Complete some part of CWatM manual

10.4.5 Soil and soil hydraulic properties

The infiltration capacity of the soil is using the Xinanjiang (also known as VIC/ARNO) model (Todini, 1996)[52]. Please find a full description of the infiltration into soil, preferential bypass flow and evapotranspiration in *Burek et al. (2020)* (https: //doi.org/10.5194/gmd-13-3267-2020).

Preferential flow, going directly to recharge groundwater, is obtained by the following relation:

$$Preferential\_flow = Infiltration \times \theta_{rel}{}^{PrefFlow\_coef}$$

where *PrefFlow_coef* is a calibration factor.

Next, modeling of unsaturated flow and transport processes can be done with the 1D Richard equation, which requires a high spatial and temporal distribution of the soil hydraulic properties.

$$\frac{\delta\Theta}{\delta t} = \frac{\delta}{\delta z}[K(\Theta(\frac{\delta h(\Theta)}{\delta z} - 1)] - S(\Theta) \text{ (1D Richard equation)}$$

where: - Θ is soil volumetric moisture content [L3/L3]
- t is time [T]
- h is soil water pressure head [L]
- K() is unsaturated hydraulic conductivity [L/T]
- z is vertical coordinate
- S is source sink term [T-1].

With the simplification of the1D Richard equation, e.g. flow of soil moisture is entirely gravity-driven and matrix potential gradient is zero, this implies a flow that is always in downward direction at a rate that equals the conductivity of the soil.

$$Q = \frac{\delta\theta}{\delta t} = -K(\theta)$$

where Q is the downward vertical flow.

The relationship can now be described with the model of Mualem (1976)[50] and with the Van Genuchten model (1980)[51] equation. Please find a full description of the soil process modeling in Burek et al. (2020) : https: //doi.org/10.5194/gmd-13-3267-2020

$$K(\Theta) = K_s(\frac{\Theta-\Theta_r}{\Theta_s-\Theta_r})^{0.5}\{1 - [1 - (\frac{\Theta-\Theta_r}{\Theta_s-\Theta_r})^{1/m}]^m\}^2 \text{ (Van Genuchten equation)}$$

Where:- K_s is saturated conductivity of the soil [cm/d-1]
- Θ, Θ_s, Θ_r are actual, maximum and residual amounts of moisture in the soil [mm]
- m is calculated from the pore-size index λ : $m = \frac{\lambda}{\lambda+1}$.

Thus, the soil hydraulic parameter Θ_s, Θ_r, λ and K_s are needed to simulated soil water transport for the Van Genuchten model. Note that the soil hydraulic parameter α is also needed to compute capillary rise.

As the soil is discretized in three layers in CWatM, the water balance of each layer *i* is:

$$\frac{\Delta m_i}{\Delta t} = Q_{i-1} - Q_i - CR_i + CR_{i+1}$$

where CR_i is the capillary rise from the layer *i* to the layer *i-1*.

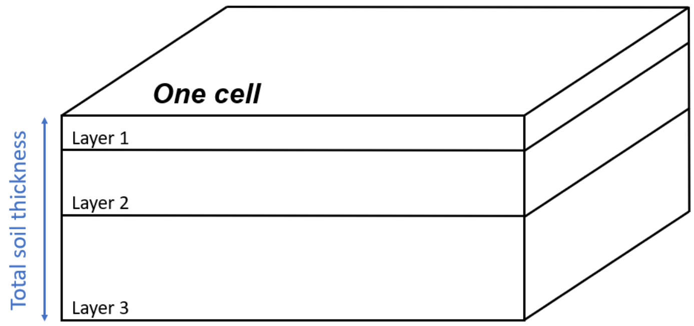

Figure 1: Scheme of the soil part as represented in CWatM.

Capillary rise is given by:

$$CR_i = min\left(K_{i-1}(\theta_{i-1}) \times \left(1 - \frac{\theta_i - \theta_{i,r}}{\theta_{i,s} - \theta_{i,r}}\right), \sqrt{K_i(\theta_{i,fc}) \times K_{i-1}(\theta_{i-1,fc})}\right),$$

where fc refers to "field capacity". Note that $\theta_{i,fc}$ is obtained by the Van Genuchten - Mualem equation:

$$\theta_{i,fc} = \theta_{i,r} + \frac{\theta_{i,s} - \theta_{i,r}}{(1 + (\alpha_i h)^n)^m}$$

where the pressure $h = 100\ cm$ and the soil hydraulic parameter α_i is the inverse of air entry suction.

Finally, once the unsaturated conductivity for each soil zone is determined, the water flux to the next zone can be estimated. At a time step of 1 day and high *K(θ)*, the vertical flux can exceed the available soil moisture. Therefore, the soil moisture equation is solved iteratively on a subdaily time step.

10.4.6 Groundwater

Using a linear reservoir representation

For groundwater modelling, maps of the recession constant [day^{-1}] and the storage coefficient [] are needed. Gleeson et al., (2011)[55]and Gleeson et al. (2014)[56]can provide data for this (GLHYMPSE). Alternatively, recession maps based on streamflow analysis over more than 3000 catchments (Beck et al., 2013) can be used.

The groundwater reservoir is filled by downward flow called "groundwater recharge" being the sum of preferential flow and percolation from the third soil layer. However, a fraction of this recharge reach directly runoff based on an "impervious" fraction on each cell. The map with impervious factor is then multiplied by a calibration factor.

The groundwater store is emptied in function of its storage following a linear model:

$$Q(t) = V(t) \, . \, k_{GW}$$

Where *Q(t)* is the baseflow [m/day] send to rivers, *V(t)* is the groundwater storage [m] at time t and k_{GW} is the groundwater recession coefficient [day^{-1}].

Groundwater feeds also soil humidity by an upward flow, called "capillary rise from groundwater", using a sub-grid representation of the topography. First, water table is computed by dividing groundwater storage *V(t)* by porosity for each cell.

The capillary rise flow from groundwater is next computed as:

$$CR_{GW} = 0.5 \times CapRiseFrac\left(1 - \frac{\theta_3 - \theta_{3,r}}{\theta_{3,s} - \theta_{3,r}}\right), \sqrt{K_3(\theta_{3,fc}) \times K_3(\theta_3)}$$

where *CapRiseFrac* is the fraction of the cell where water table more 5 is superior to the altitude. In each cell, topography is first detrended to remove regional trend, then the minimal altitude is subtracted. Then, altitude of each percentiles of cell area are computed (0,1,2,5,10,20,…100%). Altitude at 0 % means that 100 % of the cell is above the corresponding altitude, altitude at 10 % means that 90 % of the cell is above the corresponding altitude, etc… allowing to easily compute *CapRiseFrac*.

Data:

GLHYMPS—Global Hydrogeology Maps of permeability and porosity (Gleeson et al., 2014)

GSCD - Global patterns in base flow index and recession based on streamflow

observations from 3394 catchments (Beck et al., 2013)

Topography - MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets (Yamazaki et al., 2019)

Using the 2D groundwater flow equation (implemented by ModFlow)

Instead of modelling groundwater by a linear reservoir for each cell, a 2D groundwater model can be used in the aim to take into account groundwater lateral exchanges between cells. In this framework the well-known numerical model *ModFlow* is coupled with *CWatM* using the Python *Flopy* library (*McDonald and Harbaugh, 1988; Harbaugh, 2005; Bakker et al., 2016*). We used here ModFlow version 6. Thus, this version is faster because using BMI (https://github.com/MODFLOW-USGS/modflow6) to store ModFlow variables for each time step without writing them on the disk. The ModFlow model is initialized at the beginning and then variables such the simulated hydraulic heads are stored waiting the recharge for the following day coming from CWatM.

Technically, the two models are not completely coupled, CWatM is run during X days, providing inputs to groundwater, then ModFlow is run using a timestep of X days in one run, providing in return the baseflow and the capillary rise from groundwater for each cell. When water table reach the altitude of the soil bottom groundwater provides to CWatM using the Drain package of ModFlow. The upward flow is separated into baseflow and capillary rise base on the percentage of river defined for each ModFlow cell. This river network is created based on the topographic map at a finer resolution than the ModFlow resolution. Note that the ModFlow resolution can be changed (ie 100,500,1000… m). Each ModFlow cell is limited by the altitude of the surface obtained by a topographic map at finer resolution.

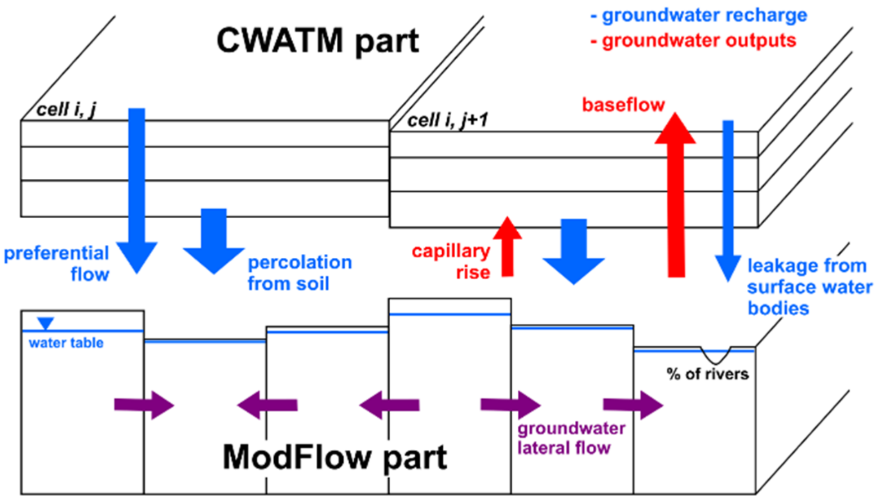

Figure 2: Scheme of the coupling between CWatM and ModFlow.

The advantages of using a physically-based groundwater model include to better represent groundwater levels, lateral flow and thus the non-local effects of pumping, as well as leakage from surface water, through river infiltration, canals, and reservoirs (these in development). River infiltration occurs when the water table is below the soil layer and equals permeability times the percentage of rivers in the cell.

10.4.7 Lakes and Reservoirs

The HydroLakes database http://www.hydrosheds.org/page/hydrolakes (Lehner et al. (2011); Messager et al. (2016), provides 1.4 million global lakes and reservoirs with a surface area of at least 10ha. CWatM differentiate between big lakes and reservoirs which are connected inside the river network and smaller lakes and reservoirs which are part of a single grid cell and part of the runoff concentration within a grid cell. Therefore, the HydroLakes database is separated into "big" lakes and reservoirs with an area ≥ 100 km2 or a upstream area ≥ 5000 km2 and "small" lakes which represent the non-big lakes. All lakes and reservoirs are combined at grid cell level but big lakes can have the expansion of several grid cells. Lakes bigger than 10000 km2 are shifted according to the ISIMIP protocol. Lake and reservoir (LR) data are specified by an id for each LR, type of LR (1 for lake, 2 for reservoir), area of LR, year of construction, and average discharge at the outlet. Water releases are based on thresholds of relative storage related to several limits, average discharge, and the calibration parameter normalStorageLimit.

For the coupled CWatM-ModFlow6 version, the leakage can occur from lakes and reservoirs as well as under rivers. Leakage occurs if the aquifer layer in the corresponding CWatM cell is unsaturated in most of the CWatM cell area. The leakage flow for each time is given by the following equation:

$$if\ water\ table\ is\ reaching\ the\ top\ of\ the\ aquifer\ in\ less\ than\ X\%$$

$$of\ the\ CWatM\ cell\ area:$$

$$Leakage = min\left(S_{surface\ water\ body}, k_{leakage} \times fraction\ area\ of\ surface\ water \right)$$

$$else: Leakage = 0$$

The upward flow from ModFlow to CWatM surface water bodies (including rivers) occurs when water table reaches the top of aquifer thanks to the ModFlow DRAIN package.

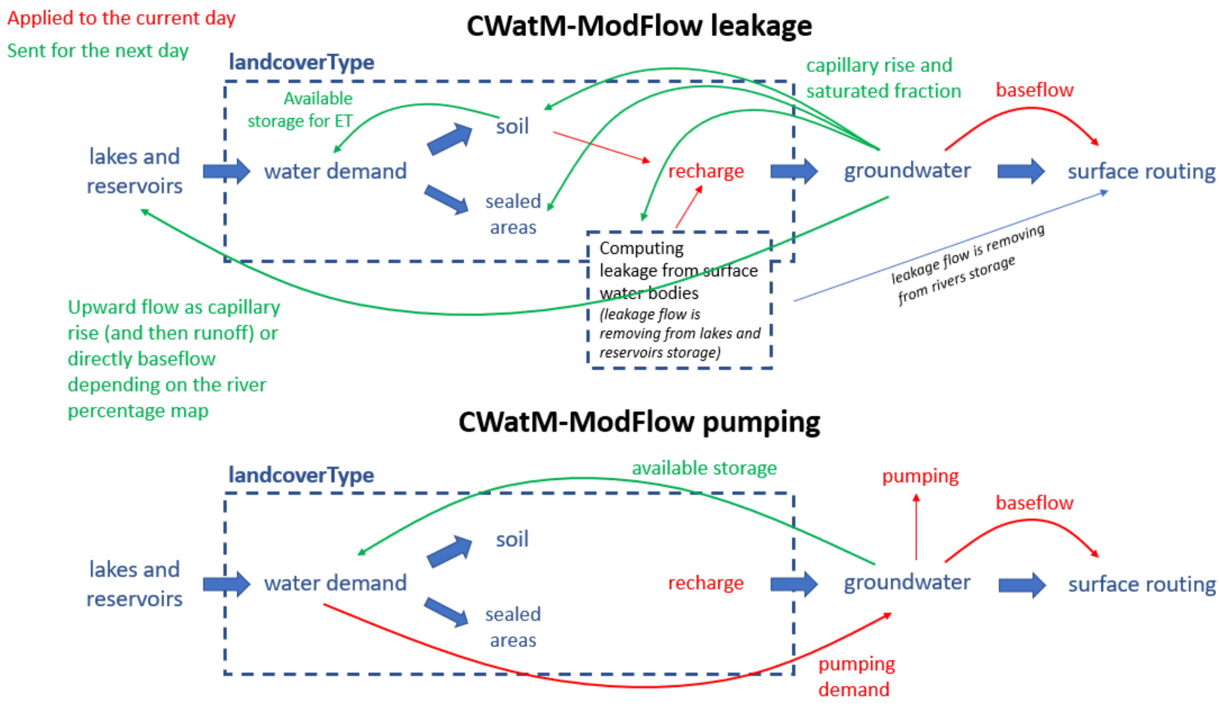

Figure 3: Scheme of the coupling between groundwater and surface water bodies (top); Scheme of the coupling between groundwater demand and groundwater pumping (bottom).

10.4.8 Water demand

Water demand come from four different sectors: domestic, industrial, livestock and irrigation. While the three first components need to be given as input, irrigation demand is computed by comparing the transpiration demand and water available to the roots in the soil.

Water demand is satisfied by either surface waters (rivers, reservoirs or lakes) or groundwater. The fraction of water demand satisfied by surface water can be estimated as (Wada et al.):

$$Surface\ abstraction = water\ demand\ \times \frac{discharge_{avg}}{discharge_{avg} + baseflow_{avg}}$$

where *avg* refers to the long-term average. When surface storage is empty water is taken from groundwater storage. If groundwater storage is also empty, water can be extracted from an unlimited fossil groundwater reservoir if using the linear groundwater reservoir (not modelled in CWatM).

Finally, while water demand occurs in each mesh, water can be extracted from the neighboring cells before to extract water stored in groundwater, and also from fossil groundwater if groundwater reservoir is empty.

An additional option called “command area” can be used to extract water from reservoir-associated meshes.

Technically, the water demand module sent the groundwater demand to the groundwater module. Then, this water is removed from the groundwater reservoir. If ModFlow coupling is used, pumping are applied thanks to the ModFlow package WEL through wells whose the location has to be defined at the beginning of the simulation. For the following time step, if ModFlow cells is not enough saturated (close from complete desaturation or if water table is too deep compared to the real deep of boreholes in the region), the groundwater demand in this cell will be prevented until the water table goes up.