Milieux environnementaux, transferts et interactions dans les hydrosystèmes et les sols

Internship Report

February - June 2022

Irrigation in the ORCHIDEE model: Evaluating inputs and results in metropolitan France

Author:
Luiza Vargas de Oliveira Heinz

Supervisor:
Agnès Durcharne
Pedro Arboleda

June 11, 2022

Contents

Abstract

Irrigation is a widespread practice that contributes to the anthropization of water and energy fluxes. Its full effect on the climate and the water cycle is not yet understood, and modelling is one of the tools available for comprehending it. The ORCHIDEE model represents the land surface's processes inside a global climate model, and an irrigation module was recently developed. The module is currently being evaluated and this study focused on its effects on the description of surface fluxes in metropolitan France.

In order to represent irrigation, data on the location of irrigated surfaces is needed. We explored four datasets that show the location and extent of irrigated areas in France: GMIA5, HID, LUH2, and Varenne de l'eau, and we compared them with country-level values from AQUASTAT. The datasets were spatially consistent, tough LUH2 has a similar area of irrigated surfaces in fewer grid cells. Both LUH2 and HID were used as input data for ORCHIDEE simulations.

Simulated irrigation increases evapotranspiration (ET) and decreases surface runoff. Irrigated volumes were greatly overestimated in France, compared to reported data in AQUASTAT and the national BNPE dataset. At the first order, the use of different meteorological data had the most prominent impact of all changes studied, though irrigation increased ET for both meteorological forcings used. Comparison of model outputs with observational ET data showed that irrigation could both reduce negative biases or increase positive ones. Given the high uncertainty in both observed and simulated values, it was not possible to determine which simulations had the most realistic ET outputs.

Résumé

L'irrigation est une pratique très répandue qui contribue à l'anthropisation des flux d'eau et d'énergie. La totalité de son effet sur le climat et le cycle de l'eau n'est pas encore compris, et la modélisation est l'un des outils disponibles pour la comprendre. Le modèle ORCHIDEE représente les processus de surfaces terrestres dans un modèle climatique global, et un module d'irrigation a été récemment développé. Ce module est actuellement en cours d'évaluation et cette étude s'est concentre sur ses effets sur la description des flux de surface en France métropolitaine.

Afin de représenter l'irrigation, des données sur la localisation des surfaces irriguées sont nécessaires. Nous avons exploré quatre jeux de données qui montrent la localisation et l'extension des surfaces irriguées en France : GMIA5, HID, LUH2 et Varenne de l'eau, et nous les avons comparés aux valeurs au niveau national provenant d'AQUASTAT. Les ensembles de données étaient spatialement cohérents, mais LUH2 présente une concetration plus forte des zones irriguées. LUH2 et HID ont été utilisés comme données d'entrée pour les simulations ORCHIDEE.

L'irrigation simulée augmente l'évapotranspiration (ET) et diminue le ruissellement. Les volumes irrigués ont été largement surestimés en France, par rapport aux données rapportées dans AQUASTAT et dans BNPE. Au premier ordre, l'utilisation de données

météorologiques différents a eu l'impact le plus important de tous les changements étudiés, bien que l'irrigation ait augmenté l'ET pour les deux forçages météorologiques utilisés. La comparaison des sorties du modèle avec les données d'observation d'ET a montré que l'irrigation pourrait à la fois réduire des biais négatifs et augmenter les positifs. Étant donné la grande incertitude des valeurs observées et simulées, il n'a pas été possible de déterminer quelles simulations avaient les sorties d'ET les plus réalistes.

1 Introduction

Irrigation is a widespread practice that covers around 18% of agricultural land, responsible for over 40% of food production [Siebert et al., 2005], contributing to the anthropization of the natural world. As it diverges large water volumes from their natural reservoirs, its effects on the water cycle and the climate are subject to research.

Irrigation aims to compensate for the soil water deficit and to improve conditions for crop growth. Water is extracted from surface or groundwater sources, decreasing their water content, and is thus often limited by its local availability. As it is applied to the soil, it can increase subsurface recharge and change local surface runoff patterns. Furthermore, the higher soil moisture content can lead to increased evaporation and the increased availability of water to plants can increase transpiration [Yin et al., 2020]. Moreover, different irrigation techniques can affect water fluxes differently. While drip and sprinkler irrigation are expected to increase significantly evapotranspiration, surface irrigation can have a considerable impact on surface runoff [Grafton et al., 2018].

The full effect of irrigation on climate is not yet known, though some studies have shown regional cooling and reduction of extreme events as possible consequences [Thiery et al., 2020]. Furthermore, the effect on the water cycle is very complex, with research indicating at times an increase or a decrease in precipitation in different regions[Al-Yaari et al., 2019] [Guimberteau et al., 2012].

One of the tools available for comprehending these effects is modelling. Global climate models (GCMs) are used to study past trends and future projections and can improve the knowledge of anthropization effects such as irrigation. GCMs often include land surface models (LSM), which are in charge of representing continental surface processes. The Coupled Model Intercomparison Project (CMIP) groups different GCMs to answer different questions related to climate and can help indicate where progress is needed in their development [Eyring et al., 2016]. In the last phase of CMIP, CMIP6, irrigation was present in only three GCMs [Al-Yaari et al., rev]. Hence, the representation of irrigation can be improved in the climate research community.

The IPSL (Institut Pierre-Simon Laplace) model is among the GCMs included in CMIP, and ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) is its LSM. ORCHIDEE can be used with prescribed meteorological forcing files or with IPSL's atmospheric model LMDZ[Boucher et al., 2020] [Krinner et al., 2005].

Irrigation was first included in ORCHIDEE in 2003 [de Rosnay, 2003], but this module was not active for CMIP6 [Cheruy et al., 2020]. Other developments have taken place since [Guimberteau et al., 2012] [Yin et al., 2020] and a new irrigation module is currently being implemented [Arboleda et al., 2021]. Moreover, previous research on ORCHIDEE indicates that irrigated zones may have stronger biases than elsewhere [Mizuochi et al., 2021], contributing to the interest in describing irrigation in the model.

As this module is new, its performance is currently being evaluated in its capacity to match observational data and its impact on different output variables is being investigated. Thus, research on its input data and parametrization is needed to improve results.

Among the current obstacles to calibrating the module, there is the need for input data. To properly model the effects of irrigation, it is necessary to know where irrigation takes place. Different databases exist[Siebert et al., 2005] [Siebert et al., 2014] [Hurtt et al., 2020] [Agreste, 2010] [FAO, 2022], at varying degrees of spatial and temporal resolution, and a good understanding of their differences is essential for choosing between them as input, as well as for comprehending their associated uncertainties. Even though overall coherence is expected between irrigation maps, different methodologies and sources are used for their creation. As a consequence, data may differ in a significant way and affect model outputs.

In this context, the objective of this internship was to evaluate and calibrate changes made to the ORCHIDEE model which incorporate irrigation [Arboleda et al., 2021], focusing on the region of metropolitan France. Different databases for irrigation distribution were investigated, as they are needed as input for the representation of irrigation in ORCHIDEE. Hence, both inputs and results of ORCHIDEE were analysed in the context of the irrigation module in France.

2 State of the art

2.1 ORCHIDEE

The ORCHIDEE model can be used either coupled with LMDZ or with meteorological forcing files. In both cases, variables such as air temperature, humidity, precipitation, incoming radiation, pressure, wind, and CO2 concentration are provided to the ORCHIDEE model, which returns information on energy and mass fluxes, such as latent heat, as well as parameters such as surface albedo [Krinner et al., 2005].

The model's calculations occur at different time steps. Energy and water balances and plant/soil carbon fluxes are computed at the same time as atmospheric physics, i.e. 15 minutes. At the daily time step, carbon allocation and soil carbon dynamics are calculated. Moreover, also at the daily time step, the routing scheme transforms the overland and groundwater flow into river discharge that eventually reaches the ocean. It can also include the evolution of vegetation biomass and leaf-area index (LAI), although this last component was not activated in this study, at a yearly time step [Boucher et al., 2020].

The calculations are made on a regular grid as illustrated in Figure 1, except for the routing scheme. Each cell has given percentages of different plant functional types (PFTs) and there are 15 of them available currently. Their proportion in a given cell will influence how the soil processes play out in the simulations, as a tropical forest is not expected to behave as a bare soil does. Eight PFTs represent forests, six stand for crops and grasslands, and one is for bare soil [Boucher et al., 2020]. As the atmospheric model does not distinguish the different PFTs, several variables, such as evapotranspiration (ET), are aggregated at the grid cell level, weighting different PFTs variables. Moreover, each cell has a single soil type based on its texture [Boucher et al., 2020].

Furthermore, in the model, the soil has a depth of two meters, which is divided into 11 layers that increase in thickness starting from the surface. The calculations of the soil processes are based on the Richards equation, describing thus the evolution of the water content on the soil. At the surface, water will infiltrate the soil and, at the bottom of the layers, the water content will, by default, become free gravitational drainage. Thus different parameters will come into play at the soil level, such as hydraulic conductivity and diffusivity [Ducharne et al., 2018].

The description of water fluxes is summarised in Figure 1. Not represented in the illustration is the precipitation (P), which is given by the LMDZ model or the atmospheric forcing. Once the precipitation reaches the surface, it can follow a few different paths. The water can go back to the atmosphere through evaporation, having infiltrated the soil beforehand, or through transpiration. Moreover, it can become surface runoff and go into what is called the fast reservoir or, after infiltrating the soil, become drainage and reach the slow reservoir. The fast and slow reservoir will feed the stream reservoir at different time delays, in order to mimic the residence times of surface and underground flows. The stream reservoir serves as a representation of the main rivers in each cell. The three reservoirs are used for calculating the river discharge in the routing

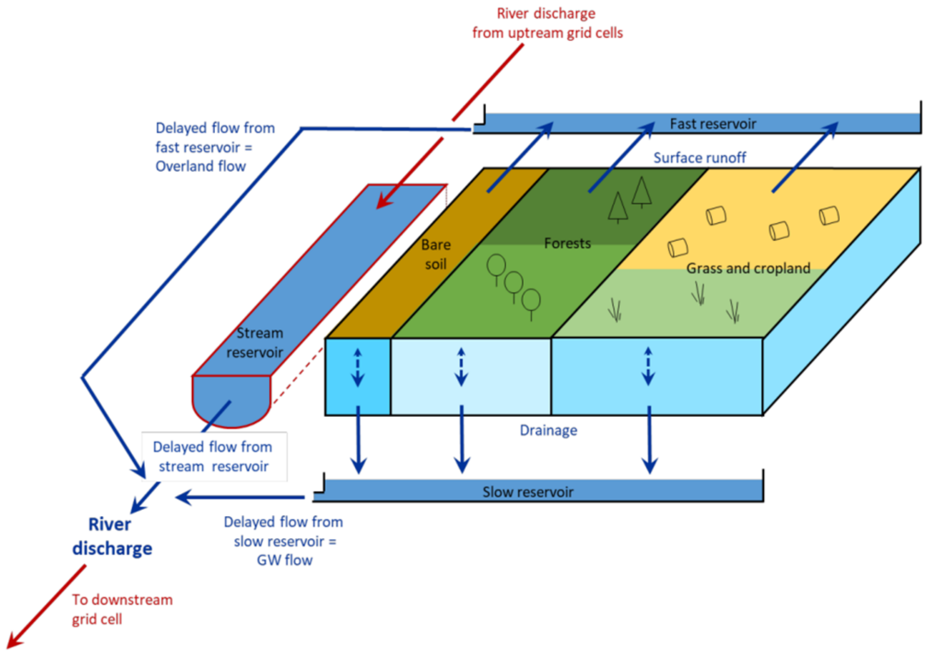


Figure 1: Simplified water fluxes in an ORCHIDEE grid cell [Arboleda et al., 2021]

scheme [Ducharne et al., 2018].

2.1.1 New Irrigation Module

Though irrigation was first introduced in ORCHIDEE in 2003 [de Rosnay, 2003], a new module was recently developed by other members of the METIS Hydrology group [Arboleda et al., 2021]. In this development, irrigation takes place for grass and crops PFTs.

For the calculations, firstly the root parameter is decided, which informs the depth of the root zone in the soil. The root zone defines which soil layers will be used for calculating the soil moisture deficit. This deficit (D) is defined in the equation below and is given by the gap between soil moisture in each layer (W) and a target soil moisture (βW_{fc}), which is in turn given by the soil moisture at field capacity W_{fc} and a parameter β that varies from 0 to 1. If the soil has already reached its target moisture, no deficit is computed and irrigation does not take place.

$$D = \sum_{i \in root\, zone} \min(0, \beta W_{fc} - W_i). \tag{1}$$

The variables D, W and W_{fc} are in kg/m and β is dimensionless. Once the deficit is computed, the irrigation requirement (I_{REQ}), in mm/dt, is calculated (Equation 2). ORCHIDEE also defines a maximum irrigation rate called I_{max} in mm/dt. As such, even if a large deficit is found, the volume of water used will be limited. Furthermore,

I_{REQ} is defined as zero if the leaf-area index (LAI) of the PFT is smaller than a chosen threshold, avoiding irrigating very young crops. The LAI is the cumulated leaf-area per unit area and can be used as an indicator for plant growth.

$$I_{REQ} = \min(I_{max}, \frac{D}{dt}) * f_{irrig} \tag{2}$$

The inclusion of I_{max} prevents having too much surface runoff, as it is produced when the rate of water to infiltrate exceeds the infiltration capacity of the soil. The irrigation rate found is multiplied by the fraction of the cell which will be irrigated f_{irrig} in m^2/m^2. The module thus requires as input a map of irrigated areas, which provides the value of f_{irrig} at different points in time and space.

Another relevant parameter is the available water in the reservoir AW, which is given in kg/m^2 and described by Equation 4 and takes into account the three reservoirs. The total water available in a reservoir is given by S_i in kg/m^2 . However, a minimum of water is required to stay in the reservoirs to preserve the ecosystems and only a fraction α_i of S_i can be extracted. Thus, a fraction ϵ_i is kept in reservoirs as the environmental flow.

$$AW = \sum_{i \in \ reservoirs} S_i \alpha_i \tag{3}$$

$$\alpha_i = 1 - \epsilon_i \tag{4}$$

$I_{REQ} * dt$ is then compared to the available water (AW) in the reservoirs which is in kg/m^2. The actual quantity of water withdrawn and applied as irrigation (I_{app}) in kg/m^2 is then given by the equation below, which is then extracted from the different reservoirs and applied to the soil column.

$$I_{app} = \min(I_{REQ} * dt, AW) \tag{5}$$

The value found for I_{app}/dt is then added to the water available at the soil surface for infiltration.

3 Materials & Methods

3.1 Observational data

Different sources of data were used in this study either in the context of result validation or input data analysis. An overview of these databases is given in Table 1, the content of which is explained in this section.

Table 1: Observational data used in the study

Database	Information Available	Unit	Interest to ORCHIDEE	Source
LUH2	AAI	ha	Input	[Hurtt et al., 2020]
HID	AEI	% of area	Input	[Siebert et al., 2014]
GMIA5	AAI and AEI	ha	Input	[Siebert et al., 2013a]
Varenne de l'eau	AAI	ha	Input	[Agreste, 2010]
AQUASTAT	AAI and AEI	ha	Input	[FAO, 2022]
	Volume	km^3/yr	Result Validation	
BNPE	Volume	km^3/yr	Result Validation	[eaufrance, 2022]
FLUXCOM	Latent Heat	MJ/m^2d	Result Validation	[Jung et al., 2019]
GLEAM	ET	mm/d	Result Validation	[Martens et al., 2017]

3.1.1 Irrigated surfaces databases

Several sources of information regarding irrigation intensity and distribution exist for metropolitan France. For ORCHIDEE to correctly model the effect of irrigation on the land surface's processes, an input database of irrigated surfaces that fairly represents reality is essential. Therefore, irrigated surfaces' datasets are analysed as possible input data.

Data exist at both national and sub-national scales, the latter with varying degrees of spatial resolution and with different temporal intervals. Different irrigation databases were investigated in the context of this study and were examined at different resolution levels, with a focus given to metropolitan France.

The data concerning irrigated surfaces is mostly found in two different formats: areas actually irrigated (AAI) and areas equipped for irrigation (AEI). As their names suggest, AAI quantifies the surfaces that were effectively irrigated in a given time, while AEI is all the area that is equipped and could be used for irrigation, regardless of whether it is or not. As such, AEI will always be higher than AAI and, for some countries, the difference between them may be significant [Siebert et al., 2005]. There is no immediate conversion method from one data type to the other, as the relation between AAI and AEI can vary spatially, and only one global map of this ratio was found during this study [Siebert et al., 2013a]. However, this map only accounts for a single year (2005) and using it indiscriminately as conversion may cause biases.

The Global Map of Irrigation version 5 (GMIA5) [Siebert et al., 2013a] provides global maps at 5 arcmin resolution for AEI and the ratio AAI/AEI. The data are static and correspond to the year 2005. Furthermore, it is the only database found that provides a spatial conversion from AEI to AAI.

Another source used for this study is the Food and Agriculture Organisation of the United Nations (FAO), which provides access to water and agriculture geospatial information in their database AQUASTAT [FAO, 2022]. Country statistics were extracted for France in three-year intervals between 1992 and 2018 for AEI, AAI, and volume of irrigation water withdrawals. These values will be compared to other sources and simulation results.

Some databases provide the evolution of irrigation distribution throughout time. The Land-Use Harmonization (LUH2) project supplies a database that harmonizes historical data with climate projections, covering the period between 850 and 2100 [Hurtt et al., 2020]. The data from LUH2 corresponds to AAI at 15 arcmin resolution. The Historical Irrigated Database (HID) also provides a time series, ranging from 1900 to 2005 [Siebert et al., 2014]. HID data corresponds to AEI and has a 5 arcmin resolution. The year 2005 was chosen for closer inspections of both LUH2 and HID, as it matches the GMIA5 data.

The French administration provides statistics on AAI at the communal level with the dataset Varenne de l'eau for 2010 [Agreste, 2010]. Other years are available at different spatial resolutions, but the year 2005, used to analyse other datasets, was not available.

3.1.2 Volume extracted for irrigation databases

As one of its result variables, the model calculates the extracted water volume, less than or equal to the irrigation needs (5), which will either come from the stream reservoir or the groundwater reservoir. As a consequence, it is possible to compare this data to water withdrawal databases and they are thus used in the context of result validation.

AQUASTAT provides information on irrigation volume since 1992 at the country level as mentioned above [FAO, 2022]. Furthermore, the national bank for quantitative water withdrawals (BNPE) provides information on the volume of water being extracted for irrigation at the point of withdrawal, [eaufrance, 2022] and records between 2008 and 2019 were analysed.

3.1.3 Energy and water fluxes databases

In order to evaluate ORCHIDEE outputs and the impact of irrigation, information on energy and water fluxes is needed. However, direct measurements of these processes at a large scale are not available and datasets that combine processing techniques with satellite recovered information were used instead. Thus, the FLUXCOM [Jung et al., 2019] and the GLEAM [Martens et al., 2017] [Miralles et al., 2011] databases were used.

The FLUXCOM initiative aims to provide data of energy fluxes based on FLUXNET eddy covariance towers, remote sensing, and meteorological data, coupled with machine learning methods, at different grid cell resolutions [Jung et al., 2019]. For the analysis presented in the following paragraphs, the data used did not contain meteorological information and used all machine learning and correction methods. FLUXCOM provides information on both carbon and energy fluxes, but only latent heat information was of interest in this study. Monthly products from 2001 to 2012 were used, with a spatial resolution of either 0.5 °or 0.5 arcmin according to the resolution of the simulation.

The Global Land Evaporation Amsterdam Model v3 (GLEAM) is a water budget model and aims to provide the land surface evaporation by using the Priestley and Taylor (PT) evaporation model, using meteorological forcing data. It utilises microwave-derived soil moisture, land surface temperature, and vegetation density, and models rainfall interception loss in its approach, focusing thus on water fluxes [Martens et al., 2017].

3.2 ORCHIDEE Simulation Analysis

3.2.1 Simulations analysed

The simulations analysed in this study were made by other members of the group METIS and the changes in parametrization are summarised in Table 2, with their respective names. These simulations were done in offline mode with two different meteorological datasets. When done at the global scale, the forcing files GSWP3 0.5°were used at a 0.5 °resolution [Kim, 2017]. NoIrrig uses the standard parametrization of CMIP6 and does not have irrigation. For LUH2_a and HID_a, the module was activated and respectively used the LUH2 and the HID database as input files for the irrigated area. As a consequence LUH2_a uses the areas actually irrigated (AAI), while HID_a uses areas equipped for irrigation (AEI), with AEI being by definition larger than AAI. Furthermore, LUH2 has yearly data that can be directly used in the model, while HID does not. As a consequence, the HID data (that has intervals between 5 and 10 years) needed to be extrapolated to cover all years. The basic parametrization defined for the irrigation module is used in LUH2_a and HID_a.

As for the other GSWP3 simulations, simulation LUH2_a was taken as a baseline and one parameter was changed at a time from LUH2_b to LUH2_f, in order to investigate the sensitivity of the model. The parameters changed are:

- The root parameter refers to the fraction of the roots in the so-called root zone. Decreasing it should decrease the need for irrigation.

- The parameter β changes the target soil moisture that irrigation aims to achieve inside the root zone (Equation 1). Lowering β should thus lower target soil moisture and irrigation needs.

Table 2: Simulations analysed and their deviations from standard parametrization

Simulation code	Irrigation	Forcing files	Database	β	Root parameter	I_{max}	α
NoIrrig	No	GSWP3 0.5°	-	-	-	-	-
LUH2_a	Yes	GSWP3 0.5°	LUH2	1	0.9	1 mm/h	0.9
HID_a	Yes	GSWP3 0.5°	HID	1	0.9	1 mm/h	0.9
LUH2_b	Yes	GSWP3 0.5°	LUH2	0.5	0.9	1 mm/h	0.9
LUH2_c	Yes	GSWP3 0.5°	LUH2	1	0.5	1 mm/h	0.9
LUH2_d	Yes	GSWP3 0.5°	LUH2	1	0.9	0.5 mm/h	0.9
LUH2_e	Yes	GSWP3 0.5°	LUH2	1	0.9	1.5 mm/h	0.9
LUH2_f	Yes	GSWP3 0.5°	LUH2	1	0.9	1 mm/h	0.5
SAFRAN_NoIrrig	No	SAFRAN 8km	-	-	-	-	-
SAFRAN_HID	Yes	SAFRAN 8km	HID	1	0.9	1 mm/h	0.9

- As described in equation 2, the model caps the amount of irrigation water that is added to the soil during one-time step at I_{max}. Its standard value is 1 mm/h and was changed to both 0.5 and 1.5 mm/h.

- The change in α refers to how much water needs to be left in the slow and stream reservoirs for them to still perform their environmental functions. The larger this fraction is, the more available water AW that can be used for irrigation (Equation 5).

For the global simulations done with GSWP3 0.5 °, only the grid cells corresponding to metropolitan France were used for the analysis

Finally, regional simulations were done in metropolitan France and the impact of forcing files was evaluated. Hence, the SAFRAN forcing files [Vidal et al., 2010], which have a resolution of approximately 8 km by 8 km, were used for simulations SAFRAN_NoIrrig and SAFRAN_HID. The simulation SAFRAN_NoIrrig uses the same parametrization as NoIrrig and SAFRAN_HID the same as HID_a.

4 Results & Discussion

4.1 Comparing different irrigation databases available for France

The different databases available for irrigated surfaces in France were studied regarding their overall values and their distribution in metropolitan France. An in-depth comparison was made for the year 2005 (or the closest available year), as well as an analysis of the temporal evolution of the different datasets. Table 3 presents the calculated sum of irrigated areas for both France and the world for the reference years, with either AAI or AEI data, according to what is available. A supplementary database, MIRCA, is shown and will be discussed further down in this section.

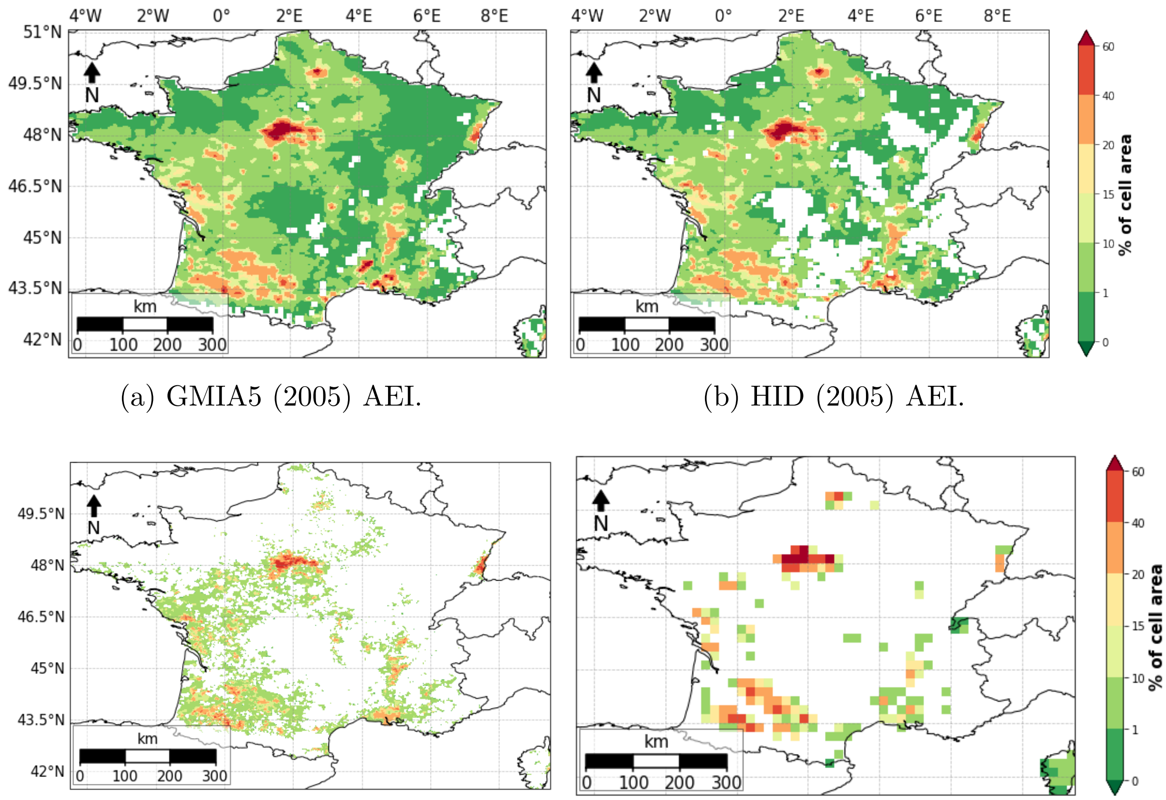
Table 3: Comparison of irrigated surfaces in the studied databases available, for France and the world

Database	Resolution	Datatype	Reference year	Irrigated areas in France		Irrigated areas in the world	
				AAI [1e6 ha]	AEI [1e6 ha]	AAI [1e6 ha]	AEI [1e6 ha]
LUH2 [Hurtt et al., 2020]	0.25°	AAI	2005	1.55	-	258	-
HID [Siebert et al., 2014]	0.0833°	AEI	2005	-	2.70	-	306
GMIA5 [Siebert et al., 2013a]	0.0833°	AEI and AAI	2005	1.68	2.89	255	308
AQUASTAT [FAO, 2022]	Country	AEI and AAI	2003 - 2007	1.51	2.64	228	309
Varenne de l'eau [Agreste, 2010]	Commune	AAI	2010	1.58	-	-	-
MIRCA [Portmann et al., 2010]	0.0833°	AHI (irrigated harvested area)	2000	1.71 (AHI)	-	312 (AHI)	-

The database GMIA5 has noticeably the highest values for both AAI and AEI in France, although not at the global scale. For other datasets, the difference between them remains under 7% for AAI and under 4% for AEI. Figure 2 shows the spatial distribution of the main databases studied. The ratio between AAI and AEI provided by GMIA5 is used whenever a conversion is made from AAI to AEI, and vice-versa. For example, this procedure is employed to have AAI values for HID.

Even if there are some differences between the data types and years presented, a qualitative analysis can be made. HID and GMIA5 have a very similar spatial distribution of AEI, the main differences seemingly being found in the least irrigated regions. LUH2 on the other hand has very few places where irrigation is present and exhibits few weakly irrigated spots, being significantly different from the previous two. Even though LUH2 has AAI data, the spatial distribution is still expected to be more similar to the HID and GMIA5.

The Varenne de l'eau database is available at the communal level and the data was first transformed into a raster of higher resolution (around 1 arcmin). When compared to HID and GMIA5 data, the data presents a similar spatial pattern, forming an S shape in the French territory with strongly irrigated areas (Figure 2 (d)). Nonetheless, some weakly irrigated areas present in GMIA5 are not counted as irrigated in Varenne



(a) GMIA5 (2005) AEI.

(b) HID (2005) AEI.

(c) Varenne de l'eau (2010) AAI.

(d) LUH2 (2005) AAI.

Figure 2: Irrigated areas in France as a percentage of grid cell area.

de l'eau. Overall, values are lower here than for HID and GMIA5, but this is expected, as Varenne de l'eau has AAI data. This data was analysed for the year 2010, as it represented the year that had both the most information and was closest to the other databases.

4.1.1 Differences between GMIA5 and HID

The total value of AEI in France was compared in greater detail between GMIA5 and HID, as a bigger difference than expected (7%) was found given HID has a large portion of its data sourced from GMIA5 [Siebert et al., 2014]. Figure 3 shows the spatial difference between HID and GMIA5 in 2005, with higher AEI in HID in southern France and lower in northern when compared to GMIA5.

The documentation for GMIA5 [Siebert et al., 2013b] indicates that the irrigable areas present in the EIDER (Entrepôt d'indicateurs et de données sur l'environnement) database [Min. de la trans. Ecolog., 2022] in the years 1997, 2000 and 2003 were used as a source for France (Table 4). When checking the database, the exact values used in GMIA5 are not immediately found. This discrepancy is explained by the authors' attempt to avoid underestimation by selecting the maximum of irrigable area reported

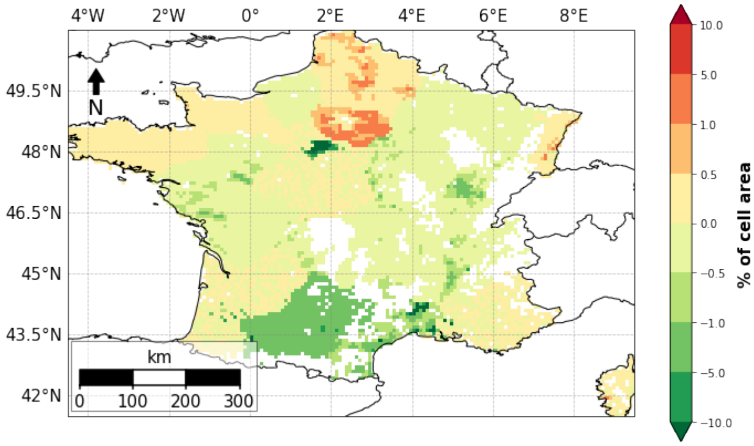


Figure 3: Difference between HID and GMIA5 in France for AEI in 2005.

by the EIDER database in the three years analysed. When selecting the highest value for each department in the three years, the exact value is obtained.

Although the HID database uses GMIA5 as one of its main sources for the year 2005, it is not the case in France . GMIA5 uses data from other years as a source for 2005 and HID looked for sources precisely from 2005 [Siebert et al., 2014]. In its supplementary information S7 [Siebert et al., 2014], EUROSTAT data is cited as the source for France in HID in 2005. The overall value for the area equipped for irrigation in HID for France is the same as in the EUROSTAT database [EUROSTAT, 2019].

Table 4: Source material for irrigated surfaces databases available for France. Data format % refers to the percentage of grid cells covered by irrigated areas.

Database	Resolution	Sources for France		Datatype	Data format	Years	Observations
LUH2 [Hurtt et al., 2020]	0.25°	1960 - 2005	FAOSTAT [FAO, 2022], GMIA5, MIRCA	AAI	%	850 - 2100	Irrigation limited to agricultural zones
		1900 - 1960	HID, GMIA5, MIRCA				
HID [Siebert et al., 2014]	0.0833°	2005	[EUROSTAT, 2019]	AEI	ha	1900 - 2005	-
		1970 - 2000	EIDER [Min. de la trans. Ecolog., 2022] and Interpolation				
		1961	FAOSTAT				
		1899 - 1960	Interpolation				
		1898	[Wilson, 1898]				
GMIA5 [Siebert et al., 2013a]	0.0833°	2005	EIDER	AEI and AAI	% and ha	2005	Maximal value in EIDER taken among the years 1997, 2000 and 2003
AQUASTAT [FAO, 2022]	Country	All	AQUASTAT Country Questionnaire	AEI and AAI	ha	1992 - 2018	-
Varenne de l'eau [Agreste, 2010]	Commune	-	-	AAI	ha	1988, 2000, 2010	-
MIRCA [Portmann et al., 2010]	0.0833°	2000	EIDER and AGRESTE	AHI (irrigated harvested area)	% and ha	2000	Same reasoning than GMIA with added agricultural efficiency data

4.1.2 LUH2

The AAI database for LUH2 has different origin sources, as it deals with a long historical period, like HID, and future projections. For the historical data (i.e. before 2015), the database HYDE is the source used. First, the values for area equipped for irrigation (AEI) after 1960 are taken from the FAOSTAT database[Goldewijk et al., 2017]

[FAO, 2022], which is available at the national level for multiple years. The AEI data then is converted to AAI using GMIA5, as it is the only available map of the ratio AAI/AEI. However, in this conversion, the ratio AAI/AEI for the year 2005 is used for all years present in LUH2.

Subsequently, the data at the national level from FAOSTAT is spatially distributed at the subnational level using the patterns of the MIRCA (monthly irrigated and rainfed crop areas around the year 2000) database [Portmann et al., 2010]. The MIRCA database does not provide AAI or AEI information, but the irrigated harvested area (AHI), i.e. AAI but taking into account harvesting efficiency, that is, how many times a year that area is harvested, explaining why it is larger than AAI. MIRCA shares the same database of EIDER as GMIA5, but MIRCA adds statistical information in France from the Agriculture and Food Ministery [Agreste, 2010]. Furthermore, the allocation procedure restricts all irrigated areas to zones already computed as cropland areas.

The MIRCA distribution is visually very similar to AAI from GMIA5 (see Annex I), even though it represents another variable, and GMIA5 itself was used to convert AEI from HYDE to AAI before accounting for multiple crops. These steps cannot thus account for the difference found in the spatial distribution of LUH2, which has very concentrated irrigated zones. Therefore, the discrepancies seem to arise from the allocation procedure which only allows irrigation to be placed where there is cropland. As the distribution of these croplands is less spread out than irrigated areas, the irrigation is forced to concentrate in these areas. Therefore, similar initial data at the national level ends up with a very different distribution at the regional level. The abruptness of the transition between irrigated and non-irrigated areas may create a problem for regional simulations, as it may not be a realistic representation of reality.

4.1.3 Water withdrawal for irrigation

Alongside the irrigated surface area, it is possible to verify the amount of water withdrawal for irrigation purposes. The "Banque nationale des prélèvements quantitatifs en eau" (BNPE) provides the volumes taken for irrigation in different water withdrawal locations, with their exact coordinates. The data analysed runs from 2008 up to 2019, but the years before 2012 have many missing data, especially in Eastern France.

This database was first compared to the AQUASTAT data at the national level, as shown in Figure 4. For this analysis, BNPE volumes were added for all of France for each year described by BNPE. As for AQUASTAT, only three years were available in this interval, 2012, 2017, and 2018, and had identical values. AQUASTAT considers this information to be valid for all years between 2008 and 2022.

For the years where data was clearly lacking in the BNPE database, BNPE underestimates the water withdrawal when compared to AQUASTAT. For the years with a more complete database, it is AQUASTAT that underestimates. BNPE data, in its most complete years, is considered to provide more information, as it provides more detailed statistics.

Subsequently, the spatial distribution of water withdrawals in BNPE was compared

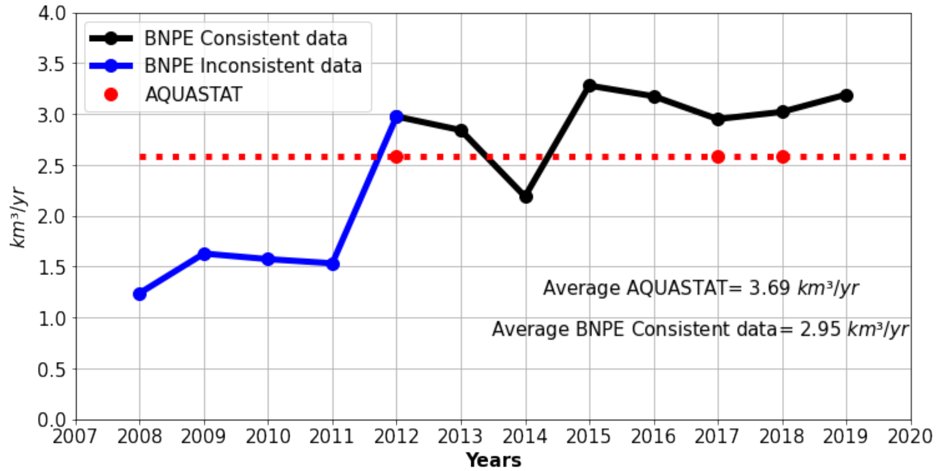
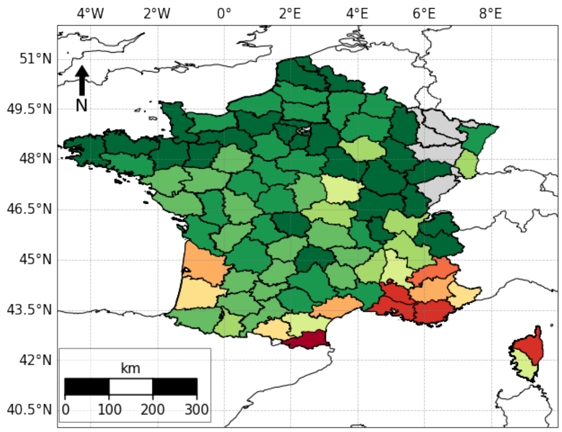


Figure 4: Total water volume withdrawal in France for irrigation purposes.

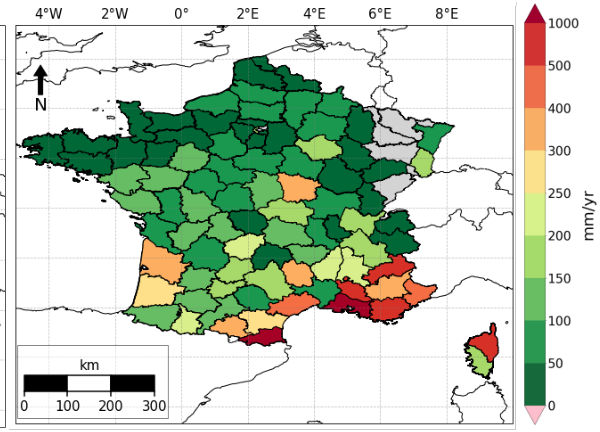
to the distribution of irrigated surfaces, specifically AAI (converted from AEI when needed), using LUH2, GMIA5, HID, and Varenne de l'eau. This analysis was done by dividing the volume by the irrigated surface for each department in mainland France to obtain an average water depth that serves as a proxy for irrigation intensity. A hypothesis is made that the water is extracted and used for irrigation in the same department. The results are found in Figure 5.

The databases GMIA5 and HID have spatial distributions that match irrigation water withdrawal very well. With the Varenne de l'eau database, there are some departments with conflicting results for water volume and irrigated surfaces, but this amounts to a negligible percentage of overall water volume. The LUH2 database, however, has considerable differences in comparison to water extraction distribution. Many departments with withdrawals in BNPE have no irrigated areas in LUH2. The sum of "orphan" BNPE withdrawals amounts to 7% of total irrigation water volume with LUH2.

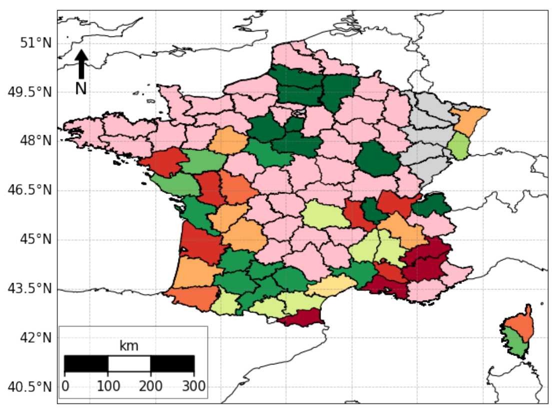
Furthermore, GMIA5, HID, and Varenne de l'eau show consistent irrigation patterns, with more intense irrigation in the south and less so in the north. This pattern is not found for LUH2, which shows high-intensity irrigation in departments such as the Loire-Atlantique, Deux Sèvres et Vienne, which is not at all observed for the other databases. The allocation method for irrigated areas in LUH2, mentioned in the previous section, might have skewed the spatial distribution to overly concentrate irrigation in areas where it is not necessarily realistic. In contrast, departments such as Var and Alpes Maritimes show significant irrigation intensity in other databases, but no AAI in LUH2. Finally, though the department of Beauce has large values for AAI and AEI, it does not have a particularly high irrigation intensity.



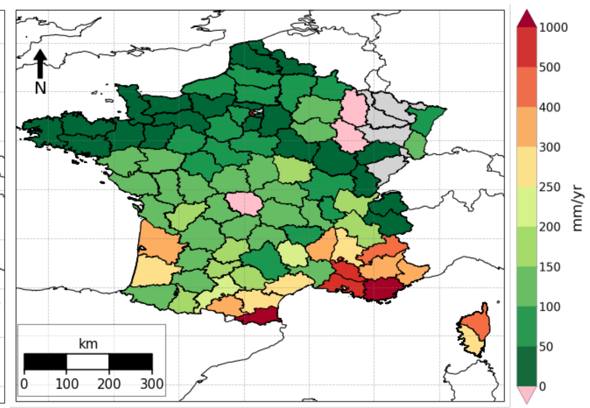
(a) GMIA5 (2005).



(b) HID (2005).



(c) LUH2 (2005).



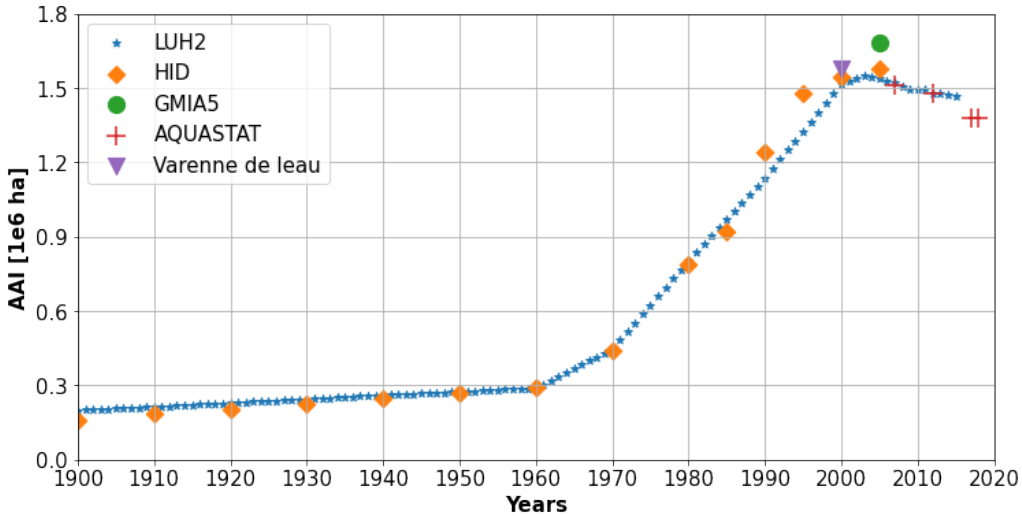
(d) Varenne de l'eau (2010).

Figure 5: Water withdrawn for irrigation in mm/yr in 2012 (BNPE) distributed at the department scale using irrigated areas from different datasets. Gray departments have no considerable irrigation, i.e. no irrigation water withdrawal according to BNPE and less than 50 ha used for irrigation. Pink departments indicate significant conflicting results between databases, i.e. BNPE data shows at least 10,000 m^3/yr destined for irrigation, while no irrigated areas are present in these departments in the surface database.

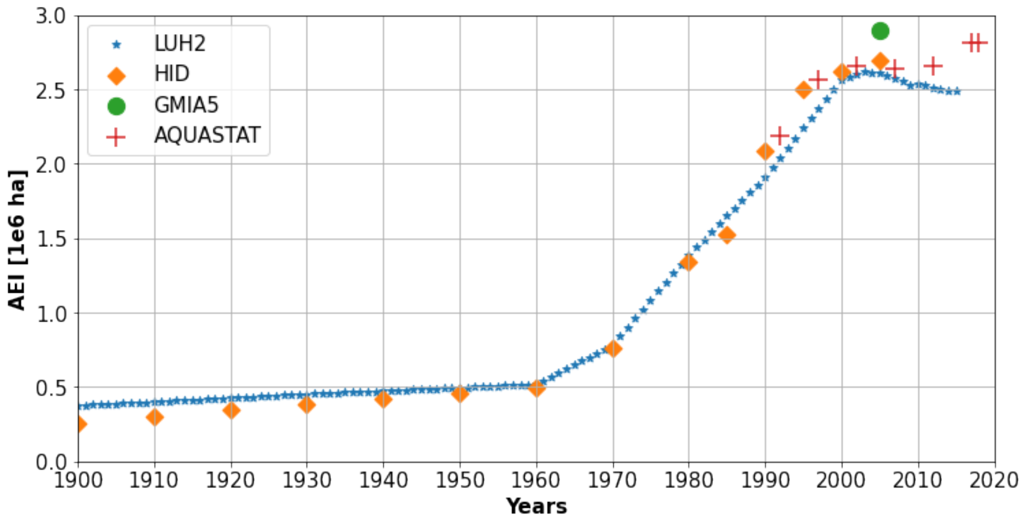
4.1.4 Temporal evolution

The total amount of irrigated surface in France, as in any other country, has evolved with time. The two databases that cover a significant amount of time are LUH2 and HID and a comparison was made since 1900, with the addition of AQUASTAT and GMIA5 in more recent years (Figure 6). As LUH2 has AAI data and HID AEI data, a conversion was necessary and done with GMIA5, similarly to what was done with HYDE [Goldewijk et al., 2017].

For the first 60 years, a linear increase is seen in both AAI and AEI, which can



(a) AAI.



(b) AEI.

Figure 6: Temporal evolution since 1900 of AAI and AEI in metropolitan France according to different irrigation databases.

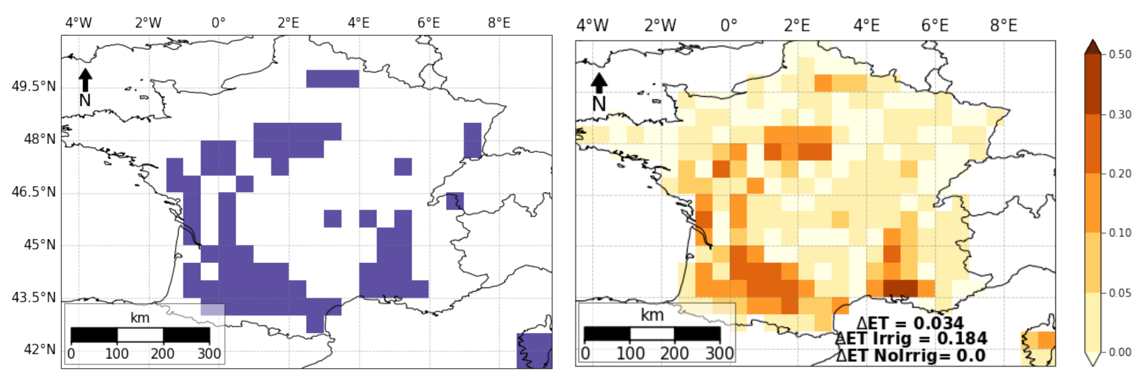
be explained by the lack of data for this period. Supplementary information on HID [Siebert et al., 2014] indicates that a linear interpolation was made between the years 1898 and 1960 for AEI in France, and, as LUH2 uses HID as a source for these years, both databases have very similar results and a linear tendance. After 1960, HID uses a combination of FAOSTAT, EIDER and EUROSTAT for France [Min. de la trans. Ecolog., 2022] [EUROSTAT, 2019] [FAO, 2022], whereas LUH2 sticks to FAOSTAT as its main source. As a result, HID shows a less linear increase and surpasses LUH2 values after 1990.

Some unexpected evolutions are found in the AQUASTAT data incorporated in the most recent years. For AEI, which has data available between 1992 and 2018, AQUASTAT values have a similar trend to HID up to 2005 and show a significant increase after. In contrast, for AAI, AQUASTAT values are indistinguishable from LUH2 for 2007 and 2012 (first years available for AQUASTAT AAI) but significantly decrease after. The general discrepancy between AAI and AEI is not necessarily false, but also hard to verify, as most historical databases offer only one type of data, and conversion methods are not widely available except for GMIA5. This finding could suggest that even though more areas are being equipped for irrigation, the amount that it is actually being used is decreasing. On the other hand, it could also be an indicator that the underestimation of irrigation practices is more pronounced for AAI than AEI. One possible hypothesis is that area being equipped for irrigation can more easily be monitored, whereas how much of it is actually used can be underreported. As the AAI data from AQUASTAT only goes back to 2007, it is difficult to further investigate this hypothesis.

4.2 Simulation results

4.2.1 Sensitivity to irrigation

The addition of irrigation increases the average ET in ORCHIDEE outputs (Figure 17). An increase of around 0.034 mm/d (around 2.3%) was found in France between simulations NoIrrig and LUH2_a and it happened in irrigated regions: ET did not change in non-irrigated regions and increased by 0.184 mm/d in irrigated ones (around 12%). A trend of increasing ET annual averages can also be observed (Figure 8 and 9).



(a) Irrigated areas. (b) $ET_{LUH2_a} - ET_{NoIrrig}$ (1980 to 2013).

Figure 7: Impact of irrigation in ORCHIDEE simulations. (a) Mask of irrigated areas for GSWP3 simulations with LUH2 used as an input. Cells in purple have an irrigated fraction bigger than zero according to LUH2. (b) Difference of ET between LUH2_a and NoIrrig between 1980 and 2013.

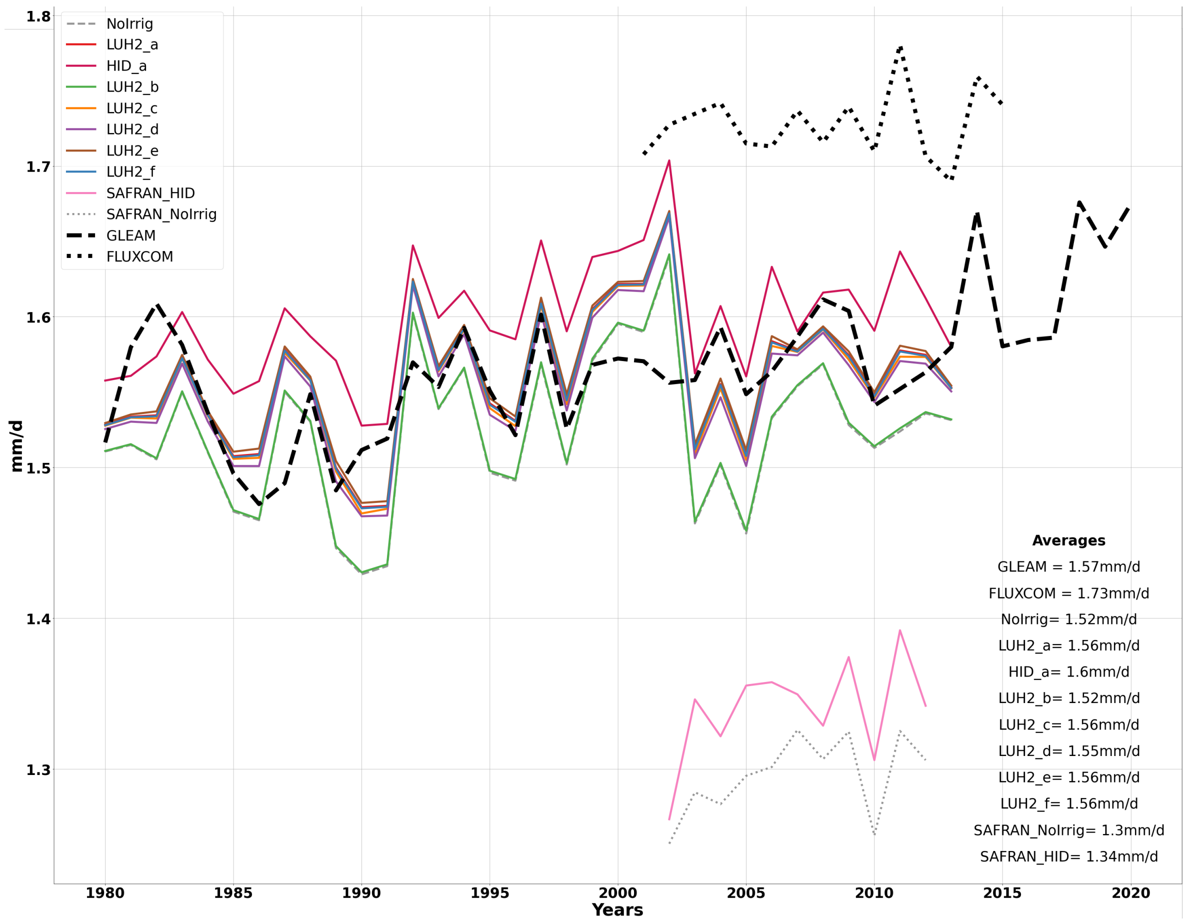


Figure 8: Average ET in simulations and databases in metropolitan France.

4.2.2 Sensitivity to parametrization and input data

Although irrigation increased ET, different parameterization choices were also shown to have a significant impact on output. Their impact on GSWP3 simulations with LUH2 as input data were studied. The change in the β coefficient, which is used to calculate soil moisture deficit (Equation 1) had a very significant influence on the results. Reducing β had the effect of decreasing irrigation needs as expected. Both the values and the spatial distribution resulting from this simulation were very similar to the NoIrrig simulation, showing the model is very sensitive to this parameter. In Figure 8, the lines for NoIrrig and LUH2_b are virtually indistinguishable. Furthermore, the average volume used for irrigation in the model (Q_{irrig}) was very low for this simulation and equaled 0.03 km^3/yr (Figure 10). This value is around two orders of magnitude below expected and can be considered an insignificant amount. Thus, changing β from 0.9 to 0.5 is essentially equivalent to stopping irrigation.

The decrease in α represented a reduction in the water availability in the simulation. Beforehand up to 90% of the water in the stream and slow reservoirs could be used for

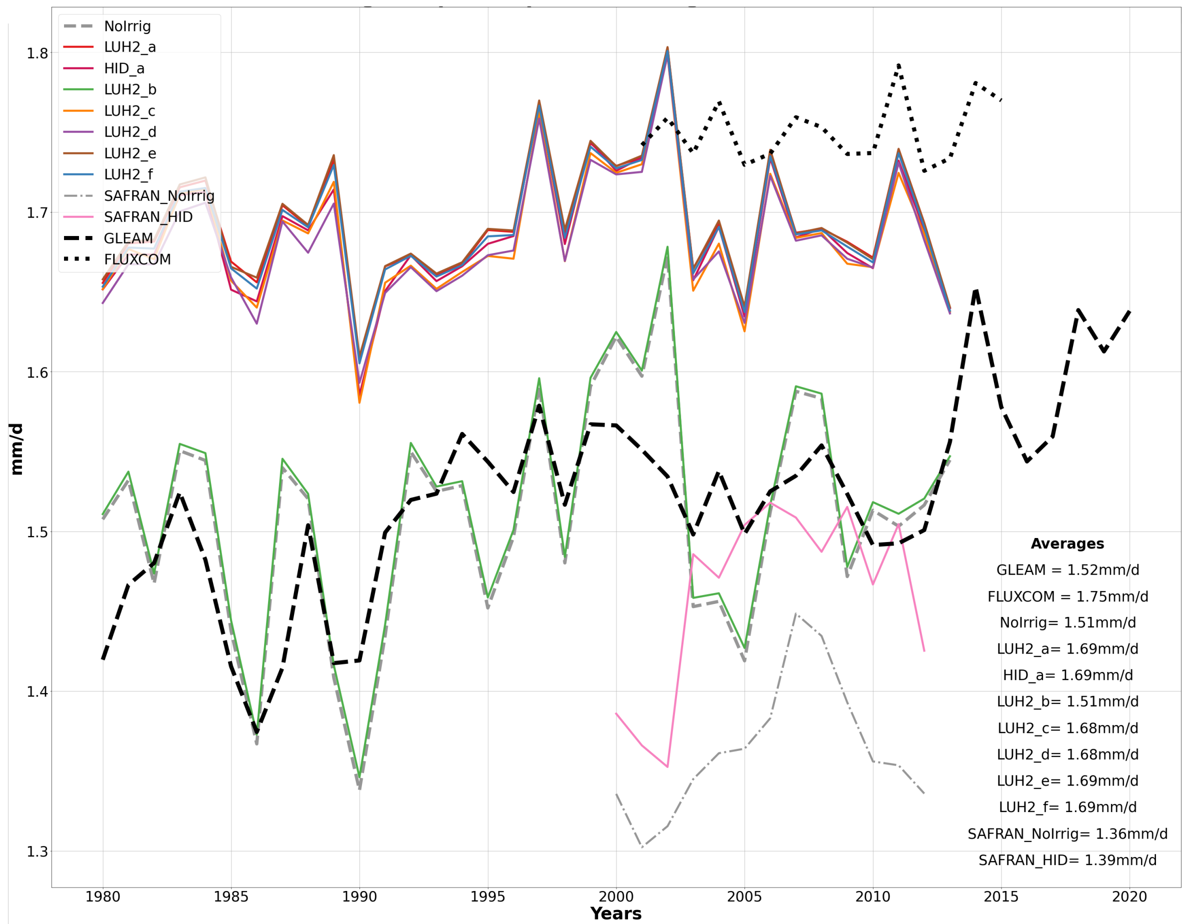


Figure 9: Average ET in simulations and databases in irrigated areas in metropolitan France. The irrigation maps used are in Annex II. All cells irrigated in LUH2 were used for GSWP3 simulations and observational products. All cells with over 5% irrigation in HID were used for SAFRAN simulations

irrigation, and for simulation LUH2_f, only 50% could. As a consequence, however, an almost imperceptible change was found in ET and Q_{irrig} decreased only by 2%. The differences were of very little statistical significance and this finding indicates that diminishing the availability of water in the model by 40% will have little influence on the irrigation amount. Hence, irrigation does not seem to be currently limited by the availability of water in France.

The variations in ET caused by changes in the maximum irrigation rate (I_{max}) were as expected: when increasing I_{max}, ET increased, and vice versa. The relative impact however was not the same, with the decrease in I_{max} by 0.5 mm/h leading to an almost 18% decrease in irrigation volume, and the increase in I_{max} by 0.5 mm/h to only an 8% increase in volume (Figure 10).

The change in the root parameter from 0.9 to 0.5 decreases the effect of irrigation by an amount comparable to the decrease in I_{max}. As a smaller root zone decreases the

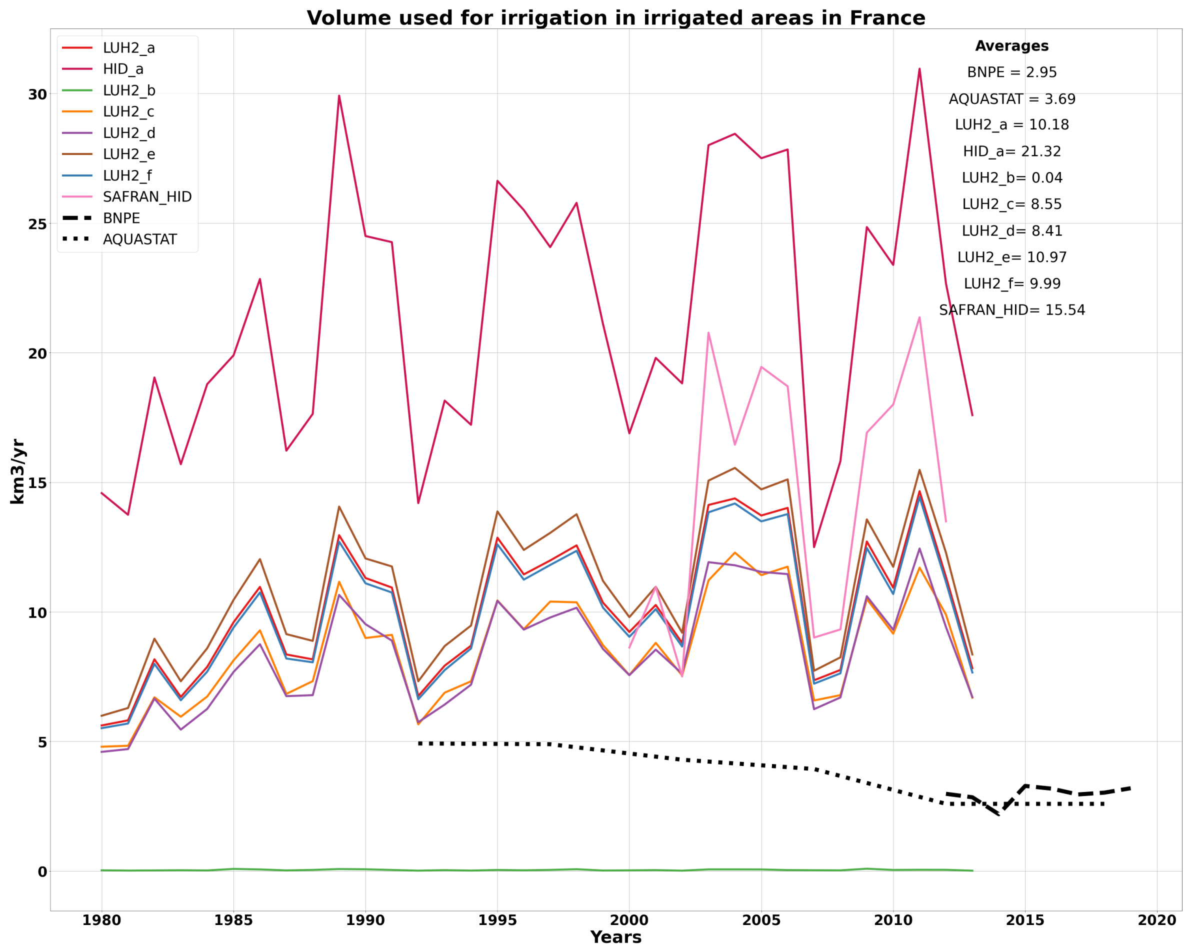


Figure 10: Volume of water extracted and used for irrigation in simulations and databases in irrigated regions in France.

need for irrigation, a decrease in ET was expected and found (around 6%).

Moreover, a few other variables were analysed and their variations were compatible with expected changes (Table 5). With added irrigation, ET, drainage, soil moisture, and LAI, and increased, whilst the amount of surface runoff and water passing through the reservoirs decreased. Total runoff refers to the sum of drainage and surface runoff. The changes are detailed in the Table 5 for the simulations NoIrrig, LUH_a, and HID_a.

All the changes observed were more intense in irrigated areas than in metropolitan France. All reservoirs decreased their water content, but the highest decrease was in the fast reservoir, most likely due to the decrease in surface runoff. The decrease in the stream and slow reservoir was a direct impact of irrigation, as water is extracted from them. The total runoff increased with irrigation, as the increase in the drainage was bigger than the decrease in surface runoff. Finally, the increase in ET, soil moisture, and LAI show that irrigation fulfilled its role in providing water to the soil in order to help vegetation growth.

Table 5: Comparison of different variables for simulations NoIrrig, LUH2_a, and HID_a for metropolitan France and its irrigated areas.

Variable	Unit	NoIrrig		LUH2_a				HID_a			
		France	Irrigated Areas	France	Difference with NoIrrig [%]	Irrigated Areas	Difference with NoIrrig [%]	France	Difference with NoIrrig [%]	Irrigated Areas	Difference with NoIrrig [%]
ET	mm/d	1,523	1,509	1,559	2,328	1,693	12,169	1,598	4,903	1,687	11,775
Soil Moisture	kg/m^2	592,470	560,568	595,715	0,548	579,413	3,362	599,479	1,183	578,623	3,221
Water in slow reservoir	kg/m^2	35,418	30,001	34,923	-1,396	27,447	-8,515	33,995	-4,019	27,137	-9,549
Water in stream reservoir	kg/m^2	5,595	6,961	5,488	-1,913	6,623	-4,857	5,349	-4,391	6,380	-8,348
Water in fast reservoir	kg/m^2	0,831	0,568	0,803	-3,363	0,468	-17,528	0,747	-10,168	0,466	-17,927
Drainage	mm/d	0,867	0,616	0,885	2,120	0,718	16,531	0,910	5,033	0,714	15,889
Surface runoff	mm/d	0,183	0,116	0,180	-1,702	0,101	-12,893	0,173	-5,781	0,100	-13,416
Total runoff	mm/d	1,050	0,732	1,065	1,453	0,819	11,887	1,083	3,146	0,814	11,264
LAI	-	1,921	1,866	1,967	2,365	2,123	13,800	2,016	4,941	2,108	13,000

Table 5 also provides insight into the effects of input data. HID has a higher value for irrigated surfaces than LUH2, as it deals with AEI and LUH2 with AAI. The model does not distinguish between the two types of entries and treats both as AAI. Adding irrigation using HID has thus similar effects to LUH2, but the changes are intensified. This intensification is most likely a combined effect of both an overall larger irrigated area and a more spread out one, with irrigation needs less concentrated.

The changes in France in variables studied were often 2 or 3 times more intense for HID_a than LUH2_a when compared to NoIrrig. The volume of water extracted for irrigation is also doubled for HID (Figure 10). However, when looking exclusively at irrigated areas, the effects shown in Table 5 have for the most part similar magnitudes in both simulations.

It is also possible to compare the simulated irrigation water volume to water withdrawal data, provided by BNPE and AQUASTAT. These databases provide, respectively, an average of 2.95 and 3.69 km^3/yr irrigation water use in France and all the simulations strongly overestimate this amount (Figure 10).

Some underestimation is expected with official statistics, in part because of small extractions. Extractions under 10,000 m^3/yr in normal zones are not accounted for by the French administration. This value changes to 7,000 m^3/yr in water-stressed zones, which represent 47% of the French territory. However, for these minor extractions to compensate for the bias of around 5 km^3/yr, approximately 500,000 of them would be necessary and exclusively used for irrigation. The total number of extraction points for all water uses (irrigation and others) is around 138,00 [Sauquet et al., 2022]. Thus, these minor extractions may contribute to some underestimation, but are not likely to fully compensate for the gap found in the data.

LUH2_a overestimates Q_{irrig} by a factor of around 3. All changes in the base parameters which intensify irrigation result in further overestimation of the irrigated volume in France, while changes that lower it give out more realistic values. When compared to reference simulation LUH2_a, simulations LUH2_c, LUH2_d, and LUH2_f result in a lower positive bias, and HID_a and LUH2_e in an increased bias.

When comparing the changes caused by the forcing files (simulations HID_a and

SAFRAN_HID), using the SAFRAN forcing files decreased the irrigation volume, but still caused the second-largest overestimation among the simulations analysed. This finding is compatible with the lower ET values observed for SAFRAN_HID when compared to HID_a. It also suggests that the overestimation of the irrigation volume in France is caused by the ORCHIDEE model or the observation uncertainties, but not by the uncertainty of the meteorological data.

Finally, the efficiency of irrigation Ef_{Irrig} was calculated by dividing the increase in ET in comparison to a simulation without irrigation (NoIrrig or SAFRAN_NoIrrig) by the volume of water used for irrigation in mm/d (Equation 6).

$$Ef_{Irrig} = 100\% * \frac{ET_{Irrig} - ET_{NoIrrig}}{Q_{irrig}} \tag{6}$$

This parameter corresponds to how much of the water extracted is effectively used and turned into ET by the cropland. The most efficient simulation was LUH2_d (76%), which represents a smaller root zone used for the calculation. The lowest efficiency found corresponded to SAFRAN_HID (53%) by a considerable margin. Furthermore, the efficiencies for LUH2_b, which decreased β, were not included, as they showed a behaviour extremely close to having no irrigation, and this skewed the results.

In principle, irrigation in ORCHIDEE resembles more closely drip or surface irrigation than sprinkler irrigation, as water is made available for infiltration at the soil surface. The irrigation efficiency at the local scale is estimated between 90 and 100% for drip irrigation, and between 50 and 85% for surface irrigation [Grafton et al., 2018]. However, at the national level [Frenken and Gillet, 2012] estimates the irrigation efficiency for France to be around 60%. As such, ORCHIDEE overestimates efficiency according to Frenken et Gillet, but underestimates according to Grafton et al. As the efficiency calculated here is at the national level, the Frenken et Gillet value should be more comparable, but it is also provided by a model with its own uncertainties.

4.2.3 Comparison to observation and sensitivity to forcing data

Figures 8 and 9 also give out information on the effect of using different forcing files and on the observational data available, which all show a trend of increasing average ET values in the last decades. The biggest difference between simulations comes from the forcing files: simulations using SAFRAN have a very strong negative bias compared to both observational products, while simulations using GSWP3 have a strong negative bias against FLUXCOM, but match GLEAM at first order. Among GSWP3 simulations, the highest values ET were found for HID_a and the lowest for NoIrrig and LUH2_b. As previously discussed, HID_a showed the most intense irrigation out of GSWP3, which increases ET. For NoIrrig and LUH2_b, they have respectively no irrigation and almost no irrigation (Figure 10), and thus have the smallest ET.

When looking at the average ET in the irrigated areas, the negative bias with FLUXCOM is reduced but, for most GWSP3 simulations, a clear positive bias appears against GLEAM. The simulations NoIrrig and LUH2_b are in this context the closest

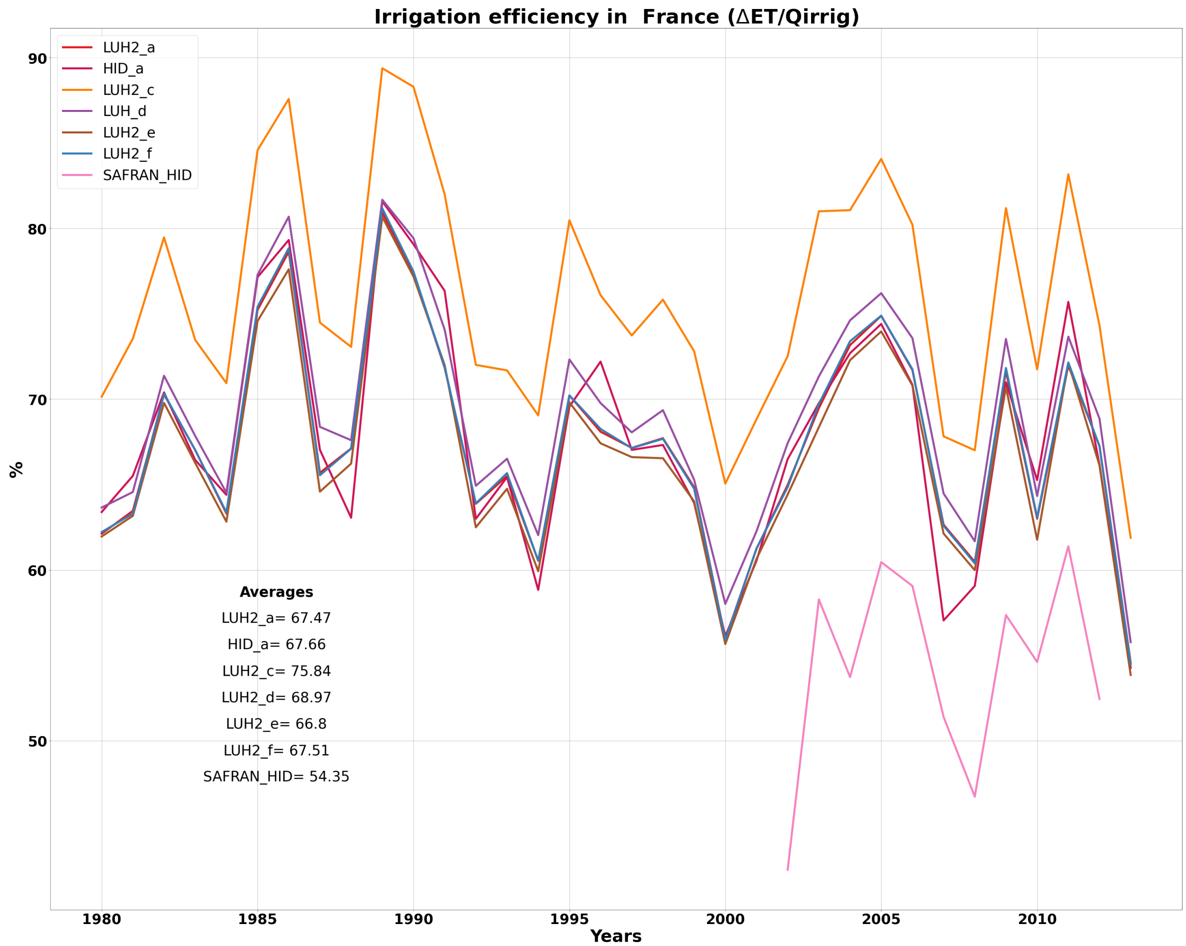


Figure 11: Irrigation efficiency in simulations calculated with Equation 6 for irrigated areas in France.

to GLEAM. As for the SAFRAN simulations, the negative bias to both observational products remains but is significantly reduced.

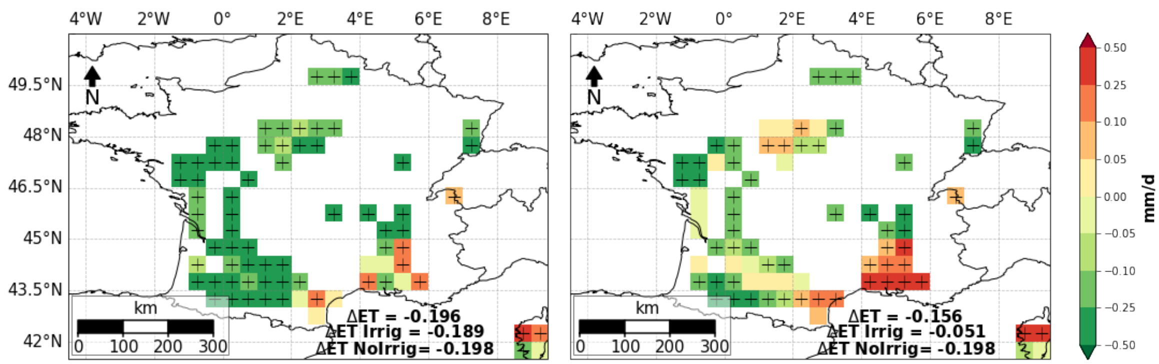
It must be noted that what constitutes an irrigated area is different for GSWP3 and SAFRAN simulations, as their very different spatial characteristics make it difficult to have a fully compatible spatial mask. For GSWP3 simulations, we selected all cells where irrigation is present according to LUH2, as LUH2 was most often used as input in GSWP3 mode (Figure 2). As for SAFRAN, all cells with more than 5% of irrigation fraction according to HID were considered irrigated areas (Annex II). This choice was made to take into account discrepancies between datasets, as LUH2 has more concentrated irrigation, and, in consequence, masks have similar total areas (GSWP3 with 15,7 million ha and SAFRAN with 14,4 million ha).

One unexpected feature found was the behaviour of the observational data and the SAFRAN simulations around the year 2003, when a major drought struck France. GSWP3 simulations show a large decrease of ET for this year, as expected for a dry year, but the observations remained relatively constant and the SAFRAN simulations showed an increase in ET.

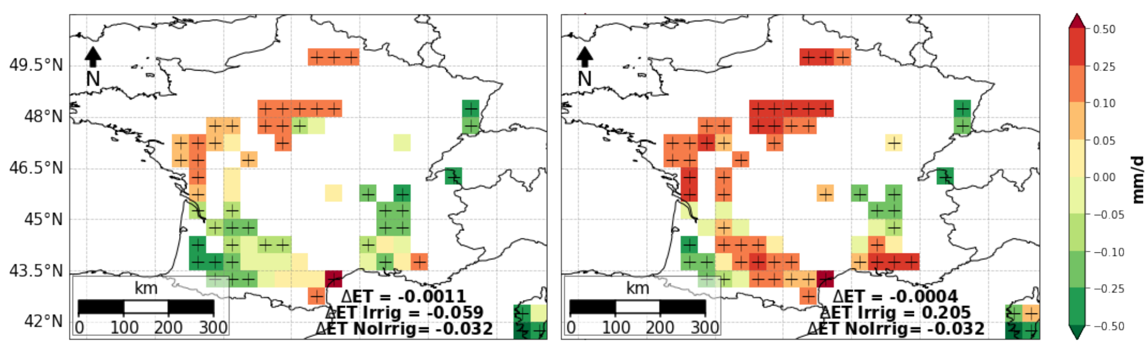
To further comprehend the significant change caused by forcing files, a comparison was made for the main variables used by ORCHIDEE. Although the values for precipitation are very similar between the two datasets, some key differences were found elsewhere (see Annex III). The air temperature is consistently higher in SAFRAN than GSWP3 and air humidity is lower, though by a small margin. Longwave downward radiation also showed some discrepancies, but the long-term means are very close. The biggest difference was observed for shortwave downwards radiation (SW_{Down}). A sharp increase in SW_{Down} was observed from 2002 to 2003 in SAFRAN, increasing the energy available to ORCHIDEE. The increase for the same year in GSWP3 was much less significant. This change might help explain the increase in ET found with SAFRAN_HID and SAFRAN_NoIrrig (Figure 8) in the same year, even though it was a year marked by drought. If ET is limited by energy in the SAFRAN simulations, the increase in SW_{Down} would allow the model to overcome the limitation, and ET would increase even in drought years.

Moreover, the distribution of ET biases against observation was also studied and the results for reference simulations NoIrrig and LUH2_a can be found in Figure 12. The differences found in Figures 8 and 9 can be observed here as well. For FLUXCOM, the addition of irrigation caused a decrease in the strongly negative bias and decreased the number of cells that showed significant differences with observations. An almost inverse pattern was found for GLEAM, as the positive bias is stronger with irrigation and more cells present significant differences. A similar analysis was made for SAFRAN_NoIrrig and SAFRAN_HID and is compatible with the decrease in biases in irrigation areas found in Figure 9 (Annex IV).

The difference between ET from GLEAM and FLUXCOM creates further uncertainties, as it is not easy to determine which simulations have the most realistic outputs. Notably, the GLEAM dataset shows higher values of ET in France close to the Mediterranean, where there is more incident solar radiation, and lower values in the northern half, where there are many irrigated areas. Thus, irrigation would only increase ET in zones where GLEAM has low values for it, increasing the positive bias. On the other hand, irrigation seems to play in the correct zones to compensate for FLUXCOM biases.



(a) NoIrrig and FLUXCOM (2001-2013). (b) LUH2_a and FLUXCOM (2001-2013).



(c) NoIrrig and GLEAM (1980-2013). (d) LUH2_a and GLEAM (1980-2013).

Figure 12: Bias de ET for simulations NoIrrig and LUH2_a. Crosses indicate statistical significance at a 0.05 level according to Students T-test. Zones colored indicate irrigated areas. Average ΔET values correspond to metropolitan France only, with ΔET_{Irrig} corresponding to irrigated areas and $\Delta ET_{NoIrrig}$ to non-irrigated areas.

5 Conclusion & Perspectives

The difficulties of representing irrigation with a numerical model begin with the lack of available data. Among what is available, an investigation was conducted to comprehend if the databases were consistent in the French metropolitan area or if they had strong differences. Overall, all datasets had consistent regions with strong irrigation, but LUH2 had a much more concentrated distribution of irrigation. When hoping to emulate regional effects of irrigation, such as the local river flow change, this characteristic may serve as a problem.

The total amount of hectares used for irrigation was consistent in France among different databases, except for a higher value for GMIA5, which resulted from a strategy to avoid underestimation. Some databases showed areas equipped for irrigation, i.e. that could be used for irrigation, and others areas actually irrigated, i.e. that are in fact used in a given year. When taking into consideration the difference between AAI and AEI, the values of all databases remain coherent.

The inclusion of an irrigation module in ORCHIDEE had visible effects in France. The main variable studied, ET, increased significantly in irrigated areas (12% for LUH_a). It also exhibited varying sensitivity to the different parameters tested, but the biggest variations were caused by the change in input data, either forcing files (GSWP3 and SAFRAN) or irrigation maps (LUH2 and HID).

Two observational products were used to assess the ET outputs, FLUXCOM and GLEAM. They have different values and spatial distributions for France and use different techniques to obtain data. Though both use observation to construct their datasets, there is a level of uncertainty and data extrapolation in them, and it is thus not possible to use them to definitely point out to a more realistic simulation.

Even with this uncertainty, simulations that used SAFRAN forcing had higher biases than those with GSWP3. SAFRAN simulations had high negative biases against both observational products, while GSWP3 simulations had mostly intermediate values for ET between FLUXCOM and GLEAM. Irrigation decreased the biases with FLUXCOM for all simulations and had varying effects on the biases with GLEAM. With GLEAM, after adding irrigation, SAFRAN simulations continued with negative biases, though decreased, and GSWP3 simulations developed a positive bias.

The comparison of the volume used to irrigate in ORCHIDEE with official statistics showed a large overestimation by the model. Though there are reasons to believe in an underestimation by official data, the overestimation is very high and indicates a failure by the model, which could be linked to irrigation efficiency.

The irrigation module presents itself as a form to improve the description of reality in the ORCHIDEE model, but it still needs adjusting. Further investigation is needed to pinpoint the best methods to do so, as the discrepancies found in France are not necessarily the same as those found in other countries. The overestimation of the volume extracted for irrigation is not reproduced everywhere, as a strong underestimation is found by the model over the globe [Arboleda et al., 2021]. Different options may include globally changing parametrization and correcting efficiencies at the national level.

Furthermore, other variables than those studied in this work are available as ORCHIDEE outputs. In future work, they could be compared against other observational data to further comprehend the impact of irrigation in the model. The Gravity Recovery and Climate Experiment (GRACE) can be used to analyse the total water storage fluctuations in the model, and river flow rate data can be compared to routing scheme outputs. Moreover, piezometric data may be used to study groundwater representation in ORCHIDEE.

The analysis made in this study can also be applied to different countries where irrigation is present. Likewise, simulations with different forcing files may help improve the irrigation module. Simulations in coupled mode may also help estimate the influence on the climate, which currently is not fully understood[Al-Yaari et al., 2019] [Al-Yaari et al., rev].

References

[Agreste, 2010] Agreste (2010). Recensement Agricole 2010. publisher: Agreste, la statistique agricole.

[Al-Yaari et al., 2019] Al-Yaari, A., Ducharne, A., Cheruy, F., Crow, W. T., and Wigneron, J.-P. (2019). Satellite-based soil moisture provides missing link between summertime precipitation and surface temperature biases in CMIP5 simulations over conterminous United States. *Scientific Reports*, 9(1):1657.

[Al-Yaari et al., rev] Al-Yaari, A., Ducharne, A., Thiery, W., Cheruy, F., and Lawrence, D. M. (in rev.). The Role of Irrigation Expansion on historical climate change: insights from CMIP6. *Earth Future*.

[Arboleda et al., 2021] Arboleda, P., Ducharne, A., Yin, Z., and Philippe Ciais, P. (2021). Exploring the effects of an improved irrigation scheme in ORCHIDEE land surface model on Europe's land surface hydrology and energy budget. Strasbourg, France.

[Boucher et al., 2020] Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N. (2020). Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth Systems*, 12(7):e2019MS002010. e2019MS002010 10.1029/2019MS002010.

[Cheruy et al., 2020] Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, ., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J.-L., Ghattas, J., Grandpeix, J.-Y., Idelkadi, A., Mellul, L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., Wang, F., and Zhao, Y. (2020). Improved near-surface continental climate in ipsl-cm6a-lr by combined evolutions of atmospheric and land surface physics. *Journal of Advances in Modeling Earth Systems*, 12(10):e2019MS002005. e2019MS002005 10.1029/2019MS002005.

[de Rosnay, 2003] de Rosnay, P. (2003). Integrated parameterization of irrigation in the land surface model ORCHIDEE. Validation over Indian Peninsula. *Geophysical Research Letters*, 30(19):1986.

[Ducharne et al., 2018] Ducharne, A., Ottlé, C., Maignan, F., Vuichard, N., Ghattas, J., Wang, F., Peylin, P., Polcher, J., Guimberteau, M., Maugis, P., Tafasca, S., Tootchi, A., Verhoef, A., and Mizuochi, H. (2018). The hydrol module of ORCHIDEE: scientific documentation [rev 3977] and on, work in progress, towards CMIP6v1.

[eaufrance, 2022] eaufrance (2022). Banque nationale des prÃĺlÃĺvements quantitatifs en eau (BNPE). publisher:.

[EUROSTAT, 2019] EUROSTAT (2019). Agri-environmental indicator - irrigation - Statistics Explained.

[Eyring et al., 2016] Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, 9(5):19371958.

[FAO, 2022] FAO (2022). AQUASTAT Core Database. publisher: Food and Agriculture Organization of the United Nations.

[FAO, 2022] FAO (2022). FAOSTAT Database. publisher: Food and Agriculture Organization of the United Nations.

[Frenken and Gillet, 2012] Frenken, K. and Gillet, V. (2012). Irrigation water requirement and water withdrawal by country.

[Goldewijk et al., 2017] Goldewijk, Klein, K., Beusen, A., Doelman, J., and Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene â HYDE 3.2. *Earth System Science Data*, 9(2):927–953.

[Grafton et al., 2018] Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., and Allen, R. G. (2018). The paradox of irrigation efficiency. *Science*, 361(6404):748–750.

[Guimberteau et al., 2012] Guimberteau, M., Laval, K., Perrier, A., and Polcher, J. (2012). Global effect of irrigation and its impact on the onset of the Indian summer monsoon. *Climate Dynamics*, 39(6):1329–1348.

[Hurtt et al., 2020] Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., HumpenÃűder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X. (2020). Harmonization of global land use change and management for the period 850â2100 (LUH2) for CMIP6. *Geoscientific Model Development*, 13(11):5425–5464.

[Jung et al., 2019] Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, 6(1):74.

[Kim, 2017] Kim, H. (2017). Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1). publisher: Data Integration and Analysis System (DIAS).

[Krinner et al., 2005] Krinner, G., Viovy, N., de Noblet-DucoudrÃĺ, N., OgÃĺe, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system: DVGM FOR COUPLED CLIMATE STUDIES. *Global Biogeochemical Cycles*, 19(1).

[Martens et al., 2017] Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernandez-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5):1903–1925.

[Min. de la trans. Ecolog., 2022] Min. de la trans. Ecolog., . (2022). EIDER Base de données régionales et départementales sur l'environnement, l'énergie, le transport, le logement et la construction.

[Miralles et al., 2011] Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15(2):453–469.

[Mizuochi et al., 2021] Mizuochi, H., Ducharne, A., Cheruy, F., Ghattas, J., Al-Yaari, A., Wigneron, J.-P., Bastrikov, V., Peylin, P., Maignan, F., and Vuichard, N. (2021). Multivariable evaluation of land surface processes in forced and coupled modes reveals new error sources to the simulated water cycle in the IPSL (Institute Pierre Simon Laplace) climate model. *Hydrology and Earth System Sciences*, 25(4):2199–2221.

[Portmann et al., 2010] Portmann, F. T., Siebert, S., and DÃűll, P. (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling: MONTHLY IRRIGATED AND RAINFED CROP AREAS. *Global Biogeochemical Cycles*, 24(1):n/a–n/a.

[Sauquet et al., 2022] Sauquet, E., Vergnes, J.-P., and Thirel, G. (2022). IDENTIFICATION DE STATIONS HYDROMETRIQUES ET PIEZOMETRES POUR LEVALUATION DES MODELISATIONS HYDROLOGIQUES.

[Siebert et al., 2005] Siebert, S., DÃűll, P., Hoogeveen, J., Faures, J.-M., Frenken, K., and Feick, S. (2005). Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences*, 9(5):535–547.

[Siebert et al., 2013a] Siebert, S., Henrich, V., Frenken, K., and Burke, J. (2013a). Global Map of Irrigation Areas version 5.

[Siebert et al., 2013b] Siebert, S., Henrich, V., Frenken, K., and Burke, J. (2013b). Update of the Digital Global Map of Irrigation Areas (GMIA) to Version 5. publisher: FAO.

[Siebert et al., 2014] Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., and Scanlon, B. R. (2014). A global dataset of the extent of irrigated land from 1900 to 2005. preprint, Global hydrology/Remote Sensing and GIS.

[Thiery et al., 2020] Thiery, W., Visser, A. J., Fischer, E. M., Hauser, M., Hirsch, A. L., Lawrence, D. M., Lejeune, Q., Davin, E. L., and Seneviratne, S. I. (2020). Warming of hot extremes alleviated by expanding irrigation. *Nature Communications*, 11(1):290.

[Vidal et al., 2010] Vidal, J.-P., Martin, E., FranchistÃĺguy, L., Baillon, M., and Soubeyroux, J.-M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system: A 50-YEAR HIGH-RESOLUTION ATMOSPHERIC REANALYSIS OVER FRANCE. *International Journal of Climatology*, 30(11):1627–1644.

[Wilson, 1898] Wilson, H. M. (1898). Relations of Irrigation to Geography. *Journal of the American Geographical Society of New York*, 30(1):1–14.

[Yin et al., 2020] Yin, Z., Wang, X. H., OttlÃĺ, C., Zhou, F., Guimberteau, M., Polcher, J., Peng, S. S., Piao, S. L., Li, L., Bo, Y., Chen, X. L., Zhou, X. D., Kim, H., and Ciais, P. (2020). Improvement of the Irrigation Scheme in the ORCHIDEE Land Surface Model and Impacts of Irrigation on Regional Water Budgets Over China. *Journal of Advances in Modeling Earth Systems*, 12(4).

Annex I - Supplementary maps for irrigation datasets

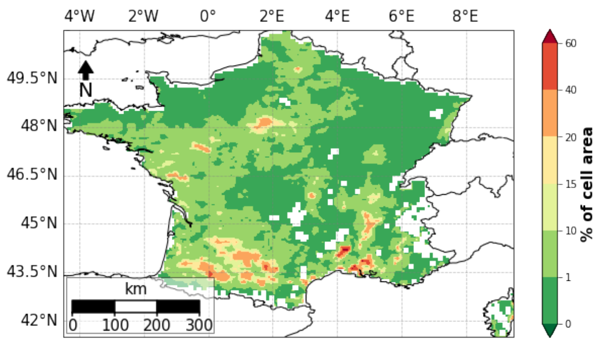


Figure 13: Areas Actually Irrigated in France in 2005 according to GMIA5.

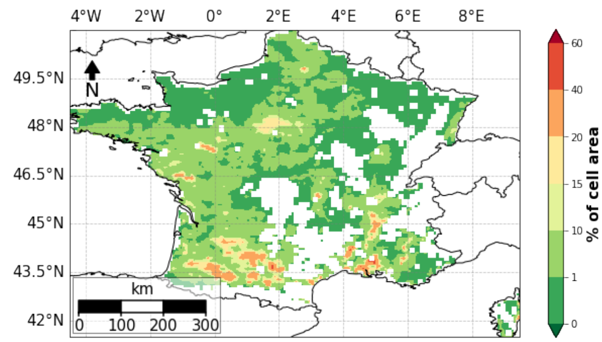


Figure 14: Areas Actually Irrigated in France in 2005 according to HID.

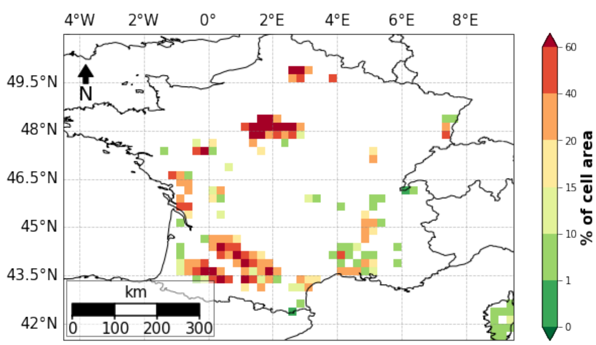


Figure 15: Areas Equipped for Irrigation in France in 2005 according to LUH2.

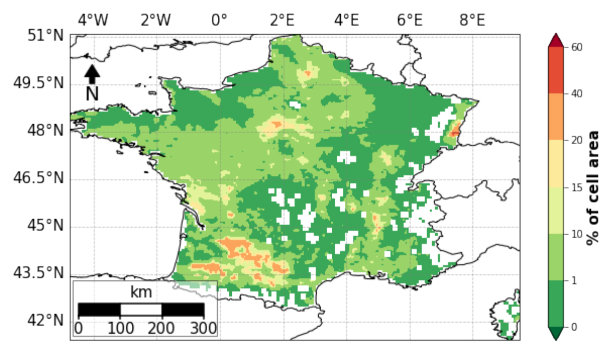


Figure 16: Irrigated harvested areas in France in 2000 according to MIRCA.

Annex II - Irrigated area masks

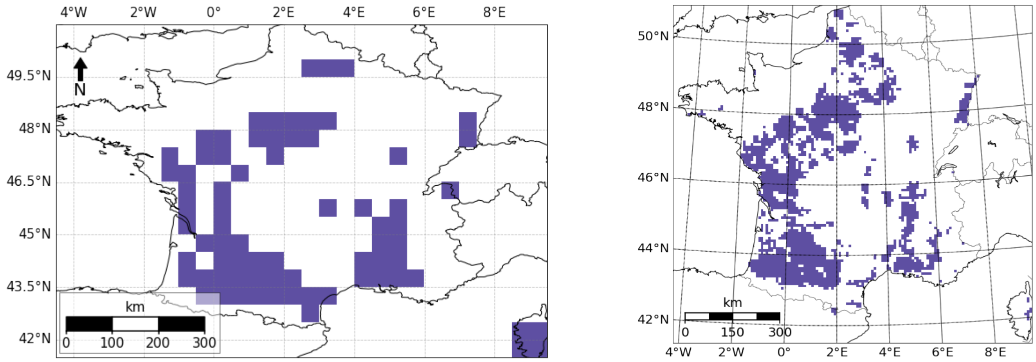
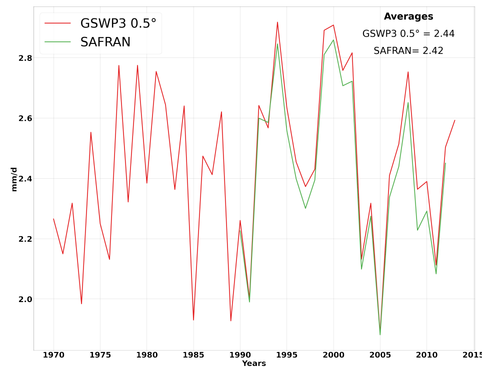
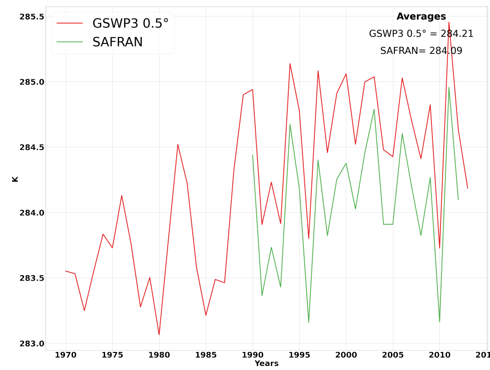


Figure 17: Irrigated areas' masks. (a) Cells with irrigation according to LUH2 in purple. Used for GSWP3 simulations. (b) Cells with irrigation fraction higher than 5% according to HID. Used for SAFRAN simulations..

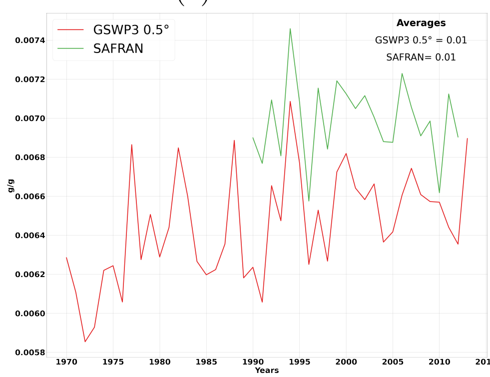
Annex III - Comparison between SAFRAN and GSWP3



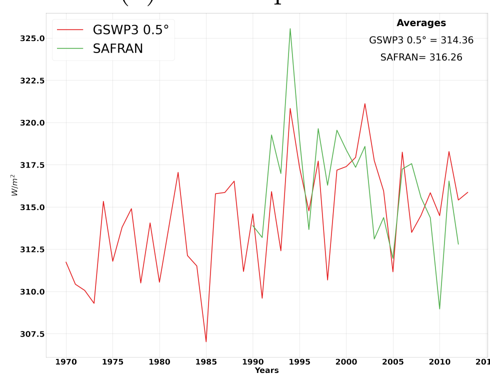
(a) Rainfall.



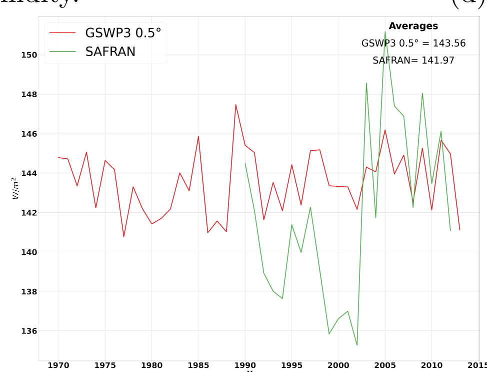
(b) Air Temperature.



(c) Air Humidity.



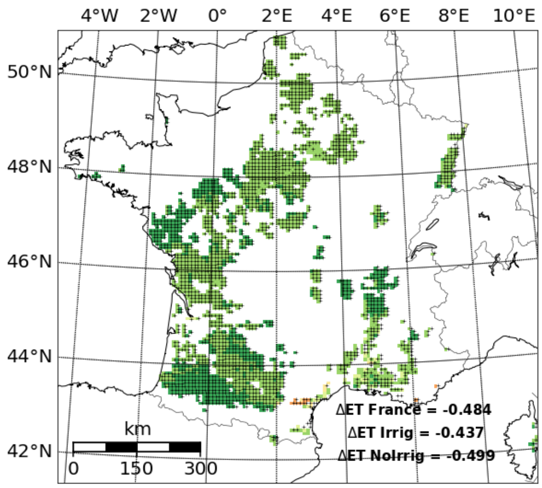
(d) LW_{Down}.



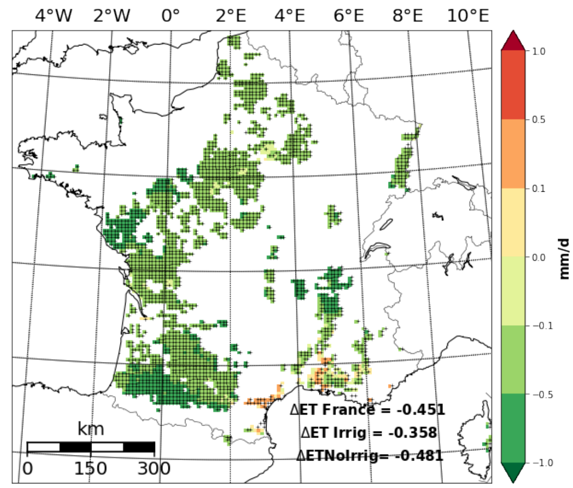
(e) SW_{Down}.

Figure 18: Comparison between average values for GSWP3 and SAFRAN forcing files in France.

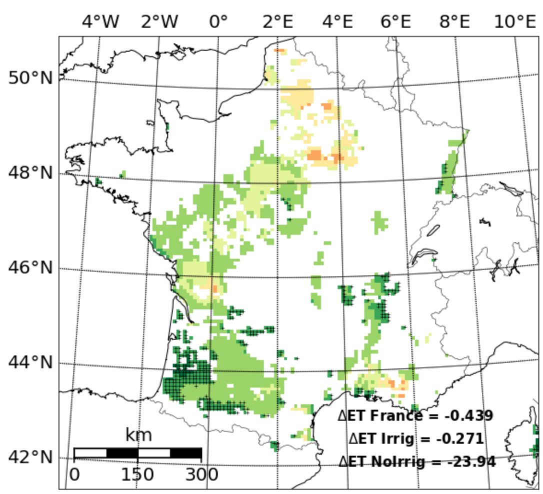
Annex IV - Biases with SAFRAN simulations



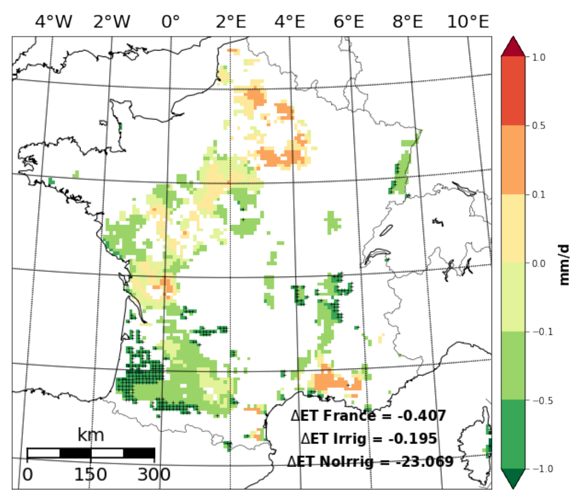
(a) SAFRAN_NoIrrig and FLUXCOM (2001-2012).



(b) SAFRAN_HID and FLUXCOM (2001-2012).



(c) SAFRAN_NoIrrig and GLEAM (2000-2012).



(d) SAFRAN_HID and GLEAM (2000-2012).

Figure 19: Bias de ET for simulations SAFRAN_NoIrrig and SAFRAN_HID. Crosses indicate statistical significance at a 0.05 level according to Students T-test. Zones colored indicate irrigated areas. Average ΔET values correspond to metropolitan France only, with ΔET_{Irrig} corresponding to irrigated areas and $\Delta ET_{NoIrrig}$ to non-irrigated areas.