Coupled groundwater and dynamic lake modelling using the Water Table Model (WTM)

Kerry L. Callaghan[1,2,3], **Andrew D. Wickert**[1,2], **Richard Barnes**[4,5,6]

[1]Department of Earth & Environmental Sciences, University of Minnesota, Minneapolis, USA
[2]Saint Anthony Falls Laboratory, University of Minnesota, Minneapolis, USA
[3]Lamont-Doherty Earth Observatory, Columbia University, New York, USA
[4]Energy & Resources Group (ERG), University of California, Berkeley, USA
[5]Electrical Engineering & Computer Science, University of California, Berkeley, USA
[6]Berkeley Institute for Data Science (BIDS), University of California, Berkeley, USA

Key Points:

- The WTM simulates water-table elevation, including both groundwater and dynamic lake components, on a continental scale.
- Combining depression-hierarchy, fill–spill–merge, and adaptive time-stepping algorithms enables efficient lake–groundwater coupling.
- Running the WTM on North America shows the importance of both model components in understanding the coupling of climate and water levels.

Corresponding author: Kerry L. Callaghan, `kerryc@ldeo.columbia.edu`

Abstract

Water stored in lakes and underground is a crucial component of the global hydrological cycle, with impacts on climate and sea level. However, long-term changes in the global distribution of this water are not well understood. Here we present the Water Table Model (WTM), which is capable of computing water-table elevation at large spatial scales and over long temporal scales. The WTM comprises two components: groundwater and dynamic lakes. The inclusion of a dynamic lake component allows us to incorporate surface-water movement and evaporation into water-table elevation estimates. We share sample results from both an artificial topography, and for the North American continent. These results indicate the close interactions between changes to water levels in lakes and the surrounding groundwater tables. The open-source code for the WTM is available on Github and Zenodo.

Plain Language Summary

Groundwater and water stored in lakes influence climate and sea level, both in the short- and long-term. However, long-term changes in groundwater and lake-water storage are not well understood. In this study, we present a new model, the Water Table Model (WTM). The WTM calculates the depth of the water table, whether it is below the land surface (groundwater) or above the land surface (lakes). The model is efficient enough that we can use it to find the water table elevation for whole continents, even over long time spans. We share a sample result for the North American continent, and show that it is important to include both the groundwater and lakes within the model.

1 Introduction

The water table represents the level below which water-saturated conditions occur. Below the water table, extensive water stores exist: the volume of groundwater stored in the upper 2 km of continental crust is approximately 22.6 million km^3 (Gleeson et al., 2016), or approximately 62 m sea-level equivalent. In addition, lake surfaces represent locations where the water table peeks above the land surface. Although lakes cover only about 3.7% of the Earth's nonglaciated land surface (Verpoorter et al., 2014), they are numerous: Verpoorter et al. (2014) recorded over 100 million lakes in their inventory. The total volume of the world's lakes is about 181.9×10^3 km^3 (Messager et al., 2016), or approximately 0.5 m SLE. This lake-water storage impacts hydrologic connectivity (Callaghan & Wickert, 2019), and therefore also sediment and contaminant transport. Groundwater provides baseflow to rivers and lakes, defines wetland locations (Fan et al., 2013), and provides a large storage of fresh water for human use (Wada, 2016). Changes in water-table elevation represent changing terrestrial water storage volumes, significantly impacting the hydrological cycle on a global scale (Ni et al., 2018; Syed et al., 2008).

High-performance computing and efficient algorithms have made continental-scale modelling of modern-day groundwater (Fan et al., 2013; Maxwell et al., 2015) and streamflow (Döll et al., 2009; NOAA, 2016) possible. Various land-surface models (e.g. Decharme et al., 2019; Koirala et al., 2014; Lawrence et al., 2019; Wiltshire et al., 2020; Yokohata et al., 2020) provide complex depictions of surface and sub-surface hydrology. These land-surface models are often coupled with climate models (e.g. Decharme et al., 2019; Lawrence et al., 2019; Yokohata et al., 2020; Zeng et al., 2002) to understand change, including changes in terrestrial hydrology, through time. However, while they may include lake components that influence local climate (Oleson et al., 2010), these models do not incorporate dynamic changes in lake-water storage through time. In addition, land-surface models generally have complex data input and calibration requirements. These limitations make it difficult to assess long-term changes in water table, both because dynamically changing lakes are not simulated and because input data may not be available prior to the beginning of the satellite record.

Here we present the Water Table Model (WTM), which couples groundwater (Section 2.1) and lake-water (Section 2.2) levels and flow across spatial scales from local catchments to the globe and over time scales of years to thousands of years (Figure 1). Input data to the WTM are commonly available for both the present day and recent geological past, and are described in sections 2.1 and 2.2. The conceptual simplicity of the model helps it to remain practical and computationally efficient over large spatial and long time scales. These large spatial and long temporal scales represent the ideal use of the WTM, which captures broad-scale natural variations in water-table elevation.

The WTM offers several advantages: (1) Simplicity – the focus of the model is on the simulation of the water table alone. There is no need to model other processes such as vadose zone processes, integrated climate factors, or streamflow. (2) Open-source model code – the source code for the WTM is available on GitHub (`https://github.com/KCallaghan/TWSM/`, v1.1.0) and Zenodo (`https://doi.org/10.5281/zenodo.3554537`, v1.1.0), for other researchers to use and peruse. (3) Inclusion of a dynamically changing lake component – lake locations are not predefined and can dynamically interact with the rest of the water table. (4) Broad applicability – the WTM can be used across spatial scales from catchment to global and can produce both transient and steady-state water-table outputs.

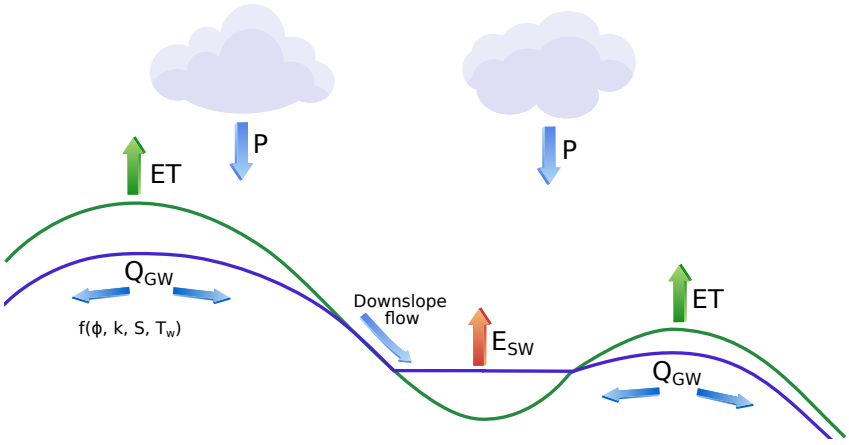

Figure 1. A schematic of the Water Table Model, including both groundwater and dynamic-lake components. Green represents the land surface and blue represents the water table after the WTM has moved water for some time. The groundwater table is recharged through precipitation (P) minus evapotranspiration (ET), including surface-water evaporation (E_{SW}) in place of land-surface evapotranspiration, as appropriate. Groundwater discharge (Q_{GW}) is computed as a function of porosity (ϕ), hydraulic conductivity (k), slope (S), and winter temperature (T_w), as well as the head gradient between adjacent cells within the discretised rectilinear domain. The lake shown in the figure receives groundwater inputs from both sides. As the water level emerges above the land surface to the left of the lake, the surface-water model component (Fill–Spill–Merge) moves this water downslope, producing a flat lake surface that partially fills the depression.

2 Model description

The WTM (Callaghan et al., 2020) simulates the depth to water table, inclusive of both groundwater and lake surfaces. Separate model components for simulation of groundwater and dynamic lakes are run in a repeated cycle to permit feedbacks between the

ground- and surface-water components of the terrestrial hydrological system. The water table is recharged by precipitation-minus-evapotranspiration ($P - ET$), and evaporation occurs where surface water is present. Exfiltrating water flows downslope and may help to fill lake basins, and seepage from lakes may raise regional groundwater tables. When climate or topography change, flow paths and hydrologic connectivity are impacted and this is reflected in changes in the water table.

The WTM captures broad natural patterns in water-table elevations, and is most appropriate for large spatial scales, from continent-spanning catchments to the globe. The WTM can simulate a steady-state water table for any given set of conditions, and can also simulate transient change in the water-table over time. For transient water-table change, the WTM is best suited to temporal scales from years to millennia.

2.1 The Groundwater Component

To simulate the groundwater table, the groundwater model component solves the 2D horizontal groundwater flow equation (Equation 4) with a finite-difference approach in a single layer of vertically integrated hydraulic conductivity (i.e. transmissivity). This 2D approach follows the conceptual framework of Fan et al. (2007); Fan and Miguez-Macho (2011) and Fan et al. (2013) and is explained in more detail in Section 2.1.2. Groundwater flow is forced by climate (i.e. recharge), topography, and sea level. This component moves physically realistic water volumes over user-selected time intervals for a user-selected number of time intervals.

The groundwater model component requires the following 2D, horizontally-distributed input arrays:

- **Topography:** Land elevation above sea level, in metres.
- **Ocean mask:** A binary mask with 1 values indicating land cells and 0 values indicating ocean cells.
- **Climatic water input:** Precipitation and, if appropriate, ice melt in metres per year.
- **Evapotranspiration:** Evapotranspiration occurring over land in metres per year.
- **Open-water evaporation:** The evaporation that will occur when there is open surface water (i.e. a lake: Appendix E).
- **Winter temperature:** Temperature during the months of December, January, and February (Northern hemisphere) or June, July, and August (Southern hemisphere) [°C].
- **Slope:** Topographic slope, which should be based on the input topography data.
- **Shallow sub-surface hydraulic conductivity – horizontal:** Horizontal hydraulic conductivity (k in Equation 6), representative of near-surface conditions, in metres per second.
- **Porosity:** Shallow sub-surface porosity (ϕ in Equations 13, 14 and 15). Porosity is unitless.
- **Starting relative water-table elevation:** The relative water-table elevation (z_{wr}) is defined as the water-table elevation minus the elevation of the land surface, in metres. Positive values indicate the presence of a lake, while negative values indicate groundwater table. This input is optional: if it is not supplied, z_{wr} will be initialised at 0 (equal to the land surface) and the model should first be run to equilibrium.

As an output, the WTM returns a 2D array of water-table elevation, z_{wr}.

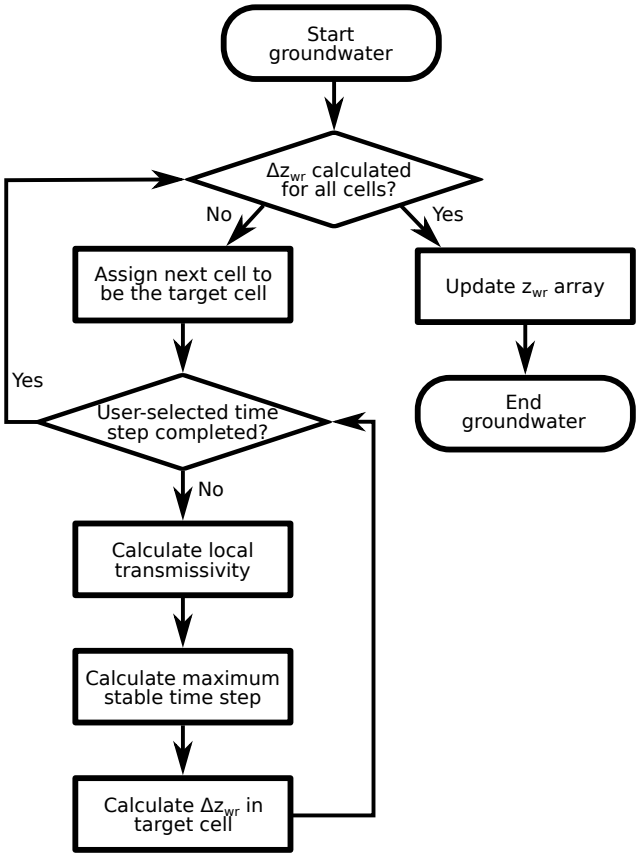

Figure 2. Steps taken by the groundwater component. Relative water-table elevation, defined as the water-table elevation minus the land-surface elevation, is abbreviated as z_{wr}. 'Target cell' refers to the current cell on which we are performing calculations. To see how the groundwater algorithm fits into the model coupling, see Figure 5.

2.1.1 Exponential decay of Hydraulic Conductivity

Hydraulic conductivity dictates the rate of lateral groundwater flow. Given the difficulty in obtaining data on the variability of hydraulic conductivity with depth at large scales, the WTM uses only a single hydraulic conductivity value in each cell of the domain, representative of near-surface conditions. This value is used from the land surface to a depth of 1.5 m, because the global soil datasets that our hydraulic conductivity inputs are based on are representative of the conditions until approximately this depth. Beyond depths of 1.5 m, we assume that the hydraulic conductivity decreases exponentially with depth. Exponential decay of hydraulic conductivity with depth is a well-documented phenomenon and is commonly used in the absence of more detailed data (Ameli et al., 2016; Cardenas & Jiang, 2010; Fan et al., 2013).

We calculate the rate of this exponential decay as an e-folding depth (f_d), dependent on both the local terrain slope and a temperature modifier to decrease hydraulic conductivity at locations where seasonal frost or permafrost occur:

$$f_d = f \times T_f, \tag{1}$$

where f is a slope-dependent term described below and T_f is a temperature-dependent modifier. We define f as:

$$f = \max\left(f_{\min}\ ,\ \frac{a}{1+bS}\right), \tag{2}$$

where S is the terrain slope; and a, b, and $f_{\min}$ are user-selected calibration constants.

The temperature-dependent modifier, T_f, is incorporated into the e-folding depth following the method used by Fan et al. (2013). When the average winter temperature drops below -5° C, we assume that seasonal frost inhibits groundwater flow. When average winter temperatures fall below -14° C, we assume that groundwater flow is affected by permafrost. This limits lateral drainage, reducing the effective hydraulic conductivity (Fan & Miguez-Macho, 2011). We define T_f as:

$$T_f = \begin{cases} 1, & \text{if } (T_C > -5°C) \\ 1.5 + 0.1T_C, & \text{if } (-14°C < T_C < -5°C) \\ \max(0.17 + 0.005T_C\ ,\ 0.05), & \text{if } (T_C < -14°C), \end{cases} \tag{3}$$

where T_C is the temperature in degrees Celsius.

2.1.2 Lateral groundwater movement

Lateral groundwater movement is computed on a cell-by-cell basis using an explicit finite difference numerical scheme. We solve for the change in head (h) in each cell using the 2D horizontal groundwater flow equation. We invoke the Dupuit–Forchheimer approximation, which posits the assumptions that flowlines are horizontal and that the hydraulic gradient is equal to the slope of the water table and does not vary with depth (Freeze & Cherry, 1979). In this way, groundwater flow is limited to two dimensions (the x and y dimensions):

$$\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right) = \frac{S_y}{T} \times \frac{\partial h}{\partial t}, \tag{4}$$

where T is the transmissivity (detailed in Section 2.1.2.2), t is the length of a single time interval, and S_y is the specific yield, here approximated as being equal to porosity. For each cell in the domain, we solve this equation in the following steps:

1. Recharge the groundwater table and evaporate surface water (Equation 5).
2. Calculate the local transmissivity (Equation 6).
3. Calculate the length of the time interval, t, to use (Equation 7).

4. Compute discharge into or out of the cell in the x and y directions during time interval t (Equation 10), hence obtaining the total change in groundwater volume within the cell (Equation 11).
5. Use the cell area and local porosity to convert the volume change to a change in water table elevation within the cell, and obtain the updated relative water table elevation z_{wr} (Equations 12, 13, 14 and 15).

We call the cell being considered at any given time the 'target cell'. Figure 2 summarises the steps taken.

2.1.2.1 Recharge and evaporation We use the precipitation (P), overland evapotranspiration (ET), and open-water evaporation (E_{SW}) input arrays to recharge the groundwater table and evaporate surface water. When surface water is present, evaporation rates typically increase. Physically, this is because actual evaporation is able to equal potential evaporation. Both physically and algorithmically, this typically acts as a feedback that slows runaway lake growth by decreasing the catchment-wide water balance as the lake surface area increases. To account for the changes in evaporation dependant on the presence of surface water, the WTM recalculates the climatic water input (Equation 5) at the beginning of each groundwater–surface-water model cycle. The climatic water input (b_C) is given by:

$$b_C = \begin{cases} \min(P - ET, 0) & \text{if } z_{wr} \leq 0 \\ P - E_{SW} & \text{if } z_{wr} > 0, \end{cases} \tag{5}$$

Equation 5 dictates how much water is added to a cell due to precipitation at each time interval. If the cell's water table is below or at the surface (groundwater) then the water added is precipitation minus evapotranspiration. If there is sufficient evapotranspiration, then no water is left to be added to the groundwater, though the surface of the earth shields the groundwater itself from evaporation. Conversely, if the cell's water table is above ground (a lake) then sufficient evaporation can actually subtract water from the cell.

We assume that vertical transit time through the unsaturated zone is short compared to the time interval for the calculation, and therefore simplify by ignoring any vertical transit time: recharge is applied directly to the water table.

2.1.2.2 Transmissivity We calculate the local depth-integrated transmissivity, based on the shallow subsurface hydraulic conductivity and the e-folding depth described in Section 2.1.1. Following Fan et al. (2013), we divide this calculation into three cases:

1. The water table lies below 1.5 m depth, where the exponential decay of hydraulic conductivity comes into play.
2. The water table lies in the shallow subsurface, above 1.5 m depth, where the unmodified input hydraulic conductivity is representative of conditions at the water table.
3. The water table lies above the land surface. In this case, hydraulic conductivity is calculated at the level of the land surface (i.e. it is identical to that for a fully saturated substrate). The dynamic lake component (Section 2.2) later moves the surface water into depressions or out of the domain as appropriate.

Based on these three cases for hydraulic conductivity, we calculate transmissivity, the depth-integrated hydraulic conductivity from $-\infty$ to z_{wr}, as:

$$T = \begin{cases} f_d \times k \times \exp\left(\frac{z_{wr}+1.5}{f_d}\right), & \text{if } (z_{wr} < -1.5\,\text{m}) \\ k \times (z_{wr} + 1.5 + f_d), & \text{if } (-1.5\,\text{m} \leq z_{wr} \leq 0\,\text{m}) \\ k \times (0 + 1.5 + f_d), & \text{if } (0\,\text{m} < z_{wr}), \end{cases} \tag{6}$$

where T is the transmissivity, f_d is e-folding depth (Equation 1), k is the shallow subsurface horizontal hydraulic conductivity, and z_{wr} is the relative water-table elevation. See Fan and Miguez-Macho (2011) and Fan et al. (2013) for more information on the derivation of these formulæ.

2.1.2.3 Time interval selection and handling We define model time intervals as "stable", "safe", or "unsafe" and use these to balance accuracy and runtime. Unsafe time intervals are those that are neither globally safe nor locally stable, and are longer than both. If the minimum "stable" time interval of all the cells is used as the global time interval, then the time interval is mathematically guaranteed to be short enough to prevent numerical instabilities in the model when computing changes in z_{wr}. However, because of the spatially variable cell size and transmissivity, different cells may have dramatically different stable time intervals. This means a single cell with a very short stable time interval can hold up the whole calculation.

Therefore, we require the user to supply a "safe" time interval, used as the global time interval, which does not have a mathematical guarantee of stability but which we can show empirically to satisfy indicators of stability such as mean water-table elevation (Figure 3). "Safe" time intervals are short enough to avoid situations where water would flow through multiple cells in the domain, rather than just to a neighbouring cell, during each time interval.

If the length of the global "safe" time interval exceeds that of the local "stable" time interval between a given target cell and its four neighbours, then several stable time intervals which collectively comprise the same time as the "safe" time interval are made for the target cell.

Ultimately, the "safe" timestep allows WTM to avoid moving water through all cells at the slowest rate while ensuring that any fluid movement between adjacent cells is still physically realistic, thereby accelerating the model.

We use a 2D von Neumann stability analysis (Charney et al., 1950; Crank & Nicolson, 1996) to calculate the maximum stable time interval for a given target cell based on local transmissivity, porosity, and cell size:

$$t_N = \frac{(L_{EW}^2 \times L_{NS}^2)}{4 \times T_{max} \times (L_{EW}^2 + L_{NS}^2) \times f_s} \times \phi_{min}, \tag{7}$$

where t_N is the von Neumann stable time interval, L_{EW} is the cellsize in the east-west direction, L_{NS} is the cellsize in the north-south direction, T_{max} is the maximum transmissivity among the target cell and its four neighbour cells, ϕ_{min} is the minimum porosity among the target cell and its four neighbours, and f_s is a factor of safety, set equal to 2. t_N is computed for every cell in the domain.

In the ideal case, we would select the smallest t_N value from the domain and set the stable time interval equal to this value for all cells. However, doing so can lead to unreasonably long compute times: a few cells with unusually high transmissivity values resulting from a combination of high hydraulic conductivity, low slope, and shallow water table can set an exceptionally short stable time interval for all cells in the domain. To prevent the problem becoming computationally untenable, we instead select an individual t_N value for each cell based on its local conditions.

Because the stability criterion is applied only to each individual target cell, this method may result in a loss of conservation of water mass. This is because the volume of water moving out of one target cell to its neighbour may not be exactly equal to the volume of water received by the neighbour if they operate on different time intervals. However, tests on the synthetic topography shown in Section 3.1 showed that mass loss on this topography with various time intervals was less than 0.1% of the summed z_{wr} across the domain. From this, we conclude that any volume loss is small relative to the total wa-

ter volumes being moved by the component. A comparison of the mean z_{wr} in a model result using various time intervals, both longer and shorter than the calculated time interval, is given in Figure 3. Although this water volume is small, users should select a time interval shorter than the von Neumann stable time interval for their region in cases where this is possible in order to obtain the most accurate result possible.

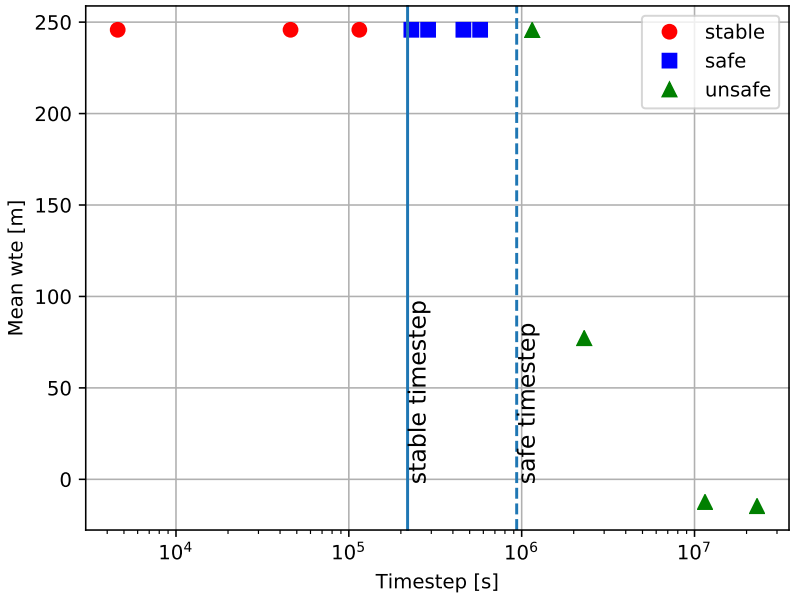

Figure 3. Mean relative water-table elevations with different time intervals in a sample model run. The *stable* time interval was 218,733 s when using a factor of safety equal to 2, and is indicated by the solid blue line. The estimated *safe* time interval was 935,609 s when using a factor of safety equal to 5, and is indicated by the dashed blue line. Red circles indicate time intervals that are shorter than the domain-wide stable time interval. Blue squares indicate time intervals that are longer than the stable time interval, so that groundwater movement may be calculated in multiple portions, but the time interval is still safe. Green triangles indicate time intervals that are unsafe. These time intervals are too long and will lead to erroneous results. Note that the leftmost green triangle does not show a significantly lower mean z_{wr} and is likely a safe time interval: it is marked unsafe as a result of the cautious selection of the safe time interval. A factor of safety equal to 1 would have resulted in the acceptance of the second-from-the-left green triangle, a time interval which, as indicated by this plot, resulted in mass loss and is clearly unsafe.

Note that even when the user-selected time interval is shorter than the domain-wide stable time interval across the model domain, minor differences will occur between results using different user-selected time intervals. This is the nature of the forward-differencing regime: a longer time interval will move water across a steeper head gradient for the full length of the time interval. These differences are small in comparison to the broad patterns in z_{wr} produced by the model (Figure 3).

While the stable time interval ensures numerical stability of the calculation, it is vital that the user also select a *safe* time interval. Any given target cell only has direct communication with its four neighbouring cells during a single time interval. This lim-

its the distance that water can move during a single time interval. In locations where there is a steep head gradient, high hydraulic conductivity, or low porosity, groundwater may realistically be able to move further than the distance to the neighbouring cell, i.e. the Courant number would be greater than 1 (Courant et al., 1956). This results in unreasonable volumes of water attempting to move cell-to-cell. The effects are most prominently visible on steep, saturated hillslopes: water attempts to move downslope unreasonably fast, resulting in large, unrealistic amounts of exfiltration. Model results in this scenario may be significantly incorrect. The selection of a safe time interval will depend on the hydrologic and topographic specifications of the input data. A visual representation of safe, unsafe, and stable time intervals is given is Figure 3.

Users can select a safe time interval by calculating the maximum velocity of water movement at their study location, and comparing this with the size of the cells in their input data, using the Darcy velocity and porosity as follows (Freeze & Cherry, 1979):

$$\Delta t_s = \frac{L_{EW}}{u * f_a}, \tag{8}$$

where Δt_s is the safe time interval, L_{EW} is the cellsize in the East-West direction, u is the velocity of the groundwater, and f_a is a factor of safety. u is given by:

$$u = -\frac{k\frac{dh}{dl}}{\phi}, \tag{9}$$

where k is the hydraulic conductivity, $\frac{dh}{dl}$ is the head gradient, and ϕ is the effective porosity.

It is not possible to know ahead of time what the maximum head gradient between two cells within the course of a model run will be. We recommend using the topographic slope of the input data as a guideline: if z_{wr} initialises at the land surface, this will represent the starting head gradient. However, steeper head gradients can occur during the course of a model run, especially immediately after the dynamic lake component described in Section 2.2 has executed. Because of this and any other factors that may result in a higher water velocity, we recommend including the factor of safety, f_a, in the calculation. In Figure 3, the safe time interval shown was obtained using a conservative factor of safety of 5, although values as low as 2.1 would have precluded the obviously unsafe time intervals. If possible, a model user should compare results at a significantly shorter time interval to confirm that there are not major differences in the broad patterns of groundwater in the result, which would be an indication that the time interval they had selected was unsafe.

2.1.2.4 Compute discharge and update water table We calculate discharge from each target cell using the elevation head, i.e. the elevation above sea level of the water table, for the target cell and its four neighbouring cells. Along coastlines, sea level provides a 0-elevation boundary condition for hydraulic head and z_{wr} is set equal to 0 at these locations. Following Fan and Miguez-Macho (2011), we use Darcy's Law and the Dupuit–Forchheimer Approximation (Freeze & Cherry, 1979) to compute the discharge from each target cell as follows:

$$\begin{aligned}
Q_N &= \frac{\left(\frac{T_N+T}{2}\right) \times (h_N - h) \times L_{EW}}{L_{NS}}, \\
Q_S &= \frac{\left(\frac{T_S+T}{2}\right) \times (h_S - h) \times L_{EW}}{L_{NS}}, \\
Q_E &= \frac{\left(\frac{T_E+T}{2}\right) \times (h_E - h) \times L_{NS}}{L_{EW}}, \\
Q_W &= \frac{\left(\frac{T_W+T}{2}\right) \times (h_W - h) \times L_{NS}}{L_{EW}},
\end{aligned} \tag{10}$$

where Q_N, Q_S, Q_E and Q_W are the groundwater discharges to the north, south, east, and west; T represents the transmissivity of the target cell and T_N, T_S, T_E and T_W are the transmissivities in the four neighbouring cells; h and h_N, h_S, h_E and h_W are the elevation heads of these five cells; L_{NS} is the distance between the centres of two cells in the north-south direction; and L_{EW} is the distance between the centres of two cells in the east-west direction.

The change in water volume in the target cell, ΔV, is given by

$$\Delta V = (Q_N + Q_S + Q_E + Q_W) \times \Delta t, \tag{11}$$

where Δt is the stable time interval in seconds.

Based on these volumes, we calculate the change in relative water table elevation (z_{wr}) in the target cell and its four neighbours, dependent on the area of each cell. Change below the land surface is modulated by the porosity in that cell. Since most water tables are relatively near the land surface, we assume that porosity remains constant with depth and that above the land surface porosity is equal to 1. By doing this, we assume that vegetation occupies negligible volume.

There are four cases to consider when calculating the final relative water-table elevation, z_{wf}, in the target cell. These are:

1. The initial relative water-table elevation, z_{wi}, is above the land surface and the water table remains above the land surface after the change in water table height. In this case, z_{wf} is calculated as
$$z_{wf} = z_{wi} + \frac{\Delta V}{A}, \tag{12}$$
where A is the area of the cell.
2. z_{wi} was above the land surface, but the water table moves below the land surface during this step. In this case, we will take the porosity, ϕ, of the subsurface into account as follows:
$$z_{wf} = \frac{z_{wi} + \frac{\Delta V}{A}}{\phi}. \tag{13}$$
3. z_{wi} was below the land surface, and z_{wf} is above the land surface. In this case, we account for porosity only in the portion filled up to the land surface:
$$z_{wf} = (z_{wi} \times \phi) + \frac{\Delta V}{A}. \tag{14}$$
4. Finally, z_{wi} and z_{wf} are both below the land surface:
$$z_{wf} = z_{wi} + \frac{1}{\phi}\frac{\Delta V}{A}. \tag{15}$$

A new array records z_{wf} for each cell. After performing these calculations at each cell in the array, the groundwater component adjusts the water table for the entire array in a single step.

2.2 The dynamic lake component

The dynamic lake component uses a parsimonious network-based approach to move surface water into depressions and compute surface-water storage within these depressions. Depressions are defined as inwardly-draining regions within the topography, where water would naturally pool without being able to flow away. The dynamic lake component proceeds in two steps: (1) it uses a digital elevation model (DEM) to build a Depression Hierarchy data structure (Barnes et al., 2020)—a directed tree of depressions and their connections, and (2) it uses the Fill–Spill–Merge method to rapidly allocate runoff to these depressions (Barnes et al., 2021), with optional infiltration occurring, and to calculate the resulting depth of surface water in all of the depressions.

2.2.1 The depression hierarchy

Understanding the topological and geographical relationships between depressions in the landscape allows us to more rapidly calculate how these depressions will trap and store water. An unfilled depression will retain water that flows into it, while a depression that is already filled with water will overflow. We use the *depression hierarchy* data structure (see Barnes et al., 2020; Barnes & Callaghan, 2019, for a full description) to aid in our computation of surface-water flow using Fill–Spill–Merge, discussed in Section 2.2.2. This depression hierarchy is scale independent, though the data may be more or less appropriate for accurate representation of depressions in the landscape.

As inputs, the depression hierarchy requires 2D arrays of topography and a binary mask indicating where land and ocean cells occur. The following are supplied as outputs:

- An array of flow directions.
- An array of labels assigning each cell in the domain to a certain depression, or to the ocean in the case of ocean-draining cells.
- A depression hierarchy, which defines the topological and geographical relationships among depressions in the landscape. This depression hierarchy stores pertinent information about each depression, including its volume and the locations and elevations of its pit (i.e., lowest elevation) and outlet cells. A full list of the recorded information is given in Appendix A.

The depression-hierarchy algorithm builds the *depression hierarchy* data structure, organized as a forest of binary trees (Barnes et al., 2020), by analysing the input topography file to determine the locations of internally-drained depressions and their catchments. Smaller depressions may be nested within larger ones. Each depression, when completely filled with water, will overflow either to the ocean or to another depression. We use these topological relationships of natural depressions in the landscape to one another to construct the depression hierarchy.

Here we modify the original depression-hierarchy code described by Barnes et al. (2020) in two critical ways to permit global calculations of water-table elevation. First, we generalize the code to allow for latitude-dependent variable cell sizes (Callaghan et al., 2020). This is necessary for work with unprojected geospatial data on a continental scale, in which the variable cell size can have a significant impact on the amount of water accommodated within a cell. Second, we include storage volume below the land surface when assessing the total capacity for water available in a depression. This is necessary for the model coupling with the groundwater component described in Section 2.1, because water tables may be below the land surface and the ground needs to become saturated before surface water begins to fill the depression.

2.2.2 Fill–Spill–Merge

The Fill–Spill–Merge (FSM) algorithm rapidly simulates the movement of surface-water downslope into depressions and ultimately assigns any surface water that exists on the landscape to a specific depression from the depression hierarchy. A depression that contains more water than it can accommodate will spill, sending any additional water to its neighbouring depression. If two neighbouring depressions are both filled, they will merge to form a larger metadepression, which will then continue to fill with water. This process continues until all surface water flows either to a depression or to the ocean. Barnes et al. (2021); Barnes and Callaghan (2020) provide a full description of the original FSM algorithm. The combination of depression-hierarchies and FSM solves the above flow routing and water distribution problem thousands of times faster than previous models (Barnes et al., 2020, 2021). Here, we add optional infiltration, discussed in section 2.2.2.1, and we allow cell size to vary with latitude, discussed in section 2.2.2.3.

FSM is time-independent, always moving surface water to its final destinations in depressions, the ocean, or out of the model domain within a single time interval. This is based on the assumption that surface-water movement is fast in comparison to that of groundwater, and that only equilibrium surface-water results are needed over the timescales we address using the WTM.

The computational efficiency of Fill–Spill–Merge—it works thousands of times faster than previous approaches—makes this a convenient algorithm to use in a variety of hydrologic studies. Understanding the location of filled depressions on the landscape reveals information about hydrologic connectivity, as well as potential water storage volume in natural depressions. The FSM algorithm is summarised in Figure 4.

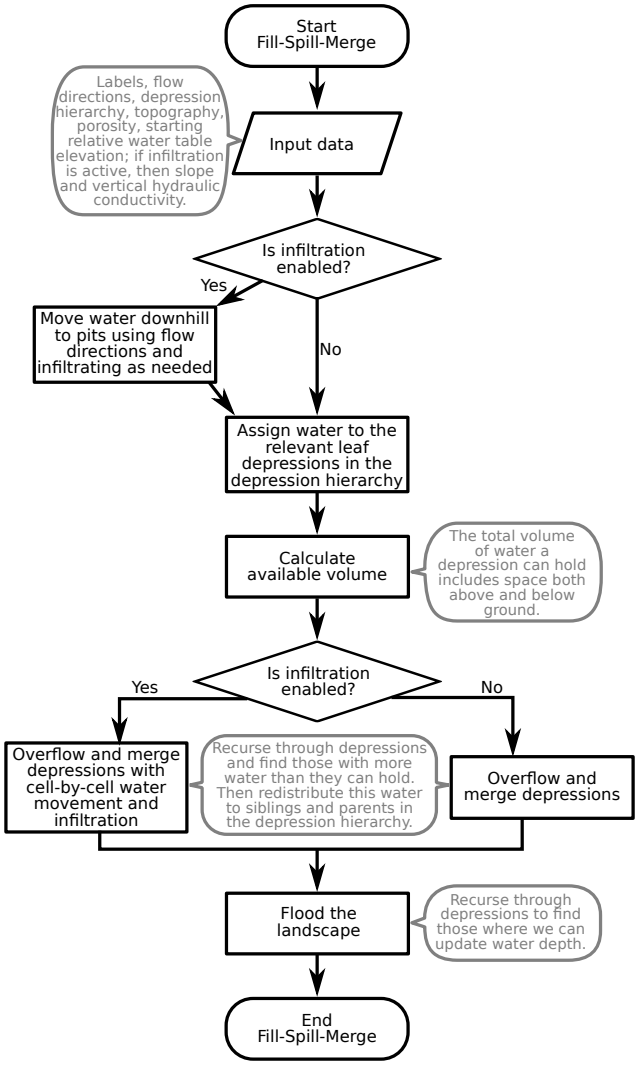

Figure 4. Steps taken by the Fill–Spill–Merge algorithm. We modified the Barnes et al. (2021) algorithm to include optional infiltration during downslope flow. Additionally, seepage will always occur beneath standing water during the 'flood the landscape' step (i.e. a lake will never represent a perched aquifer: there is a single water table at the lake surface). Figure 5 shows how FSM fits into the model coupling.

The following 2D, horizontally distributed arrays are required as inputs for FSM. Note that several of these are the same inputs required by the groundwater component:

- **Labels:** The depression-hierarchy algorithm gives each cell a number, here called its 'label', that associates it with a particular depression.
- **Topography:** Land elevation above sea level, in metres.
- **Porosity:** Shallow subsurface porosity.
- **Starting relative water table elevation (z_{wr}):** The elevation of the water-table relative to the land surface, as output by the groundwater algorithm. In cells containing surface water, we will refer to this water as 'runoff'.

In addition, FSM requires the **Depression hierarchy:** The data structure created by the depression-hierarchy algorithm, detailing all of the depressions and their topological relationships to one another.

If the infiltration option in Fill–Spill–Merge is enabled, additional required inputs are:

- **Flow directions:** The 2D array of flow directions created by the depression hierarchy algorithm.
- **Slope:** Topographic gradient, which should be computed from the input topographic data.
- **Shallow subsurface hydraulic conductivity – vertical:** Vertical hydraulic conductivity for infiltration calculations.

The output of FSM is an array showing the updated z_{wr}, after infiltration has (optionally) occurred and surface water has either flowed into depressions to form lakes or exited the domain.

2.2.2.1 Infiltration Here, we add an optional infiltration component to FSM. This is only recommended for cases in which the input data have a high enough resolution to resolve hillslopes and river channels that wholly occupy distinct individual cells. When using coarser resolution input data, a single pixel will contain sections of both river network and hillslope, and the model will not have sufficient information about the transit routes and times of water across these different zones, themselves determined by drainage density and hillslope geometry, to realistically simulate infiltration. When input-data resolution becomes high enough to differentiate these hillslope and channel components of the landscape, the infiltration component adds an additional element of realism to the model.

In cases in which the user does not want infiltration to occur during surface flow, the land surface will be treated as impermeable in order to simulate rapid evacuation of surface water from each cell via river networks. The algorithm will bypass its subroutine to compute cell-to-cell flow of water in order to speed calculations. Using the labels array and depression hierarchy data created by the depression hierarchy algorithm, FSM will move water directly from each surface-water-containing cell to the relevant depression in the hierarchy.

When the infiltration option is enabled, the FSM algorithm first moves surface water downslope cell-by-cell, using the flow directions generated by the depression hierarchy. As the water moves downslope, some may infiltrate; the remainder continues along the flowpath until it flows into the ocean, out of the domain, or into a pit cell (that is, the cell within a depression that has the lowest elevation). Similarly, when water overflows from one depression to another later in the algorithm, the water will move cell-to-cell and infiltration will be computed (see 'Overflow and merge depressions with cell-by-cell water movement and infiltration' in Figure 4).

Our method for managing infiltration can accommodate both saturation-excess and infiltration-excess overland flow, while omitting the dynamics of the unsaturated (i.e. vadose) zone. When surface water flows into a cell, the following possibilities exist:

1. The cell is fully saturated, i.e. the groundwater table is at the land surface. In this case, no infiltration is possible, and saturation-excess overland flow (Dunne & Black, 1970) will occur.
2. The cell is not fully saturated. Some water will infiltrate and some will continue to flow downslope as infiltration-excess overland flow (Horton & Htrata, 1955). The amount of infiltration that occurs is dependent on infiltration rate and residence time of water in a cell, discussed further below.

When water reaches a cell that is not fully groundwater-saturated, we calculate the time t that it takes the water to cross the cell as a function of the distance travelled by the water from cell to cell (L) and the depth-integrated flow velocity (u). We compute this velocity using Manning's equation (Eq. B1), making it a function of slope (S) and flow depth (h). This flow depth decreases as the water crosses the cell due to infiltration at a rate governed by the saturated vertical hydraulic conductivity (k_{sat}); for simplicity, we do not consider transient wetting and drying effects in the unsaturated zone. We limit the topographic slope, S, to a minimum value of 10^{-6} to allow movement over flat cells in the DEM. We calculate L based on the directions of travel between the two cells (north–south, east–west, or diagonal), and the latitude of the cells.

We obtain the transit time t for each cell following the derivation in Appendix B:

$$t = \left[h_0 - \left(h_0^{5/3} - \frac{5}{3} \frac{n}{S^{1/2}} k_{\text{sat}} L \right)^{3/5} \right] \Bigg/ k_{\text{sat}}. \tag{16}$$

From this travel time, we next want to find the amount of water that will infiltrate. There are three possible solutions for the potential total amount of water infiltrated, I_{pot}:

$$I_{\text{pot}} = \begin{cases} h_0 & \text{if } h_0^{5/3} \leq \frac{5}{3} \frac{n}{S^{1/2}} k_{\text{sat}} \\ 0 & \text{if } k_{sat} = 0 \\ k_{\text{sat}} t & \text{otherwise.} \end{cases} \tag{17}$$

In the first case, the entire column of water that enters the cell can infiltrate before it crosses. (Indeed, you may already have noted that the solution to Equation 16 becomes undefined in this case.) In the second case, a vertical hydraulic conductivity of 0 means that no water is able to infiltrate. In the third, some of the surface water infiltrates as it crosses the cell, and the remainder continues to the next cell.

Converting I_{pot} to the actual amount of infiltration that occurs, I, requires consideration of the space available to accommodate infiltration water. Combining Eq. 17 with the amount of groundwater space available in the cell, given by $-\phi z_{wr}$ where ϕ is the subsurface porosity (assumed constant) and z_{wr} is the relative water table elevation, provides the general solution:

$$I = \min\left(-\phi z_{wr} \ , \ I_{\text{pot}}\right). \tag{18}$$

This amount of infiltrated water is then subtracted from the flow depth, h. If $h > 0$ as the water exits the cell, then it continues onwards to the next downslope cell.

2.2.2.2 Seepage When a lake is present in a depression, we allow the water column to instantaneously seep into the subsurface until either (a) the full subsurface is saturated or (b) no surface water remains. The WTM does not simulate any perched water tables; a lake surface represents the water table with complete saturation up to that elevation.

2.2.2.3 Variable cell areas When performing computations using a latitude–longitude grid, cells at higher latitudes have smaller areas than cells at lower latitudes due to the roughly spherical shape of the Earth. The same volume of water at each of these two latitudes would translate to a different thickness of ground- or surface water in a cell. We conserve water volume within the model and account for this variable cell area when calculating z_{wr}.

2.2.2.4 Updating the relative water table elevation If a depression is completely full, its water level is set to the outlet level of the depression. If a depression is only partially full, we calculate its water level using the volume of water and the elevations and areas of each cell following equation 6 from Barnes et al. (2021):

$$z_{wr} = \frac{V_w}{\sum_i A_i} + \frac{\sum_i A_i z_i}{\sum_i A_i}, \tag{19}$$

where z_{wr} is the relative water table elevation, V_w is the volume of water in the depression, A_i is the area of cell i in the depression that contains water, and z_i is the elevation of cell i. See Appendix C for the derivation of this formula.

2.3 Model coupling

The WTM couples the two model components by running the groundwater and dynamic lake components sequentially in a repeated cycle. Figure 5 demonstrates the steps followed within the coupled model. Figure 6 shows an example topography with depressions and the conditions at each step through the WTM.

A user may elect to perform either a steady-state or a transient model run. For a steady-state model run, the user should monitor the amount of change in z_{wr} from one model cycle to the next. Values indicating the total change in the array are output in a text file. The total change in z_{wr} should approach an asymptote as it nears steady-state. For these model runs, a single set of climate and topography input arrays are required. An example of the progression of a model run as it moves towards steady-state is shown in Figure 7.

For transient runs, the WTM requires two sets of input arrays, which define the state of climate and topography at both the beginning and the end of the model run period. These parameters will then be adjusted from starting state to the ending state using a linear interpolation throughout the model run. If the topography changes during the time period to which the model is being applied, the depression hierarchy will be recalculated accordingly. The WTM also requires a starting z_{wr} array representing water-table conditions at the beginning of the modelled period. The user will select the total real-world length of time for which to run the model.

3 Sample model runs

3.1 Synthetic data

We provide sample results on a synthetic dataset to make the results of the WTM easy to visualise. The topography consists of two hills and two depressions with a two-dimensional sinusoidal shape (see Figure 8a). To focus on the interaction of topography with the WTM, we used constant values across the entire domain for the other inputs: $0.03\,\mathrm{m\,yr^{-1}}$ precipitation, $0.0\,\mathrm{m\,yr^{-1}}$ evapotranspiration, $0.5\,\mathrm{m\,yr^{-1}}$ open water evaporation, e-folding depth of 100 m, and a winter temperature of 0°C. Hydraulic conductivity was set to $0.0001\,\mathrm{m\,s^{-1}}$ and porosity to 0.25. The water table was initialised at the land surface, i.e. $z_{wr} = 0$.

Figure 8 shows the results of the model when applied to the synthetic topography. We use this visually simple example to compare results of running the groundwater com-

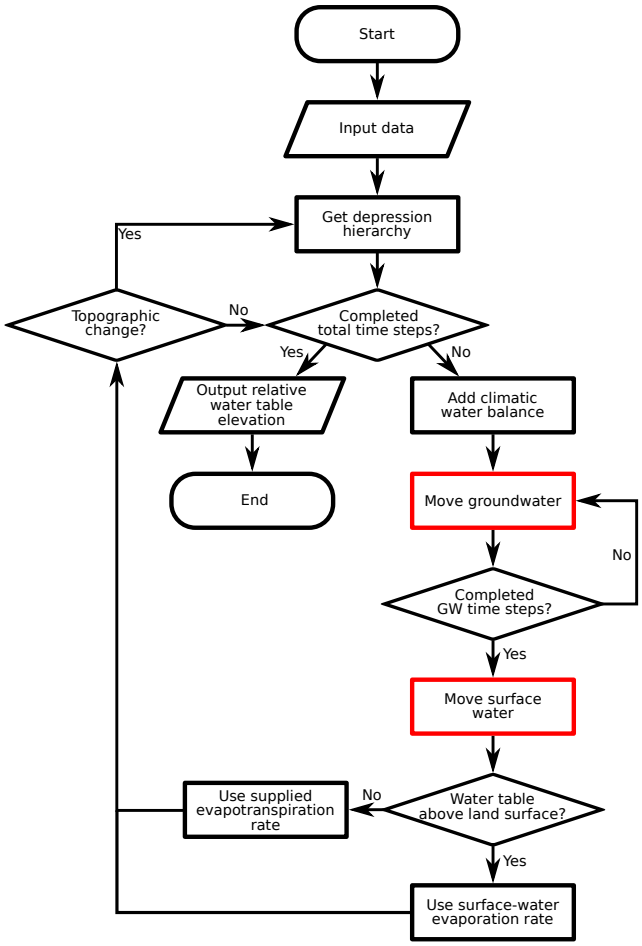

Figure 5. Steps taken in the coupled model. The two red boxes indicate the two main modelling components of our coupling. Figure 2 details the 'move groundwater' step and Figure 4 details the 'move surface water' step.

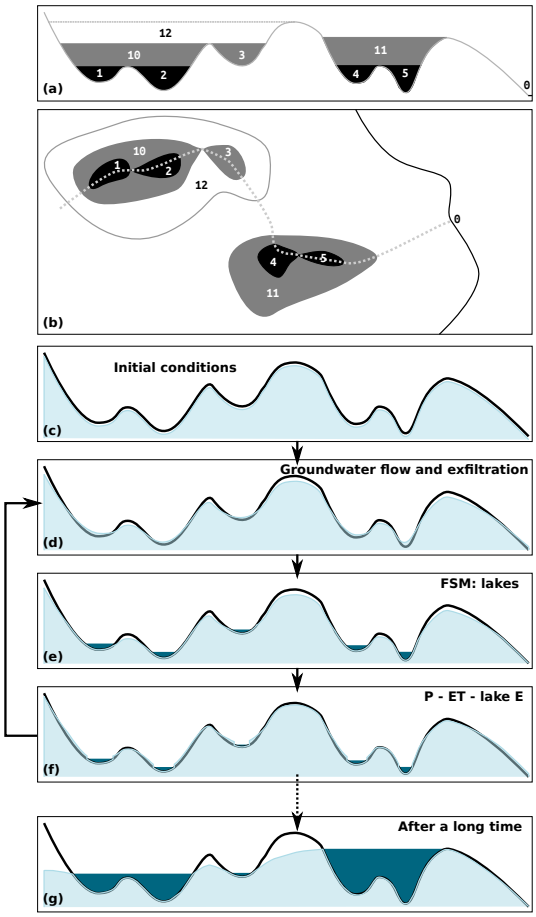

Figure 6. Water movement in the groundwater and dynamic lake components of the WTM. (a) A cross section across a portion of the landscape, showing the depressions and their labels within the depression hierarchy. (b) Map view of the depressions from (a). (c) Initial relative water table elevation conditions: the water table is at the land surface. (d) The groundwater component has executed. The water table is deeper below the hilltops and exfiltration has occurred on hillsides. (e) FSM has executed. Surface water is now distributed into lakes at the bottom of depressions. (f) P-ET is added over land surfaces, and P-E, which may be negative, is added over the lake surfaces. (g) After steps (d) to (f) have been repeated many times, we reach the steady-state result that is the output of the model. Higher precipitation towards the right-hand side of the figure has resulted in a higher water level in those depressions. Depressions 1 and 2 have merged and their metadepression, 10, has partially filled. Depression 3 has partially filled. Depressions 4 and 6 have merged into their metadepression 11, which has completely filled with water. Any overflow from depression 11 has flowed into the ocean. Figure modified from Barnes et al. (2021).

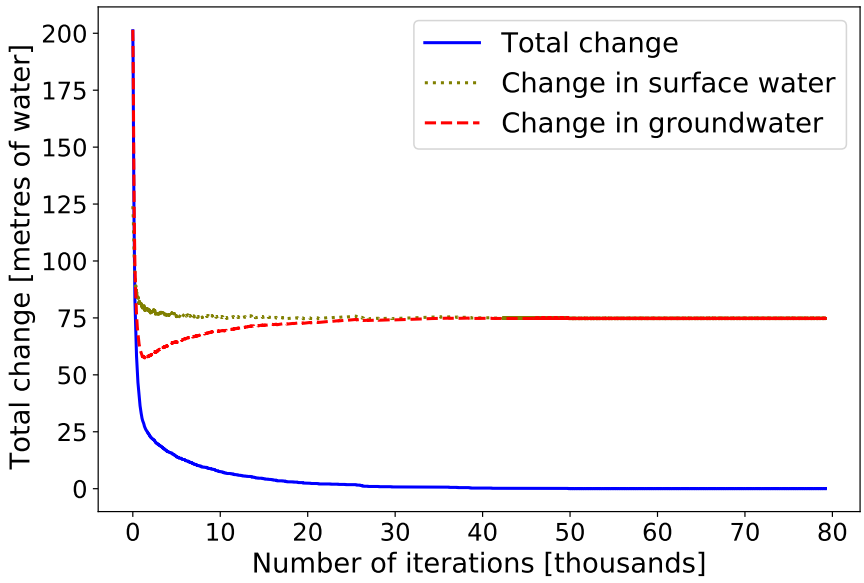

Figure 7. The absolute value of the summed change in the relative water table elevation array per iteration, showing the progression towards equilibrium. Both the total change and the slope of the line decrease the longer the model run continues. This was the result of a model run on artificial topography featuring sinusoidally-shaped depressions, discussed in Section 3.1.

ponent in isolation, to results from the full coupled WTM. We also provide an example of water depths in completely filled depressions produced using RichDEM's complete-depression-filling command (Barnes et al., 2014a, 2014b; Barnes, 2016), for comparison to the partially-filled depressions that result from the WTM's surface-water evaporation.

The results emphasize the importance of including the dynamic lake component in the WTM. The z_{wr} values obtained using the coupled model include not only the surfaces of lakes in the depressions themselves, but also a regionally higher water table due to the impact that the lake water has on surrounding relative water table elevations. When the lake surface is not present, the low z_{wr} at the floor of the depression results in continued exfiltration from the hillslopes, and hence a lower regional water table. The results also show that the high open-water evaporation, and the ability of lakes to recharge local groundwater, have kept lake levels low in comparison with a fully-filled depression. These factors permit internal drainage of the lakes.

3.2 Real-world model run: North America

Sample results from a real-world site are provided for the North American continent. Details on the input data used are given in Appendix D.

Using equations 8 and 9 discussed in Section 2.1.2, we computed the fastest transit time that we would expect to see across a cell to determine the safe time interval. Estimating the head gradient with the topographic gradient, and using a factor of safety of 3, we obtained a safe time interval of 216,228 s. Therefore, we performed the model run for North America using a time interval of 216,000 s (equivalent to 2.5 days). FSM executed after every 146 groundwater time intervals, i.e. once per year.

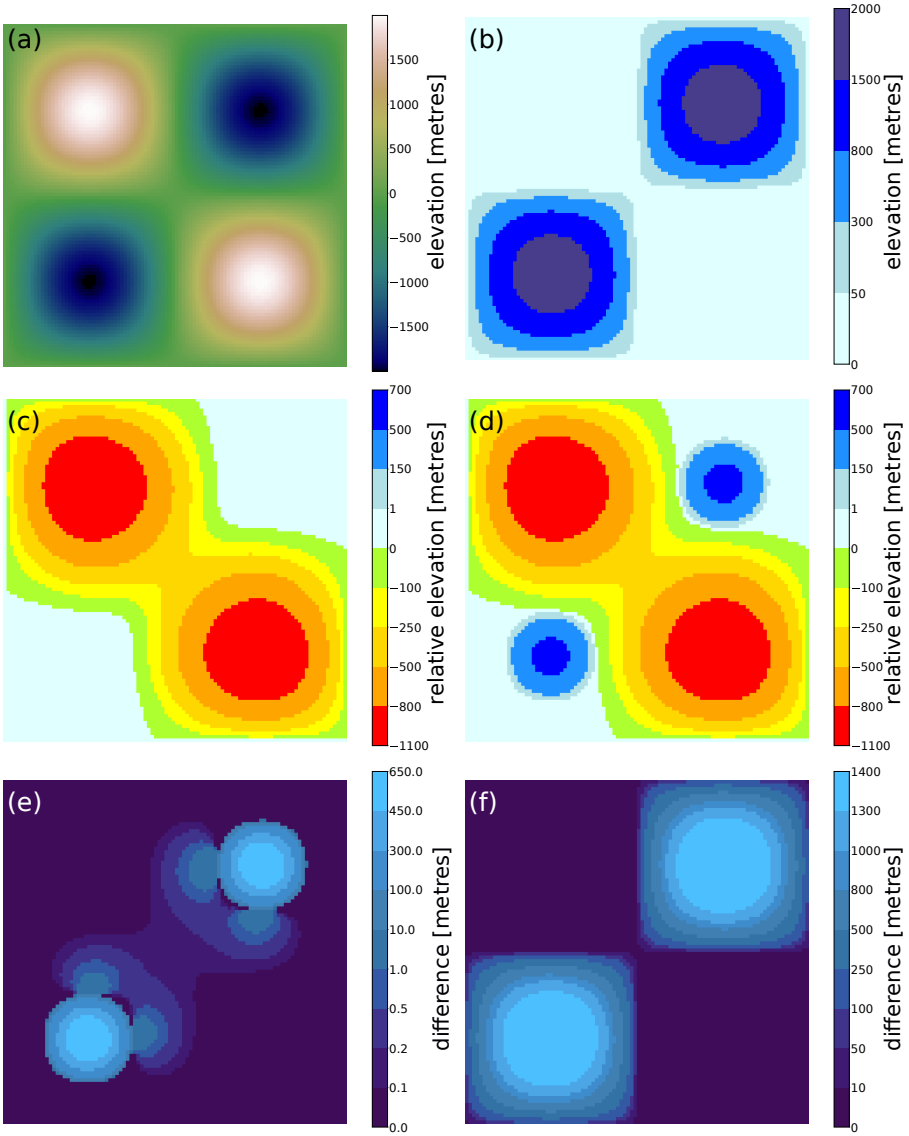

Figure 8. The results of the model on an artificial topography, with two sinusoidal depressions and two sinusoidal hills. (a) shows the artificial topography used in these sample results. (b) shows the water-table elevation resulting from a flood-fill algorithm, used to completely fill the depressions. In (c), only the groundwater component of the WTM was used. The maximum relative water table elevation is 0, since all surface water is assumed to run off or evaporate. Results are shown in metres relative to the land surface. In (d), the full WTM coupled model was used, so the two depressions contain surface water. While the biggest difference between these two is seen in the depressions themselves, (e) indicates that the actual differences are further-reaching: in the case where surface water is present, the local to regional groundwater table is also affected. (f) shows the difference between only the surface-water portion of the coupled model, and the result of a flood fill. This emphasizes the fact that the depressions were not completely filled by the coupled model, partially as a result of the surface-water evaporation term.

Figure 9 shows the sample model results for the continent of North America. Results of the full coupled model are compared with results of the groundwater component run in isolation, with no surface water being allowed to accumulate; simplifying the WTM in this way makes it identical to the Fan et al. (2013) model. When surface water is allowed to accumulate in the coupled model, additional water is stored in lakes, but local water tables are also higher in the surrounding cells (Figure 8) due to the higher elevation head of the water table within these lakes.

Both groundwater-only and coupled WTM simulations (Figure 9) capture broad climate-driven patterns in z_{wr} at a continental scale. The drier climate in the west results in deeper water tables while wetter climates in the north and east result in shallower water tables. Variable geology and topography add detail to this overall pattern driven by the climatic gradient.

The results from the coupled WTM simulation in Figure 9b show the importance of accounting for surface water in the model. North America contains tens of thousands of lakes. Many of these are reflected in the results, ranging from large water bodies, such as the Great Lakes, to smaller features such as internally-drained basin and range depressions, calderas, and fluvial lakes. These features are highly sensitive to the input data, particularly surface-water evaporation, and therefore improvements to the input data will result in more accurately simulated lakes.

Figure 10 compares the results of the WTM simulation with observed data across North America. Fan et al. (2007) gathered over 500000 observations of groundwater depth across the continent. These are compared to WTM groundwater depths in Figure 10a. Simulated water table depths tend to be shallower than observations, a trend that has already been observed by Fan et al. (2013) and Maxwell et al. (2015). The trend in simulated groundwater depth appears similar to these prior works, though a higher proportion of very shallow groundwater in this work likely results from the inclusion of the dynamic lakes module and associated groundwater seepage. A lack of very shallow water tables in the observed data may be because the observed data is likely to suffer from selection bias as a result of well location. Additionally, it includes anthropogenic impacts on the water table (e.g. pumping), while the simulated water table does not.

Lake depths were compared to the Kourzeneva et al. (2012) lake dataset. This comparison is shown in Figure 10b. The Kourzeneva et al. (2012) dataset includes full lake bathymetry for some major lakes, but only mean depth for most smaller lakes, making direct comparisons between the simulated and observed data difficult. Additionally, bathymetry is non-identical between the Kourzeneva et al. (2012) dataset and topography used for the WTM simulation: the Great Lakes bathymetry for the simulation was based on GEBCO Bathymetric Compilation Group (2020) data. Improvements in observed data in the future will enable us to better test simulated results.

4 Conclusions

The WTM computes changing water tables and terrestrial water storage across spatial and temporal scales. Using the depression hierarchy data structure and Fill–Spill–Merge to efficiently route surface water, we are able to create a continental-scale simulation of water table that includes locations where the water table is above the land surface. The WTM is ideal for modelling changes in the water table under changing climatic conditions, and the simple input requirements mean that it can be used for the distant past or for the future as climate changes.

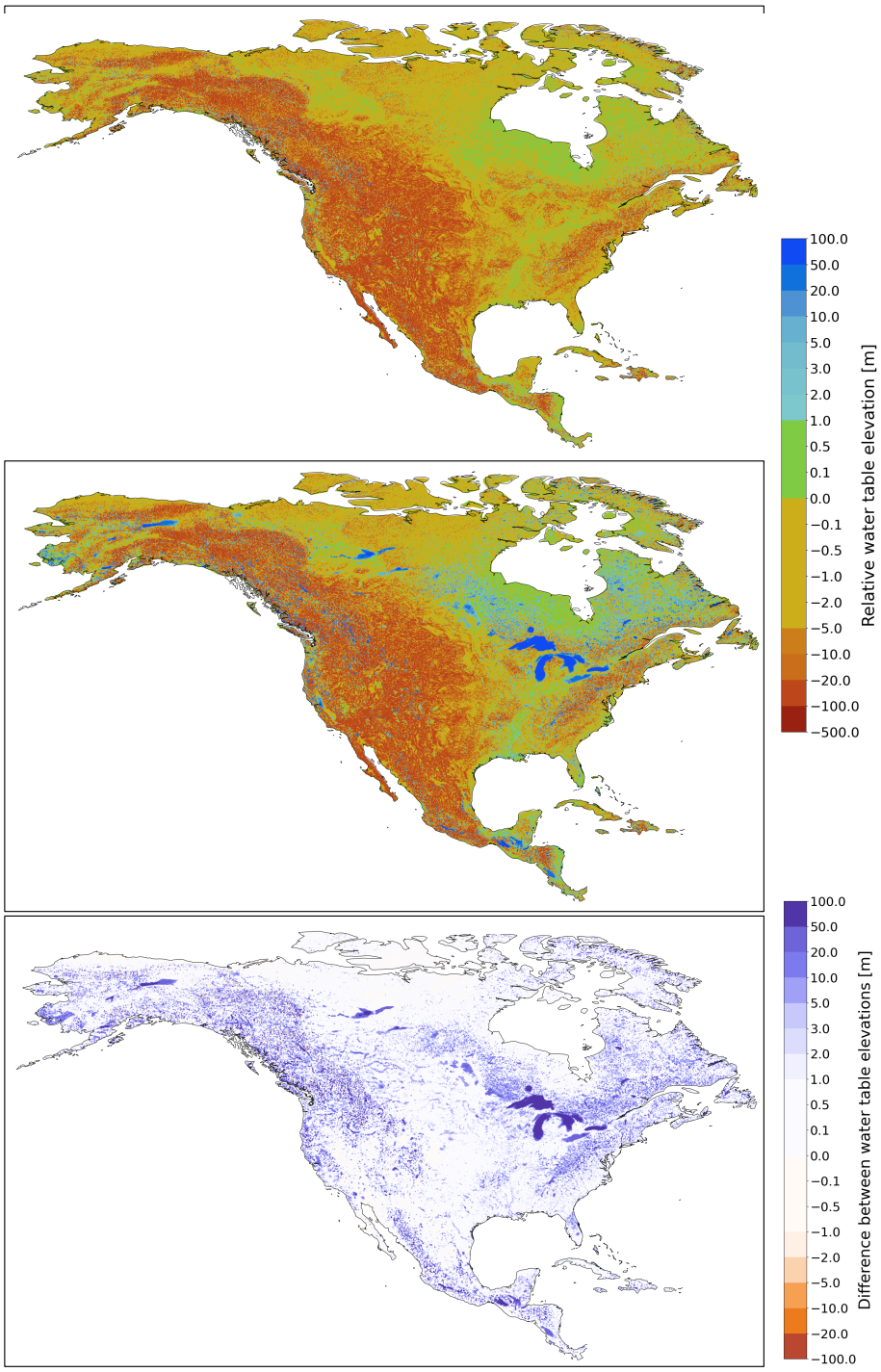

Figure 9. A comparison of the fully coupled WTM with the groundwater component alone for the North American continent. Model runs were performed using modern-day data as detailed in Appendix D. Shown here is the equilibrium water table elevation relative to the land surface. (a) shows the result of the groundwater component alone, where surface water is all assumed to run off or evaporate. (b) shows the result of the fully coupled model (WTM), including lakes. (c) shows the difference between the two.

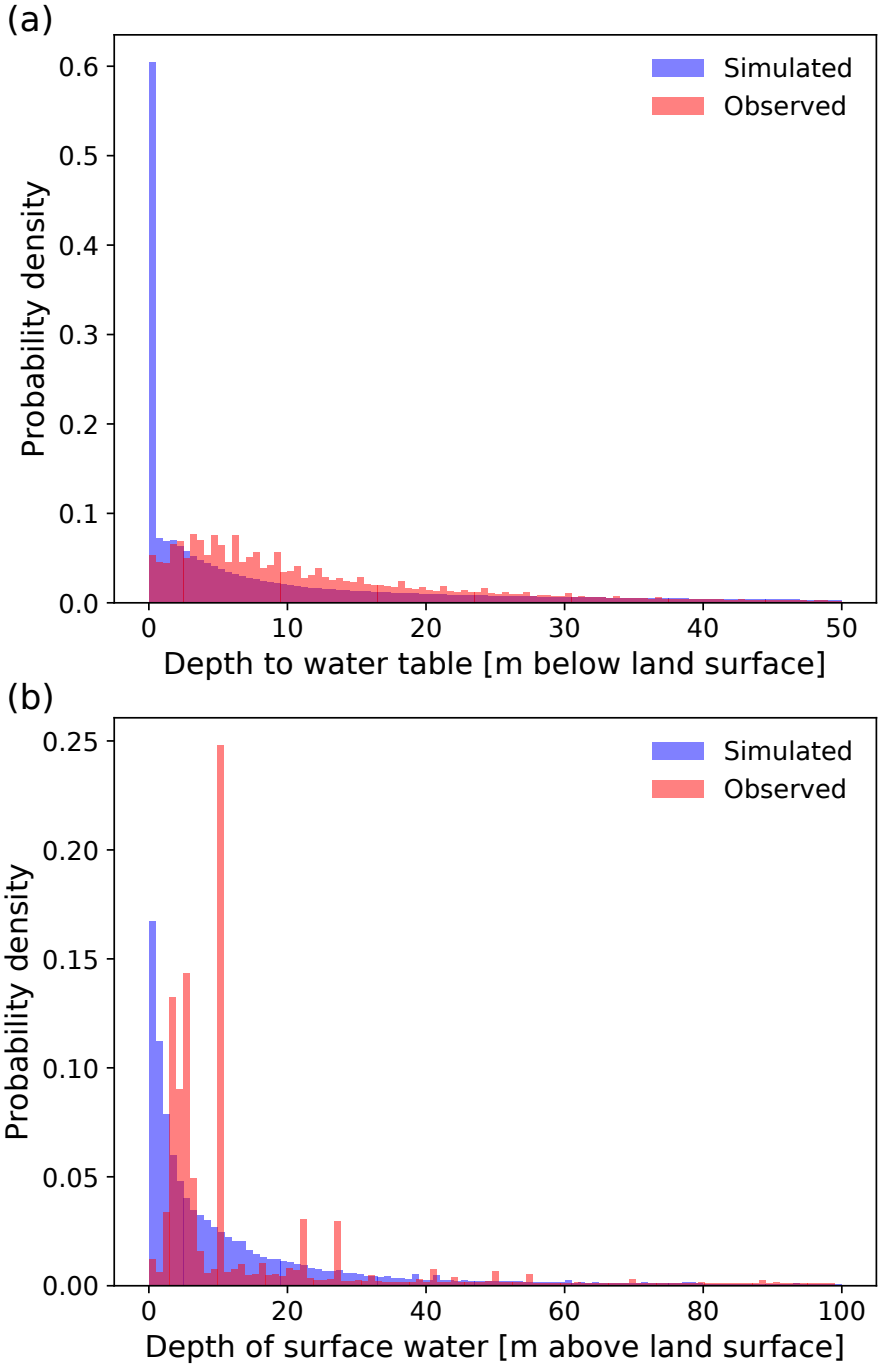

Figure 10. A comparison of the results of the WTM simulation with real-world data. (a) shows simulated and observed groundwater depths. (b) shows simulated and observed lake depths.

Appendix A Depression hierarchy data

The depression hierarchy stores data about each depression that may be useful for a variety of applications. Aside from the relationships of depressions to one another, data on the size, elevation, and location of each depression is also stored. A *leaf* depression refers to the lowest level of depressions in the hierarchy. These depressions do not contain any subdepressions. Any depression that is not a leaf depression will contain two subdepressions, referred to as its *children*. The depression is the *parent* of its children. For each depression, the data recorded comprise:

- The location of its pit cell (the lowest cell in the depression).
- The elevation above sea level of the pit cell.
- The location of its outlet cell, the lowest point at which the depression will spill over to a neighbouring depression or to the ocean.
- The elevation above sea level of the outlet cell.
- The label of the depression (a unique number).
- The number of cells in the depression.
- For leaf depressions, a list of all of the cell indices for cells that the depression contains.
- The area of the depression.
- The above-ground volume of the depression.
- The total volume of the depression, including any space in unsaturated ground.
- The difference between the above-ground and total volumes of the depression.
- The volume of water currently contained in the depression.
- The label of the depression that is its parent in the binary tree structure.
- The label of the leaf-level depression into which the depression overflows.
- The label of the upper-level depression into which the depression overflows. In some cases, this will be the same as the above item.
- For non-leaf depressions, the labels of the two children of the depression in the binary tree structure.
- Whether this depression overflows into a depression which ultimately overflows to the ocean.
- If this depression ultimately overflows into the ocean, then a list of any depressions which link to the ocean via this one.

Appendix B Transit time across a cell

To calculate the amount of infiltration that happens while water is in transit across a cell, we must consider the total time the water takes to cross the cell. The more time that the water spends in a cell, the longer it will have to infiltrate. Water will take longer to flow across cells that are larger or have shallower slopes, or when the water depth, and hence its flow velocity, is smaller.

We use Manning's equation to estimate the time taken for flow to cross a cell.

$$u = \frac{1}{n} R_h^{2/3} S^{1/2}, \tag{B1}$$

where u is the mean (i.e. vertically averaged) velocity of the surface water moving across the cell, n is the Gauckler–Manning coefficient, R_h is the hydraulic radius, and S is the slope. By default, we set Manning's n to a value of $0.05\, m^{-1/3} s$. We make the assumption that the height of water in the cell, h, is much smaller than the cell width. This allows us to simplify the hydraulic radius to equal h:

$$u = \frac{1}{n} h^{2/3} S^{1/2}. \tag{B2}$$

Because S and n are both constants, for convenience we will combine them in constant k_0, where

$$k_0 = \frac{S^{1/2}}{n}, \tag{B3}$$

so that

$$u = k_0 h^{2/3}. \tag{B4}$$

The next step is to consider the infiltration rate,

$$\frac{\mathrm{d}h}{\mathrm{d}t} = -k_{\mathrm{sat}}. \tag{B5}$$

By separating variables, integrating, and defining $h = h_0$ at $t = 0$, we obtain:

$$h = h_0 - k_{\mathrm{sat}} t. \tag{B6}$$

We substitute Eq. B6 into Eq. B4 and use the definition of velocity as the time derivative of position to set up the final equation to integrate:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = k_0 \left(h_0 - k_{\mathrm{sat}} t\right)^{2/3}. \tag{B7}$$

where L is the displacement in an arbitrary orientation. By separating variables and solving via u substitution, we obtain:

$$\begin{aligned} L &= k_0 \int_0^{t_i} (h_0 - k_{\mathrm{sat}} t)^{2/3} dt \\ &= -\frac{3}{5}\frac{k_0}{k_{\mathrm{sat}}}(h_0 - k_{\mathrm{sat}} t)^{5/3} + c, \end{aligned} \tag{B8}$$

where c is the constant of integration. Defining $L = 0$ when $t = 0$ (i.e. that the clock starts when the water first touches the cell margin), we obtain:

$$c = \frac{3}{5}\frac{k_0}{k_{\mathrm{sat}}} h_0^{5/3} \tag{B9}$$

This gives the distance crossed by the water as:

$$L = \frac{3}{5}\frac{k_0}{k_{\mathrm{sat}}}\left(h_0^{5/3} - (h_0 - k_{\mathrm{sat}} t)^{5/3}\right) \tag{B10}$$

We now know the distance that the water must travel from one side of the cell to the other. We want to know the amount of time that this transit takes, because this is the amount of time that the water has to infiltrate within the cell. Rearranging to solve for the transit time and substituting S and n back in gives

$$t = \left[h_0 - \left(h_0^{5/3} - \frac{5}{3}\frac{n}{S^{1/2}} k_{\mathrm{sat}} \Delta L\right)^{3/5}\right] \Big/ k_{\mathrm{sat}}. \tag{B11}$$

Appendix C Calculating lake-water levels

The depression hierarchy is used to distribute water to the appropriate depressions. Within each depression, we need to determine the geographic distribution and depth of this water. We do so by calculating the elevation of the water surface within a given depression. In cases where a depression is filled to the brim, this elevation is simply the elevation of the outlet cell. In cases where the depression is not filled to the brim, a calculation needs to be performed as follows:

V_w: Total volume of water contained in the depression

592 z_w: Elevation of the water surface
593 z_i: Topographic elevation in cell i
594 A_i: Area of cell i

We know the volume of water, topographic elevation of each cell, and area of each cell, and need to find the elevation of the water surface. We know that the volume of the depression can be calculated with:

$$\begin{aligned} V_w &= \sum_i A_i \left(z_w - z_i\right) \\ &= \sum_i A_i z_w - \sum_i A_i z_i \\ &= z_w \sum_i A_i - \sum_i A_i z_i \end{aligned}$$

We then separate z_w and algebraically rearrange.

$$z_w = \frac{V_w}{\sum_i A_i} + \frac{\sum_i A_i z_i}{\sum_i A_i}$$

595 All terms on the right-hand side are known.

596 Appendix D Model input data: North America

597 Here we provide the data sources that we used for each of the input arrays required
598 by the WTM for the continent of North America.

599 We obtained topographic data from the GEBCO 2020 grid (GEBCO Bathymet-
600 ric Compilation Group, 2020). We modified this topography using lake bathymetry from
601 the Global Lake Database (Kourzeneva et al., 2012), using all included lakes except for
602 the Great Lakes, which are already included in GEBCO 2020, and the Great Salt Lake.
603 We updated the bathymetry of the Great Salt Lake using data from Tarboton (2017).
604 We created the slope input files using this topography and GRASS GIS (Neteler et al.,
605 2012).

606 We obtained precipitation and evapotranspiration data from the Terraclimate dataset
607 (Abatzoglou et al., 2018). We averaged monthly data from Terraclimate, applying a weight-
608 ing based on month length, over a total of 30 years, from 1981 to 2010 inclusive to ob-
609 tain annual averages. The spatial resolution of the Terraclimate data was resampled from
610 1/24 degrees to 30 arcseconds using a bivariate spline approximation.

611 We used the ERA5 reanalysis monthly mean 0.25 degree latitude-longitude grid
612 data for winter temperature (European Centre for Medium-Range Weather Forecasts,
613 2019). The data is a long-term annual average, based on monthly averages from 1979
614 to 2018 inclusive. We used monthly temperatures from December, January and Febru-
615 ary for the Northern hemisphere. We resampled the data from 0.25 degree resolution to
616 a 30 arcsecond resolution using the topography and an adiabatic lapse rate of 5° C/km,
617 representative of a wet adiabatic lapse rate (Peirce et al., 1998).

618 Hydraulic conductivity and porosity values are both based on the hybrid STATSGO/FAO
619 soil texture database available at `https://ral.ucar.edu/solutions/products/wrf`
620 `-noah-noah-mp-modeling-system` (last accessed: 10 November 2020), which give 12
621 different soil texture categories. We converted these to hydraulic conductivity and poros-
622 ity values using the representative values suggested by Clapp and Hornberger (1978). The
623 value for silt was not provided in Clapp and Hornberger (1978), and was estimated based
624 on other nearby values and the range of possible values given by Earle (2015). Similarly,
625 the value for bedrock was selected from the range given by Earle (2015). The value for
626 'organic materials' was taken from the value listed as 'peat' from Fan et al. (2007). We

accounted for anisotropy between vertical and horizontal hydraulic conductivity, using the anisotropy ratios listed by Fan et al. (2007).

We calculate evaporation of surface water using the classic Penman (1948) equation, modified following Hersbach (2011) to account for variable water-surface roughness due to wind-driven waves:

$$E = \frac{R_n + (c_p \rho_a u_*^2)/(\Delta_{P,T} u)}{\rho_w \Delta H_{\text{vap}} + P c_p \rho_w (R_v/R_a)} \left(e_{\text{sat}} - e_a\right). \tag{D1}$$

Here, E is the rate of open-water evaporation, R_n is net solar radiation, c_p is the specific heat capacity of air at constant pressure, ρ_a is air density, u_* is wind shear velocity, $\Delta_{P,T}$ is the gradient in temperature–pressure space of the liquid-to-vapor phase transition for water, u is wind velocity (typically at 2 meters elevation above the surface), ρ_w is water density, ΔH_{vap} is the latent heat of vaporization of water, P is atmospheric pressure, $R_v/R_a = 1/0.622$ is the ratio of the gas constants of water vapor and air, e_{sat} is water vapor pressure at saturation, and e_a is water vapor pressure. Appendix E holds our derivation.

The open-water evaporation calculations were based on data from TerraClimate (Abatzoglou et al., 2018) and the GEBCO Bathymetric Compilation Group (2020) elevation data set. The open-water evaporation rates were calculated from monthly climatic data from 1958 to 1970, inclusive.

Calibration constants for the e-folding depth were set to $a = 100$, $b = 150$, and $f_{\min} = 2.5$, following Fan et al. (2013).

Appendix E Surface-water evaporation

We calculate open-water evaporation by solving and applying the Penman Equation (Dingman, 1994) alongside the Charnock (1955) expression for the roughness length over open water as a function of wind-induced waves. This evaporation rate overrides the input evapotranspiration rate wherever the water table crops out above the surface, forming an exposed water body (Fig. 5).

The Penman (1948) Equation combines radiative, sensible, and latent heat transfer to solve for evaporation. Though it is well-established (Finch & Calver, 2008; Valiantzas, 2006; Vörösmarty et al., 1998; Zotarelli & Dukes, 2010), we choose to include a brief derivation of the Penman equation due to (1) the central role played by evaporation in our study; (2) the fact that most derivations center on the Penman–Monteith equation (Monteith, 1965), which involves plant transpiration that is not relevant to our application to lakes; and (3) our inclusion of a wind-speed-determined roughness length to modulate wind-driven turbulent energy transfers, which seems reasonable to include but that we have not found in our review of the literature. Here we use variable nomenclature that is more common to thermodynamics than to hydrology.

E1 Penman Equation (general form)

The Penman Equation relates evaporation rate (E), which is a latent-heat flux, to net-radiation flux (R_n: incoming and outgoing shortwave and longwave) and sensible heat flux due to turbulent atmospheric heat transfer ($Q_{H,s}$, where subscript H indicates enthalpy and s indicates that it is sensible):

$$E = \frac{R_n - Q_{H,s}}{\rho_w \Delta H_{\text{vap}}}. \tag{E1}$$

Here, ρ_w is water density, and ΔH_{vap} is latent heat of vaporization of water. These terms in the denominator act to convert the energy fluxes [$\mathrm{W\,m^{-2}}$] into evaporation [$\mathrm{m\,s^{-1}}$].

E2 Input data products

Inputs for our solution come from the TerraClimate and GEBCO_2020 datasets. TerraClimate (Abatzoglou et al., 2018) comprises monthly 2.5-arcminute (∼5 km N–S) gridded data products for:

- Incoming solar (shortwave) radiation
- Monthly averaged minimum and maximum daily temperatures
- Wind speed
- Vapor pressure

GEBCO_2020 (GEBCO Bathymetric Compilation Group, 2020) is a 30-arcsecond (∼1 km N–S) global gridded topographic and bathymetric data set. We resample this to 2.5 arcminutes to match the resolution of TerraClimate.

E3 Net radiation

In the field, acquiring net radiation requires paired upward- and downward-facing pyranometers and pyrgeometers to measure incoming and outgoing shortwave and longwave radiation. Here we use a combination of calculations and remotely sensed data products to assemble a solar-radiation data product at an appropriate resolution for our contiental-scale modeling example.

TerraClimate (Abatzoglou et al., 2018) provides the incoming shortwave radiation flux, $R_{\mathrm{in},s}$. Outgoing shortwave radiation equals the incoming radiation times the surface albedo α. Therefore, net shortwave radiation, $R_{n,s}$, is given by

$$R_{n,s} = (1 - \alpha) R_{\mathrm{in},s}. \tag{E2}$$

We use $\alpha = 0.06$ as characteristic of open water.

We lack data on net longwave radiation, $R_{n,l}$, but know that (1) outgoing longwave flux is proportional to surface temperature via the Stefan–Boltzmann Law and (2) that incoming longwave radiation is related to greenhouse gases in the atmosphere that absorb and re-emit this outgoing radiation. We therefore follow and modify the approach taken by Zotarelli and Dukes (2010) in approximating the surface temperature by the maximum and minimum air-temperature values, and using vapor pressure and cloudiness to estimate the impact of greenhouse gases on longwave absorption and re-radiation:

$$R_{n,l} = \sigma \frac{T_{\mathrm{max}}^4 + T_{\mathrm{min}}^4}{2} \left(0.34 - 0.00014 e_a^{1/2}\right) \mathcal{C}. \tag{E3}$$

Here, σ is the Stefan–Boltzmann constant, T is temperature in Kelvin, e_a is the near-surface atmospheric vapor pressure, and $\mathcal{C}$ is what we choose to call the "cloud function".

We can estimate the value of the cloud function by the difference between the clear-sky solar radiation, $R_{\mathrm{in},s,\mathrm{CS}}$, and the solar radiation received at the land surface, $R_{\mathrm{in},s}$. To compute the clear-sky solar radiation, we first compute the top-of-atmosphere (i.e., extraterrestrial) solar radiation ($R_{\mathrm{in},s,\mathrm{TOA}}$): see `sunpos.py` from Wickert (2020). We then modify it based on elevation (Zotarelli & Dukes, 2010), which determines the atmospheric thickness above a particular location:

$$R_{\mathrm{in},s,\mathrm{CS}} = \left(0.75 + 2 \cdot 10^{-5} z\right) R_{\mathrm{in},s,\mathrm{TOA}}, \tag{E4}$$

where z, as in the main text, is surface elevation in meters.

This method works only where sufficient incoming solar radiation exists to produce a meaningful difference between $R_{\mathrm{in},s,\mathrm{TOA}}$ and $R_{\mathrm{in},s}$. Based on our tests, a reasonable

cutoff incoming value of solar radiation is $15\,\mathrm{W\,m^{-2}}$.

$$\mathcal{C} = \begin{cases} 1.35\frac{R_{\mathrm{in},s}}{R_{\mathrm{in},s,\mathrm{TOA}}} - 0.35 & \text{if } R_{\mathrm{in},s,\mathrm{TOA}} \geq 15 \\ \frac{\left[1.35\frac{R_{\mathrm{in},s}}{R_{\mathrm{in},s,\mathrm{TOA}}} - 0.35\right]}{\left[\right]_{15\text{–}20}}, & \text{otherwise} \end{cases} \tag{E5}$$

where the lower term equals the average of the upper term where $15 < R_{\mathrm{in},s,\mathrm{TOA}} < 20$. This is an obvious kludge for the sake of generating a proof-of-concept model outputs, and generates a reasonable but inaccurate cloud-function value for the polar regions.

The final step is straightforward. Net radiation flux is simply the sum of the net shortwave and longwave fluxes:

$$R_n = R_{n,s} + R_{n,l} \tag{E6}$$

E4 Sensible heat flux

The Penman equation derivation for sensible heat flux, $Q_{H,s}$, results in the following equation (Dingman, 1994):

$$Q_{H,s} = \frac{K_H u}{\Delta_{P,T}} \left[\frac{E}{K_E u} - (e_{\mathrm{sat}} - e_a) \right]. \tag{E7}$$

Here, u is wind speed. $\Delta_{P,T}$ is the slope of the water liquid-to-vapor phase transition at the air temperature, T_a, which is nominally taken at two meters above the surface. Likewise, e_{sat} is the saturation water vapor pressure at the air temperature, whereas e_a is the actual water vapor pressure. K_H and K_E are coefficients of turbulent conductance $[\mathrm{kg\,m\,s^{-1}\,K^{-1}}]$ for sensible heat and water vapor (i.e., latent heat), respectively.

These conductance coefficients are defined based on ratios of heat (K_H) and water vapor (K_E) transfer to momentum transfer. (Dingman, 1994):

$$\begin{aligned} K_H =& \frac{D_H}{D_M} c_p \rho_a \left(\frac{u_*}{u}\right)^2, \\ K_E =& \frac{D_{\mathrm{WV}}}{D_M} \frac{\Delta\rho_a}{P\rho_w} \frac{R_a}{R_v} \left(\frac{u_*}{u}\right)^2. \end{aligned} \tag{E8}$$

Here, D_H is thermal diffusivity in air, D_M is diffusivity of momentum, and D_{WV} is diffusivity of water vapor. For a stable atmosphere, which we assume, the same turbulent eddies result in the transfer of heat, momentum, and water vapor. Therefore, $D_H/D_M = D_{WV}/D_M = 1$. This simplifies E8 to:

$$\begin{aligned} K_H =& c_p \rho_a \left(\frac{u_*}{u}\right)^2, \\ K_E =& \frac{\rho_a}{P\rho_w} \frac{R_a}{R_v} \left(\frac{u_*}{u}\right)^2. \end{aligned} \tag{E9}$$

To restate the variable definitions from the main text for convenience: c_p is the specific heat capacity of air at constant pressure, ρ_a is air density; u_* is wind shear velocity, u is measured wind velocity (typically at 2 meters elevation above the surface), ρ_w is water density, P is atmospheric pressure, and $R_a/R_v = 0.622$ is the ratio of the gas constants of air and water vapor.

E5 Full Penman Equation

Combining Equations E1 and E7 and solving for evaporation results in the common full form of the Penman Equation (cf. Dingman, 1994):

$$E = \left[R_n + \left(\frac{K_H u}{\Delta_{P,T}} \right) (e_{\mathrm{sat}} - e_a) \right] \Big/ \left[\rho_w \Delta H_{\mathrm{vap}} + \left(\frac{K_H}{K_E} \frac{1}{\Delta_{P,T}} \right) \right]. \tag{E10}$$

Substituting in the definitions of coefficients K_H and K_E, we obtain Equation D1.

E6 Variable water-surface roughness

The u_* term in the diffusivity of momentum, D_M, may be evaluated by solving for the boundary-layer velocity profile given by the logarithmic Law of the Wall, in which

$$u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right). \tag{E11}$$

Here, $\kappa = 0.407$ is von Kármán's constant and z_0 is a surface roughness length. It is then possible to solve for u_* by knowing the wind velocity – u at a known elevation, z_1, which is typically 2 m above the surface – and the surface roughness length scale.

When wind flows over open water, it generates waves, thereby making this roughness length itself a function of wind speed. This makes Eq. E11 nonlinear, thereby adding a complexity not included in models of evaporation over land.

To address this problem, we turn to Charnock (1955), who found a quadratic relationship between z_0 and u_*. Hersbach (2011) define z_0 by showing that it depends on kinematic viscosity, ν, in light winds and on a Charnock (1955) relationship for strong winds:

$$z_0 = K_\nu \frac{\nu}{u_*} + K_{\text{wave}} \frac{u_*^2}{g}, \tag{E12}$$

where the coefficients $K_\nu = 0.11$ and $K_{\text{wave}} \approx 0.018$. We then substitute this expression for z_0 into Eq. E11 and solve for u_* using the known u at elevation z_1:

$$u_* = \kappa u \Big/ \ln\left(\frac{z_1}{K_\nu \nu / u_* + K_{\text{wave}} u_*^2 / g}\right). \tag{E13}$$

With our single known wind speed at z_1=2 meters elevation (Abatzoglou et al., 2018), we can solve this equation for u_* in one of two ways. First, we can use a numerical root finder. We implement this using the `root_scalar` method within Scipy (?, ?; Virtanen et al., 2020) (see `https://github.com/umn-earth-surface/TerraClimate-potential-open-water-evaporation`). The second option is to derive an analytical solution. This is possible for the original Charnock (1955) relationship using a Lambert W function, but is not possible for the form given by Hersbach (2011). Roots to Equation E13 exist for wind velocities less than approximately 55 m s^{-1}.

Acknowledgments

RB was supported by the Department of Energy's Computational Science Graduate Fellowship (Grant No. DE-FG02-97ER25308) and, through the Berkeley Institute for Data Science's PhD Fellowship, by the Gordon and Betty Moore Foundation (Grant GBMF3834) and by the Alfred P. Sloan Foundation (Grant 2013-10-27).

KLC and ADW were supported by the National Science Foundation under grant no. EAR-1903606.

Ying Fan Reinfelder and Gonzalo Miguez-Macho graciously provided the source code used in Fan et al. (2013) and spent time explaining the concept, code, and required inputs.

This collaboration resulted from a serendipitous meeting at the Community Surface Dynamics Modeling System (CSDMS) annual meeting, which RB had attended on a CSDMS travel grant.

Complete, well-commented source code, an associated makefile, and correctness tests are available from `https://doi.org/10.5281/zenodo.3554537` (v1.1.0).

The authors declare that they have no conflict of interest.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, *5*, 1–12. doi: 10.1038/sdata.2017.191

Ameli, A. A., McDonnell, J. J., & Bishop, K. (2016). The exponential decline in saturated hydraulic conductivity with depth: a novel method for exploring its effect on water flow paths and transit time distribution. *Hydrological Processes*, *30*(14), 2438–2450. doi: 10.1002/hyp.10777

Barnes, R. (2016). Parallel Priority-Flood Depression Filling For Trillion Cell Digital Elevation Models On Desktops Or Clusters. *Computers and Geosciences*, *96*, 56–68. Retrieved from `http://dx.doi.org/10.1016/j.cageo.2016.07.001` doi: 10.1016/j.cageo.2016.07.001

Barnes, R., Callaghan, K., & Wickert, A. (2021). Computing water flow through complex landscapes, Part 3: Fill-Spill-Merge: Flow routing in depression hierarchies. *Earth Surface Dynamics*, *9*, 105–121. doi: https://doi.org/10.5194/esurf-9-105-2021

Barnes, R., & Callaghan, K. L. (2019). Depression Hierarchy Source Code. *Zenodo*. doi: https://doi.org/10.5281/zenodo.3238558

Barnes, R., & Callaghan, K. L. (2020). Fill-Spill-Merge Source Code. *Zenodo*. doi: https://doi.org/10.5281/zenodo.3755142

Barnes, R., Callaghan, K. L., & Wickert, A. A. D. (2020). Computing water flow through complex landscapes, Part 2: Finding hierarchies in depressions and morphological segmentations. *Earth Surface Dynamics*, *8*(2), 431–445. doi: https://doi.org/10.5194/esurf-2019-34

Barnes, R., Lehman, C., & Mulla, D. (2014a). An efficient assignment of drainage direction over flat surfaces in raster digital elevation models. *Computers \& Geosciences*, *62*, 128–135. Retrieved from `http://dx.doi.org/10.1016/j.cageo.2013.01.009` doi: 10.1016/j.cageo.2013.01.009

Barnes, R., Lehman, C., & Mulla, D. (2014b). Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models. *Computers \& Geosciences*, *62*, 117–127. Retrieved from `http://dx.doi.org/10.1016/j.cageo.2013.04.024` doi: 10.1016/j.cageo.2013.04.024

Callaghan, K. L., Barnes, R., & Wickert, A. D. (2020). *Water Table Model (WTM): Source Code.* doi: 10.5281/zenodo.4265369

Callaghan, K. L., & Wickert, A. D. (2019). Computing water flow through complex landscapes - Part 1: Incorporating depressions in flow routing using FlowFill. *Earth Surface Dynamics*, *7*(3), 737–753. doi: https://doi.org/10.5194/esurf-7-737-2019

Cardenas, M. B., & Jiang, X. W. (2010). Groundwater flow, transport, and residence times through topography-driven basins with exponentially decreasing permeability and porosity. *Water Resources Research*, *46*(11), 1–9. doi: 10.1029/2010WR009370

Charney, J. G., FjÖrtoft, R., & Neumann, J. V. (1950). Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, *2*(4), 237–254. doi: 10.3402/tellusa.v2i4.8607

Charnock, H. (1955). Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, *81*(350), 639–640. doi: 10.1002/qj.49708135026

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water resources research*, *14*(4).

Courant, R., Friedrichs, K., & Lewy, H. (1956). *On the partial difference equations of mathematical physics* (Tech. Rep.). New York: Courant institute of mathematical sciences.

Crank, J., & Nicolson, P. (1996). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Advances*

in Computational Mathematics, *6*(1), 207–226. doi: 10.1007/bf02127704

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J. P., Alias, A., ... Voldoire, A. (2019). Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1207–1252. doi: 10.1029/2018MS001545

Dingman, L. (1994). *Physical hydrology.* New York: Macmillan Pub. Co.

Döll, P., Fiedler, K., & Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, *13*(12), 2413–2432. doi: 10.5194/hess-13-2413-2009

Dunne, T., & Black, R. D. (1970). An experimental investigation runoff production in permeable soils. *Water Resources Research*, *6*(2), 478–490.

Earle, S. (2015). *Physical Geology.* Victoria, B. C.. Retrieved from `https://opentextbc.ca/geology/`

European Centre for Medium-Range Weather Forecasts. (2019). *ERA5 Reanalysis (Monthly Mean 0.25 Degree Latitude-Longitude Grid).* doi: https://doi.org/10.5065/P8GT-0R61

Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. doi: 10.1126/science.1229881

Fan, Y., & Miguez-Macho, G. (2011). A simple hydrologic framework for simulating wetlands in climate and earth system models. *Climate Dynamics*, *37*(1), 253–278. doi: 10.1007/s00382-010-0829-8

Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., & Robock, A. (2007). Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations. *Journal of Geophysical Research Atmospheres*, *112*(10), 1–17. doi: 10.1029/2006JD008111

Finch, J., & Calver, A. (2008). Methods for the quantification of evaporation from lakes. *Report*(October), 47.

Freeze, R. A., & Cherry, J. A. (1979). *Groundwater.* Englewood Cliffs, New Jersey: Prentice-Hall.

GEBCO Bathymetric Compilation Group. (2020). *GEBCO_2020 Grid.* Retrieved from `https://www.gebco.net/data_and_products/gridded_bathymetry_data/` doi: doi:10.5285/a29c5465-b138-234d-e053-6c86abc040b9

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., & Cardenas, M. B. (2016). The global volume and distribution of modern groundwater. *Nature Geoscience*, *9*(2), 161–164. doi: 10.1038/ngeo2590

Hersbach, H. (2011). Sea surface roughness and drag coefficient as functions of neutral wind speed. *Journal of Physical Oceanography*, *41*(1), 247–251. doi: 10.1175/2010JPO4567.1

Horton, R. E., & Htrata, T. (1955). Erosional development of streams and their drainage basins, hydrophysical approach to quantitative morphology. *Nihon Ringakkai Shi/Journal of the Japanese Forestry Society*, *37*(9), 417–420. doi: 10.11519/jjfs1953.37.9{_}417

Koirala, S., Yeh, P. J., Hirabayashi, Y., Kanae, S., & Oki, T. (2014). Global-scale land surface hydrologic modeling with the representation of water table dynamics. *Journal of Geophysical Research*, *119*(1), 75–89. doi: 10.1002/2013JD020398

Kourzeneva, E., Asensio, H., Martin, E., & Faroux, S. (2012). Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *64*(1). doi: 10.3402/tellusa.v64i0.15640

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., ... Zeng, X. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. doi:

10.1029/2018MS001583

Maxwell, R. M., Condon, L. E., & Kollet, S. J. (2015). A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development*, *8*(3), 923–937. doi: 10.5194/gmd-8-923-2015

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*, 1–11. doi: 10.1038/ncomms13603

Monteith, J. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology*(19), 205–234.

Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012, 5). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling and Software*, *31*, 124–130. doi: 10.1016/j.envsoft.2011.11.014

Ni, S., Chen, J., Wilson, C. R., Li, J., Hu, X., & Fu, R. (2018). Global Terrestrial Water Storage Changes and Connections to ENSO Events. *Surveys in Geophysics*, *39*(1), 1–22. doi: 10.1007/s10712-017-9421-7

NOAA. (2016). National Water Model: Improving NOAA's Water Prediction Services. (August), 2. Retrieved from `http://water.noaa.gov/documents/wrn-national-water-model.pdf`

Oleson, K., Lawrence, D., Bonan, G., Flanner, M., Kluzek, E., Lawrence, P., ... Thornton, P. (2010). Technical Description of version 4.0 of the Community Land Model (CLM). *NCAR Technical Note NCAR/TN-478+STR*(April), 257.

Peirce, J. J., Weiner, R. F., & Vesilind, P. A. (1998). *Environmental Pollution and Control* (Fourth ed ed.). Butterworth-Heinemann.

Penman, H. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *193*(1032), 120–145.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., & Wilson, C. R. (2008). Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resources Research*, *44*(2). doi: 10.1029/2006WR005779

Tarboton, D. (2017). Great Salt Lake Bathymetry. *HydroShare*. Retrieved from `http://www.hydroshare.org/resource/582060f00f6b443bb26e896426d9f62a`

Valiantzas, J. D. (2006). Simplified versions for the Penman evaporation equation using routine weather data. *Journal of Hydrology*, *331*(3-4), 690–702. doi: 10.1016/j.jhydrol.2006.06.012

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, *41*(18), 6396–6402. doi: 10.1002/2014GL060641

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*.

Vörösmarty, C. J., Federer, C. A., & Schloss, A. L. (1998). Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling. *Journal of Hydrology*, *207*(3-4), 147–169. doi: 10.1016/S0022-1694(98)00109-7

Wada, Y. (2016). Modeling Groundwater Depletion at Regional and Global Scales: Present State and Future Prospects. *Surveys in Geophysics*, *37*(2), 419–451. doi: 10.1007/s10712-015-9347-x

Wickert, A. D. (2020). Potential open water evaporation from TerraClimate. *Zenodo*. doi: https://doi.org/10.5281/zenodo.4391500

Wiltshire, A. J., Carolina Duran Rojas, M., Edwards, J. M., Gedney, N., Harper, A. B., Hartley, A. J., ... Smout-Day, K. (2020). JULES-GL7: The Global Land configuration of the Joint UK Land Environment Simulator version 7.0 and 7.2. *Geoscientific Model Development*, *13*(2), 483–505. doi:

10.5194/gmd-13-483-2020

Yokohata, T., Kinoshita, T., Sakurai, G., Pokhrel, Y., Ito, A., Okada, M., ... Emori, S. (2020). MIROC-INTEG-LAND version 1: A global biogeochemical land surface model with human water management, crop growth, and land-use change. *Geoscientific Model Development*, *13*(10), 4713–4747. doi: 10.5194/gmd-13-4713-2020

Zeng, X., Shajkh, M., Dai, Y., Dickinson, R. E., & Myneni, R. (2002). Coupling of the Common Land Model to the NCAR Community Climate Model. *Journal of Climate*, *15*(14), 1832–1854. doi: 10 .1175 / 1520 -0442(2002) 015⟨1832 : COTCLM⟩2.0.CO;2

Zotarelli, L., & Dukes, M. (2010). Step by step calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method). *Institute of Food and ...*, 1–10. Retrieved from `https://edis.ifas.ufl.edu/pdffiles/AE/AE45900.pdf`