Geosci. Model Dev., 17, 2141–2164, 2024
https://doi.org/10.5194/gmd-17-2141-2024

Geoscientific Model Development
Open Access

Model description paper

Validation of a new global irrigation scheme in the land surface model ORCHIDEE v2.2

Pedro Felipe Arboleda-Obando[1], **Agnès Ducharne**[1], **Zun Yin**[2], **and Philippe Ciais**[3]

[1]Laboratoire METIS (UMR 7619), IPSL, Sorbonne Université, CNRS, EPHE, 75005 Paris, France
[2]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, 08540 New Jersey, USA
[3]Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CNRS-CEA-UVSQ, Gif-sur-Yvette, 91191 Essonne, France

Correspondence: Pedro Felipe Arboleda-Obando (pedro.arboleda_obando@upmc.fr)

Received: 16 June 2023 – Discussion started: 14 July 2023
Revised: 4 January 2024 – Accepted: 12 January 2024 – Published: 15 March 2024

Abstract. Irrigation activities are important for sustaining food production and account for 70 % of total global water withdrawals. In addition, due to increased evapotranspiration (ET) and changes in the leaf area index (LAI), these activities have an impact on hydrology and climate. In this paper, we present a new irrigation scheme within the land surface model ORCHIDEE (ORganising Carbon and Hydrology in Dynamic EcosystEms)). It restrains actual irrigation according to available freshwater by including a simple environmental limit and using allocation rules that depend on local infrastructure. We perform a simple sensitivity analysis and parameter tuning to set the parameter values and match the observed irrigation amounts against reported values, assuming uniform parameter values over land. Our scheme matches irrigation withdrawals amounts at global scale, but we identify some areas in India, China, and the USA (some of the most intensively irrigated regions worldwide), where irrigation is underestimated. In all irrigated areas, the scheme reduces the negative bias of ET. It also exacerbates the positive bias of the leaf area index (LAI), except for the very intensively irrigated areas, where irrigation reduces a negative LAI bias. The increase in the ET decreases river discharge values, in some cases significantly, although this does not necessarily lead to a better representation of discharge dynamics. Irrigation, however, does not have a large impact on the simulated total water storage anomalies (TWSAs) and its trends. This may be partly explained by the absence of nonrenewable groundwater use, and its inclusion could increase irrigation estimates in arid and semiarid regions by increasing the supply. Correlation of irrigation biases with landscape descriptors suggests that the inclusion of irrigated rice and dam management could improve the irrigation estimates as well. Regardless of this complexity, our results show that the new irrigation scheme helps simulate acceptable land surface conditions and fluxes in irrigated areas, which is important to explore the joint evolution of climate, water resources, and irrigation activities.

1 Introduction

Irrigation seeks to increase crop yields by reducing plant water stress (Siebert and Döll, 2010; Klein Goldewijk et al., 2017) and supports about 43 % of the world's food production on about 20 % of the arable land (Siebert and Döll, 2010; Grafton et al., 2017). The beneficial effects of irrigation on food production, population, and economic growth have dramatically pushed the increase in irrigated areas during the 20th century from 28 Mha in 1850 to 276 Mha in 2000 (Klein Goldewijk et al., 2017; Siebert et al., 2015). As a consequence, by the year 2000, irrigation accounted for 70 % of the total water withdrawn (between 2657 and 3594 $km^3\,yr^{-1}$). The consumptive water use, i.e., the part of the withdrawn water that actually becomes evapotranspiration (ET) and does not flow to surface supplies and groundwater, represents half of that volume (between 1021–1598 $km^3\,yr^{-1}$, which is around 1.7 % of total continental ET of $75.6 \times 10^3\,km^3\,yr^{-1}$, according to Jung et al., 2019) and represents around 90 % of the total consumptive water use by human activities (Pokhrel et al., 2016; Döll et al., 2012; Hoogeveen et al., 2015).

Water abstraction and corresponding ET increase have a direct impact on the water and energy balances and on surface and subsurface hydrology (Döll et al., 2012; Taylor et al., 2013; Vicente-Serrano et al., 2019). The atmosphere also reacts to these changes in land surface fluxes, for instance, with regional increases or decreases in rainfall rate or decreases in temperature extremes (Lo and Famiglietti, 2013; Guimberteau et al., 2012b; Cook et al., 2015; Al-Yaari et al., 2019; Thiery et al., 2020). Thus, it was recently shown that climate models better capture historical trends in evapotranspiration if they account for irrigation and its expansion, although the resulting cooling effect is too strong if irrigation is not limited by water availability (Al-Yaari et al., 2022). Finally, with the acceleration of climate change, the irrigation water demand is likely to increase, not only by expansion of the irrigated area but also by increasing temperature and changing precipitation variability (Wada et al., 2013). All of these impacts and effects have promoted the inclusion of irrigation inside land surface models (LSMs), which represent the continental branch of the hydrologic cycle in the Earth system models (Pokhrel et al., 2016).

Besides LSMs, global hydrology models (GHMs) also represent irrigation at a global scale. Originally, GHMs were developed to assess water resource availability and water use. In GHMs, irrigation demand is equal to the increase in the ET due to irrigation (i.e., water that becomes evapotranspiration). This ET increase is estimated as the differences between crop-specific potential ET and actual ET with no irrigation (Siebert and Döll, 2010; Mekonnen and Hoekstra, 2011; Wada and Bierkens, 2014; Chiarelli et al., 2020). Following Allen et al. (2006), the crop-specific potential ET is defined as $\mathrm{ET_c} = k_\mathrm{c} \cdot \mathrm{ET_0}$, where parameter k_c depends on the crop type and growing stage, and $\mathrm{ET_0}$ is the reference crop ET, corresponding to the atmospheric evaporative demand. Some models also consider conveyance losses and return flows to rivers and aquifers; i.e., they consider the total water withdrawal (water demand plus losses), using empirical ratios of irrigation efficiency (ratio of ET increase to water withdrawal) or specific rules according to the irrigation technique (Rost et al., 2008; Jägermeyr et al., 2015). The advantage of calculating the withdrawn volume is that it allows comparison and validation with datasets of reported values, for example, the Food and Agriculture Organization (FAO) AQUASTAT dataset (Frenken and Gillet, 2012). GHMs explicitly represent water supply sources (Döll et al., 2012) and allow the estimation of non-sustainable groundwater used for irrigation (Wada et al., 2012). Some GHMs also simulate water allocation (use of water by type of source), based on rules that use information of local infrastructure and environmental flow estimations (Siebert et al., 2010; Hanasaki et al., 2008a).

LSMs do not generally use potential evapotranspiration (PET) to estimate irrigation demand. The reason is that LSMs do not deduce ET from daily PET input data but from the surface energy balance at hourly and subhourly time steps. This difference raises consistency issues between empirical PET formulas and potential ET rates in LSMs (Barella-Ortiz et al., 2013). Some LSMs prescribe irrigation rates estimated offline (Lo and Famiglietti, 2013; Cook et al., 2015), but most of the LSMs and some GHMs estimate irrigation demand by calculating a deficit such as a soil moisture deficit between actual and a target soil moisture (Haddeland et al., 2006; Hanasaki et al., 2008a; Leng et al., 2014; Pokhrel et al., 2015; Jägermeyr et al., 2015). Some LSMs, which benefit from a physically based description of surface runoff and drainage, can explicitly calculate return flow, but conveyance losses are not explicitly included (Yin et al., 2020; Leng et al., 2017). In addition, irrigation shortage due to water availability is not always well represented in those LSMs (and GHMs) that include this feature, as some of them include a virtual infinite reservoir to fulfill irrigation demand (Ozdogan et al., 2010; Leng et al., 2014; Pokhrel et al., 2012). This virtual reservoir may represent fossil groundwater use and water table depletion, which is important in some areas like the High Plains (in the USA) and India (Pokhrel et al., 2015; Leng et al., 2017; Felfelani et al., 2021). Water allocation is commonly based on a stream water supply first rule (Guimberteau et al., 2012b), with some exceptions that use the global groundwater inventory from Siebert et al. (2010) (Leng et al., 2017; Felfelani et al., 2021). These rather simple irrigation schemes are used in land surface–atmosphere simulations to assess irrigation effects on climate (Puma and Cook, 2010; Lo and Famiglietti, 2013; Guimberteau et al., 2012b; Lo et al., 2021) but not on water resources assessment.

ORCHIDEE (ORganising Carbon and Hydrology in Dynamic EcosystEms), the LSM of the IPSL (Institut Pierre-Simon Laplace) Earth system model (Krinner et al., 2005; Boucher et al., 2020), has been used to assess irrigation effects on climate. First attempts to crudely represent irrigation were based on potential evaporation and potential transpiration for a generic crop type (de Rosnay et al., 2003; Guimberteau et al., 2012b). This irrigation scheme restrains irrigation according to available water and includes simple allocation rules. Recently, ORCHIDEE-CROP, a version of the model that includes a crop phenology module, improved the irrigation scheme by representing flood and paddy irrigation techniques and was tested in an offline mode in China (Yin et al., 2020). These improvements open the possibility for assessing irrigation effects on water resources. This is important, as there is evidence that some modeling biases within ORCHIDEE in offline and coupled modes are correlated to the surface equipped for irrigation (Mizuochi et al., 2021).

Here, we present evidence on the effect of irrigation on the reduction in the modeling biases in some key variables, like ET and leaf area index (LAI), and on river discharge and total water storage anomalies (TWSAs). After describing the ORCHIDEE model and the global irrigation scheme, we set the parameter values using short simulations. We perform a sensitivity analysis and a simple parameter tuning to fit observed irrigation rates. We then perform long simulations, and we compare irrigation estimates to observations and corresponding variability due to parameter values and input maps. We validate irrigation estimates by reported values, and we assess the spatial variability in the modeling bias. Then we assess the modeling bias against observed datasets using a factor analysis, with and without irrigation, for ET and LAI. In large basins with extensive irrigation activities, we compare simulated and observed values of discharge and total water storage anomalies (TWSAs). We also show some results on the correlation between the irrigation bias and some landscape descriptors as a first step to improve the realism of the scheme. Finally, we discuss the results, and we present the main conclusions and perspectives.

2 Model description

2.1 ORCHIDEE v2.2

ORCHIDEE describes the fluxes of mass, momentum, and heat between the surface and the atmosphere (Krinner et al., 2005). Here we use version 2.2, which is close to the version used for CMIP6 (corresponding to 2.0). Version 2.0 has been described in many papers (Cheruy et al., 2020; Boucher et al., 2020; Tafasca et al., 2020), and version 2.2 only adds a few minor bug corrections. We summarize the main characteristics of the model that mediate in the simulation of irrigation.

In each grid cell, vegetation is represented by a mosaic of up to 15 plant functional types (PFTs), including generic C_3 and C_4 crops, as well as generic C_4 grasses, and tropical, boreal, and temperate C_3 grasses. The PFTs fractions are described by the LUHv2 dataset (Lurton et al., 2020), and each PFT is characterized by a specific set of parameters applied to the same set of equations (Boucher et al., 2020; Mizuochi et al., 2021). Plant phenology is controlled by the STOMATE module, which couples photosynthesis and the carbon cycle and computes the evolution of the leaf area index (LAI), and all of these processes depend on the CO_2 atmospheric concentration (Krinner et al., 2005).

A specialized version of ORCHIDEE has been proposed by Wu et al. (2016) and evaluated by Müller et al. (2017) to better describe temperate crops, with phenology thresholds based on accumulated degree days after the sowing date, improved carbon allocation to reconcile the calculations for leaf and root biomass and grain yield, and nitrogen limitation related to fertilization. It was not used in this work owing to a lack of ubiquitous parameters at the global scale, so C_3 and C_4 crops are simply assumed to have the same phenology as natural grasslands but with higher carboxylation rates and adapted maximum possible LAI (Krinner et al., 2005). The crop growing season depends on mean annual air temperature, as detailed in Krinner et al. (2005). In cool regions, it starts after a predefined number of growing degree days, while in warm regions, it starts a predefined number of days after soil moisture has reached its minimum during the dry season. In intermediate zones, the two criteria have to be fulfilled. The end of the growing season also depends on temperature and water stress and on leaf age.

Roots constitute an important link between the carbon and the water balance. In each PFT, root density decreases exponentially with depth, and the parameter that controls the decay is PFT-dependent. It is worth noting that the root density profile is constant in time and goes down to the bottom of the soil column, set at 2 m, but forest PFTs have much denser roots than crop and grass PFTs, especially in the bottom part of the soil (Wang et al., 2018). The resulting root density profile is combined with the soil moisture profile and a water stress function to define the water stress factor of each PFT on transpiration (Tafasca et al., 2020) and to estimate the water uptake for transpiration (de Rosnay et al., 2002).

Evapotranspiration is represented by a classical aerodynamic approach and is composed of snow sublimation, interception loss, bare soil evaporation (E), and transpiration (T). The first two proceed at a potential rate, while bare soil evaporation is limited by the upward diffusion of water through the soil, and transpiration is controlled by a stomatal resistance, which depends on soil moisture and vegetation parameters. The vegetation types are grouped into three soil columns according to their physiological behavior: high vegetation (eight forest PFTs); low vegetation (six PFTs for grasses and crops); and bare soil. While the energy balance is calculated for the whole grid cell (Boucher et al., 2020), a separate water budget is calculated independently for each soil column in order to prevent forest PFTs from depriving the other PFTs of soil moisture.

Vertical soil water flow is represented by a 1-D Richards equation coupled to a mass balance, and lateral flow between cells and soil columns is neglected (de Rosnay et al., 2002; Campoy et al., 2013). Here, soil depth is set at 2 m and discretized into 22 layers to finely model the lower layers implicated in drainage. Infiltration is simulated as a sharp wetting front, based on the Green and Ampt model (Tafasca et al., 2020; D'Orgeval et al., 2008). The resulting increase in top soil moisture is redistributed by the Richards equation. The bottom boundary condition assumes free drainage, equal to the hydraulic conductivity of the deepest node. The saturated hydraulic conductivity decreases with depth, but roots increase the hydraulic conductivity near the surface (D'Orgeval et al., 2008). Soil parameters are a function of soil texture (Tafasca et al., 2020), and the spatial distribution is taken from the Zobler (1986) map.

A routing scheme transfers surface runoff and drainage from land to the ocean through a cascade of linear reservoirs (Ngo-Duc et al., 2007; Guimberteau et al., 2012a). Each grid cell is split into subbasins, according to a 0.5° flow direction map. Three reservoirs are considered inside every subbasin, representing groundwater, overland, and river reservoir, and each one presents a distinct residence time (Ngo-Duc et al., 2007). The groundwater reservoir collects drainage from the soil column, while the overland reservoir collects surface runoff. Both reservoirs are internal to each subbasin and flow to the stream reservoir, which also collects streamflow from the upstream basins and contributes to large-scale routing across subbasins and grid cells. Note that there are two surface reservoirs, overland representing the headwater streams, and a river reservoir representing large rivers.

The water and energy budgets and the routing scheme are computed at the same 30 min time step, while the carbon and plant phenology processes in STOMATE are solved with a daily time step.

2.2 Irrigation scheme

The irrigation scheme (Fig. 1) is based on the flood irrigation representation from Yin et al. (2020), but it includes some changes in the parameterization to run at a global scale. The flood irrigation technique (which consists of adding water to the soil surface to achieve a certain soil moisture content) is chosen for global simulations, as it is the most used technique (Jägermeyr et al., 2015; Sacks et al., 2009).

First, the scheme defines a root zone depth in the crop and grass soil column, based on the cumulative root density (CRD), ranging from 0 at the soil surface to 1 at the soil bottom; the root zone comprises all soil layers with a CRD below a user-defined threshold, $\mathrm{Root_{lim}}$. When the threshold is set to 0.9, the root zone includes 90 % of the root system. For a 2 m soil column with 22 layers, and an exponential root density decay of 4 (default value for crops and grasses in ORCHIDEE), this threshold defines a root zone depth of 0.65 m, encompassing 11 soil layers.

We can then define a soil moisture deficit D (mm) in the root zone as the sum of the difference between actual soil moisture and a soil moisture target in all layers of the root zone, as follows:

$$D = \sum_{i \in \text{Root zone}} \max(0, \beta\, W_i^{\text{fc}} - W_i), \tag{1}$$

where W_i and W_i^{fc} (both in mm) are the actual and field capacity soil moisture in soil layer i, respectively, and β is a user-dependent parameter that controls the target value with respect to field capacity (shown in Fig. S10 with soil texture). When soil moisture drops below the target, irrigation is triggered. To prevent irrigation when there is no plant development, for example, during winter, we set the deficit D to zero if all crops and grasses are below a certain LAI threshold, $\mathrm{LAI_{lim}}$. By doing so, we overlook irrigation used to enhance germination and tend to underestimate irrigation amounts.

The irrigation requirement I_{req} (mm s^{-1}) is calculated as

$$I_{\text{req}} = f_{\text{irr}} \min(D/\Delta t, I_{\max}), \tag{2}$$

where f_{irr} is the fraction of irrigated surface (–) in the grid cell, defined by a map of irrigated fractions. The map that prescribes the irrigated fraction may change every year, but note that we do not separate the irrigated area into a separate soil column; i.e., the soil column includes crops (both irrigated and rainfed) and grasses. $I_{\max}$ is a user-defined maximum hourly irrigation rate (mm h^{-1}). This third threshold is used to avoid excessive runoff production when the deficit is larger than the infiltration capacity of the soil. Therefore, the deficit is fulfilled progressively over the subsequent time step. The effective irrigation (I; see below) is uniformly applied over the crop and grass soil column. Therefore, care must be taken by the model that the irrigated fraction is not greater than the soil column. If the fraction of the irrigated surface is much smaller than the crop and grass soil column, then irrigation will eventually be spread over a larger area than the actual irrigated surface. This particular case (an important difference between irrigated fraction and soil column fraction) could likely result in an overestimation of the amount of irrigation (mainly because the water put on the surface will not be sufficient to reach the soil moisture target; see Fig. S9). Besides, the fraction of irrigation water that actually evaporates could be larger than in reality. The latter could lead to an overestimation of the evapotranspiration increase, especially in areas that are energy-controlled (Puma and Cook, 2010), and an overestimation of irrigation efficiency.

Irrigation can be withdrawn from three routing reservoirs, but the effective water availability, A_{w} (mm), also depends on the facility to access surface water and groundwater, and it can be reduced to preserve environmental flows as follows:

$$A_{\text{w}} = f_{\text{sw}}\,(a_1\, S_1 + a_2\, S_2) + f_{\text{gw}}\, a_3\, S_3. \tag{3}$$

In this equation, S_j (mm) is the volume storage in each routing reservoir, with index j equal to 1, 2, and 3 for the stream, overland, and renewable groundwater (i.e., shallow aquifers that are recharged by drainage at the soil bottom) reservoirs, respectively. To prevent the complete depletion of these reservoirs, which all feed streamflow and support aquatic ecosystems, we mimic an environmental flow regulation by reducing the available volume, owing to a user-defined parameter a_j, to between 0 and 1. It is set here at 0.9 for all three reservoirs, so as to keep at least 10 % of the available water at each time step. The facility to irrigate from surface water reservoirs (S_1 and S_2) and groundwater reservoir (S_3) is accounted for by factors f_{sw} and f_{gw}, also ranging between 0 if the reservoirs cannot be used and 1 if they are fully accessible. In the present application, these factors represent the fraction of irrigated areas that are equipped for irrigation with surface and groundwater, respectively, following the global map of Siebert et al. (2010). We do not consider irrigation from nonconventional sources (e.g., wastewater and water from desalination plants). This map assumes that a grid cell is either equipped for groundwater irrigation or for surface water irrigation, so $f_{\text{sw}} + f_{\text{gw}} = 1$.

Eventually, the actual irrigation I (mm s^{-1}) is estimated at each time step by comparing I_{req}, i.e., the demand, to water

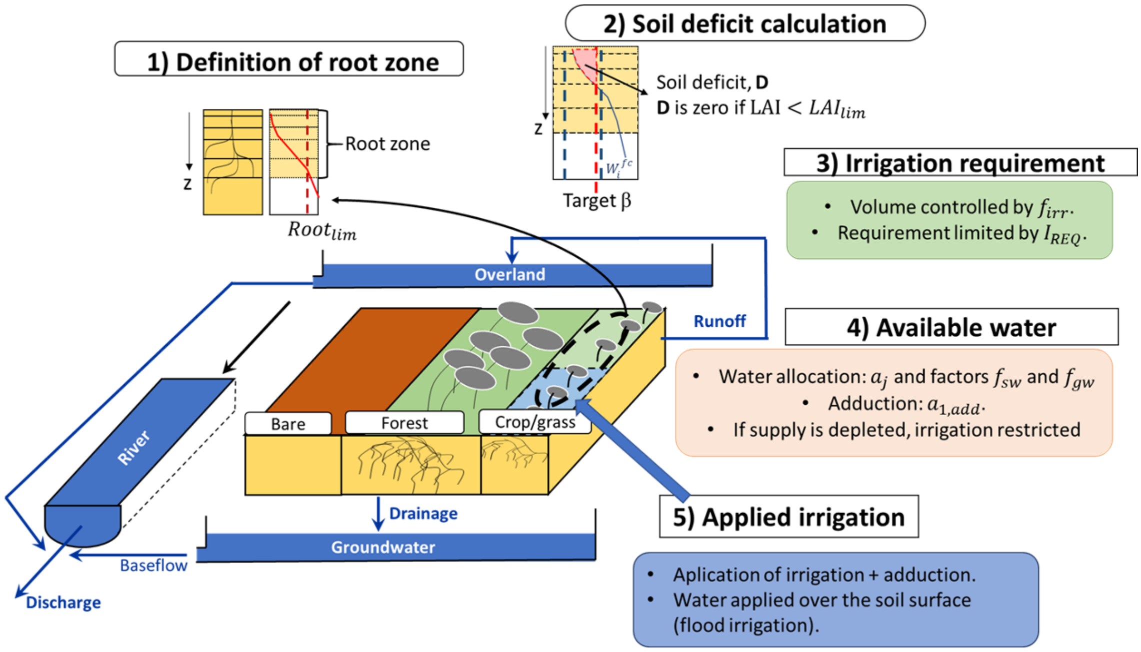

Figure 1. ORCHIDEE model and new irrigation scheme. See the text for an explanation of the parameters.

availability A_w (mm), i.e., the supply, as follows:

$$I = \min(A_\mathrm{w}/\mathrm{d}t, I_\mathrm{req}). \tag{4}$$

If we assumed that water abstraction Q_j from each natural reservoir due to irrigation withdrawal is simply proportional to the available water in each of them, then it would be given by the following equations, with the sum of the three right-hand side terms being equal to I:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = -Q_1 = -\frac{f_\mathrm{sw}\, a_1\, S_1}{A_\mathrm{w}} I \tag{5}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = -Q_2 = -\frac{f_\mathrm{sw}\, a_2\, S_2}{A_\mathrm{w}} I \tag{6}$$

$$\frac{\mathrm{d}S_3}{\mathrm{d}t} = -Q_3 = -\frac{f_\mathrm{gw}\, a_3\, S_3}{A_\mathrm{w}} I. \tag{7}$$

But we chose to implement an additional constraint for surface water withdrawals, which are withdrawn from the stream reservoir (corresponding to large rivers) at a higher priority. This new constraint leads to the definition of the revised set of equations, where the total surface water availability is $A_\mathrm{sw} = f_\mathrm{sw}\,(a_1\, S_1 + a_2\, S_2)$. as follows:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = -Q_1 = -\min\left(\frac{A_\mathrm{sw}}{A_\mathrm{w}} I\,,\ \frac{f_\mathrm{sw}\, a_1\, S_1}{\Delta t}\right) \tag{8}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = -Q_2 = -\min\left(\frac{A_\mathrm{sw}}{A_\mathrm{w}} I - Q_1\,,\ \frac{f_\mathrm{sw}\, a_2\, S_2}{\Delta t}\right) \tag{9}$$

$$\frac{\mathrm{d}S_3}{\mathrm{d}t} = -Q_3 = -\frac{f_\mathrm{gw}\, a_3\, S_3}{A_\mathrm{w}} I. \tag{10}$$

The sum of Q_1, Q_2, and Q_3 still equals I.

When $I_\mathrm{req} - I > 0$, i.e., there is a deficit and the water supply cannot satisfy the irrigation demand, then the scheme may adduct water from the neighboring grid cell with the largest streamflow volume. The choice of water adduction was introduced in Guimberteau et al. (2012b) but was disabled due to the coarse modeling resolution (grid cell larger than 100×100 km in size). Here we use a similar parameterization, but we add a user-defined parameter to take into account the facility to access distant river reservoirs:

$$\frac{\mathrm{d}S_{1,\mathrm{add}}}{\mathrm{d}t} = -Q_{1,\mathrm{add}} = -\min\left(I_\mathrm{req} - I\,,\ \frac{a_{1,\mathrm{add}}\, S_{1,\mathrm{add}}}{\mathrm{d}t}\right). \tag{11}$$

In this equation, water adduction $Q_{1,\mathrm{add}}$ from the largest river reservoir in the neighboring grid cell $S_{1,\mathrm{add}}$, will depend on the facility of access represented by the factor $a_{1,\mathrm{add}}$. This factor can range between 0 if there is no adduction and 1 if the distant river reservoir is fully accessible for water adduction.

The irrigation water, $I + Q_{1,\mathrm{add}}$, is finally added at the soil surface for infiltration, thus resembling a flood or drip irrigation technique. We note that irrigation is not restricted to an optimal period during the day but may be triggered at any moment. It may lead to an overestimation of evapotranspiration (Ozdogan et al., 2010). We do not represent dams operation in this simulation, even if they play an important role in modulating the temporal dynamics of surface water and ensure water supply for irrigation in many large river basins (Pokhrel et al., 2016; Hanasaki et al., 2008a).