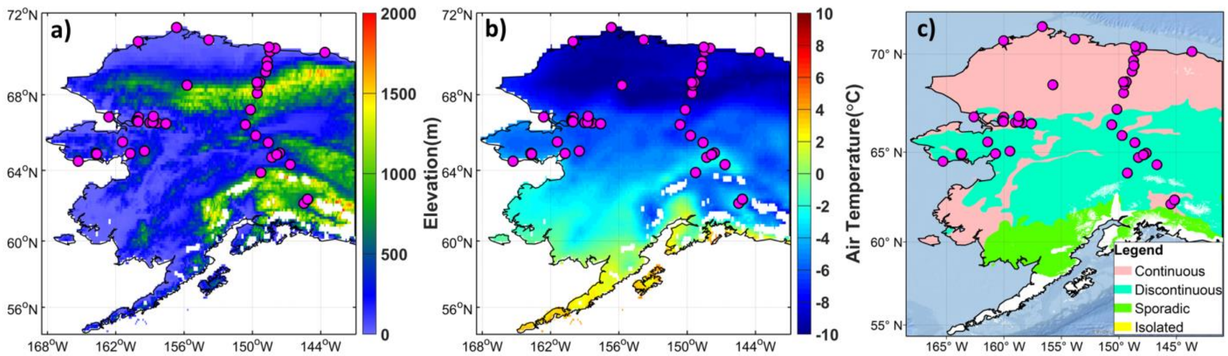


Figure 1 – (a) Elevation data underlying GEOS-5, (b) air temperature at 2m above the ground extracted from MERRA-2 for the Alaska domain and (c) a permafrost extent map categorized by four types, i.e., Continuous (90-100%), Discontinuous (50- 90%), sporadic(10- 50%) and isolated patches (0 - 10%) [*Brown et al.*, 2002], obtained from the National Snow and Ice Data Center. Regions in white in (a) and (b) denote glaciers. Magenta dots indicate the locations of in situ permafrost sites used in this study.

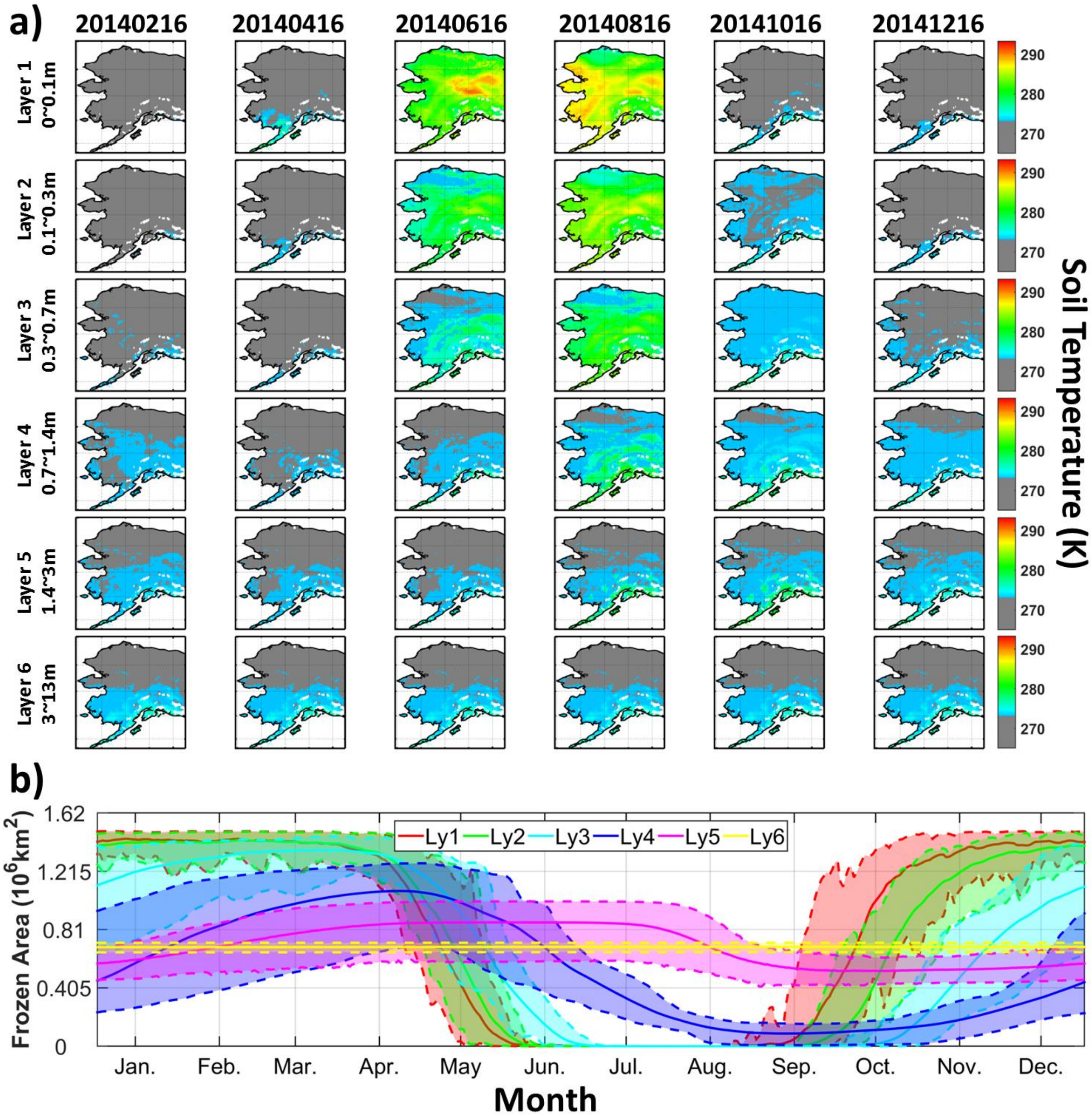


Figure 2 – (a) Example of modeled soil temperature for 6 dates in 2014. Gray color indicates frozen soil (temperature equal to or below 273.15K). (b) 35-year climatology of frozen area, with shaded area representing the range associated with inter-annual variability. Dashed lines indicate the maximum and minimum across the 35 years.

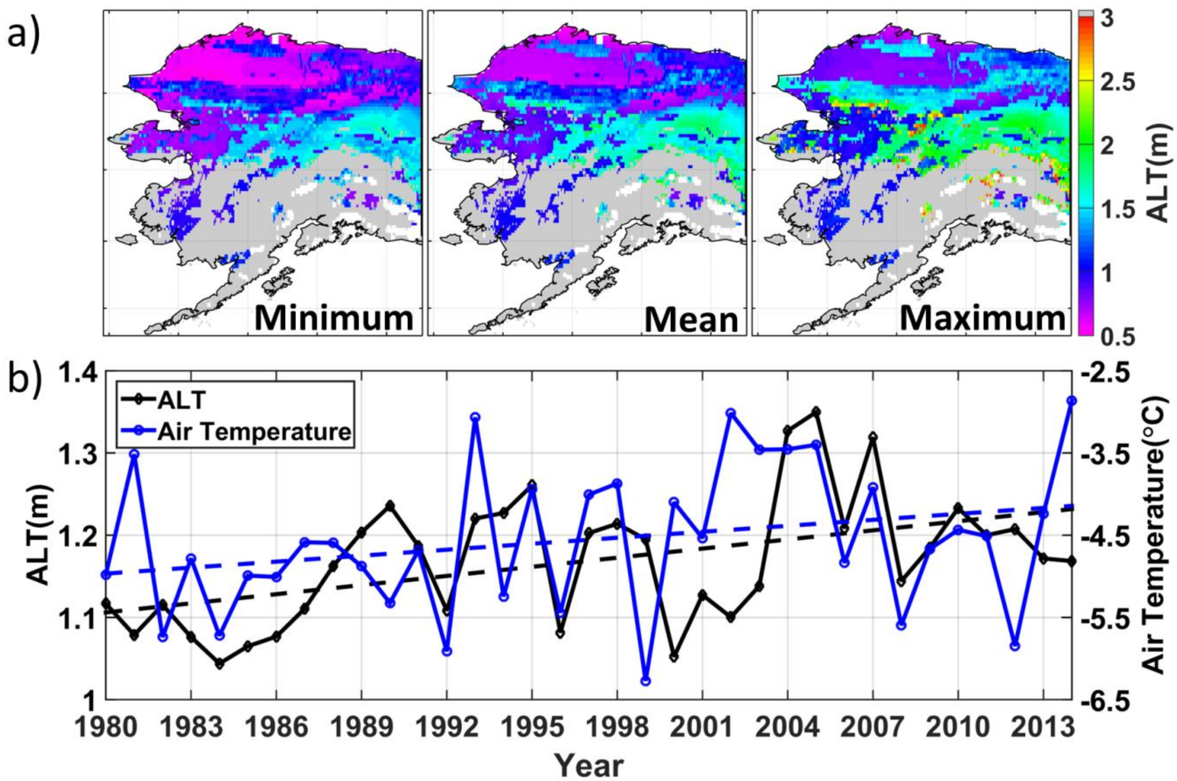


Figure 3 – (a) 35-year minimum, mean, and maximum of the annual ALT. The light gray color indicates permafrost-free areas. (b) Spatial mean of the annual ALT (black) and the annual mean 2-m air temperature (blue). Dashed lines are linearly fitted trend lines for the two variables.

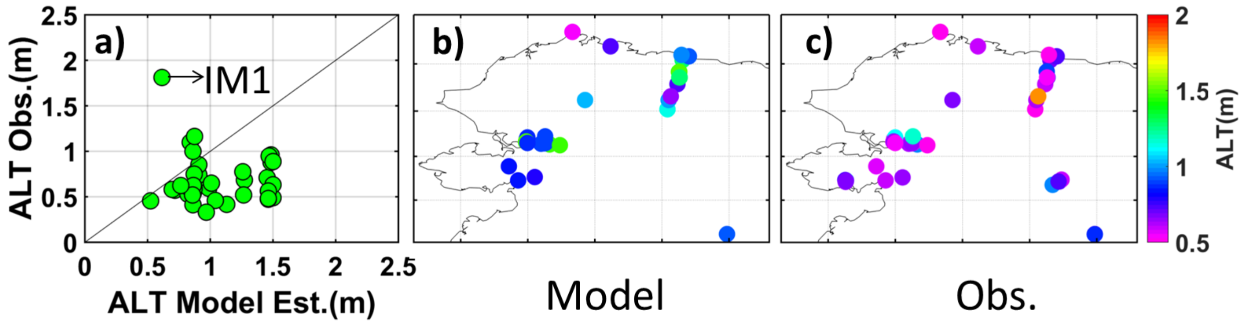


Figure 4 – (a) Multi-year mean of simulated (abscissa) vs. observed (ordinate) ALT. (b), (c) Maps of the multi-year mean ALT from (b) the model simulation and (c) the in situ observations.

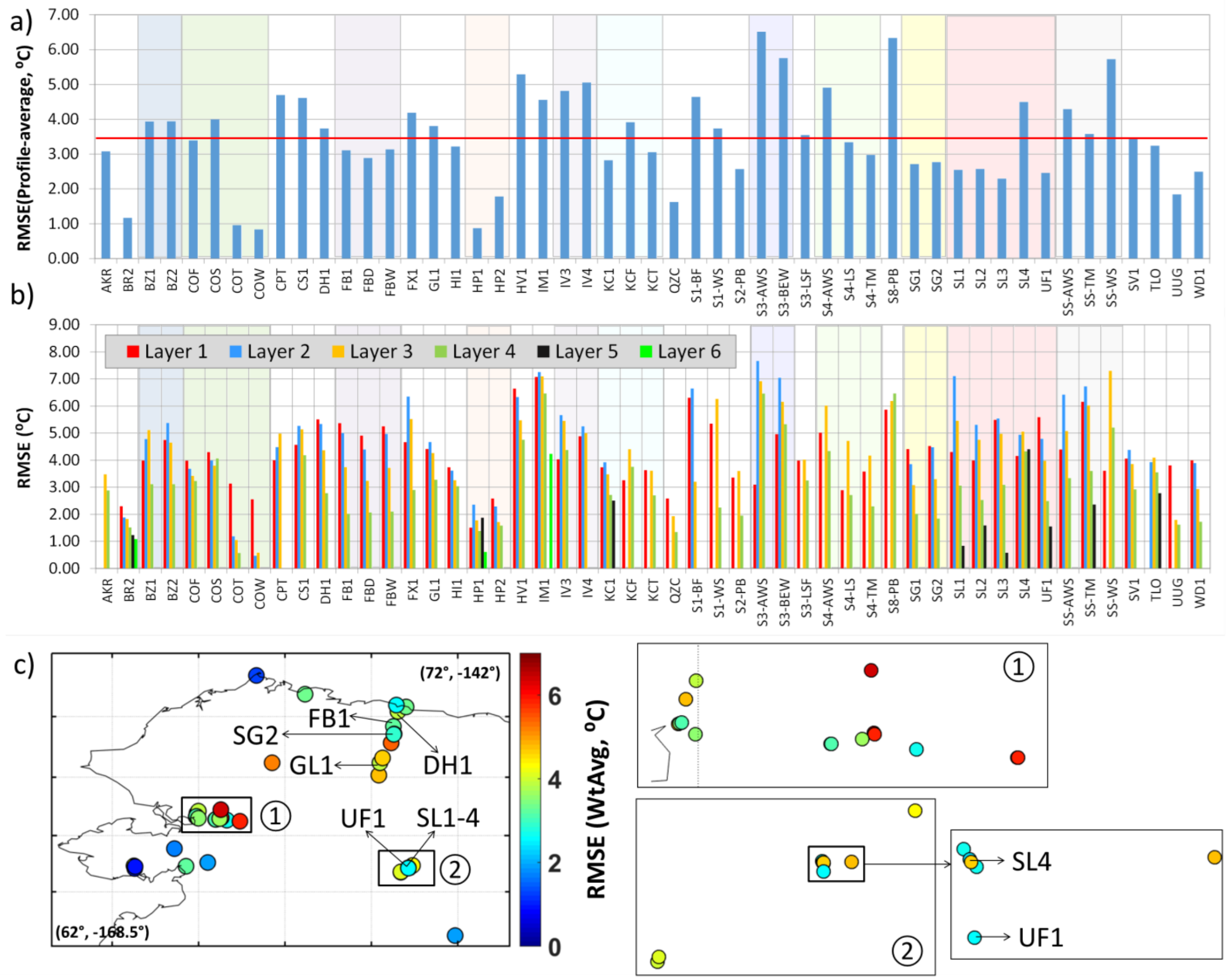


Figure 5 – (a) Profile-average RMSE for soil temperature estimates from the baseline simulation at 51 sites across Alaska. (b) As in (a) but for the RMSE of each soil layer. Background shading in (a) and (b) indicates sites that are within the same 9-km model grid cell. (c) Map of the profile-average RMSE for soil temperature. Note that symbols overlap for sites that are close to each other. Two overlapping areas (denoted ① and ②) are zoomed in for details.

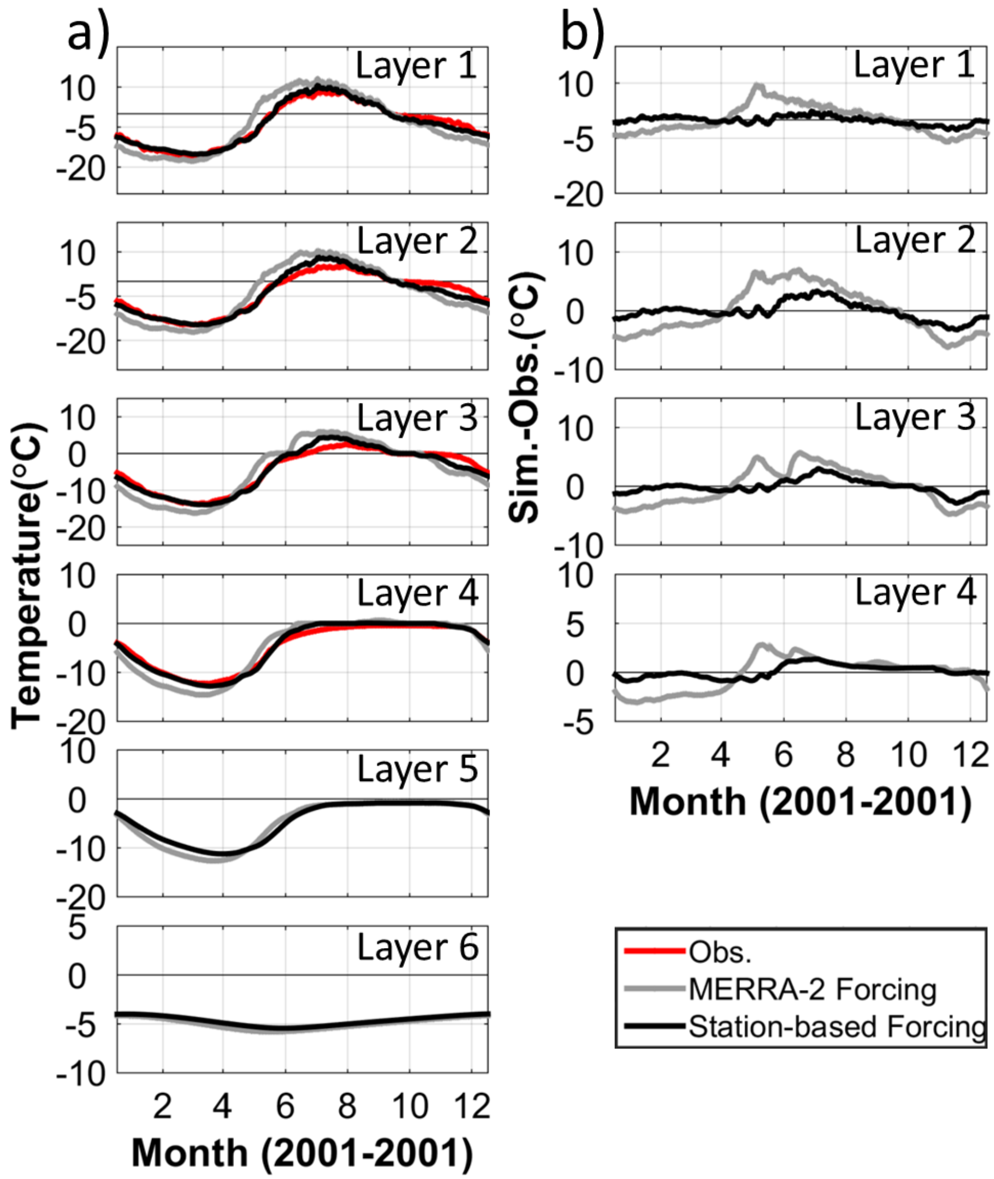


Figure 6 – (a) Comparison of multi-year mean seasonal cycles of observed (red) and simulated soil temperature results at DH1 with original MERRA-2 forcing fields (in gray) and station-based forcing (in black). Differences between simulations and observations for top four layers are shown in panel (b).

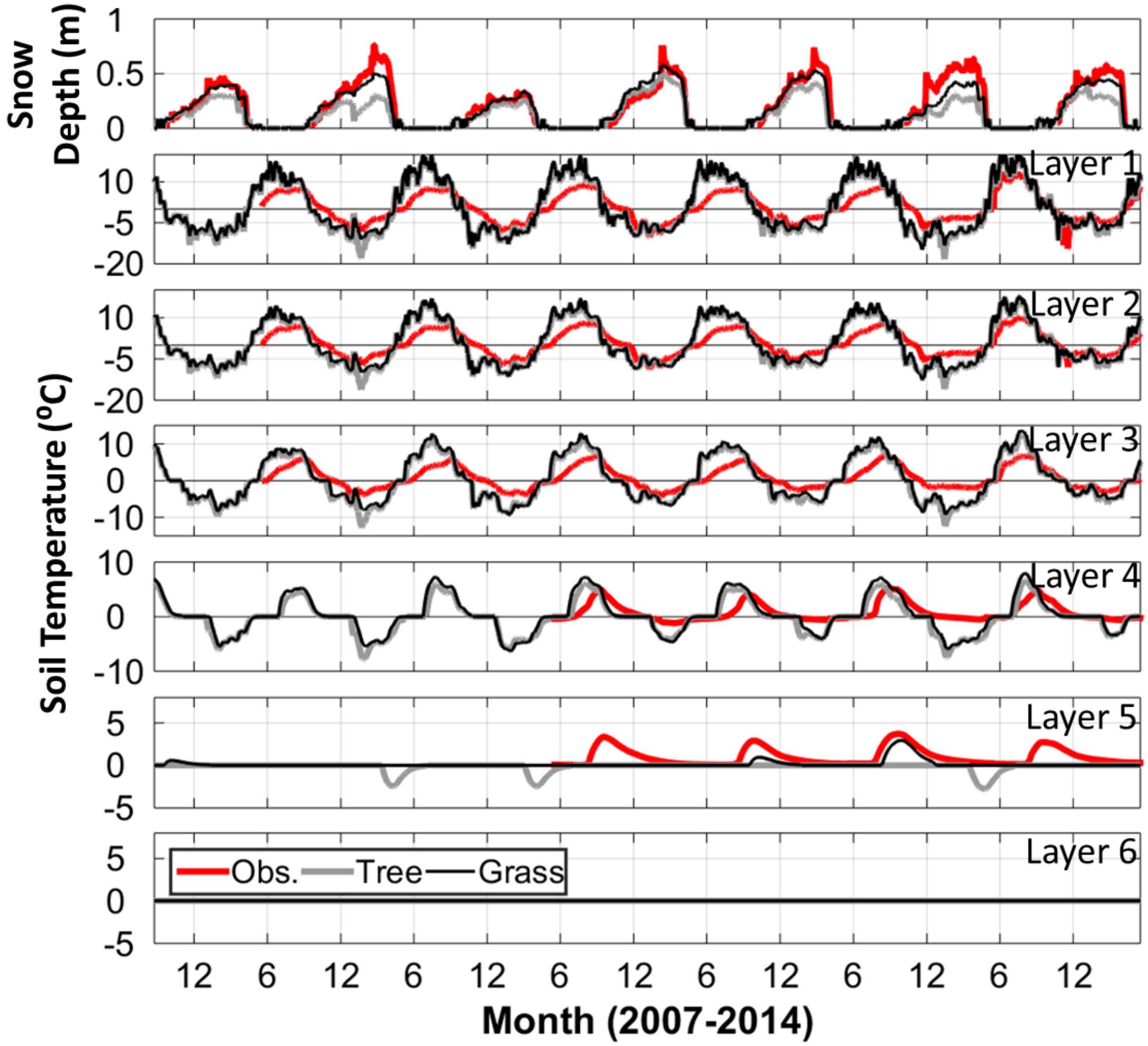


Figure 7 – Comparison of observed (red) and simulated soil temperature results at UF1 with original global land cover (denoted “Tree” in gray) and grassland (denoted “Grass” in black) in accordance with local surface conditions. Top panel shows the observed and simulated snow depth for each of the two experiments.

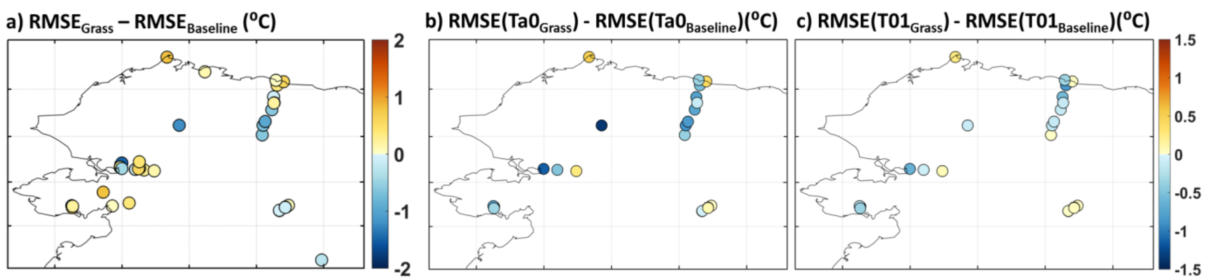


Figure 8 – (a) Difference of profile-average RMSE between the "Grass" experiment and the baseline results. Blue colors (negative values) indicate model improvements whereas orange and red colors (positive values) indicate model degradation. (b) Difference in RMSE of temperature offset along the air-to-shallow soil gradient (Ta0) between the two experiments. (c) Difference in RMSE of temperature offset along the shallow-to-deep soil gradient (T01) between the two experiments.

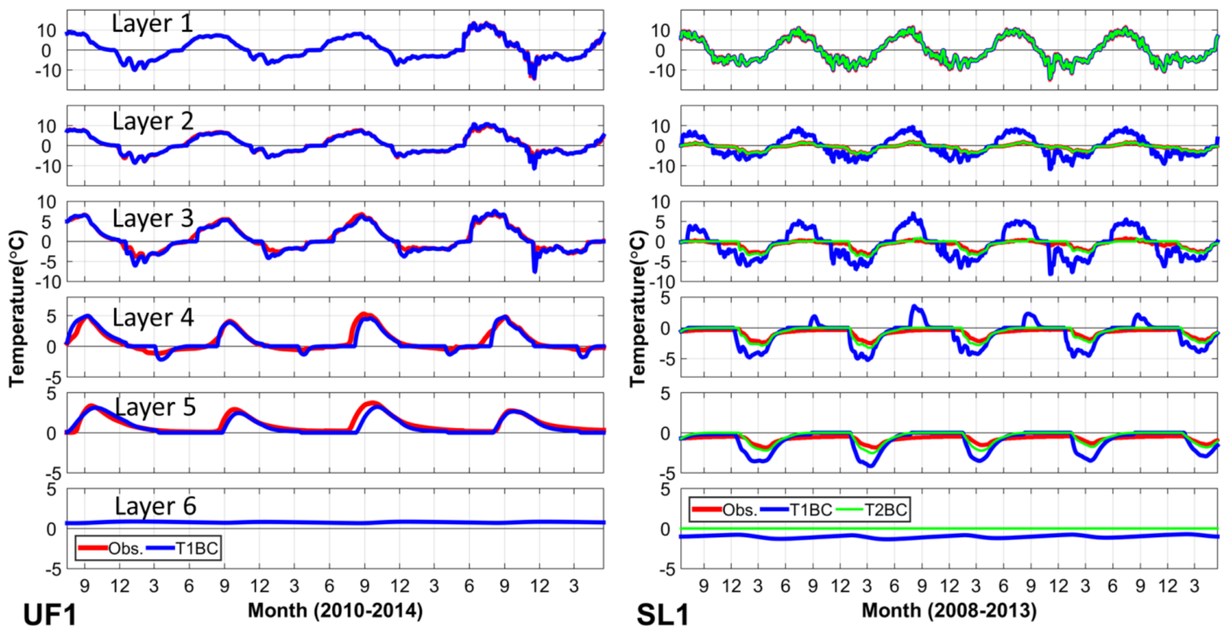


Figure 9 – Comparison of observed (red line) and simulated (blue line) soil temperature where observations are used to prescribe the top layer temperature (denoted T1BC) at UF1 and SL1. For SL1, simulation results from T2BC (green line) in which soil temperatures at both the 1st and the 2nd layer were prescribed to observations are also shown.

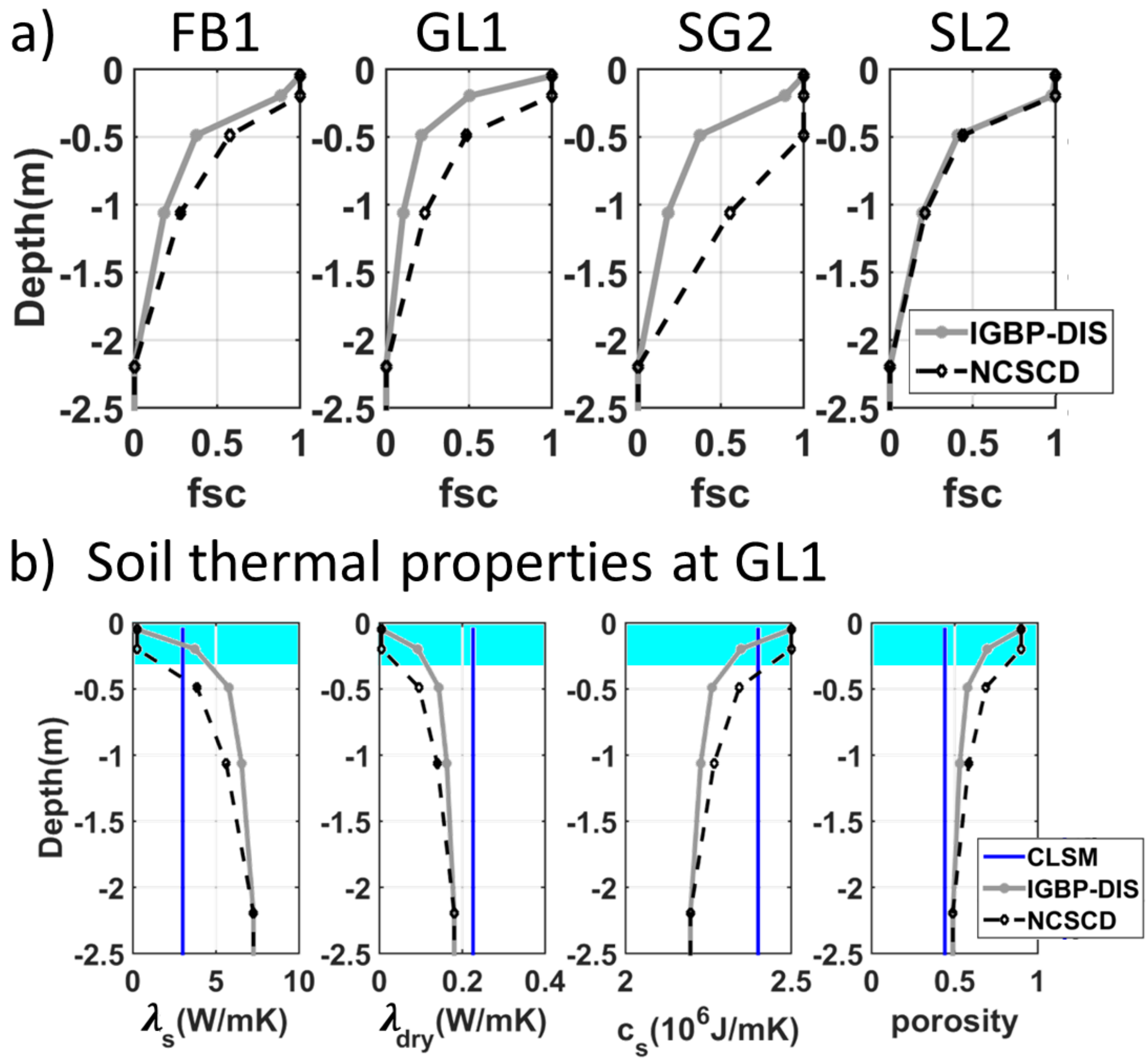


Figure 10 – (a) Vertical profiles of soil carbon fraction (fsc) based on IGBP-DIS and NCSCD at sites FB1, GL1, SG2 and SL2. Profiles at SL1, SL3 and SL4 are identical to SL2. The gray profile is based on IGBP-DIS. The black dash profile is derived using NCSCD. The cumulative carbon storage profile for polar and boreal soils as identified in *Zinke et al.* [1986] was used to calculate the vertical profile. (b) Example of the associated soil thermal properties at site GL1, including the thermal conductivity for soil solids (λ_s), the thermal conductivity for dry soil (λ_{dry}), the specific heat capacity of soil (c_s) and soil porosity. Blue line represents the default values originally used in CLSM. Cyan shading indicates the extent of the top two model layers.

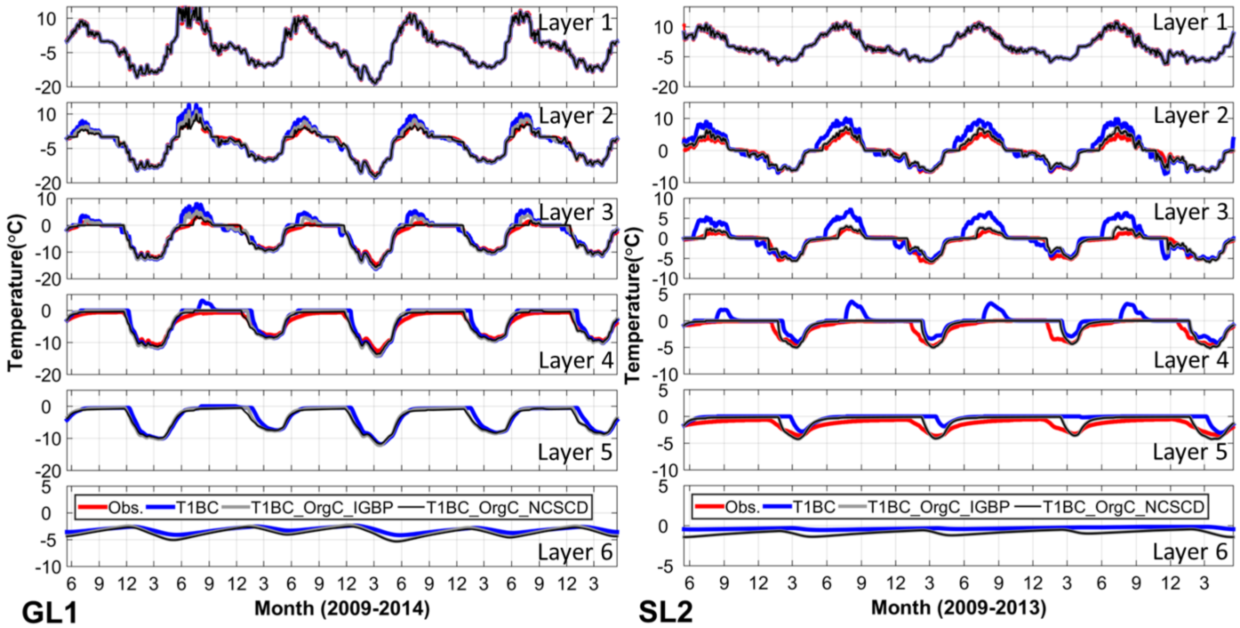


Figure 11 – Simulation results at GL1 and SL2 for baseline T1BC experiment in which soil temperature in the top layer was prescribed from in situ observations, as well as from two T1BC simulations (T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD) that incorporate organic carbon content profiles derived from the two carbon datasets (IGBP-DIS and NCSCD).

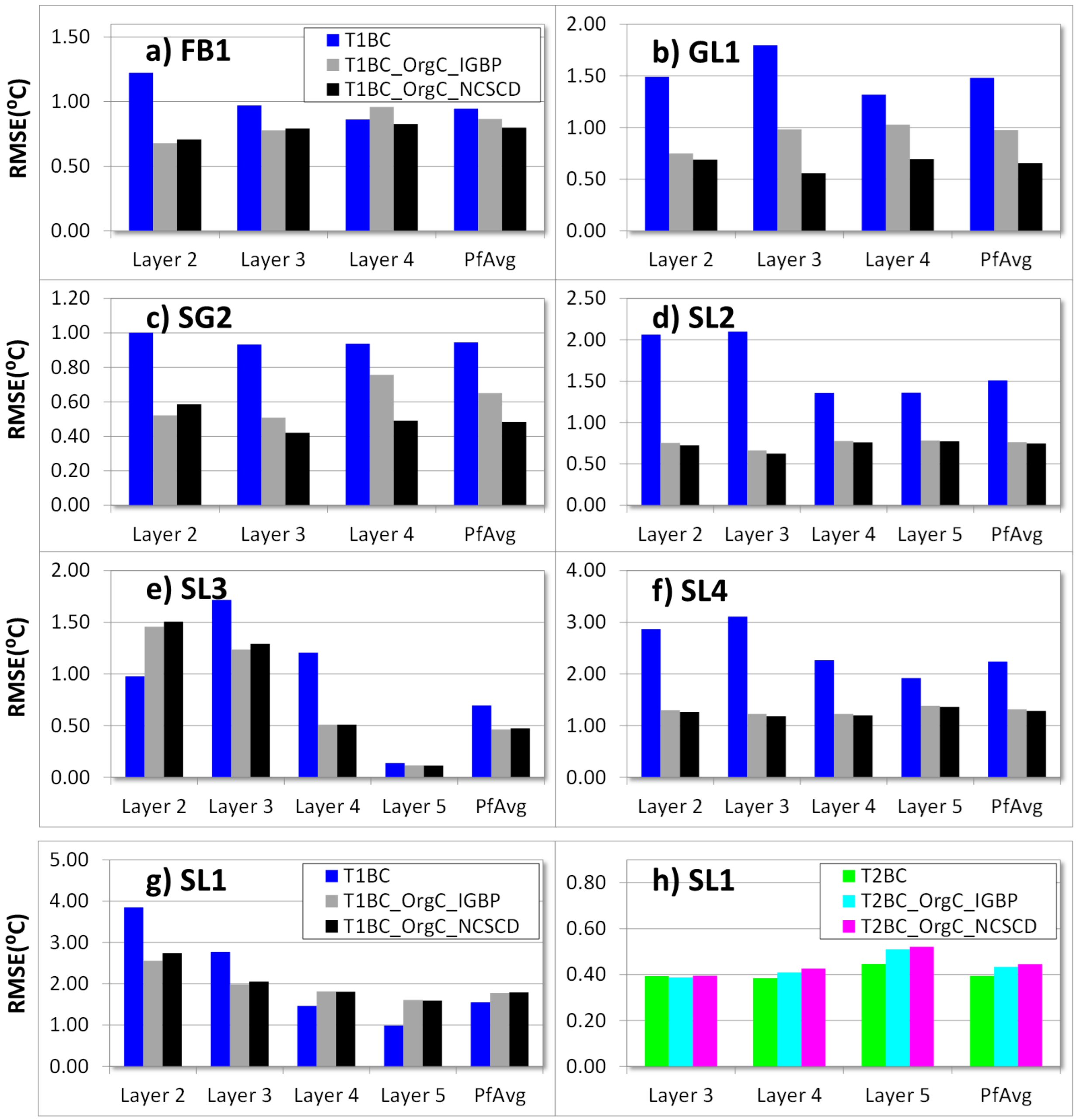


Figure 12 – RMSE (^{o}C) of soil temperature for individual model layers and the profile-average RMSE (PfAvg) at FB1, GL1, SG2, SL2, SL3, SL4, and SL1 from the baseline T1BC simulation and from the two T1BC simulations incorporating organic carbon content profiles (T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD). For SL1, RMSEs for the baseline T2BC simulation and from the two T2BC simulations using the carbon datasets are also shown.

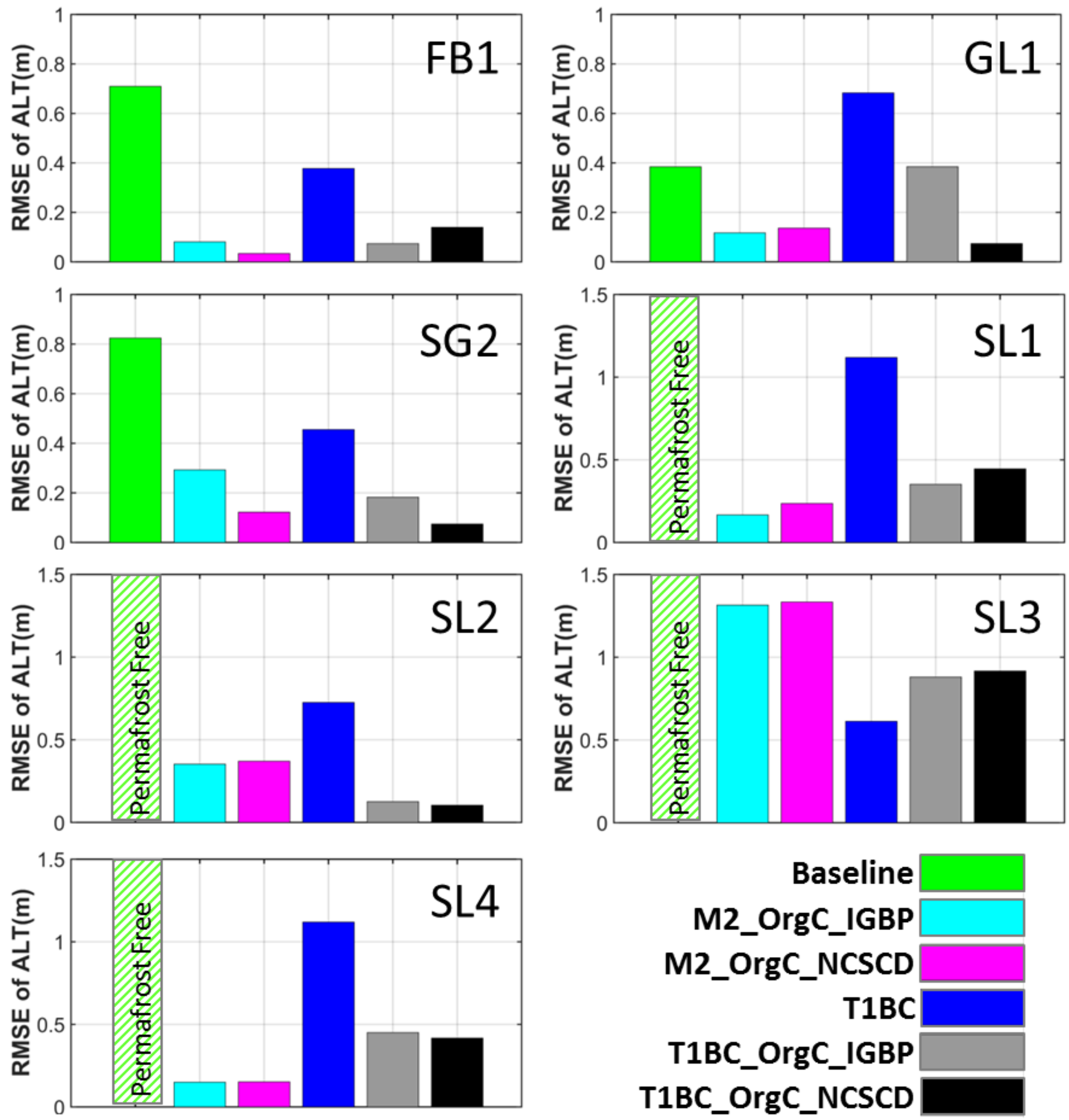


Figure 13 – The RMSEs of annual ALT from different experiments at the seven testing sites, including three simulations with MERRA-2 forcing (i.e. Baseline, M2_OrgC_IGBP and M2_OrgC_NCSCD) and three simulations with prescribed top soil temperature (i.e. T1BC, T1BC_OrgC_IGBP and T1BC_OrgC_NCSCD). Baseline simulation results indicate that SL1, SL2, SL3 and SL4 are all permafrost free and thus the RMSE for these sites are null.

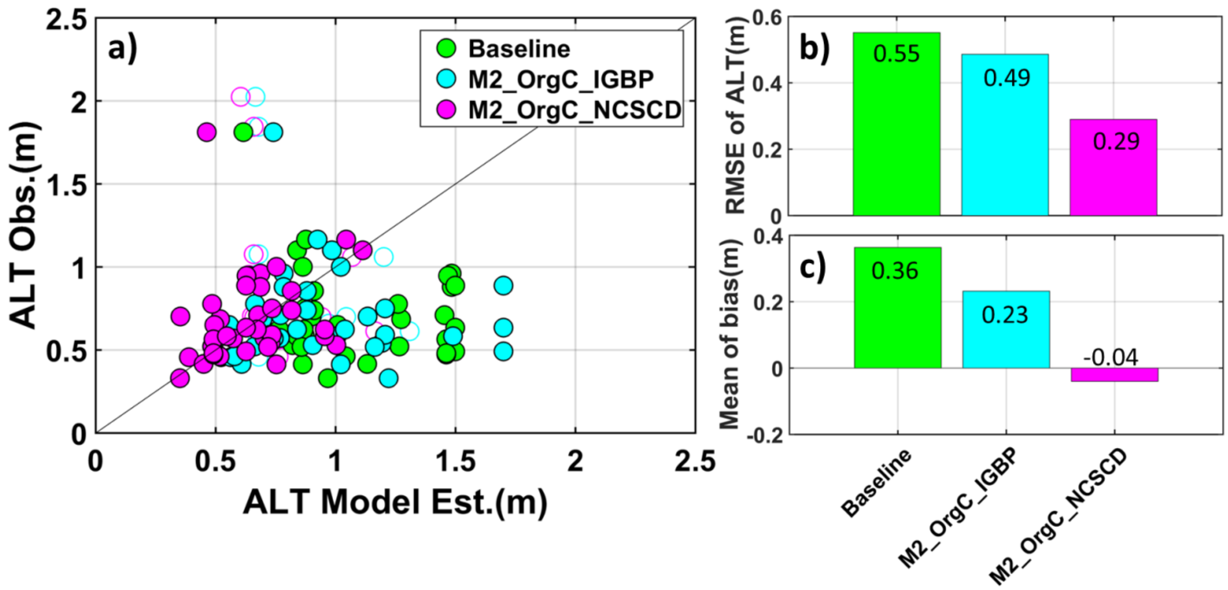


Figure 14 – (a) Multi-year mean of estimated ALT from three simulations driven by MERRA2 forcing vs. observed ALT at sites across Alaska, including baseline simulation and the two simulations incorporating organic carbon impacts (M2_OrgC_IGBP and M2_OrgC_NCSCD). Open cycles represent sites that baseline simulation show permafrost-free (thus no corresponding green dots) whereas the simulations with carbon impacts do not, and are not used for calculation of RMSE and bias. (b) RMSE of the multi-year mean of ALT from the three experiments. (c) Mean of bias of the multi-year mean of ALT from the three experiments.

Figure 1.

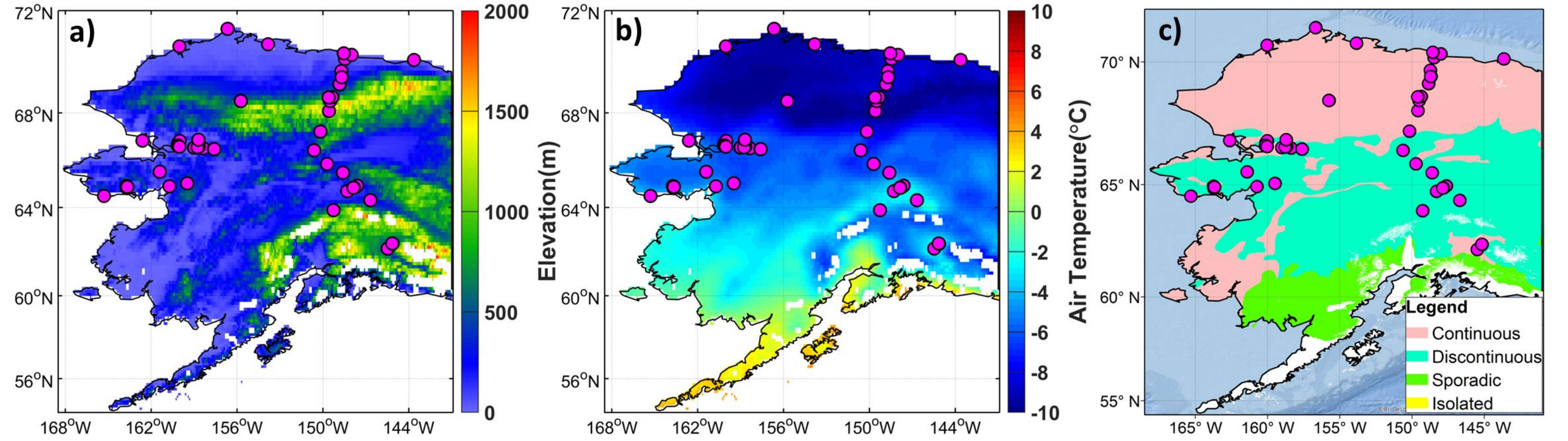
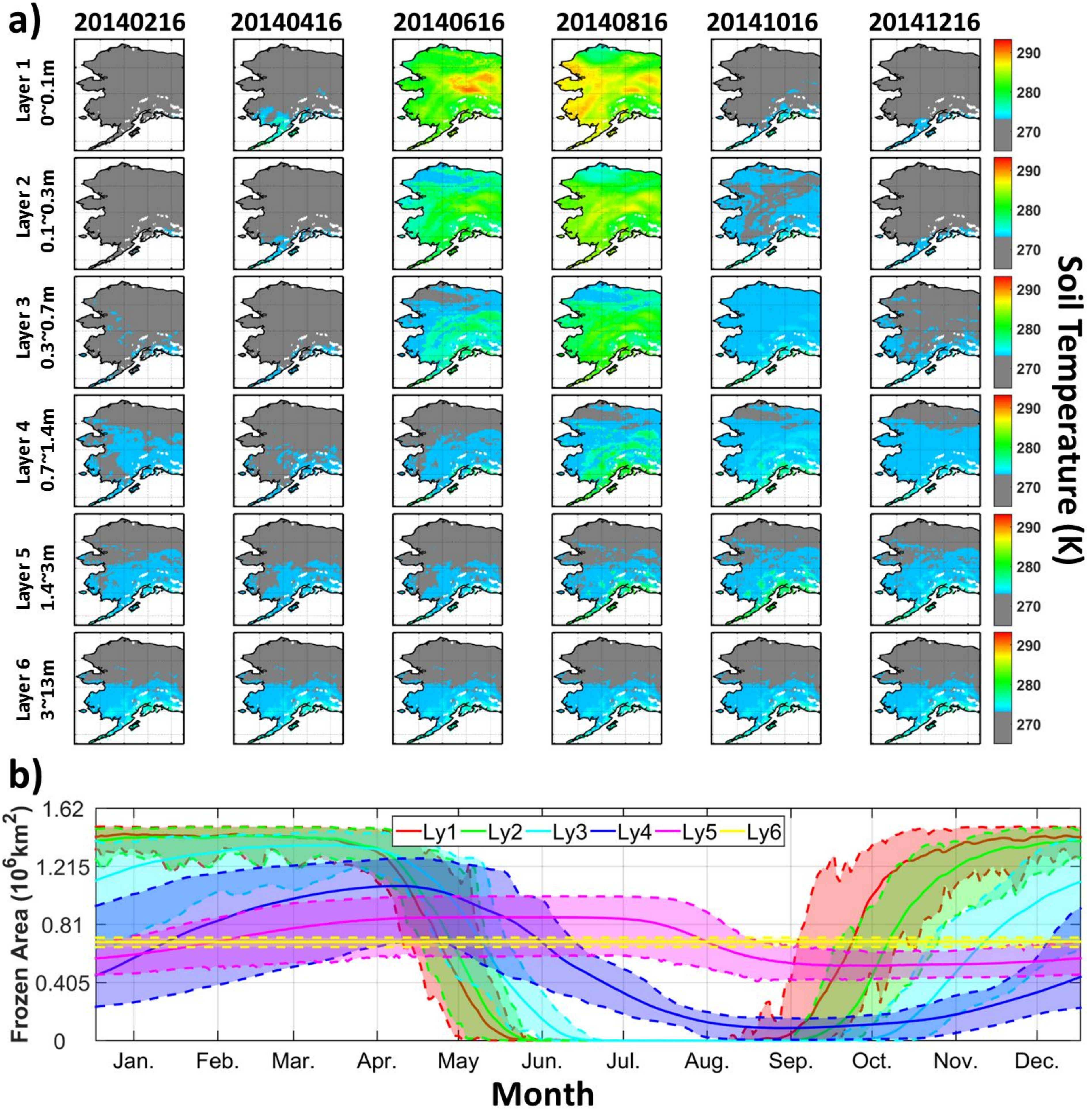
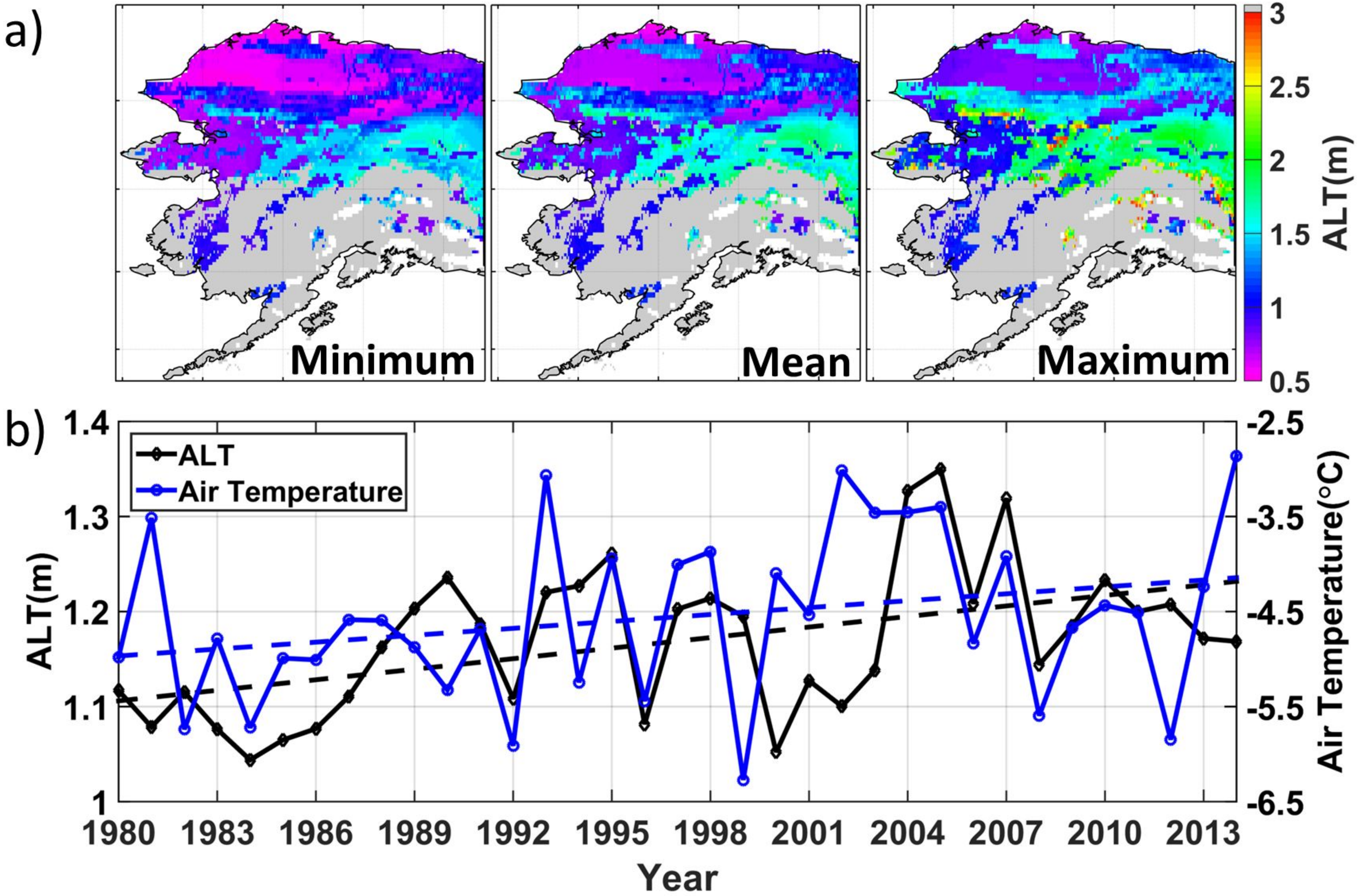

a)
b)
c)
72°N
68°N
64°N
60°N
56°N
168°W
162°W
156°W
150°W
144°W
2000
1500
1000
500
0
Elevation(m)
10
8
6
4
2
0
-2
-4
-6
-8
-10
Air Temperature(°C)
70° N
65° N
60° N
55° N
165° W
160° W
155° W
150° W
145° W
Legend
Continuous
Discontinuous
Sporadic
Isolated

Figure 2.



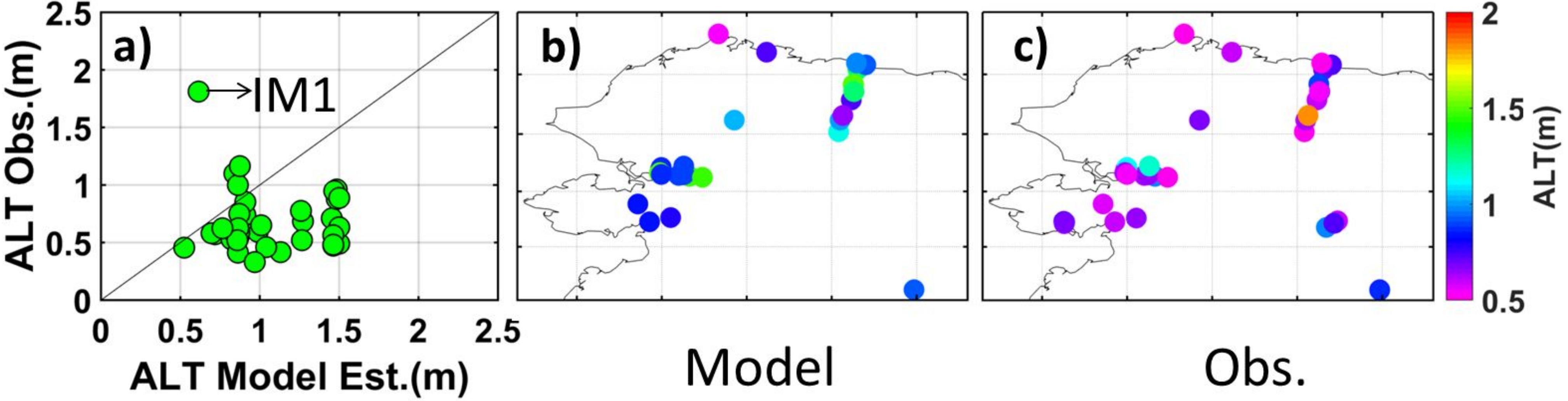
a)
20140216
20140416
20140616
20140816
20141016
20141216
Layer 1 0~0.1m
Layer 2 0.1~0.3m
Layer 3 0.3~0.7m
Layer 4 0.7~1.4m
Layer 5 1.4~3m
Layer 6 3~13m
290
280
270
Soil Temperature (K)
b)
Frozen Area (10^6 km^2)
1.62
1.215
0.81
0.405
0
Ly1
Ly2
Ly3
Ly4
Ly5
Ly6
Jan.
Feb.
Mar.
Apr.
May
Jun.
Jul.
Aug.
Sep.
Oct.
Nov.
Dec.
Month

Figure 3.



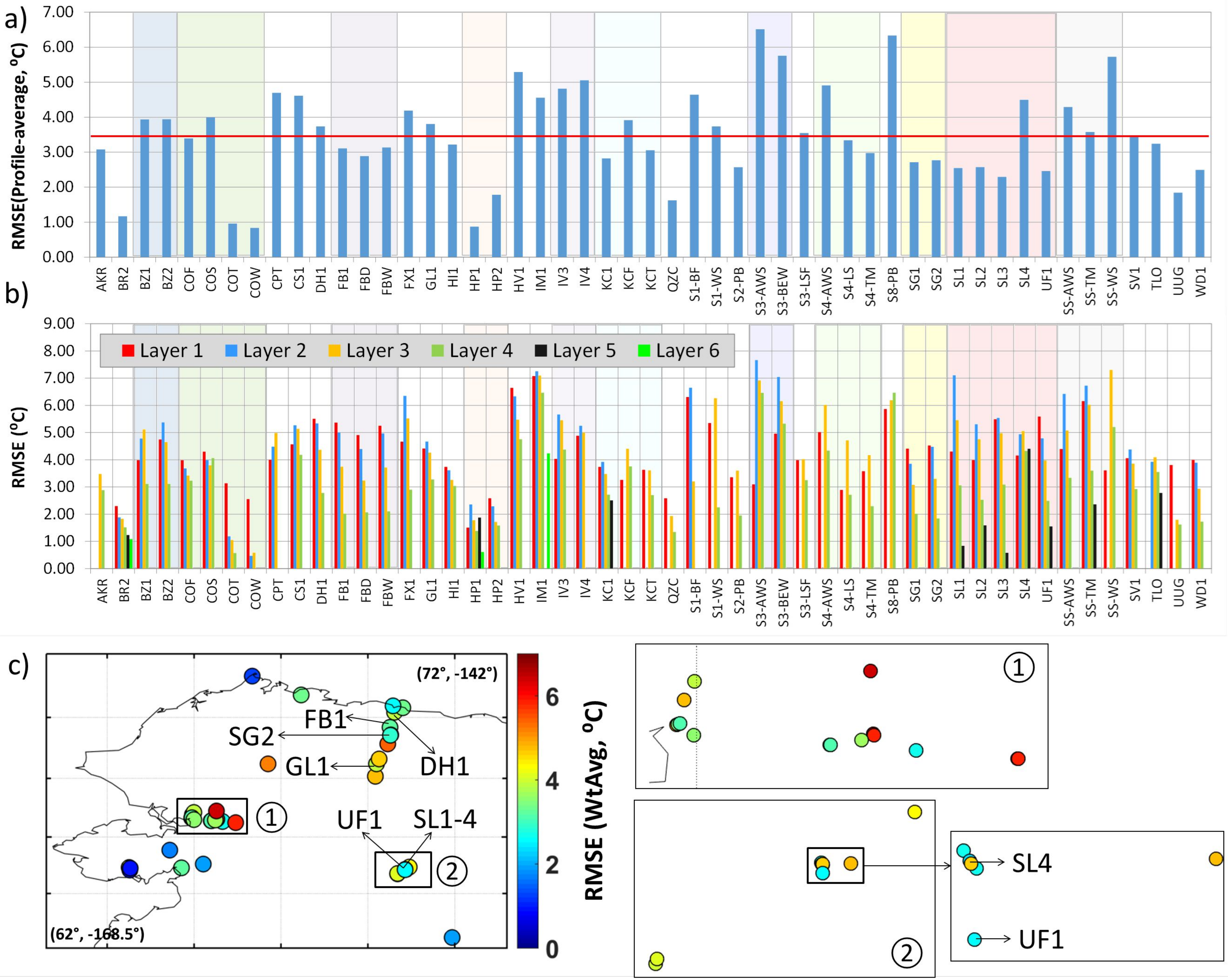
a)
Minimum
Mean
Maximum
ALT(m)
3
2.5
2
1.5
1
0.5
b)
ALT
Air Temperature
ALT(m)
Air Temperature(°C)
Year
1.4
1.3
1.2
1.1
1
-2.5
-3.5
-4.5
-5.5
-6.5
1980
1983
1986
1989
1992
1995
1998
2001
2004
2007
2010
2013

Figure 4.



a)
IM1
ALT Obs.(m)
ALT Model Est.(m)
0
0.5
1
1.5
2
2.5
b)
c)
Model
Obs.
ALT(m)

Figure 5.



a)
RMSE(Profile-average, °C)
b)
RMSE (°C)
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6
AKR BR2 BZ1 BZ2 COF COS COT COW CPT CS1 DH1 FB1 FBD FBW FX1 GL1 HI1 HP1 HP2 HV1 IM1 IV3 IV4 KC1 KCF KCT QZC S1-BF S1-WS S2-PB S3-AWS S3-BEW S3-LSF S4-AWS S4-LS S4-TM S8-PB SG1 SG2 SL1 SL2 SL3 SL4 UF1 SS-AWS SS-TM SS-WS SV1 TLO UUG WD1
c)
(72°, -142°)
(62°, -168.5°)
FB1
SG2
GL1
DH1
UF1
SL1-4
①
②
RMSE (WtAvg, °C)
SL4
UF1

Figure 6.

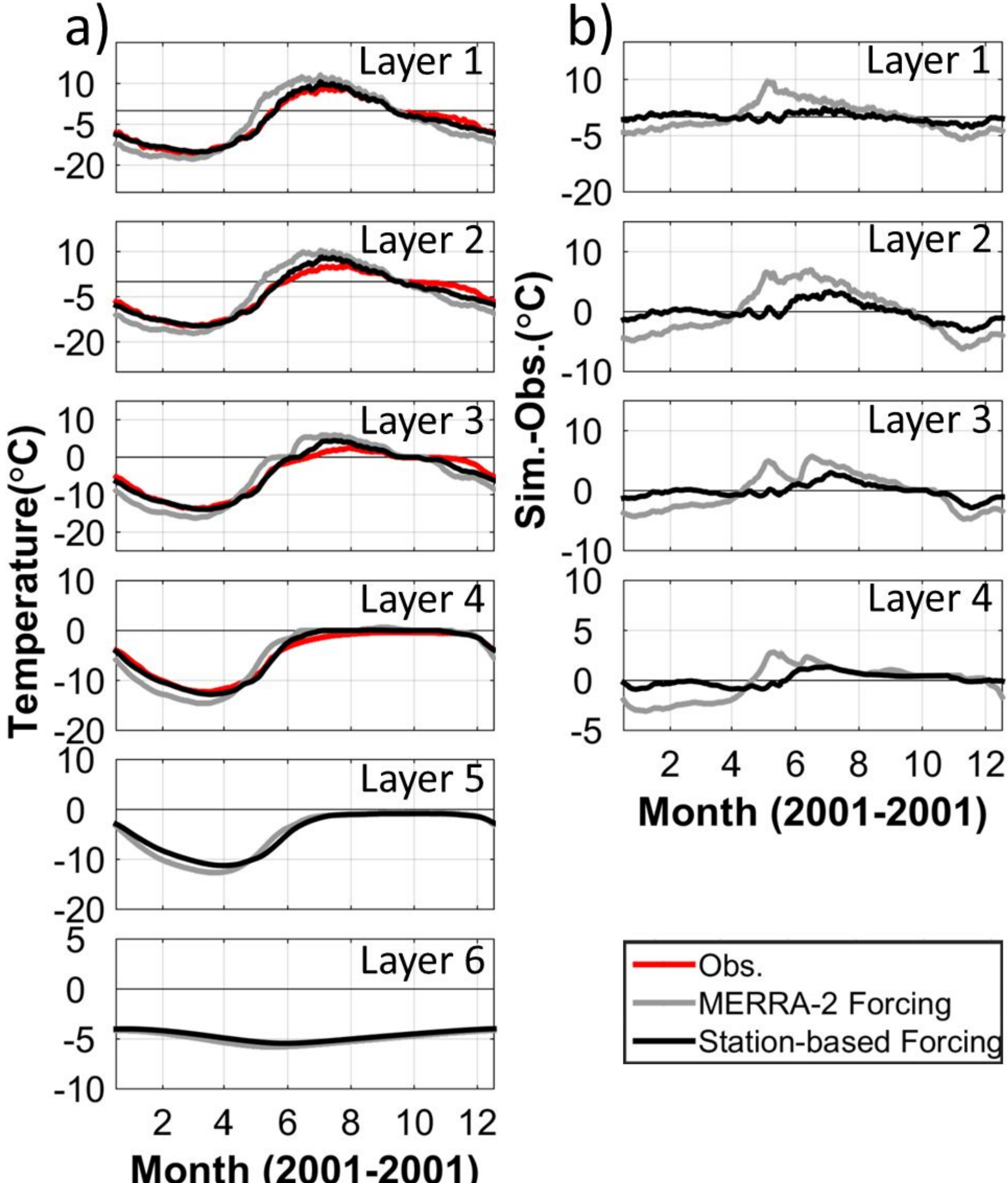
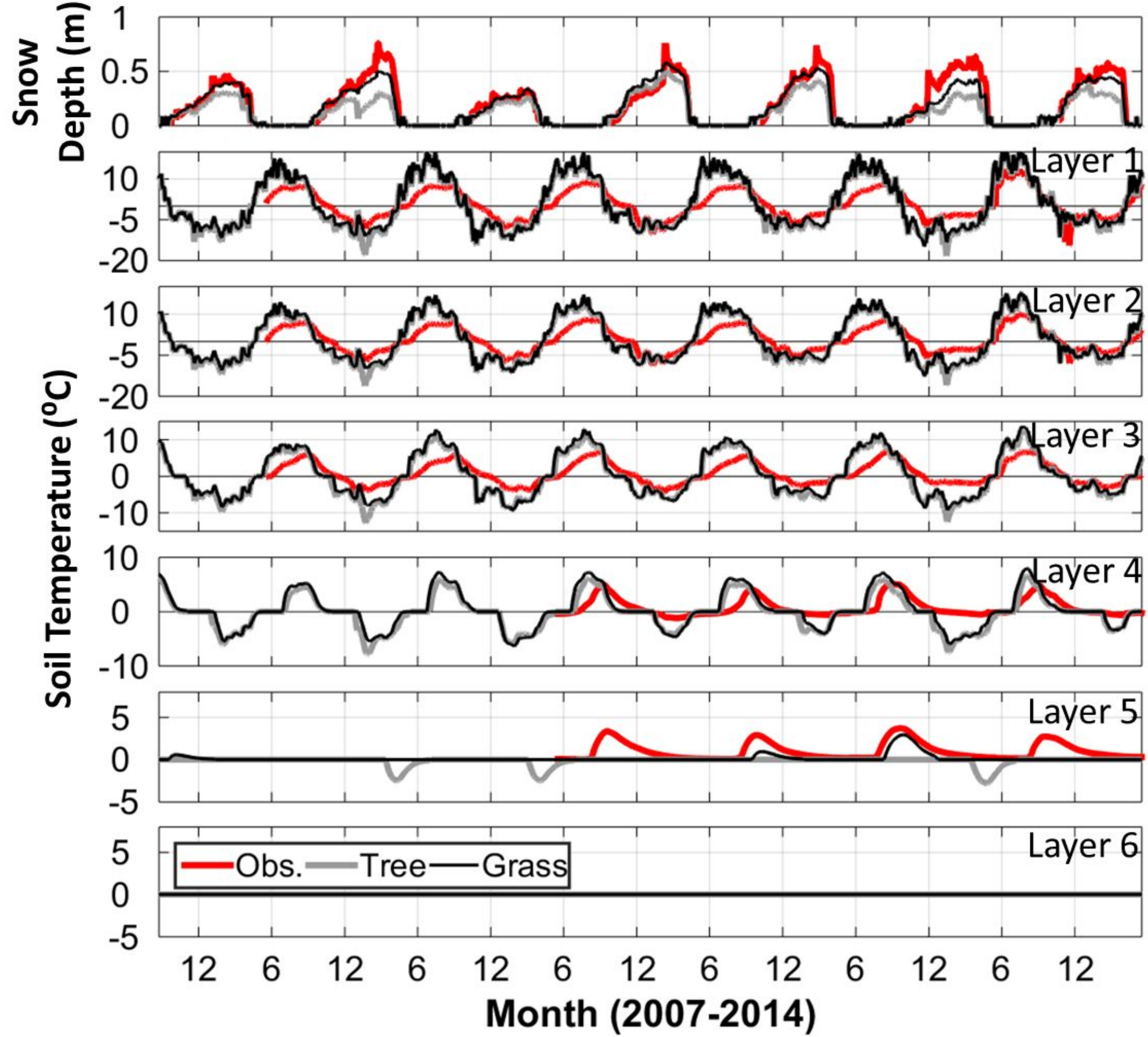
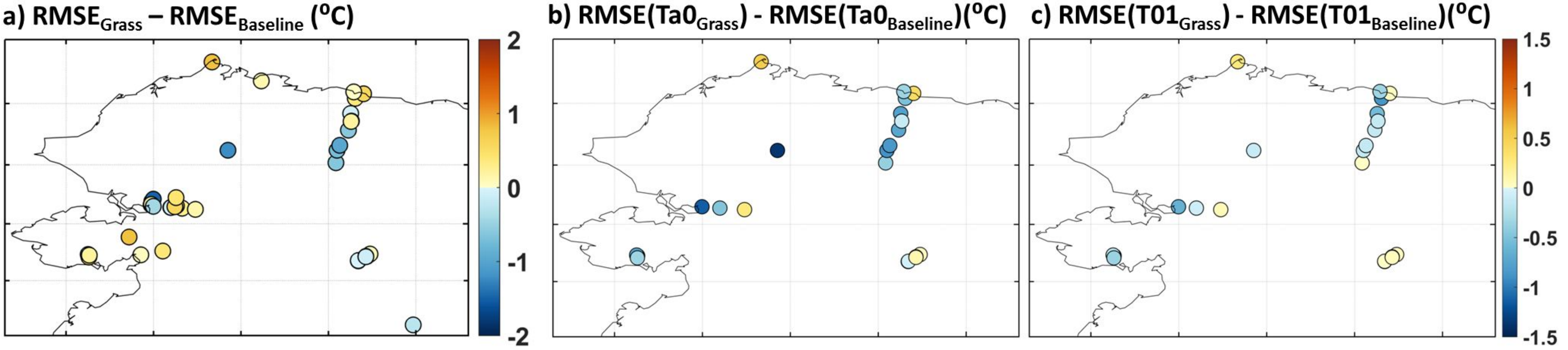

a)
b)
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6
Temperature(°C)
Sim.-Obs.(°C)
Month (2001-2001)
Month (2001-2001)
Obs.
MERRA-2 Forcing
Station-based Forcing

Figure 7.



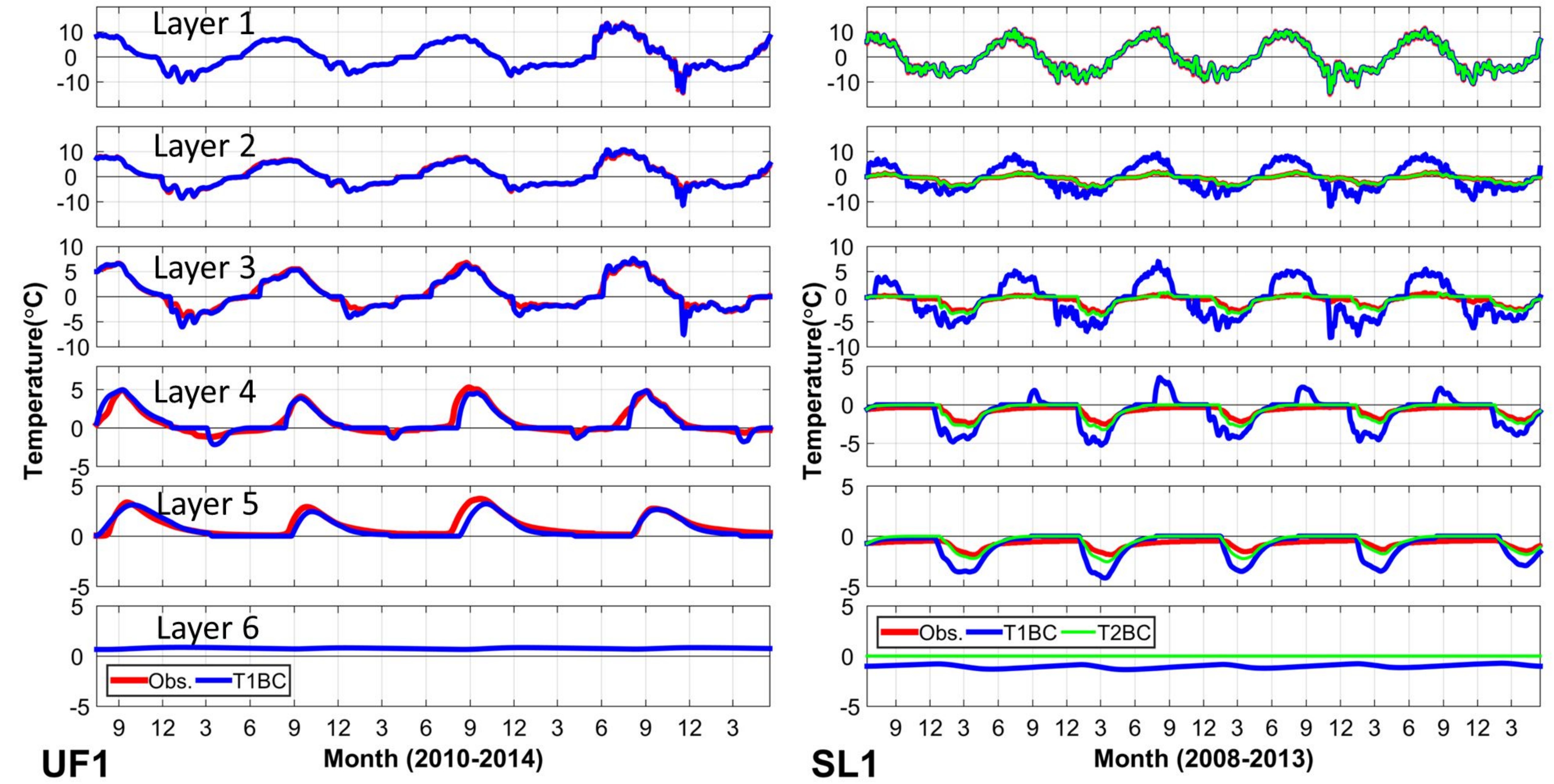
Snow Depth (m)
1
0.5
0
Layer 1
10
-5
-20
Layer 2
Layer 3
0
-10
Layer 4
Layer 5
5
Layer 6
Soil Temperature (ºC)
Obs.
Tree
Grass
12
6
Month (2007-2014)

Figure 8.



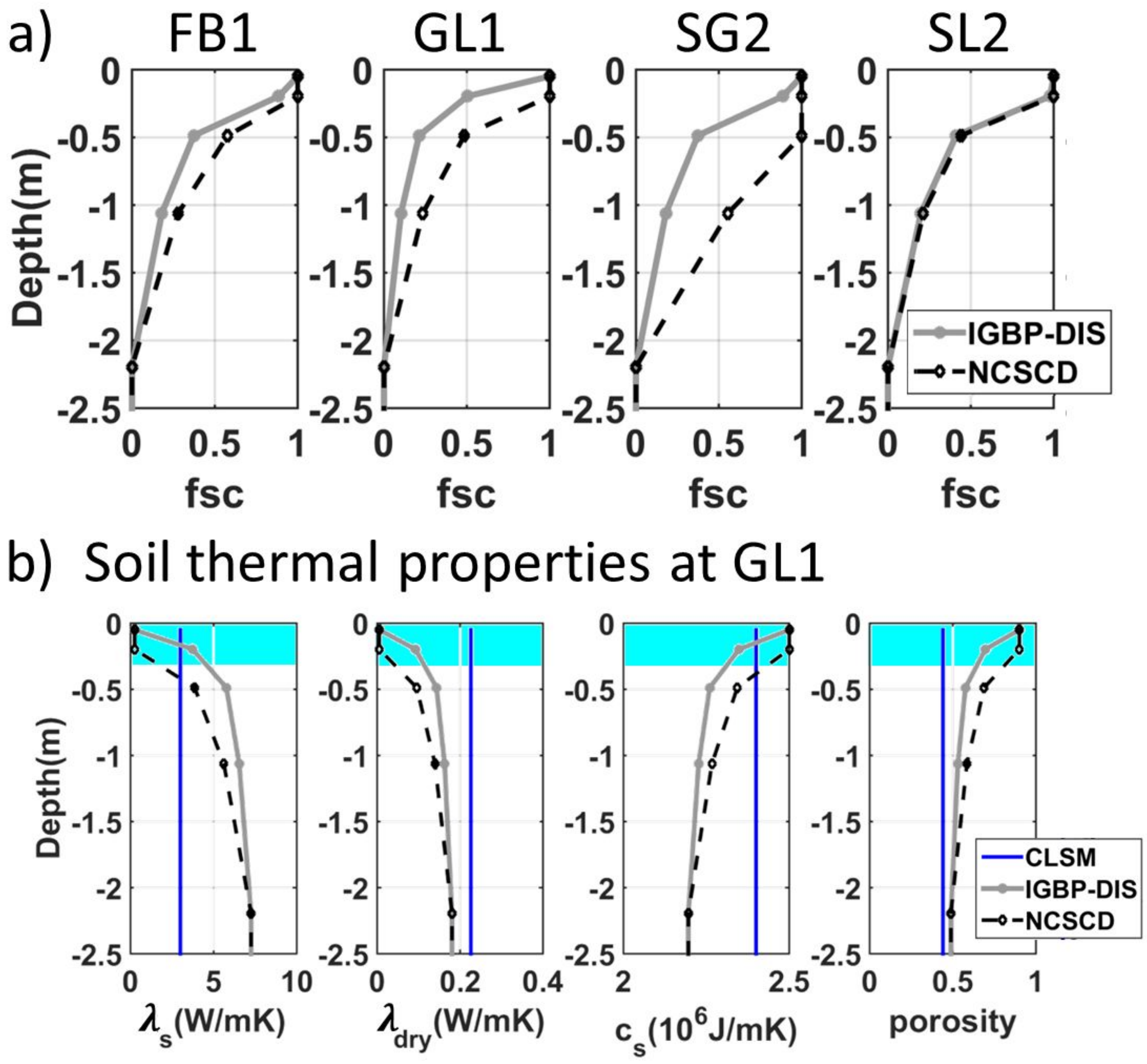
a) RMSE$_{Grass}$ – RMSE$_{Baseline}$ (^oC)
b) RMSE(Ta0$_{Grass}$) - RMSE(Ta0$_{Baseline}$)(^oC)
c) RMSE(T01$_{Grass}$) - RMSE(T01$_{Baseline}$)(^oC)
2
1
0
-1
-2
1.5
1
0.5
0
-0.5
-1
-1.5

Figure 9.



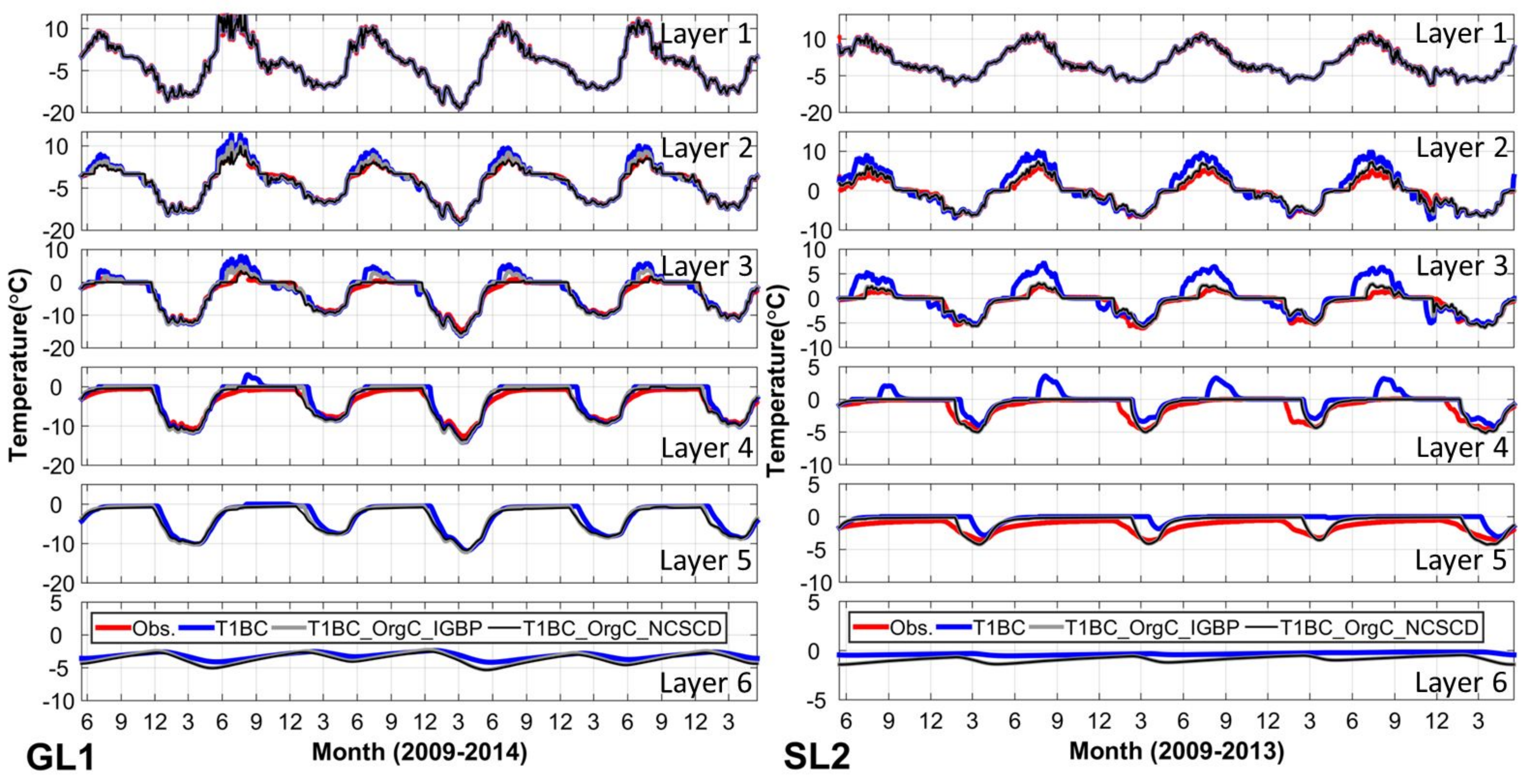
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6
Obs.
T1BC
T2BC
Temperature(°C)
Month (2010-2014)
Month (2008-2013)
UF1
SL1

Figure 10.



a)
FB1
GL1
SG2
SL2
Depth(m)
fsc
IGBP-DIS
NCSCD
b) Soil thermal properties at GL1
λ_s(W/mK)
λ_{dry}(W/mK)
c_s(10^6J/mK)
porosity
CLSM

Figure 11.



Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6
Temperature(°C)
Obs.
T1BC
T1BC_OrgC_IGBP
T1BC_OrgC_NCSCD
Month (2009-2014)
GL1
Month (2009-2013)
SL2

Figure 12.

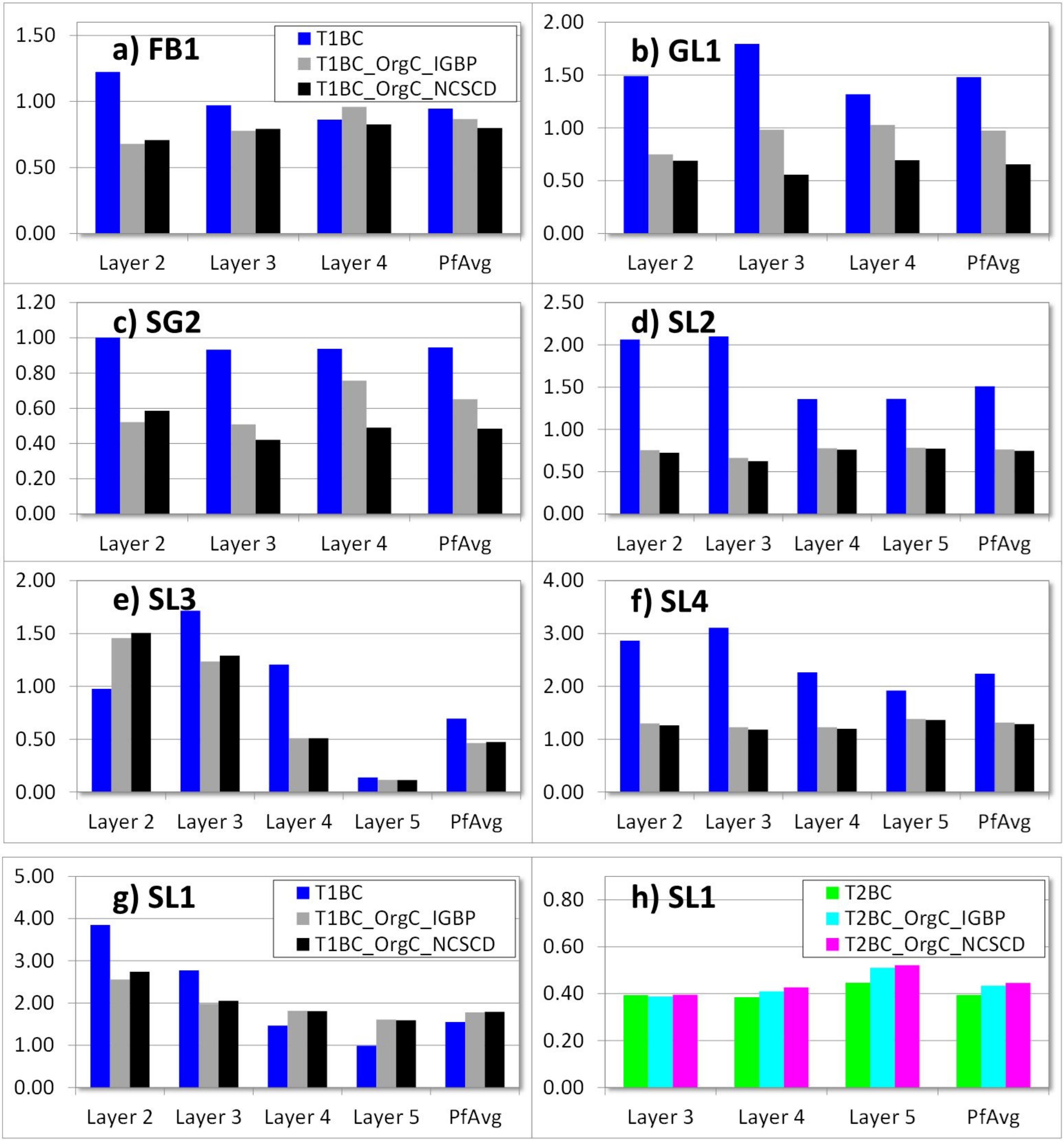
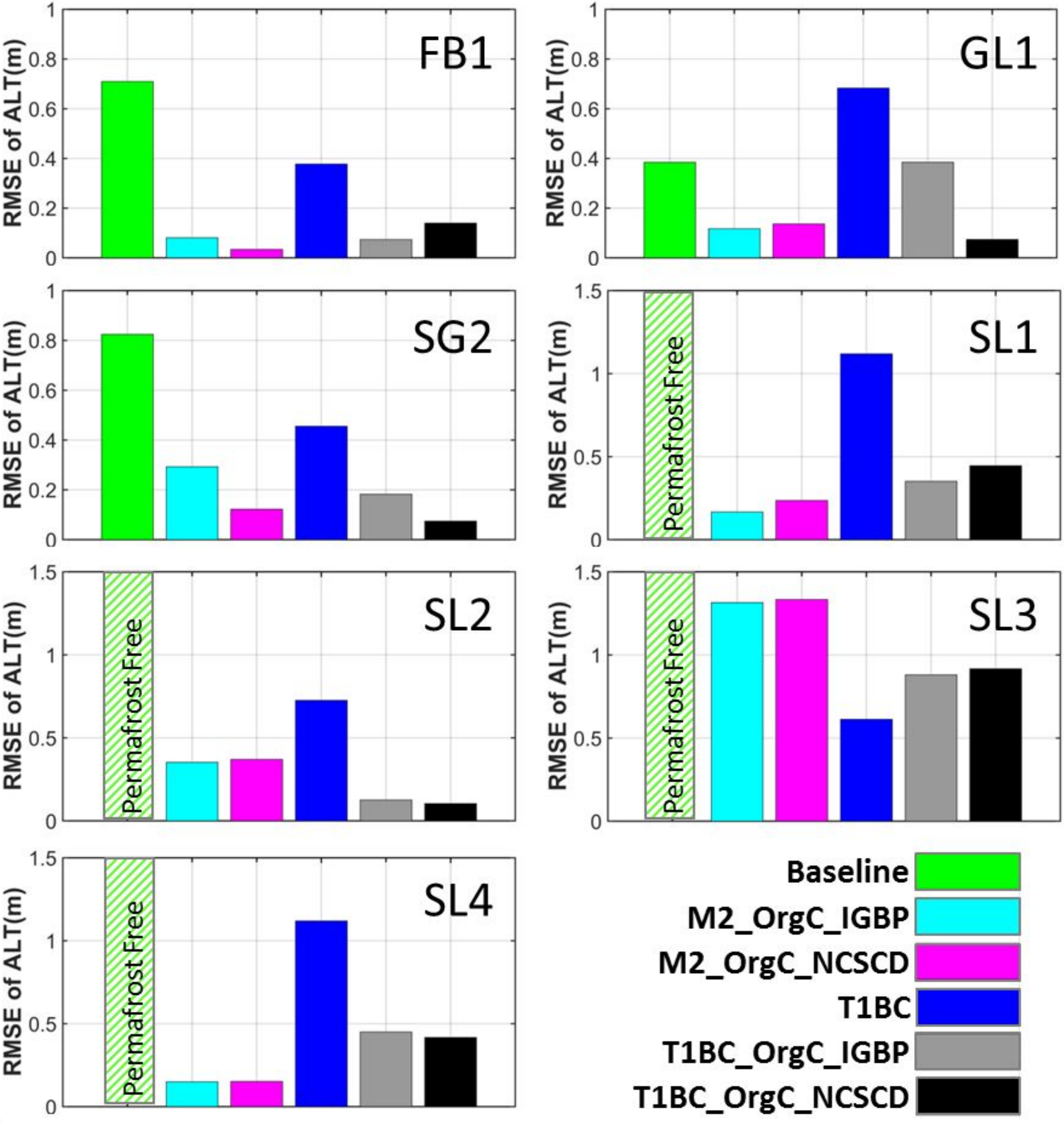
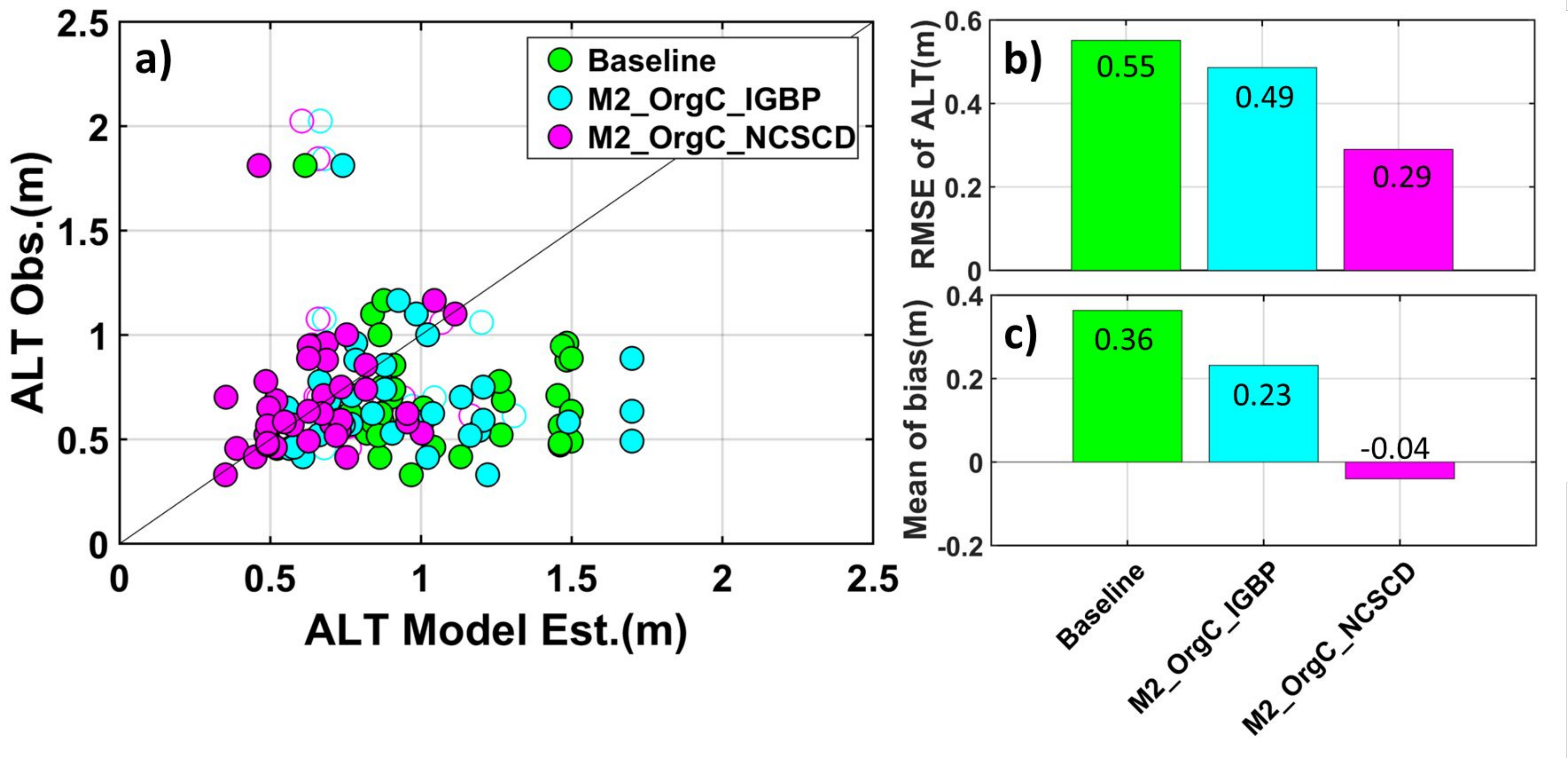

a) FB1
b) GL1
c) SG2
d) SL2
e) SL3
f) SL4
g) SL1
h) SL1
RMSE(ºC)
T1BC
T1BC_OrgC_IGBP
T1BC_OrgC_NCSCD
T2BC
T2BC_OrgC_IGBP
T2BC_OrgC_NCSCD
Layer 2
Layer 3
Layer 4
Layer 5
PfAvg

Figure 13.



FB1
GL1
SG2
SL1
SL2
SL3
SL4
RMSE of ALT(m)
Permafrost Free
Baseline
M2_OrgC_IGBP
M2_OrgC_NCSCD
T1BC
T1BC_OrgC_IGBP
T1BC_OrgC_NCSCD

Figure 14.



a)
b)
c)
Baseline
M2_OrgC_IGBP
M2_OrgC_NCSCD
ALT Obs.(m)
ALT Model Est.(m)
RMSE of ALT(m)
Mean of bias(m)
0.55
0.49
0.29
0.36
0.23
-0.04