Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale

GUO-YUE NIU AND ZONG-LIANG YANG

Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas

(Manuscript received 20 September 2005, in final form 14 January 2006)

ABSTRACT

The presence of ice in soil dramatically alters soil hydrologic and thermal properties. Despite this important role, many recent studies show that explicitly including the hydrologic effects of soil ice in land surface models degrades the simulation of runoff in cold regions. This paper addresses this dilemma by employing the Community Land Model version 2.0 (CLM2.0) developed at the National Center for Atmospheric Research (NCAR) and a simple TOPMODEL-based runoff scheme (SIMTOP). CLM2.0/SIMTOP explicitly computes soil ice content and its modifications to soil hydrologic and thermal properties. However, the frozen soil scheme has a tendency to produce a completely frozen soil (100% ice content) whenever the soil temperature is below 0°C. The frozen ground prevents infiltration of snowmelt or rainfall, thereby resulting in earlier- and higher-than-observed springtime runoff. This paper presents modifications to the above-mentioned frozen soil scheme that produce more accurate magnitude and seasonality of runoff and soil water storage. These modifications include 1) allowing liquid water to coexist with ice in the soil over a wide range of temperatures below 0°C by using the freezing-point depression equation, 2) computing the vertical water fluxes by introducing the concept of a fractional permeable area, which partitions the model grid into an impermeable part (no vertical water flow) and a permeable part, and 3) using the total soil moisture (liquid water and ice) to calculate the soil matric potential and hydraulic conductivity. The performance of CLM2.0/SIMTOP with these changes has been tested using observed data in cold-region river basins of various spatial scales. Compared to the CLM2.0/SIMTOP frozen soil scheme, the modified scheme produces monthly runoff that compares more favorably with that estimated by the University of New Hampshire–Global Runoff Data Center and a terrestrial water storage change that is in closer agreement with that measured by the Gravity Recovery and Climate Experiment (GRACE) satellites.

1. Introduction

Frozen soil occupies 55%–60% of the land surface of the Northern Hemisphere in winter (Zhang et al. 1999). It has tremendous impacts on ecosystem diversity and productivity. Seasonal freezing and thawing of soil can affect the decomposition of organic substances and the migratory patterns and physiology of biota living in the soil. Greenhouse gases released from the land surface can be dramatically increased after the spring thaw. Frozen soil also plays an important role in the climate system by altering soil thermal and hydrological properties. Freezing of soil water delays the winter cooling of the land surface, and thawing of the frozen soil delays the summer warming of the land surface (Poutou et al. 2004). Frozen soil also affects the snowmelt runoff and soil hydrology by reducing the soil permeability. Runoff from the Arctic river systems is about 50% of the net flux of freshwater to the Arctic Ocean (Barry and Serreze 2000). This is a large percentage when compared to the freshwater inputs to the tropical oceans, where freshwater input is dominated by precipitation. Runoff affects ocean salinity and sea ice conditions (McDonald et al. 1999; Peterson et al. 2002). The degree of surface freshening can affect the global thermohaline circulation (Aagaard and Carmack 1989; Broecker 1997). A realistic representation of the thermal and hydraulic properties of frozen soil will benefit global and regional climate studies.

Land surface models (LSMs) for use in climate studies showed much larger scatter in simulating runoff and soil moisture in spring than in other seasons at Valdai, Russia, due to the uncertainties in representing the ef-
fects of frozen soil on infiltration and runoff generation (Luo et al. 2003). Even when given the same volume of snowmelt water as input, different LSMs produced totally different runoff in terms of timing and magnitude. Earlier LSMs that did not explicitly solve soil ice content parameterized the hydraulic conductivity as a step function of soil temperature and switched off infiltration for subfreezing temperatures (Xue et al. 1991; Yang and Dickinson 1996). This treatment failed to produce the spring peaks of soil moisture due to the underestimated infiltration of snowmelt water (Robock et al. 1995; Xue et al. 1996). Xue et al. (1996) improved their LSM’s ability to simulate the spring peaks of soil moisture by gradually decreasing the hydraulic conductivity at a rate of 10% per degree for subfreezing temperatures following SiB2 (Sellers et al. 1996). Pitman et al. (1999) implemented an explicit representation of the hydrological and thermal effects of soil ice in their LSM but found that the representation degraded runoff simulation in a large-scale river basin. They suggested LSMs should not include the effects of soil ice on runoff until more observations at various scales are available.

Field studies in the literature also revealed conflicting results about the effects of frozen soil on infiltration and runoff generation. The field studies in local-scale open areas showed that the soil infiltration capacity is normally reduced by the presence of ice, which may generate considerable surface runoff and decrease the underlying groundwater recharge (Dunne and Black 1971; Kane and Stein 1983; Bengtsson et al. 1992; Stadler et al. 1996). However, Russian laboratory and field studies in 1960s and 1970s showed that there are weak or no clear effects of frozen soil on infiltration and runoff generation especially in forested areas. Koren (1980) used these results to formulate the frozen ground component of a rainfall–runoff model. More recent field studies in forested areas supported the Russian studies (Shanley and Chalmers 1999; Nyberg et al. 2001; Lindstrom et al. 2002; Bayard et al. 2005). Shanley and Chalmers (1999) showed that the effects of frozen soil on runoff are scale dependent. There was no significant correlation between seasonal runoff ratios and ground frost depth for the 15 yr of record from the Sleepers River watershed, United States, with an area of 111 km², while the increased runoff due to frozen ground was observed occasionally in its 0.59-km² agricultural subcatchment. Lindstrom et al. (2002) also concluded that there were no clear effects of frozen soil on the timing and magnitude of runoff from an analysis of 16-yr data in a 0.5-km² watershed in northern Sweden. Researchers (Stadler et al. 1997; Stähli et al. 1999; Nyberg et al. 2001) have demonstrated that soil structure, air-filled porosity, ice content, and the number of freeze-thaw cycles are the governing factors affecting the infiltration capacity of frozen soil. Even at very local scales, recent laboratory and field studies using dye tracer techniques (Flury et al. 1994; Stadler et al. 2000; Stähli et al. 2004) revealed that water can infiltrate into deeper soil through preferential pathways where air-filled macropores exist at the time of freezing.

One-dimensional numerical models using the fully coupled heat and mass balance equations (Flerchinger and Saxton 1989; Zhao and Gray 1997; Cox et al. 1999; Koren et al. 1999; Stähli et al. 2001; Cherkauer and Lettenmaier 2003; Hansson et al. 2004) showed a variety of ways to parameterize the hydraulic properties of frozen soil. Most of these models introduced the concept of supercooled soil water, which is the liquid water that coexists with ice over a wide range of temperatures below 0°C, by applying the freezing-point depression equation. Some of the models (Flerchinger and Saxton 1989; Cox et al. 1999; Hansson et al. 2004) assumed that the freezing–thawing process is similar to the drying–wetting process with regard to the dependence of the soil matric potential on the liquid water content. This assumption leads to a very low infiltration rate or even upward water movements resulting in ice heave in surface layers (Hansson et al. 2004). Spaans and Baker (1996) demonstrated the validity of this assumption. But it may be only applicable to very local scale, fine-textured soil. However, other modelers (Zhao and Gray 1997; Koren et al. 1999; Stähli et al. 2001; Cherkauer and Lettenmaier 2003) proposed quite different schemes to compute hydraulic properties as a means to produce greater infiltration rates. Stähli et al. (2001) proposed two separate domains for the water infiltration into frozen soil: the low-flow domain where water flows through the liquid water film absorbed by the soil particles and the high-flow domain where water flows through the air-filled macropores. Koren et al. (1999) assumed that frozen soil is permeable due to soil structural aggregates, cracks, dead root passages, and worm holes, and used the total water content (frozen and unfrozen) in the Clapp–Hornberger relationships in a land model for use in a weather prediction model. Cherkauer and Lettenmaier (2003) assumed that surface water tends to find areas of higher infiltration capacity as it flows across a frozen surface. They split their model domain into 10 bins, each having different ice content that they derived from the observed spatial distribution of soil temperature, to increase the infiltration rate in a macroscale hydrologic model. This treatment performs fairly well in large-scale cold-region rivers (Su et al. 2005).

The National Center for Atmospheric Research
(NCAR) Community Land Model (CLM) (Oleson et al. 2004) introduced a frozen soil scheme in which the freezing–thawing processes are analogous to those in snow (appendix A). The ice fraction of a soil layer becomes 100% when the heat content is sufficient to freeze all the liquid water. In addition, the hydraulic conductivity for frozen soil is parameterized as a function of the liquid water using the Clapp–Hornberger relationships (Clapp and Hornberger 1978). When there is no liquid water, the soil permeability becomes so low that most of the snowmelt water flows laterally as surface runoff, resulting in a much earlier and higher-peaked spring runoff than that estimated by the University of New Hampshire–Global Runoff Data Center (UNH-GRDC). In this study, we modified the frozen soil scheme by introducing the supercooled soil water and a fractional permeable area, which is parameterized as an exponential function of the ice content to increase infiltration rate.

2. Model description

In this study, we use a modified version of CLM2.0, in which a simple TOPMODEL-based runoff scheme (SIMTOP; appendix B) (Niu et al. 2005) was implemented into the standard CLM2.0 (Bonan et al. 2002). CLM2.0 computes soil temperature and soil water in 10 soil layers to a depth of 3.43 m. The thermal and hydraulic properties of frozen soil of the standard CLM2.0 are described in detail in appendix A. Here we describe the modifications to the CLM2.0 frozen soil scheme.

a. Implementation of the supercooled soil water

When soil water freezes, the water closest to soil particles remains in liquid form due to the absorptive and capillary forces exerted by soil particles. This supercooled soil water at subfreezing temperatures is equivalent to a depression of the freezing point. We show in the followings how a freezing-point depression equation is derived and how this equation relates to that in Koren et al. (1999).

When ice is present, soil water potential remains in equilibrium with the vapor pressure over pure ice. When neglecting the soil water osmotic potential, the soil water matric potential \( \psi(T) \) (mm) for each soil layer is

\[
\psi(T) = \frac{10^3 L_f (T - T_{frz})}{gT},
\]

where \( T \) and \( T_{frz} \) are soil temperature and freezing point (K), respectively (Fuchs et al. 1978); \( L_f \) is the latent heat of fusion (J kg\(^{-1}\)); and \( g \) is the gravitational acceleration (m s\(^{-2}\)). Spaans and Baker (1996) demonstrated that freezing–thawing processes are similar to drying–wetting processes with regard to the dependence of the soil matric potential on liquid water content. The soil matric potential as a function of liquid water content, \( \psi(\theta_{liq}) \), is

\[
\psi(\theta_{liq}) = \psi_{sat} \left( \frac{\theta_{liq}}{\theta_{sat}} \right)^{-b},
\]

where \( \theta_{sat} \) and \( \theta_{liq} \) are porosity and the partial volume of liquid water, \( \psi_{sat} \) (mm) is the saturated soil matric potential depending on the soil texture, and \( b \) is the Clapp–Hornberger parameter. By equating \( \psi(T) \) in Eq. (1) to \( \psi(\theta_{liq}) \) in Eq. (2), we derive the expression for the freezing-point depression equation:

\[
\theta_{liq,max} = \theta_{sat} \left( \frac{10^3 L_f (T - T_{frz})}{gT} \right)^{-1/b}.
\]

where \( \theta_{liq,max} \) is the maximum liquid water when the soil temperature is below the freezing point. Several researchers (Flerchinger and Saxton 1989; Zhao and Gray 1997; Cox et al. 1999; Cherkauer and Lettenmier 2003) have used this equation in the studies of water flow in frozen soil. The above equation defines the upper limit of the liquid water content for subfreezing temperatures. Additional water may be ice depending on the available energy. Thus, the liquid water content for the next time step \((N + 1)\) follows:

\[
\theta_{liq}^{N+1} = \min(\theta_{liq,max}, \theta_{liq}^{N}),
\]

where \( \theta_{liq}^{N} \) is the total volumetric soil moisture at time step \( N \), including liquid water content and ice content. In the new formulation for frozen soil, Eq. (A10) in CLM2.0 for ice content \( \theta_{ice}^{N} \) is then modified as

\[
\theta_{ice}^{N+1} = \min(\theta_{ice}^{N} - \theta_{liq}^{N+1}, \theta_{ice}^{N} + \theta_{ice}^{R_{frz}} \Delta t).
\]

Koren et al. (1999) proposed an alternative method for representing the maximum supercooled soil water by iteratively solving the following equation, a variant of the freezing-point depression equation:

\[
(1 + 8\theta_{ice})^2 \psi_{sat} \left( \frac{\theta_{liq,max}}{\theta_{sat}} \right)^{-b} = \frac{10^3 L_f (T - T_{frz})}{gT},
\]

where the \((1 + 8\theta_{ice})^2\) term accounts for the increased interface between soil particles and liquid water due to the increase of ice crystals.

Figure 1a shows that there is a considerable amount of liquid water (0.1 m\(^3\) m\(^{-3}\) or more) when the soil temperature is well below 0°C for various clay percentages. When the soil is assumed at saturation, hence \( \theta_{ice} = \theta_{sat} - \theta_{liq,max} \), Eq. (6) produces about 0.06 m\(^3\) m\(^{-3}\) more liquid water than does Eq. (3). The parameters \( b \),
\(\psi_{\text{sat}}\) and \(\theta_{\text{sat}}\) used in computing Fig. 1a are determined as functions of sand and clay percentages (Cosby et al. 1984; Oleson et al. 2004). In the calculation, we assume the sand percentage is 10\%, that is, the actual value in the W-3 watershed of the Sleepers River. Constrained by the freezing-point depression equations [both Eqs. (3) and (6)], CLM2.0/SIMTOP produces much more liquid water than without the constraint over a wide range of subfreezing temperatures (Fig. 1b). The pattern of the relationship between the soil liquid water and soil temperature as shown in Fig. 1 is consistent with the detailed observations of Nyberg et al. (2001).

b. Parameterization of the soil hydraulic properties of a GCM grid cell

The permeability of frozen soil is mainly controlled by the air-filled macropores and the ice content in the soil, which is primarily controlled by atmospheric forcing but can be also affected by subgrid distributions of terrain (height and slope), vegetation, and snow cover. Given an area of the size of a GCM grid-cell, there exist patches that are permeable and patches that are impermeable. The snowmelt water in impermeable areas may flow laterally to permeable areas. We assume that the effects of these impermeable and permeable areas on infiltration can be linearly aggregated. We also assume that both the fractional permeable and impermeable areas share the same total soil moisture of the grid cell. With these assumptions, the water flux within the soil of a GCM grid can be expressed as

\[
q = (1 - F_{\text{frz}})q_u + F_{\text{frz}}q_{\text{frz}},
\]

where \(F_{\text{frz}}\) is the fractional impermeable (frozen) area, and \(q_u\) and \(q_{\text{frz}}\) are water flux in the unfrozen and frozen areas, respectively. Assuming that \(q_{\text{frz}} = 0\), we have

\[
q = (1 - F_{\text{frz}})q_u = -(1 - F_{\text{frz}})k_u \frac{\partial(\psi_u + \zeta)}{\partial z},
\]

where \(k_u\) and \(\psi_u\) are the hydraulic conductivity and the matric potential for unfrozen soil. Comparing the above equation to Eq. (A6), we have

\[
k = (1 - F_{\text{frz}})k_u = (1 - F_{\text{frz}})k_{\text{sat}} \left( \frac{\theta}{\theta_{\text{sat}}} \right)^{2b+3},
\]

\[
\psi = \psi_u = \psi_{\text{sat}} \left( \frac{\theta}{\theta_{\text{sat}}} \right)^{-b},
\]

where \(\theta = \theta_{\text{sat}} + \theta_{\text{ice}}\) is the total grid-cell volumetric soil moisture. The above two equations describe that frozen soil at a GCM-grid scale is permeable but with reduced infiltration and percolation rates by a factor of \((1 - F_{\text{frz}})\).

We then parameterize the fractional impermeable area as a function of soil ice content at a layer:

\[
F_{\text{frz}} = e^{-\alpha(\theta_{\text{sat}} - \theta_{\text{sat}})} - e^{-\alpha},
\]

where \(\alpha = 3.0\) is an adjustable scale-dependent parameter. As shown in Fig. 2, Eq. (11) results in a smaller fractional impermeable area than the linear function, \(\theta_{\text{ice}}/\theta_{\text{sat}}\). The weaker-than-linear relationship is configured to account for the enhanced infiltration rate due to snowmelt water flowing from impermeable areas to permeable areas. Equation (11) has a pattern similar to the relationship between the ice content and the impermeable area fraction used by Koren et al. (1999). Note that Koren et al. (1999) used the impermeable area fraction to modify their surface runoff production, while this work extends the modification to soil hydraulic properties. The surface saturated area expressed by Eq. (B2) is then modified to

\[
F_{\text{sat}} = (1 - F_{\text{frz}})F_{\text{max}} e^{-0.5/\chi} + F_{\text{frz}}.
\]
where \( f \) is the decay factor, which can be determined through sensitivity analysis or calibration against the hydrograph recession curve, and \( z \) is the grid-cell-mean water table depth, which is computed as a function of liquid soil water (Niu et al. 2005).

3. Test in the Sleepers River watershed

Subcatchment W-3 (8.4 km\(^2\)) of the Sleepers River watershed (111 km\(^2\)), located in the highlands of Vermont, provides five years of meteorological and hydrological data taken for 1969–74, which were used to evaluate the runoff schemes (Stieglitz et al. 1997; Warrach et al. 2002; Niu et al. 2005). The W-3 topography is characterized by rolling hills, and the soil is dominated by silty loams (10% clay and 10% sand). The vegetation types are approximately one-third grassland, one-third coniferous forest, and one-third deciduous forest. More details are provided by Stieglitz et al. (1997) and Warrach et al. (2002). The Sleepers River watershed receives about 1100 mm of precipitation annually, distributed fairly evenly throughout the year. Annually, 20%–30% precipitation falls as snow. Snow cover persists on average from early December to early April. The average January temperature is \(-8^\circ C\). The annual maximum frost depth ranges from 5 to 40 cm depending on land-cover types (Shanley et al. 1999).

We conducted three experiments: one with the CLM2.0/SIMTOP as described in appendixes A and B as a control run (CTRL) and two other runs with CLM2.0/SIMTOP and the modifications described in section 2b but equipped with two different freezing-point depression equations, one with Eq. (3) (NEW) and one with Eq. (6) (KOREN). We used an arbitrary initial condition with a relatively wet (0.4 m–3 m–3) and warm (283 K) soil. In deep winter, CTRL simulated almost 100% ice fraction (the ratio of ice to the total soil water), while NEW and KOREN largely reduced the ice fraction to about 70% and 50%, respectively, due to inclusion of the supercooled soil water (Fig. 3a). Because the new parameterizations for hydraulic conductivity and soil matric potential greatly increase the soil permeability in the new schemes, the water content of the upper 50 cm of soil dramatically increases from as low as 90 mm to around 170 mm (Fig. 3c) due to the increased infiltration rate (Fig. 3b) during snowmelt
events in winter. There are no obvious differences in the simulated snow surface temperature (Fig. 3d) and snow depth (Fig. 3e) between CTRL and NEW, indicating that the modified frozen soil scheme (NEW) does not affect the modeled snow surface temperature and snow depth. In the cold seasons of 1969–70 and 1973–74, when there was less snow, the frozen depth is relatively greater. It is also shown that NEW has negligible effects on the modeled frozen depth (Fig. 3e). Figure 4 shows the profiles of the total soil water, liquid water, and ice fraction. The two new runs (NEW and KOREN) produce more liquid water (Fig. 4b) and less ice fraction (Fig. 4c) in the upper 0.5 m than does CTRL. KOREN produces even more liquid water (Fig. 4b) and less ice (Fig. 4c) than NEW because of the $(1 + 8\theta_{ice})^2$ term in Eq. (6). CTRL results in an odd profile of the total soil water (Fig. 4a) because of the discontinuity of the soil matric potential at freezing point produced by Eq. (A8).

The CLM2/SIMTOP produces much higher runoff...
peaks and lower recession-period runoff during the snowmelt events in both winter and spring (Figs. 5a and 5c) due to the lower permeability of the frozen soil. The lower recession-period runoff is associated with a deeper water table, which results from the extremely low liquid soil water. However, in the new schemes, the introduction of the supercooled soil water and the higher soil permeability leads to more soil water and

Fig. 4. The 1973–74 winter-averaged profiles of (a) total soil water, i.e., liquid water plus ice, (b) liquid water, and (c) ice fraction.

Fig. 5. Comparison of modeled and observed runoff: (a), (b) 1969–70 and (c), (d) 1973–74. CTRL denotes CLM2.0/SIMTOP and NEW denotes the simulation with the modified frozen soil scheme. Also shown on top of each panel are the RMSE and correlation coefficient (R).
thus a shallower water table, which in turn produces more subsurface runoff in the recession period (Figs. 5b and 5d).

Given the fact that CLM2.0/SIMTOP with Eq. (6) produces more liquid water than does CLM2.0/SIMTOP with Eq. (3), one wonders whether the new parameterizations of the soil hydraulic properties as described in section 2b could be dropped in the KOREN run. Additional experiments show that the CLM2.0/SIMTOP equipped with Eq. (6) only still produces lower-than-observed runoff in recession period.

Because NEW and KOREN produce similar simulations of the total soil water storage and runoff, we will just use the NEW scheme in the global tests.

4. Test in the six largest river basins in cold regions

To drive the model at a global scale, we used the Global Land Data Assimilation System (GLDAS) 1° × 1° 3-hourly, near-surface meteorological data for the years 2002–04 (Rodell et al. 2004a). These forcing data are observation-derived fields including precipitation, air temperature, air pressure, specific humidity, shortwave and longwave radiation, and wind speed. The reason we chose the GLDAS forcing data is that they cover the same period during which the terrestrial water storage change measured by the Gravity Recovery and Climate Experiment (GRACE) satellites is available. The vegetation and soil parameters at 1° × 1° were interpolated from the higher-resolution raw data of CLM2.0, which were also used in the studies of Bonan et al. (2002) and Niu et al. (2005). The model-simulated 3-yr-averaged runoff and snow depth were validated against the monthly UNH-GRDC runoff climatology and the U.S. Air Force Environmental Technical Applications Center (USAF-ETAC) snow depth climatology, respectively. The UNH-GRDC monthly composite runoff dataset combined observed river discharge with output from a water balance model that was driven by observed meteorological data. This dataset preserves the accuracy of the observed discharge measurements and maintains the spatial and temporal distribution of simulated runoff, thereby providing the “best estimate” of terrestrial runoff over large domains (Fekete et al. 2000). The monthly USAF-ETAC global snow depth climatology (Foster and Davy 1988) was compiled from ground-referenced measurements of snow depth. The six largest river basins in cold regions, that is, the Lena, Yenisei, Mackenzie, Ob, Churchill–Nelson, and Amur River basins, selected in this study are mostly in Russia and North America, where Foster and Davy (1988) rated the snow depth data as having high confidence levels.

To reduce the uncertainties induced by the initial conditions of the soil moisture and temperature, we first ran the model for three years from 2002 to 2004 and saved the model prognostic variables including soil moisture and temperature at the end of the model run. We used the saved model prognostic variables as the initial conditions for another 3-yr run from 2002 to 2004.

We conducted two experiments: one with CLM2.0/SIMTOP (CTRL) and one with CLM2.0/SIMTOP equipped with the modifications to the hydraulic properties described in section 2b and the freezing-point depression Eq. (3) (NEW). In both the experiments, the runoff decay factor \( f = 4.0 \, \text{m}^{-1} \). The 3-yr averaged runoff simulated from the NEW experiment greatly improves over that from CLM2.0/SIMTOP for all the six river basins (Fig. 6). CTRL consistently produced earlier and higher peaks in spring (March, April, and May) and less runoff in summer than the UNH-GRDC runoff climatology. This is mainly because the low soil permeability induced by the excessive surface soil ice content enhances surface runoff in spring. However, the snow depth simulated by CTRL is very similar to the USAF-ETAC snow depth climatology except for Yenisei River and Ob River basins, where snowmelt occurs even later than observations (Fig. 6). This excludes the possibility that the higher and earlier peaks of the simulated runoff result from earlier snowmelt. The modified scheme (NEW) largely improves the runoff simulation both in magnitude and seasonality (Fig. 6). However, these modifications to the frozen soil scheme do not affect the simulation of snow depth (Fig. 7).

As mentioned before, the UNH-GRDC runoff dataset is a composite product of runoff simulated by a water balance model constrained by those disaggregated from the observed river discharges including R-ArcticNet (see http://www.r-arcticnet.sr.unh.edu/v3.0/main.html). Although a no-time-delay assumption is applied when the gauge-observed discharge is distributed uniformly over a catchment, the resulting runoff fields over a large river basin may approximate the real runoff when there are adequate gauges within the large river basin. This may explain why UNH-GRDC runoff occurs earlier than the R-ArcticNet river discharges with greater values in spring (especially in May) consistently in Lena, Yenisei, Mackenzie, and Ob Rivers (Fig. 6). The cold-season (December–March) runoff simulated by the modified scheme (NEW) is about twice as large as the GRDC runoff, but it is still less than the estimates derived from the R-ArcticNet river discharges most obviously in Yenisei, Mackenzie, and Ob Rivers. Ye et al. (2003) and Yang et al. (2004) provided an explanation that the higher winter R-ArcticNet
River discharges may result from a combined effect of natural changes and human activities such as operations of power plant releasing water in winter.

We also compared the model-simulated soil water storage change to the GRACE-derived terrestrial water storage change in the six river basins. The recent launch of the National Aeronautics and Space Administration (NASA)’s GRACE satellites on 17 March 2002 provides modelers with a new constraint for developing LSMs. GRACE satellites measure the earth’s gravity field with enough precision to infer changes in terrestrial water storage over sufficiently large regions (Tapley et al. 2004; Wahr et al. 2004). Rodell et al. (2004b) have demonstrated that GRACE is useful to estimate basin-scale evapotranspiration when combined with precipitation and runoff data. In this study,
we used two GRACE datasets using different filtering algorithms (Chen et al. 2005; Seo and Wilson 2005). These two datasets contain 20 months starting from August 2002 to July 2004 with four months missing in between. Because GRACE measures the variation, not the absolute value, of the terrestrial water storage, we selected the same 20 months of the modeled data as those of GRACE to compute the anomalies of the water storage in the six river basins.

Both the GRACE-derived and the modeled data show positive water storage anomalies in winter and spring and negative anomalies in summer and fall (Fig. 8). The positive anomalies can be interpreted as increasing snow water stored on the ground in winter and spring. In spring, snowmelt water in the modified scheme (NEW) infiltrates into deeper soil layers, but most of it is immediately removed through surface run-off in CTRL. Thus, the soil water storage in late spring (April–June) simulated by the modified scheme (NEW) is greater than that by CTRL. The maximum anomaly in water storage simulated by the modified scheme (NEW) occurs one month later than that by CTRL (Fig. 8). This change appears to be favorable when compared to the GRACE-derived water storage variability. However, the simulated water storage variability still shows a higher anomaly in winter and spring and a lower anomaly in summer and fall.

The wintertime terrestrial water storage is mainly controlled by the following two processes: water drainage from deep soil layers and sublimation from the ground and canopy snow surface. The overestimated soil water change in winter may be caused by the model’s underestimated snow surface sublimation and/or underestimated soil water drainage. It is most likely that the snow water might be overestimated because CLM neglects such key processes as wind-blown snow
and interception of snowfall by the canopy (Niu and Yang 2004), both of which can increase the snow surface exposed to the air and thus enhance the amount of sublimation.

The water storage in river systems is also a critical factor that affects the variability of the terrestrial water storage especially for large river basins. Snowmelt water that flows to local river systems may still remain in a large river basin for a period of time depending on the area of the basin and local slopes and contribute to the total water storage. However, the river water storage and its interactions with soil water storage are not explicitly represented in CLM2.0 although a river routing submodel is included in the model to compute river discharges. To reduce the effects of river water storage on the timing of the total water storage variations, we compared the modeled water storage with the GRACE-derived in three relatively small river basins in cold regions (Fig. 9). The modified scheme (NEW) still improves the simulation of the water storage variability in all the three small river basins, most obviously in the spring of 2004.

On the other hand, overestimation of the summer evapotranspiration can result in lower soil water storage in summer and thus may contribute to a higher water storage anomaly in winter. The uncertainties in the atmospheric forcing (e.g., precipitation and radiation) can also introduce errors in simulating the water storage variations. This indicates that the land-model simulated energy and water fluxes could be further improved when constrained by the GRACE water storage change data in conjunction with other observed runoff and snow-water data.

Human activities that affect river discharges are additional sources of uncertainties in the modeling of the water storage variability. Reservoir regulations in the
Yenisei River basin (Yang et al. 2004) and the Lena River basin (Ye et al. 2003) have significantly altered the river discharges by retaining water in summer and releasing water in winter, thereby resulting in less runoff in summer and more runoff in winter. An additional experiment that produces less runoff in summer and more runoff in winter by adjusting the runoff decay factor \([f\text{ in Eqs. (B2) and (B3)}]\) from 4.0 to 2.0 \(m^{-1}\) further improves the simulation (Figs. 8 and 9).

5. Tests of two alternative hydraulic properties

The proposed frozen soil scheme as described in previous sections consists of three parts that are closely related to each other. First, the amount of (supercooled) liquid water is governed by the freezing-point depression equation. This part is necessary in that it defines an upper limit at which soil water can remain in liquid form for temperatures below the freezing point of pure water. Without this part, as is the case in CLM2.0/SIMTOP, soil water tends to become ice whenever the soil temperature is below the freezing point of pure water. Second, the soil hydraulic properties are defined as a function of a ratio of total liquid and ice) soil moisture to soil porosity. The reason to use the total liquid and ice soil moisture here is tied intimately to the next part. Third, the hydraulic conductivity is multiplied by the fractional permeable area, which integrates the effects of soil macropores and (supercooled) liquid water on infiltration at the soil surface and percolation between the soil layers. All these three parts are integral to the whole framework, which is shown to greatly improve the runoff simulation by increasing the infiltration capacity of frozen soil. This notwithstanding, it is of interest to see whether equally good simulations can result if the second and third parts of the new frozen soil scheme are represented differently.

Indeed, there are two other major forms of parameterizations of soil hydraulic properties that have appeared in the literature. In the first form, the soil hydraulic properties are expressed as a function of a ratio of (supercooled) liquid water only to soil porosity (Flerchinger and Saxton 1989, hereafter referred to as FS; Cox et al. 1999; Hansson et al. 2004):

\[
k = k_{sat}(\theta_{liq}/\theta_{sat})^{2b+3},
\]

\[
\psi = \psi_{sat}(\theta_{liq}/\theta_{sat})^{-b},
\]

In the second form, the soil hydraulic properties are expressed as a function of a ratio of (supercooled) liquid water to effective soil porosity, defined as soil porosity minus ice content (Zhao and Gray 1997, hereafter ZG):

\[
k = 10^{-E_{hloe}k_{sat}(\theta_{liq}/\theta_{sat})^{2b+3}},
\]

\[
\psi = \psi_{sat}(\theta_{liq}/\theta_{sat})^{-b},
\]

where \(E\) is the impedance factor accounting for the effect of ice on soil permeability, and \(E = 7.0\).

Both forms do not employ the concept of the fractional permeable area. To evaluate the hydrologic impacts of using the above two forms in CLM2.0/SIMTOP, we conducted two separate runs. In the experiment to test the FS scheme, we replaced Eqs. (9) and (10) with Eqs. (13) and (14), respectively. In the experiment to test the ZG scheme, we replaced Eqs. (9) and (10) with Eqs. (15) and (16), respectively. Both schemes produce earlier- and higher-than-observed runoff in all the six river basins (Fig. 10). Additionally, both schemes produce a much smaller infiltration capacity in the soil than the proposed scheme as described in section 2b. We conclude that replacing the second and third parts in the proposed frozen soil scheme with either FS or ZG fails to produce realistic simulations of runoff. The ZG scheme with a smaller impedance factor (e.g., \(E = 2.0\)) did a good job (figures not shown). However, the impedance factor is smaller than the value reported in literature by a factor of 3.5–5.0 (Takata 2002).

These tests have demonstrated that the second and third parts in the proposed scheme are superior to either the FS or ZG schemes. In the proposed scheme, hydraulic properties are expressed by the total soil water content, thereby greatly increasing the infiltration capacity of the frozen soil. Furthermore, given a frozen ground as large as a GCM grid cell, it is not uncommon to see the presence of macropores and/or snowmelt water flowing from impermeable to permeable areas. As such, it appears to be reasonable to assume a fraction of the frozen ground that is permeable. The fractional permeable area is parameterized such that it allows the infiltration capacity of the frozen ground to decrease more slowly than Eqs. (13) or (15) do.

6. Conclusions

Soil hydrologic and thermal properties are altered dramatically by the presence of soil ice. This study shows how a land surface model that explicitly represents the hydrologic effects of soil ice can improve the simulations of runoff and soil water storages by employing CLM2.0/SIMTOP. A frozen soil scheme, in which the soil ice content is solely determined by the available energy, as in CLM2.0/SIMTOP, should be modified to take into consideration the hydrologic effects of the supercooled soil water and relax the dependence of hydraulic properties on the soil ice content.
We proposed a modified frozen soil scheme that can produce more accurate magnitude and seasonality of runoff and soil water storage. These modifications include 1) an introduction of supercooled soil water by implementing the freezing-point depression equation and 2) a new parameterization of the hydraulic conductivity and soil matric potential for frozen soil under the assumption that a fractional permeable area exists in a GCM grid cell, through which water can infiltrate via air-filled macropores and unfrozen pores.

The comparison between CLM2.0/SIMTOP and the modified frozen soil schemes has been conducted using observed data from a small watershed and from the six largest river basins in cold regions (Lena, Yenisei, Mackenzie, Ob, Churchill–Nelson, and Amur) using the GLDAS observation-derived atmospheric forcing data. We draw conclusions as follows: 1) CLM2.0/SIMTOP produces excessive soil ice content, which results in little or no liquid water in the soil, and hence extremely low permeability. As such, the model produces spring runoff earlier and greater than the UNH-GRDC observations. 2) The modified frozen soil scheme produces monthly runoff that compares more favorably with the UNH-GRDC estimates and simulates a terrestrial water storage change that is more consistent with the GRACE-derived estimates. 3) The supercooled soil water governed by the freezing-point depression equation is a must in the proposed scheme, which effectively reduces the excessive presence of ice content in CLM2.0/SIMTOP. 4) The proposed parameterization of the hydraulic properties for frozen soil under the assumption that a fractional permeable area exists in a GCM grid cell is shown to be superior to the other schemes reported in the literature. 5) The GRACE-derived terrestrial water storage change provides us with a new constraint for developing land models because it offers insights into the processes that control the terrestrial water storage variability.

Acknowledgments. This work was funded by NASA Grant NAG5-10209, NAG5-12577, and NOAA Grant NA03OAR4310076. The authors thank Lindsey E. Gulden for her comments and suggestions. Matthew Rodell is thanked for providing us with the GLDAS observation-derived forcing data. Marc Stieglitz is thanked for providing us with the forcing data and parameters of the Sleepers River Watershed. Jianli Chen and K.-W. Seo are thanked for providing us with the GRACE-derived data. The comments from three anonymous reviewers essentially led to the improvement of this paper. The computing resources are provided by the Texas Advanced Computing Center (TACC).

APPENDIX A

The Coupled Heat Transport and Water Flow in the NCAR CLM

a. Heat transport

The principle of energy conservation in the form of continuity equation is invoked as
\[ C \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left( \lambda \frac{\partial T}{\partial z} \right) + \rho_{\text{ice}} L_f \frac{\partial \theta_{\text{ice}}}{\partial t}, \quad (A1) \]

where \( t \) and \( z \) are time and height above some datum in the soil column (positive upward), respectively, and \( C \) and \( \lambda \) are the volumetric heat capacity and the thermal conductivity, respectively. The second term on the right-hand side is the rate of energy released from freezing or consumed by melting. Here \( \rho_{\text{ice}} \) (917 kg m\(^{-3}\)) and \( \theta_{\text{ice}} \) are the density and partial volume of ice content, respectively, and \( L_f \) is the latent heat of fusion \((0.3336 \times 10^6 \text{ J kg}^{-1})\).

The volumetric heat capacity \( C \) (J m\(^{-3}\) K\(^{-1}\)) is from de Vries (1963) and depends on the volumetric heat capacities of the soil matrix, \( C_{\text{soil}} \), liquid water, \( C_{\text{liq}} \), and ice constituents, \( C_{\text{ice}} \):

\[ C = C_{\text{soil}}(1 - \theta_{\text{sat}}) + C_{\text{ice}} \theta_{\text{sat}} + C_{\text{liq}} \theta_{\text{liq}}, \quad (A2) \]

where \( \theta_{\text{sat}} \) is the porosity. \( \theta_{\text{liq}} \) is the partial volume of liquid water.

Soil thermal conductivity \( \lambda \) (W m\(^{-1}\) K\(^{-1}\)) is from Farouki (1981):

\[ \lambda = K_e \lambda_{\text{sat}} + (1 - K_e) \lambda_{\text{dry}}, \quad (A3) \]

where \( K_e \) is the Kersten number and \( \lambda_{\text{dry}} \) is the thermal conductivity of dry soil as a function of soil bulk density (see Oleson et al. 2004 for detail). The saturated thermal conductivity \( \lambda \) (W m\(^{-1}\) K\(^{-1}\)) depends on the thermal conductivities of the soil matrix, \( \lambda_{\text{soil}} \), liquid water, \( \lambda_{\text{liq}} \), and ice constituents, \( \lambda_{\text{ice}} \):

\[ \lambda_{\text{sat}} = \lambda_{\text{soil}}(1 - \theta_{\text{sat}}) \lambda_{\text{liq}} \lambda_{\text{ice}}, \quad (A4) \]

where the thermal conductivity of soil matrix varies with the sand and clay content, and \( \lambda_{\text{liq}} = 0.59 \text{ W m}^{-1} \text{ K}^{-1} \) and \( \lambda_{\text{ice}} = 2.29 \text{ W m}^{-1} \text{ K}^{-1} \).

**b. Water transport**

The conservation of liquid water for one-dimensional vertical water flow in the soil is expressed as

\[ \frac{\partial \theta_{\text{liq}}}{\partial t} = - \frac{\partial q}{\partial z} - E - \frac{R_{\text{fm}}}{C}, \quad (A5) \]

where \( q \) is the soil water flux (mm s\(^{-1}\)), \( E \) is the evapotranspiration rate, and \( R_{\text{fm}} \) is the melting or freezing rate.

The soil water flux \( q \) is described by Darcy’s law:

\[ q = -k \frac{\partial (\psi + z)}{\partial z}, \quad (A6) \]

where \( k \) is the hydraulic conductivity (mm s\(^{-1}\)), \( \psi \) is the soil matric potential (mm), and \( z \) is the depth from the soil surface.

The hydraulic conductivity and the soil matric potential vary with volumetric soil water and soil texture based on the work of Clapp and Hornberger (1978) and Cosby et al. (1984). In frozen soil, the hydraulic conductivity and soil matric potential vary with the partial volume of liquid water:

\[ \psi = \begin{cases} 
\psi_{\text{sat}}(\theta_{\text{liq}}/\theta_{\text{sat}})^{-b} & T > T_{\text{frz}} \\
10^3 \frac{L_f(T - T_{\text{frz}})}{gT} & T \leq T_{\text{frz}},
\end{cases} \quad (A8) \]

where \( \theta_{\text{sat}} \) (mm s\(^{-1}\)) and \( \psi_{\text{sat}} \) (mm) are the saturated hydraulic conductivity and the saturated soil matric potential depending on the soil texture, \( b \) is referred to as the Clapp–Hornburger parameter, and \( g \) is the gravitational acceleration (m s\(^{-2}\)).

The conservation of the partial volume of ice is

\[ \frac{\partial \theta_{\text{ice}}}{\partial t} = R_{\text{fm}}, \quad (A9) \]

where \( R_{\text{fm}} = H_{\text{frz}}/(\rho_{\text{ice}} L_f \Delta z) \), where \( H_{\text{frz}} \) (W m\(^{-2}\)) is the energy for freezing (positive) or melting (negative). The ice content for the next time step is

\[ \theta_{\text{ice}}^{N+1} = \min(\theta_{\text{ice}}^N, \theta_{\text{ice}}^N + R_{\text{fm}} \Delta t), \quad (A10) \]

where \( \theta_{\text{ice}} = \theta_{\text{cc}} + \theta_{\text{liq}} \) is the total volumetric soil water content.

The energy for freezing or melting \( (H_{\text{frz}}) \) is assessed from the energy excess or deficit needed to change soil temperature to freezing point \( T_{\text{frz}} \):

\[ H_{\text{frz}} = C \Delta z \frac{T_{\text{frz}} - T_{\text{frz}}^{N+1}}{\Delta t}, \quad (A11) \]

where \( T_{\text{frz}}^{N+1} \) is the layer temperature resulting from all the other processes except for phase change, and \( \Delta z \) and \( \Delta t \) are the layer depth and time step. In freezing phase (when \( \theta_{\text{liq}} > 0 \) and \( T_{\text{frz}}^{N+1} > T_{\text{frz}} \), where \( T_{\text{frz}} = 273.16 \text{ K} \)), \( H_{\text{frz}} \) is limited by the latent energy released from freezing all the liquid water in a layer within a time step, that is, \( L_f \theta_{\text{liq}} \rho_{\text{liq}} \Delta z/\Delta t \) (W m\(^{-2}\)), where \( \rho_{\text{liq}} \) is the liquid water density (1000 kg m\(^{-3}\)). In melting phase (when \( \theta_{\text{ice}} > 0 \) and \( T_{\text{frz}}^{N+1} < T_{\text{frz}} \)), \( H_{\text{frz}} \) is limited by the latent energy consumed for melting all the ice in a layer within a time step, \( L_f \theta_{\text{ice}} \rho_{\text{ice}} \Delta z/\Delta t \) (W m\(^{-2}\)). The residual energy that may not be consumed by melting or released from freezing is used to warm or cool the soil layer.
APPENDIX B

A Simple TOPMODEL-Based Runoff Scheme

The runoff scheme used in the simulations presented here is a simple TOPMODEL-based runoff model (Niu et al. 2005). In SIMTOP, the saturated hydraulic conductivity \( K_{\text{sat}} \) can either be defined as a function of soil texture as in climate models or decay exponentially with soil depth as in TOPMODEL applications. In this study, \( K_{\text{sat}} \) is defined as a function of soil texture, because it is commonly defined in this way in the land model community.

The surface runoff,

\[
R_s = F_{\text{sat}} Q_{\text{wat}} + (1 - F_{\text{sat}}) \max[0, (Q_{\text{wat}} - I_{\max})],
\]

where \( Q_{\text{wat}} \) is the input of water (sum of rainfall, dewfall, and snowmelt) incident on the soil surface, and \( I_{\max} \) is the soil infiltration capacity dependent on soil texture and moisture conditions. The saturated fraction, \( F_{\text{sat}} \), is parameterized as

\[
F_{\text{sat}} = \begin{cases} 
  F_{\text{max}} e^{-0.5/\theta_e} & \theta_e \geq 0.05 \\
  1.0 & \theta_e < 0.05 
\end{cases}
\]

where \( F_{\text{max}} \) is the potential or maximum saturated fraction for a grid cell, and \( F_{\text{max}} \) is defined as the cumulative distribution function (CDF) of the topographic index when the grid-mean water table depth is zero, that is, the percent of the pixels with topographic index larger than its grid-cell or catchment-averaged value. Here \( \theta_e = \theta_{\text{sat}} - \theta_{\text{ice}} \) is the effective porosity. The decay factor, \( f_e \), can be determined through sensitivity analysis or calibration against the hydrograph recession curve; \( z_{ve} \) is the grid-cell–mean water table depth.

The subsurface runoff is parameterized as

\[
R_{sb} = R_{sb, \text{max}} e^{-f_e z_e},
\]

where \( R_{sb, \text{max}} \) is the maximum subsurface runoff when the grid-cell averaged water table depth is zero. Here \( R_{sb, \text{max}} = 1.0 \times 10^{-4} \text{ mm s}^{-1} \).

REFERENCES


