HYDROLOGICAL PROCESSES

Published online 21 December 2014 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10398

A hydrologic routing model suitable for climate-scale simulations of arctic rivers: application to the Mackenzie River Basin

R.E. Beighley,1* K. Eggert,2 C.J. Wilson,2 J.C. Rowland2 and H. Lee3

1 Civil and Environmental Engineering, Northeastern University, Boston, MA, USA
2 Los Alamos National Laboratory, Los Alamos, NM, USA
3 Civil and Environmental Engineering, University of Houston, Houston, TX, USA

Abstract:

In this study, the Hillslope River Routing (HRR) model was modified for arctic river basin applications and used to route surface and subsurface run-off from the Community Land Model (CLM) in the Mackenzie River Basin (MRB) for the period 2000–2004. The HRR modelling framework performs lateral surface and subsurface run-off routing from hillslopes and channel/floodplain routing. The HRR model was modified here to include a variable subsurface active layer thickness (ALT; permafrost) to enable subsurface water to resurface, a distributed surface storage component to store and attenuate the rapid generation of snowmelt water, compound hillslopes to account for the low relief near rivers and floodplains, and reservoir routing to complete the total surface and subsurface water storage accounting. To illustrate the new HRR model components, a case study is presented for the MRB. The basin is discretized into 5077 sub-basins based on a drainage network derived from the global digital elevation model (DEM) developed from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor on board NASA’s Terra satellite and river widths extracted from LandSat images. The median hillslope land area is 68.5 km2 with a flow length of 2.8 km. Gridded CLM surface and subsurface run-offs are remapped to the HRR model’s irregular sub-basins. The role of each new model component is quantified in terms of peak annual streamflow (magnitude and timing) at select locations and basin-wide total water storage anomalies. The role of distributed surface storage is shown to attenuate the relatively rapid generation of snowmelt water, impact the annual peak hydrograph (reduced peaks by >30% and detailed peak by >20 days), and account for 20% of the monthly total water storage anomalies averaged over the year and ranging from 14 to 25% (−10 to 30 mm) throughout the year. Although additional research is needed to dynamically link spatially distributed ALT to HRR, the role of ALT is shown to be important. A basin-wide, uniform 1 m ALT impacts the annual peak hydrograph (reduced peaks by 9% and detailed peak by 8 days) and trends in total water storage anomalies. Copyright © 2014 John Wiley & Sons, Ltd.

KEY WORDS arctic hydrology; hydrologic routing; Mackenzie River Basin

Received 13 March 2014; Accepted 4 November 2014

INTRODUCTION

This paper describes the modification and application of a topographically based river routing model, Hillslope River Routing (HRR), described in Beighley et al. (2009 and 2011), for simulating Arctic river discharges to study the impacts of climate change. Further, the work provides a framework for using a detailed topographic landscape with a coarser climate model such as the land component of US National Center for Atmospheric Research’s (NCAR) Community Earth System Model (Gent et al., 2011). The Mackenzie River Basin (MRB) was chosen for this study as a result of its representation of arctic hydrology, large size (1.8 M km2) and availability of streamflow data from the Water Survey of Canada (WSC).

Mechanistic river basin models attempt to provide a representation of run-off processes suitable for understanding river system response to changes in land cover and climate conditions for a basin. The approach is to describe run-off processes in such a way that changes in the system can be represented by changes in run-off process parameters. Typical river basin run-off models include hillslope surface and subsurface run-off processes as contributing water to the river system. As scales become finer, constrained by the limits of the underlying terrain data, the stream network of the basin is more completely described. This leads to a more highly resolved interaction between hillslope and river processes. In the arctic, the effects of melting snow and thawing permafrost as a function of both seasonal dynamics and long-term climate change are of particular interest. Further, in a basin such as

*Correspondence to: R.E. Beighley, Civil and Environmental Engineering, Northeastern University, Boston, MA, USA.
E-mail: r.beighley@neu.edu
the Mackenzie Basin, much of the water flowing from the hillslopes becomes stored by ponding in the low-gradient regions in valley bottoms. This increase in water storage has an important effect on overall basin run-off.

The overall goal of this work is to improve the simulation of river hydrographs (timing and peaks) at scales more commonly used for river analyses (i.e. individual river reaches) in global scale hydrologic models such as the US Department of Energy (DOE) and National Science Foundation (NSF) sponsored Community Land Model (CLM). This work was part of a larger effort to examine effects of scale on arctic hydrologic process modelling. Although several of the global to continental scale terrestrial hydrologic models contain lateral transfer components (e.g. CLM Lawrence et al., 2011; GLDAS with source-to-sink routing Zaitchik et al., 2010; CHARMS Goteti et al., 2008; THMBv2 Coe et al., 2008; NLDAS Mitchell et al., 2004; GLDAS Rodell et al., 2004; CLMS Koster et al., 2000, TRIP Oki and Sud, 1998; LSX Costa and Foley, 1997; WBM/WTM Vorosmarty et al., 1996; VIC-3L Liang et al., 1994 and Branstetter and Erikson 2003), their modelling frameworks are predominately devoted to grid-based, vertical column physics and provide limited hydraulic routing realism. In these models, the terrestrial surface is often discretized using the grid of the atmospheric model component, where river systems returning water to the oceans are commonly mapped to the latitude/longitude grid of the relatively coarse-scale land model (e.g. commonly ≥0.25°). Significant advancements focused specifically on improving hydraulic routing realism in large-scale applications have also been made (e.g. MOSART Li et al., 2013; CaMa-Flood Yamazaki et al., 2011; LISFLOOD-FP Biancamaria et al., 2009; HRR Beighley et al., 2009; and TRIP Decharme et al., 2008).

For the present work, the Pafstetter (PFAF) basin delineation procedure (Verdin and Verdin, 1999) is used to subdivide the land surface into elements based on natural stream network topology. This process is recursive in that finer resolution is accomplished by applying the same delineation process to the subunits of the previous iteration until a sufficient average resolution is reached or until the resolution of the underlying terrain data is reached. The resulting irregular elements are then mapped to the grid of the atmospheric/hydrologic model to provide area weighted fluxes of water to drive the HRR model.

In the arctic, spring and summer run-off occurs as the snowpack melts and as ice in the shallow soils thaws. To accommodate these processes, the HRR model was modified to include a variable shallow subsurface active layer thickness (ALT) and a surface storage component to store and attenuate the rapid generation of snowmelt water. To better represent surface slopes, especially near the relatively flat river channels, and to increase the spatial resolution of model units, the HRR model was modified to have a compound hillslope (i.e. upperslope and lowerslope). The Mackenzie Basin also contains thousands of lakes, but three specific lakes (Lake Athabasca, Great Slave Lake, and Great Bear Lake) have significant, individual impacts on river flows. To accommodate these lakes, a reservoir routing module is also added to the HRR model. The coupling of
CLM hydrology with HRR routing plus details for the HRR modifications are the focus of this paper. Note that, in this study, daily precipitation and snowmelt are obtained from CLM. However, the variable subsurface ALT is uniform throughout the basin and fixed in time. Future efforts will dynamically link the ALT to CLM.

**STUDY AREA**

This modelling effort is focused on Canada’s largest basin, the MRB, which represents approximately 20% of the country. The MRB has a contributing land area of approximately 1.8 M km² and generally flows from south to north ultimately discharging into Beaufort Sea, part of the Arctic Ocean (Figure 1). The basin has thousands of small lakes and three clearly defined large lakes: (ordered from south to north) Lake Athabasca, Great Slave Lake, and Great Bear Lake (Figure 1). Although Lake Athabasca and Great Bear Lake have large drainage areas (115 000 and 275 000 km², respectively), nearly half the MRB’s drainage area (971 000 km²) is routed through Great Slave Lake. Table I provides summary characteristics for the three lakes. In this study, explicit lake routing is only included for these lakes. The role of smaller lakes is represented through a surface storage compartment described in the succeeding texts.

<table>
<thead>
<tr>
<th>Lake</th>
<th>Contributing area (km²)</th>
<th>Surface area (km²)</th>
<th>Volume (km³)</th>
<th>Surface elev. range (masl)</th>
<th>Discharge range (1000 m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake Athabasca</td>
<td>275 000</td>
<td>7850</td>
<td>204</td>
<td>207.5–210.5</td>
<td>0–3</td>
</tr>
<tr>
<td>Great Slave Lake</td>
<td>971 000</td>
<td>27 000</td>
<td>1580</td>
<td>156.1–157.1</td>
<td>1–13</td>
</tr>
<tr>
<td>Great Bear Lake</td>
<td>115 000</td>
<td>30 800</td>
<td>2240</td>
<td>155.6–156.6</td>
<td>0.4–0.7</td>
</tr>
</tbody>
</table>

Table II. Summary statistics for 27 streamflow gauges shown in Figure 2, where the discharge statistics are based on daily data over the entire study period (2000–2004)

<table>
<thead>
<tr>
<th>Station no.</th>
<th>Drainage area (km²)</th>
<th>Mean (m³/s)</th>
<th>Median (m³/s)</th>
<th>Min (m³/s)</th>
<th>Max (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10JE002</td>
<td>341</td>
<td>6.4</td>
<td>6.4</td>
<td>6.1</td>
<td>6.6</td>
</tr>
<tr>
<td>07SA004</td>
<td>1520</td>
<td>7</td>
<td>4</td>
<td>0.8</td>
<td>105</td>
</tr>
<tr>
<td>07EA007</td>
<td>1690</td>
<td>29</td>
<td>11</td>
<td>2</td>
<td>255</td>
</tr>
<tr>
<td>10GC003</td>
<td>2050</td>
<td>11</td>
<td>2.9</td>
<td>0.1</td>
<td>274</td>
</tr>
<tr>
<td>10CB001</td>
<td>2160</td>
<td>27</td>
<td>11</td>
<td>4</td>
<td>520</td>
</tr>
<tr>
<td>07LB002</td>
<td>3160</td>
<td>20</td>
<td>18</td>
<td>14</td>
<td>32</td>
</tr>
<tr>
<td>07AA002</td>
<td>3873</td>
<td>81</td>
<td>27</td>
<td>8</td>
<td>603</td>
</tr>
<tr>
<td>10KB001</td>
<td>7400</td>
<td>152</td>
<td>122</td>
<td>32</td>
<td>1690</td>
</tr>
<tr>
<td>10FA002</td>
<td>9270</td>
<td>46</td>
<td>19</td>
<td>0.3</td>
<td>381</td>
</tr>
<tr>
<td>10HA004</td>
<td>11 200</td>
<td>119</td>
<td>0.0</td>
<td>0.0</td>
<td>1020</td>
</tr>
<tr>
<td>07TA001</td>
<td>13 900</td>
<td>43</td>
<td>42</td>
<td>27</td>
<td>67</td>
</tr>
<tr>
<td>10GB006</td>
<td>20 200</td>
<td>81</td>
<td>10</td>
<td>0.0</td>
<td>1980</td>
</tr>
<tr>
<td>10CD001</td>
<td>20 300</td>
<td>193</td>
<td>76</td>
<td>15</td>
<td>2660</td>
</tr>
<tr>
<td>07RD001</td>
<td>26 600</td>
<td>137</td>
<td>134</td>
<td>65</td>
<td>233</td>
</tr>
<tr>
<td>07CD001</td>
<td>30 792</td>
<td>98</td>
<td>85</td>
<td>31</td>
<td>424</td>
</tr>
<tr>
<td>10JA002</td>
<td>32 100</td>
<td>123</td>
<td>121</td>
<td>70</td>
<td>226</td>
</tr>
<tr>
<td>07GJ001</td>
<td>50 300</td>
<td>297</td>
<td>165</td>
<td>24</td>
<td>3490</td>
</tr>
<tr>
<td>07LE002</td>
<td>50 700</td>
<td>307</td>
<td>293</td>
<td>143</td>
<td>621</td>
</tr>
<tr>
<td>07OB001</td>
<td>51 700</td>
<td>89</td>
<td>36</td>
<td>2</td>
<td>675</td>
</tr>
<tr>
<td>10MC002</td>
<td>70 600</td>
<td>723</td>
<td>236</td>
<td>76</td>
<td>6370</td>
</tr>
<tr>
<td>07DA001</td>
<td>132 585</td>
<td>455</td>
<td>324</td>
<td>75</td>
<td>2930</td>
</tr>
<tr>
<td>10JC003</td>
<td>146 400</td>
<td>576</td>
<td>575</td>
<td>375</td>
<td>676</td>
</tr>
<tr>
<td>10ED002</td>
<td>275 000</td>
<td>2401</td>
<td>1310</td>
<td>205</td>
<td>12 600</td>
</tr>
<tr>
<td>07KC001</td>
<td>293 000</td>
<td>2012</td>
<td>1800</td>
<td>762</td>
<td>8850</td>
</tr>
<tr>
<td>07NB001</td>
<td>606 000</td>
<td>3129</td>
<td>3000</td>
<td>1400</td>
<td>6630</td>
</tr>
<tr>
<td>10GC001</td>
<td>1027 000</td>
<td>6847</td>
<td>6030</td>
<td>2040</td>
<td>19 800</td>
</tr>
<tr>
<td>10LC014</td>
<td>1679 100</td>
<td>9077</td>
<td>6100</td>
<td>2600</td>
<td>30 000</td>
</tr>
</tbody>
</table>


Daily streamflow data from 27 gauges were used to assess the model results. Figure 1 shows the locations of the gauges, and Table II lists their summary statistics. Gauges with drainage area ranging from 341 km² to 1.68 M km² were used, with approximately three gauges within each PFAF level 1 model unit. Within each level 1 unit, the three gauges represent land areas ranging from small to large (i.e. in unit 1, the three gauges drain areas of 7400, 70 600, and 1.68 M km²). The gauge data will be used to calibrate and validate simulated streamflow.

**MODELS**

**CLM**

The CLM version 4 was used to provide surface and subsurface run-off inputs. Specifically, the formulation of the CLM used in this study is CLM4SP, where SP stands for the CLM4 version with prescribed climatological satellite phenology. To access the model results, refer to [http://www.earthsystemgrid.org/search;jsessionid=BDB102AF90A2C293F23698BD370674D2?Experiment=CLM4SP](http://www.earthsystemgrid.org/search;jsessionid=BDB102AF90A2C293F23698BD370674D2?Experiment=CLM4SP). CLM provides the vertical column hydrology, as well as the energy and mass fluxes between the atmosphere and land surfaces. It calculates these fluxes using the land surface projection of the atmospheric model grid. Run-off to the oceans is calculated by the river transport model (RTM, described by Branstetter and Erikson, 2003) that gathers the water not evaporated, redistributed in the soil or lost to deeper groundwater into a river system network that is delineated on the edges of the CLM grids. The RTM does not route the river water through the network but uses a calibrated delay to achieve the hydrograph timing. The CLM simulates permafrost dynamics, in particular, the thickness of the active layer, which provides the transmissivity for subsurface water flow from the CLM grid elements. The CLM ALT was not directly used as part of this current effort but could provide an additional coupling between HRM and CLM in future work. In this work, ALTs were based on values approximated from the literature (e.g. Nixon and Taylor, 1998; Oelke et al., 2004).

**HRR**

Routing of CLM surface and subsurface run-off is performed using one-dimensional flow equations. In particular, the methods used are the following: (a) the kinematic wave method for surface and subsurface run-off from plane elements and (b) Muskingum–Cunge routing for channels and floodplains. These methods are well proven and described by numerous authors (Chow, 1959; Mahmood and Yevjevich, 1975; Maidment, 1992; Das, 2004; Samani and Shamsipour, 2004).

Relative to the methods describing HRR in Beighley et al. (2009 and 2011), this application introduces four modifications: 1 multi-sloped hillslope; 2 shallow ALT to represent the permafrost boundary (ALT); (3) surface storage compartment \( (S_d) \) and drainage rate \( (S_r) \) to represent the cumulative storage of all the small lakes and depressions distributed throughout the landscape; and (4) lake/reservoir routing model to explicitly account for the largest lakes. Figure 2 illustrates the hillslope components.

![Figure 2. Linkage between CLM and HRR models, where CLM’s surface run-off \( Q_{over} \) (\( e_{CLM} \)) and subsurface drainage \( Q_{drain} \) (\( e_{so,CLM} \)) are remapped and uniformly distributed along HRR hillslopes as surface \( (e_{HRR}) \) and subsurface \( (e_{so,HRR}) \) run-off; HRR hillslope elements showing the simulated surface and subsurface surfaces (dashed lines): shallow impermeable layer (i.e. depth of permafrost) with ability to generate Dunne flow, and unconfined subsurface system that cannot generate Dunne flow.](image)

![Figure 3. Storage-discharge relationships for Lake Athabasca, Great Slave Lake, and Great Bear Lake; maximum values are 204 km³ and 2400 m³/s for LA; 1542 km³ and 10000 m³/s for GSL; and 2200 km³ and 700 m³/s GBL.](image)
now available in HRR. Note that ALT, $S_d$, and $S_r$ are uniformly distributed throughout the hillslope but can vary between hillslopes. The parameterization of the hillslope terms are discussed in the next section. The reservoir routing is performed using a level pool approach, where storage is a function of lake discharge (i.e. Modified Puls). Figure 3 shows the storage-discharge relationships used in this study. In this application, the storage-discharge relationships (Figure 3) were approximated from stage-discharge relationships obtained from Andrishak and Hicks (2006) for Lake Athabasca, Gibson et al. (2006) for Great Slave Lake, and MacDonald et al. (2004) for Great Bear Lake. Lake storages corresponding to each stage (i.e. lake level) were determined using the constant surface areas and total storage volumes listed in Table I.

In terms of model parameterization, most of the parameters are based on the underlying terrain data (i.e. river/hillslope lengths and slopes) and the fundamental model assumption that each sub-basin (i.e. model unit) is an open book with two hillslopes draining one channel element, which can receive flow from upstream elements (Beighley et al., 2009). Figure 4 illustrates these concepts. There are three critical routing parameters: hillslope surface roughness ($N$), subsurface horizontal conductivity ($K_h$), and channel/floodplain roughness ($n$), which can be linked to soils or land cover data but generally require calibration to account for scale impacts. For example, in this study, the average hillslope length is 3 km ranging from <1 to almost 80 km. For the longer flow lengths, water is assumed to remain as lateral run-off (i.e. subsurface or surface flow uniformly distributed over the hillslope) for the entire hillslope length. In reality, it is likely that the water travels some distance as lateral run-off before discharging into a small channel. This channel would then convey the water the remaining hillslope length to the main channel. Thus, $k$ and $K_h$ must represent the combined flow velocities of the entire model unit’s hillslope length. Channel and floodplain roughness are less sensitive to scale. The addition of a shallow, impermeable layer (i.e. permafrost depth) provides the ability to generate Dunne flow that helps to limit the subsurface flow lengths but does not remove the scale effects.

**METHODOLOGIES**

**Landscape representation**

For this study, the topography is based on the global digital elevation model (DEM) derived from NASA’s Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER DEM). The ASTER DEM has horizontal postings of approximately 30 m with a reported mean elevation error of approximately 8.3 m based on analyses in 24 regions with individual regional mean errors ranging from 3.3 to 15.0 m (Tachikawa et al., 2011http://asterweb.jpl.nasa.gov/gdem.asp). Although the data are posted at 30-m intervals, Tachikawa et al. (2011) suggest that the optimal horizontal resolution is between 90 and

![Figure 4. HRR model’s hillslope and channel approximations](image)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Hillslope length (km)</th>
<th>Hillslope slope (%)</th>
<th>Lower hillslope slope (%)</th>
<th>Channel length (km)</th>
<th>Channel slope (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>&lt;1</td>
<td>0</td>
<td>0</td>
<td>&lt;1</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>2.8</td>
<td>3.7</td>
<td>0.3</td>
<td>26</td>
<td>1.4</td>
</tr>
<tr>
<td>Mean</td>
<td>3.0</td>
<td>7.5</td>
<td>1.9</td>
<td>63</td>
<td>1.7</td>
</tr>
<tr>
<td>Max</td>
<td>80</td>
<td>56</td>
<td>33</td>
<td>1780</td>
<td>13</td>
</tr>
</tbody>
</table>

Table III. Summary statistics for model unit characteristics
120 m that minimizes the standard deviation of vertical DEM errors. Given the size of the watershed and the suggested horizontal resolution, the DEM was averaged to 90 m, where the elevations for the 90-m DEM grid cells were the mean of the underlying nine 30-m grid cell elevations. Using the 90-m DEM, flow directions and flow accumulation datasets were created using ArcGIS Hydrology tools. A threshold area of ~20 km² was used to define the initial drainage network. Finally, the MRB was delineated to PFAF level 4. Figure 1 shows the MRB delineated to PFAF level 1 (i.e. units 1 to 9 per the methods described by Verdin and Verdin, 1999). The appeal of the PFAF delineation is that it results in a computational basin discretization that is topographically faithful to actual basin topography. In addition, the PFAF approach provides a sub-basin numbering system for the sub-basins that contain the topological information for forming a description of the drainage basin. Further, by resulting in a system of linked sub-basins, the method provides an inherent computational parallelization for surface water modelling. At level 4, the 1.8 M km² MRB is divided into 5077 sub-basins. Table III summarizes the landscape characteristics corresponding to the individual sub-basins referred to herein as model units.

Building on the level 4 river network, catchments contributing to each river reach and the 90 m DEM, hillslope, and channel lengths and slopes were determined (Table III). Channel length (Lc) was determined directly from the river network (Figure 4) using the GIS to calculate the length of the individual reach segments. For each catchment, the hillslope length (Lp) was determined by approximating the two hillslopes draining to the river segment (Figure 4) as two equally sized rectangles. Thus, Lp is half the catchment area divided by the length of the channel (Lp = Ao/2/Lc), where the local catchment area (Ao) is determined using ArcGIS. For each reach, channel slope (Sc) was determined by averaging the pixel-to-pixel slope along the river reach, where the pixel-to-pixel slope was the distance from a given pixel centre to the centre of the next downstream channel pixel divided by their difference in elevation. The hillslope slope (Sp) was approximated as the mean ground slope within a given catchment, where ground slope was determined using the ArcGIS slope function and the 90 m digital terrain model. Overall, the basin relief is low with a mean model unit hillslope slope of only 0.075 m/m and median slope of 0.037 m/m. Given that the hillslope dimensions are based on the open book approximation, it is not possible to sample an actual hillslope flow element. Thus, to subdivide the hillslope length as shown in Figure 2, the lower segment’s slope was assumed to be the unit’s mean slope minus one standard deviation, with a minimum value condition of 0.0001 m/m. The length of the lower slope was set to 33% of the derived hillslope length (0.33Lp). Future research will specifically investigate methods to estimate the hillslope components from all possible flowpaths within a given model unit. The intent of the previous assumptions is to build and test the model’s capabilities while providing a reasonable estimate for representing the structure of the landscape.

Given that HRR requires channel/floodplain cross sections, a combination of two approaches was used. For the main river (refer to Figure 1), river channel extents were extracted from the Landsat images using the automated feature extraction software program GeniePro (Brumby et al., 1999; Perkins et al., 2005) commercially available from Observera, Inc. GeniePro allows for the automated delineation of the full extent of a river channel (wetted and dry exposed portions of the channel bed) using all spectral bands based on a user defined set of training data. This accounts for all rivers draining more than approximately 100 000 km² and approximately 12 500 km of river length. For the smaller drainage areas (A_u < 100 000 km²), a rectangular section was assumed, and the relationship between width and drainage area was developed based on measurements of channel widths from Landsat images at 72 locations throughout the basin:

\[ w = 2.07A_u^{0.48} \quad R^2 = 0.51 \]  

where \( w \) is the combined channel/floodplain width (m) and \( A_u \) is the upstream drainage area (km²). It is also interesting to note that the coefficients and exponents developed here produce similar widths as the relationships presented in Beighley and Gummadi (2011) (\( w = 1.96A_u^{0.41} \)) for areas <10 000 km² in the Amazon Basin, illustrating the strong dependence on drainage area. Thus, the relationship developed here with limited data is assumed to be reasonable and used to approximate channel width for tributary channels. Although channel widths are critical for simulating water surface elevations, the effect of channel width error on simulated discharge is minimal (i.e. <10%) (Gummadi, 2008). River widths for all rivers greater than approximately 30–60 m are currently being developed using RivWidth (Pavelsky and Smith, 2008).

Linking CLM and HRR models

For this study, the daily CLM4SP model results are extracted from the 0.9° latitude by 1.25° longitude gridded output files for the period 2000 through 2004 to generate surface and subsurface run-off within the HRR model, with irregular model units corresponding to PFAF level 4. Figure 2 illustrates the connections between CLM and HRR. Note that, HRR does not currently provide any information back to CLM. For example, the daily model results for surface run-off (Qover), which includes snowmelt, and subsurface drainage (Qdrain), which integrates all soil layers, were extracted from each monthly CLM output file (e.g. clm4.clm2.h1.1999-12-12-00000.nc)
obtained from the NCAR website (http://www.cesm.ucar.edu/models/cesm1.0/clm/index.shtml). The run-off terms from CLM grid cells were then remapping to HRR model units and assumed to be uniformly distributed along the unit’s hillslope surface and subsurface, respectively.

Figure 5 shows the comparison between CLM grid and irregular HRR model units. Given the differences in model resolution (324 CLM pixels vs 5077 HRR units) shown in Figure 5, the CLM results were assigned to HRR model units based on the distance between HRR model unit centroid location and CLM pixel centres. The CLM pixel with the shortest straight line distance was used to assign CLM output to the HRR model unit. Note that approximately 60% of the HRR model units are completely contained within a single CLM grid cell for which the previous remapping technique provides exact area matching. For the remaining units, we assume that land areas associated with the HRR model units extending into or out of a given CLM grid cell are balanced. For example, a HRR model unit with its centroid assigned to one CLM pixel may have some land area overlapping another CLM pixel and, similarly, that CLM pixel is likely linked to other HRR model units that overlap additional CLM pixels. To check for area conservation, we cumulated the land area of all HRR model units associated with each CLM pixel and compared HRR area with CLM pixel area. The results show that the mean conservation of CLM land area is 99.7%. Although this approach does not provide exact area remapping, it eliminates any temporal averaging that may result from using an area-weighted approach to split area between multiple CLM grid cells.

Model calibration and sensitivity analysis

The surface run-off and channel routing parameters hillslope surface roughness \( (N) \) and channel/floodplain roughness (Manning’s \( n \)) were approximated based on the mostly vegetated/forested land cover of the basin (approximately 86%, Tong et al., 2010). Thus, values of 1.0 and 0.08 for \( N \) and \( n \), respectively, were used. Note that, the channel/floodplain roughness is an approximation intended to represent conditions within the channel (e.g. \( n = 0.05 \)) and on the floodplain (e.g. \( n = 0.1 \) to 0.15). Also, in the model, the Manning’s \( n \) value is increased by a factor of 3 to account for sinuosity not captured by the DEM-derived river flowpath. Lengths, slopes, and widths for all the hillslopes and channels were based on the river network derived from the ASTER DEM and assumptions described in the preceding texts. For the Muskingum–Cunge method, the reference discharge was assumed to be the mean daily discharge. Given the limited effects of reference discharge on simulating daily streamflow and the relationship between mean daily discharge and drainage area for the gauge data, the reference discharge
was approximated as 1% of the upstream drainage area for each river reach. Future efforts will explore the use of the variable parameter method (Ponce and Yevjevich, 1978). The reservoir storage-discharge relationships are shown in Figure 3. While HRR has the capability of utilizing active zone thicknesses calculated by the CLM, that coupling is reserved for future work. Nixon and Taylor (1998) reported ALTs in the Mackenzie Delta and basin uplands ranging from 38 to 182 cm, depending on location and site specific factors. Oelke et al. (2004) shows a 23-year (1980–2002) mean ALT of 142 cm for areas of continuous and discontinuous permafrost in the Mackenzie Basin. For the present work, an average of 1.0 m was assumed throughout the basin for the depth of the permafrost layer \((D_p)\). Future efforts will directly couple the space-time varying thicknesses estimated from the CLM.

The remaining three model parameters were considered in the calibration process and used in the sensitivity analysis: subsurface horizontal conductivity \((K_h)\), depth of surface storage \((S_d)\), and surface storage drain rate \((S_r)\). The calibration of the model was performed by adjusting \(K_h, S_d,\) and \(S_r\), to maximize the mean index of agreement \((\bar{d})\):

\[
\bar{d} = \frac{1}{k} \sum_{i=1}^{k} \left[ 1 - \frac{\sum_{j=1}^{m} (P_{ij} - O_{ij})^2}{\sum_{j=1}^{m} (|P_{ij} - \bar{O}_i| + |O_{ij} - \bar{O}_i|)^2} \right]
\]

where \(P_{ij}\) is the predicted discharge on day \(j\) at gauge \(i\), \(O_{ij}\) is the observed discharge on day \(j\) at gauge \(i\), \(\bar{O}_i\) is the mean observed daily discharge over the study period (2000–2004), \(m\) is the number of daily discharges in the study period \((m = 1827)\), and \(k\) is the number of gauge locations \((k = 27)\). A sensitivity analysis was also performed to show the impacts of changing the calibrated values of \(K_h, S_d,\) and \(S_r\). The simulated parameter variability was ±25% for each individual parameter (i.e. six additional model runs). Future efforts will focus on developing new methods to measure/estimate these values and their corresponding uncertainties.

The index of agreement was used in this study rather than the commonly used Nash–Sutcliffe efficiency because the calibration focused only on the routing parameters (i.e. could not alter the availability of water, which was provided from CLM). Although the index of agreement is similar to the Nash–Sutcliffe efficiency, it is less sensitive to systematic model overprediction or underprediction (Willmott et al., 1985). For example, at some gauge locations, the temporal correlation between predicted and observed discharge was good, but the magnitude of the discharges was consistently biased high or low. However, the bias could not be eliminated by changing only the routing parameters. To improve simulated discharges at these locations, calibrating both CLM and HRR would be required.

Although not used for model calibration, monthly total water storage anomalies (TWSA) from HRR and CLM are compared with TWSA derived from NASA’s Gravity Recovery and Climate Experiment (GRACE) satellite mission. For this analysis, GRACE TWSA for the Mackenzie Basins were derived using the Center for Space Research (CSR) Release 5 (RL05) data that were decorrelated (Duan et al., 2009) and smoothed using a 3-degree Gaussian filter (Guo et al., 2010). The TWSA are represented in terms of equivalent water height anomalies (Wahr et al., 1988) and are computed based on the mean TWS determined from the monthly measurements obtained during 2003, which is the first complete year of available measurements from the GRACE satellite.

**RESULTS AND DISCUSSION**

**Calibration results**

To calibrate the model, HRR parameters for reservoir routing, split slope planes, surface storage, and a specified ALT (1.0 m) were set as discussed in the preceding texts, and the calibration was performed by adjusting \(K_h, S_d,\) and \(S_r\) to maximize the mean index of agreement \((\bar{d})\). This process was performed by running the model in batch mode where individual and combinations of parameters were systematically changed over a pre-set range. Results were then analyzed, and parameter ranges were refined based on model performance. This process was repeated until the optimal parameters were found. The resulting parameter values were 0.019 m/s, 1.25 to 12.5 cm, and 1.1 mm/day for \(K_h, S_d,\) and \(S_r\), respectively. Note that \(S_d\) was spatially distributed based on the inverse of plane slope with a lower and upper limit of 1.25 and 12.5 cm, respectively. To release water from the surface store, the depth of water above \(S_d\) was used as the height \((H)\) above a rectangular weir (i.e. surface storage discharge per unit width of plane \(= 1.074H^{1.5} [\text{m}^3/\text{s/m}]\)). In addition, the calibrated value of \(K_h\) is much higher than values reported in the literature as a result of scaling effects. Here, the mean hillslope length is 3.0 km, and the subsurface flow travels that entire distance before discharging into a stream. However, it is likely that subsurface flow would discharge into a small tributary before travelling the entire length simulated in the model. The calibrated \(K_h\) value therefore represents the effective velocity of the model unit that combines both subsurface and some degree of channel flow.

The calibrated model results are summarized in Table IV (model M23). The mean index of agreement \((\bar{d})\) based on comparisons between simulated and measured discharges at the 27 gauge locations listed in Table II was 0.58. Figure 6 shows the agreement for each location. Grouping the results by region generally shows better agreement in
the south ($d = 0.61$) and middle ($d = 0.61$) parts of the basin as compared with the northern ($d = 0.43$). This is likely because of the sensitivity of the ALT in the northern most region of the basin, where ALT is smallest. This point is discussed in more detail in later sections. The locations

Table IV. Index of agreement ($d$) for all gauges; outlet gauge (10LC014; 1.7 M km$^2$) for the model sensitivity analysis consisting of 24 model simulations encompassing all combinations of four critical processes (reservoirs vs. no reservoirs; surface storage vs. no surface storage; single slope vs. split slope; and active layer thicknesses (ALT) of 0.0, 1.0 and 1000 m).

<table>
<thead>
<tr>
<th>M</th>
<th>Mass outlet</th>
<th>NSE outlet</th>
<th>$d$ outlet</th>
<th>$d$ all</th>
<th>$d$ res</th>
<th>$d$ upper</th>
<th>$d$ middle</th>
<th>$d$ lower</th>
<th>Reservoir routing</th>
<th>Surface storage</th>
<th>Hillslope slope</th>
<th>ALT (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.23</td>
<td>-0.51</td>
<td>0.73</td>
<td>0.33</td>
<td>0.43</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>N</td>
<td>N</td>
<td>Single</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1.23</td>
<td>-0.46</td>
<td>0.74</td>
<td>0.33</td>
<td>0.44</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>N</td>
<td>N</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.23</td>
<td>-0.46</td>
<td>0.74</td>
<td>0.33</td>
<td>0.44</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>N</td>
<td>N</td>
<td>Single</td>
<td>1000</td>
</tr>
<tr>
<td>4</td>
<td>1.23</td>
<td>-0.41</td>
<td>0.75</td>
<td>0.34</td>
<td>0.44</td>
<td>0.31</td>
<td>0.32</td>
<td>0.34</td>
<td>N</td>
<td>N</td>
<td>Split</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1.21</td>
<td>-0.29</td>
<td>0.76</td>
<td>0.34</td>
<td>0.45</td>
<td>0.31</td>
<td>0.33</td>
<td>0.33</td>
<td>N</td>
<td>N</td>
<td>Split</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1.16</td>
<td>-0.26</td>
<td>0.76</td>
<td>0.34</td>
<td>0.45</td>
<td>0.31</td>
<td>0.33</td>
<td>0.33</td>
<td>N</td>
<td>N</td>
<td>Split</td>
<td>1000</td>
</tr>
<tr>
<td>7</td>
<td>1.23</td>
<td>0.59</td>
<td>0.92</td>
<td>0.53</td>
<td>0.62</td>
<td>0.53</td>
<td>0.54</td>
<td>0.47</td>
<td>N</td>
<td>Y</td>
<td>Single</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1.22</td>
<td>0.60</td>
<td>0.92</td>
<td>0.54</td>
<td>0.63</td>
<td>0.55</td>
<td>0.56</td>
<td>0.46</td>
<td>N</td>
<td>Y</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1.22</td>
<td>0.60</td>
<td>0.92</td>
<td>0.54</td>
<td>0.63</td>
<td>0.55</td>
<td>0.56</td>
<td>0.46</td>
<td>N</td>
<td>Y</td>
<td>Single</td>
<td>1000</td>
</tr>
<tr>
<td>10</td>
<td>1.23</td>
<td>0.59</td>
<td>0.92</td>
<td>0.53</td>
<td>0.62</td>
<td>0.54</td>
<td>0.55</td>
<td>0.46</td>
<td>N</td>
<td>Y</td>
<td>Split</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1.20</td>
<td>0.64</td>
<td>0.91</td>
<td>0.575</td>
<td>0.69</td>
<td>0.61</td>
<td>0.61</td>
<td>0.43</td>
<td>N</td>
<td>Y</td>
<td>Split</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1.02</td>
<td>0.69</td>
<td>0.89</td>
<td>0.55</td>
<td>0.75</td>
<td>0.63</td>
<td>0.53</td>
<td>0.35</td>
<td>N</td>
<td>Y</td>
<td>Split</td>
<td>1000</td>
</tr>
<tr>
<td>13</td>
<td>1.22</td>
<td>0.18</td>
<td>0.81</td>
<td>0.37</td>
<td>0.70</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>Y</td>
<td>N</td>
<td>Single</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1.22</td>
<td>0.21</td>
<td>0.81</td>
<td>0.37</td>
<td>0.71</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>Y</td>
<td>N</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1.22</td>
<td>0.21</td>
<td>0.81</td>
<td>0.37</td>
<td>0.71</td>
<td>0.31</td>
<td>0.31</td>
<td>0.33</td>
<td>Y</td>
<td>N</td>
<td>Single</td>
<td>1000</td>
</tr>
<tr>
<td>16</td>
<td>1.23</td>
<td>0.25</td>
<td>0.83</td>
<td>0.38</td>
<td>0.71</td>
<td>0.31</td>
<td>0.32</td>
<td>0.34</td>
<td>Y</td>
<td>N</td>
<td>Split</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1.20</td>
<td>0.31</td>
<td>0.83</td>
<td>0.38</td>
<td>0.67</td>
<td>0.31</td>
<td>0.33</td>
<td>0.33</td>
<td>Y</td>
<td>N</td>
<td>Split</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>1.15</td>
<td>0.34</td>
<td>0.83</td>
<td>0.37</td>
<td>0.67</td>
<td>0.31</td>
<td>0.32</td>
<td>0.33</td>
<td>Y</td>
<td>N</td>
<td>Split</td>
<td>1000</td>
</tr>
<tr>
<td>19</td>
<td>1.23</td>
<td>0.70</td>
<td>0.93</td>
<td>0.56</td>
<td>0.80</td>
<td>0.53</td>
<td>0.54</td>
<td>0.47</td>
<td>Y</td>
<td>Y</td>
<td>Single</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>1.22</td>
<td>0.68</td>
<td>0.92</td>
<td>0.56</td>
<td>0.79</td>
<td>0.55</td>
<td>0.56</td>
<td>0.46</td>
<td>Y</td>
<td>Y</td>
<td>Single</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>1.22</td>
<td>0.68</td>
<td>0.92</td>
<td>0.56</td>
<td>0.79</td>
<td>0.55</td>
<td>0.56</td>
<td>0.46</td>
<td>Y</td>
<td>Y</td>
<td>Single</td>
<td>1000</td>
</tr>
<tr>
<td>22</td>
<td>1.23</td>
<td>0.68</td>
<td>0.92</td>
<td>0.56</td>
<td>0.79</td>
<td>0.54</td>
<td>0.55</td>
<td>0.46</td>
<td>Y</td>
<td>Y</td>
<td>Split</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>1.20</td>
<td>0.62</td>
<td>0.88</td>
<td>0.583</td>
<td>0.74</td>
<td>0.61</td>
<td>0.61</td>
<td>0.43</td>
<td>Y</td>
<td>Y</td>
<td>Split</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>1.02</td>
<td>0.61</td>
<td>0.84</td>
<td>0.54</td>
<td>0.70</td>
<td>0.63</td>
<td>0.53</td>
<td>0.35</td>
<td>Y</td>
<td>Y</td>
<td>Split</td>
<td>1000</td>
</tr>
</tbody>
</table>

For comparison, M3 represents the base HRR model, and M23 represents the fully modified HRR model, which includes reservoir routing, surface storage, split hillslopes, and an ALT = 1.0 m

Figure 6. Relationship between model performance ($d$) and drainage area (A) for the calibrated model (M23 in Table IV) based on the data from the 27 gauges listed in Table II; data points grouped based on location in the basin and regulation impacts as listed in Table II.

The locations impacted by lakes matched gauge data better than non-regulated locations ($d = 0.74$) as a result of the prescribed storage-discharge relationships used in the model (Figure 6). The location with the best agreement ($d = 0.88$) was the outlet gauge, which integrates the largest drainage area and all three lake models.

Although there is variability in model performance between gauges, Figure 6 shows that there is a positive relationship between drainage area and the index of agreement ($d$). It is important to note that the forcings for HRR are from the CLM model with a resolution of 0.9° latitude by 1.25° longitude, which means that each grid cell is greater than 10$^4$ km$^2$. Figure 6 shows that for drainage areas greater than about 10$^4$ km$^2$, the agreement is better ($d = 0.65$; only 2 of 18 or 11% with $d < 0.5$) than the agreement for smaller areas ($d = 0.47$; 4 of 9 or 44% with $d < 0.5$). These results are consistent with the resolution of CLM used here and suggest that although the local-scale landscape features are captured in the HRR model (i.e. median hillslope scale is 68.5 km$^2$) that local-scale variability in precipitation and snowmelt is important.

Exploring results at individual gauge locations (Figure 7) highlights four key findings. One, agreement
of daily discharges for individual gauges shows that short
duration peaks are not captured well for gauges draining
large areas (Figure 7a–c). For the outlet gauge, the model
underestimates the peak discharges in 2001, 2002, and
2004 by 25, 31, and 36%, respectively, but is within
approximately 10% of the annual peak in 2000 and 2003
that have smaller, short duration peaks associated with the
annual flood wave. Similar findings are seen at the other
gauges. The year-to-year change in model performance is
likely because of a combination of uncertainty in CLM
output (i.e. specifically the accumulated snowpack and
resulting melt rates) and threshold effects from the newly
added model elements related to surface storage, reservoir
routing and ALT. For example, Figure 7e shows model
results that are sensitive to the depth of surface storage. In
2000, 2003, and 2004, run-off from CLM exceeds the
surface storage capacity and produces a rapid discharge
response similar to the pattern observed in the gauge data.
However, in 2001 and 2002, the storage capacity appears
to be just at or slightly above capacity that does not result
in a burst of discharge as seen in the other years. This
point supports not only the model formulation but also the
importance of estimating the correct storage capacity.

The second finding is that the model tends to
underestimate peak flood discharge in the northern half
of the basin while overestimating in the southern half of
the basin as a result of ALT assumptions. For example,
comparing model results for the two largest gauges
(Figure 7a and b) in 2000 and 2003 shows that the model
overestimated the flood peak at gauge 10GC001

Figure 7. Comparison between the calibrated model (model 23) and gauge data for the period 2000–2004 with mean daily flows for that period to the
right of the vertical dashed line: (a) near the outlet of the basin, (b) along the main channel, (c) a main stem tributary, and (d–f) three smaller secondary
tributaries; shaded box in panel (c) highlights a region of poor model performance discussed in text; refer to Figure 1 and Table II for gauge information.
(1.0 M km²), which drains the southern half of the basin but matched the peak at the outlet gauge 10LC014 (1.7 M km²), indicating the model underestimated discharges from the northern half of the basin. This south to north pattern is likely a function of the assumed 1 m ALT throughout the basin. Comparing model M22–M24 in Table IV, which only represent changes in ALT, shows that the maximum index of agreement ($\delta$) for the northern gauges is obtained when the ALT is set to zero, while the maximum $\delta$ for the southern gauges is obtained when the ALT is set to 1000 m. Although the range of 0 and 1000 m is extreme, these results suggest that the ALT is likely $>$1 m in southern region and $<$1 m in the northern region. These results are consistent with published patterns of ALT (Nixon and Taylor, 1998) and support the need to fully couple the routing model with CLM to access time-varying ALT.

The third point is that the rapid drop in discharge followed by a secondary, small peak evident at several locations in Dec/Jan (refer to highlighted region in Figure 7c) is not well represented in the model. The secondary peak is captured in the model, but the rapid drop in discharge to below base flow levels is not (refer to Figure 7a, b, c). The near vertical drop in discharge is generally captured at gauge 10HA004, which drains 11 200 km² but is no longer evident at gauges draining larger areas. Future efforts will explore the physical mechanisms causing this discharge pattern. Explanations for this could include a non-linear release rate of the surface storage and more rapid subsurface response, which would reduce the falling limb of the snowmelt driven hydrograph more rapidly such that the start of the rainfall induced hydrograph is more defined. For the two larger gauges, which are impacted by reservoirs, it may be possible to improve the storage-discharge relationships approximated from elevation and discharge information for low flows/storage conditions where hydraulic controls may lead to sudden decreases in outflow.

The final point to note from Figure 7 is that the linked CLM-HRR model is able to provide reasonable results for select locations draining as few as 1500 km², which is much less than the resolution of CLM. Although model performance is generally less favourable for drainages less than 10 000 km², Figures 7e and 7f show that the model captures the general discharge pattern but does not capture all the variability seen at gauge 07HA004-M in the northern portion of the basin, which is likely because of a combination of overestimating both surface storage and ALT that further attenuates the discharge response. The contrast between these two smaller gauges is also interesting. Gauge 07HA004-M shows a very flashy hydrograph compared with the smooth response at 07SA004-S. This contrast can be explained by their watershed characteristics: 07HA004-M has a mean ground slope of 33% compared with 4% for 07SA004-S. The impact of slope is important to note and supports the inclusion of the split slope plane in HRR to better represent the subhillslope-scale landscape characteristics.

In addition to streamflow, TWSA and individual water storage anomalies (WSA), relative to the mean of the monthly total water storage or individual water storages for 2003, from CLM-HRR and GRACE are compared. Figure 8 shows the mean monthly TWSA’s and WSA’s for the 2003–2004 period, which is the period of overlap with GRACE. Three points are evident: The magnitudes of TWSA agree with the GRACE estimates; the model’s peak TWSA recession (May to June) starts roughly 1 month after the drop estimated from GRACE (Apr to May); and the model does not capture the secondary peak in TWSA observed in the summer. The lag in peak TWSA recession occurs during the snowmelt period (Figure 8b). Thus, this lag is most sensitive to CLM snow processes, which are discussed later. Given that both CLM-HRR and GRACE have similar patterns and magnitudes in TWSA suggests that the water balance within and the transfer rates between major storages are well represented. The secondary peak in summer TWSA is potentially because of a combination of precipitation not captured by the models, the uniform, constant surface storage release rate, and the lack of evaporation from the

![Figure 8. Mean monthly: (a) total water storage anomalies (TWSA), and (b) individual water storage anomalies (WSA) for surface, subsurface, snow, and combined rivers and lake storages from CLM-HRR and GRACE for the period 2003–2004](image)
surface storage modelled by HRR. At the peak of summer, it is likely that the surface storage drain rate is higher than at the start of snowmelt. A higher drain rate would lead to a more rapid decrease in TWSA followed by the increase of TWSA in summer/early fall as a result of precipitation (Cassano and Cassano, 2010). As simulated, the constant drain rate may be masking the change of source water from snowmelt to rainfall.

Relative to the gauge data, the final model produces approximately 12% more annual run-off averaged over the 27 gauges but 20% more run-off when compared with the outlet gauge. Given the importance of snow accumulation and melt in the MRB, Figure 9 shows the comparison between monthly snow water equivalent (SWE) simulated by CLM and estimated by the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) instrument on the NASA’s Aqua satellite (Tedesco et al., 2004). For the study period, there are 2 years of overlap with AMSR-E (2003–2004). Figure 9 shows good agreement in timing but over estimates in SWE magnitudes by approximate 24%. The additional SWE in March may also contribute to the 1-month lag in TWSA recession in Figure 8 and discussed in the preceding texts. Additional run-off bias may be explained by the decoupled model setup (i.e. CLM simulates evaporation for its water stores, but the surface storage simulated by HRR is not subjected to evaporation). Given that CLM solves the mass and energy balance, it is not possible to assume that additional energy is available for evaporation in HRR. This challenge requires fully coupling the water balance and routing models.

**Sensitivity analysis**

To illustrate the sensitivity of model results to parameter uncertainty, additional model simulations were performed where subsurface horizontal conductivity ($K_h$), depth of surface storage ($S_d$), and surface storage drain rate ($S_r$) were changed individually by ±25%. Simulated discharges were then compared with the calibration model (M23) at gauge 10ED002-L. The gauge is located in the middle of the basin, drains 275,000 km$^2$, and is not impacted by any of the three large lakes (refer to Figure 7c). The results show that the model is most sensitive to changes in surface storage drain rate and depth of surface storage (Table V). For example, changing $S_d$ resulted in a −5.3 to +7.8% change in peak discharge at the gauge. This is similar to changing $S_r$ that resulted in a −6.0 to +7.1% change in peak discharge. However, changing $K_h$ only resulted in a −2.6 to +2.0% change in peak discharge. In all cases, impacts on annual run-off and trends in TWSA’s are minimal (0.0 to 0.4 mm and −0.03 to 0.03 mm/month). Only changes to $S_d$ resulted in changes in the timing of peak discharge (−2 to +1 days). For all three parameters, decreases of 25% resulted in more change in peak discharge as compared with increases of 25% indicating that their roles in the model are non-linear and impacted by thresholds. For example, increasing $S_d$ beyond the available surface run-off provides no additional attenuation. Similarly, although increasing $S_r$ provides more subsurface run-off, it also makes more surface storage available prior to peak snowmelt that provides capacity for run-off attenuation.

To illustrate these points, Table V shows the comparison between the final calibrated model (M23) and the ±25% changes for each of the three parameters. Changing $S_d$ impacts the rising limb of the hydrograph (i.e. changes peak discharge and time to peak by several days) and the very end of the recession limb (i.e. shifts storage from the start to the end of the flood). Changing $S_r$, resulted in the largest impacts on the hydrograph relative to the other two parameters with the largest changes around the peak and tail of the flood hydrograph. The changes in the flood peak resulting from $S_r$ and $K_h$ tend to be mostly vertical given that they both represent flow velocities. In contrast, $S_d$ shifts the flood peak in both timing and magnitude (Table V).

**Role of individual hydrologic processes**

Although the model was calibrated, the intent of calibration was to provide a reasonable model such that the newly added model components can be assessed. In this section, we compare results for the 24 models listed in Table IV. The models provide all possible combinations of the four model components: reservoirs versus no reservoirs; surface storage versus no surface storage; single slope versus split slope; and ALTs of 0.0, 1.0, and 1000 m. For reference, the models are compared with model M3 (no reservoirs, no surface storage, single slope, and an ALT of 1000 m). Model M23 represents the final calibrated model with all of the new
components (reservoirs, surface storage, split slope, and an ALT of 1.0 m).

Figure 10a shows the simulated hydrograph at the outlet gauge location for 2003 with an ALT of 1 m for single (M2) and split slope (M5) hillslopes relative to the base model with no new components (M3); all of which do not include reservoir routing or surface storage. Although there are very small differences between ALT values of 1 m (M2) and 1000 m (M3; base) for the single sloping plane, the split slope plane (M5 and M6) shows a small reduction in Q_p but a decrease of 2 and 6% (4 and 12 mm) in annual run-off and an increasing trend in TWSA’s of 0.6 and 1.7 mm/month (Table V). The impact on TWSA with minimal impact on Q_p (averaging 1–2% over 2000–2004) is because of the reduced slope of the hillslope near the channel slowing the subsurface flow velocities and storing water in the active layer. This point is supported by the results for models M1 and M4 (ALT=0 m), where all the run-off is on the surface, and there is no build-up of subsurface storage (i.e. no trend in TWSA).

The impact of surface storage is shown in Figure 10b, which decreases the peak discharge by 33 to 56% and slows the time of peak discharge by 20 to 41 days (M7–M12) as expected. Figure 11 shows how the combined effects of surface storage and the split slope planes amplify the effect of ALT. When surface storage (S_d and S_r) is included, snowmelt is transferred to subsurface flow via the surface storage drain rate. When the ALT is 1000 m, there is no resurfacing of subsurface flow, and when combined with the reduced slope near the river on split slope hillslope (M12), there is substantial build-up of subsurface water storage (annual run-off reduced by 17% or 34 mm and trend in TWSA increased by 4.7 mm/month). This point highlights the importance of ALT in the Mackenzie Basin, which has significant portion of lands with low relief (65% of the basin has ground slope less than 5%; 31°< 2.5°).

When reservoir routing for the three largest lakes is included, the peak discharges are reduced by 14 to 19% (Figure 10c, Table V). However, there is essentially no change in the time of the peak as a result of the magnitude
and short duration of snowmelt and the nearly linear relationships between storage and discharge for the lakes (Figure 3). The lack of change to peak timing as a result of reservoir routing is an important finding and supports the need for another mechanism to slow the response by approximately 1 month to agree with gauge data (Figure 7).

Combining surface storage, reservoir routing, and split slopes provides reductions in $Q_p$ (42 to 60%) beyond that of each individual mechanism and slows the peak

Figure 10. Summary of model results at the outlet gauge ($A = 1.7 \, M\, km^2$) for 1 year (2003) with a uniform, constant active layer thickness (ALT) of 1.0 m, single and split slope hillslopes, and combinations of reservoir and surface storage processes: (a) without reservoirs or surface storage, (b) surface storage without reservoirs, (c) reservoirs without surface storage, and (d) with reservoirs and surface storage; base model (M3) corresponds to original HRR model without modifications presented here, refer to Tables IV, V for details on each model number (M#)

Figure 11. Effects of active layer thickness (ALT) on (a) the mean flood hydrograph at the outlet gauge ($A = 1.7 \, M\, km^2$) and (b) the mean change in basin-wide total water storage anomalies (TWSA), for the period 2000–2004
response by 28 to 42 days (Figure 10d, Table V M19–M24). Similar to the surface storage only models (Figure 11, M10–M12), the ALT variations on the split slope (M22–M24) show significant changes in TWSA with increasing ALT. The final calibrated model (M23) results in a 51% decrease in Qp, 36 day delay in time of peak, 3% (5 mm) reduction in annual run-off, and a 0.5 mm/month trend in TWSA. Collectively, the new model components provide mechanisms to attenuate the rapid snowmelt signal while producing discharge hydrographs that generally agree with gauge streamflow measurements and TWSA’s consistent with GRACE data.

For an average ALT of 1 m, surface storage, lake routing, and reduced hillslope slopes near rivers account for approximately 19 to 34, 10 to 15, and 3 to 10%, respectively, of the overall reduction (51%) in the annual flood peak. These results are consistent with Bowling and Lettenmaier (2010) who found that 80% of snowmelt goes into storage. Branstetter and Erikson (2003) also concluded that including storage processes, in particular the effects of impoundments, results in significantly improved agreement between measured and predicted streamflow by global climate models.

The inclusion of an active layer provides a mechanism for run-off to be attenuation but also the ability to resurface when the ALT fills. This generation of additional surface run-off leads to a faster flood response and larger peak discharges (Table V). Figure 11 shows the results for ATLs of 0, 1, and 1000 m. The use of 1000 m is intended to represent an unlimited ALT. Comparing model results for 21 and 22 shows that for the single slope hillslopes with 1 or 1000 m ALTs, the reduction in Qp is 44% for both cases indicating that surface run-off via saturation from below did not occur for the 1 m ATL. However, comparing model results for 23 and 24 shows that for the split slope hillslope, there is a difference in Qp reduction for 1 m (51%) versus 1000 m (60%) ALT. There is no difference in Qp reduction for the single slope case, but there is a difference in Qp reduction for the split slope case indicating that an ALT of 1 m on the split slope produces surface run-off from below. These results are consistent for models M11–M12, which also include surface storage providing a mechanism to transfer snowmelt to the ALT. Thus, the reduction in slope along the lower 33% of the hillslope combined with a mechanism to transfer snowmelt to the ALT results in subsurface flow depth greater than 1 m, produces surface run-off, increases peak discharge by 9%, and reduces the time of the peak by 6 days (model M23 vs M24). Note that, as implemented here, excess subsurface discharge generated when the ALT capacity is exceeded is added to surface run-off directly and not routed through the surface storage module. Future efforts will explore methods for coupling the surface storage with the excess subsurface flow.

Distribution of water storage

One of the unique aspects of the linked HRR-CLM model is its ability to track all water storages upstream of any river reach. Figure 12 shows the distribution of individual water storages contributing to the active fraction of the mean monthly TWS changes averaged over the entire MRB for the period 2000–2004. Note that, in terms of absolute water storage, the subsurface, lakes and combined surface, rivers, and snow account for approximately 73, 26, and 1%, respectively, and is controlled by the assumed maximum groundwater store in CLM (5000 mm of water). The distribution of monthly storage changes in Figure 12 was determined by dividing the absolute value of the monthly changes for individual components by the sum of monthly absolute value changes.
from all components. The relative contribution for each component varies throughout the year but on average surface, subsurface, snow, channels, and lakes represent 11, 29, 41, 15, and 4 of the active TWS changes. Although snow represents the largest range in monthly anomalies, the lack of snow for roughly 4 months reduces the monthly mean to only 41%. The combined role of surface waters (surface storage, rivers, and lakes) accounts for an appreciable fraction of the active TWS changes (30%) ranging from 13 to 60% throughout 2000–2004. Looking specifically at surface storage, the maximum monthly changes from each year, the basin-wide, mean monthly storage change is 29 mm. The role of surface storage in the MRB’s active water storage signal supports the need for routing of lateral run-off.

CONCLUSIONS
This goal of this study is to improve the representation of landscape topography in a climate simulation. Here, we link the HRR and CLM models in the MRB to demonstrate the practicality of using a topographically defined land surface for routing purposes. In this application, the models are not coupled. Surface and subsurface excesses from CLM are passed to the HRR model for routing. Ideally, the land surface stores and fluxes simulated in HRR-CLM would be mapped back onto the atmospheric model grid using the same algorithm used for mapping downward fluxes to HRR to ensure a balance of the of water and energy budgets (i.e. potential for additional surface water evaporation from the surface storages in HRR). This approach would allow a more accurate representation of the earth’s land surface in coupled climate model simulations.

For this study, the HRR model was modified to account for key arctic river basin processes included in the simulation of all water stores and fluxes. Specifically, four new components are incorporated into HRR: 1 multi-sloped hillslope; 2 a shallow ALT to represent the permafrost boundary (ALT); (3) surface storage compartment \((S_d)\) with a corresponding drainage rate \((S_r)\) to represent the cumulative storage of all the small lakes and depressions distributed throughout the landscape; and (4) a lake routing model to explicitly account for large lakes. The HRR model with the four new components was tested using data from 27 streamflow gauges and is shown to compare reasonably well with gauged streamflow and GRACE TWSA given the resolution (0.9° by 1.25°) of the CLM excesses passed to HRR. For example, there is a significant positive relationship between model performance and drainage area, which is likely because of the resolution of the CLM model used here. At the smaller scales, run-off and snowmelt are muted by the spatial averaging. Future efforts will explore quantitative measures to define area thresholds for the applicability of model results.

Additionally, the results of this study are impacted by the assumed ALT. Here, a uniform ALT of 1 m was used. However, comparing simulated and measured streamflow suggests an ALT <1 m in the northern half of the basin and ALT >1m in the southern half are likely more appropriate. Future efforts will investigate spatial and temporal variations in ALT. Ideally, ALT would be fully coupled to the CLM model.

Using the calibrated model parameters, a modelling experiment is presented to quantify the role of each new model component individually and in various combinations. The primary finding is that surface storage ranging from 1.25 to 12.5 cm with a drain rate of approximately 1 mm/day is required to slow the rapid snowmelt by approximately 1 month and further reduce peak discharge by roughly 30%. This surface storage is required even though the HRR model is simulating lake routing, reduced landscape slopes near river floodplains, and an ALT. In addition, the use of discrete surface storage depths and uniform release rates lead to threshold effects. If snowmelt and run-off do not exceed the assumed surface storage, the resulting streamflow response is greatly attenuated. If they do exceed storage, a large, rapid streamflow response is produced. Thus, results are more sensitive to underestimating the surface storage as compared with overestimating. Once the storage estimate exceeds available snowmelt and run-off, it can be increased significantly with no additional impact on streamflow. Similarly, release rate impacts the available surface storage at the time of peak snowmelt. If the release rate is too large, there tends to be available storage capacity and streamflow is attenuated. Thus, once the combination of surface storage and release rates are able to accommodate all the rapid snowmelt, further changes to the parameters that lead to increases in storage result in only minimal changes in simulated streamflow. Future efforts will explore methods for estimating surface storage and non-linear release rates to limit the resulting threshold effects.

Building on the HRR model’s ability to track all water stores and fluxes within any subcatchment of averaged over any watershed, the individual components of the TWS signal are quantified. In the Mackenzie Basin, the combined role of surface waters (surface storage, rivers, and lakes) accounts for an appreciable fraction of the TWS change (30%) ranging from 13 to 60% over period 2000–2004. The contribution of surface waters supports the need for routing of surface run-off and streamflow. The overall intent of this study is to highlight the role of lateral run-off processes in large river basins and support future research on coupling lateral and vertical process models applicable for continental to global scales.
ACKNOWLEDGEMENTS

This work was funded by the DOE Office of Science, Biological and Environmental Research program, Regional Climate Impacts and Next Generation Ecosystem-Arctic projects, and NASA’s Terrestrial Hydrology Program (Grant No. NNX12AQ36G and NNX14AD82G). We acknowledge Garrett Altmann and Mary Greene of the Earth and Environmental Sciences Division at the Los Alamos National Laboratory for their contributions to LandSat image processing and determining river widths.

REFERENCES


Colorado USA: NASA DAAC at the National Snow and Ice Data Center 10.5067/AMSR-E/AE_DYSNO.002.


