\begin{document}

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

% \Author[affil]{given_name}{surname}

\Author[1]{Pedro F.}{Arboleda-Obando}
\Author[1]{Agnès}{Ducharne}
\Author[2]{Zun}{Yin}
\Author[3]{Philippe}{Ciais}

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

%% The [] brackets identify the author with the corresponding affiliation. 1, 2, 3, etc. should be inserted.

%% If an author is deceased, please mark the respective author name(s) with a dagger, e.g. "\Author[2,$\dag$]{Anton}{Smith}", and add a further "\affil[$\dag$]{deceased, 1 July 2019}".

%% If authors contributed equally, please mark the respective author names with an asterisk, e.g. "\Author[2,*]{Anton}{Smith}" and "\Author[3,*]{Bradley}{Miller}" and add a further affiliation: "\affil[*]{These authors contributed equally to this work.}".

\correspondence{Pedro F. Arboleda-Obando (pedro.arboleda_obando@upmc.fr)}

\runningtitle{TEXT}

\runningauthor{TEXT}

\firstpage{1}

\maketitle

\begin{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 on 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. It restrains actual irrigation according to available freshwater by including a simple environmental limit and using allocation rules depending 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 US (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 of 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 (TWSA) and its trends. This may be partly explained by the absence of non-renewable 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 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.

\end{abstract}

\introduction
Irrigation seeks to increase crop yields by reducing plant water stress \citep{Siebert2010a,Goldewijk2017}, and supports about 43\% of the world's food production on about 20\% of arable land \citep{Siebert2010a,Grafton2017}. The beneficial effects of irrigation on food production, population and economic growth, have dramatically pushed the increase of irrigated areas during the 20th century, from 28 Mha in 1850 to 276 Mha in 2000

\citep{Goldewijk2017,Siebert2015}. As a consequence, by the year 2000, irrigation accounted for 70\% of the total water withdrawn (between 2657 and 3594 km^3/year). 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/year, around 1.7\% of total continental ET of 75.6x10^3 km^3/year according to \citet{Jung2019}) and represents around 90\% of the total consumptive water use by human activities \citep{Pokhrel2016,Doll2012,Hoogeveen2015}.

Water abstraction and corresponding ET increase have a direct impact on the water and energy balances, and on surface and subsurface hydrology \citep{Doll2012,Taylor2013,VicenteSerrano2019}. Atmosphere reacts as well to these changes on land surface fluxes, for instance with regional increases/decreases in rainfall rate, or decreases in temperature extremes \citep{Lo2013,Guimberteau2012a,Cook2015a,Al-Yaari2019,Thiery2020}. 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 \citep{AlYaari2022}. 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 \citep{Wada2013}. All 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 \citep{Pokhrel2016}.

Besides LSMs, global hydrology models (GHMs) also represent irrigation at global scale. Originally, GHMs were developed to assess water resources availability and water use. In GHMs, irrigation demand is equal to the increase of 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 \citep{Siebert2010a,Mekonnen2011,Wada2014,Chiarelli2020}. Following \citet{AllenG.2006}, the crop specific potential ET is defined as $ET_c = k_\text{c}\cdot ET_\text{0}$, where parameter k_c depends on crop-type and growing stage, and 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), by using empirical ratios of irrigation efficiency (ratio of ET increase to water withdrawal) or specific rules according to the irrigation technique \citep{Rost2008,Jagermeyr2015}. The advantage of calculating the withdrawn volume is that it allows comparison and validation with datasets of reported values, for example the FAO AQUASTAT dataset \citep{AgriInfo2011}. GHMs explicitly represents water supply sources \citep{Doll2012} and allow the estimation of non-sustainable groundwater used for irrigation \citep{Wada2012}. 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 \citep{Siebert2010,Hanasaki2008}.

LSMs do not use in general a PET to estimate irrigation demand. The reason is that LSMs do not deduce ET from daily PET input data but from surface energy balance at hourly and subhourly time steps. This difference raises consistency issues between empirical PET formulas and potential ET rates in LSMs \citep{Barella-Ortiz2013}. Some LSMs prescribe

irrigation rates estimated offline \citep{Lo2013,Cook2015a}, but most of LSMs and some GHMs estimate irrigation demand by calculating a deficit, for instance, a soil moisture deficit between actual and a target soil moisture \citep{Haddeland2006,Hanasaki2008,Leng2014,Pokhrel2015,Jagermeyr2015}. 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 \citep{Yin2020,Leng2017}. In addition, irrigation shortage due to water availability is not always well represented in those LSMs (and GHMs) including this feature, as some of them include a virtual infinite reservoir to fulfill irrigation demand \citep{Ozdogan2010,Leng2014,Pokhrel2012}. This virtual reservoir may represent fossil groundwater use and water table depletion, which is important in some areas like US high plains and India \citep{Pokhrel2015,Leng2017,Felfelani2020}. Water allocation is commonly based on a stream-water-supply first rule \citep{Guimberteau2012a} with some exceptions that use the global groundwater inventory from \citet{Siebert2010} \citep{Leng2017,Felfelani2020}. These rather simple irrigation schemes are used in land surface-atmosphere simulations to assess irrigation effects on climate \citep{Puma2010,Lo2013,Guimberteau2012a,Lo2021} but not on water resources assessment.

ORCHIDEE, the LSM of the IPSL (Institut Pierre Simon Laplace) Earth system model \citep{Krinner2005,Boucher2020} 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 \citep{DeRosnay2003,Guimberteau2012a}. This irrigation scheme restraints 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 offline mode in China \citep{Yin2020}. These improvements open the possibility to assess irrigation effects on water resources. This is important, as there is evidence that some modelling biases within ORCHIDEE in offline and coupled modes are correlated to the surface equipped for irrigation \citep{Mizuochi2021}.

Here, we present evidence on the effect of irrigation on reduction of modelling biases in some key variables like ET and leaf area index (LAI), and on river discharge and total water storage dynamics (TWSA). After describing the ORCHIDEE model and the global irrigation scheme, we set the parameter values by 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 of the modelling bias. Then we assess the modelling 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 (TWSA). 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.

\section{Model description}
\label{sec2}

\subsection{ORCHIDEE v2.2}\label{sec2_1}
ORCHIDEE describes the fluxes of mass, momentum, and heat between the surface and the atmosphere \citep{Krinner2005}. Here we use version 2.2, which is close to the version used for CMIP6 (corresponding to 2.0). Version 2.0 has been largely described in many papers \citep{Cheruy2020,Boucher2020,Tafasca} 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 C3 and C4 crops, as well as generic C4 grasses, and tropical, boreal and temperate C3 grasses. The PFTs fractions are described by the LUHv2 dataset \citep{Lurton2020}, and each PFT is characterized by a specific set of parameters, applied to same set of equations \citep{Boucher2020,Mizuochi2021}. 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), all these processes depending on CO$_2$ atmospheric concentration \citep{Krinner2005}.

A specialized version of ORCHIDEE has been proposed by \citet{Wu2016} and evaluated by \citet{Muller2017} to better describe temperate crops, with phenology thresholds based on accumulated degree days after 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 by lack of ubiquitous parameters at global scale, so that C3 and C4 crops are simply assumed to have the same phenology as natural grasslands, but with higher carboxylation rates and adapted maximum possible LAI \citep{Krinner2005}. The crop growing season depends on mean annual air temperature, as detailed in \citet{Krinner2005}. 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 \citep{Wang2018}. 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 \citep{Tafasca} and to estimate the water uptake for transpiration \citep{DeRosnay2002}.

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 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 \citep{Boucher2020}, 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 \citep{DeRosnay2002,Campoy2013}. Here, soil depth is set to 2 meters, and discretized into 22 layers ~~here~~ to finely model lower layers implicated in drainage. Infiltration is simulated as a sharp wetting front based on the Green and Ampt model \citep{Tafasca,DOrgeval2008}. 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 \citep{DOrgeval2008}. Soil parameters are a function of soil texture \citep{Tafasca}, and the spatial distribution is taken from the \citet{Zobler1986} map.

A routing scheme transfers surface runoff and drainage from land to the ocean through a cascade of linear reservoirs \citep{Ngo-Duc2007,Guimberteau2012}. Each grid cell is split into sub-basins according to a 0.5° flow direction map. Three reservoirs are considered inside every sub-basin, representing groundwater, overland, and river reservoir, and each one presents a distinct residence time \citep{Ngo-Duc2007}. 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 sub-basins and grid cells. Note that there are two surface reservoirs, overland representing the headwater streams, and river reservoir representing large rivers.

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

\subsection{Irrigation scheme}\label{sec2_2}
The irrigation scheme (Figure \ref{fig01}) is based on the flood irrigation representation from \citet{Yin2020}, but it includes some changes in the parameterization to run at 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 \citep{Jagermeyr2015, Sacks2009a}.

Firstly, 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, 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:

\begin{equation}\label{eq4_2}
D = \sum_{i \in \text{Root zone}} \max(0, \beta \, W_i^{fc} - W_i),
\end{equation}

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, LAI_lim. By doing so, we overlook irrigation used to enhance germination, and tend to underestimate irrigation amounts.

\begin{figure*}[t]
\includegraphics[width=1.0\textwidth]{Figures/Fig01.png}
\caption[ORCHIDEE model and new irrigation scheme.]{ORCHIDEE model and new irrigation scheme. See text for explanation of parameters.}
\label{fig01}
\end{figure*}

The irrigation requirement I_req [mm/s] is calculated as:
\begin{equation}\label{eq4_1}
I_\text{req} = f_\text{irr} \, \min(D / \Delta t, I_\text{max}),
\end{equation}

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]. 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 irrigated surface is much smaller than the crop and grass soil column, irrigation will eventually be spread over a larger area than the actually irrigated surface. This particular case (an important difference in irrigated fraction and soil column fraction) could likely result in 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 Figure 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 \citep{Puma2010} 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:

$$\label{eq4_3}
A_w = f_{sw} \, (a_1 \, S_1 + a_2 \, S_2) + f_{gw} \, a_3 \, S_3.$$

In this equation, S_j [mm] is the volume storage in each routing reservoir, with index j equal 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 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, between 0 and 1. It is set here to 0.9 for all three reservoirs, such 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 \citet{Siebert2010}. We do not consider irrigation from non-conventional 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 $f_{sw} + f_{gw} = 1$.

Eventually, the actual irrigation I [mm/s] is estimated at each time step by comparing I_req, i.e. the demand, to water availability A_w [mm], i.e. the supply:

$$\label{eq4_4}
I = \min(A_w / dt , I_\text{req}).$$

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

$$\label{eq4_5}
\begin{eqnarray}
\frac{dS_1}{dt} &=& -\, Q_1 = - \frac{ f_{sw} \, a_1 \, S_1}{A_w} \, I \\
\frac{dS_2}{dt} &=& -\, Q_2 = - \frac{ f_{sw} \, a_2 \, S_2}{A_w} \, I \\
\frac{dS_3}{dt} &=& -\, Q_3 = - \frac{ f_{gw} \, a_3 \, S_3}{A_w} \, I
\end{eqnarray}$$

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

$$\label{eq4_5_bis}
\begin{eqnarray}
\frac{dS_1}{dt} &=& -Q_1 = -\min \left(\frac{A_{sw}}{A_w} \, I \,,\,\frac{ f_{sw} \, a_1 \, S_1}{\Delta t} \right) \\
\frac{dS_2}{dt} &=& -Q_2 = -\min \left(\frac{A_{sw}}{A_w} \, I - Q_1 \,,\, \frac{f_{sw} \, a_2 \, S_2}{\Delta t} \right) \\
\frac{dS_3}{dt} &=& -Q_3 = - \frac{f_{gw} \, a_3 \, S_3}{A_w} \, I
\end{eqnarray}$$

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

When $I_\text{req} - I > 0$, i.e. there is a deficit and the water supply cannot satisfy the irrigation demand, the scheme may adduct water from the neighboring grid cell with the largest streamflow volume. The choice of water adduction was introduced in \citet{Guimberteau2012a}, but was disabled due to the coarse modelling resolution (grid-cell larger than 100 x 100 km of 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:
\begin{equation}\label{eq4_6}
\frac{dS_\text{1,add}}{dt} = -Q_\text{1,add} = -\min{(I_\text{req} - I \,,\, \frac{a_\text{1,add} \, S_\text{1,add}}{dt})}
\end{equation}

In this equation, water adduction $Q_\text{1,add}$ from the largest river reservoir in the neighboring grid cell $S_\text{1,add}$, will depend on the facility of access represented by the factor $a_\text{1,add}$. 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, $I+Q_\text{1,add}$, 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 \citep{Ozdogan2010}. We do not represent dams operation in this simulation, even if they play an important role to modulate the temporal dynamics of surface water and assure a water supply for irrigation in many large river basins \citep{Pokhrel2016, Hanasaki2008}.

\section{Data description}\label{sec3}

\subsection{Input data for ORCHIDEE}\label{sec3_1}
Firstly, ORCHIDEE is run at global scale in offline mode. We run the model for the period 1970 - 2013, but we leave the first 10 years as warm-up, and we focus our analysis on the period 1980 - 2013. We use the GSWP3 \citep{VanDenHurk2016} as meteorological forcing (\url{http://hydro.iis.u-tokyo.ac.jp/GSWP3/}) with a resolution of 0.5 degrees. We also run short simulations for the sensitivity analysis and parameter tuning (see below in Section \ref{sec4_sa_tu}). We prescribe the irrigated surfaces in transient mode, i.e. irrigated surfaces may change every year, based on the Historical Irrigation Dataset (HID) from \citet{Siebert2015} and on Land Use Harmonization 2 (LUHv2) dataset from \citet{Hurtt2020}.

HID presents a map every 10 years before 1980 and every 5 years after at 5 arcmin resolution, and for each year, we use the nearest map in time to avoid data interpolation, LUHv2 presents a map every year with a 0.25 degrees resolution. The main difference between the HID and LUHv2 maps is that HID prescribes the area that is equipped for irrigation (AEI), while LUHv2 prescribes the area that is actually irrigated (AAI). As a result, the HID dataset has a greater irrigated surface (3.0 $10^6 km^2$ for HID, 2.5 $10^6 km^2$ for LUHv2 at global scale around 2000). It also means that AAI should be included in the AEI if the two datasets shared a similar spatial distribution. But this is not the case, as the two

datasets rely on different information sources, processed with different methods \citep{Oliveira2022}.

The performed simulations use uniform parameters over irrigated areas, main changes between simulations are summarized in Table \ref{table02}. As a reference, we use a simulation with no irrigation, called NoIrr, while simulation Irr with irrigation activated, uses parameter values according to results from the sensitivity and tuning analysis (Section \ref{sec4_sa_tu}) and the HID maps. The simulation Irr_NoTuned also activates irrigation and uses the HID dataset as input, but it uses a priori parameter values. This latter simulation does not consider the conclusions from the sensitivity analysis, and for instance does not activate irrigation withdrawal from the overland reservoir nor adduction.

We run additional simulations to assess the uncertainty of the simulated irrigation amount and the influence of the most sensitive parameters, according to the sensitivity analysis: the impact of the desactivation of adduction in our scheme is considered in simulation Irr_NoAdd, the effects of changes in the β value are considered in simulations Irr_NoTuned, Irr, and Irr_Beta, and finally the effect of changes in the I_{max} value is considered in simulation Irr_Imax. All these simulations use HID to prescribe irrigated areas. We analyze the effect of large differences in prescribed irrigated areas on irrigation amounts by using the LUHv2 dataset as input (simulation Irr_LUH) and using the same parameter values as the Irr simulation.

\begin{table*}
\caption{Simulations with inputs and parameter values. In brackets, the units of the parameter, [-] means that the parameter corresponds to a fraction and does not have a unit. In bold the change in parameter values with respect to the Irr simulation.}
\label{table02}
\begin{tabular}
{|>{\centering\arraybackslash}m{2.6cm}
|>{\centering\arraybackslash}m{1.9cm}
|>{\centering\arraybackslash}m{1.8cm}
|>{\centering\arraybackslash}m{1.2cm}
|>{\centering\arraybackslash}m{2.3cm}
|>{\centering\arraybackslash}m{1.6cm}
|>{\centering\arraybackslash}m{2.1cm}|}
\hline

Simulation	**Irrigation**	**Irrigated surfaces**	**β [-]**	**a_i [-]**	**I_{max} [mm/h]**	**Adduction [-]**
NoIrr	**No**	–	–	–	–	–
Irr_NoTuned	Yes	HID	**1.0**	**0.9,0.0,0.9**	**1.0**	**0.0**
Irr	Yes	HID	0.9	0.9,0.9,0.9	3.0	0.05
Irr_LUH	Yes	**LUHv2**	0.9	0.9,0.9,0.9	3.0	0.05
Irr_NoAdd	Yes	HID	0.9	0.9,0.9,0.9	3.0	**0.0**
Irr_Beta	Yes	HID	**0.75**	0.9,0.9,0.9	3.0	0.05

Irr_Imax   & Yes   & HID   & 0.9   & 0.9,0.9,0.9   & \textbf{1.0}   & 0.05 \\ \hline
\end{tabular}
\end{table*}

\subsection{Validation datasets and landscape descriptors}\label{sec3_2}
The validation of the new irrigation scheme and its effect on the model bias is focused on five variables: evapotranspiration, leaf area index, discharge, irrigation withdrawal and total water storage anomalies. We also use two landscape descriptors datasets (see below).

\medskip
\textbf{Irrigation water withdrawals -}
We use two datasets: first, we compare the simulated irrigation rates with values from the FAO-AQUASTAT database (\url{https://www.fao.org/aquastat/en/}) reported in \citet{AgriInfo2011} for irrigation volumes around 2000. AQUASTAT is based on reported values at the country scale, so it does not inform on seasonal values or their spatial distribution. In countries with a lack of information, data is completed using modelling outputs to estimate the plant requirement, and country-level ratios of irrigation efficiency to calculate the irrigation water withdrawal \citep{Hoogeveen2015}. While the plant requirement corresponds to the increase of evapotranspiration, the irrigation water withdrawal is the volume that is abstracted from the natural reservoirs, and includes the losses and return flows.

 We also use the spatially explicit information of irrigation water withdrawal around the year 2000 from \citet{Sacks2009a}. This reconstruction uses national-level census data, primarily from AQUASTAT, with maps of croplands by crop type, areas equipped for irrigation, and climatic water deficit. The result is a gridded map with a resolution of 0.5 degrees.

\medskip
\textbf{Evapotranspiration - }
We use two datasets: the first product is GLEAM v3.3a, which combines satellite-observed values of soil moisture, vegetation optical depth, and snow-water equivalent, reanalysis of air temperature and radiation, and a multisource precipitation product at 0.25⁰ resolution \citep{Martens2017}. The second dataset is FLUXCOM \citep{Jung2019}, which merges Fluxnet eddy covariance towers with remote sensing (RS) and meteorological (METEO) data using machine learning algorithms at 0.5⁰ resolution. Here we use RS+METEO products, specifically the averages of RS+METEO$_\text{WFDEI}$ and RS+METEO$_\text{CRUNCEP\,v8}$, to cover the analysis period.

\medskip
\textbf{Leaf area index - }
We use the LAI3g dataset \citep{Zhu2013} climatological values for the period 1983-2015. This dataset applies a neural network algorithm on satellite observations of the Normalized Difference Vegetation Index (NDVI) 3g to estimate LAI at 5 arcmin resolution.

\medskip
\textbf{River discharge - }

We use monthly data from the Global Runoff Data Centre (GRDC,
\url{https://www.bafg.de/GRDC-/EN/Home/homepage_node.html}) in 14 large basins with
strong irrigation activities. We choose the station nearest to the river mouth that also has
data available for the study period (Figure S4 shows the basins and its corresponding
discharge station). Basin boundaries were delineated with the flow directions map used by
ORCHIDEE (section \ref{sec2_1}).

\medskip
\textbf{Total water storage anomalies - }
We compare the total water storage anomalies (TWSA) from our simulations with three
different monthly products of TWSA based on GRACE (Gravity Recovery and Climate
Experiments) observations based on global mascon solutions, that are suitable for
hydrologic applications \citep{Scanlon2016a}: CSR \citep{Save2016}, GRC Tellus, called
here TELLUS \citep{Emry2014} and NASA GSFC \citep{Loomis2019}. CSR has a spatial
resolution of 0.25 degree, while TELLUS and NASA GSFC have a resolution of 0.5 degrees.
As the differences between products at the large river basin scale are small, we use the
average value of the three products. All the products cover the period from april 2002 to the
end of the simulation in 2014.

\textbf{Landscape descriptors -}
We compare the simulation results with two landscape descriptors which are linked to
irrigation and may contribute to the irrigation bias. We use the fraction of irrigated rice
around year 2000 from MIRCA2000 \citep{Portmann2010} (see spatial distribution and focus
on Southeast Asia in Fig. S5), and the location and volume of major dams based on the
Global Reservoir and Dams dataset, GRanD \citep{Lehner2011}. GRanD contains
information on the maximum storage capacity and main use of dams with reservoirs larger
than 0.1 ha. Here we consider dams that have irrigation as their main purpose.

\subsection{Data processing and analysis}\label{sec3_3}
We aggregate and interpolate all the observed data to the 0.5 degrees spatial resolution of
the ORCHIDEE simulations. For ET, we mask GLEAM and the simulated data according to
FLUXCOM, which does not cover all the continents, so all the comparisons are made over
the same grid cells with available information. For LAI, we exclude grid-cells with no data in
LAI3g from the analysis. We compare grid cell values and zonal average values. The
statistical significance of the mean difference between observed and simulated time series is
assessed with a Student's t-test at the 5 \% significance level.

We use the simulated discharge from the grid cell that best matches the watershed area
upstream of the discharge station. In addition, we only use time steps with data available
from observations, so that both time series agree. For TWSA, we compare observed and
simulated basin averages. As ORCHIDEE gives the total water storage (TWS) value, we
normalize the time series with the mean value of the NoIrr simulation for the period 2002-
2008, the same as the observed products. In this way, the effect of irrigation over TWS is
observed in the simulated time series.

In addition to direct comparison at the grid cell, zonal or basin scale, we perform a factor
analysis to reveal relationships between modelling bias and landscape descriptors. We use
the fraction of irrigated areas around 2000 from HID, as well as the fraction of irrigated rice

from MIRCA2000, both interpolated to the ORCHIDEE resolution. We categorized grid cells into six classes by irrigation fraction levels based on the two datasets, following \citet{Mizuochi2021}: Class 1: 0\%, Class 2: 0 to 5\%, Class 3: 5 to 10\%, Class 4: 10 to 20\%, Class 5: 20 to 50\% and Class 6: 50 to 100\%.

We also performed a comparison between the average basin-scale irrigation bias and the volume capacity of dams used for irrigation within the basin, according to GRanD. We use the Pearson's correlation coefficient (r) as a metric for the correlation analysis.

\section{Sensitivity analysis and parameter tuning}\label{sec4_sa_tu}

\subsection{Sensitivity analysis}
Short simulations were run to assess the sensitivity of the irrigation amount at the global scale to different parameter values, assumed to be uniform in all irrigated areas. We used GSWP3 as meteorological forcing, and the LUHv2 from \citet{Hurtt2020} to prescribe the irrigated surfaces (see Section \ref{sec3_1}). We ran a total of 23 simulations with varying parameters, plus a reference simulation with no irrigation. All of them were run with the same initial conditions for three years (1998 - 2000), and a comparison of irrigation amount and ET increase was performed for the year 2000. By using the last simulation year, we reduce the effect of the common initial conditions on the simulation results, and the year 2000 corresponds to the values given in AQUASTAT and \citet{Sacks2009a}. Note that we use a single meteorological forcing dataset and compare our estimates to a single set of observed AQUASTAT data for the period around the year 2000. We choose to compare our estimates for the year 2000 because this year is commonly used as the reference period in the literature concerning the estimation of the amount of irrigation on a global scale (see, e.g., Pokhrel et al., 2016, Table 2). The choice of the year 2000 is mainly due to the existence of more complete reported or observed values for that year, as well as simulated estimates. We use the same reference period to compare our results with independent data. A brief description of each parameter as well as the unit, range, and values used in the sensitivity analysis is shown in Table \ref{Table01}.

\begin{table*}
\caption{Parameters of the irrigation module, brief description, range and values used in the sensitivity analysis. Values in bold correspond to the reference value. Note that, for parameter a_i, the three reservoirs share the same value.}
\label{Table01}
\begin{tabular}{
 | >{\raggedright\arraybackslash}m{2.1cm}
 | >{\raggedright\arraybackslash}m{6.4cm}
 | >{\raggedright\arraybackslash}m{2.0cm}
 | >{\raggedright\arraybackslash}m{1.4cm}
 | >{\raggedright\arraybackslash}m{2.6cm} | }
 \hline
\textbf{Parameter} & \textbf{Description} & \textbf{Unit} & \textbf{Range} & \textbf{Values} \\ \hline
β & Controls the soil moisture target, equal to $\theta_{fc}*\beta$ & Fraction, no units & {[}0-∞{]} & 0.6, 0.8, \textbf{0.9}, 1.0, 1.2, 1.4 \\ \hline

a_j & Controls the fraction in reservoir available for irrigation, the complement being the fraction left for ecosystems & Fraction, no units & {[}0-1{]} & 0.1, 0.5, \textbf{0.9}, 1.0 \\ \hline
$a_{1,add}$ & Controls the fraction in stream reservoir available for adduction & Fraction, no units & {[}0-1{]} & 0.0, \textbf{0.05}, 0.1, 0.2 \\ \hline
I_{max} & Maximum irrigation rate per hour & mm/h & {[}0-∞{]} & 0.5, 1.0, \textbf{3.0}, 5.0, 7.0 \\ \hline
$Root_{lim}$ & Defines if a soil layer is part of the root zone & Cumulated relative root density & {[}0-1{]} & 0.0, 0.5 \textbf{0.9}, 1.0 \\ \hline
LAI_{lim} & Minimum LAI in crops and grasses PFTs to trigger irrigation & m\textasciicircum{}2/m\textasciicircum{}2 & {[}0-∞{]} & 0.0, \textbf{0.1}, 0.3, 0.5, 1.0 \\ \hline
\end{tabular}
\end{table*}

We change the value of one parameter at a time (once-at-a-time screening, see \citet{Mishra2009a,Song2015b}) and then we observe its effect on the irrigation rate and on the increase in evapotranspiration. We tried to include the full range of parameters, but it is worth noting that in some cases, values were restricted to ensure an expected behavior. In the case of β, we set values around 1.0 (target equal to the field capacity soil moisture) as it seems a plausible target for flood irrigation, but note that values higher than 1.4 or lower than 0.6 are possible. Theoretically the upper limit is infinite, but values above 1.5 may exceed the saturated soil moisture for some soil textures, the lower limit is zero (see Table S4). For adduction, we set parameter values under 0.2 (20\% of streamflow available for adduction at every time-step), which seem high enough to represent water adduction in large river basins \citep{Leng2015}. In the case of the LAI$_\text{lim}$ and I_{max}, upper values were selected a priori. Those values in bold in Table \ref{Table01} are called reference values afterwards. The reference parameter values are intended to maximize the irrigation amount, as preliminary tests (not shown) performed with a priori values exhibited an underestimation of irrigation rates at global scale. The reference values do not change if not explicitly required by the once-at-a-time screening method.

\begin{figure*}[hb!]
\includegraphics[width=16cm]{Figures/Fig02.png}
\caption{Sensitivity of global irrigation volumes and increase of evapotranspiration (km^3) to changes in parameter values, for the year 2000 using short simulations. Secondary y-axis correspond to ET increase values compared to the simulation with no irrigation. Note that the y-axis scales differ between parameters.}
\label{fig02}
\end{figure*}

Figure \ref{fig02} shows that β is the parameter with the strongest effect on the global mean irrigation rate, followed by the cumulated root density threshold Root_lim, I_max, and the fraction of stream storage available for adduction $a_\text{1,add}$. The fraction of water storage left for the ecosystems (called Environmental in the figure, a_j) has a more limited effect, suggesting that in many irrigated areas, there is enough water from surface and groundwater to fulfill the irrigation requirements. Finally, the LAI limit, LAI_lim, to trigger irrigation has a weaker effect

than the other parameters. In the case of ET increase (Fig. \ref{fig02}, blue line), the sensitivity to the different parameters is similarly hierarchized, although the magnitude is not necessarily the same. Also note that the effect on irrigation efficiency (i.e. ratio of ET increase to irrigation amount) is different for β values higher than 1.0 and for Root_lim values higher than 0.5. This implies that the fraction of irrigation water that becomes runoff or deep drainage is more important.

\subsection{Parameter tuning}

\begin{figure*}[hb!]
\includegraphics[width=12cm]{Figures/Fig03.png}
\caption{Calibration of β value with \citet{Sacks2009a} dataset as observed value, using outputs from the short simulations. Bias in total irrigation volume in km^3 by β value (a), boxplot of the bias of irrigation rates in gridcells in mm/d by β value (b).}
\label{fig03}
\end{figure*}

The sensitivity analysis showed that β has the strongest effect on the simulated irrigation amount, and that these effects can induce changes on the irrigation efficiency. Therefore, we explored in more detail its behavior to set an value. Note that we used the chosen reference values for the other parameters. We compared the irrigation rate estimated by ORCHIDEE for 2000 in the short tests with the observed irrigation from \citet{Sacks2009a} (Fig. \ref{fig03}) using total irrigation volume at global scale, and irrigation difference at grid-cell scale. When comparing the irrigation water amount at global scale (in km^3 for the year 2000, Fig. \ref{fig03}-a) we observe that a value of 1.2 maximizes the irrigation and minimizes the irrigation bias. When we assess the distribution of bias using grid-cell values (in mm/d, Fig. \ref{fig03}-b) we observe that for β equal to 0.8, 0.9, or 1, the bias distribution is centered around 0, while it starts to move up for values 1.2 and 1.4. This behavior can be slightly different depending on the irrigated fraction (see Fig S9). For simplicity here and as a tradeoff between the underestimation of irrigation volume and the spatial distribution of bias, we choose to set β to 0.9 in all irrigated areas. We decided to use the reference values for the other parameters, as they play a minor role according to the sensitivity analysis, and the reference values does not minimize the irrigation amount.

After this analysis we underline four points. First, this process does not correspond to a proper calibration, as we assumed uniform parameter values, the number of simulations is low and the observed data is sparse. The objective of the sensitivity analysis and parameter tuning was to identify key parameters and reduce the underestimation of irrigation by tuning the uniform parameter values. Second, our scheme does not include conveyance losses although application losses and return flows are represented. As ORCHIDEE determines the water partitioning, some model flaws in hydrologic processes like infiltration or bare soil evaporation could bias the effect in return flows, in the increase of ET, and ultimately in the irrigation efficiency. Third, although the once-at-a-time method is suitable given the computational cost of running an ORCHIDEE simulation, it also has drawbacks and limitations in its analysis \citep{Song2015b}, for instance its qualitative nature, and lack of quantification of individual interaction between parameters. Fourth, we use the LUHv2 map, which represents the areas actually irrigated, AAI (a lower value than the areas equipped for

irrigation, AEI, which is used in other datasets). We do not consider the effect of prescribed data uncertainty nor the effect of meteorological forcing in this analysis.

\section{Results}\label{sec4}

\subsection{Validation of irrigation water withdrawals}\label{sec4_1}
Irrigation from the Irr simulation is estimated at 0.049 mm/d (2452.5 km^3/year) around the year 2000 (Fig. \ref{fig4}-a). This estimation is in the lower part of other studies which range between 2465 and 3755 km^3/year \citep{Pokhrel2016} and is lower than AQUASTAT estimation of 2735.1 km^3/year around the year 2000 \citep{AgriInfo2011}. The results suggest that the proposed scheme is adequate to simulate the reported estimations of irrigation despite the underestimation (-10\% than the 2735.1 km^3/year from AQUASTAT around 2000). We note that this estimate is also higher than that of the old irrigation scheme of \citet{Guimberteau2012a}, but the scheme proposed here can still benefit from a more robust parameter tuning.

\begin{figure*}[!h]
\includegraphics[width=1.0\textwidth]{Figures/Fig04.png}
\caption{Total water withdrawal for irrigation in the Irr simulation, yearly average for 1998-2002 (a), groundwater irrigation withdrawn for irrigation, yearly average for 1998-2002 , difference in water withdrawn for irrigation between Irr (yearly average, 1998 - 2002) and AQUASTAT values \citep{AgriInfo2011} at country level in km3/year (b), difference in water withdrawn for irrigation between Irr (yearly average, 1998 - 2002) and dataset from \citet{Sacks2009a} (c).}
\label{fig4}
\end{figure*}

At the country-scale, Fig. \ref{fig4}-b shows that the irrigation module underestimates water withdrawals in the main hotspots of irrigation, i.e. India, China, and the USA, while it overestimates irrigation rates in Africa, East Europe and Latin America (see also Fig. S1). Such reduced contrasts between highly and weakly irrigated countries could indicate a limitation of the irrigation scheme to represent local irrigation strategies, as our scheme uses global uniform values for all the parameters. Comparison with the estimation from \citet{Sacks2009a} (see Fig. S1) supports this result (-0.004 mm/d, -219.2 km3/y, Fig. \ref{fig4}-c) and allows us to identify the areas where the irrigation bias is the strongest. In India, the Indus basin presents a strong underestimation, as well as in the Northern part of the Ganges-Brahmaputra basin. In China, there is a more widespread underestimation. That is also the case in the west part of the US Great Plains. The other regions present in general an overestimation of irrigation withdrawals, which is especially important in some small areas in Africa, in Eastern Europe and north to the Caspian sea, and in some areas of central Asia. Finally, we note that within a country, it is possible to observe areas with positive and negative bias, for instance in the USA or India. This could also be partially explained by the use of globally uniform values, as there could be important local differences on irrigation strategies within the same country, and it remarks the need to assess the irrigation bias at different scales (See Fig. S11).

\subsection{Variability of the irrigation rates due to parameter values and input data}\label{sec4_2}
The global annual irrigation volumes (Fig. \ref{fig5}-a) show a large uncertainty across the simulations due to changes in the parameter values (for instance, -24.7\% between Irr_NoTunned and Irr) , but note that the change on irrigation rates at gridcell scale can have a strong spatial heterogeneity within a country (\ref{fig5}-c) for instance in India or the USA. The parameter set used in the Irr simulation manages to increase the irrigation rate and to markedly reduce the irrigation bias when compared to the Irr_NoTuned simulation at global scale, even if locally we may observe both an increase or a decrease on the irrigation rate in the same country, for instance in China (with a marked north-south difference, (\ref{fig5}-c, Irr_NoTuned-Irr) or the Indus river basin in Pakistan and India (see (\ref{fig5}-c, Irr_NoTuned-Irr). Also, a positive trend in the annual irrigation volume is observed in all simulations. It is caused by the increase in irrigated area, observed in both HID and LUHv2 datasets (see simulations Irr and Irr_LUH). The irrigated area has been identified by \citet{Puy2021} as the main driver of irrigation water withdrawal, and the increase of the prescribed irrigated area in the simulations explains in part the positive trend in the irrigation rate (see Fig. S6).

\begin{figure*}[p]
\includegraphics[width=1.0\textwidth]{Figures/Fig05.png}
\caption{Time series of globally averaged irrigation rates simulated by ORCHIDEE (a). Map of the standard deviation of mean irrigation rates from all simulations in mm/d (b). Maps of mean difference between Irr simulation and others, for the period 1980 - 2013 in mm/d (c). Blank areas correspond to grid-cells with no irrigated areas}
\label{fig5}
\end{figure*}

Based on the mean annual values (Fig. \ref{fig5}-c), the β parameter has the largest effect on the mean irrigation rate (-22.3\% when β decreases from 0.9 to 0.75), followed by the change of input map from HID to LUHv2 (-19.7\%), a lower I_{max} (-16.5\%) and, finally, no adduction (-15.7\%). From a spatial point of view, the overall reduction in irrigation due to the above changes is not homogeneous and large areas may even display an increased irrigation rate. The exception is the β parameter, which shows an overall reduction in irrigation with a lower parameter value, except in the Indus basin. The Indus river basin is a region that depends on both surface and groundwater for irrigation, and the irrigation demand is one of the most important worldwide \citep{Laghari2012}. In Irr_Beta, a lower β induces a reduction of water demand in the upper areas of the Indus river basin, increasing the river discharge downstream. More surface water supply in the middle and lower parts of the basin can increase irrigation in these areas, even if the water demand also decreases, because the irrigation deficit i.e. the difference between demand and supply, is still high despite the demand reduction. We advance here that the propagation of water supply through the river system can explain part of the heterogeneity in the response to other parameter changes.

To estimate the interannual irrigation rate variablity, we calculate the coefficient of variation (ratio of standard deviation to the mean, Fig. \ref{fig5}-b). We use the pluriannual mean irrigation rate from all simulations with irrigation activated from table \ref{table02}. Results

highlight a certain homogeneity, but we can identify at least two distinct areas: an area of low variability (around 0.25) in Southern Asia, some areas in the Mediterranean and North and South America, and an area of high variability (around 0.75) in Northern Europe, North America, Africa and Australia, with some points where the coefficient of variation is the highest (over 2.0), especially in Africa. The use of global parameter values could explain the relative homogeneity of the coefficient of variation, while regional differences like climate variablity and irrigation water demand could explain the existence of these two variability classes.

\subsection{Factor analysis: correlation of modelling biases and irrigation classes}\label{sec4_3}
Figure \ref{fig06}-a shows the bias of ET by class of irrigated fraction at grid-cell scale, when we compare ORCHIDEE simulations with FLUXCOM dataset. It shows that the activation of irrigation reduces the ET bias in those areas with high irrigation fractions (also see S7 for the spatial distribution). For the comparison with GLEAM in Fig. \ref{fig06}-b, it shows that the activation of irrigation induces a positive bias in those areas with irrigation. When comparing absolute ET values by irrigation class (Fig. \ref{fig06}-c), we observe that NoIrr and GLEAM are similar in all the classes except for 0 and All, i.e. no irrigated fraction and all grid cells. It means that the differences between NoIrr and GLEAM comes from non-irrigated areas. This could suggest a limitation in GLEAM to represent the effects of irrigation on ET rates as this product does not respond to the presence of irrigated areas. On the other hand, Irr and FLUXCOM boxplots are similar for classes 10-20, 20-50, and 50-100. A seasonal assessment on zonal average values for irrigated areas support this suggestion (Fig. S7). Thus, from now we prioritize FLUXCOM for our analysis regarding the ET bias.

\begin{figure*}[p]
\includegraphics[width=1.0\textwidth]{Figures/Fig06.png}
\caption{Factor analysis of ET bias with FLUXCOM against irrigated fraction classes (a), and with GLEAM (b). Mean ET values of simulations and observed products against irrigated fraction classes (c). LAI bais with LAI3g against irrigated fraction classes (d) and mean LAI values of simulations and observed product against irrigated fraction classes (e).}
\label{fig06}
\end{figure*}

A similar analysis for the LAI bias and classes of irrigated fraction (Fig \ref{fig06}-d) shows an increase in the LAI difference between ORCHIDEE and LAI3g in the Irr simulation. Also, for all classes, the positive bias in the NoIrr simulation is exacerbated in the Irr simulation, except for the most intensively class (class 50-100), which reduces the negative bias when comparing NoIrr and Irr simulations. It is worth noting that the class 50-100, where irrigation is more important, is the single one with a negative bias in NoIrr, and this negative bias is partially reduced when irrigation activities are included (see Fig. S8 for spatial distribution and zonal average values). This is due to less water stress and thus more photosynthesis and biomass production, which is coherent with the decrease of ET bias for this class. When comparing absolute values between the simulations and the observed product (Fig. \ref{fig06}-e), we observe that irrigation activation within ORCHIDEE does not significantly change the distribution of LAI values at global scale. These results are coherent with

changes that irrigation induces on water fluxes and reservoirs, as well as on water and energy budget (see Fig. S2 and S3).

\subsection{Effect of irrigation on TWSA and river discharge}\label{sec4_5}
We now focus on the average TWSA value at the basin scale (Fig. \ref{fig09}). Activation of irrigation induces small changes in TWSA, which is coherent with changes in TWS between both simulations (Figure S2). For instance, we observe higher peaks in Huang He when irrigation is activated. Low values also become lower for the Irr simulation in Huang He basin. In the Ganges river basin, low values are lower as well in the Irr simulation than in NoIrr. The changes in water pathways and related residence times that explains changes in TWS between Irr and NoIrr (transfer of water from a reservoir with rapid flows like the streamflow to the soil, with a slow flow), also could explain these changes in TWSA dynamics at large basin scale. Other basins, like the Nile river basin or the Amu-Darya, show little effect between both simulations, even if extreme peaks values can be overestimated (during 2007 in the Nile) or underestimated (during 2005 and 2006 in Amu-Darya). But note that the model is unable to follow the GRACE trends, in basins with negative trends (for instance Huang He, Indus or Ganges) or positive trends (Murray river between 2011 and 2014).

\begin{figure}[p]
\centering
\includegraphics[width=1.0\textwidth]{Figures/Fig09.png}
\caption{Comparison of TWSA between ORCHIDEE simulations and GRACE datasets in large basins with strong irrigation activities.}
\label{fig09}
\end{figure}

The correct simulation of river discharge (Fig. \ref{fig10}) is another challenge in ORCHIDEE and other LSMs \citep{Oki1999,Ducharne2003,Guimberteau2012,Koirala2014,Cheruy2020}. Irrigation plays an important role to reduce the average values when we compare NoIrr and Irr simulations (see for example the Nile, or the Indus rivers, these results are coherent with those from \citet{DeGraaf2014}). The main effect of irrigation over the seasonal variations is that peak discharge can occur before in the Irr simulation (for instance Missouri river or Yellow river), or that the decrease after the peak is more rapid and low values are lower in Irr than in NoIrr (for instance Colorado river or the Danube river). These changes are due to the triggering of irrigation during spring and summer, and the corresponding ET increase. It is worth noting that the Irr simulation does not necessarily reduce the discharge bias against GRCD data compared to the NoIrr simulation, with the exception of the Danube river (see Table S1 for some goodness-of-fit metrics for observed and simulated discharge values).

\begin{figure}[p]
\centering
\includegraphics[width=1.0\textwidth]{Figures/Fig010.png}
\caption{Comparison of observed and simulated river discharge in large basins with strong irrigation activities.}

\label{fig10}
\end{figure}

\subsection{Factor analysis: correlation of irrigation biases and landscape descriptors}\label{sec4_6}

We compare biases and errors in irrigation estimates with landscape descriptors that could help explain these modelling errors. We also seek a perspective to increase the realism of the irrigation scheme and reduce the error in irrigation estimation. For the irrigation bias, classes with a high fraction of irrigated paddy rice, (for instance class 20-50 or 50-100), exhibit a higher bias than classes with small fractions (Fig \ref{fig11}-a). The spatial distribution of irrigated paddy rice is concentrated in southeast Asia, and includes the most irrigated river basins worldwide (see Fig. S5). At the large basin scale (see values in Table S2), the irrigation bias also correlates well with the capacity of dams used for irrigation (Fig. \ref{fig11}-b) if we retire a single outlier corresponding to the Nile river basin (r value without the outlier is -0.55).

\begin{figure}[!h]
\centering
\includegraphics[width=16cm]{Figures/Fig011.png}
\caption{Factor analysis of irrigation rate bias with data from \citet{Sacks2009a}, against irrigated paddy rice classes (a). Basin average value of irrigation bias against dams capacity (b). Map of β that minimizes the irrigation bias, according to the short simulations (c) and corresponding map of minimum irrigation bias according to the β value in (c), in mm/d (d).}
\label{fig11}
\end{figure}

The correlation between paddy rice and irrigation bias suggests the need to explicitly represent paddy irrigation at global scale. Thus, we add an assessment of the β value and irrigation bias, using the short simulations used on the parameter tuning (see section \ref{sec4_sa_tu}). We use all simulations with changes on the β parameter from Table \ref{Table01}. Then, we build a composite map of the β value that minimizes the irrigation bias at a grid-cell scale (Fig. \ref{fig11}-c) and then we show the corresponding irrigation bias as compared to \citet{Sacks2009a} dataset. The results roughly show at least two classes for β: the first with values of 1.2 and 1.4 (for instance in China and north India) and the second with values of 0.6. Using at least two β values is not enough to reduce the irrigation bias at global scale, but it has an important effect on the spatial distribution of the irrigation bias in Southern Asia, the region with the most paddy rice area. These results suggest that the β parameter should have at least two values, 1.3 in areas with paddy rice, and 0.6 in the rest of irrigated areas. But note that the data used for this analysis correspond to a single year, i.e. year 2000. Also, regional characteristics, like more than one harvest of paddy rice due to optimal climate conditions, are not taken into account in this analysis, but could also help to explain the irrigation underestimation in our estimations \citep{Yin2020}.

\section{Discussion}\label{sect5}

In this study, we implemented a new global irrigation scheme inside the ORCHIDEE land surface model based on previous work from \citet{Yin2020} in China. While we found a reduction in some modelling biases when irrigation is activated, we also identified at least four types of limitations in our modelling framework that can affect the estimates of irrigation or the effects of irrigation on other variables inside the land surface model:
\begin{enumerate}
\item The irrigation scheme exhibits some shortcomings that may bias the estimated irrigation amount: the use of a single irrigation technique; simplified rules to trigger irrigation and allocate the available water; the joint representation of rainfed and irrigated crops within the same soil column; the non-representation of conveyance losses, although losses due to return flows are represented.
\item The parameter tuning is overly simplistic. As a first step, we considered globally uniform parameters, which is overly simplistic, although spatially distributed values would allow us to better describe the local features of irrigation systems, as shown by the spatial variations in optimized β map, and the dependence of the local irrigation bias on the fraction of paddy rice.
\item We also use a single meteorological forcing dataset and a single year to characterize observed irrigation values. This contributes to biasing the parameter adjustment process by taking uncertain data (meteorological forcing and reference irrigation) as certain.
\item The ORCHIDEE model exhibits many uncertainties that are not related to the irrigation scheme, but ultimately impact the irrigation withdrawals and efficiency (defined here as the ratio of additional ET due to irrigation to water withdrawal) and the temporal dynamics of irrigation. One particular uncertainty comes from the overestimation of bare soil evaporation \citep{Cheruy2020}, that we presently try to correct in ORCHIDEE. Other uncertainties result from the inherent simplifications of any model. In ORCHIDEE, they include the use of a single soil texture in each grid cell, of only two kinds of crops with simplified phenology and crop calendars, and the choices made to simulate infiltration and evaporative processes.
\end{enumerate}

These shortcomings and limitations could induce positive or negative biases in the simulated regional irrigation amounts; this as a result of differences in regional landscape, hydro-climatic conditions and local irrigation practices not well represented or absent in our scheme. For example, the missing representation of paddy irrigation induces under-irrigation in paddy rice areas, the joint representation of rainfed and irrigated crops induces over-irrigation in areas with other crop types and irrigation techniques, and the simplistic parameter tuning could tend to minimize the overall net bias, while increasing regional biases. These limitations (some shared with other global LSMs) call for further model developments that aim at a better representation of the water supply (fossil groundwater and water adduction to list two mentioned in the results) and the water demand (a separate water budget for irrigated areas, the inclusion of other irrigation techniques, new irrigation rules such as irrigation before sowing or interruption of irrigation before harvest). In addition to the improvements noted here that focus on model developments, the irrigation representation can be improved by using new input datasets and regional parameter values to include local practices (if these datasets exist at the coarse model resolution in the global domain, and for historical period or future scenarios). For instance, to prescribe regional β values, or to prescribe the start and end of the growing season.

The model estimates the irrigation water demand by calculating a soil moisture deficit according to a user-defined soil moisture target. Besides, it constrains the actual irrigation rate by the available water supply. The water supply takes into account the facility to access surface or underground water sources according to local infrastructure, and environmental restrictions. Note that this environmental restriction is a simplification compared to the complex methods used in the real world to estimate environmental flow requirements, and other more robust approaches exist (for instance in \citet{Hanasaki2008}, providing monthly environmental flow requirements). Strict environmental requirements could reduce the surface water supply, thus the irrigation rate \citep{Hanasaki2008a}.

For the facility to access the water sources, we use two static factors based on local infrastructure, while water allocation is dynamic and can change according to water availability \citep{DeGraaf2014} as well as economic and societal aspects \citep{DOdorico2020}. The irrigation scheme also allows the adduction of water from neighboring grid cells, which can be important in areas of China and India \citep{Laghari2012,Yin2020b}, where surface water is intensively used. This representation of water adduction, however, is very simple, and could be improved by including human water management and dams operation, as in \citet{Zhou2021a}, where the supply and demand network is operated as a system, taking into account some constraints like topography and environmental flow \citep{hanasaki2018}.

On the water demandFinally, we observed that the conditions to trigger and stop irrigation, although controlled by four parameters, may seem too simple in our scheme, especially compared to specialized irrigation models, the new irrigation scheme in LSM CLM5 \citep{yao2022}, which includes multiple irrigation techniques, or the ISBA LSM \citep{Druel2022}, which implement complex sets of rules to represent different irrigation strategies. Some rules could change the moment when irrigation is triggered and increase the amount (for instance allowing irrigation some days before the crop emergence) or decrease itthe irrigation amount (for instance, preventing irrigation during maturity of the crop, shortening the growing season, or preventing continuous irrigation during more than a certain number of days). Implementing these sets of rules for irrigation strategies in ORCHIDEE is feasible, for instance the definition of the growing season (with trigger of irrigation before sowing and stop before harvesting) could be based on the prescription of start and ending dates as done by Yin et al., 2020, or could use the phenology information simulated by the model (as in the version used here, or using a crop-specific module as in Wu et al., 2016). But defining the set of rules and parameter values would need a careful tuning and evaluation process, with local data at sub-yearly scale.

Despite thesethis limitations, the evaluated irrigation scheme produces acceptable estimations of yearly irrigation withdrawals on a global mean basis, but it underestimates irrigation volumes in areas of China, India and the US (the most irrigated areas). Our estimations are affected by the uncertainty on global parameter values assumed to be uniform, and on the map of irrigated fractions \citep{Puy2021}. We show that the lack of paddy rice irrigation could contribute to the underestimation of irrigation in southern Asia, as the paddy technique needs the inundation of the field and maintains a saturated soil at least

during 80\% of the crop duration \citep{DeVrese2018}. The irrigation module of LSM MATSIRO, called MAT-HI and HiWG-MAT \citep{Pokhrel2012, Pokhrel2015}, already implemented an explicit representation of paddy rice irrigation, by setting a higher soil moisture target for rice than for other crops. An explicit paddy representation was also implemented in ORCHIDEE-CROP \citep{Yin2020} at a regional scale, by implementing a pond for paddy rice and using a water level target, but it uses detailed crop information not easy to access at global scale. A surrogate approach in our simpler irrigation scheme could be to use at least two β values, one for paddy rice and another one for other crops, as suggested by the composite map of β values minimizing the irrigation bias.

An outcome of our study is to reveal that the GLEAM values do not exhibit a significant sensitivity of ET to the presence of irrigated areas. This suggests that GLEAM is not suitable for estimating ET rates in irrigated areas. For instance, coupled simulations using CLM4 in northern India showed as strong modelling underestimation of ET rates, even with no irrigation \citep{Fowler2018}. When we compare the simulations with the FLUXCOM product, the activation of irrigation leads to a reduction of the negative evapotranspiration bias, but the use of a single soil column in ORCHIDEE for both rainfed and irrigated crops could induce an overestimation of ET increase (See Fig. S11, in some cases the irrigation efficiency by country is too high). The ET bias improvement is particularly substantial in heavily irrigated areas, where the simulated LAI is also improved by irrigation (which reduces there the negative LAI bias). These results show the benefits of including an irrigation scheme to partially reduce some modelling biases, especially in intensively irrigated areas, and are coherent with the multivariate evaluation of ORCHIDEE done in \citet{Mizuochi2021}.

ET and LAI are two important drivers of land-atmosphere coupling via water, energy and momentum transfer \citep{Seneviratne2010,Greve2019}, but there is evidence that the effects on ET and LAI due to human land-cover change and landscape management are not monotonic \citep{Sterling2013}. The sensitivity of these drivers to irrigation calls for further studies in coupled mode to explore the joint evolution of climate, land surface fluxes, and the use of water resources. Some studies focuses on the effects of irrigation on climate and land surface fluxes for the historical climate \citep{Boucher2004,Sacks2009a,Puma2010,Guimberteau2012a,Cook2015a,Thiery2017a,Al Yaari2022}, but to the best of our knowledge, that is not the case for the future climate under different scenarios.

In contrast to the effects on ET and LAI, the effect of irrigation on land surface hydrology is rather weak. For discharge, activation of irrigation logically reduces river discharge, because of surface and groundwater withdrawal for irrigation. This reduction does not necessarily improve the model performance to fit observed values, with the exception of the Danube river basin. Multiple causes could explain the incorrect simulation of discharge dynamics in ORCHIDEE, even when irrigation is activated. For instance, uncertainties resulting from the atmospheric forcing are not assessed here, while they are known to affect the yearly and seasonal values of discharge \citep{Guimberteau2012,Decharme2019}. Also, a wrong ET estimation, errors in snow dynamics, and the lack of permafrost representation contribute to the mismatches \citep{Cheruy2020}. Finally, a lack of representation of other anthropogenic processes like dam management (Fig. \ref{fig11}), and water withdrawal for other economic

sectors and other uses could explain the differences in seasonal discharge dynamics between ORCHIDEE and observed data in some basins \citep{Pokhrel2016}.

Effect of irrigation on simulated TWSA is weak. In some large river basins, we observed increases in low values in areas with significant surface water supply. But even when irrigation is activated, ORCHIDEE is not able to follow the trends exhibited by GRACE datasets, for instance in Huang He and Indus river basin, two heavily irrigated areas where water depletion has been related to groundwater pumping for irrigation \citep{Rodell2018,Yin2020}. There are probably multiple causes for inability of LSMs to capture large negative decadal water storage trends \citep{Scanlon2018}, starting with the underestimation of irrigation rates at country-level and grid-cell scale (Fig. \ref{fig4}). Glacier loss misrepresentation in ORCHIDEE could also explain part of the differences to observed negative trends in some basins, for instance in the Indus and Ganges basins, that depend on water flow from the Himalaya mountains \citep{Rodell2018}. And of course, errors in the partitioning between the different water fluxes in ORCHIDEE \citep{Cheruy2020,Mizuochi2021} contribute to the problems in both simulations (NoIrr and Irr).

We also underline the lack of fossil groundwater abstraction in ORCHIDEE as a very likely cause to the underestimation of irrigation rates and TSWA trend mismatch. Fossil groundwater, also called non renewable groundwater, is important in semiarid areas like Pakistan and Middle East, and contributes nearly 20\% to gross irrigation water demand for the year 2000 \citep{Wada2012}. As the irrigation scheme represents abstractions from shallow aquifers but not from fossil sources, it probably restrains irrigation too often due to a supply shortage, thus could have problems fitting the negative trend in those areas with heavy groundwater use, as already reported by \citet{Yin2020} for China. But we must add that the estimation of fossil groundwater use is challenging. For instance, an assessment of the TWSA trends of residuals between our simulation and GRACE shows differences with estimates of groundwater depletion from \citep{Wada2012} in some countries (see Table S3). Underestimation of irrigation rates, and uncertainties arising from fluxes partitioning and from meteorological data would also affect the estimations of fossil groundwater abstraction. So far, we cannot explain to which extent each one of these possible causes participates in the misrepresentation of GRACE TWSA trends by ORCHIDEE.

Our results show that the new irrigation scheme helps simulate acceptable land surface conditions and fluxes in irrigated areas for ET and LAI, but they also show that inclusion of irrigation alone is not necessarily sufficient for a good fit between the simulated values of TWSA and discharge and observed products. Including additional anthropogenic processes could help to reduce some of these biases. For instance, dams management and fossil irrigation withdrawal could increase the water supply in some basins during dry months or years, thus increasing irrigation amount in areas with high irrigation demand and water supply shortage. At the same time, these processes may have an impact on river discharge dynamics and could help to represent the misrepresentation of TWSA trends in some areas.

\conclusions %% \conclusions[modified heading if necessary]

We implemented a global irrigation scheme within ORCHIDEE LSM, with a simple representation of environmental restriction, water allocation rules based on local

infrastructure, and water adduction from non-local water reservoirs. We compared the irrigation estimates to reported values of irrigation withdrawal, and then we compared the outputs with and without irrigation to observed products of ET, LAI, TWSA and discharge. Our results highlight how the inclusion of irrigation can reduce some modelling biases, especially on ET and LAI, but they also underline the difficulties to represent irrigation on a large scale by using a simple scheme and limited information.

The model could still benefit from improvements on parameter tuning by explicitly representing paddy rice irrigation. Paddy irrigation could decrease irrigation bias in areas of southern Asia by increasing the irrigation demand. Dam management representation and inclusion of non-renewable groundwater use could also reduce negative biases in some heavily regulated basins by increasing the water supply. These three aspects could change the spatial distribution of the ET and LAI increases within the model. For TWSA and discharge, the inclusion of processes like dams management or fossil groundwater use could help to represent observed seasonal dynamics and trends that the model is not currently able to represent.

Finally, we remember that LSMs are commonly used in coupled mode with climate models and irrigation can have an impact on some atmospheric variables via changes on latent heat flux and leaf area index. Thus, the results obtained here encourage the use of coupled simulations to explore the joint evolution of climate under the ongoing climate change (for historical and especially for future periods), water resources, and irrigation activities. While there is an increasing literature body that explores the coupling of irrigation and climate for the historical period, to the best of our knowledge that is not the case for future scenarios. Coupled climate simulations for future scenarios could help to foresee potential changes on the joint long-term evolution of water resources use and climate, and might help to identify possible social consequences.