WBM Documentation

Authors (alphabetical):
Stanley Glidden
Danielle Grogan
Richard Lammers
Alex Prusevich
David Simons
Wilfred Wollheim
Shan Zuidema

Last Updated: 2022-02-11

Table of Contents

1. WBM overview

Overall goal of WBM (Mission Statement, Research Priorities)

To simulate all the world's water.

We achieve this by developing a tool to help us explore and understand drainage basin-scale hydrological and material transport processes both historical and in the future.

WBM Overview and key publications

The University of New Hampshire Water Balance Model (WBM) is a process-based, modular, gridded hydrologic model that simulates spatially and temporally varying water volume and material transport across a wide range of spatial domains. WBM represents all major land surface components of the hydrological cycle, and tracks fluxes and balances between the atmosphere, aboveground water storages (e.g. snowpack, glaciers), soil, vegetation, groundwater, and runoff (**Figure 1**). A digitized river network connects grid cells, enabling simulation of flow through the river and groundwater systems. Direct human influences include domestic, industrial, and agricultural (irrigation and livestock) water demand and use, the impacts of impervious surfaces, and hydro-infrastructure (dams, reservoirs, canals, inter-basin transfers). The model is also the hydrological core of the Framework for Aquatic Modeling of the Earth System (FrAMES), which predicts water temperature, nutrient fluxes (Stewart et al. 2011, 2013; Samal et al. 2017, Wollheim et al. 2008), and chloride fluxes (Zuidema et al, 2018). The model has an embedded water routing scheme, including constituent transport.

WBM is modular and can operate at a wide range of spatial scales from local watersheds at 120 m grid cells (e.g. Stewart et al. 2011) to global freshwater systems at ½ degree grid cells (e.g. Grogan et al. 2017; Wisser et al. 2010). WBM accepts hydrologic, land use/land cover, water management, and water demand inputs from other models and data sources, such as glacier melt models (Huss and Hock 2015; Rounce et al. 2019) and econometric models (Zaveri et al. 2016) and has provided boundary conditions for the SIMPLE economic model (Liu et al. 2017).

WBM accounts for the operation of dams and reservoirs (Wisser et al. 2010), inter-basin hydrological transfers (Zaveri et al. 2016), and agricultural water use from irrigation (Grogan et al. 2015, 2017; Grogan 2016; Wisser et al. 2010, Zaveri et al. 2016). Additionally, WBM modules have been developed recently, and include the use of sub-grid elevation band distributions derived from a high-resolution elevation dataset to improve handling of snowpack in mountainous regions.

The model has been applied to address a variety of hydrologic questions over many different regions across the globe including:

Global   Grogan et al. 2017; Grogan 2016; Wisser et al. 2008, 2009, 2010; Fekete et al. 2006; Vörösmarty et al. 2000, 2010.

Arctic	Bring et al. 2017; Shiklomanov et al. 2013; Rawlins et al. 2003, 2005, 2006a,b, 2009.
Asia	Zaveri et al. 2016; Grogan et al. 2015; Douglas et al. 2006; Groisman et al. 2018.
Africa	Vörösmarty et al. 2005.
South America	Vörösmarty et al. 1989.
North America	Zuidema et al. 2018; Samal et al. 2017; Stewart et al. 2011, 2013; Vörösmarty et al. 1998.
Tropics	Douglas et al. 2005, 2007.
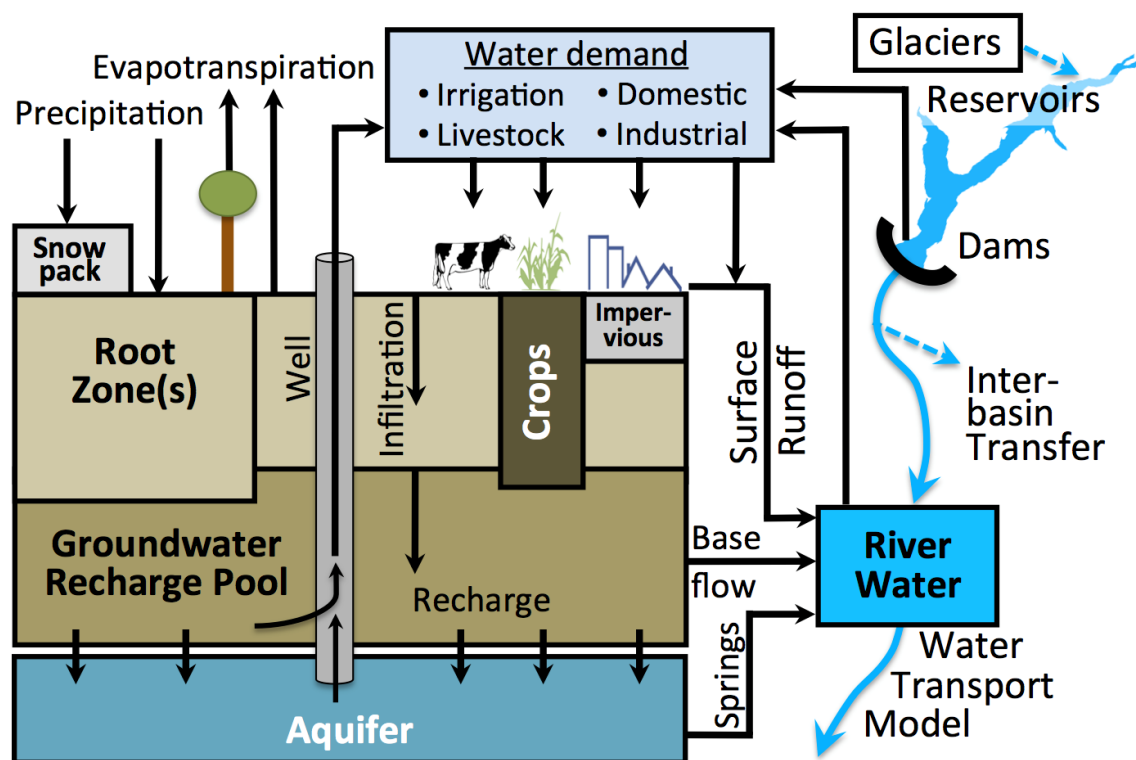
Figure 1-1: Major elements of the Water Balance Model

References

Bring, A, A. Shiklomanov, R.B. Lammers (2017) Pan-Arctic river discharge: prioritizing monitoring of future climate change hotspots, *Earth's Future*, doi:10.1002/2016EF000434.

Douglas, E.M., Sebastian, K., Vörösmarty, C.J., Wood, S., & Chomitz, K.M. (2005) The Role of Tropical Forests in Supporting Biodiversity and Hydrological Integrity. *SSRN Electronic Journal*, http://doi.org/10.2139/ssrn.757186.

Douglas, E.M., Niyogi, D., Frolking, S., Yeluripati, J.B., Pielke Sr., R.A., Niyogi, N., Vörösmarty, C.J., and Mohanty, U.C. (2006), Changes in moisture and energy fluxes due to agricultural land use and irrigation in the Indian Monsoon Belt, *Geophys. Res. Lett.*, 33, L14403, doi:10.1029/2006GL026550.

Douglas, E.M., Wood, S., Sebastian, K., Vörösmarty, C.J., Chomnitz, K.M., Tomich, T.P. (2007) Policy implications of a pan-tropic assessment of the simultaneous hydrological and biodiversity impacts of deforestation, *Water Resour Manage*, 21: 211. https://doi.org/10.1007/s11269-006-9050-2.

Fekete, B. M., J. J. Gibson, P. Aggarwal, and C. J. Vörösmarty (2006) Application of isotope tracers in continental scale hydrological modeling, *Journal of Hydrology* 330, 444-456.

Grogan, D.S., D. Wisser, A. Prusevich, R.B. Lammers, S. Frolking (2017) The use and re-use of unsustainable groundwater for irrigation: A global budget, *Environmental Research Letters*, 12(3), 034017, doi: 10.1088/1748-9326/aa5fb2.

Grogan, D.S., F. Zhang, A. Prusevich, R.B. Lammers, D. Wisser, S. Glidden, C. Li, S. Frolking (2015) Quantifying the link between crop production and mined groundwater irrigation in China, *Science of the Total Environment*, 511:161-175; doi:10.1016/j.scitotenv.2014.11.076.

Grogan, D.S. (2016) Global and regional assessments of unsustainable groundwater use in irrigated agriculture (doctoral dissertation). Available at Doctoral Dissertations. 2. http://scholars.unh.edu/dissertation/2.

Groisman, P.; Bulygina, O.; Henebry, G.; Speranskaya, N.; Shiklomanov, A. et al. (2018) Dry land belt of Northern Eurasia: Contemporary environmental changes and their consequences. *Environ. Res. Lett.*, 13 115008.

Huss, M. and R. Hock (2015) A new model for global glacier change and sea-level rise. *Frontiers in Earth Sciences* 3:54, doi: 10.3389/feart.2015.00054.

Liu, J., T. Hertel, R. Lammers, A. Prusevich, U. Baldos, D. Grogan, S. Frolking (2017) Achieving Sustainable Irrigation Water Withdrawals: Global Impacts on Food Production and Land Use, *Environmental Research Letters*, 12(10):104009, http://stacks.iop.org/1748-9326/12/i=10/a=104009.

Rawlins, M.A., R.B. Lammers, S. Frolking, B.M. Fekete, C.J. Vörösmarty (2003) Simulating Pan-Arctic Runoff with a Macro-Scale Terrestrial Water Balance Model, *Hydrological Processes*, 17(13):2521-2539

Rawlins, M.A., K.C. McDonald, S. Frolking, R.B. Lammers, M. Fahnestock, J.S. Kimball, C.J. Vörösmarty (2005) Remote sensing of snow thaw at the pan-Arctic scale using the SeaWinds scatterometer, *Journal of Hydrology*, 312(1-4):294-311.

Rawlins, M.A., C.J. Willmott, A. Shiklomanov, E. Linder, S. Frolking, R.B. Lammers, C.J. Vörösmarty (2006a) Evaluation of Trends in Derived snowfall and rainfall across Eurasia

and linkages with Discharge to the Arctic Ocean, *Geophysical Research Letters*, Vol. 33, L07403, doi:10.1029/2005GL025231.

Rawlins, M.A., Frolking, S., Lammers, R.B., Vörösmarty, C.J. (2006b) Effects of Uncertainty in Climate Inputs on Simulated Evapotranspiration and Runoff in the Western Arctic, *Earth Interactions*, 10:1-18, doi: 10.1175/EI182.1.

Rawlins, M.A., M. Steele, M.C. Serreze, C.J. Vörösmarty, W. Ermold, R.B. Lammers, K.C. McDonald, T.M. Pavelsky, A. Shiklomanov and J. Zhang (2009) Tracing freshwater anomalies through the air-land-ocean system: A case study from the Mackenzie River basin and the Beaufort Gyre, *Atmosphere-Ocean*, 47(1), 79–97, doi:10.3137/OC301.2009.

Rounce et al. 2019 HiMAT paper

Samal, N.R., Wollheim, W.M., Zuidema, S., Stewart, R.J., Zhou, Z., Mineau, M.M., Borsuk, M.E., Gardner, K.H., Glidden, S., Huang, T., Lutz, D.A., Mavrommati, G., Thorn, A.M., Wake, C.P., & Huber, M. (2017) A coupled terrestrial and aquatic biogeophysical model of the Upper Merrimack River watershed, New Hampshire, to inform ecosystem services evaluation and management under climate and land-cover change, *Ecology and Society*, 22(4), 18. https://doi.org/10.5751/ES-09662-220418.

Shiklomanov A.I., R.B. Lammers, D.P. Lettenmaier, Y.M. Polischuk, O. Savichev, L.C. Smith, A.V. Chernokulsky (2012) Hydrological changes: historical analysis, contemporary status and future projections, pp. 111-154, in Groisman, P.Ya. and G. Gutman (eds.) *Regional Environmental Changes in Siberia and Their Global Consequences*, Springer Environmental Science and Engineering, Springer, Dordrecht, 357 pp.

Stewart, R.J., W.M. Wollheim, M.N. Gooseff, M.A. Briggs, J.M. Jacobs, B.J. Peterson, and C.S. Hopkinson (2011) Separation of river network–scale nitrogen removal among the main channel and two transient storage compartments, *Water Resour. Res.*, 47, W00J10, doi:10.1029/2010WR009896.

Stewart, R.J., W.M. Wollheim, A. Miara, C.J. Vörösmarty, B. Fekete, R.B. Lammers, B. Rosenzweig (2013) Horizontal Cooling Towers: Riverine Ecosystem Services and the Fate of Thermoelectric Heat in the Contemporary Northeast U.S., *Environmental Research Letters*, 8:025010, doi:10.1088/1748-9326/8/2/025010.

Vörösmarty, C.J., B. Moore, M.P. Gildea, B. Peterson, J. Melillo, D. Kicklighter, J. Raich, E. Rastetter, and P. Steudler (1989) A continental-scale model of water balance and fluvial transport: Application to South America. *Global Biogeochemical Cycles* 3:241-65.

Vörösmarty, C.J., C.A. Federer and A. Schloss (1998) Potential evaporation functions compared on U.S. watersheds: Implications for global-scale water balance and terrestrial ecosystem modeling. *Journal of Hydrology* 207: 147-69.

Vörösmarty, C.J. P. Green, J. Salisbury, R.B. Lammers (2000) Global Water Resources: Vulnerability from Climate Change and Population Growth, *Science*, 289:284-288, July 14, 2000.

Vörösmarty, C.J., E.M. Douglas, P.A. Green, C. Revenga (2005) Geospatial indicators of emerging water stress: An application to Africa. *Ambio*. 34: 230-236.

Vorosmarty, C.J., P.B. McIntyre, Prusevich A.A., et al. (2010). Global threats to human water security and river biodiversity, *Nature* 467(7315): 555-561.

Wisser D, S Frolking, EM Douglas, BM Fekete, CJ Vörösmarty, AH Schumann (2008) Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, *Geophysical Research Letters*, 35, L24408, doi:10.1029/2008GL035296.

Wisser D, S Frolking, EM Douglas, BM Fekete, AH Schumann, CJ Vörösmarty (2009) The significance of local water resources captured in small reservoirs for crop production – A global-scale analysis. *Journal of Hydrology*, doi:10.1016/j.jhydrol.2009.07.032.

Wisser D, BM Fekete, CJ Vörösmarty, AH Schumann (2010) Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), *Hydrology and Earth System Science*, 14, 1-24.

Wollheim, W.M., C.J. Vorosmarty, A.F. Bouwman, P.A. Green, J. Harrison, E. Linder, B.J. Peterson, S. Seitzinger, and J.P.M. Syvitski (2008). Global N removal by freshwater aquatic systems: a spatially distributed, within-basin approach. *Global Biogeochemical Cycles*. GB2026, doi:10.1029/2007GB002963.

Zaveri, E., D.S. Grogan, K. Fisher-Vanden, S. Frolking, R.B. Lammers, D.H. Wrenn, A. Prusevich, R.E. Nicholas (2016) Invisible water, visible impact: Groundwater use and Indian agriculture under climate change, *Environmental Research Letters*, 11, 084005.

Zuidema, S. W.M. Wollheim, M.M. Mineau, M.B. Green, R.J. Stewart (2018) Chloride impairment in a New England river network: regional assessment using a dynamic watershed transport model. *Journal of Environmental Quality*. doi:10.2134/jeq2017.11.0418.

Core water balance functions

2. Potential evapotranspiration

2.1 Hamon PET

Potential evapotranspiration, PET, is the maximum amount of water that can be lost from soil through combined evaporation and transpiration, assuming no shortage of soil water. It provides an upper bound on non-irrigated actual evapotranspiration and is used as a baseline reference for calculating irrigated evapotranspiration.

WBM can use the Hamon method (Hamon,1963) to calculate PET [mm]. This is the least data-intensive method, and it was found to estimate global average PET as well as other, more data-intensive methods. Additionally, Vorosmarty (1998) found that amongst the reference-surface PET methods, the Hamon method produced both the lowest mean annual error and the smallest bias when compared to observation data.

$$PET = 330.2 \, \Lambda \, \rho_{sat} \tag{2.1-1}$$

where

Λ = day length, expressed as a fraction of a 12-hour period

$$\rho_{sat} = 2.167 \frac{P_{sat}}{T+273.15} \text{ [g m}^{-3}\text{]} \tag{2.1-2}$$

T = daily mean temperature [°C]

$$P_{sat} = \begin{cases} 0.61078 \, e^{\frac{T}{t+237.3}} & if \ 0 \le T \\ 0.61078 \, e^{\frac{T}{t+265.5}} & if \ 0 > T \end{cases} \text{ [kg m}^{-1}\text{s}^{-2}\text{]} \tag{2.1-3}$$

References:

Hamon, W. R. (1963) Computation of direct runoff amounts from storm rain- fall, International Association of Hydrological Sciences Publications, 63, 52-62.

2.2 Penman-Montieth PET

WBM can calculate potential evapotranspiration (ET_p [mm d^{-1}]) using derivatives of the combination equations pioneered by Penman (1948) and Monteith (1965) as described in Dingman (2002). Penman-Monteith potential evapotranspiration (ET_p) is given by equation 2.2-1 below, and is calculated for soil area of each pixel at a daily time-step.

$$\text{ET}_p = \frac{\Delta \cdot (K - G - L_O) + \rho_a \cdot c_a \cdot C_{at} \cdot e_a^*(1 - h_a)}{\rho_w \cdot \lambda_v[\Delta + \gamma \cdot (1 + C_{at}/C_{can})]} \qquad \text{2.2-1}$$

Variables in the above equation are defined along with methods of derivation in Table 1.

2.2.1 FAO Drainge Paper No. 56 Evapotranspiration

An alternative implementation of potential evapotransipiration that utilizes the Penman-Monteith fomulation of Allen et al. (1998) is also implemented in WBM. The model solves potential evapotransipiration using the form presented by Zotarelli et al. (2018) in equation 2.2-2 below.

$$\text{ET}_p = DT \cdot (K - G - L_O) + PT \cdot TT \, (e_a^* - e_a) \qquad \text{2.2-2}$$

Variables in the above equation along with methods of derivation are provided in Table 1.

Table 1: Definitions of terms used in evapotranspiration calculation. Compiled by Dingman [2002] unless stated otherwise.

Term	Units	Description	Formulation
Δ	kPa K^{-1}	slope of the ratio between saturation vapor pressure and air temperature (in K)	$\dfrac{2508.3}{[T_a + 237.3]^2} \exp\left(\dfrac{17.3 \cdot T_a}{T_a + 237.3}\right)$
T_a	°C	Mean air temperature in degrees centigrade	Input
K	MJ m^{-2} h^{-1}	Net incoming solar radiation (From ?)	$K_{CS}\left(0.803 - 0.34\,k_{cld} - 0.458\,k_{cld}{}^2\right)(1-a)$
K_{cs}	MJ m^{-2} h^{-1}	Clear sky radiation	Estimated from extraterrestrial solar radiation
k_{cld}	-	Cloud/shielding factor	$(0.9 * f_{cloud})^3$
α	-	Albedo	Input
G	MJ m^{-2} h^{-1}	Ground heat flux	Input (if available otherwise 0)
L_0	$MJ\,m^{-2}\,h^{-1}$	Net out-going long-wave radiation (From Allen et al. 1998)	$4.903e^{-9} \times (T_a + 273.15)^4 \vert$ $\times \left(0.34 - 0.14\sqrt{e_a^* h_a}\right)$ $\times \left(1.35 \times \dfrac{K + 0.1}{K_{CS} + 0.1} - 0.35\right)$
e_a^*	kPa	Saturation water vapor pressure	$0.6108 \cdot exp\left(\dfrac{17.27 \cdot T_a}{T_a + 237.3}\right)$ at $T_a \geq 0$ $0.6108 \cdot exp\left(\dfrac{21.87 \cdot T_a}{T_a + 265.5}\right)$ at $T_a < 0$
e_a	kPa	Water vapor pressure	$\dfrac{h_s\, P_a}{0.378\,h_s + 0.622}$
h_a	-	Relative air humidity (fraction) Note: In order fixing input data errors, the minimum allowed h_a is set to 0.1	Relative humidity / 100 or ${e_a}/{e_a^*}$
h_s	kg kg^{-1}	Specific air humidity	Input

e_a	kPa	Actual water vapor pressure	$e_a = h_a \cdot e_a^*$

Table 1 (Continued): Definitions of terms used in evapotranspiration calculation. Compiled by Dingman [2002] unless stated otherwise.

Term	Units	Description	Formulation
γ	kPa K^{-1}	Psychrometric constant	$\dfrac{c_a \cdot P_a}{0.622 \cdot \lambda_v}$
ρ_a	kg m^{-3}	Density of air	$\dfrac{P_a}{T_a R_a}$
c_a	MJ kg^{-1} K^{-1}	Heat capacity of air	$1.00 \times 10^{-3} MJ\ kg^{-1}\ K^{-1}$
P_a	kPa	barometric air pressure	Input or $\dfrac{100}{0.288 \cdot (T_a + 273.15)}$
λ_v	MJ kg^{-1}	latent heat of vaporation of water	$2.50 - 2.36 \times 10^{-3} \cdot T_a$
C_{at}	m h^{-1}	atmospheric conductance	$\dfrac{v_a}{6.25\left[ln\left(\dfrac{\frac{z_h}{z_0}}{z_{veg}} + \dfrac{1-z_d}{z_0}\right)\right]^2}$
ρ_w	kg m^{-3}	density of water	1000 kg m^{-3}
k	-	Von Karmon's constant	0.4
$z_{[x]}$	m	m: height of v_a measurement, d: zero-plane displacement, 0: roughness height	$z_d = 0.7 \cdot$ height of vegetation (z_{veg}), $z_0 = 0.1 \cdot z_{veg}$
v_a	M hr^{-1}	average wind speed (at z_m)	measured

Table 1 (Continued): Definitions of terms used in evapotranspiration calculation. Compiled by Dingman [2002] unless stated otherwise.

Term	Units	Description	Formulation
C_{can}	m h^{-1}	canopy conductance	$0.5 \cdot C_{leaf}$
LAI	-	leaf area index	Input
C_{leaf}	m h^{-1}	Stewart's [1988] estimate of stomatal leaf conductance	$C_{leaf}^* \cdot f_K(K_{in}) \cdot f_\rho(\Delta\rho_v) \cdot f_T(T_a)$
C_{leaf}^*	m h^{-1}	Maximum stomatal conductance	Input
f_K	-	Stewart's [1988] stomatal conductance dependance on incoming solar radiation	$\dfrac{12.78\ K_{in}}{11.57\ K_{in} + 104.4}$
f_ρ	-	Stewart's [1988] stomatal conductance dependance on vapor pressure deficit	$\max(1 - 66.6\ \Delta\rho_v, 0.2328)$
f_T	-	Stewart's [1988] stomatal conductance dependance on temperature	$\dfrac{T_a\ (40 - T_a)^{1.18}}{691}$
$\Delta\rho_v$	kg m^{-3}	Vapor pressure deficit	$\dfrac{e}{T_a R_a} - \dfrac{e^*}{T_a R_a}$
DT	-	Zotarelli delta term	$\dfrac{\Delta}{\Delta + \gamma(1 + 0.34 v_a)}$
PT	-	Zotarelli psi term	$\dfrac{\gamma}{\Delta + \gamma(1 + 0.34\ v_a)}$
TT	-	Zotarelli temperature term	$\left(\dfrac{900}{T_a + 273.15}\right) v_a$

References:

Allen RG, Pereira LS, Raes D, and Smith M (1998) Crop evapotranspiration: Guidelines for computing crop water requirements. FAO irrigation and Drainage Paper No. 56. https://www.kimberly.uidaho.edu/water/fao56/fao56.pdf

Dingman, S. L. (2002), *Physical hydrology*, 2nd ed., x, 646 p. pp., Prentice Hall, Upper Saddle River, N.J.

Monteith, J.L., 1965. Evaporation and environment: the state and movement of water in living organisms. Symposium of the Society for Experimental Biology 19, 205–224.

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

Stewart, J., 1988. Modelling surface conductance of pine forest. Agricultural and Forest Meteorology 43, 19–35. https://doi.org/10.1016/0168-1923(88)90003-2

Zotarelli, L., Dukes, M.D., Romero, C.C., Migliaccio, K.W., 2018. Step by Step Calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method) 10.

3. Actual evapotranspiration (AET)

3.1 Vegetation AET

Actual evapotranspiration (AET) from vegetated land areas is a function of the potential evapotranspiration (PET, see Section 2), soil moisture, and soil properties. If soil moisture is sufficient, then AET = PET. Otherwise, PET is modified by a soil drying function, $g(W_s)$. The amount of water that can be drawn out of the soil moisture pool depends on the current soil moisture, and the available water capacity (soil water between wilting point and field capacity).

Available water capacity, $Wcap$ [mm], indicates the portion of the soil moisture storage pool within the grid cell that is held against gravity drainage. Available water capacity is determined by taking the difference between the field capacity, $Fcap$ [-], and the wilting point, Wpt [-], each expressed as fractions of the total depth. This difference is then scaled by the total rooting depth, Rd [mm], to determine the depth in mm of water the grid cell can accommodate before gravity drainage (equation 3.1-1).

$$W_{cap} = R_d\left(F_{cap} - W_{Pt}\right) \quad \text{[mm]} \tag{3.1-1}$$

Field capacity, wilting point, and rooting depth are all input from global datasets based on soil and vegetation type. Alternatively, available water capacity $Wcap$ can be input directly into the model instead of calculated.

The drying function $g(W_s)$ estimates AET as a fraction of PET based on the present soil moisture content (W_s [mm]) relative to $Wcap$ through an empirical constant α [-] and is given by equation 3.1-2. The default value of 5.0 provides a match to the drying curve of Pierce (1958); however the coefficient α can be adjusted to calibrate the model based on regionally unique combinations of soil properties, vegetation, and climate.

$$g(W_s) = \frac{1-e^{-\frac{\alpha W_s}{W_{cap}}}}{1-e^{-\alpha}} \tag{3.1-2}$$

A plot of the drying function for three values of α is given in Figure 3.1.

AET is calculated wherever soil water capacity is defined according to equation 3.1-3.

$$AET = \begin{cases} 0 & \text{if } W_{cap} = 0 \\ g(W_s)(PET - P_t - M) & \text{if } P_t + M < PET \end{cases} \quad \text{[mm]} \tag{3.1-3}$$

where P_t is throughfall and M is snowmelt discussed in Sections 5 and 4, respectively. Equation 3.1-3 assumes any available latent energy first evaporates incident precipitation prior to being withdrawn from soils.
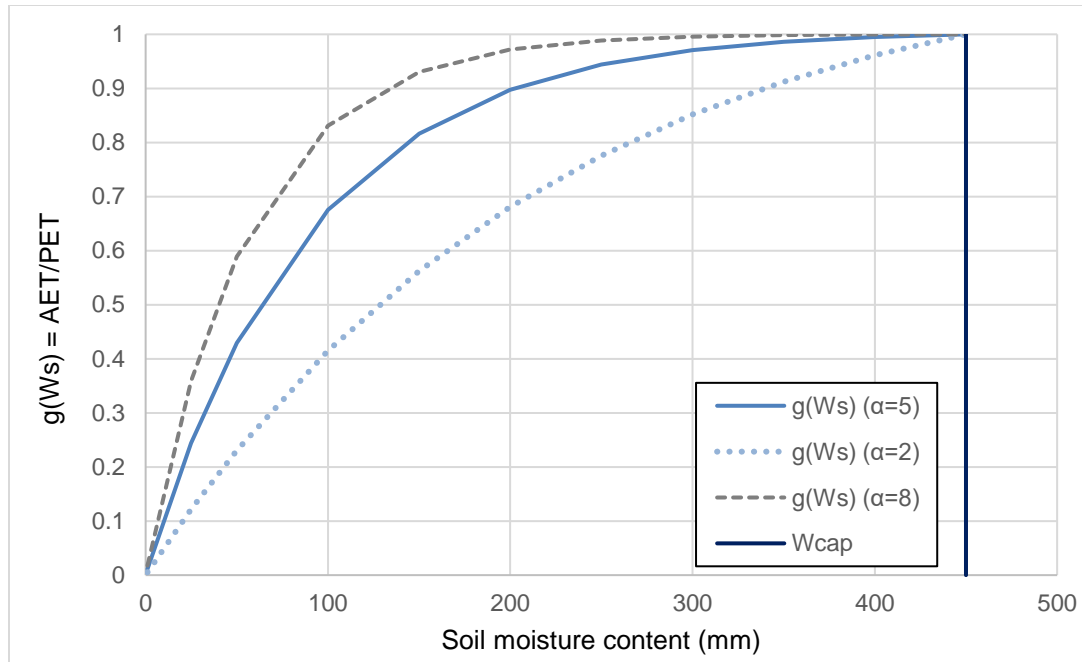
$$P_a = P + M_s - I_c \quad \text{[mm]} \tag{3.1-4}$$

Figure 3.1: Example soil moisture drying function g(Ws) relating actual evapotranspiration to potential evapotranspiration for a soil with 450 mm available water capacity and three values of the empirical soil drying parameter α.

3.2 Open water evapotranspiration

Open-water evaporation rate (E_{OW} [mm d^{-1}]) can either be input to WBM as a separate data input, which is widely available in global reanalysis meteorological data such as MERRA-2 (Gelaro et al. 2017), or can be scaled relative to calculated PET.

Open-water evaporation applies to the surface area of the river network (as estimated at the beginning of each timestep), to the area of surface impoundments, and to water stored on plant canopies (see Section 5).

The Hamon (1963) equation is described above (Section 2.1), and compares favorably to the Bowen-Ratio Energy Balance method for open water surfaces, even when measurements are potentially impacted by limited fetch (Rosenberry et al., 2007).

References:

Hamon, W.R., 1963. Computation of direct runoff amounts from storm rainfall, International Association of Hydrological Sciences, 63,52-62.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate, 30(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Pierce, L. T. (1958). Estimating Seasonal and Short-Term Fluctuations in Evapo-Transpiration from Meadow Crops. Bulletin of the American Meteorological Society, 39(2), 73–78. https://doi.org/10.1175/1520-0477-39.2.73

Rosenberry, D.O., Winter, T.C., Buso, D.C., Likens, G.E., 2007. Comparison of 15 evaporation methods applied to a small mountain lake in the northeastern USA. Journal of Hydrology 340, 149–166. https://doi.org/10.1016/j.jhydrol.2007.03.018

4. Snow
4.1   Snowpack and snow water equivalent

WBM models precipitation, P, as snowfall P_s[mm], and tracks snowpack, S_p [mm], and snowmelt, M [mm].
When mean daily temperature, T [°C], is below the snowfall threshold T_s [°C], precipitation is treated as snow. When mean daily air temperature, T [°C], is above the snowmelt threshold, T_m [°C], a portion of the snow is melted.

For regions with large orographic gradients, the elevation distribution of each model grid cell is calculated from a 30-meter DEM, resulting in binned elevation categories of ΔH vertical bands which are also called elevation or snow bands. The range and size of the snow bands can be chosen by a user, and the default range is from 0 to 5000 m elevation with band size of 250 m. A temperature lapse rate, L [°C/km], is applied to the mean daily temperature, T [°C] at the reference elevation, H_{ref} for each binned elevation category (band), resulting in an adjusted mean temperature, T_e [°C], for the portion of each grid cell in elevation band category e.

$$T_e = T + \frac{L}{1000}\left(H_e - H_{ref}\right) \tag{4-1}$$

The reference elevation for each temperature dataset is usually provided through Geopotential data layer which can be converted to the elevation by dividing it by gravity constant $g = 9.80665$ m/sec^2. Alternatively it can be calculated from the aforementioned 30-meter DEM dataset as an average elevation in the spatial extent of each pixel of the temperature dataset. Keep in mind that temperature dataset pixel sizes are specific to that dataset and depend on its resolution and projection.
Precipitation rates are assumed to be equal across all elevation bands e, such that $P_e = P$ [mm/day]. If sub-grid elevation snow processes are not used, the same snow processes apply to the entire grid cell.

Snow water equivalent (SWE) is updated through timesteps of length dt in elevation bin e as:
$$\frac{dS^e}{dt} = P_s^e - M^e \tag{4-3}$$

$$P_s^e = \begin{cases} P \ if \ T^e < T_s \\ 0 \ if \ T_s \leq T^e \end{cases} \text{[mm d}^{-1}\text{]} \tag{4-4}$$

$$M^e = \begin{cases} 2.63 + 2.55 \ T^e + 0.0912 \ T^e P \ if \ T_m < T^e \\ \qquad\qquad 0 \ if \ T^e \leq T_m \end{cases} \text{[mm d}^{-1}\text{]} \tag{4-5}$$

Total SWE in snowpack S_p, [mm/d] in the grid cell at each time-step is the sum of all SWE values at each elevation band e multiplied by the corresponding fraction of grid cell area represented by elevation bin e, f^e:

$$S_p = \sum_{e=1}^{n} S_p^e f^e \tag{4-6}$$

Variables controlling SWE accumulation include the snowfall threshold Ts, with a default value of -1 °C; the snow melt threshold Tm, with a default value of 1 °C; and L is the lapse rate, with a default value of -6.4 °C/km. Both Te and L can be constants for the whole simulation domain, or they can be a spatially variable gridded input layer.

4.2   Excess snowpack accumulations

At high elevations and cold climates it is a common case that annual snowfall exceeds annual snowmelt volume. In the natural systems the excess snowpack converts to ice and triggers glacial dynamics (growth, flow, and melt at lower elevations). WBM accounts for glacier areas in a separate module, but pixels with partial glacier areas are still processes through its snowpack/snowmelt module (see previous section). That causes the problem of infinite snow accumulation. To address this problem WBM combines the following sequence of steps:

1.   Glacier area is placed to the highest elevation bands within each pixel (grid cell).

2.   At the date of annual snowpack minimum the snowbands are shifted downward. The date of annual snowpack minimum is assumed to be August 15 in the Northern hemisphere and February 15 in the Southern hemisphere.

3.   The snowpack in excess of threshold (e.g. 5000 mm of snow water equivalent, SWE) is shifted downstream by the flow direction network to the next pixel at the dates of snowpack minimum.

The above steps are executed in order until the snow accumulation problem gets eliminated. I.e. some pixels (grid cells) need to use step (1) only, some steps (1)-(2), and some all three steps to solve excess snowpack accumulation problem.

5. Canopy interception of precipitation

Rainfall interception by vegetation can be significant for many land covers such as all forest types and some others. Intercept water on the vegetation canopy does not reach soil, evaporates and makes an additional contribution to the total evapotranspiration flux. The canopy intercept does not apply to snow which is assumed to be part of the total snowpack that shares common snow sublimation process.

WBM uses canopy rain interception formulations similar to those adopted in VIC model following monograph of [*Dickinson*, 1984]. The canopy water balance is given as following

$$\frac{dW_i}{dt} = (P - P_t) - E_c, \quad where \ W_i \leq W_i^{max} \tag{5-1}$$

where W_i is intercept canopy water storage (mm), t is time (d), P and P_t are rain precipitation and throughfall respectively (mm/d), E_c is evaporation of the intercept canopy water (mm/d). Note that the quantity in the round brackets of the RHS of eq. (5-1) is the canopy intercept as the rainfall water not reaching the ground. The canopy water storage is limited by its capacity W_i^{max} which is found to be proportional to the Leaf Area Index (LAI)-

$$W_i^{max} = C_{LAI} * LAI \tag{5-2}$$

where C_{LAI} is canopy interception coefficient (mm) which can vary from 0.15 by the BROOK90 [*Dingman*, 2002] to 0.25 in VIC model or a value of 0.2 mm as suggested in [*Dickinson*, 1984].

The canopy water evaporation rate E_c (mm/d) is defined as a simplification of the form presented by [*Deardorff*, 1978; *Dickinson*, 1984]

$$E_c = E_{ow} * \left(\frac{W_i}{W_i^{max}}\right)^{\frac{2}{3}} \tag{5-3}$$

WBM simplifies eq. (5-3) by neglecting aerodynamic resistance, and assuming open water evaporation rates instead of a specific evaporation rate calculated for fully wet leaf surfaces. Furthermore, WBM uses a Eulerian approximation of W_i from the previous timestep to estimate canopy evapotranspiration.

Throughfall (P_t) is calculated as rainfall that exceeds storage capacity and canopy evapotranspiration according to equation 5-4.

$$P_t = \begin{cases} P + W_i - E_c - W_i^{max} & if \ W_i^{max} < P + W_i - E_c \\ 0 & if \ W_i^{max} > P + W_i - E_c \end{cases} \tag{5-4}$$

Canopy interception storage (W_i) is then updated according to equation 5-1.

References

Deardorff, J. W. (1978), Efficient Prediction of Ground Surface-Temperature and Moisture, with Inclusion of a Layer of Vegetation, *Journal of Geophysical Research-Oceans*, *83*(Nc4), 1889-1903.

Dickinson, R. E. (1984), Modeling evapotranspiration for three-dimensional global climate models, *Climate Processes and Climate Sensitivity*, *Geophysical Monograph 29*, 58-72.

Dingman, S. L. (2002), *Physical hydrology*, 2nd ed., x, 646 p. pp., Prentice Hall, Upper Saddle River, N.J.

6. Soil moisture

Soil moisture balance, *Ws* [mm], is calculated with an accounting system that tracks a grid cell's water inputs, water outputs, and soil moisture pool holding capacity. The soil moisture pool depth is determined by the rooting depth. Inputs come in the form of precipitation as throughfall, P_t [mm d^{-1}], and as snow melt, M_s [mm d^{-1}]. Water intercepted by the canopy and ultimately evaporated, E_c, reduces how much precipitation reaches the soil (Section 5). Output is via actual evapotranspiration, *AET* [mm d^{-1}] (Section 3.1) and gravity drainage called soil surplus S [mm d^{-1}]. Soil moisture can be calculated for individual sub-pixel scale units defined by land-cover or crop type. The calculations presented below are repeated for each crop type being simulated. WBM uses perl Data Language slicing to improve performance of the set of equations. Fluxes leaving the root zone (S and *AET*) are summed according to pixel fraction for each land-cover type.

Change in soil moisture is calculated by equation 6-1.

$$\frac{dW_s}{dt} = P_t + M - AET - S$$

Throughfall (P_t) is discussed in Section 5, snow melt in Section 4, and actual evapotranspiration in Section 3. Soil surplus water S equals any water infiltrating soil in excess of available water capacity (equation 6-2).

$$S = \begin{cases} W_s^{k-1} + P_t + M - AET - W_{cap} & if\ W_{cap} < W_s^{k-1} + P_t + M - AET \\ 0 & if\ W_{cap} > W_s^{k-1} + P_t + M - AET \end{cases} \quad \text{[mm d}^{-1}\text{]} \quad (6\text{-}2)$$

7. Runoff

Runoff in WBM consists of storm runoff, surface runoff, baseflow, and irrigation runoff. The combined surface runoff and baseflow exit the terrestrial portion of each pixel, and are collected in a river network that allows the water to be transported downstream, the details of which will be discussed in Section 8.

7.1 Surface Runoff

When water inputs to a grid cell exceed the daily evapotranspiration and available water capacity then gravity drainage is initiated, defined in WBM as surplus water S [mm d^{-1}] leaving the root zone (Section 6). A fraction $(1 - \gamma$ [-]) of this surplus becomes quickflow, interpreted as representing flow through shallow soils and near stream surface runoff, R_s [mm d^{-1}]. Note The remaining fraction (γ [-]) of the surplus percolates to groundwater (I_P [mm d^{-1}]), either the shallow groundwater storage pool, W_g [mm d^{-1}] or to aquifers W_{Aqf} [m d^{-1}]. The groundwater percolation fraction (γ) defaults to 0.5, and is generally robust in the range of 0.4 to 0.6 (Zuidema et al. 2018, Stewart et al. 2011), but may vary beyond these ranges in specific contexts (Zuidema et al. 2020).[1]

---

[1] γ is a percolation fraction, setting how much of the surplus enters the groundwater pool. In Vörösmarty et al. (1998), γ indicates a surface runoff fraction, setting how much of the surplus becomes surface runoff.

Surface runoff is retained in a surface runoff retention pool (W_{SRP} [mm]) (called rainfall runoff detention pool in Wisser et al. (2010)) prior to draining to the stream network. Drainage from the surface runoff retention pool (R_{SRP} [mm d^{-1}]) follows a tank drain formulation:

$$R_{SRP} = C_{SRP} \sqrt{2\,G\,W_{SRP}} \qquad \text{[mm d}^{-1}\text{]} \qquad (7.1\text{-}1)$$

Where C_{SRP} is a unitless discharge coefficient of the surface runoff retention pool and includes unit conversions, and G is gravitational acceleration. A plot illustrating how R_{SRP} varies with W_{SRP} is provided as Figure 7.1.
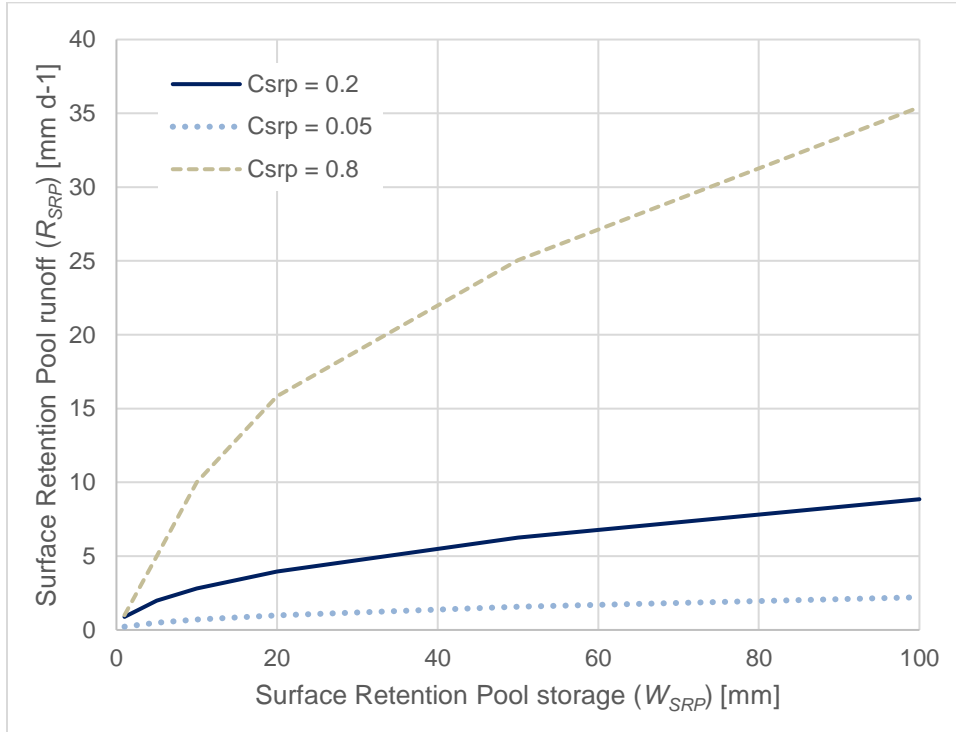

Figure 7.1: Calculated runoff from the surface retention pool across a range of values of storage within the pool for three reasonable values of the drain parameter C_{SRP}.

There is an upper limit, T_SRP [mm], imposed on the storage volume in the surface runoff retention pool. This limit captures the response of over-filled surface topographic depressions. When the volume of the surface runoff retention pool exceeds this limit, then the over-flow water, R_{EXC} [mm d^{-}], is immediately moved to the river. This helps to capture flashy hydrodynamic responses more accurately during extreme events (Zuidema et al., 2020).

The balance of the surface runoff retention pool W_{SRP} is expressed as:

$$\frac{dW_{SRP}}{dt} = R_S - R_{RSP} - R_{Exc} \qquad (7.1\text{-}2)$$

The balance of the surface runoff retention pool is calculated as a split operator in three stages that introduce inputs (1), calculate runoff (2), and then remove any remaining storage within the pool via over-flow water (3):

(1) $W_{SRP}^1 = W_{SRP}^k + R_s$ (7.1-3)

(2) $W_{SRP}^2 = W_{SRP}^1 - R_{SRP}$ (R_{SRP} is calculated using W_{SRP}^1)   (7.1-4)

$$R_{Exc} = \begin{cases} T_{SRP} - W_{SRP}^2 & if\ W_{SRP}^2 > T_{SRP} \\ 0 & if\ W_{SRP}^2 \le T_{SRP} \end{cases} \tag{7.1-5}$$

(3) $W_{SRP}^{k+1} = W_{SRP}^2 - R_{Exc}$ $\qquad\qquad\qquad\qquad\qquad\qquad$ (7.1-6)

Where W_{SRP}^k and W_{SRP}^{k+1} are the storage in the surface retention storage pool at the previous and present time-step, respectively. The threshold for storage in the surface runoff retention pool is set to 1,000 mm by default, meaning that unless otherwise specified as a non-default value, the storage in the surface retention pool is highly unlikely to be limited anywhere on the Earth's surface.

7.2 Irrigation Runoff

For irrigated croplands, a separate surface storage pool W_{Irr} is maintained to separate differing water inputs for irrigated and non-irrigated portions of pixels. The balance of this pool and runoff from irrigated portions of pixels (R_{Irr}) is calculated identically to surface runoff retention pool; however, the upper limit to surface retention does not apply, and there is no excess surface runoff (e.g. R_{Exc}) calculated for irrigated areas; the balance of W_{Irr} is calculated in only stages 1 and 2 above.

7.3 HBV Direct Recharge

WBM has the option to also introduce direct recharge (I_D), following the method of Hydrologiska Byråns Vattenbalansavdelning (HBV - Bergström and Lindström, 2015). Direct recharge simulates immediate recharge of slow response groundwater pools during precipitation events, likely through direct connections to groundwater via macro-pore flow, and is calculated prior to soil balance calculation as:

$$I_D = (P_t + M) * \left(\frac{W_S}{AWC}\right)^{\beta_D}$$

where $P_t + M$ [mm d^{-1}] is effective precipitation incident to the soil surface (following canopy interception), W_S [mm] is water storage in the soil or root zone, AWC [mm] is available water capacity of the soil, and β_D is the HBV direct recharge shape parameter (Bergström and Lindström, 2015). If direct recharge is calculated by WBM, effective precipitation infiltrating to soil (I_S [mm d^{-1}]) is calculated as:

$I_S = I - I_D$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (7.3-1)

Otherwise, if direct recharge is not calculated, then:

$I_S = I$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (7.3-2)

Direct recharge is added to soil percolation (I_P) to calculate total groundwater recharge (I_G):

$I_G = I_S + I_D$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (7.3-3)

However, if direct recharge is not calculated, then total groundwater recharge consists soil of soil percolation:

$I_G = I_S$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (7.3-4)

7.4 Groundwater

Groundwater recharge (I_G) is the sum of soil percolation and direct recharge. Groundwater is represented as both a shallow groundwater pool, and optionally as aquifers which can be represented in three different

ways.[2] We refer to the shallow groundwater pool, and interpret this pool as representing the hydrodynamic response of subsurface water responding to recharge events and generating baseflow conceptualized as residing in shallow alluvial aquifers proximal to streams. Aquifer representations are described in Section 9. Where aquifers are represented (they are optionally represented in none, in part of, or over the entire model domain), soil percolation is partitioned to a fraction recharging shallow groundwater (I_{SGW}), and a fraction recharging aquifers (I_{Aqf}) by:

$$I_{Aqf} = \gamma_{Aqf}\, I_G \tag{7.4-1}$$
$$I_{SGW} = \left(1 - \gamma_{Aqf}\right) I_G \tag{7.4-2}$$

where γ_{Aqf} [-] is the aquifer percolation fraction, and defaults to zero when aquifers are not defined. I_{Aqf} is directed to aquifers (Section 9), and I_{SGW} represents recharge to the shallow groundwater represented as a simple retention pool.

Water drains from the groundwater storage pool (W_{SGW} [mm]) to streams through baseflow (R_{SGW} [mm d^{-1}]), at a rate defined by the hydrodynamic groundwater response constant (β [d^{-1}]).

$$R_{SGW} = \beta W_{SGW} \quad [\text{mm d}^{-1}] \tag{7.4-3}$$

The total change in groundwater is then the percolation from surplus, (i.e., recharge), minus the loss to baseflow.

$$\frac{dW_{SGW}}{dt} = I_{SGW} - \beta W_{SGW} \tag{7.4-4}$$

β [d^{-1}] is an empirical constant that defaults to 0.0167 [d^{-1}] meaning that typical baseflow recession has a time-scale of 60 days by default.

7.5 Storm runoff

Impervious?

Storm runoff directs water to streams immediately with no lag in time. Storm runoff is generated as melt and precipitation on impervious or open water surfaces, as well as runoff that exceeds the surface runoff retention pool limit (R_{Exc}).

All precipitation and melt on open-water surfaces is considered open-water storm runoff (R_{ow} [$mm\ d^{-1}$]).

$$R_{ow} = f_{ow}(P + M_s) \quad [\text{mm d}^{-1}] \tag{7.5-1}$$

where f_{ow} is the fraction of pixel area covered by open water. Impervious areas prevent water from entering soils and increases overland runoff. If provided with an impervious area map, WBM calculates overland runoff in impervious areas, R_imp [$mm\ d^{-1}$] as:

$$R_{Imp} = C_{imp} f_{imp}(P_t + M_s) \quad [\text{mm d}^{-1}] \tag{7.5-2}$$

where C_{imp} [-] is the hydrologically connected impervious fraction, a unitless scalar for impervious surfaces that determines the fraction of precipitation over impervious areas that is directly routed to streams, f_{imp} is the pixel area fraction covered by impervious surfaces. Precipitation incident to impervious surfaces include calculation of canopy interception to account for vegetation co-located with imperviousness. WBM assumes a relationship for directly connected imperviousness from Alley and Veehuis (1983) that assumes that degree of impervious connectedness scales non-linearly with the fraction of impervious cover (f_{imp}) in each pixel:

---

[2] Drainage from aquifers add additional runoff above the runoff generated by the mechanisms described here. Types of drainage vary by the form of aquifer representation, and are described in Section 9.

$$C_{imp} = f_{imp}{}^{0.4} \tag{7.5-3}$$

Total storm runoff is the sum of storm, and open-water runoff and excess surface runoff:

$$R_{storm} = R_{ow} + R_{imp} + R_{Exc} \quad [\text{mm d}^{-1}] \tag{7.5-4}$$

7.6 Total Runoff

The total amount of water that exits the terrestrial portion of the pixel and enters the stream network (total runoff, R_t [mm d^{-1}]) is the sum of the surface retention pool release, irrigation retention pool release, baseflow, and storm runoff:

$$R_t = R_{SRP} + R_{Irr} + R_{SGW} + R_{storm} \quad [\text{mm d}^{-1}] \tag{7.6-1}$$

References

Stewart, R. J., Wollheim, W. M., Gooseff, M. N., Briggs, M. A., Jacobs, J. M., Peterson, B. J., & Hopkinson, C. S. (2011). Separation of river network–scale nitrogen removal among the main channel and two transient storage compartments. Water Resources Research, 47(1). https://doi.org/10.1029/2010WR009896

Wisser, D., Fekete, B. M., Vörösmarty, C. J., & Schumann, A. H. (2010). Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H). Hydrology and Earth System Sciences, 14(1), 1–24. https://doi.org/10.5194/hess-14-1-2010

Zuidema, S., Wollheim, W., Mineau, M. M., Green, M. B., & Stewart, R. J. (2018). Controls of Chloride Loading and Impairment at the River Network Scale in New England. Journal of Environmental Quality, 47(4), 839–847. https://doi.org/10.2134/jeq2017.11.0418

Zuidema, S., Grogan, D., Prusevich, A., Lammers, R., Gilmore, S., & Williams, P. (2020). Interplay of changing irrigation technologies and water reuse: example from the upper Snake River basin, Idaho, USA. Hydrology and Earth System Sciences, 24(11), 5231–5249. https://doi.org/10.5194/hess-24-5231-2020

8. River routing

WBM has three options for calculating hydrologic routing of water through a river network. The river network is represented as a 1-dimensional cell-table (directed graph) where each subsequent entry is guaranteed to be on a separate flow-path, or is downstream of all preceding entries. WBM checks for circularity in the river network and is prevented from running if found.

8.1 Hydraulic geometry

Related to routing are a series of properties that describe the hydraulic geometry of stream channels. WBM incorporates both downstream and at-a-station stream geometry relationship assumptions to calculate river width, depth, and velocity from discharge. WBM assumes that each grid cell has a single representative stream reach and calculates a rolling average of annual mean discharge for each reach in a simulation over the previous five-years of a simulation. The long-term mean discharge, $\bar{Q}$, [m^3/s] is then used to estimate the long-term mean depth, $\bar{z}$, [m], width, $\bar{y}$, [m], and velocity, $\bar{u}$, [m/s] using downstream hydraulic geometry relations and scalers from (Park, 1977):

$$\bar{z} = \eta \bar{Q}^v$$

$$\bar{y} = \tau \bar{Q}^{\phi}$$
$$\bar{u} = \delta \bar{Q}^{\epsilon}$$

Instantaneous estimates of the three variables (z [m], y [m], and u [m/s] for depth, width and velocity, respectively) are given as functions of instantaneous Q [m³/s] and mean discharge $\bar{Q}$ [m³/s], scaled by appropriate at-a-station hydraulic geometry exponents (S.L. Dingman, 2009).

$$z = \bar{z} \left(\frac{Q}{\bar{Q}} \right)^{f}$$

$$y = \bar{y} \left(\frac{Q}{\bar{Q}} \right)^{b}$$

$$u = \bar{u} \left(\frac{Q}{\bar{Q}} \right)^{m}$$

In the above equations, parameters $\eta, \nu, \tau, \phi, \delta, \epsilon, f, b$ and m are all user defined variables set to defaults found in (Leopold & Maddock, 1953; Park, 1977).

8.2 Flow accumulation

The simplest routing routine employed by WBM is flow accumulation, where all incoming runoff and upstream discharge is immediately moved to the next downstream pixel.

8.3 Muskingum

In the case where simulations use a coarse spatial resolution (e.g., half a degree of latitude and longitude) such that river flow likely remains within the grid cell on a daily time step, WBM can use the Muskingum flow routing option. Unfortunately, Muskingum routing does not account for residual in-stream water storage and other anthropogenic or natural water abstractions from streamflows, and WBM will exit if there is this identified mismatch in routines. Muskingum routing has limitations on pixel size and time steps requiring cell's Courant number (i.e. fraction of cell size travelled by the flood wave during time step Δt) to be much less than 1. These limitations prohibit use of Muskingum routing in many WBM model settings. In cases where necessary conditions for using Muskingum routing are met, this method is preferred over Linear Reservoir Routing (LRR) because it is derived from a simplification of hydraulics accounting for non-uniform flow across the reach during changes in flow. LRR assumes uniform instream storage and flow within a grid cell.

WBM's Muskingum flow routing option estimates the flow rate and water level in each grid cell's stream segment using a distributed flow routing model based on the Saint-Venant partial differential equations for one-dimensional flow. Specifically, this is the Muskingum-Cunge kinematic wave model that approximates the solution to the Saint-Venant partial differential equations (Maidment, 1992). These equations require six assumptions:

1. Flow from grid j to grid j + 1 is one-dimensional,
2. The stream length through the grid cell is significantly larger than the flow depth,
3. Vertical acceleration and vertical changes in pressure are negligible,
4. Water density is constant,
5. Channel bed and banks are immobile, and
6. Channel bottom slope is small, less than 15%.

Additionally, WBM assumes a rectangular channel bed and no loss of water from the channel to groundwater.

The Muskingum-Cunge solution estimates the outflow, Q_{j+1}^{t+1} [m^3 s^{-1}], at time t+1 and grid cell j+1, as a linear combination of three known inflows and outflows. These are:

1) the inflow of the current time step and previous grid cell, Q_j^{t+1} [m^{3}s^{-1}],

2) outflow of the previous time step and current grid cell, Q_{j+1}^t [m^{3}s^{-1}], and

3) inflow from the previous time step and adjacent upstream grid cell, Q_j^t [m^{3}s^{-1}]:

$$Q_{j+1}^{t+1} = C_0 Q_j^{t+1} + C_1 Q_{j+1}^t + C_2 Q_j^t \qquad (8.1\text{-}1)$$

The coefficients C_0 [-], C_1 [-], and C_2 [-], are defined such that:
$$C_0 + C_1 + C_2 = 1 \qquad (8.1\text{-}2)$$

and if any of these three coefficients are less than 0, they are reset to 1, 0, and 0, respectively. The coefficients are unitless functions of the Courant number, C, and Reynolds number, D:

$$C_0 = \frac{-1+C+D}{1+C+D} \qquad (8.1\text{-}3)$$

$$C_1 = \frac{1+C-D}{1+C+D} \qquad (8.1\text{-}4)$$

$$C_2 = \frac{1-C+D}{1+C+D} \qquad (8.1\text{-}5)$$

Both C and D depend on riverbed geometry, and are defined as:

$$C = U_w V_m \frac{\partial t}{L} \qquad [-] \qquad (8.1\text{-}6)$$
$$D = \frac{Y_m}{S_0 U_w L} \qquad [-] \qquad (8.1\text{-}7)$$

where U_w [m^3 s^{-1}] is the speed of wave propagation (also referred to as the wave celerity), V_m is the mean annual fluid velocity [m s^{-1}] defined below, L is the river length in the grid cell [m], dt [s] is the time step length (daily), Y_m is the mean flow depth [m], and S_0 is the riverbed slope [m km^{-1}]. These variables are defined as:

$$U_w = \left(1 + \frac{\frac{2}{3}\sigma}{\sigma+1}\right) V_m \qquad (8.1\text{-}8)$$

where the shape parameter σ = 2 [-],
$$V_m = \frac{Q_m}{Y_m W_m} \qquad (8.1\text{-}9)$$

where Q_m is the mean annual discharge in the river segment [m^{3}s^{-1}], and W_m is the corresponding mean annual channel width [m]:

$$W_m = \tau Q_m^{\varphi} \qquad (8.1\text{-}10)$$

where τ [-] and φ [-] are constants 8.0 and 0.58, respectively (Knighton, 1998).

28

Parameter Y_m is calculated as:

$$Y_m = \eta Q_m^v \hspace{7cm} (8.1\text{-}11)$$

Where η and v are empirical constants of 0.25 and 0.4, respectively (Knighton, 1998), and S_0 is an input to the model that defaults to 0.1.

River length L is calculated as (Fekete et al. 2001):

$$L = \frac{N\sqrt{A_c}}{1-0.077\log(A_c)}$$
$$(8.1\text{-}12)$$

Where A_c [m²] is the area of the grid cell and N is direction factor that depends on whether flow crosses the pixel in cardinal or ordinal directions.

$$N = \begin{cases} 1 & \text{if pixel drains to N, S, E, or W} \\ \dfrac{1}{\sin\left(\frac{\pi}{4}\right)} & \text{if pixel drains to NE, SE, NW, SW} \end{cases} \hspace{2cm} (8.1\text{-}13)$$

As the discharge is calculated for each time step within a grid cell, the discharge value is stored so that it can be used to determine the mean annual discharge in future calculations. Mean annual discharge reflects a rolling average of the previous five years of mean annual discharge. Grid cells which are defined as open water (e.g., lakes) use the flow accumulation routing scheme, in which water is transported immediately between the grid cell and the open water outlet point. In this case, the coefficients C_0, C_1, and C_2 are redefined to equal 1, 0, and 0, respectively. Routing delays on open water bodies are simulated by WBM's reservoir operations (Section 11).

8.4 Linear reservoir routing

The linear reservoir routing (LRR) method implemented by WBM reflects a common approach for simple routing schemes (Dingman, 2002, p429). LRR provides a dampened routing response like Muskingum; however, does not provide any delay in the onset of the flood wave propagation.

Let us consider continuity (mass conservation) for surface water storage (river) as a partial differential equation

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \hspace{6cm} (8.2\text{-}1)$$

where the first term is rate of rise of flow cross sectional area (for an assumed rectangular channel), A, the second term is flow, Q, gradient through the grid cell. Its differential form with introduction of (reservoir) storage, S, inflow, Q_{in}, and outflow, Q_{out}, it can be transformed to a full differential form

$$\frac{dS}{dt} = Q_{in} - Q_{out} \hspace{5cm} (8.2\text{-}2)$$

If cell surface water is considered to be an ideal reservoir then the change in storage is a function of outflow only, i.e.

$$S = f(Q_{out}) \hspace{6cm} (8.2\text{-}3)$$

which has a common form of (e.g. hydrograph)-

$$S = K Q_{out}{}^n \hspace{6cm} (8.2\text{-}4)$$

29

LRR is a special case when the power term is equal to 1 and the equation (4) becomes a linear relation between storage and outflow.

Next step in formulation of LRR is finding the scaling coefficient K in the equation (4). Let us assume constant velocity and uniform water volume distribution within its travel time reach which leads to the following system of equations

$$\begin{cases} S_{total} = S_{out} + S_{in} \\ S_{out} = Q_{out}\,\Delta t \\ S_{in} = Q_{out}\,\frac{\Delta l}{U_w} \end{cases} \qquad (8.2\text{-}5)$$

where S_{total} is total storage (volume) of water, i.e. S in equations (1-4), that has to be distributed between this pixel, S_{in}, and outflow volume to the downstream pixel, S_{out}; Q_{out} is outflow rate (discharge) from the given cell, Δt is time step, Δl is this cell river reach, and U_w is wave celerity. Note that term $\Delta l / U_w$ is time for the flood wave to propagate along the cell river reach. The total storage is usually composed of the following terms

$$S_{total} = S_{in}^0 + (Q_{in} + R + Q_{abs} - w\Delta l E_{ow} + D_{RIV})\,\Delta t \qquad (8.2\text{-}6)$$

where S_{in}^0 is instream water storage in the cell from the previous time step, Q_{in} is inflow, R is runoff rate (converted to volumetric flow in m^3 s^{-1}), and Q_{abs} is collective water abstraction within a pixel that may include human water use withdrawals and returns. Open water evaporation and exchange with local aquifers (if simulated) also affect storage within the reach. Solving equation (5) for Q_{out} in regard to variables that are known or can be evaluated in the model, i.e. S_{total} and U_w, yields-

$$Q_{out} = \frac{S_{total}}{\Delta t + \frac{\Delta l}{u}} = \frac{1}{1 + \frac{\Delta l}{u\Delta t}}\frac{S_{total}}{\Delta t} = \frac{C}{1+C}\frac{S_{total}}{\Delta t} \qquad (8.2\text{-}7)$$

where C is cell's Courant number $C = U_w\frac{\Delta t}{\Delta l}$ which is a fraction of river reach within cell travelled by the flood wave during time step Δt.

Equation (7) represents a linear relation of storage with outflow indicating a linear (n=1) solution for equation (4) above.

References:

Dingman, S. L. (2008). Physical hydrology. Long Grove, IL: Waveland Press Inc.

Fekete, B. M., Vörösmarty, C. J., & Lammers, R. B. (2001). Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error. Water Resources Research, 37(7), 1955–1967.

Knighton, D. (1998) Fluvial Forms and Processes: A New Perspective. Oxford University Press, Inc., Arnold, London.

Leopold, L. B., & Maddock, T. (1953). The Hydraulic Geometry of Stream Channels and Some Physiographic Implications (Geological Survey Professional Paper No. 252) (p. 64). Washington, D. C.: United States Geological Survey. Retrieved from https://doi.org/10.3133/pp252

Maidment, D.R. (1992) Handbook of Hydrology. McGraw-Hill, Inc., Columbus, Ohio

Park, C. C. (1977). World-wide variations in hydraulic geometry exponents of stream channels: An analysis and some observations. Journal of Hydrology, 33(1), 133–146. https://doi.org/10.1016/0022-1694(77)90103-2

9. Groundwater

All WBM simulations utilize the shallow groundwater pool to simulate hydrodynamic response of baseflow. We conceptualize the shallow groundwater pool (Section 7.4), as representing groundwater flowpaths that are entirely contained within the pixel. To simulate the effects of regional aquifers, WBM has three options that may be used.

9.1 Low resolution: unparameterized aquifers

At the lowest level of parameterization, WBM can simulate water extractions from an unlimited unsustainable groundwater pool, in addition to the shallow groundwater pool that is explicitly represented. The state of the unsustainable pool is not simulated directly within WBM, i.e., there is no accounting of the volume of water in this imaginary pool. Rather, when water extractions are needed, water can be withdrawn and added to the soil or other WBM water stock subject to irrigation demand parameter values (Section 13).

For the purposes of calculating total water storage (TWS), the amount of water taken from this pool is accumulated daily, providing an estimate of water extracted from unsustainable groundwater sources. Total water storage is an output of WBM that sums all water stores in a pixel.

Other than water extractions, there is no interaction between unsustainable groundwater and other water pools within WBM; there is no recharge to and no baseflow from unsustainable groundwater.

9.2 Medium resolution: lumped aquifers

WBM can account for large aquifers using a lumped aquifer representation with unidirectional vertical movement, to provide a tradeoff between the parsimony of unparameterized aquifers (Section 9.1), and the quasi-three dimensional representation discussed below (Section 9.3). Lumped aquifers can be represented over all or portions of the model domain. Recharge percolating through the root zone is proportioned between shallow groundwater (γ_{SGW}) and the deeper (lumped) aquifer ($\gamma - \gamma_{SGW}$) at each pixel overlying an identified aquifer. Additionally, inflows from the surface flow network can be specified as point-based losing reaches that infiltrate directly to the aquifer (bypassing the shallow groundwater pool); flows to aquifers via river sinks (I_{snk}) are parameterized as a fraction of daily flow (χ_{snk}) by equation 9.2-1.

$$I_{snk} = \Sigma_j^n \chi_{snk}^j Q^j \tag{9.2-1}$$

where Q^j is the is discharge at river sink location j. Outflows from the aquifer occur as springs represented as points with head-dependent conductance similar to drains in MODFLOW (Harbaugh, 2005). Average head within the lumped aquifer head is calculated as:

$$h = \frac{S_A}{C_A} * Z_A + Z_0, \tag{9.2-1}$$

where h is aquifer head (m), S_A is the volume stored within the aquifer (km^3), C_A is the capacity of the aquifer (km^3) (so the ratio of $\frac{S_A}{C_A}$ is the fractional storage), Z_A is the aquifer thickness (m) and Z_0 is the base elevation (m). Drainage through individual springs (Q_{spr}) is calculated as:

$$Q_{spr} = K_{spr}(h - Z_{spr}) \tag{9.2-2}$$

where K_{spr} is an individual spring's conductance ($m^2\,d^{-1}$), and Z_{spr} is the elevation of each spring (m). Q_{spr} is then summed for all individual springs in calculating the storage balance of the lumped aquifer, and Q_{spr} is added to the river network at individual locations. All recharge to and abstractions from the aquifer are summed through the previous day and mass balance of the aquifer is updated at a daily time-step using a Runge-Kutta 3(2) order (Bogacki-Shampine) scheme. Under this split operator solution, water percolating to and pumped from the aquifer is assumed to influence aquifer volume following a one day lag. The single-day lag is expected to underestimate percolation travel-times through the unsaturated zone and the far-field hydrodynamic response of the aquifer to changes in pumping. The volume of water represented by the lumped aquifer model is assumed not to interact with shallow or root zone water (i.e. head is assumed to remain below the base of these zones) and fluxes from the aquifer to these zones are neglected.

9.3 High resolution: MODFLOW aquifers

Formulations of the ModFlow Groundwater module in WBM

WBM provides a finite-difference solution to groundwater head following the formulations of MODFLOW to provide spatially resolved head in circumstances where aquifer head-dependent aquifer exchange with surface water, or characterizing depth to groundwater for abstractions are needed. ModFlow Groundwater module in WBM replicates a few basic groundwater-related functionality of the USGS MODFLOW software [*Langevin et al.*, 2017] in order to apply to regional and Global scale modeling [*de Graaf et al.*, 2015; *de Graaf et al.*, 2017].

Equations for groundwater flow continuity. Case of homogeneous horizontal flow (eq. (8-14) in [*Dingman*, 2002]) is given as a common conductivity/diffusivity equation

$$S_s \frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(K_h \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_h \frac{\partial h}{\partial z}\right) + w \tag{9.3-1}$$

where h is head (m), K_h is the hydraulic conductivity (m d^{-1}) with time units in days (d), S_s is specific storage (m^{-1}), and w is external flux (withdrawal or sink) per unit volume (d^{-1}). Natural aquafers usually have anisotropic conductivity so that K_h has to be referenced by the flow direction along all three axes. Discretization of flow equation (1) is done by application/multiplication to a cell finite volume (V) which converts it a continuity equation (Note: water volume is used as the conservation term instead of mass assuming that water is incompressible liquid). E.g. for *x*-direction term in equation (1) becomes

$$q_i = V_{i,j,k}\left[\frac{\partial}{\partial x}\left(K_h \frac{\partial h}{\partial x}\right)\right]_i = K_{i-\frac{1}{2}}A_{i-\frac{1}{2}}\frac{\Delta h_{i-\frac{1}{2}}}{\Delta x_{i-\frac{1}{2}}} + K_{i+\frac{1}{2}}A_{i+\frac{1}{2}}\frac{\Delta h_{i+\frac{1}{2}}}{\Delta x_{i+\frac{1}{2}}} \tag{9.3-2}$$

where indices (i,j,k) are discrete notations for (x,y,z) axes, q_i is volumetric *x*-direction flux (m^3/d) into the cell (i,j,k) with volume $V_{i,j,k}$, A is cell side wet/saturated surface area, and indices $\pm1/2$ refer to the opposite orthogonal sides of the cell to a given flux direction. Head gradient is defined as head separation pressure (difference) from outside to the cell (distance in the denominator is always positive)-

$$\frac{\Delta h_{i-\frac{1}{2}}}{\Delta x_{i-\frac{1}{2}}} = \frac{h_{i-1} - h_i}{x_i - x_{i-1}} \tag{9.3-3}$$

Accordingly the left-hand side term in the equation one becomes

$$Q_{i,j,k} = V_{i,j,k}\left[S_s \frac{\partial h}{\partial t}\right]_{i,j,k} = V_{i,j,k}S_{s(i,j,k)}\frac{\Delta h_{i,j,k}}{\Delta t} \tag{9.3-4}$$

where $Q_{i,j,k}$ is total water flux (m^3/d) into the cell (i,j,k).

Following special notations used in the MODFLOW formulations [*McDonald and Harbaugh*, 1988] for groundwater flux numerical implementation, a cell conductance term C (m^2/d) can be introduced. It significantly simplifies equation notations. E.g. for *i-1/2* boundary it is

$$C_{i-\frac{1}{2}} = K_{i-\frac{1}{2}}A_{i-\frac{1}{2}}\frac{1}{\Delta x_{i-\frac{1}{2}}} \tag{9.3-5}$$

Combining equations (1)-(5) yields

$$Q_{i,j,k} = C_{i-\frac{1}{2}}\Delta h_{i-\frac{1}{2}} + C_{i+\frac{1}{2}}\Delta h_{i+\frac{1}{2}} + C_{j-\frac{1}{2}}\Delta h_{j-\frac{1}{2}} + C_{j+\frac{1}{2}}\Delta h_{j+\frac{1}{2}} + C_{k-\frac{1}{2}}\Delta h_{k-\frac{1}{2}} + C_{k+\frac{1}{2}}\Delta h_{k+\frac{1}{2}}$$
$$+ W_{i,j,k}$$
(9.3-6)

where W is volumetric external flux (m³/d) into the cell (withdrawal or sink). Equation (9.3-6) is the fundamental grounds water flow equation to be solved in WBM.

Formulations for the flow equation (9.3-6) parameters- Time and head dependent conductance parameters and head to storage relation of the flow equation has to be evaluated on each time step of the groundwater flow simulations. While formulations for these are given above in equations (9.3-4) and (9.3-5), an important practical modification come by the following substitutions [*Dingman*, 2002]

$$S_y = H S_s$$
(9.3-7)

$$T = H K_h$$
(9.3-8)

where S_y is storage coefficient (m m⁻¹), and T is transmissivity (m² d⁻¹). The reason for using storage coefficients in place of specific storage is that S_y values are "dimensionless" and relevant to sediment porosity (unconfined aquifers) or to a deformation modulus (confined aquifers) which are primary material parameters. The use of transmissivity allow much larger vertical dimensions of a cell (> 10 m) as hydraulic conductivity K_h can significantly vary [*Ingebritsen and Manning*, 1999]. These substitutions yield

$$Q_{i,j,k} = A_{k(i,j,k)}S_{y(i,j,k)}\frac{\Delta h_{i,j,k}}{\Delta t} = \Delta x_{i,j,k}\Delta y_{i,j,k}S_{y(i,j,k)}\frac{\Delta h_{i,j,k}}{\Delta t}$$
(9.3-9)

$$C_{i-\frac{1}{2}} = T_{i-\frac{1}{2}}b_{i-\frac{1}{2}}\frac{1}{\Delta x_{i-\frac{1}{2}}}$$
(9.3-10)

where A_k is cell area or area orthogonal to z-direction (k indices), and b is the saturated thickness of the cell interface. Note that eq. (9.3-10) does not apply to cell conductance C in z-direction that has still to be calculated using eq. (9.3-5).

Permeability and conductivity- Permeability (P, m²) is a universal geometrical metric for connected pore network in rock or sediment materials. In case of fluid substance to be water, permeability relates to conductivity (K, m/sec) through water density (ρ, kg/m³), Earth gravity constant (g, kg/m³), and the dynamic viscosity of water (μ, kg/(m·sec)) as following:

$$K = P\frac{\rho\,g}{\mu}$$
(9.3-11)

In turn, both are known to decease with depth [*Ingebritsen and Manning*, 1999] exponentially as size and density of pore network tend to decrease with depth. We use formulations of [*Miguez-Macho et al.*, 2007] that use e-folding value, f, as the exponent factor for the conductivity decrease with depth, z, as following

$$K_h(z) = K_h^{ref}e^{\frac{-(z-z_{ref})}{f}}$$
(9.3-12)

where K_h^{ref} is a hydraulic conductivity at the reference depth, z_{ref}. Surface (zero depth) is most commonly used as the reference depth, so that $z_{ref}=0$ which some simplifies the equation (9.3-12).

On Global and regional spatial scales, which is the most common case of WBM applications, a generalized morphology of aquifers is used with relatively thick aquifer layers where variability of hydraulic conductivity can be significant and highly dependent on vertical location of the wetted portion of the layer in the simulation finite volume cell so that the transmissivity equation (9.3-8) becomes inapplicable. That problem can be addressed by using transmissivity integral over the wetted vertical portion of the cell from

depth z_0 to z_1 (the former is the head depth in case of unconfined aquifer) as it was done in [*de Graaf et al.*, 2015]

$$T = \int\limits_{D(z_0)}^{D(z_1)} K_h^{ref} e^{\frac{-z}{f}} dz, \quad \text{where} \quad z_{ref} = 0 \tag{9.3-13}$$

e-folding factor: *e*-folding factor, *f*, is assumed to be an inverse function of surface roughness [*Fan et al.*, 2007] expressed as average tangent slope, β, of the terrain

$$f = \frac{a}{1 + b\beta} \quad \text{for} \quad \beta \leq 0.16; \quad f = 5\,m \quad \text{for} \quad \beta > 0.16 \tag{9.3-14}$$

The values of coefficients *a* and *b* depend on the sediment or bedrock type. In [*Fan et al.*, 2007] these are

       Regolith: *a* = 120 m and *b* = 150

       Bedrock: *a* = 20 m and *b* = 125

Averaged values used in [*de Graaf et al.*, 2015] are

       Averaged: *a* = 100 m and *b* = 150

Hydraulic conductivity at the surface and storage coefficient: The parameters K_h^{ref} and S_y in 5 arc degree resolution for Global and regional simulations is reported in [*de Graaf et al.*, 2015; *de Graaf et al.*, 2017] where it is, in turn, derived from the Global lithological map (GLiM) [*Hartmann and Moosdorf*, 2012].

9.3.2 Numerical solution for the MODFLOW groundwater flow in WBM using EFDM solver

The flow equation (9.3-1) with variable coefficients decomposed for finite volumes for the numerical solution (e.g. in the re-arranged form of eq. (9.3-6)) becomes a well know system of second order partial differential equations (PDF) which presently do not have explicit solution for relatively large spatial domains (> N*100 cells) using existing computing resources and technologies. USGS groundwater software package MODFLOW has few choices of implicit backward difference method (IBDM) for the PDF including the most advanced Newtonian solver/method which strength is in ability to work with asymmetric matrix option that often result from boundary condition constraints. The core formulation of the IBDM methods comes to composing a vector equation [*Harbaugh et al.*, 2000] to be solved iteratively until the precision (delta) of the approximate solution is reached

$$[A]\{h\} = \{q\} \tag{9.3-15}$$

where $[A]$ is a (N x N) matrix of the PDF coefficients, $\{h\}$ and $\{q\}$ are N vectors of cell heads and cell fluxes accordingly. N is number of cells in the simulation domain (2D or 3D).

Main advantage of using IBDM solvers is their stability for a wide range and variability of the flow parameters, spatial domain resolutions, and steady-state and transient cases including coarse time steps. Unfortunately, despite all aforementioned advantages the IBDM solvers are resource prohibitive and cannot be used for groundwater flow simulations coupled with WBM because these are very computationally intensive that require sub-hour clock time to be solved for each time step of transient simulations on a high end multicore server for an average size spatial domain of 100k cells or higher [*Hughes and White*, 2013].

Another alternative to solve PDF is the ***explicit forward difference method*** (EFDM) which main limitations are in stability resulting in strict constraints mostly relevant to a cell size (needs to be relatively large), transient simulation time steps (need to be relatively small), and inability of this method to solve steady state problem. But the main advantage of this method is small computational load, and so speed and CPU efficiency. The validity of the EFDM solution can be predicted using the Frederick's stability criteria [*Crank*, 1975] which has to be less than 1/2. We found, typical WBM spatial and temporal domains used in very much all our simulations satisfy the Frederick's stability criteria that in our formulation can be expressed as following

$$r = \frac{\overline{T} \, \Delta t}{L^2} \approx 0.000004 \leq \frac{1}{2} \text{ for } 6'' \text{ and } \approx 0.0004 \leq \frac{1}{2} \text{ for } 1 \text{ km grids} \tag{9.3-16}$$

where $\overline{T} = \frac{T}{S_y}$ and L is a linear scale of cells (size).

The "forward difference" of EFDM method assumes that the head vector $\{h\}$ and matrix of coefficients $[A]$ represents state of the previous time step, *n-1*, of the computational stencil and a resulting flux vector $\{q\}$ is projected "forward" on the time scale as it represents current time step, *n*. And, finally, "explicit" term of the method refers to the temporal relation between left- and ring-hand vectors of eq. (15) which are not on the same time temporal (time step) point and, so, are "explicit" (direct) approximate solution not requiring iterations to solve. These assumptions for the EFDM method can be expressed by modified equation (9.3-15) for the approximate solution

$$[A]^{n-1}\{h\}^{n-1} \approx \{q\}^n \tag{9.3-17}$$

where superscript *n* is a time step of the simulation.

Equations (9.3-17) combined with stability/convergence criteria (9.3-16) checked on each time step are used as a PDF solver of groundwater flows in WBM.

Boundary and initial conditions

Boundary conditions for the WBM MODFLOW groundwater problem are treated in a much similar manner as in any other fluid dynamics or heat/diffusion transfer numerical problem. Boundary conditions are assessed at each time step of the simulation, and they can be subdivided into two groups relevant to (1) the aquifer geometry, (2) source terms, and (3) special cases-

1. Geometry boundary conditions:

    a. Each aquifer vertical layer is masked by nodata value in each cell where its thickness is equal to zero.

    $$(C, W, h, q)_{i,j,k} = \emptyset \text{ where } \Delta z_{i,j,k} = 0 \tag{9.3-18}$$

    b. Conductance, C, is set to zero at the non-ocean boundary of the aquifer

    $$C_{(i,j,k)\pm\frac{1}{2}} = 0 \text{ where } \left\{ \Delta z_{(i,j,k)\pm 1} = 0 \ or \ (C, W, h, q)_{(i,j,k)\pm 1} = \emptyset \right\} \tag{9.3-19}$$

    c. Horizontal conductance, $C_{i,j\pm 1/2}$, is set to half of the opposite cell side conductance at the non-simulated aquifer boundary (optional). This boundary is defined as the simulated river network (watershed) boundary where aquifer extends outside of it to adjacent cells that are not simulated by WBM

$$C_{(i,j,k)\pm\frac{1}{2}} = 0.5\ C_{(i,j,k)\pm 1\pm\frac{1}{2}} \quad \text{where}\ \left\{\Delta z_{(i,j,k)\pm 1} \neq 0\ \ and\ \ (C, W, h, q)_{(i,j,k)\pm 1} = \emptyset\right\} \tag{9.3-20}$$

This boundary condition is optional. It superposes condition (9.3-19) when enabled. The reduction coefficient 0.5 in equation (9.3-20) is arbitrary chosen, but it has to be < 1 for stability reasons.

d. Head of side-adjacent cells that are ocean or endorheic lake is set to the ocean or endorheic lake elevation which is surface elevation, H^s

$$h_{(i,j,k)\pm 1} = H_{i,j,k}^s \quad \text{where}\ \ [(i,j,k) \pm 1] \ni [Ocean\ or\ Endorheic\ Lake] \tag{9.3-21}$$

e. Head of the uppermost aquifer layer in open water cells (e.g. lakes, reservoirs, etc.) is set to the lake elevation which is surface elevation, H^s

$$h_{i,j,k} = H_{i,j,k}^s \quad \text{where}\ \ (i,j,k) \ni Lakes \tag{9.3-22}$$

f. Head of the uppermost aquifer layer above surface elevation is removed as baseflow, Q^{bf} (m/day) and its negative value refers to a drain from the aquifer (negative storage change)

$$Q_{i,j,k}^{bf} = -\frac{S_y\left(h_{i,j,k} - H_{i,j,k}^s\right)}{\Delta t} \quad \text{where}\ \ h_{i,j,k} > H_{i,j,k}^s \tag{9.3-23}$$

2. <u>Source terms</u>, *W*, where terms (a)-(c) are related to Natural processes, and terms (d)-(f) are related anthropogenic systems:

   a. $\boldsymbol{Q_{i,j,k}^P}$ (*positive*)- Percolation of excess soil moisture to shallow groundwater storage and further down to aquifer (WBM models shallow groundwater storage and deep groundwater aquifers as separate entities). Percolation aquifer recharge is controlled by core WBM functions/processes.

   b. $\boldsymbol{Q_{i,j,k}^{RIV}}$ (*positive | negative*)- Groundwater to surface water exchange process. It is replicated from River Package (RIV) of MODFLOW software [*Harbaugh*, 2005] and [*de Graaf et al.*, 2015; *de Graaf et al.*, 2017]. This is described in a dedicated section below.

   c. $\overline{\boldsymbol{Q_{\iota,j,k}^{Spr}}}$ (*negative*)- Natural high volume springs and karst outlets (e.g. Covington-Weaver springs of the Upper Snake River Basin (USRB)). Outflow from these springs is naturally controlled by localized system of high conductance channels in the aquifer which are not modeled in relatively coarse scale WBM domains. So, presently we use a simplified approach of sub-uniformly removal of required volumes of water from the entire aquifer where these springs are located. Sub-uniformly water removal term refers to uniform removal with the exclusion of dry cells and cells with insufficient storage. Water from the latter is still subtracted making them dry, and the remainder is re-distributed among other wet cells.

   d. $\sum_A \boldsymbol{Q_{i,j,k}^A}$ (*negative*)- Abstractions, Q^A, of aquifer water for each water use sector simulated by WBM:
      - *A*=**1** - Irrigation
      - *A*=**2** - Industrial
      - *A*=**3** - Domestic
      - *A*=**4** - Livestock

e. $\sum_{Irr} \boldsymbol{Q_{i,j,k}^{Irr}}$ (*positive*)- Percolation from of irrigation water losses (portion of inefficient irrigation water): return flows and irrigation water losses in the delivery network (canals)

- ***Irr=1*** - Percolation from water delivery networks (canals).
- ***Irr=2*** - Percolation of excess (inefficient) water from irrigation process such as flood or sprinkle application

WBM irrigation process functions control cell-level volumes of inefficient irrigation water percolation to the aquifer.

f. $\sum_{A} \boldsymbol{Q_{i,j,k}^{a}}$ (*positive*)- Percolation, return fluxes, from non-irrigation water abstractions simulated by WBM. These include:

- ***a=1*** - Industrial
- ***a=2*** - Domestic
- ***a=3*** - Livestock

g. $\boldsymbol{Q_{i,j,k}^{MAR}}$ (*positive*)- Forced river sinks and Managed Aquifer Recharge (MAR). For the same reason as described above in (c), the water volume of sinks is uniformly added to all top layer aquifer cells.

The listed above source components can be summarized as the cumulative source term for eq. (9.3-6) used in WBM

$$W_{i,j,k} = Q_{i,j,k}^{P} + Q_{i,j,k}^{RIV} + \overline{Q_{i,j,k}^{Spr}} + \sum_{A} Q_{i,j,k}^{A} + \sum_{Irr} Q_{i,j,k}^{Irr} + \overline{Q_{i,j,k}^{MAR}} \qquad (9.3\text{-}24)$$

where terms on the right-hand side correspond to the source terms list (2) above. Hat above a term indicates fraction of the total aquifer value uniformly distributed across all cells of the aquifer.

3. <u>Special cases</u> include following-
   a. Dry cell problem for horizontal fluxes is treated in a similar manner as in MODFLOW-NWT formulation [*Niswonger et al.*, 2011] given by eq. (9.3-9). There are some difference in the WBM implementation-

- Instead of quadratic smoothing for zeroing of drying cell conductance, WBM uses exponential smoothing used iteratively until negative drying cell, $(i,j,k)^{d}$, storage is eliminated (Fig. 1)

$$C_{(i,j,k)^{d} \pm \frac{1}{2}} \mathrel{*}= k_{exp}, \qquad \text{repeat while } S_{(i,j,k)^{d}}^{\Delta h} < 0 \qquad (9.3\text{-}25)$$

where k_{exp} is exponent smoothing factor which has to be $k_{exp} < 1$.
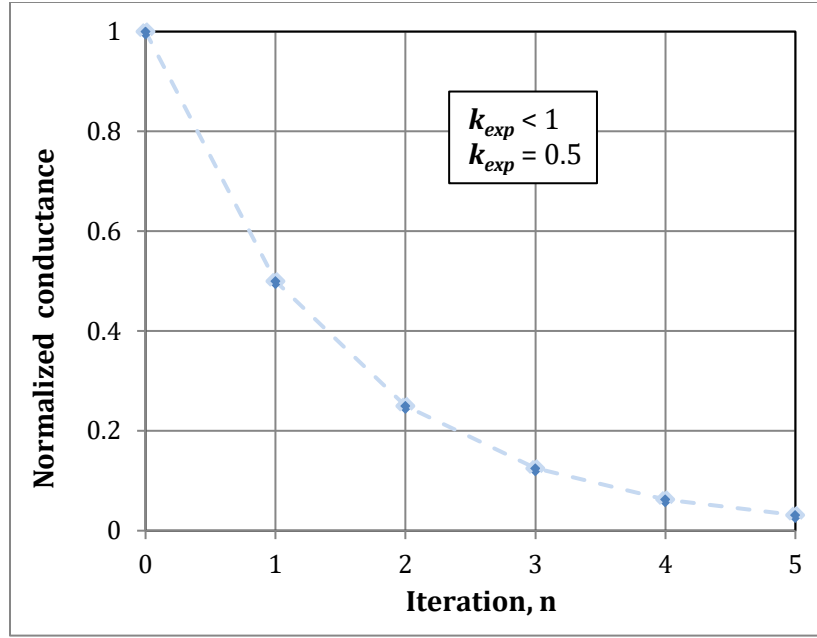
Figure 9-1. Exponential smoothing of drying cell conductance

- Setting k_{exp} to zero shuts conductance from the drying cell as it is done in BCF, LPF, and HUF Packages of MODFLOW [*Niswonger et al.*, 2011].

The advantage of using exponential smoothing approach is that it uses only one control parameter (k_{exp}) instead of three in the quadratic scheme which are, in turn, cell size dependent. The disadvantage is that the exponential smoothing does not result in "true" dry state of the cell and it leaves very small amount of residue water in the cell which is acceptable for WBM simulation purposes.

b. Dewatering of dry cell in vertical fluxes is done same way as in MODFLOW-NWT formulation [*Niswonger et al.*, 2011] given by eq. (9.3-10).

Groundwater to surface water exchange process (RIV submodule)

RIV submodule in WBM is used to simulate groundwater to surface water (streams) fluxes that closely replicates the RIV module of the MODFLOW package. This water flux/exchange process in the context of regional and Global modeling is described in [*de Graaf et al.*, 2015; *de Graaf et al.*, 2017] and implemented in WBM as following-

$$Q^{RIV} = C^{RIV} * min(HRIV - HBOT, HRIV - h) \qquad (9.3-26)$$

which is equivalent to the equation (11) in [*de Graaf et al.*, 2015]. Notations in the above equation are: C^{RIV} is river bed conductance (m²/day); *HRIV* (m) is average river head elevation; *HBOT* (m) is the bottom elevation of the river bed; *h* (m) is groundwater head elevation; and so *HRIV- HBOT* term corresponds to the instantaneous steam depth calculated in WBM using Manning equation [*Dingman*, 2002]. The river bed conductance is calculated as

$$C^{RIV} = \frac{1}{BRES} * P_{chn} * L_{chn} \qquad (9.3-27)$$

where P_{chn} and L_{chn} are stream wetted perimeter approximated by the channel width and channel length correspondingly, and their product makes a river stream area in the grid cell, *BRES* is bed resistance (day⁻

40

[1]) which constitutes inverse percolation rate of a standing water body. For instance, *BRES* value of 10 day[-1] means that 1/10 of the water body depth percolates during one day time period.

In the papers [*de Graaf et al.*, 2015; *de Graaf et al.*, 2017] it was noted that the groundwater head in all water grid cells (e.g. lakes and reservoirs) stays at the lake surface elevation. So this implies that *BRES* in these grid cells has to be set to value 1 day[-1] forcing the groundwater head to meet this condition.

The drain (DRN) package of the MODFLOW is similar to RIV formulation as in equation 9.3- (26) above with the difference that it is applied to small streams with width under 25 m [*de Graaf et al.*, 2017] and it does not allow draining water into the aquifer, if the head is below the stream surface elevation. At this time we do not use it and apply equation (9.3-26) to all streams in the WBM simulated river network.

Production of the river head elevation- It is important to use a good quality land surface elevation data and the river elevation data. The former can be produced from the high resolution DEM (we use composite 30 m SRTM DEM) as the simple or area weighted grid cell average. And the latter can be derived as the minimum elevation of the same high resolution DEM in the STN grid cell. But errors and uncertainties of the source high resolution DEM do never result in a river elevation profile consistently

Using/Running ModFlow module in WBM: Inputs, Controls, and Output

Implementation of described above formulations for 3D groundwater flows source terms and boundary conditions is available in WBM version 18.8.0 and later on mater branch, but not yet available in the public open source branch. The latter has simplified aquifer models in Lumped and Virtual aquifer modules (see other WBM documentation).

The WBM ModFlow module gets turned on by choosing "ModFlow" aquifer type in the "Aquifers" input of the WBM initialization file or DB record of WBM run IDs. Note that only one type of aquifer model can be used in a given WBM run, and so choosing ModFlow type will not allow running Lumped or Virtual aquifers in this simulation.

The list of required (no defaulted input or **DEFAULT** input, if loaded to the WBM installation) and optional (defaulted input) WBM ModFlow inputs is itemizes and explained below-

1. **DEM – (required):** Digital elevation model for the WBM simulation domain (defined by flow direction STN grid) which refers to an average cell elevation above sea level in meters. It is used to define geometry of the aquifer top which is assumed to be at the land surface. **DEFAULT** WBM input layer for DEM is topography dataset of 500 m resolution from NASA Blue Marble project (https://visibleearth.nasa.gov).

2. **STN_DEM – (required):** Digital elevation model for the river network (STN) segments which refers to an average river or stream surface elevation above sea level in meters. This dataset is required for modeling of groundwater to surface water (rivers and lakes) exchange process. Theoretically, it is a minimum elevation in a grid cell since the water is expected to be at the lowest elevations within the grid cell. But due to inconsistencies in high resolution DEMs, the resulting river elevation profiles are always uneven and jagged, and so the STN elevation of a downstream pixel is often some higher than the one from upstream. This problem has been addressed by works of W. Schwanghart with co-authors who developed a TopoToolbox-2 MATLAB software that can perform smoothing of river network elevation profiles using a quantile regression techniques [*Schwanghart and Scherler*, 2014; 2017]. Alternatively, a mathematician from UNU Research Computing Center (RCC), Dr. Robert Carrier, has developed a smoothing algorithm based on methods of linear programming in C for a similar quantile regression approach (unpublished work). Both utility software programs are available at the UNH

WSAG as WBM data pre-processing tools. No **DEFAULT** input is provided for this key.

3. **Aqf_Thick – (required):** Total aquifer thickness of all layers in meters. Note that top of the aquifer is land surface elevation defined by the "DEM" gridded layer. **DEFAULT** WBM input layer for Aqf_Thick is 5 minute resolution dataset from [*de Graaf et al.*, 2017].

4. **Conf_Thick – (optional):** Thickness in meters of confining aquifer (top layer) as illustrated on Fig.3 of [*de Graaf et al.*, 2017]. Defaulted value is zero (no confining aquifer layer). And the **DEFAULT** input is 5 minute resolution dataset from [*de Graaf et al.*, 2017].

5. **WT_DEM_B – (optional):** Initial water table (head) elevation (above sea level in meters) for unconfined→confined aquifer (bottom layer) where it is present. Defaulted value is STN_DEM layer assuming the aquifer head to be in equilibrium with the river level.

6. **WT_DEM_T – (optional):** Initial water table (head) elevation (above sea level in meters) for confining aquifer (top layer) where it is present. Defaulted value is STN_DEM layer assuming the aquifer head to be in equilibrium with the river level.

7. **Aqf_Sy – (optional):** Specific yield (S_y) for unconfined aquifer (unblocked portion of the bottom layer). Defaulted value is 0.1, and the **DEFAULT** input is 5 minute resolution dataset from [*de Graaf et al.*, 2017].

8. **Aqf_SyCnfd – (optional):** Specific yield (S_y) for confining aquifer (top layer where defined). Defaulted value is 0.001.

9. **Aqf_Ph – (optional):** Near surface horizontal hydraulic permeability in Log10 units of m^2. Defaulted value is 13.1, and the **DEFAULT** input is 5 minute resolution dataset from [*de Graaf et al.*, 2017]. Keep in mind that the hydraulic permeability gets converted to horizontal conductivity and then scaled to a given depth using e-folding factor (*f*) as described in the previous sections.

10. **Aqf_Kh_B – (optional):** Horizontal conductivity of confined aquifer (bottom layer). Defaulted value is 10.6 m/day.

11. **Aqf_Kv_B – (optional):** Vertical conductivity of confined aquifer (bottom layer). Defaulted value is 1.6 m/day.

12. **Aqf_KhConf – (optional):** Horizontal conductivity of confining aquifer (top layer). Defaulted value is 0.008 m/day.

13. **Aqf_KvConf– (optional):** Vertical conductivity of confining aquifer (top layer). Defaulted value is 0.0008 m/day.

14. **eFolding– (optional):** E-Folding depth in meters. Defaulted value is 120 m. **DEFAULT** input is 5 minute resolution dataset calculated from elevation roughness following [*Fan et al.*, 2007].

15. **aqfInfiltFr– (optional):** Fraction of surplus water available for infiltration to an aquifer. Defaulted value is 0.5.

16. **BRes– (optional):** Riverbed resistance in days used in RIV sub-module. Defaulted value is 10 days.

17. **SLE– (optional):** Sea level elevation in meters used at exorheic basin outlets. Defaulted value is 0 meters. Note- some STN subsets are clipped for regions that are not actually connected to the World ocean (e.g. Caspian sea tributaries or upstream sections of clipped river networks such as Upper Snake River Basin that clips part of Columbia river just below Twin Falls in Idaho). The SLE elevation is assumed to be constant elevation of the aquifer water table outside of the STN spatial domain.

18. **Aqf_BndFlx– (optional):** Flag to allow groundwater fluxes outside of the STN spatial domain. Defaulted value is 0 (do not allow).

19. **Aqf_Clump– (optional):** Flag to clump virtual/clumped aquifers to a single body. Defaulted value is 0 (do not allow). Note- water component tracking and other aquifer statistics are saved in a set of spreadsheets (one per separate or segmented aquifer). Separate aquifer is defined as a result of 4-component 2D segmentation, and represents a continuously connected aquifer pixels in the STN domain. The 4-componet segmentation refers to side-by-side pixel connectivity only (corner-by-corner connectivity is not included). Tracking water components in a large number of separate aquifers is a computationally intensive task. By default, the tracking and statistics is done for each separate aquifer. And, if such statistics is not needed in the output and model run time is a concern, then setting "Aqf_Clump" flag to 1 will cause clumping all separate aquifers to just one which will result in much faster model execution times.

20. **MF_steps– (optional):** Number of steps per WBM time step, Δt, in the EFDM solver. Defaulted value is 1. Note- end of the year summary of the model log file contains the highest value of the Frederick's stability criteria (FSC) for the EFDM solver or the model quits if FSC becomes larger than 0.5. In order to have EFDM groundwater flow solution to be stable and the FSC to keep reasonable low (the smaller the FSC is the better), a WBM user can set a larger number of "MF_steps" improving the solver stability, but in expense of the model performance.

Validation of ModFlow module in WBM: test runs

A validation of ModFlow module in WBM is done over 6 arc minute continental US subset of STN-06 river network, using MERRA-2 climate drivers, MIRCA-2000 irrigated crop set, simulated aquifer geometries and parameters from [*de Graaf et al.*, 2017]. A 200 yearlong spinup (20 times of 1980 decade) has been used to equilibrate and stabilize groundwater aquifers. Then the groundwater head depth for 5 Corn Belt stated has been checked against observed USGS groundwater gauges using 4 sites per each state evenly distributed around state center point. Due to high spatial variability of simulated aquifer geometries and parameters, a 15 km search distance has been used to find the best match around the USGS gauging station.

The validation below is directed toward matching patterns of variability of groundwater head levels as coarse scale hydrological model data cannot be directly compared to groundwater observations due to (a) differences in coarse scale land topography input and actual elevation of the gauging site; and (b) simulated aquifer geometry and parameters used in the model input which does not represent actual aquifer parameters. But the seasonal and inter-annual variability of the head depth changes can be compared to WBM model outputs as a validation metric. The pattern match requires scaling of the input

data so that the variance of both observed and simulated data should match. We used the following scaling transformation for the input data

$$H_s = S_{factor}\left(H_a - H_a^{avg}\right) + H_a^{avg} + S_{offset}$$ (9.3-27)

where H_a, H_a^{avg}, H_s and are actual, actual mean, and scaled head depth, S_{factor} and S_{offset} are scaling factor and offset are found by minimization the difference between observed and simulated data variability. The interpretation of scale factor is a bias in the sediment specific yield integrated over the aquifer vertical dimension, and the offset is solely reflect local topography difference relative to coarse resolution DEM used in the modeling (6 arc minute resolution which is about 10 km pixel size). Higher scale factor indicate higher specific yield of the actual sediments as comparted to the simulated one, and vice versa, the lover scale indicate lower yield of the gauge wells as expected.
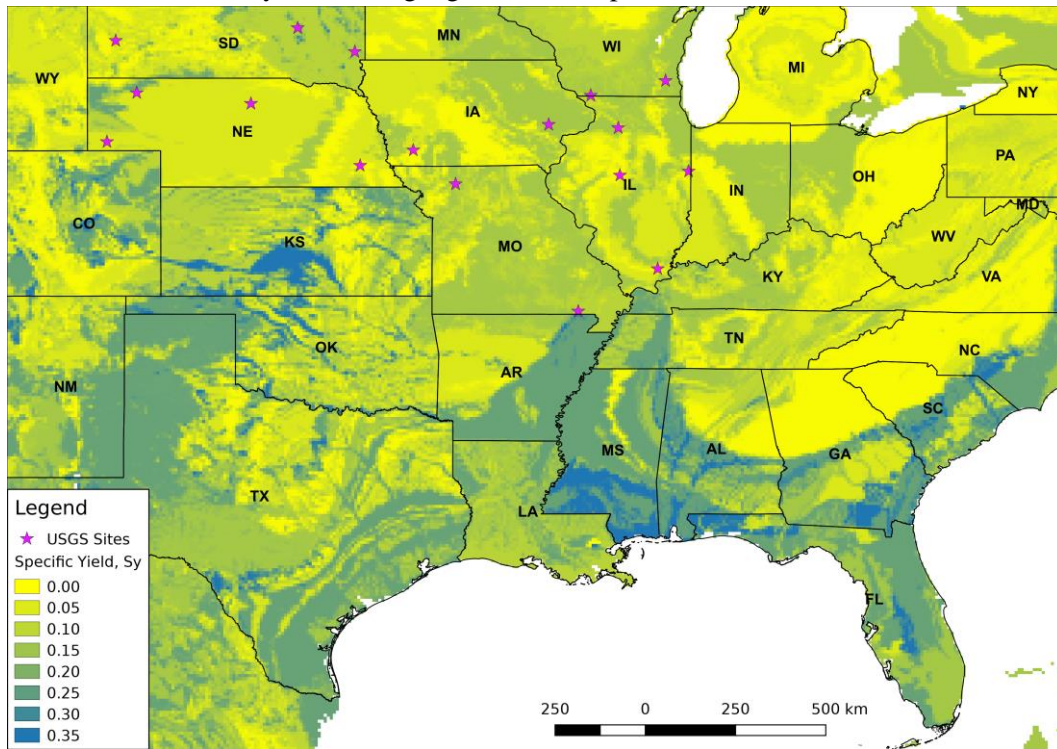

Figure 9-2. Specific yield used in the validation ModFlow-WBM simulation. The data citation is [de Graaf et al., 2017].
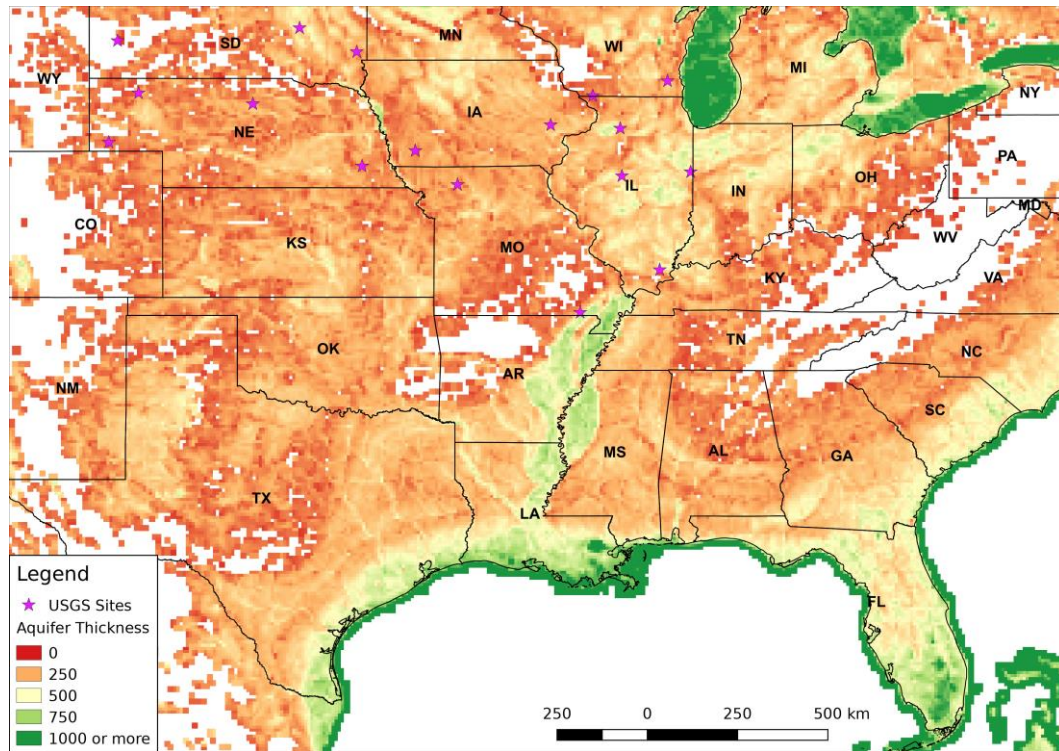
Figure 9-3. Aquifer thickness used in the validation ModFlow-WBM simulation. The data citation is [de Graaf et al., 2017].

The mismatch of the modeled aquifer geometries (Figure 1) and properties (e.g. specific yield, Figure 2) as compared to the "official" USGS aquifer map (Figure 3) is obvious. And the latter can explain the need for the pattern match by eq. (25) instead of direct model comparison to the head depths in USGS gauging wells. Also it can help in understanding of the values of the scale factors and pattern timing shifts as these are discussed for some individual sites presented below.
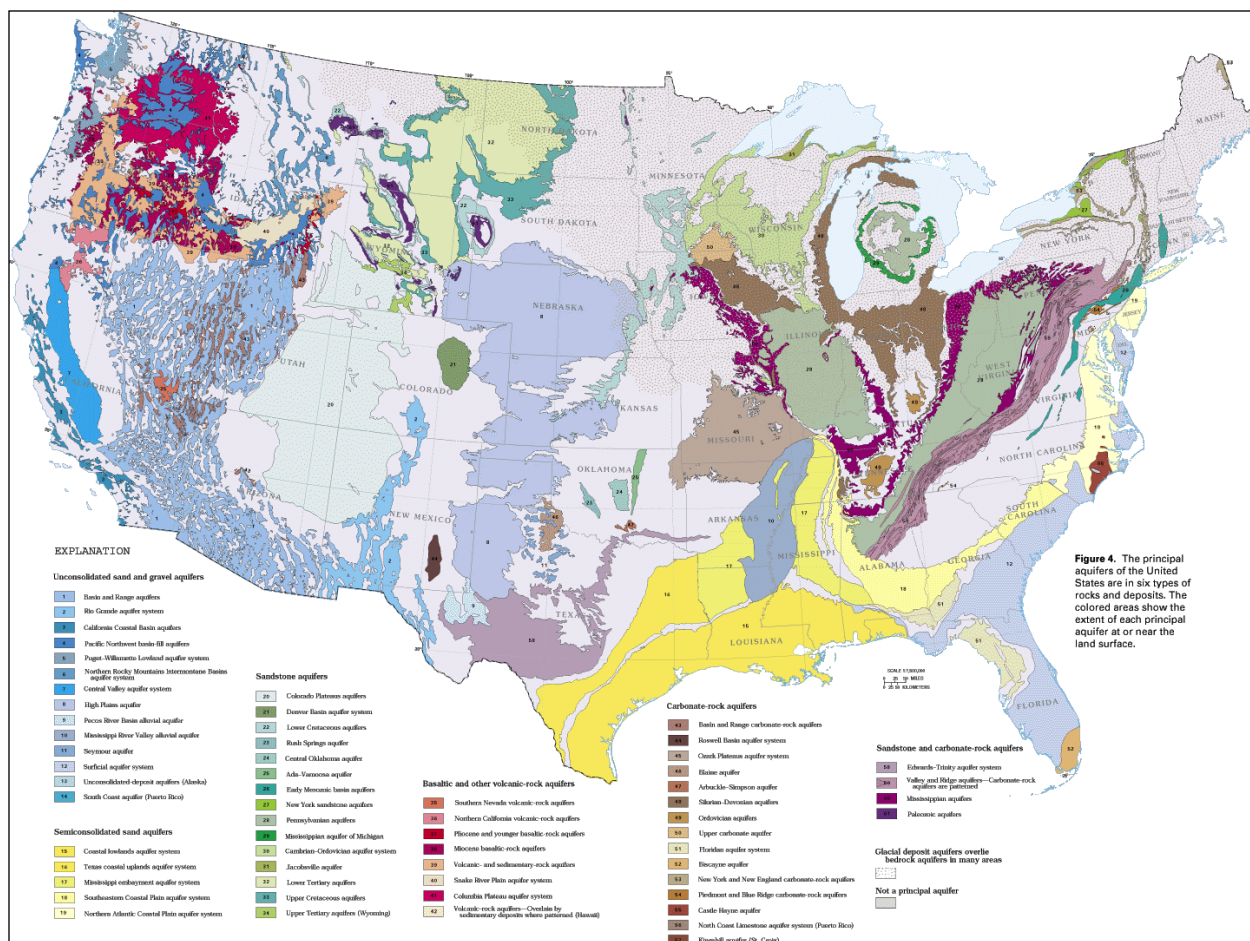
Figure 9-4. USGS map of continental US aquifers.

The validation results are shown and interpreted below grouped by the US states that cover most of the agricultural Corn Belt.

References

Crank, J. (1975), *The mathematics of diffusion*, 2d ed., viii, 414 p. pp., Clarendon Press, Oxford, Eng.

de Graaf, I. E. M., E. H. Sutanudjaja, L. P. H. van Beek, and M. F. P. Bierkens (2015), A high-resolution global-scale groundwater model, *Hydrology and Earth System Sciences*, *19*(2), 823-837.

de Graaf, I. E. M., R. L. P. H. van Beek, T. Gleeson, N. Moosdorf, O. Schmitz, E. H. Sutanudjaja, and M. F. P. Bierkens (2017), A global-scale two-layer transient groundwater model: Development and application to groundwater depletion, *Advances in Water Resources*, *102*, 53-67.

Dingman, S. L. (2002), *Physical hydrology*, 2nd ed., x, 646 p. pp., Prentice Hall, Upper Saddle River, N.J.

Fan, Y., G. Miguez-Macho, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, *Journal of Geophysical Research-Atmospheres*, *112*(D10).

Harbaugh, A. W. (2005), MODFLOW-2005, the U.S. Geological Survey modular groundwater model— The ground-water flow process, *Report Rep.*, 253 pp, USGS, Reston, VA.

Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model - User Guide to Modularization Concepts and the Ground-Water Flow Process, *Report Rep. 2000-92*.

Hartmann, J., and N. Moosdorf (2012), The new global lithological map database GLiM: A representation of rock properties at the Earth surface, *Geochem. Geophys. Geosys.*, *13*(12).

Hughes, J. D., and J. T. White (2013), Use of General Purpose Graphics Processing Units with MODFLOW, *Groundwater*, *51*(6), 833-846.

Ingebritsen, S. E., and C. E. Manning (1999), Geological implications of a permeability-depth curve for the continental crust, *Geology*, *27*(12), 1107-1110.

Langevin, C. D., J. D. Hughes, E. R. Banta, R. G. Niswonger, S. Panday, and A. M. Provost (2017), Documentation for the MODFLOW 6 Groundwater Flow Model, *Report Rep. 6-A55*, Reston, VA.

McDonald, M. G., and A. W. Harbaugh (1988), A modular three-dimensional finite-difference ground-water flow model, *Report Rep. 06-A1*.

Miguez-Macho, G., Y. Fan, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation, *Journal of Geophysical Research-Atmospheres*, *112*(D13).

Niswonger, R. G., S. Panday, and M. Ibaraki (2011), MODFLOW-NWT, A Newton formulation for MODFLOW-2005, *Report Rep. 6-A37*, 44 pp, USGS.

Schwanghart, W., and D. Scherler (2014), Short Communication: TopoToolbox 2-MATLAB-based software for topographic analysis and modeling in Earth surface sciences, *Earth Surface Dynamics*, *2*(1), 1-7.

Schwanghart, W., and D. Scherler (2017), Bumps in river profiles: uncertainty assessment and smoothing using quantile regression techniques, *Earth Surface Dynamics*, *5*(4), 821-839.

10. Glacier melt water

WBM can use output from a glacier dynamics model (e.g., Huss and Hock, 2015; Rounce et al., in press) as an input to WBM. The glacier dynamics model simulates glacier mass balance for all glaciers in the global Randolf Glacier Inventory (RGI Consortium, 2014), and estimates liquid water discharge from each glacier outlet on a monthly basis. We assume daily glacier discharge is constant through each month.

To avoid double-counting precipitation and runoff over the glacier area, each WBM grid cell is assigned a glaciated fraction (0 for non-glacial regions). Precipitation is reduced linearly by this fraction, thereby reducing runoff and effectively removing the glaciated area from the hydrological simulation. We assume the glacier occupies the highest elevation bands within each grid cell. Each glacier has a single designated outlet location; it is from this location that glacial discharge enters the WBM river system. While a single glacial area may intersect multiple river basins, each glacier discharged to only one basin. Glacier melt water, either as a single runoff unit or as multiple components (runoff as ice melt vs precipitation) can be tracked in WBM; see section **Primary Component Tracking** below.

If the glacier simulations provide a time series of glaciated area, WBM has a pre-processing tool that rasterizes this changing glacier area, allowing WBM to allocate land within each grid cell to glaciated vs non-glaciated fractions dynamically over time.

11 Hydro infrastructure
11.1 Reservoirs

Water Release from Controlled Large Reservoirs

Dams and reservoirs are an integral part of simulated river networks. It is a challenge to develop generic mathematical functions for dam operating rules because water release from large reservoirs is controlled by people based on the primary use of water stored with the reservoir. Furthermore, many hydrological factors, such as seasonal variance of water inflow, forecasts of extreme floods or droughts, upstream snow storage, interact with the timing and needs for reservoir storage. Normal operation of individual dams is generally unknown, so models must rely on limited available outflow data, dam locations, and limited physical characteristics of the reservoir's hydro-infrastructures [*Lehner et al.*, 2011]. The goal for WBM is to develop a simple, but still realistic model for dam operating rules through mathematical functions which are based on the minimum possible set of input parameters.

In order to design a mathematical model for managed reservoirs we incorporate critical principles in dam operations. First, dams are constructed for specific use purposes and accordingly optimized for an operational regime that normally corresponds to an average annual flow at the given location. The key considerations in such a design are bathymetry of the reservoir and its potential water storage, average annual river flow over a historical time period, and inflow hydrograph. For water balance modeling of such large managed dams, we assume that optimal operating rule parameters are based on long-term averages of stream flow and maximum capacity of the reservoir. We assume that the optimal water storage must be below its maximum capacity and water release should be maintained at an average annual discharge level as much as water storage allows. On the other hand in cases of high-flow time periods when storage approaches its maximum capacity, the discharge is likely to exponentially increase to prevent overtopping the dam. Two fundamental principles of controlling water release from large dams are considered in our model:

1. **Dam operation at and below optimal capacity** By design, reservoir storage targets maintenance of an average annual flow as long as possible, but should never be below some minimal regulatory flow as effective storage becomes critically depleted. We found that a logarithmic function can reasonably address such a behavior by maintaining average annual flow within a wide range of available water volumes in the reservoir at and below its optimal storage.
2. **Dam operation above optimal capacity** At water levels above optimal reservoir storage, rapidly increasing rates of release are needed to prevent overtopping of the dam. We find that an exponential increase in the water release prevents dam overflow.

Based on these two logical considerations we combine logarithmic and exponential dam operating functions that are quasi-continuously spliced at the optimal designed reservoir storage level (Figures 11.1-1 through 11.1-5). Parameterization of this bi-functional reservoir operating model makes use of the following quantities.

1. *Equilibrium reservoir storage* depends mainly on dam purpose and use. We assume that this storage, S_e, corresponds to designed optimal water level and, thus, the reservoir water release corresponding to this equilibrium level is equal to an average annual discharge for most dams, i.e. $Q = Q_e = Q_{Av}$ at $S = S_e$. Considering continuity of water release functions and the assumption

that discharge is continuously, and positively related to storage $\left(\frac{\partial Q}{\partial S} > 0\right)$, the value of optimal reservoir storage can be used as a splicing point for logarithmic and exponential sections of this bi-functional water release model, i.e. $[S_e, Q_e] = F_{Log} \cap F_{Exp}$.

2. *Minimum allowed reservoir release* is mandated to maintain some flow within a river.
3. *Logarithmic water release function for medium and low storage levels* is parameterized by two scaling parameters to control the curvature and slope of the logarithmic water release function (at $S < S_e$).
4. *Exponential water release function for high storage levels* is also parameterized by two scaling parameters to control the exponential rate at which discharge gets increased as reservoir storage approaches its maximum capacity (at $S \geq S_e$).

Values for the above parameters are selected to simulate operating rules for human-controlled dams specific to each dam's, or each purpose. WBM recognizes 5 purposes for dam operations (Table 11.1-1). Average annual discharge Q_{av} and reservoir maximum storage capacity S_{max} are used in the formulation as reference values for nondimensionalization. The value for Q_{av} in WBM is calculated over past 5 full years of the simulation to alleviate a problem of long-term discharge trends due to climate change in the catchment area and temporary changes in annual flows due to construction of hydro-infrastructure upstream such as new dams, or changes in human water use.

Using the described above assumptions, the model for water release from controlled reservoirs is described by the following transversal function for $Y = \frac{Q}{Q_{av}}$ as a function of $X = \frac{S}{S_{max}}$:

$$
\begin{cases}
F_{Log} \Rightarrow Y = Y_0 + a\, ln(1 + c\, X) & \text{at } X < X_e \\
F_{Exp} \Rightarrow Y = B + b\, (X - X_b)^p & \text{at } X \geq X_e
\end{cases}
\tag{11.1-1}
$$

where variables and constants are all dimensionless, i.e. Y and Y_0 are reservoir release and minimum allowed release normalized by average annual discharge, X and X_e are present and equilibrium water storage normalized by maximum reservoir storage capacity. Coefficients c and p are independent parameters, and a, B, b, and X_b are derived coefficients and an offset parameter to match curve slopes (first derivative) and the F_{log} and F_{exp} meeting (equilibrium) point. The latter should be calculated from condition of $(X_e, Y_e) = F_{Log} \cap F_{Exp}$, i.e. both segments of the model must meet at this point with the same first derivative. Using substitutions $d = X_e - X_b$ and $q = a \frac{c}{(1+c\, X_e)}$ and matching first derivatives at (X_e, Y_e) point we have implicit equation for d:

$$
\frac{d^{p-1}}{d^p - (1 - X_e + d)^p} + \frac{q}{p}\frac{1}{Y_1 - Y_e} = 0
\tag{11.1-2}
$$

where Y_1 is a hypothetical discharge when the reservoir is full ($Y = 1$). After a value for d is found from solving implicit equation (1.2) the values for B and b follows:

$$
a = \frac{Y_e - Y_0}{ln(1 + cX_e)}
\tag{11.1-3}
$$

$$b = \frac{Y_1 - Y_e}{(1 - X_e + d)^p - d^p} \qquad (11.1\text{-}4)$$

$$B = X_e - b\, d^p \qquad (11.1\text{-}5)$$

Reservoirs with low regulatory capacity (R_c), the ratio between annual mean flow and the reservoir maximum capacity, below 0.1, which equates to a capacity of about 1 month of average annual flow, cannot be adequately replicated by this model. For dams with R_c less than 0.1, variance in seasonal hydrology results in water release to similar to inflow during most of the year, meaning reservoir effective storage is low during dry periods or completely full during high discharge seasons. Models for water release from uncontrolled dams (Section # 2) can be used instead for reservoirs with low R_c.

Parameterization of the controlled reservoir operating model by dam purpose

The formulation for large, controlled reservoirs permits unique parameterizations that follow common flow and supply regimes. Most of large dams are built to serve one or more purposes in using and controlling water resources [*Lehner et al.*, 2011]. Selection of 6 parameters controls the operational behavior of controlled dams in WBM. Each dam input to WBM is identified with a specific purpose (if no purpose is given, it is simulated as an uncontrolled dam). Parameters controlling dam operation are specified by purpose, and/or by individual dams allowing the user to select typical operational parameters for entire classes of dams, or specifying unique parameterizations for dams where more detailed data is available. Default values for each of the major classes of operation recognized by WBM are presented in Table 11-1, and discussed in the following paragraphs.

Table 11-1. Suggested parameters for reservoir operating model by dam use

Dam Purpose	Y_0	Y_1	X_e	Y_e	C	p	Comment
Generic	0.2	5	0.80	1.0	4	6	Works for most of dams
Flood control	0.2	5	0.20	1.0	100	170	Low optimal storage
Hydroelectric	0.2	5	0.85	1.0	200	6	High storage, uniform discharge
Irrigation(LRO)*	0.1	5	0.70	0.1	1	3	Filling operations, off-season
Irrigation(HRO)*	0.2	5	0.85	1.0	200	6	Release operations, irrigation season
Water supply	0.1	5	0.70	0.1	1	6	High storage, min discharge

* Irrigation dam parameters vary throughout the between low release operations (LRO) and high release operations (HRO). See "*Irrigation*" section below.

Generic- Many dams have multiple uses combining self-exclusive requirements or those are not reported in the available databases of dam inventories [*Lehner et al.*, 2011]. For instance hydropower generation and flood control may conflict in the optimal water level in the reservoir storage. In these cases we can suggest using a "generic" type of dam use with some average values for the model parameters (Figure 1.1).
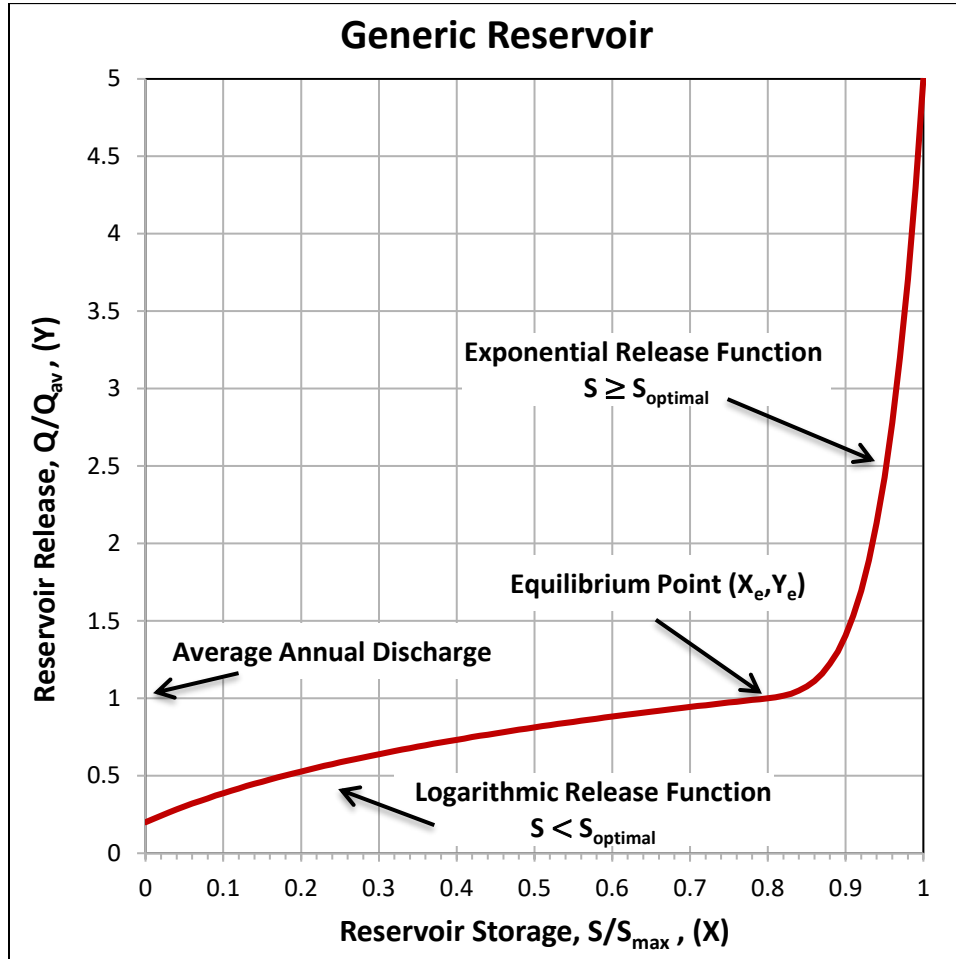
Figure 11-1. Relation between reservoir water release and storage for a generic dams use.

Flood control- These reservoirs are supposed to maintain low water storage so that a reserve of their capacity would be available to accommodate as much water as possible during upstream flood events. The behavior of flood control dams is simulated with a very low optimal storage level (20 %) and increased water release when accumulation of water exceeds it. X_e parameter has to be low in this case (Figure 1.2).
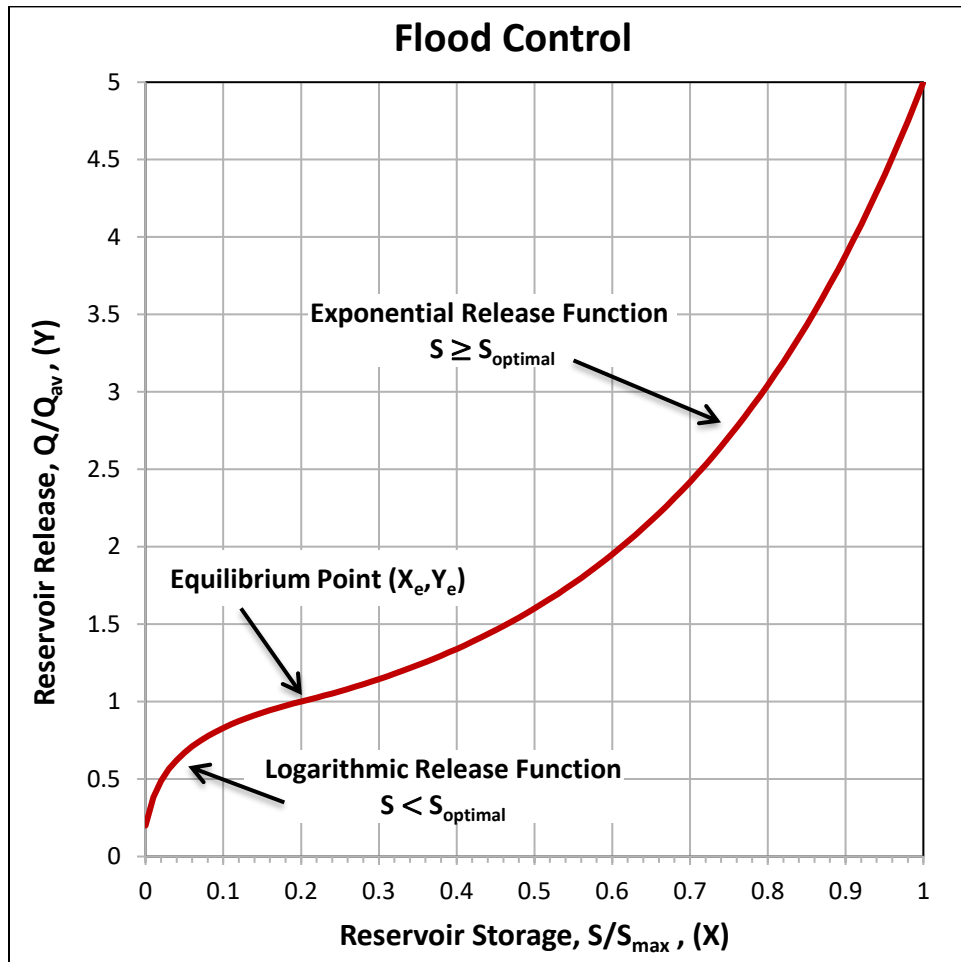

Figure 11-2. Reservoir release curve for flood control dams. See parameters in Table 11-1.

Hydroelectric- Gravitational potential energy of released water needs to be maximized. A high optimal level of water storage (e.g. 90 %) with minimal margin for the cases of seasonal high inflow into the reservoir (e.g. spring snow melt or monsoon season). At the same time during low reservoir refill periods (e.g. dry season) the outflow discharge needs to be maintained at a uniform value to continue production of electricity. This can be modeled by high values for *c* parameter (Figure 1.3).
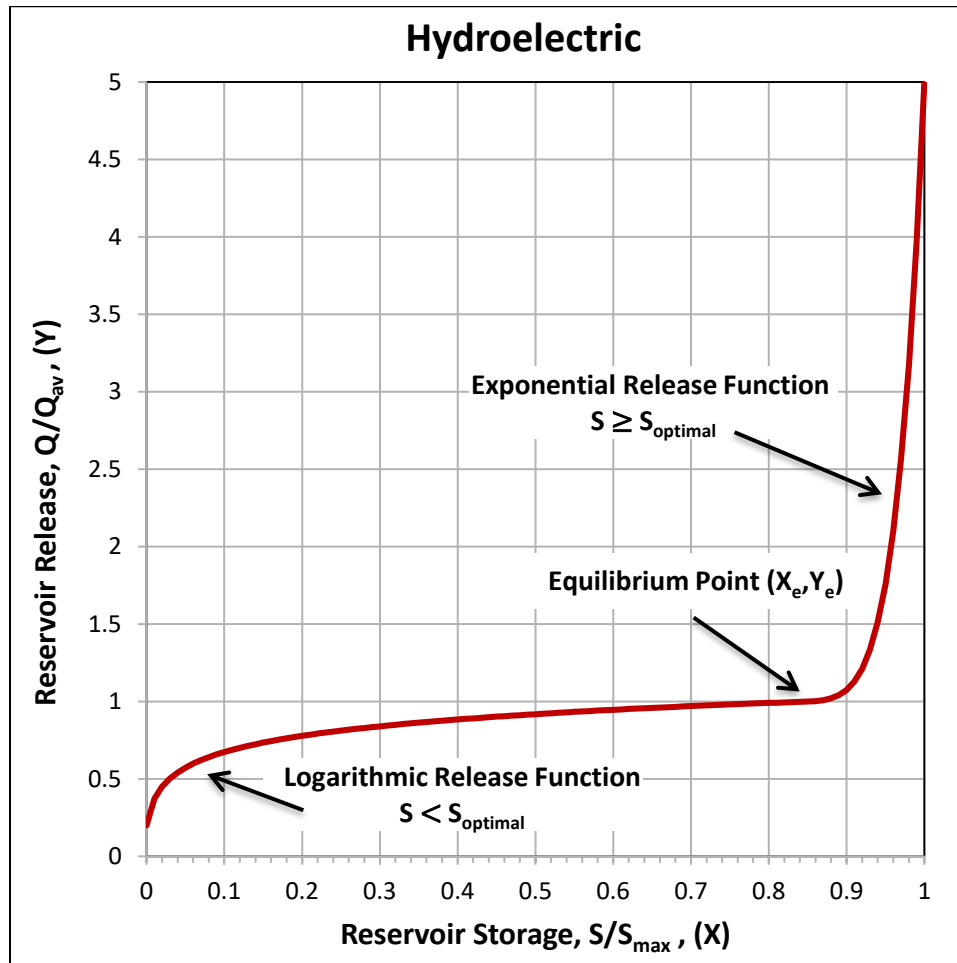

Figure 11-3. Reservoir release curve for hydroelectric dam use. See parameters in Table 11-1.

Irrigation- These reservoirs maximize utilization of the reservoir storage for irrigation by maintaining high water storage in the reservoir outside of the irrigation season and high water release when during the irrigation season (assumes provisioning to downstream irrigation). This is achieved by adapting the water release curve to local irrigation demand (Figure 11-4). Long term daily averages of irrigation demand frequency is input to WBM as the daily probability density function of annual irrigation demand. We use linear interpolation of water release curves between low release operations (LRO) during days with no irrigation demand to high release operation (HRO) during days with maximum irrigation demand. The linear interpolation is done as following:

$$[M] = [M^{LRO}] + ([M^{HRO}] - [M^{LRO}]) \frac{F_{irr}}{F_{irr}^{max}} \qquad (11.1\text{-}6)$$

where vector $[M] = [Y_0, Y_1, X_e, Y_e, c, p]$ with superscripts *LRO* and *HRO* referring to low and high release operation water release curve/regime, F_{irr} and F_{irr}^{max} are daily irrigation frequency and its typical maximum value correspondingly. We suggest using value of 0.05 for F_{irr}^{max}, but it should not be higher than 0.075 for stability reasons.
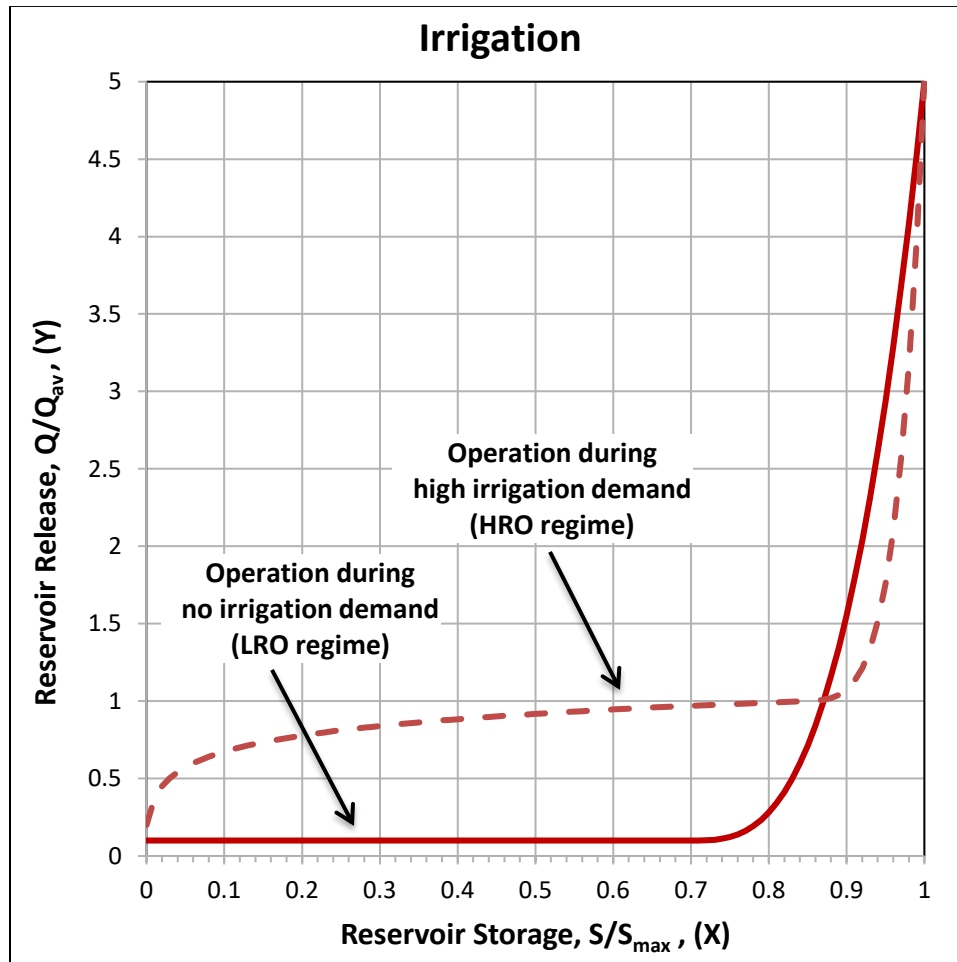
Figure 11-4. Reservoir release curve for irrigation dam purposes. See parameters in Table 11-1 and eq. (11.1-6).

Water supply- These reservoirs are built with intent to maximize utilization of inflow water by minimizing outflow. This would result in high water storage in the reservoir which can be withdrawn agricultural/irrigation, industrial, and domestic use directly from the reservoir. Low values for Y_e parameter can simulate such type of dam use (Figure 1.3).
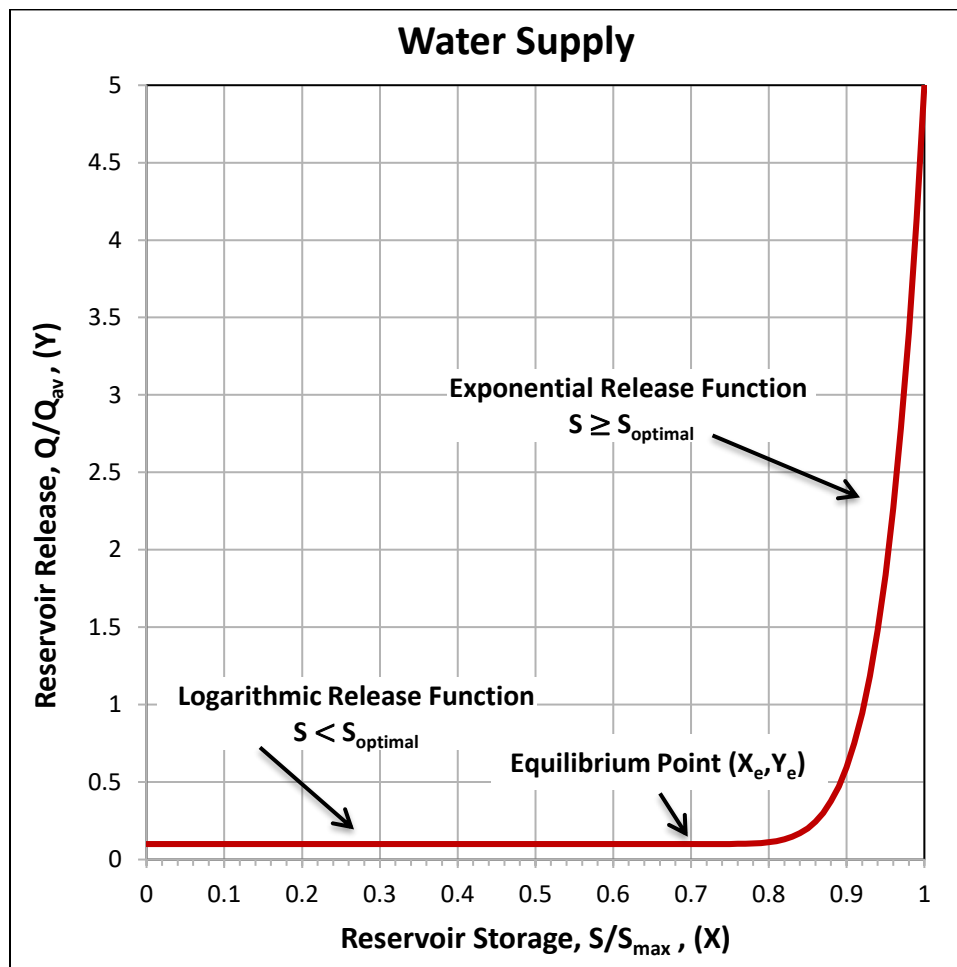

Figure 11-5. Reservoir release curve for water supply dam purposes. See parameters in Table 11-1.

Reservoir capacity input into WBM is assumed to be effective maximum storage that excludes non-effective (dead) storage. Initialization of reservoir storages is done to 100 % of their reported capacity. Each class of dam is input with a minimum and maximum capacity. If a dam is input with a specific purpose, but is less than the minimum capacity, it is represented as an uncontrolled small reservoir. Similarly, if a dam is identified as an uncontrolled reservoir, but exceeds the maximum capacity for small uncontrolled reservoirs, it is simulated as a generic large, controlled reservoir. The default for simulating reservoirs as small, uncontrolled reservoirs is 1 km^3 of capacity (1000 mcm).

11.2 Water Release from Uncontrolled Small Reservoirs

Small reservoirs are usually uncontrolled or rarely (seasonally) controlled by operating personnel. By design those are mostly spillway overflow flood control and small volume storage dams where effective length of crest (gate width) often matches or close to natural river stream width during its average annual flow [*United States. Bureau of Reclamation.*, 1987]. The crest of spillways is commonly ogee shaped and a discharge over them is given by the Rehbock equation [*Khatsuria*, 2005]:

$$Q = C_D \; \frac{2}{3} \; L \; \sqrt{2g \; H_e{}^3} \qquad\qquad (11.2\text{-}1)$$

where Q is reservoir release (discharge), L is effective length of dam crest, g is gravity acceleration, and H_e is water head on the crest. Coefficient C_D depends on water approach velocity and head to dam weir (height) ratio. For relatively deep dams and slow water approach velocities it takes value of $\pi / (\pi + 1) \approx 0.611$ as derived from potential flow theory [*Khatsuria*, 2005]. So substitution of constants in metric units into equation (11.2-1) yields a log-linear form:

$$\log Q = \frac{3}{2}\log H_e + \; \log(1.804 \; L) \qquad\qquad (11.2\text{-}2)$$

Head of the crest $H_e = \frac{S_e}{A_r}$ is a function of reservoir area, A_r, and effective storage above crest, S_e.

Considering very small regulatory capacity of small reservoirs, inflow discharge cannot be removed from daily time series calculations, and reservoir water balance takes form of first-order nonlinear ordinary differential equation:

$$\frac{dS}{dt} = \; Q_{in} - k \; S_e^{\frac{3}{2}} \qquad\qquad (11.2\text{-}3)$$

where dimensional constant $k = C_D \; \frac{2}{3} \; L \; \sqrt{\frac{2g}{A_r{}^3}}$. WBM utilizes a solution to equation 11.2-3 demonstrated by the US Army Corps of Engineers in a technical document HDC-111-3/3 [*United States. US Army Corps of Engineers.*, 1987] where an empirical relation has been obtained from measurements over ten varying spillway design structures:

$$Q = Q_d \left(\frac{H_e}{H_d}\right)^{1.6} \qquad\qquad (11.2\text{-}5)$$

where subscript d refers to dam designed quantities which we assume is equivalent to long term annual averages from WBM. From (11.2-5) we can suggest that spillway dams have effective storage as a function of reservoir surface area and head height:

$$S_e = H_e A = H_e A_0 (1 + \alpha \; H_e) \qquad\qquad (11.2\text{-}6)$$

where α is reservoir flood area rate (m^{-1}), and A_0 is the reservoir area at crest level. Equation (11.2-5) and (11.2-6) can be combined yielding:

$$
\begin{cases}
Q = Q_d \left(\dfrac{\sqrt{1+\beta S_e}\, -1}{\sqrt{1+\beta S_d}-1} \right)^{1.6} & \text{for } \alpha \geq 0 \\[4mm]
Q = Q_d \left(\dfrac{S_e}{S_d} \right)^{1.6} & \text{for } \alpha = 0
\end{cases}
\tag{11.2-7}
$$

where $\beta = \frac{4\alpha}{A_0}$. Equation (11.2-7) is used in WBM. The flood area rate α depends on the reservoir size and geographic properties of the watershed. For small reservoirs with spillway dams it is likely to be in the range of 0.2 to 0.4 m^{-1}, e.g. the reservoir area increases by about 1/3 with 1 m of its stage rise. But the flood area rate is likely to be very small ($\alpha \approx 0$) for any reservoirs with an artificial abutment (e.g. concrete, earth, stone, etc.). A value of 0.3 is assumed as a default in WBM.
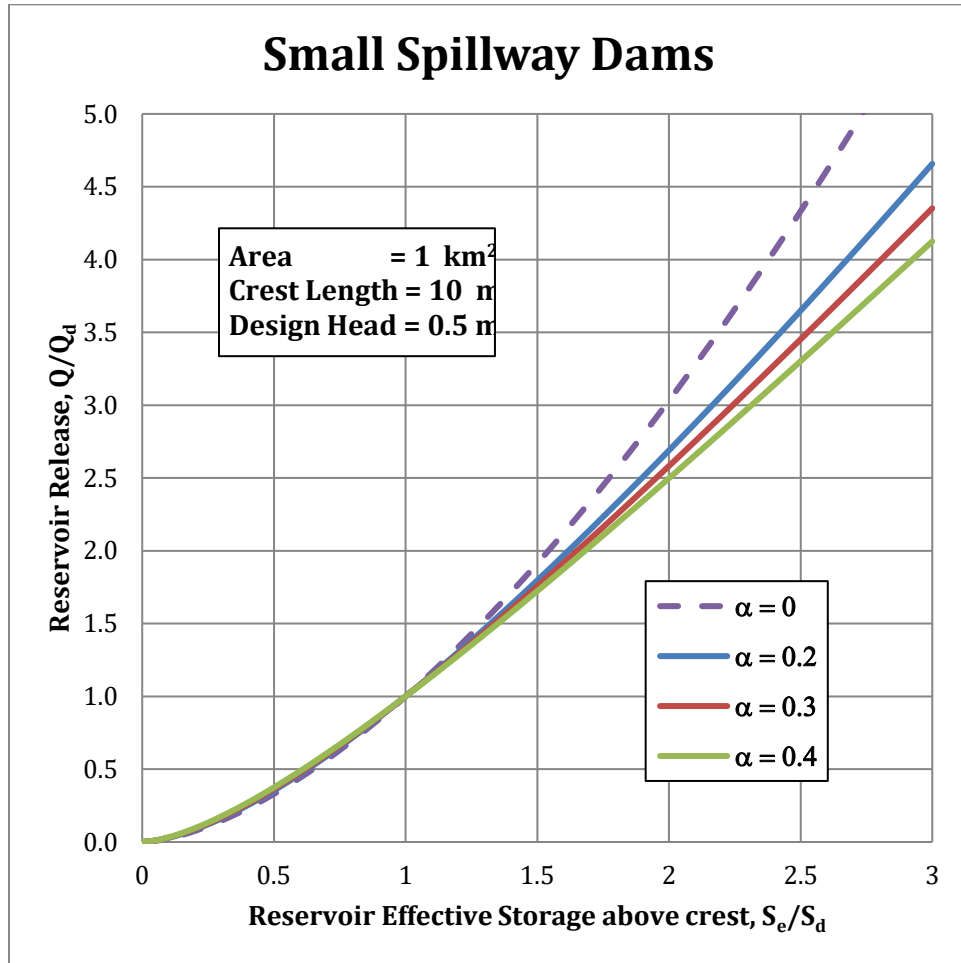

Figure 11-6. Discharge from spillway dams by equation (11.2-7).

References

Khatsuria, R. M. (2005), *Hydraulics of spillways and energy dissipators*, xx, 649 p. pp., Marcel Dekker, New York.

Lehner, B., et al. (2011), High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, *Frontiers in Ecology and the Environment*, *9*(9), 494-502.

United States. Bureau of Reclamation. (1987), *Design of small dams*, 3rd ed., xliii, 860 p. pp., U.S. Dept. of the Interior For sale by the Supt. of Docs., U.S. G.P.O., Washington, D.C.?

United States. US Army Corps of Engineers. (1987), Spillways*Rep.*, 100-111 pp, Coastal and Hydraulics Laboratory, Vicksburg, Mississippi.

12. Inter-basin Transfers

A global database of inter-basin transfers has been developed by Richard B Lammers, and used in Zaveri et al (2016) and Liu et al (2017):

WBM simulates transfers of water between hydrologic basins by moving water across basin divides from one river location to another. We simulate both existing inter-basin transfers - transfers with infrastructure that was completed prior to 2006 – and future potential transfers. Future potential transfers were determined by literature review of government and NGO proposals. For all inter-basin transfers (completed and proposed), five parameters are used to simulate the transfer. These are: the donor/from latitude and longitude, the recipient/to latitude and longitude, a minimum allowed flow, a maximum allowed flow, and a rule for flow volumes between the minimum and. In some cases, maximum allowed flow is based on published reported annual transfer capacities. In addition to the reported latitudes and longitudes of the transfers, we grid cell based locations for each transfer, which in some cases are different than the reported location because they were adjusted to ensure they linked to the correct rivers within the STN-30p network version 6.02. The completed transfers are implemented in the year that construction was completed; proposed transfers are turned on at their proposed completion date, as there is no set date for completing construction of these transfers.

The volume of water transferred through each canal is calculated as:

$$D = \begin{cases} 0 & if\ Q_d \leq Q_{min} \\ (Q_d - Q_{min}) * \frac{P}{100} & if\ Q_{min} > Q_d \geq Q_{max} \\ Q_{max} & if\ Q_d > Q_{max} \end{cases} \tag{12-1}$$

where D [m³s⁻¹] is the amount of water diverted through the canal, Q_d [m³s⁻¹] is the donor river discharge, Q_{min} [m³s⁻¹] is the minimum flow parameter, Q_{max} [m³s⁻¹] is the maximum flow parameter, and P is the percent flow parameter.

The transfer volume, D, is corrected to D_{corr} for small transfer volumes:

$$D_{corr} = 0\ if\ D < 0.01 \tag{12-2}$$

Evaporation from open water along the canals is removed from the transfer volume:

$$D_{corr_e} = \begin{cases} D_{corr} - E\ if\ (D_{corr} - E) > 0.001 \\ 0 & if\ (D_{corr} - E) \leq 0.001 \end{cases} \tag{12-3}$$

where D_{corr_e} [m³s⁻¹] is the transfer volume corrected for evaporation, and E [m³s⁻¹] is the evaporation volume:

$$E = L * W * FWE \tag{12-4}$$

where L [m] is the length of the canal (listed in Table S8 where published data is available, or calculated based on a straight line between to/from points), FWE is free-water evaporation [mm/day] which can be calculated through various free-water evaporation models or by scaled calculated potential evapotranspiration by the Hamon method; and W [m] is the width of the canal:

$$W = \begin{cases} \tau * D_{corr}^{\varphi}\ if\ (\tau * D_{corr}^{\varphi}) \geq 0.01 \\ 0 & if\ (\tau * D_{corr}^{\varphi}) < 0.01 \end{cases} \tag{12-5}$$

where τ (8.0) and φ (0.58) are held constant (Park, 1977).

Water is transferred on a daily time step. Several of the lengthy inter-basin transfers were split into multiple transfer segments for the purpose of the simulation. This allowed for water to be released and/or stored along the canal route, from where it can be accessible for irrigation withdrawals.

References:

Park C 1977 World-wide variations in hydraulic geometry exponents of stream channels – Analysis and some observations *J Hydrol 33* 133-146

Zaveri E., Grogan D.S., Fisher-Vanden K., Frolking S., Lammers R.B., Wrenn D.H., Prusevich A., Nicholas R.E. Invisible water, visible impact: Groundwater use in Indian agriculture under climate change. *Environ. Res. Lett.* **11** (2016) doi:10.1088/1748-9326/11/8/084005

Water extractions

13. Irrigation
13.1 Irrigation water demand

Definitions:
Net irrigation water demand is the amount of water required by crops to achieve the crops' potential evapotranspiration. In addition, net irrigation water demand includes the amount of water required to maintain flood levels within rice paddies. Inefficiencies in the water delivery and application systems are not included.

Gross irrigation water demand is the amount of water required to meet net irrigation demand, plus the water lost through inefficiencies in water delivery and application.

Net irrigation water is the amount of irrigation water used by crops, not including losses due to inefficiencies. This water volume is less than net irrigation water demand when the demand is not completely fulfilled.

Gross irrigation water is the amount of irrigation water used by crops, *including* losses due to inefficiencies. This water volume is less than gross irrigation water demand when the demand is not completely fulfilled.
In WBM, crops extract water from the soil moisture pool each day of the crop's growing season. Given sufficient water in the soil moisture pool, the amount of water used by each crop is the crop potential evapotranspiration, $PETc$ [mm]:

$$PET_c = k_c \cdot PET_0$$
(13.1-1)

where PET_0 [mm] is a reference evapotranspiration, and k_c [-] is a crop-specific, time-varying scalar. This method follows the FAO-recommended crop-modeling methodology outlined in Allen et al (1998). Here, we use the Penman-Monteith method for estimating PET_0 (Allen et al, 1998).

If soil moisture levels fall below a crop-specific threshold, SMT_c [mm], then irrigation water is called for. Soil moisture threshold SMT_c for crop c is:

$$SMT_c = CDF_c \cdot RD_c \cdot AW_{cap}$$
(13.1-2)

where CDF_c [-] is a crop depletion factor, RD_c [mm] is the crop's root depth, and $AWcap$ [-] is the soil's available water capacity.

When soil moisture is below SMT_c, then the time step's net irrigation water demand, $I_{net,t}$, is the difference between the current soil moisture and field capacity:

$$I_{net,t} = \begin{cases} Fcap - SM_t \ if \ SM_t \leq SMT_c \\ 0 \ if \ SM_T > SMT_c \end{cases}$$

$$(13.1\text{-}3)$$

where $Fcap$ [mm] is the soil's field capacity, and SM_t [mm] is the soil moisture at time t. Annual net irrigation water demand is the sum of all daily net irrigation water demands through the year.

Alternative irrigation water demand method:
Instead of using the crop-specific soil moisture threshold, WBM can be set to a "daily irrigation" mode, in which irrigation water demand, $I_{net,t}$, is equal to the difference between soil moisture content and field capacity each day:

$$I_{net,t} = Fcap - SM_t$$

$$(13.1\text{-}4)$$

This demand causes water to be extracted from water sources each day. However, this water is then stored in a "virtual" storage pool until the soil moisture reaches the crop-specific soil moisture threshold SMT_c; then water is moved from the virtual storage to the soil moisture pool. *This option was developed to solve the problem of requiring large amounts of water on the same day. The daily method spreads the demand out.*

For a given irrigation system efficiency, I_{eff} [-], gross irrigation water demand, I_{gross} [mm] , is:

$$Igross = \frac{I_{net}}{I_{eff}}$$

$$(13.1\text{-}5)$$

where $I_{eff} \in (0,1)$.

$$(13.1\text{-}6)$$

Gross irrigation water demand is calculated differently when process-based irrigation systems are represented. See the section Irrigation Technology Method for the explanation.

Default parameter values:
Default values for k_c, CDF, and RD for 26 different crop categories are from Siebert and Döll (2010).

13.2 Irrigation water extraction

13.2.1 Irrigation efficiency method

In the irrigation efficiency method, water is extracted for irrigation to meet the gross irrigation water requirement, $I_{gross,}$ described in section **Irrigation Water Demand**. There are several options for (a) from where to take water, and (b) how much water to take.

Water sources:
There are 6 categories of water sources in WBM:
1. Surface water: this includes water stored in the river network and water in reservoirs. Surface water can be abstracted from the local pixel as well as neighboring pixels.
2. Small irrigation reservoirs (aka farm ponds): this is an optional parameterization for WBM.
3. Shallow groundwater: this is the water in the shallow groundwater pool; it is typically considered a "sustainable" water source.
4. Unsustainable groundwater: this is an unlimited source of water that is not simulated directly within WBM, i.e., there is no accounting of the volume of water in this imaginary pool. Rather, when water is needed in excess of surface and groundwater supplies, additional water can be drawn from this unlimited pool and added to the soil or other WBM water stock.
5. Aquifer water: this is water in the lumped aquifer pool, which replaces unsustainable groundwater in pixels where lumped aquifers are simulated.
6. MODFLOW aquifer water: this is water in the gridded aquifer field simulated by the MODFLOW WBM module, and is substituted for unsustainable groundwater where distributed aquifers are simulated.

For simulations using lumped (5) or distributed (6) aquifers underlying only part of the spatial domain, unsustainable groundwater (4) can be used outside of defined aquifers.

WBM implements a "search distance" for water when extraction is called for, allowing a given grid cell to search and access surface water from other grid cells within that distance representing canal networks common in regions with irrigated agriculture and dense anthropogenic uses. The default search distance is 100 km; this parameter can be adjusted in the input file and can be different for each water demand category (irrigation vs livestock, domestic, and industrial water demands).

If no priority order or target ratio between water sources is given, then by default WBM will extract water in this order:
1. Small irrigation reservoirs (if simulated)
2. Shallow groundwater within the grid cell
3. River storage within the grid cell
4. River storage from largest river within the search distance
5. Unsustainable groundwater, or aquifer water

The priority order between within-grid-cell shallow groundwater and river storage (steps 2 and 3) can be changed in the input file.

Alternatively, a target ratio of extraction between surface water and groundwater (sw:gw ratio) can be provided. In this case, the order of extraction is:
1. Small irrigation reservoirs (if simulated)
2. Shallow groundwater within the grid cell, with an upper limit of the target amount of groundwater to extract based on the input sw:gw ratio.
3. River storage within the grid cell, with an upper limit of the target amount of groundwater to extract based on the input sw:gw ratio.
4. If the irrigation water demand has not been fulfilled, take additional water from the within-grid-cell shallow groundwater pool (in excess of target sw:gw ratio).
5. River storage from largest river within the search distance
6. Unsustainable groundwater, or aquifer water

This order attempts to balance achieving the target sw:gw ratio while only resorting to unsustainable water sources once all sustainable sources have been exhausted.

Water extraction from rivers cannot exceed a specified fraction of the river storage + flow volume; this specified fraction is 80% of river storage + flow, and will be user defined in future versions of the model.

As an optional parameter, a limit can be placed on how much unsustainable groundwater to extract (range: 0 to 1). This parameter scales the unsustainable groundwater extraction by the value given; e.g., if 1 unit of unsustainable water is called for and the parameter is 0.5, then only 0.5 units are extracted.

13.2.2 Irrigation technology method

Irrigation technology in the UNH Water Balance Model (WBM) is a process-based alternative to the prior conceptual formulation where non-beneficial fates were specified as a fraction of gross irrigation (Grogan et al., 2017; Wisser et al., 2010, 2008). The process-based formulation redistributes inefficient irrigation water via surface runoff flows, groundwater percolation, and evaporation during both delivery and application stages. The system explicitly represented non-consumptive losses using technology specific parameters applied to proportions of irrigated croplands operating each technology. Losses during delivery were calculated from conveyance surface area (as a fraction of irrigated cropland), daily open water evaporation, and percolation. Conveyance methods included pipes with no evaporation or percolation, and open conveyances such as canals and ditches that percolate at a fraction of local infiltration rates and evaporate from their surfaces. Incidental losses during application follow Jägermeyr et al. (2015) and use the distribution uniformity parameter that described excess water needed to satisfy net irrigation demand based on the type of technology, either drip, sprinkler, or flood. The distribution uniformity parameter defaults to values originally estimated for surface/flood, sprinkler, and direct/drip agriculture (Jägermeyr et al., 2015).

The process of calculating non-beneficial use (N) and non-consumptive returns (L) via application of irrigation water is performed throughout the WBM time-step cycle. Following calculation of net crop water demand (I_{net}), additional delivery and application requirements are calculated accounting for technology specific inefficiencies. Then, an initial estimate of delivered water is based on estimated water availability and if available water is determined to be insufficient to meet demand (plus inefficient use and loss), all associated irrigation fluxes are scaled downward linearly by the provisional irrigation supply factor (X_{irr}). At this stage, WBM performs the river routing calculation, and estimates of provided water are updated according to actual water availability. Finally, excess water introduced to irrigated crop fields is partitioned between non-beneficial evaporation, non-consumptive runoff, and non-consumptive percolation. What follows is a more detailed description of each of these steps. Unless specified otherwise, all calculations described in this section are distributed spatially across irrigated crop areas as grid operations.

WBM can run any number of individual technologies simultaneously using data of irrigated land fraction for which each of the technologies is used

$$\begin{cases} \sum_i f_i^{d,irr} = 1 \\ \sum_i f_i^{a,irr} = 1 \end{cases}$$

(13.2.2-1)

where $f_i^{d,irr}$ and $f_i^{a,irr}$ are fraction of land served by technology i within irrigated land, and superscripts d and a denotes delivery and application technology group, respectively.

Irrigation Delivery

Inefficient fluxes from conveyances rely on calculated daily open water evaporation rates (function of air temperature, humidity, and wind speed), and percolation rates of saturated soil. These rates are spatially and temporally distributed to the fraction of surface area of the irrigation delivery system ($f_i^{d,A}$) relative of the irrigated area (A^{irr}, m²) for each i delivery technology. These non-beneficial fluxes are calculated at each pixel on each day crops demand irrigation water. Crop water demand functionality of WBM is

described by Grogan *et al.* (2017). We assume that there is no surface runoff from any irrigation water delivery technology.

Evaporation of delivery water (N_{evap}^d) is calculated for days when irrigation demand is required as

$$N_{evap}^d = A_{fw} E_{fw} \qquad (13.2.2\text{-}2)$$

where E_{fw} is evaporation rate from free water surface (m/d), and A_{fw} is a weighted calculation of the pixel area undergoing free water evaporation through irrigation delivery systems:

$$A_{fw} = A^{irr} \sum_i^n f_i^{d,irr} f_i^{d,A} \varepsilon_i^{evap} \qquad (13.2.2\text{-}3)$$

where $f_i^{d,A}$ (-) is the fraction of area relative to irrigated area that irrigation delivery systems occupy on the ground, and ε_i (-) is a parameter that describes the fraction of an irrigation delivery technology that experiences free-surface evaporation. For the ε_i^{evap} parameter we suggest using values approaching 1.0 for ditch and canals (because both have water surface exposed for evaporation), and approaching 0.0 for pipe delivery technology as the only water exposed to air for evaporation in pipes consists of pipe leakage. All parameters can be spatially explicit.

Percolation is calculated from unlined irrigation conveyance (canal or ditch) benthic surface in a method similar to the calculation for evaporation.

$$L_{perc}^d = A_{perc} P_{perc} \qquad (13.2.2\text{-}4)$$

where P_{perc} is percolation rate from the base of an irrigation delivery system to saturated soil, and A_{perc} is a weighted calculation of the pixel area undergoing saturated canal percolation under irrigation delivery systems:

$$A_{perc} = A^{irr} \sum_i^n f_i^{d,irr} f_i^{d,A} \varepsilon_i^{perc} \qquad (13.2.2\text{-}5)$$

where ε_i^{perc} fraction of canal area to which percolation is applied by technology i. For the ε_i^{perc} parameter we suggest using 1.0 for ditch (no lining at the bottom of the ditch), a value representing the fraction of canal bottom areas in the domain that are un-lined (e.g. ~ 1 for canals assuming 100 % of bottom area are exposed to percolation), and zero for pipe delivery technology as its water is isolation from percolation in pipes.

Both N_{evap}^d and L_{perc}^d are scaled by the actual supply factor (X_{irr}). It should be noted that L_{perc}^a is introduced to the model at the location of the irrigated fields and not explicitly at the locations of canals. Furthermore, water that percolates beneath canals is considered a non-consumptive loss associated with irrigated agriculture.

Irrigation Application

Process-based modelling of irrigation water losses by application technology is implemented following an approach similar to Jägermeyr *et al.* (2015). Differences between the two approaches reflect additional processes introduced here, as well as accommodating unique structures of the two hydrologic models. The first stage of estimating inefficient fluxes during application of irrigation water is to estimate inefficient runoff from excess application, which follows calculation of crop irrigation requirement, and concurrent with estimation of inefficient delivery fluxes N_{evap}^d and L_{perc}^d. Excess irrigation supply (I^a), analogous to the *Application Requirements (AR)* parameter of Jägermeyr *et al.* (2015), is calculated for

each crop group (k, which can be either specific crop functional groups or pre-processed average land-cover groups described below):

$$I^a = \sum_i^n \sum_k^m \begin{cases} \max\left(0.5 S_{AWC}^k \overline{DU}_i - W_{irr} - L_{perc}^{rice}, 0.0\right) & where\ I^{demand,k} > 0 \\ 0 & where\ I^{demand,k} = 0 \end{cases} \quad \text{(13.2.2-6)}$$
(S6)

where S_{AWC}^k is a grid of crop (k) specific available water capacity (mm) that accounts for soil properties, $\overline{DU}_i$ is the application technology specific distribution uniformity coefficient (Jägermeyr et al., 2015), W_{irr} is the storage in the irrigation runoff retention pool (whose balance is calculated like the surface retention surface runoff pool of WBM, but applies only to the irrigated pixel fraction), and L_{perc}^{rice} is percolation associated with rice paddies, which is calculated separately (Grogan et al., 2017) and only applies over pixels with identified rice paddy, and $I^{demand,k}$ is the crop group specific irrigation demand. Existing storage in the irrigation runoff retention is subtracted assuming that irrigation requirements are reduced by whatever volume exists in pixels above field capacity assuming that existing excess volume in the irrigation retention pool is shared by all crops at a given pixel. Soil porosity defining soil saturation above field capacity is not presently a parameter input to WBM; therefore, we estimate the volume of additional water above field capacity that saturates soil as $0.5 S_{AWC}^k$. The distribution uniformity parameter ($\overline{DU}$) is a fraction of the crop field to which this soil saturation applies. $\overline{DU}$ for flood irrigation is close to 1 (all the soil in a crop area gets saturated) while for sprinkler irrigation about half of the possible saturation volume is actually applied. In the case of drip irrigation, a very small amount of water goes above W_{cap} and so $\overline{DU}$ is very low.

A fraction (ε_{mist}) of water delivered to irrigated crop fields can be lost non-beneficially above crop canopy from enhanced evaporation of, for instance, sprinkler mists. The flux of mist enhanced evaporation (N_{mist}^a) is calculated for each technology (i):

$$N_{mist}^a = (I^a + I^{demand,k})\varepsilon_{mist} \quad \text{(13.2.2-7)}$$

Parameterization of ε_{mist} depends on local climate and specifics of sprinkler technology such that they can vary widely from 0 to 40%, with most analyses estimating losses to be less than about 5% (Bavi et al., 2009; McLean et al., 2000; Uddin et al., 2010).

Application and delivery inefficiencies are summed to net irrigation demanded by crops to estimate a provisional gross irrigation flux (G^*):

$$G^* = I^{demand} + I^a + N_{mist}^a + N_{evap}^d + L_{perc}^d \quad \text{(13.2.2-8)}$$

A variety of functions are associated with sourcing available irrigation water in WBM, which yield a fraction of available water (X_{irr} where $X_{irr} = 1$ indicates complete availability) from the distribution of sources (Section 13.2.1). Where water supply is less than complete ($X_{irr} < 1$), all terms above are reduced linearly to utilize available supply via:

$$I^{demand} *= X_{irr} \quad \text{(13.2.2-9)}$$
$$I^a *= X_{irr} \quad \text{(13.2.2-10)}$$
$$N_{mist}^a *= X_{irr} \quad \text{(13.2.2-11)}$$
$$N_{evap}^d *= X_{irr} \quad \text{(13.2.2-12)}$$
$$L_{perc}^d *= X_{irr} \quad \text{(13.2.2-13)}$$

Actual gross irrigation (G) is calculated following routing later in the time-step, and small deviations between estimated and actual water availability are accounted for in subsequent timesteps.

Following routing through the stream network, the water balance of irrigation retention pool (W_{ret}) is updated using a stable solution and follows a conceptual order of flux priorities. The change in volume of W_{ret} is governed by the differential equation:

$$\frac{dW_{ret}}{dt} = I^{atm} + I^a - N^a_{evap} - L^a_{perc} - L^a_{rnff} \tag{13.2.2-14}$$

where I^{atm} is water incident to irrigated crop fields from natural precipitation or melt, N^a_{evap} is non-beneficial evaporation from saturated soil surface, L^a_{perc} is percolation from saturated soils to groundwater, and L^a_{rnff} is surface runoff from saturated soil. The stock of W_{ret} at the end of the timestep is calculated in four independent steps (denoted by superscripts):

1) $W^1_{ret} = W^0_{ret} + I^{atm} + I^a$

2) $N^a_{evap} = \min\left(A_{irr}\overline{DU} \times E_p, W^1_{ret}\right)$

   $W^2_{ret} = W^1_{ret} - N^a_{evap}$

3) $L^a_{perc} = \min\left(A_{irr}\overline{DU} \times P_{perc}, W^2_{ret}\right) \tag{13.2.2-15}$

   $W^3_{ret} = W^2_{ret} - L^a_{perc}$

4) $L^a_{rnff} = \min\left(A_{irr}\beta_{surf} \times \sqrt{2g} \times \frac{W^3_{ret}}{A_{irr}}, W^3_{ret}\right)$

   $W_{ret} = W^3_{ret} - L^a_{rnff}$

where W^0_{ret} is the stock of the water retention pool at the end of the previous timestep, E_p is the potential evapotranspiration (mm/d), β_{surf} is the parameter describing the rate of leakage from the irrigation (and surface) retention pools, and g is the constant of gravitational acceleration. The order of updating the irrigation retention pool gives first precedence to non-beneficial evaporation, and lowest precedence to surficial runoff. Therefore, we consider the irrigation water balance to be conservative with respect to non-beneficial losses, and we expect that non-consumptive losses may be marginally higher.

14. Livestock water demand and extraction

Input data

Input data for livestock water use are: average daily temperature, livestock density for each livestock category, service water per head, and two growth parameters. All livestock data and methods are from FAO (2006); the same temperature inputs are used here as in the irrigation water use section.

Method

Daily livestock water, L_w, for each livestock type is calculated each day as:

$$L_w = I_l + s_l \cdot T_m + SW_l \cdot D_l \tag{14-1}$$

where

I_l is an intercept parameter for livestock type l

s_l is a slope parameter for livestock type l [-]

T_m is the daily mean temperature, with a minimum value of 0 [°C]

SW_l is the daily service water volume required per animal

D_l is the density of livestock type l in the grid cell

Additionally, a growth rate can be applied to each livestock category to represent increases in population over the default circa year 2005 density data.

Consumptive vs non-consumptive use:

Livestock is assumed to consume 5% of water demand, with the remaining 95% returning to the system via runoff. *Note: these assumptions are hard-coded into the model.*

Table 14-1. Default global parameters (*Note: these parameters are being updated in the FAO GLEAM project; values reported here represent the current state of the model prior to the GLEAM project*):

Livestock	SlopeValue, s_j	InterceptValue, I_l	ServiceWater, SW_l	AnimalGrowthRate
buffalo	0.345	16.542	5	0.001863
cattle	0.345	16.542	5	0.001863
goats	0.215	4.352	5	0.003731
pigs	1.4575	-6.14	25	0.000309
poultry	0.019	0.1823	0.09	0.13397
sheep	0.57	-0.35	5	0.003

References:

FAO 2006. Livestock's Long Shadow: Environmental Issues and Options. http://www.fao.org/3/a0701e/a0701e.pdf

15. Domestic and industrial water demand and extraction

Input data

Data inputs for domestic and industrial water use are: domestic per capita water use, industrial per capita water use, and population density. Time series of domestic per capita water use, *DWpp* and industrial per capita water use, *IWpp,* are from Liu et al (2017). Annual population density projections are from the IIASA decadal projections (IIASA, 2007) under the B2 scenario.

Method

In WBM, the domestic and industrial sectors use water each day. Domestic water use, *Dw* [mm], is:

$$Dw = A \cdot DWpp \cdot D_{pop} \tag{15-1}$$

And industrial water use, *Iw* [mm] is:

$$Iw = A \cdot IWpp \cdot D_{pop} \tag{15-2}$$

where

A [km^2] is the area of the grid cell

DWpp [mm/d] is the domestic water use per capita

IWpp [mm/d] is the industrial water use per capita

D_{pop} [persons km^{-2}] is the population density

Note: There is no climate dependence in the above equation.

Default parameter values:

A global time series of *DWpp* was developed by Liu et al (2017). See Liu et al (2017) SI page 15 for details.

References:

IIASA, 2007. Greenhouse gas initiative (GGI) scenario database.

Liu J., Hertel T., Lammers R., Prusevich A., Baldos U., **Grogan D.S.,** Frolking, S. Achieving sustainable irrigation water withdrawals: Global impacts on food security and land use. *Environ. Res. Lett.* **12** (2017) 104009, doi:10.1088/1748-9326/aa88db.

16. Tracking

WBM tracks water (and constituents) from each given source (water source components in each individual grid cell) through flows and stocks within the model. Stocks include river storage, small and large reservoir storage, groundwater storage, runoff and irrigation storage pools, rice paddy flood waters, and soil moisture. Flows are runoff and baseflow, infiltration, recharge, river discharge, water discharge from reservoirs, evaporation, evapotranspiration, inter-basin transfers, water extracted for human water use, and return flows. The same tracking algorithm applies to all water source components. For any water component c in water storage stock S at time t:

$$S_c^k = \frac{(S_c^{k-1} \cdot S^{k-1}) + \Sigma_i(I_{c,i} \cdot I_i) - \Sigma_i(S_c^k O_j)}{S^k} \qquad (16\text{-}1)$$

where S_c^k is the fraction of stock S composed of component c at time k. S^k is the total volume of stock S at time k. I_i are inflows to and O_j are outflows from stock S, with $I_{c,i}$ the fractions of the ith flow composed of component c all at time-step k. Component stocks (S_c^k) are updated throughout the timestep such that solution is split into multiple operators as the various fluxes impact each stock.

All stocks and flows are considered well-mixed, so that the flows out of a stock have the same fractional water source components as the stock itself. All stocks are initialized with $S_c = 1$ for a default component c. See tracking options below for a description of the default components.

Depending on application for which tracking is being used, managing tracked components through spinup may need different assumptions. WBM provides two options for managing components through spinup:
1) Tracking occurs through spinup, and the model simulation period begins with stocks mixtures reflecting mixtures at the end of spinup.
2) All stocks are reset to at the beginning of the simulation period.

Option 1 is appropriate in identifying the most representative characterization of components within any stock. Option 2 is appropriate when accumulating the flux of a specific component during a dynamic simulation.

WBM Tracking Categories:
WBM currently has four types of water components that can be tracked:

1. Primary source components
   Primary source components are: rainwater, snow melt, glacier melt, and unsustainable groundwater. The default initialization category here is rainwater. Glacier melt can only be tracked if glacier water is provided as a model input.

2. Runoff source components
   Runoff source components are: surface runoff, baseflow, glacier melt, and snow melt. Note, this category includes a significant exception to Eq. (1): glacier melt and snow melt water that enter the groundwater pool and become baseflow are re-categorized as baseflow once the water enters the river system. The default initialization category here is rainwater.

3. Human use components

Human use components are: irrigation water return flows, domestic/industrial/livestock water return flows (all one category), relict water, and pristine water. The default category here is pristine water.

4. Runoff land mask components
   Runoff land mask components are defined by an input layer identifying different land grid cells as different sources. Runoff generated by each land category is then tracked through the system. Examples of land categories include political boundaries and land cover categories.
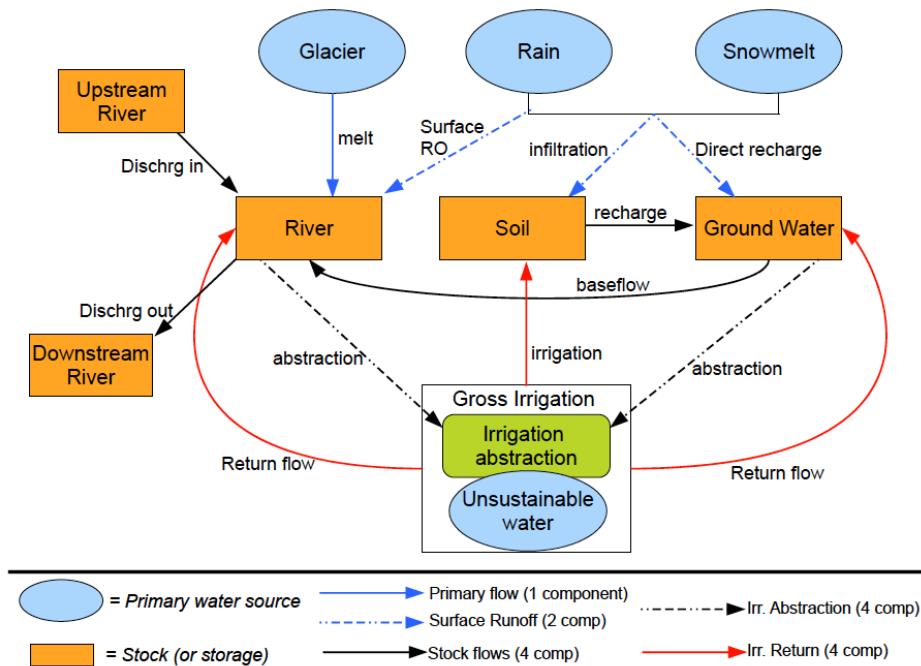

Figure 16-1. Example of tracking primary source components through WBM stocks and flows.

Water quality

17. Water Temperature

WBM calculates stream temperature using a volumetric weighted average of inputs, with adjustments made due to temperature equilibration with the atmosphere and due to radiative forcing.

Surface and baseflow runoff water temperature

WBM calculates runoff temperatures from each grid cell from volume-weighted mixtures of precipitation equilibrated with autoregressive integrated N-day moving average (ARIMA) of N previous day's daily air temperatures, and snowmelt, which is assigned a temperature of 0°C. The ARIMA weighted temperatures assume that water stored within soil or shallow groundwater equilibrate to average air temperature over different time windows. Furthermore, baseflow runoff is calculated as an average between the runoff temperatures are provided as a weighted average of N-day ARIMA of daily air temperatures and base layer temperature (BLT) that is an input to the system that represents the temperature of deep groundwater contributing to baseflow though modulated through the hydrodynamic response in the shallow groundwater pool. Generally a spatially explicit dataset of mean annual temperature is used as an input for the BLT temperature which is a ground temperature at depth of about 6 m where influence of seasonal air temperatures can be neglected. As such, there is considerable variation in seasonal surface runoff temperatures whereas shallow groundwater temperatures has a much lesser seasonal variablity. Impervious and open-water storm runoff is assumed to be in equilibrium with daily mean wet-bulb air-temperature.

The ARIMA temperature (T_A^N, C) of N-day moving window is calculated as

$$T_A^N = \sum_{i=0}^{\infty} \frac{(N-1)^i}{N^{i+1}} T_a^i \qquad (17\text{-}1)$$

where i is an index of the day prior to present and T_a^i is the air temperature at the day i. The ARIMA model is a simplified but effective way to account for heating/cooling inputs to a top layer of land from atmosphere which, in turn, transfers to the water in contact with the layer. Physically it represents a temperature of a fluid or solid body that receives daily portions of heat equivalent of 1/N of the body mass at that day's temperature which equilibrates with the cumulative body temperature and then it loses the heat equivalent of 1/N of the body mass at the mixed body temperature as shown in Figure 17-1.
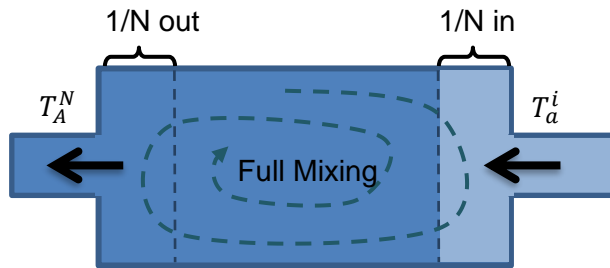

Figure 17-1. Autoregressive integrated N-day moving average (ARIMA) model schematics. Each day a heat equivalent of 1/N of the body mass at T_a^i temperature is added replacing heat of equivalent volume at previous day mixed and fully equilibrated temperature.

A smaller moving average window corresponds to a larger relative amount of daily mixing heat additions, and, thus, reducing the signal of previous days heating/cooling history. By default WBM uses 5-day moving window for the surface runoff temperature ($N_{sr} = 5$), and 15-day moving average for the shallow base flow temperature ($N_{bf} = 15$). The moving window day-interval values are chosen to correspond to a typical 10 and 150 cm soil layer heat propagation lag times from ambient air temperature according to GIPL soil temperature model (Jafarov et al., 2012; Wisser et al., 2011). We note that the current

implementation of landscape water temperature in runoff differs from the weighted daily averages of incident precipitation used in prior studies (Stewart et al. 2013, Samal et al. 2017); the current formulation approximates the effect of soil water changing temperatures through conductive processing following precipitation. Essentially the ARIMA model is a simplified model of the integral soil temperature of a given depth. Since the baseflow is formed as a mix of water from different soil or bedrock depths sources, the base flow temperature (T_{bf}) is, in turn, calculated as a weighted average of deeper shallow ground water (> 6 m deep) that has a value of long term mean annual air temperature (T_a^{av}) and calculated daily top soil layer temperature (T_A^N) using a weighting factor for the latter (w_{tl}) as following

$$T_{bf} = (1 - w_{tl}) * T_a^{av} + w_{tl} * T_A^N \qquad (17\text{-}2)$$

WBM uses a default value of 0.59 for the weighting parameter w_{tl}. This parameter and lengths of ARIMA running averages were found empirically by minimizing the error of simulated and observed runoff water temperatures from the data of Hubbard Brook site of the Long Term Ecological Research (LTER) network (Figure 17-2). While we find that this parameter combination works reasonably well over many study catchments in temperate regions, updating these values for region-specific studies if advisable.
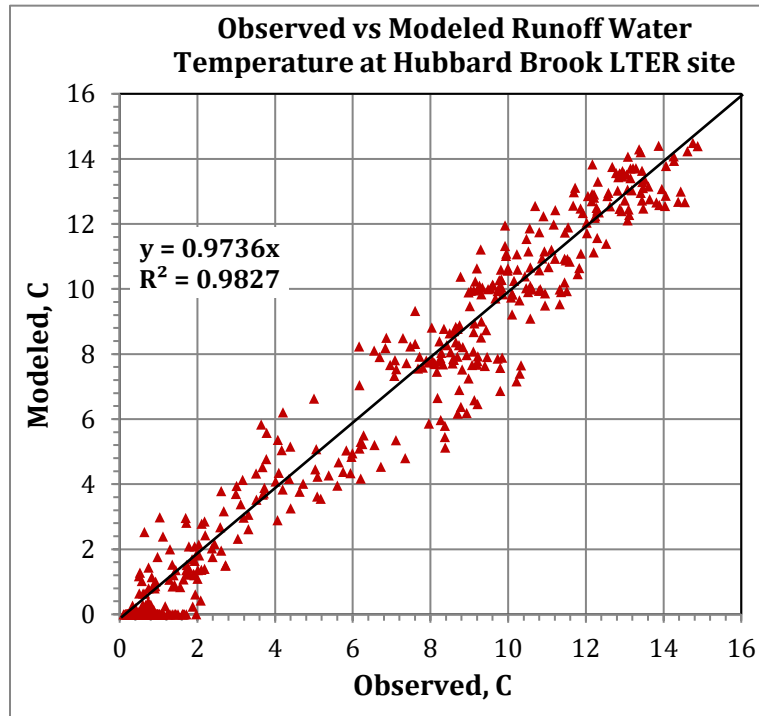

Figure 17-2. WBM simulated runoff temperature validation and results of parameter optimization using observational data from Hubbard Brook site of the Long Term Ecological Research (LTER) network for hydrological years (October through September) 2012-2014.

Streamflow (rivers and reservoirs) water temperature

Streamflows water temperatures are adjusted during discharge routing using the river temperature re-equilibration model RTRM (Stewart et al., 2013) that follows an approach based on a combined empirical and deterministic approach outlined in (Dingman, 1972). This method is appropriate for large scale applications, including lakes and large rivers (Morse, 1972) and is based on the theory of equilibrium

temperature; the temperature at which there is no net exchange of energy with the atmosphere (Edinger et al., 1968; Morse, 1972; Webb et al., 2003). The model uses wind speed, air temperature, weather conditions (clear/cloudy), relative or specific humidity, and incoming solar radiation to predict water temperatures. The in-stream equilibrium temperature (T_e, °C) and resulting water temperature (T_w, °C) of any given river reach is determined as (Dingman, 1972) and adjusted to a simulation time step:

$$T_e = T_a + \left[\frac{E_R - E_O}{\chi_E}\right]$$ (17-3)

$$T_w = (T_o - T_e)exp\left(-\frac{\chi_E}{\rho_w\, C_w\, h}\ \min_{\Delta t}\left(\frac{L}{u}\right)\right) + T_e$$ (17-4)

where T_a is the local air temperature (°C), E_R is the net incoming solar radiation (KJ m^{-2} d^{-1}), E_O is the heat loss rate when $T_w = T_a$ (KJ m^{-2} d^{-1}), χ_E is the energy exchange coefficient (KJ m^{-2} d^{-1} °C^{-1}), T_w is the resulting water temperature (°C), T_o is the initial water temperature of inflowing water from upstream [?] (°C), L is the length of the river grid cell (m), ρ_w is the density of water (kg m^{-3}), C_w is the specific heat of water (KJ kg^{-1} °C^{-1}), h is water depth (m), and u is the stream velocity (m d^{-1}), Δt (d^{-1}) is a simulation time step.

Notes for equation (17-4):

- Minimum operator in equation (17-4) controls exposure time while water travels through the grid cell which should not exceed the length of simulation time step to prevent double counting of water heating during its routing downstream.
- Water depth h is assumed do not exceed 20 m reservoir and lake depth which is an empirical limit to the active mixing surface layer indicated by typical lake thermocline (REF).

Values for E_O and χ_E are determined using linear functions based on data in New England rivers across various weather conditions and wind speeds (u_a) as follows (Dingman, 1972):

Clear:

$$E_O = 105 + 23u_a$$ (17-5)

$$\chi_E = 35 + 4.2\,u_a$$ (17-6)

Cloudy:

$$E_O = -73 + 9.1\,u_a$$ (17-7)

$$\chi_E = 37 + 4.6\,u_a$$ (17-8)

We found that the described above method yield systematic overestimation of instream water temperatures. The source of error is apparent as air humidity is ignored which controls equilibrium water temperature in contact with atmospheric air. So, WBM applies air humidity correction to equilibrium water temperature (T_e) following known thermodynamic formulation for dew point (wet bulb) temperature (Van Wylen et al., 1994)-

$$T_e^h = \frac{237.3 * \ln\frac{e_s\, rh}{610.78}}{7.5 * \ln 10 - \ln\frac{e_s\, rh}{610.78}}$$ (17-9)

where vapor pressure (e_s, Pa) is a function of relative humidity (rh, fraction)-

$$e_s = 610.78 * e^{\frac{17.27 * T_e}{T_e + 237.3}}$$ (17-10)

.

WBM water temperature calculation functions also have a correction to the net incoming solar radiation (E_R) for a canopy shading of streams which can be very considerable for small streams where they cross landscapes with high canopy forest during vegetation seasons with high values of Leaf Area Indices (LAI). Canopy shading of river water surfaces reduces solar radiation heating. It affects only portion of

river beds along their bank at the distance of the canopy heights assuming quasi-average 45° sun inclination throughout the daylight period and regardless to river bank orientation. In addition, the density of canopy also controls amount of radiation that can penetrate the vegetation cover. The latter is accounted by using normalized Leaf Area Index (LAI) in its annual time series. Putting together both canopy height and LAI the equation used for canopy shading factor (f_{shade}, fraction) is

$$f_{shade} = \overline{LAI} * \max_1 \left(\frac{H_c}{W_s} \right) \tag{17-11}$$

where $\overline{LAI}$ (unitless) is normalized LAI index between its annual min and max values, H_c (m) is canopy height, and W_s (m) is stream width. The default dataset for the canopy height is from (Simard et al., 2011). The canopy shading factor f_{shade} is added to cloud fraction correction to the unobstructed net incoming solar radiation for the water temperature calculation inputs.

Combining temperature of local runoff and streamflow routing

At each pixel, initial temperature at the beginning of the timestep is calculated as the volume weighted average of upstream inputs, local runoff in the current time step, and storage remaining in the stream reach following routing from the previous timestep (S_R):

$$T_0 = \left(S_R T_w^{k-1} + \sum_j^n Q_j^k T_{w,j}^k \right) \Big/ Q^* \tag{17-12}$$

where T_w^{k-1} is stream calculated at the end of the previous timestep, Q_j^k is the discharge flowing into the cell from upstream pixel j, and $T_{w,j}^k$ is the temperature of the jth upstream cell, and Q^* is the total flow at the pixel prior to calculating any retention in the cell from routing. Equilibrium temperature T_e is calculated early in the time-step, whereas the calculation of stream temperature is calculated during WBM's routing function call.

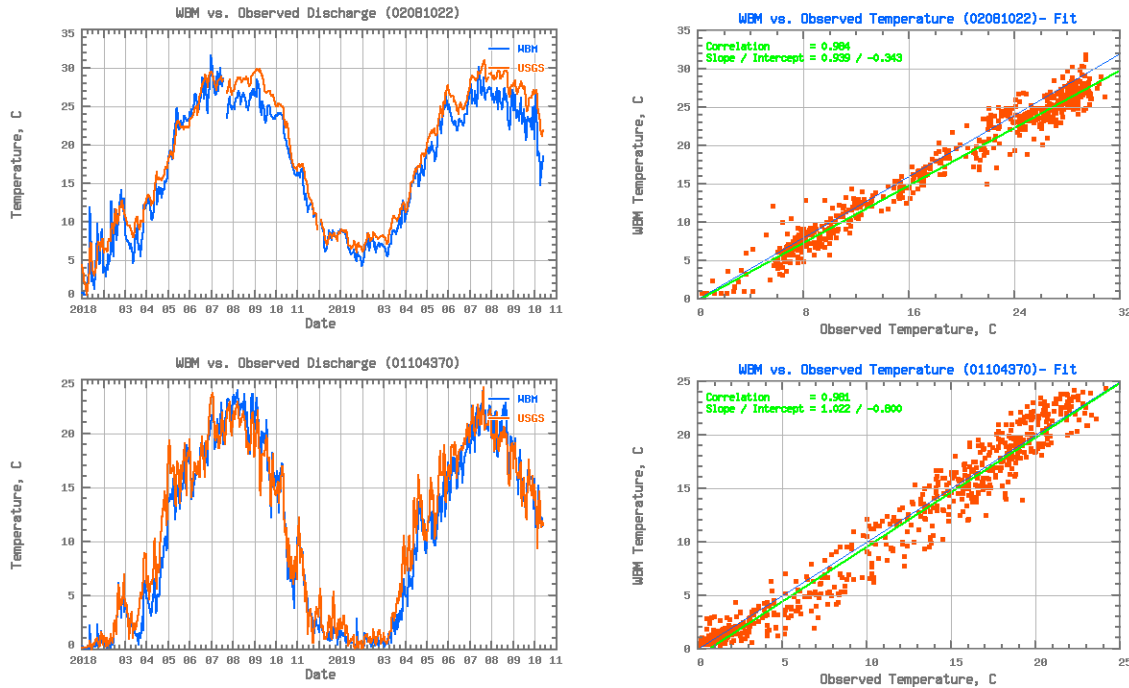We found a satisfactory match of WBM calculated and USGS observed water temperatures (Figure 17-3).

Figure 17-3. A typical match of simulated and observed river temperature in the Eastern US.

- **(Top) Large catchment river example: Station # 02081022, Roanoke River Near Oak City, NC, catchment area 22695 km^2, Nash-Sutcliffe coefficient is 0.94, R^2 is 0.984.**
- **(Bottom) Small catchment river example: Station # 01104370, Stony Brook near Weston, MA, catchment area 26 km^2, Nash-Sutcliffe coefficient is 0.95, R^2 is 0.981.**

References

Dingman, S.L., 1972. Equilibrium Temperatures of Water Surfaces as Related to Air Temperature and Solar-Radiation. Water Resources Research, 8(1): 42-&.

Edinger, J. E., Duttweiler, D. W., & Geyer, J. C. (1968). The Response of Water Temperatures to Meteorological Conditions. Water Resources Research, 4(5), 1137–1143. https://doi.org/10.1029/WR004i005p01137

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

Simard, M., Pinto, N., Fisher, J.B. and Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res., 116(G4): G04021.

Stewart, R.J., Wollheim, W.M., Miara, A., Vörösmarty, C.J., Fekete, B., Lammers, R.B. and Rosenzweig, B., 2013. Horizontal cooling towers: riverine ecosystem services and the fate of thermoelectric heat in the contemporary Northeast US. Environmental Research Letters, 8(2): 025010.

Van Wylen, G.J., Sonntag, R.E. and Borgnakke, C., 1994. Fundamentals of classical thermodynamics. Wiley, New York, xii, 852 p. pp.

Webb, B. W., Clack, P. D., & Walling, D. E. (2003). Water–air temperature relationships in a Devon river system and the role of flow. Hydrological Processes, 17(15), 3069–3084. https://doi.org/10.1002/hyp.1280

Wisser, D., Marchenko, S., Talbot, J., Treat, C. and Frolking, S., 2011. Soil temperature response to 21st century global warming: the role of and some implications for peat carbon in thawing permafrost soils in North America. Earth System Dynamics, 2(1): 121-138.

18. Nitrogen routing

DIN is loaded to the river network from both point source (DIN_{PS} [kg day^{-1}]) based on wastewater treatment plant effluent and non-point sources (DIN_{NPS} [kg day^{-1}]) based on human land use. For non-point source loading, by default WBM utilizes an empirical DIN loading function that was originally developed for the Ipswich River watershed located in northeast Massachusetts (Wollheim et al., 2008). This sigmoidal function relates the fraction of human land use upstream (both developed and agriculture) with the concentration of DIN in runoff (C_{DIN}^{NPS}) [g L^{-1}]. Specifically, C_{DIN}^{NPS} is calculated as,

$$C_{DIN}^{NPS} = \frac{Asym}{1 + e^{\frac{(Xmid - HLU)}{scale}}} \tag{18-1}$$

where $Asym$ [g L^{-1}] is the maximum concentration found in runoff, HLU [-] if the fraction of both developed and agricultural land use, $scale$ [-] determines the range of HLU at which concentration rises, and $Xmid$ is the inflection point of that curve. $Xmid$ depends on runoff ($runoff$) and has an intercept ($Xmid_b$) and a slope ($Xmid_m$).

$$Xmid = Xmid_b + Xmid_m \cdot \log(runoff) \tag{18-2}$$

Parameters $Asym$, $Xmid_b$, and $Xmid_m$ default to values reported in Wollheim et al. (2008), but are accepted as input parameters when locally available information is available, or for the purposes of model calibration.

Grid cells containing a waste water treatment plants (WWTP) receive DIN loading as,

$$L_{DIN}^{WWTP} = WWTP_{loadRate} * Pop_{WWTP} * (1 - R_{Trmt}) \quad [kg/d] \tag{18-3}$$

where daily nitrogen load influent to the treatment plant ($WWTP_{loadRate}$) [kg/Person/d], population served by each plant (Pop_{WWTP} [P]) for each treatment plant is read into the model and interpolated linearly between years of known service population. Nitrogen removal for treatment plants (R_{Trmt}) are values input from a lookup-table relating removal rate to treatment type (Table 18-1). The data used for waste water treatment plants in the USA is available through the US Environmental Protection Agency Clean Water Needs Survey (CWNS) data (USEPA, 2016) and includes plant coordinates in longitude-latitude, population served, and treatment type.

Table 18-1: Nitrogen removal fractions for each process type for waste water treatment plants following .

Process Type	Removal fraction (-)
Primary	0.1
Secondary	0.5
Secondary (Advanced) / Tertiary	0.8

Concentration of DIN in local runoff (C_{DIN}^{Local} [g L^{-1}]) entering the stream network adds the flux from WWTP to concentration associated with NPS loading via Equation 18-4.

$$C_{DIN}^{Local} = C_{DIN}^{NPS} + \frac{L_{DIN}^{WWTP}}{(1000\,A\,runoff)} \tag{18-4}$$

Where A is pixel area in m^2, and runoff has units of mm d^{-1}.

Stream nitrate concentration is calculated in two steps. Prior to calculating the concentration in the stream during the current time-step, evapoconcentration of DIN from evaporation from the river network is calculated. Stream concentration from the prior timestep $\left(C_{DIN}^{stream^{t-1}} \right)$ is scaled upwards by the flux of network evaporation by Equation B.54.

$$C_{DIN}^{stream^1} = C_{DIN}^{stream^{t-1}} \left[\frac{A^{stream} E^{stream} dt + S}{S} \right] \tag{18-5}$$

In Equation 18-5, $C_{DIN}^{stream^1}$ is an intermediate solution of stream DIN concentration prior to the routing, A^{stream} [m^2] is the surface area of open water, E^{stream} [$m\,d^{-1}$] is the evaporation rate from open water surfaces, and S [m^3] is the flow storage of unrouted streamwater. During this step, the mass of DIN that is removed from the surface network from abstraction for human uses is calculated for verifying whole basin DIN mass balance.

Stream DIN concentration is then advanced during routing in two steps that account for 1) new inputs to the network (Equation 18-6), and 2) in-stream DIN removal. Stream DIN concentration after adding local inputs $\left(C_{DIN}^{stream^2}\ [g\,L^{-1}]\right)$ is given by equation 18-6.

$$C_{DIN}^{stream^2} = \frac{1}{Q}\left(\sum_{j=0}^{n} Q^j C_{DIN}^{stream^2} + (1000\ A\ runoff\ dt)\ C_{DIN}^{Local} + \frac{V_{stream}}{dt} C_{DIN}^{stream^1}\right) \qquad (18\text{-}6)$$

where Q is discharge within the reach during the time-step, and Q^j is the discharge from the n reaches upstream draining to the respective pixel.

Then stream DIN concentration at the end of the time-step is calculated in Equation 18-7.

$$C_{DIN}^{stream} = C_{DIN}^{stream^2}(1 - R) \qquad (18\text{-}7)$$

The proportion of DIN removed within each grid cell by physical and biogeochemical processes (R [-]) is calculated following the temperature corrected first-order uptake methods of the Stream Solute Workshop (1990). R is calculated using the efficiency loss model (Mulholland et al. 2008) with an uptake velocity (v_f) [m day^{-1}] that varies with both in-channel water temperature and DIN concentration. Removal is calculated by Equation 18-8,

$$R = 1 - exp\left(-\frac{v_f}{H_L}\right) \qquad (18\text{-}8)$$

where the uptake velocity (v_f [$m\,d^{-1}$]) and the hydraulic load (H_L [$m\,d^{-1}$]) are given by Equations 18-9 and 18-10.

$$V_f = \frac{86400\ s\,d^{-1}}{100\ cm\,m^{-1}} 10^{\left(int + slope\,\cdot\,\log\left(1e^6\mu g\,g^{-1}\,C_{DIN}\right)\right)} \cdot X_{Temp} \qquad (18\text{-}9)$$

$$H_L = \frac{d}{\tau} = \frac{d}{\Delta l/u} = \frac{Q}{A} \qquad (18\text{-}10)$$

In Equations 18-9, int [log cm s^{-1}] is the uptake velocity intercept (value of -2.975; Mulholland et al. 2008), and $slope$ [-] is the efficiency loss slope (slope of the uptake velocity vs. concentration, value of -0.493; Mulholland et al. 2008). Conversion between cm s^{-1} and m d^{-1} and μg L^{-1} and g L^{-1} and needed to relate the units of the empirical relationships from Mulholland et al. (2008) and the native units in WBM. These conversions are dropped in the derivation below. A water temperature correction (X_{Temp} [-]) is given by Equation 18-11.

$$X_{Temp} = Q_{10}^{\left(\frac{(T_w - Tref)}{10}\right)} \qquad (18\text{-}11)$$

In Equation 18-11, Q_{10} is the factor (default of 2) of increase for every 10 degrees difference of water temperature (T_w) from a reference temperature (T_{ref}) that = 20°C based on the data in Mulholland et al. 2008. In Equation 18-10, d is water depth (but is limited to 20 m to prevent unrealistically high H_L for reservoirs considering not all areas of deep reservoirs will be effective at denitrification), and the τ is the reach residence time calculated as the reach length (m) divided by the daily flow velocity u (m d^{-1}). The reach residence time τ is limited to the time-step length to prevent unrealistically high values of removal from being calculated.

Reach scale velocity, depth, and temperature are estimated based on runoff and storage within reaches at the beginning of the timestep, and thus H_L and several terms for v_f can be calculated efficiently prior to the networked routing calculation. However, because v_f is dependent on C_{DIN}, DIN removal in rivers (R_{River}) must be calculated as a network operation. Prior to the network calculation, a denitrification coefficient (χ_{denit}) [-] is calculated. First, we expand the definition of v_f in equation 18-8 (Equation 18-12):

$$R = 1 - \exp\left(-\frac{10^{int + slope \cdot \log(C_{DIN}) \cdot X_{temp}}}{H_L}\right) \qquad (18\text{-}12)$$

Then, after expanding powers and logs:

$$R = 1 - \exp\left(-\frac{(10^{int} C_{DIN}^{slope} X_{temp})}{H_L}\right) \qquad (18\text{-}13)$$

Terms are then rearranged:

$$R = 1 - \exp\left(-\frac{10^{int} X_{temp}}{H_L} C_{DIN}^{slope}\right) \qquad (18\text{-}14)$$

and the first term is precomputed as a denitrification coefficient (χ_{denit}):

$$\chi_{denit} = \frac{10^{int} X_{Temp}}{H_L} \qquad (18\text{-}15)$$

Then the denitrification coefficient (χ_{denit}) and efficiency loss slope ($slope$) are passed to functions performing the downstream network calculation of C_{DIN} while simultaneously calculating river removal R_{River} according to equation 18-16.

$$R_{River} = 1 - \exp\left(-\chi_{denit}\left(C_{DIN}^{stream\, slope}\right)\right) \qquad (18\text{-}16)$$

The default parameterization in WBM provided above represents the permanent DIN removal from the stream network by denitrification, but int and $slope$ are parameters that can be input to represent assimilation or local estimates of DIN removal processes.

For river reaches within reservoirs, an alternative to the in-stream denitrification is available. Grid cells containing reservoirs remove DIN following the empirical relationship developed by Seitzinger et al. (2002), which relates the fraction of DIN removed within the waterbody to hydraulic load and utilizes the same water temperature correction factor as the efficiency loss model.

$$R_{reservoirs} = \left(0.88453 \cdot H_L^{-0.3677}\right) \cdot X_{Temp}/365.26 \qquad (18\text{-}17)$$

In 18-17 A_{rsvr} [m^2] is the surface area of the reservoir, H_L [m day^{-1}] is the hydraulic load calculated by Equation 18-18:

$$H_L = Q_{rsvr}/A_{rsvr} \qquad (18\text{-}18)$$

where Q_{rsvr} is discharge out of the reservoir, and assuming the reservoir surface area is equal to the reservoir benthic surface area. Dividing by 365.26 converts the original removal rate from Seitzinger et al. (2002) for a daily time-step.

References

Mulholland, P.J., Helton, A.M., Poole, G.C., Hall, R.O., Hamilton, S.K., Peterson, B.J., Tank, J.L., Ashkenas, L.R., Cooper, L.W., Dahm, C.N., Dodds, W.K., Findlay, S.E.G., Gregory, S.V., Grimm, N.B., Johnson, S.L., McDowell, W.H., Meyer, J.L., Valett, H.M., Webster, J.R., Arango, C.P., Beaulieu, J.J., Bernot, M.J., Burgin, A.J., Crenshaw, C.L., Johnson, L.T., Niederlehner, B.R., O'Brien, J.M., Potter, J.D., Sheibley, R.W., Sobota, D.J., Thomas, S.M., 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature 452, 202–205. https://doi.org/10.1038/nature06686

Park C 1977 World-wide variations in hydraulic geometry exponents of stream channels – Analysis and some observations. *J Hydrol* **33** 133-146

Seitzinger, S.P., Styles, R.V., Boyer, E.W., Alexander, R.B., Billen, G., Howarth, R.W., Mayer, B., Van Breemen, N., 2002. Nitrogen retention in rivers: model development and application to watersheds in the northeastern USA. Biogeochemistry 57, 199–237.

Siebert S and Döll P (2010) Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *J. Hydrol.* 384, 198-217, doi:10.1016/j.jhydrol.2009.07.031.

Stream Solute Workshop, 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. Journal of the North American Benthological Society 95–119.

USEPA, 2016. Clean Watersheds Needs Survey 2012: Report to congress (No. EPA-830-R-15005).

Wollheim, W., Peterson, B.J., Thomas, S.M., Hopkinson, C.S., Vorosmarty, C.J., 2008. Dynamics of N removal over annual time periods in a suburban river network.

Appendix – History of the WBM code and its variations

The Water Balance Model has had several variations and names during its evolution. Code names (with principal coders) have included:

|	
WBM	Water Balance Model (Vörösmarty, Fekete, Prusevich)
P/WBM	Permafrost Water Balance Model (Holden)
PWBM	Pan-Arctic Water Balance Model (Lammers)
PWBM	Permafrost Water Balance Model (Rawlins)
WBMPlus	WBM with irrigation and reservoirs added (Wisser)
WBM_TrANS	Rewrite of WBM in Perl/PDL (Prusevich)
FrAMES	Framework for Aquatic Modeling of the Earth System; includes WBMplus (Wollheim, Stewart, Zuidema)
WBMsed	WBM with sediment module (Cohen)
WBM-TP2M-ReDS	WBM with power plants (Miara)
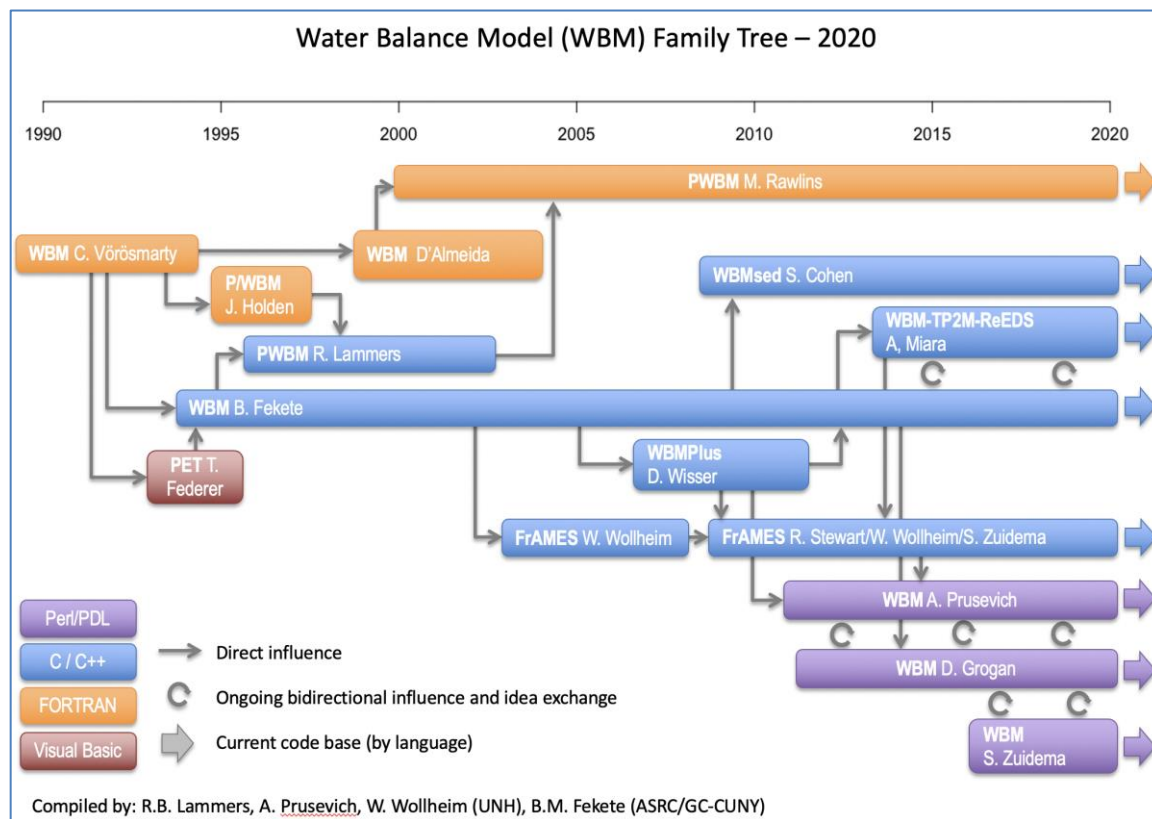
Figure A-1. WBM family tree.

New versions of the code base (re-writes):

Charles Vörösmarty – original model code (FORTRAN) – 1989-1994
  Jonathan Holden – Permafrost WBM (P/WBM) (FORTRAN) – 1995-96

Balazs Fekete – object oriented version (C++) – 1994-present

    Dominik Wisser – many code additions (C++) – 2007-2011

    Wil Wollheim – biogeochemical additions to WBM (FRAMES) (C++) – 2001-2018

        Rob Stewart – additional water quality functions (FrAMES) (C++) – 2008-2018

        Shan Zuidema – Additional modifications (FrAMES) (C++) – 2015-present

Richard Lammers – Re-write of code using Holden P/WBM as guide (C) – 1997-2003

Cassiano D'Almeida – Re-write WBM (FORTRAN) – 1999-2004

    Michael Rawlins – Build PWBM onto D'Almeida version (FORTRAN)

                with input from Holden and Lammers code bases – 2000-present

Alex Prusevich – Re-code to emphasize human water use (Perl/PDL) – 2010-present

    Danielle Grogan – added new functions to WBM (Perl/PDL) – 2011-present

    Shan Zuidema – added new functions to WBM (Perl/PDL) – 2016-present