Drainage assessment of irrigation districts: on the precision and accuracy of four parsimonious models

Pierre Laluet[1], Luis Olivera-Guerra[1,*], Víctor Altés[2,3], Vincent Rivalland[1], Alexis Jeantet[4], Julien Tournebize[4], Omar Cenobio-Cruz[5], Anaïs Barella-Ortiz[5], Pere Quintana-Seguí[5], Josep Maria Villar[3], Olivier Merlin[1]

[1]Centre d'Etudes Spatiales de la Biosphère (CESBIO), Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France
[2]isardSAT, Doctor Trueta 113, Barcelona, Spain
[3]Department of Environment and Soil Sciences, University ofe Lleida, Lleida, Spain
[4]University of Paris-Saclay, INRAE Jouy-en-Josas – Antony, UR HYCAR, Antony, France
[5]Observatori de l'Ebre, Universitat Ramon Llull – CSIC, Roquetes, Spain
*now at: Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ-UPSACLAY, UMR 8212, IPSL, Gif-sur-Yvette, France

Correspondence: pierre.laluet@gmail.com

Abstract. In semi-arid irrigated environments, agricultural drainage is at the heart of three agro-environmental issues: it is an indicator of water productivity, it is the main control to prevent soil salinization and waterlogging problems, and it is related to the health of downstream ecosystems. Crop water balance models combined with subsurface models can estimate drainage quantities and dynamics at various spatial scales. However, such models' precision (capacity of a model to fit the observed drainage using site-specific calibration) and accuracy (capacity of a model to approximate observed drainage using default input parameters) have not yet been assessed in irrigated areas. To fill the gap, this study evaluates four parsimonious drainage models based on the combination of two surface models (RU and SAMIR) and two subsurface models (Reservoir and SIDRA) with varying complexity levels: RU-Reservoir, RU-SIDRA, SAMIR-Reservoir, and SAMIR-SIDRA. All models were applied over two sub-basins of the Algerri-Balaguer irrigation district, northeastern Spain, equipped with surface and subsurface drains driving the drained water to general outlets where the discharge is continuously monitored. Results show that RU-Reservoir is the most precise (average KGE ($Q^{0.5}$) of 0.87), followed by SAMIR-Reservoir (average KGE ($Q^{0.5}$) of 0.79). However, SAMIR-Reservoir is the most accurate model for providing rough drainage estimates using the default input parameters provided in the literature.

1 Introduction

In the context of ongoing global changes, semi-arid irrigated areas especially face multiple challenges. First, agricultural water productivity is a critical issue in regions where water resources are under increasing pressure (FAO, 2021). Second, one-third of the world's irrigated land is affected by the soil salinization issue, which is likely to bring a significant loss of arable lands (Singh et al., 2019). Third, non-point pollution is another issue in irrigated areas, with return flows that may contain high concentrations of nutrients (García-Garizábal et al., 2012) and/or pesticides (Abdi et al., 2021).

Agricultural drainage is at the heart of the above three challenges (water productivity, soil salinization, and non-point pollution). More than 20% of the total irrigated lands in the world are equipped with drainage systems, including open ditches or buried drains (Schultz et al., 2007). Drainage systems are generally installed to prevent waterlogging during heavy rainfall (through a sudden rise in the water table), to facilitate salt leaching (particularly when the irrigation water has high salt concentrations), and to maintain a low water table to avoid salt accumulation in the root zone by capillary rise (particularly when the groundwater has high salt concentrations). Moreover, the drained water quantity and quality are strong indicators of agricultural water productivity and the possible impact of nitrates, salts, and pesticide concentration on downstream ecosystems (Blann et al., 2009).

Measuring drainage discharge is an effective way of monitoring the quantity and quality of drainage to help address the three challenges mentioned above. However, the proportion of drained irrigation districts equipped with such instruments is very low. In this context, estimating the drained water in irrigated areas, including those not instrumented, is of major importance. Some work has been done in this direction over rainfed areas, focusing on modeling the quantity and quality of the drained water discharged (e.g., Negm et al., 2017), or on developing drainage scenarios that integrate changes in agricultural practices (e.g., Tournebize et al., 2004) or in climatic conditions (e.g., Golmohammadi et al., 2020; Jeantet et al., 2022). However, only a few studies have dealt with the quantitative estimation of drainage in semi-arid irrigated areas. At the field scale, Ale et al. (2013) compared the ability of the physically-based DRAINMOD (Skaggs et al., 2012) and ADAPT (Gowda et al., 2012) models to simulate monthly drainage in a drip-irrigated plot in the US. The determination coefficient between simulated and observed drainage was 0.90 and 0.85 for ADAPT and DRAINMOD, respectively, for data during seven years. More recently, Feng et al. (2021) simulated the daily drainage with the physically-based Hydrus-2D model in a furrow-irrigated plot and obtained a Nash-Sutcliffe model efficiency coefficient (NSE) value of 0.91 and 0.94 for the calibration and validation year respectively. At a larger spatial scale, Cavero et al. (2012) simulated the monthly drainage of three Mediterranean irrigated catchments (mainly surface irrigation) located in Spain, Algeria, and Turkey, ranging from 4000 to 10000 hectares. They used the crop water balance model APEX (Gassman et al., 2009) coupled with DRAINMOD over two hydrological years and obtained Root-mean-square deviation (RMSD) ranging from 3.4 mm to 25.3 mm per month. On a much larger scale, Wen et al. (2020) simulated the monthly drainage over 19 sub-basins in a 1.2 million-hectare irrigation district in northern China using an empirical approach based on water table observations. The results showed a mixed performance with an average NSE of 0.64, a standard deviation of 0.21 for the 2-year calibration period, an average NSE of 0.34, and a standard deviation of 0.44 for the 2-year validation period. In the same irrigation district, Chang et al. (2021) simulated the

annual drainage discharge under different management scenarios using the semi-empirical SAHYSMOD model (Oosterbaan et al., 2005).

Although the literature on agricultural drainage is still limited for semi-arid irrigated areas, there are many scientific papers on drainage estimation in humid regions where waterlogging problems are common. The vast majority of them use physically-based models such as DRAINMOD (Moursi et al., 2022; Muma et al, 2015), RZWQM2 (Ma et al., 2012; Xian et al., 2017; Jiang et al. 2020), MACRO (Larsbo et al., 2005; Jarvis and Larsbo, 2012), SWAP (van Dam et al., 2008), HydroGeoSphere (De Schepper et al., 2015), and FLUSH (Turunen et al., 2013; Nousiainen et al., 2015). Generally, these models are implemented at the plot scale and represent the intermediate processes (e.g., macropore infiltration, deep seepage, water redistribution in the soil profile, rooting distribution, lateral flows) involved in drainage at the daily or hourly time step. They rely on a lot of information for model parameterization, implying a detailed knowledge of the studied site and potentially numerous parameters to calibrate. For example, Ma et al. (2012) recommend for the RZWQM2 model an independent measurement for 11 parameters and a calibration for 11 others out of a total of 24 parameters (the two remaining ones being taken from the literature). In fields where intensive measurement campaigns have been conducted, these models can simulate the observed drainage well at hourly, daily, weekly, or monthly scales. However, the application of such models to poorly monitored basins remains limited due to the need for site-specific calibration (using drainage measurements) to set their relatively numerous input parameters.

Henine et al. (2022) proposed a simple semi-empirical drainage model, RU-SIDRA, to generalize a drainage model for various agricultural conditions. It combines a surface model (RU) to simulate the daily recharge and a subsurface model (SIDRA) (Lesaffre and Zimmer, 1987; Bouarfa and Zimmer, 2000) to convert the simulated recharge into daily drainage discharge. RU is a water balance model based on a simplified version of the FAO-56 method (Allen et al., 1998) and relies only on a single sensitive parameter. SIDRA is based on the resolution of a semi-analytical formula derived from the Boussinesq physical equation (Boussinesq, 1904), leading to two main sensitive parameters. The robustness of RU-SIDRA was evaluated by Jeantet et al. (2021) on 22 non-irrigated French fields and sub-basins over 200 hydrological years. It was found that RU-SIDRA performs as well as physically-based models in reproducing daily drainage and is as robust as the latter from one hydrological year to another. However, Jeantet et al. (2021) emphasized a limitation of RU-SIDRA associated with the empirical nature of the RU model. Indeed, in RU, the start of drainage, occurring in autumn in the non-irrigated sites studied by Jeantet al. (2021), is not systematically well reproduced from year to year. This is potentially due to its poor representation of the processes governing the variation of the soil water stock (e.g., root growth or evapotranspiration). Over irrigated areas, this difficulty is expected to be further exacerbated by the impact of summer crops. In addition, as the physically based models described previously, RU-SIDRA also relies on a site-specific calibration step using rarely available observed drainage data.

Finding the right balance between the simplicity required for a drainage model to be easily applicable over large areas and the complexity needed to ensure a sufficient realism of intermediate processes and, hence, the robustness of drainage estimates in time remains challenging. In this context, this study seeks to address the following questions:

- Can parsimonious models with different degrees of complexity precisely reproduce the daily drainage in a semi-arid irrigated context with site-specific calibration?

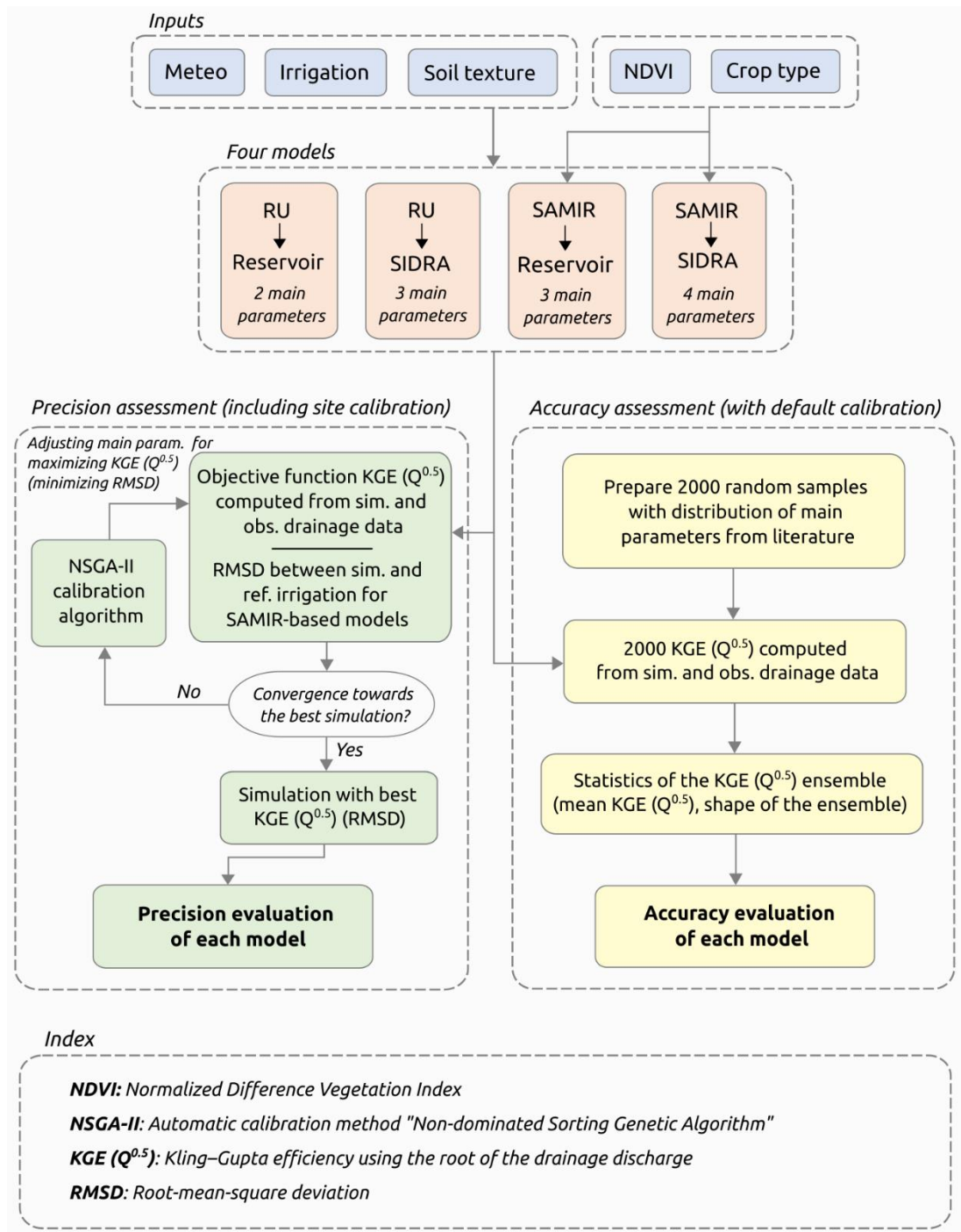
- Can such models with default calibration (with parameter values provided in the literature) reproduce drainage quantities and dynamics, even roughly?

In this context, we evaluated the precision and accuracy of several parsimonious models based on the RU-SIDRA formalism. By precision and accuracy, it is meant the capacity of each model to predict the drainage after site-specific calibration (using drainage measurements) and by setting the model input parameters to default values found in the literature (without using drainage measurements), respectively. Precision evaluation aims to investigate the models' strengths and weaknesses by calibrating and validating them over the same period. The accuracy evaluation aims to determine i) whether it is possible to estimate drainage when no in situ drainage data is available for calibration (which is the case for most irrigation districts), i.e., under non-optimal calibration conditions, and ii) which of the models evaluated performs best under these conditions.

To cover a range of modeling complexity, we investigate the SAMIR model (Simonneaux et al., 2009) as an alternative to the RU model. SAMIR is more complex than RU as it simulates more processes (e.g., root growth, vegetation development, evaporation, vegetation cover, specific crop water needs, and stress resistance) while remaining parsimonious with only two parameters integrating most of the sensitivity for the recharge simulation (Laluet et al., 2023). We also investigate the Reservoir model as an alternative to the SIDRA model. Reservoir is a fully empirical model driven by a single parameter. It is a simplified version of a module recently incorporated in the SASER (Safran-Surfex-Eaudysee-Rapid) hydrological model (Quintana-Seguí et al., 2017; Vergnes and Habets, 2018; David et al., 2011) by Cenobio-Cruz et al. (2023) who showed its ability to satisfactorily reproduce the low flows (groundwater discharge) observed at 53 hydrological stations in France and Spain.

The combination of both recharge (RU and SAMIR) models and both subsurface (Reservoir and SIDRA) models generate four drainage models by order of increasing complexity: RU-Reservoir (two main parameters), RU-SIDRA (three main parameters), SAMIR-Reservoir (three main parameters), SAMIR-SIDRA (four main parameters). In this study, the precision and accuracy evaluations of the four drainage models are carried out in two sub-basins of the Algerri-Balaguer irrigation district located in the Ebro basin, northeast of Spain. These two sub-basins are instrumented with flow meters that continuously measure the daily drainage discharge into the main drains.

2 Material and methods

135 The overall methodology to assess both the precision and accuracy of RU-Reservoir, RU-SIDRA, SAMIR-Reservoir, and SAMIR-SIDRA is presented in the flowchart of Fig. 1. First, the study site and data used are presented (Section 2.1), followed by a description of the four models (Section 2.2). The following sections describe the site-specific calibration strategy used for evaluating the precision of the models (Section 2.3) and the selection of input parameter ranges used for evaluating the accuracy of the models (Section 2.4).

140 **Figure 1.** Flowchart of the proposed methodology for precision and accuracy evaluation of the drainage simulated by four parsimonious models with different levels of complexity.

2.1 Study area and data

2.1.1 Study area

The Algerri-Balaguer (AB) irrigation district is located in northeast Spain, 20 km north of Lleida. It is characterized by a semi-arid continental Mediterranean climate with an average annual reference evapotranspiration (ET_0) of 1027 mm and precipitation of 380 mm (2000-2021). AB has an area of 8100 hectares of cropland with mainly corn, barley, wheat, fruit trees, and alfalfa. 6800 hectares are equipped for irrigation with sprinklers for annual crops and drip systems for fruit trees. An overview of the AB area is shown in Fig. 2. For an extensive description of the irrigation district in terms of soil, geology, crops, irrigation, and drainage system, the reader is referred to Altés et al. (2022).

In 1998, modernization works were carried out in the AB district, including flattening the land for plot consolidation and installing irrigation systems and a drainage network. The drainage network consists of surface (open ditches) and subsurface (buried pipes) drains (Altés et al., 2022). Field drains (underground perforated plastic pipes) are connected to collectors (underground concrete pipes larger than field drains). These collectors, in turn, are connected to main drains (either larger underground concrete pipes or open ditches). The main drains ultimately convey the water to general outlets (green dots in Fig. 2). During irrigation implementation, the collectors and main drains were installed in the first few years. Since then, field drains have been installed progressively at the initiative of each farmer according to his needs. We have no precise information on the surface that has been equipped with field drains or on their spacing. From field observations, we know that the surface drains were dug at a depth of approximately 2 m, as were the main drains.

Two of these main drains have been equipped with CTD-10 sensors (Meter Group Inc., Pullman, WA, USA) that continuously measure the water level. They collect drainage water from areas of 116 and 2050 hectares each, forming two sub-basins, AB1 and AB2 (see Fig. 2). These areas correspond to the topographic basins formed by the main drains at the CTD-10 sensors locations and were computed using the QGIS software with a 2 m resolution DEM provided by the Cartographic and Geological Institute of Catalonia. Table 1 shows, for AB, AB1, and AB2, the area and percentages of the main crop types. Figure 2.b shows the land cover of AB1 and AB2 for 2021.

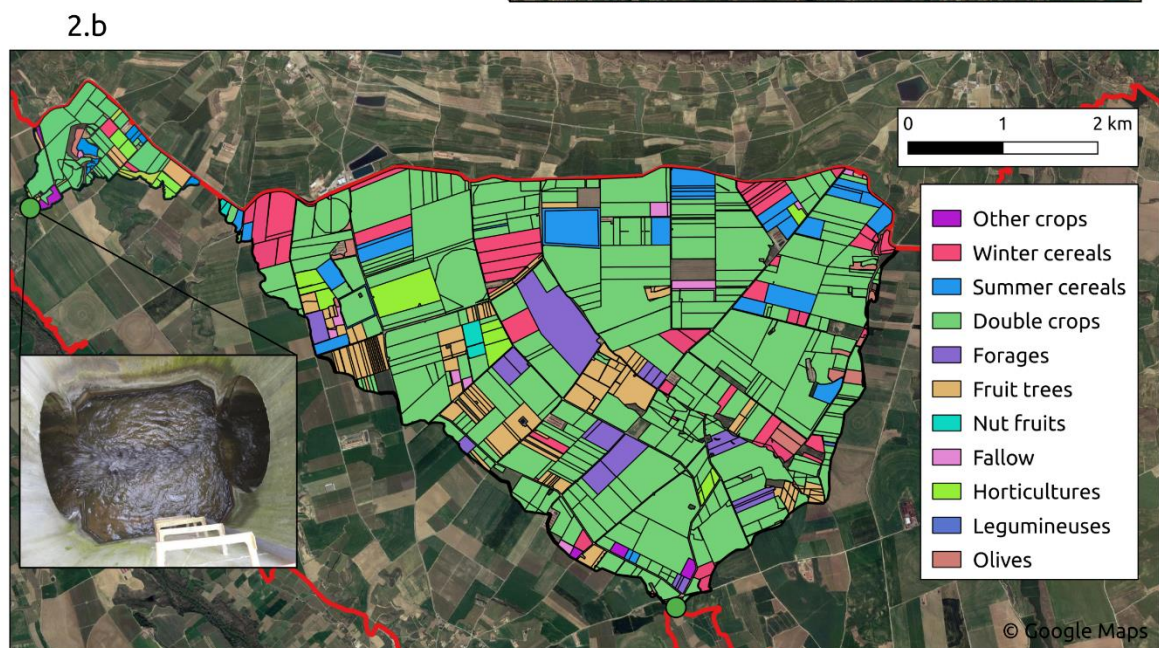
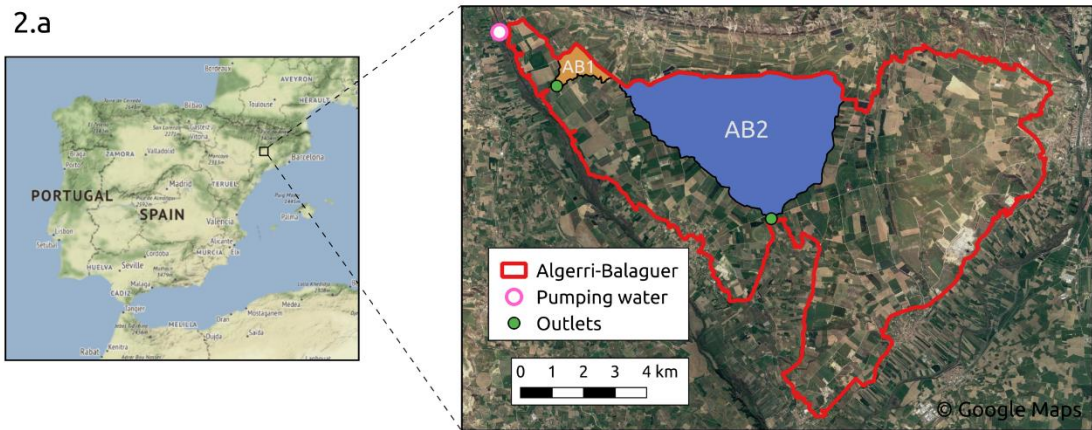



Figure 2. The AB irrigation district with the two monitored sub-basins AB1 and AB2 and their outlets, as well as the location of the pumping station for irrigation (coordinates: 797157/4636983.4 ETRS89/UTM zone 31N) (a); Zoom on the two sub-basins with their land use for the year 2021 and a picture of inside the AB1 outlet where water level is measured before being converted into drainage discharge (b).

Table 1. Irrigated surfaces of AB, AB1, and AB2, and percentage of surface area occupied by their main crops in 2021 and 2022.

	Irrigated surface (hectares)	Year	Percentage of surface occupied by...		
			Double crop (mostly wheat or barley in winter and maize in summer)	Summer cereal (mainly maize)	Others (alfalfa, winter cereals, olives, fruit trees, etc.)
AB	6800	2021	58 %	8 %	34 %
		2022	69 %	7 %	24 %
AB1	116	2021	59 %	6 %	35 %
		2022	69 %	6 %	25 %
AB2	2050	2021	72 %	6 %	22 %
		2022	78 %	4 %	18 %

175

2.1.2 Description of the data used in this study

The water levels measured in the AB1 and AB2 outlets are obtained hourly and converted into daily discharge using the Manning-Strickler equation and the knowledge of the main drain hydraulic characteristics (Altés et al., 2022). The available drainage data used herein cover the period from February 2021 to October 2022 (21 months) for AB1 and from May 2021 to October 2022 (18 months) for AB2. The observed drainage at AB1 was 29 mm in 2021 and 26 mm in 2022, while it was 61 mm in 2021 and 46 mm in 2022 at AB2.

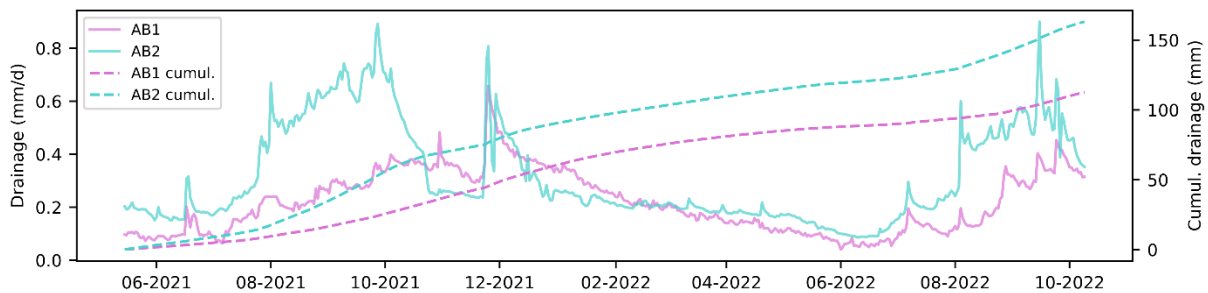



Figure 3. Bottom: Daily drainage data of AB1 and AB2 from May 2021 to October 2022.

The irrigation data consists of the daily flow of water pumped from a river next to the AB district (see Fig. 2), which is the only supply for the irrigation network. They are provided by the Automatic Hydrological Information System of the Ebro Basin (SAIH). Pumping flow data are aggregated to the weekly scale to consider the potential delay of several days between pumping and application in the field. 5.8% of the volume is removed to account for evaporation loss and leakage, based on a comparison between the water pumped from the river and irrigation data from water meters (Olivera-Guerra et al., 2023).

Soil texture is obtained from the 250-meter resolution SoilGrids product (Hengl et al. 2017; Poggio et al. 2021). It is relatively uniform over the AB area and corresponds to a silty clay loam soil (Jahn et al., 2006).

Meteorological data are obtained from five stations belonging to the Catalan Meteorological Station Network. Two are located within the AB district, and the three others are located around the area at a maximum distance of 5 km. The mean and standard deviation of the instantaneous measurements of precipitation and ET_0 made by the five stations are very low. Therefore, the spatial average of precipitation and ET_0 measurements is used as forcing at the scale of the AB sub-basins.

2.2 Description of the four models

The four models evaluated herein result from the combination of two water balance models (RU and SAMIR) and two drainage discharge models (Reservoir and SIDRA). Their main characteristics are listed in Table 2.

Table 2. Description of the four models used.

	RU	SAMIR	Reservoir	SIDRA
Description	Simple water balance model	FAO-2Kc-based crop water balance model	Exponential emptying of a reservoir	Semi-analytical formula derived from the Boussinesq equation
Inputs	ET_0, rainfall, irrigation, soil texture type	ET_0, rainfall, soil texture, land cover, $NDVI$	Recharge	Recharge, drainage network characteristics, soil texture type
Parameters	2 crop parameters 3 soil parameters	7 crop parameters 2 soil parameters 2 irrigation parameters	1 depletion coefficient parameter	2 drainage network geometry parameters 2 soil parameters
Sensitivity	1 param. most sensitive (S_{inter}) (Henine et al., 2022; Chelil et al., 2022)	2 param. most sensitive (Zr_{max} and a_{Kcb}) (Laluet et al., 2023)	1 param. sensitive (k)	2 param. most sensitive (K and μ) (Henine et al., 2022; Chelil et al., 2022)

2.2.1 SAMIR

205   The SAMIR model (Simonneaux et al., 2009) is a FAO-56 double crop coefficient-based model (FAO-2Kc) (Allen et al., 1998) designed to simulate the crop water balance components for daily ET estimation and crop water requirements by considering the plant and soil water status. It uses i) meteorological forcing variables to calculate ET_0 (calculated using the Penman-Monteith equation), ii) precipitation, iii) crop and soil parameters to calculate soil reservoir properties, as well as plant and soil resistance to water stress, and iv) Normalized Difference

210   Vegetation Index ($NDVI$) to drive plant development, obtained from the Sentinel-2 satellites with a resolution of 10 m and a revisit of five days.

The daily water balance equation simulated with SAMIR is:

$$Dr_t = Dr_{t-1} + ET_t - P_t - I_t + R_t \tag{1}$$

where Dr is the root zone depletion, ET is the actual evapotranspiration, P is the precipitation, I is the irrigation,

215   and R is the underground recharge. Every term is expressed in mm for the day t (and t-1 for Dr). ET is estimated by multiplying two crop coefficients to ET_0 as follows:

$$ET_t = (Kcb_t \cdot Ks_t + Ke_t \cdot Kr_t) \cdot ET_{0_t} \tag{2}$$

where $ET_0 \cdot Kcb \cdot Ks$ is the water transpired by plants (T, mm), and $ET_0 \cdot Ke \cdot Kr$ is the soil evaporation (E, mm). Kcb (-) is the basal crop coefficient governing the potential crop transpiration. It is estimated from a linear relationship

220   with $NDVI$. Ks (-) is the water stress coefficient reducing the potential transpiration, Ke (-) is the potential soil evaporation coefficient, and Kr (-) is the evaporation reduction coefficient.

Kr is calculated with a pedotransfer function using clay and sand fractions that were derived and evaluated over a variety of sites (Lehmann et al., 2018; Merlin et al., 2016) and recently implemented into SAMIR by Amazirh et al. (2021).

225   Ks is calculated based on the daily computation of the water balance in the root-zone layer, as follows:

$$Ks_t = \frac{TAW_t - Dr_t}{TAW_t(1-p)} \tag{3}$$

where Dr is calculated from the daily water balance according to Eq. (1), TAW (mm) is the maximum available water in the root zone, and p (-) is the fraction of TAW that a crop can extract without facing water stress. Allen et al. (1998) suggest that p controls the water depth threshold below which irrigation should be triggered to avoid

230   crop water stress by keeping Dr smaller than $TAW \cdot p$ (and thus Ks equal to 1). TAW is estimated as follows:

$$TAW_t = (SM_{FC} - SM_{WP}) \cdot Zr_t \tag{4}$$

where SM_{FC} ($\text{m}^3\,\text{m}^{-3}$) is the soil moisture at field capacity and SM_{WP} ($\text{m}^3\,\text{m}^{-3}$) the soil moisture at the wilting point, both derived from the soil texture by applying the pedotransfer function proposed by Román-Dobarco et al. (2019). Zr (mm) is the rooting depth that varies between a minimum value (set to 100 mm for annual crops) and a crop-

235   dependent maximum value (reached at the maximum $NDVI$ of the simulated field).

A spatialized version of SAMIR at the plot scale was recently developed and is used in this study to simulate recharge at AB1 and AB2. More details on the methodology behind this spatialization can be found in Olivera-Guerra et al. (2023).

To simulate plot-scale irrigation using SAMIR, we used the method proposed by Olivera-Guerra et al. (2023), which consists of inverting two time-varying SAMIR irrigation parameters from the irrigation data measured at the pumping station. By applying the inverted parameters to each irrigated field and aggregating the resulting simulations over the entire AB district, the simulated irrigation volumes and timing were close to those measured at the pumping station (RMSD < 0.70 mm d^{-1} on average for six irrigation seasons). The values of the SAMIR irrigation parameters found by Olivera-Guerra et al. (2023) for the AB district were used herein for 2021 and 2022. The plot-scale irrigations simulated by SAMIR are then averaged for AB1 and AB2 to be used as forcing in the RU model (as RU is not spatialized).

2.2.2 RU

RU is a water balance model designed to simulate the recharge of the water table. It is one of the components of the RU-SIDRA model introduced by Henine et al. (2022). In contrast to SAMIR, RU has not been designed to precisely reproduce *ET* by simulating plant phenology or processes related to evaporation. Its purpose is to reproduce the correct amount of recharge to be converted into drainage with the SIDRA model. While a detailed description of the model can be found in Henine et al. (2022) and its evaluation in Jeantet et al. (2021), only a general overview is provided here.

RU uses as input precipitation, irrigation, ET_0, and soil texture type (for default parameter values). It comprises a module simulating the net infiltration (P_{net}, mm) and a soil reservoir module transforming P_{net} into recharge R.

P_{net} is calculated as follows:

$$P_{net_t} = P_t + I_t - CET_t \tag{5}$$

With *CET* (mm) being the corrected *ET*, which is computed as follows:

$$CET_t = \begin{cases} ET_{0_t} \cdot e^{-\frac{S_{RFU} - S_t}{S_t}} & if\, S_t < S_{RFU} \\ ET_{0_t} & if\, S_t \geq S_{RFU} \end{cases} \tag{6}$$

where S_t (mm) is the current water level in the soil reservoir on day t and S_{RFU} (mm) is a water level threshold triggering the plant water stress limiting ET_0. S_{RFU} is $0.4 \cdot S_{inter}$ with S_{inter} (mm) being a threshold in the soil reservoir that triggers recharge, analogous to SM_{FC} in SAMIR.

The soil reservoir module is designed to simulate recharge (R) depending on three stages related to the amount of water in the soil reservoir (S):

$$Stage\ 1: S_t < S_{inter};\ R_t = 0 \tag{7}$$

$$Stage\ 2: S_t \in [S_{inter}; S_{max}];\ R_t = \beta \cdot P_{net_t} \tag{8}$$

With S_{max} (mm) being the water level below which recharge occurs with a reduction coefficient β (−) and calculated as follows:

$$S_{max} = S_{inter} + S_{IDS} \tag{9}$$

270   Where S_{IDS} (mm) is the intense drainage season reservoir level, reached at the season period where drainage is most important due to large amounts of precipitation and/or irrigation (Jeantet et al., 2021). During this time, the level of the reservoir is higher than that of S_{inter}. Henine et al. (2022) and Chelil et al. (2022) found that S_{IDS} and β are not significantly sensitive. Based on the values used in Jeantet et al. (2021), S_{IDS} was set to 20 mm and β to 0.33.

275     $Stage\ 3: S_t > S_{max}; R_t = P_{net_t}$ (10)

Note that RU is not spatialized, implying that a single simulation is performed for AB1 and AB2 separately with average forcings and parameters.

2.2.3 SIDRA

SIDRA is a physically-based model designed to calculate the drainage flow of a drained plot or sub-basin. It is
280   based on the resolution of a semi-analytical formula derived from the Boussinesq equation, which leads to Eq. (11) and Eq. (12). For a complete description of SIDRA, readers are referred to Tournebize et al. (2004), Henine et al. (2022), and Zimmer et al. (2023).

First, the water table level variation under the influence of recharge (R) and drainage is computed as follows:

$$\frac{dh_t}{d_t} = \frac{R_t - K\frac{h_t^2}{L^2}}{C\mu}; h_{t+1} = h_t + \frac{dh_t}{d_t} \tag{11}$$

285   where h is the water table at the midpoint between drains (m), K is the horizontal hydraulic conductivity (m d^{-1}), μ the drainable porosity (m^3 m^{-3}), and C is a water table shape factor (-) equal to 0.904. h is bounded between 0 and 1.5 (the average drain depth assumed at AB). Eq. (11) is based on the assumption that the water table is flat, which is consistent with the flat topography of the AB district. The drainage of the water table into buried pipes or open drains, leading to drainage flow Q, is calculated as follows:

290     $Q_t = AK\frac{h^2}{L^2} + (1-A)R_t$ (12)

where L is the half of the drain spacing (m) and A is a water table shape factor (-) equal to 0.896. As we do not have precise information on the location of the field drains and the proportion of the surface equipped with them at AB, L can be either calibrated with drainage data or set to a value frequently found in the literature (generally between 3 and 12 m).

295   Boussinesq's equation assumes that buried pipes and open drains rest on an impermeable layer, meaning that the entire water table could be drained after a given period without rain or irrigation. The hydrogeological

configuration of the AB district, with the presence of a shallow impervious layer, allows us to assume that this condition is respected.

SIDRA is not spatialized; therefore, when combined with SAMIR (which is spatialized, unlike RU), it uses the average daily recharge from all the simulated plots of AB1 and AB2 as input.

2.2.4 Reservoir

Reservoir (Cenobio-Cruz et al., 2023) is a conceptual model designed to reproduce the delay between the recharge and the water table draining into a river, buried pipes, or open drains. The idea of this model is that a reservoir filled by recharge is drained according to a linear relationship between the water level Z (mm) and a depletion coefficient ω (-). It can be expressed as follows:

$$Q_t = Z_t \cdot \omega \tag{13}$$

Cenobio-Cruz et al. (2023) incorporated a reservoir size parameter to simulate overflow and generate quick flows. We decided not to include this parameter to make the Reservoir model as simple as possible.

2.2.5 Initialization of the state variables

The four models, RU-Reservoir, RU-SIDRA, SAMIR-Reservoir, and SAMIR-SIDRA, require initialization of their state variables. These variables were initialized with a 12-month spin-up simulation using data from 2020. SAMIR initializes the depletion parameters of the soil and surface reservoirs, RU the water level in the soil reservoir, SIDRA the water table level, and Reservoir the water level in the conceptual reservoir.

2.2.6 Sensitivity of model parameters

Laluet et al. (2023) conducted an extensive sensitivity analysis of ET and recharge simulated by SAMIR over various agro-pedoclimatic conditions. Two of the nine parameters were found to dominate the model sensitivity: a_{Kcb} governs the relationship between $NDVI$ and Kcb (related to T demand), and Zr_{max} is the maximum rooting depth that governs the size of the root zone reservoir.

Henine et al. (2022) and Chelil et al. (2022) analyzed the sensitivity of the RU-SIDRA parameters for drainage simulation on two different plots. They both showed that for the RU model, the S_{inter} parameter controls most of the model sensitivity and that for SIDRA, it is K and μ.

2.3 Strategy for evaluating the models' precision

2.3.1 Calibration strategy

It is reminded that precision means a model's ability to approximate the observed drainage data as closely as possible using site-specific calibration (using drainage measurements). The most sensitive parameters of the four models were calibrated using an automatic calibration algorithm. Since we do not have information on the half of the drain spacing on AB, we calibrated the L parameter, bringing the number of parameters to be calibrated for RU-SIDRA to four and for SAMIR-Reservoir to five. Indeed, although Chelil et al. (2022) showed that L is not

very sensitive when it varies between 3.5 m and 6 m, its uncertainty within the AB district is large enough for it to be significantly sensitive. The other less sensitive parameters are fixed at the default values given in the literature. To analyze the variability of the parameters from one hydrological year to the next, we split the data into 12 months named "2021 period" and a period with the remaining months (9 months for AB1, 6 months for AB2) called "2022 period". Analysis of the variability of the values of the calibrated parameters between the two periods provides information on the predictive capacity of the models. If the values are close from one period to another, this suggests that the model robustness is high. If they are not, this indicates a low level of robustness.

The calibration method used is the multi-objective, Non-dominated Sorting Genetic Algorithm (NSGA-II) (Deb et al., 2002). For the case of SAMIR, which simulates both irrigation and recharge, we use a multi-objective method to ensure that the calibration of a_{Kcb} and Zr_{max} parameters does not significantly modify the simulated irrigation. Therefore, for both SAMIR-Reservoir and SAMIR-SIDRA, drainage and irrigation are optimized together, whereas, for both RU-Reservoir and RU-SIDRA, only drainage is optimized (an averaged irrigation is given as forcing in this case as RU cannot simulate irrigation). NSGA-II is one of the most widely used multi-objective algorithms. It implements a fast, non-dominant sorting approach to discriminate solutions based on dominance and Pareto optimality. It provides a set of optimal non-dominated solutions (set of parameters), allowing the user to choose the best solution according to his priorities. In the SAMIR case, the best solution would be the one that simulates the most precise drainage, provided that it simulates irrigation consistent with the observed data at the pumping station. Readers are referred to Deb et al. (2002), Bekele and Nicklow (2007), and Shafii and De Smedt (2008) for a detailed description of the algorithm.

2.3.2 Parameters distribution for calibration

NSGA-II requires a distribution provided by the user for each calibrated parameter. The distribution and references used in this study are provided in Table 3.

The distribution of a_{Kcb} of the SAMIR model is based on Laluet et al. (2023a), who obtained a_{Kcb} values for 37 agricultural seasons (mainly maize and wheat, being widely present in the AB district) from the linear relationship $Kcb = a_{Kcb} \cdot NDVI + b_{Kcb}$. Knowing the value of $NDVI$ at bare soil (where Kcb is zero) and at full vegetation (where Kcb is equal to Kcb_{max}), a_{Kcb} and b_{Kcb} can be inferred.

The distribution of Zr_{max} is derived from tables provided by Allen et al. (1998) and Pereira et al. (2021). This study uses the mean and standard deviation of Zr_{max} for maize, being the most present and irrigated crop type at AB1 and AB2.

The distribution of S_{inter} of the RU model is taken from Jeantet et al. (2021), who calibrated this parameter on drainage discharge in situ data from 22 French drained sites.

The distribution of K and μ of the SIDRA model is also based on Jeantet et al. (2021), who derived the mean and standard deviation of these parameters from field measurements performed on 15 silty soils in France, similar to the soil texture of AB1 and AB2. The half of the drain spacing parameter L distribution was chosen to be wide (uniform distribution between 4 m and 60 m) to consider that some plots are not drained at AB1 and AB2. As a comparison, Chelil et al. (2022) used a uniform distribution between 3.5 m and 6 m for fully drained sites.

The distribution of the ω parameter of the Reservoir model is derived from Cenobio-Cruz et al. (2023), who calibrated this parameter on discharge flow data of about 25 catchments in northern Spain.

2.3.3 Metrics used for calibration and validation

For drainage calibration, the objective function Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) is used:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\delta-1)^2} \qquad (14)$$

370 where r is the Pearson's correlation coefficient, α the bias component, and δ the ratio of the discharge variance:

$$\alpha = \frac{m_s}{m_o} \ \ and \ \ \delta = \frac{\sigma_s}{\sigma_o} \qquad (15)$$

where m and σ are the mean and standard deviation, respectively. Subscripts s and o represent the simulated and observed flow, respectively.

Since KGE tends to place more weight on high flows (Santos et al., 2018), we used KGE ($Q^{0.5}$). KGE ($Q^{0.5}$) is the
375 KGE calculated from the squared roots of simulated and observed drainage, allowing the weights between high and low flows to be more balanced. Following Jeantet et al. (2021), we consider simulations as "excellent" when KGE ($Q^{0.5}$) is larger than 0.8, "very good" when it is larger than 0.7, "good" when it is larger than 0.6, "acceptable" when it is between 0.5 and 0.6, and "unsatisfactory" when it is below 0.5. Furthermore, to get a reference in mind, a KGE of -0.41 is equivalent to having a simulation performance equal to the average of the observed data (Knoben
380 et al., 2019).

The RMSD objective function is used for irrigation calibration:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{j}(\hat{y}_i - y_i)^2}{j}} \qquad (16)$$

where j is the number of days in the simulated time series, i is one day of the time series, $\hat{y}_i$ is the simulated time series, and y_i is the reference time series. We consider irrigation to be well simulated when the RMSD calculated
385 between the irrigation measured at the pumping station and the one simulated by SAMIR for all the plots in AB is below 0.70 mm d^{-1}. This value corresponds to the RMSD found on average from 2017 to 2021 by Olivera-Guerra et al. (2023).

Table 3. Distributions of the main parameters calibrated with the NSGA-II algorithm to evaluate the models' precision, and associated references.

	Parameter	Distribution	Reference
SAMIR	a_{Kcb} (-)	Normal (mean: 1.45; std. dev.: 0.12)	Laluet et al. (2023)
	Zr_{max} (mm)	Normal (mean: 1000; std. dev.: 170)	Allen et al. (1998)
			Pereira et al. (2021)
RU	S_{inter} (mm)	Normal (mean: 138; std. dev.: 53)	Jeantet et al. (2021)
SIDRA	K (m d^{-1})	Lognormal (mean: 0.99; std. dev.: 2.53)	
	μ (-)	Lognormal (mean: 0.018; std. dev.: 2.19)	
	L (m)	Uniform (low: 4; high: 60)	
Reservoir	k (-)	Normal (mean: 0.02; std. dev.: 0.05)	Cenobio-Cruz et al. (2023)

2.4 Strategy for evaluating the models' accuracy

Complementarily to the evaluation of the models' precision, this study also aims to assess the accuracy of the four models. It is reminded that by accuracy, we mean the ability of a model to approximate the observed data as closely as possible by relying only on default values given by the literature for its main parameters, i.e., without any site-specific calibration step.

To this end, for each of the four models, 2000 sets of their most sensitive parameters are generated randomly using a Monte Carlo sampling, with the distributions presented in Table 3, except for the half of the drain spacing L. Indeed, for the accuracy evaluation, we consider a situation where we have no information on the geometry of the drainage network and therefore on L. In this hypothetical situation we don't know that a portion of the surface is not drained, resulting potentially in high L values. Therefore, we use a value of 6 meters for the accuracy evaluation, which is a value frequently found in the literature (Jeantet et al., 2021). In addition, to focus only on the drainage accuracy evaluation, the irrigation obtained with SAMIR during the precision evaluation step is injected into SAMIR as a forcing. The KGE ($Q^{0.5}$) obtained with a simulation performed with average default parameters is calculated, and the ensemble generated by the 2000 Monte-Carlo simulations is analyzed.

2.5 Complexity of models' calibration

The models involving SAMIR are more complex to calibrate than those involving RU. This is due in particular to the fact that i) SAMIR-based models involve input data that are not always readily available at the required fine resolution (in particular land use maps), ii) SAMIR is spatialized and therefore requires more computing resources (several hours of computation for 2000 simulations on the AB district with 8 GB RAM and 4 CPUs running in parallel). The RU-based models are simpler to calibrate than the SAMIR-based ones because RU requires only

meteorological data as input, they are not spatialized and demand fewer computing resources (a few minutes to run 2000 simulations on the AB district with the same computing configuration).

415 The two subsurface models are simple and require few computing resources. SIDRA-based models are slightly more complex to calibrate than Reservoir-based models as they require two drainage network characteristic parameters (half of the drain spacing and depth of drains) as input. However, they are not the most sensitive parameters in the SIDRA model (Henine et al., 2022; Chelil et al., 2022).

We see a gradient in terms of the level of complexity and expertise required to calibrate the models: from SAMIR-
420 SIDRA to RU-Reservoir. The four models are coded in Python, as is the NSGA-II calibration algorithm provided by the Python package spotpy.

3 Results and discussion

3.1 Precision evaluation

This section will first give a quick overview of the irrigation simulated by SAMIR and used as a forcing by RU. The precision of the drainage simulated by each model is then presented. Finally, we will explain why the model precision differs between the four models and provide some recommendations and perspectives.

3.1.1 Irrigation simulated by SAMIR

430 Table 4 shows for SAMIR-Reservoir and SAMIR-SIDRA the RMSD obtained between simulated and observed irrigation at the pumping station over all the irrigated plots of AB, resulting from the NSGA-II multi-objective calibration performed for AB1 and AB2. The average RMSD obtained is 0.35 mm d^{-1} for 2021 and 0.66 mm d^{-1} for 2022, in line with the quality criteria defined previously (< 0.70 mm d^{-1}). The average amount of irrigation simulated for the AB1 sub-basin is 592 mm on average for the period from May to October 2021 and 693 mm for
435 the period from May to October 2022. For AB2, it is 619 mm in 2021 and 720 mm in 2022. These amounts are fully consistent given the amounts of irrigation measured at the pumping station and the proportions of surface used for double crops and summer cereals in each AB1 and AB2 sub-basins compared to those in the entire AB district (see Table 2).

440 **Table 4.** RMSD between daily simulated and observed irrigation, seasonal (between May and October for each year) cumulated simulated irrigation obtained with site-calibrated SAMIR-Reservoir and site-calibrated SAMIR-SIDRA separately, and seasonal in situ irrigation over the whole AB district (including non-irrigated plots).

	Irrigation obtained for AB1 calibration				Irrigation obtained for AB2 calibration				In situ irrigation (entire AB district)	
	2021		2022		2021		2022		2021	2022
	RMSD (mm d^{-1})	Average amount (mm)	RMSD (mm d^{-1})	Average amount (mm)	RMSD (mm d^{-1})	Average amount (mm)	RMSD (mm d^{-1})	Average amount (mm)	Average amount (mm)	Average amount (mm)
SAMIR-Reservoir	0.37	595	0.69	693	0.34	621	0.62	719	509	587
SAMIR-SIDRA	0.32	590	0.65	694	0.38	617	0.66	720		

3.1.2 Which model is more precise?

445 Figures 4 and 5 compare the simulated and observed drainage for the AB1 and AB2 sub-basins, respectively, and the 2021 and 2022 periods. Table 5 shows that among the 16 combinations of model/sub-basin/period, nine have KGE ($Q^{0.5}$) considered as "excellent", four "very good", two "good", and one "acceptable", and there is no "unsatisfactory" simulation. RU-Reservoir stands out as the most precise model (mean KGE ($Q^{0.5}$) = 0.87), followed by SAMIR-Reservoir (0.79), RU-SIDRA (0.76), and SAMIR-SIDRA (0.68). From these results, two

450 highlights stand out:

i) The models based on Reservoir are more precise in terms of drainage simulations than those based on SIDRA. This is particularly true for the AB1 sub-basin.

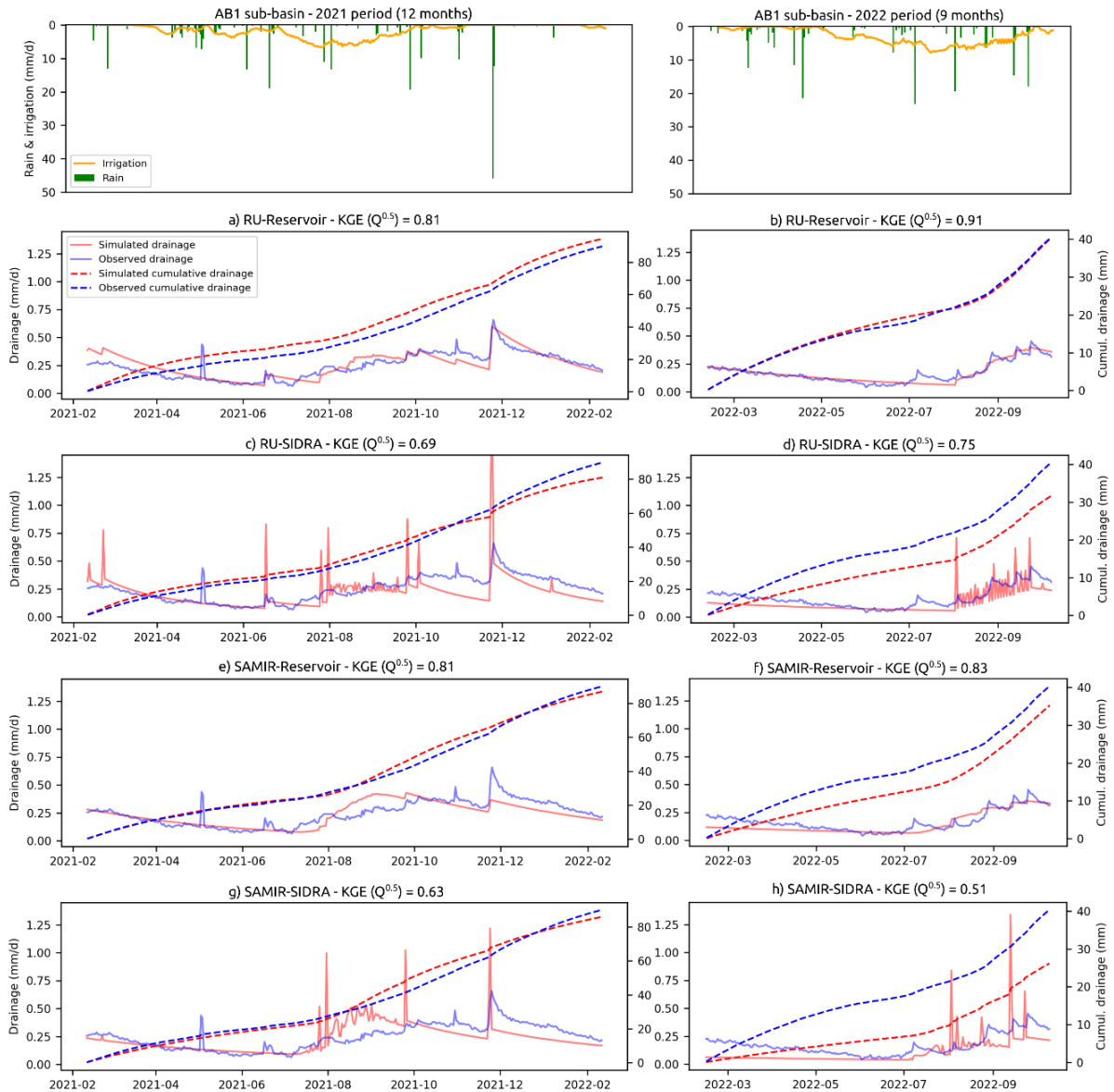
ii) The models based on RU are more precise in terms of drainage simulations than those based on SAMIR.

455

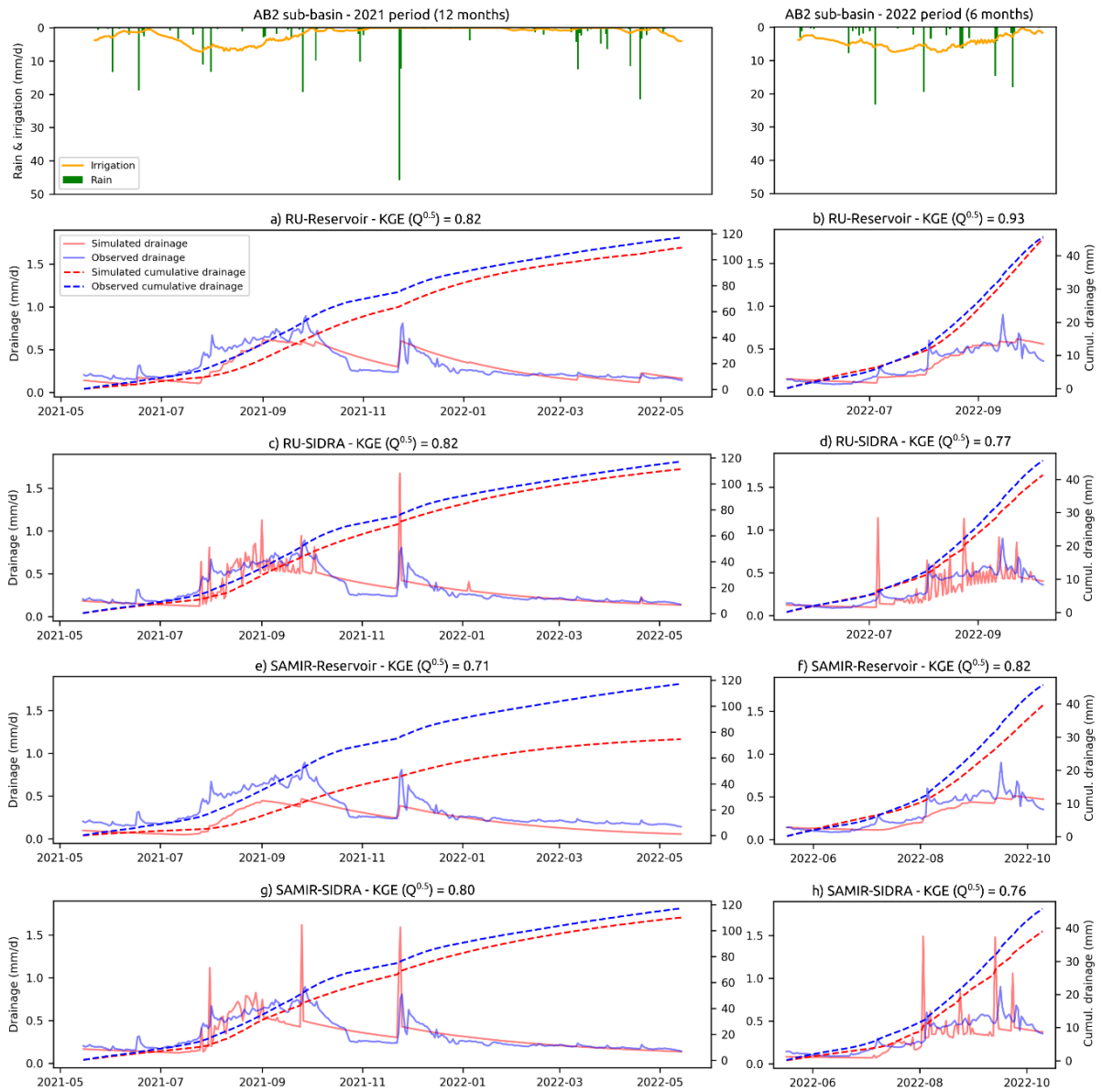
Figure 4. Daily and cumulated drainage of the AB1 sub-basin simulated by the four site-calibrated models for 2021 (left) and 2022 (right) periods. Plots at the top show the observed precipitation and the simulated irrigation.

460 **Figure 5.** Same as Fig. 4 but for AB2 sub-basin.

Table 5. KGE ($Q^{0.5}$) obtained with NSGA-II calibration for the AB1 and AB2 sub-basins, the four models, and both study periods. Values above 0.8 are considered as "excellent", between 0.7 and 0.8 as "very good", between 0.6 and 0.7 as "good", and between 0.5 and 0.6 as "acceptable".

	KGE ($Q^{0.5}$) for AB1		KGE ($Q^{0.5}$) for AB2	
	2021 period	2022 period	2021 period	2022 period
RU-Reservoir	0.81	0.91	0.82	0.93
RU-SIDRA	0.67	0.75	0.82	0.77
SAMIR-Reservoir	0.81	0.83	0.71	0.82
SAMIR-SIDRA	0.63	0.51	0.80	0.76

Table 6. Parameter values obtained with NSGA-II calibration for the AB1 and AB2 sub-basins, the four models, and both study periods.

	Parameters	AB1 / 2021 period	AB1 / 2022 period	AB2 / 2021 period	AB2 / 2022 period
RU-Reservoir	S_{inter} (mm)	13	76	16	27
	k (-)	0.015	0.008	0.014	0.007
RU-SIDRA	S_{inter} (mm)	10.5	94	28	51
	K (m d^{-1})	0.58	1.30	1.08	0.75
	u (-)	0.07	0.13	0.14	0.11
	L (m)	29	42	38	39
SAMIR-Reservoir	a_{Kcb} (-)	1.34	1.23	1.29	1.23
	Zr_{max} (mm)	1087	1130	1087	1004
	k (-)	0.009	0.004	0.012	0.005
SAMIR-SIDRA	a_{Kcb} (-)	1.36	1.19	1.23	1.24
	Zr_{max} (mm)	1072	1104	943	1079
	K (m d^{-1})	0.55	0.58	0.91	0.70
	u (-)	0.10	0.10	0.17	0.09
	L (m)	36	38	32	43

470

3.1.3 Why are Reservoir-based models more precise?

The differences in the formalism of Reservoir and SIDRA and the low responsiveness of the AB1 and AB2 hydrosystems explain the better performance of the site-calibrated RU-Reservoir and SAMIR-Reservoir models. Indeed, SIDRA was designed to simulate flow peaks followed by relatively steep recession curves. In contrast, Reservoir does not simulate peaks and generates flow according to a depletion coefficient ω that can be very low. However, the measured drainage dynamics for AB1 and AB2 are representative of low-responsive hydrosystems. This can be seen in Fig. 4 and 5, where the 40 mm rainfall in December 2021 generates a peak of less than 1 mm d^{-1} for both sub-basins, followed by a smooth recession curve with a discharge that never reaches 0 during the hydrological year. In comparison, the data used by Jeantet et al. (2021) from 22 French experimental sites, accounting for nearly 200 hydrological years, show winter peaks exceeding 20 mm d^{-1} in most of the studied years, followed by steep recession curves where 0 flow is reached in a few weeks. Therefore, the Reservoir model is favored by the low responsiveness of the AB1 and AB2 hydrosystems, especially for AB1. The SIDRA model shows a better precision for AB2 than AB1 because AB2 is more responsive with larger amounts of discharge. In addition, the values of the Reservoir depletion coefficient ω are low at our sites (0.009 on average) compared to those obtained by Cenobio-Cruz et al. (2022) (0.02 on average), which again reflects the relatively low responsiveness of the studied area.

The soil does not explain this low responsiveness since the mean calibrated values of K and μ (0.81 m d^{-1} and 0.11 m^3 m^{-3}, respectively; see Table 6) are consistent with the order of magnitude found in Jeantet et al. (2021) for a similar soil type (K from 0.1 to 1.8 m d^{-1}, and μ from 0.05 to 0.08 m^3 m^{-3}). The explanation seems to lie in the fact that the surfaces of AB1 and AB2 are not fully equipped with field drains. The transfer time from the recharge location to the main drain would be longer on non-equipped plots. The values of the calibrated half of the drain spacing L support this assumption, with an average of 37 m (see Table 6), being a value that could represent an average between low L values for plots equipped with field drains and high L values for plots that are not.

3.1.4 Why are RU-based models more precise?

To understand the better performance of the site-calibrated RU-Reservoir and RU-SIDRA models, in comparison with SAMIR-Reservoir and SAMIR-SIDRA, respectively, it is again necessary to look at the differences in formalism between RU and SAMIR. Indeed, RU, unlike SAMIR, has a stage during which a reduction factor β of 0.33 is applied to the recharge (Eq. (8); this stage is triggered by the water level in the soil reservoir). In AB1 and AB2, this stage is triggered, particularly during the irrigation season, and results in the spread of recharge amounts. This benefits RU-SIDRA, which, owing to the lower recharge events simulated by RU, generates lower peaks consistent with the drainage observations. This process is well illustrated in Fig. 4.d, where RU-SIDRA simulates numerous small peaks during the irrigation period, allowing better matching of the observations. In contrast, in Fig. 4.h, SAMIR-SIDRA simulates larger and fewer peaks.

Furthermore, the S_{inter} values of RU obtained through calibration are very low (17 mm on average for the 2021 period and 61 mm for the 2022 period; see Table 5) compared with those obtained by Jeantet et al. (2021) (138 mm on average). This is because the RU-based models are not spatialized and use average irrigation derived from

the irrigation simulated by SAMIR, whether for intensely irrigated corn plots generating a lot of recharge or non-irrigated plots generating no recharge. RU-based models simulate less recharge by simulating an average plot using average irrigation. To compensate for this, the NSGA-II optimization algorithm finds low S_{inter} values to reduce the reservoir size and, therefore, the ET, increasing the recharge. This explains the low S_{inter} values retrieved over AB1 and AB2.

3.1.5 Variability of calibrated parameter values between the two periods analyzed

We can see from Table 6 that the calibrated values of most parameters vary between the two periods for a given sub-basin. These variations indicate a lack of predictive capacity of the models, at least between the two periods analyzed (using the parameters values obtained from calibration on the first period the models fail to predict the second period). We believe this is due to the semi-empirical nature of the models. Indeed, parameter values vary between the two periods to compensate for the fact that physical processes (e.g., lateral subsurface flows, root growth, evapotranspiration) are either too empirically simulated or neglected.

3.1.6 Recommendations and perspectives

Based on the results obtained over AB by the four models with varying complexity levels, we recommend using RU-Reservoir when drainage data are available for calibration. The simplest model can reproduce the drainage observed at the AB1 and AB2 outlets fairly well. The RU-Reservoir model efficiently combines the performance of RU (allowing a better temporal distribution of recharge than SAMIR) and Reservoir (offering drainage simulations with relatively low responsive dynamics), which fits perfectly with the present study. However, if a study concerns a more responsive hydrosystem with larger peaks and steeper recession curves, we recommend using RU-SIDRA. Indeed, for the AB2 sub-basin, which is slightly more responsive than AB1, RU-SIDRA shows during the 2021 period (Fig. 5.c) a precision comparable to that of RU-Reservoir (Fig. 5.a). This leads us to assume that RU-SIDRA could be more appropriate than RU-Reservoir for even more responsive hydrosystems.

To support the strength of these recommendations and to ensure that the models can be ultimately used as decision support tools with confidence, we believe it is necessary to improve the robustness of the models by better simulating certain physical processes (e.g., lateral subsurface flows, root growth, evapotranspiration). Indeed, the variability of calibrated parameter values between the two periods indicates their limited robustness.

Since RU is not spatialized and has a very simple formalism, the models based on it are more straightforward to run and require fewer resources than those based on SAMIR. However, they do not consider the spatial heterogeneity in irrigation generally encountered in irrigated sub-basins, resulting in very low S_{inter} calibrated values to simulate enough recharge. Spatialization of the RU-based models with irrigation data for each plot would lead to higher calibrated S_{inter} values that are more consistent with those proposed in Jeantet et al. (2021). However, unlike SAMIR, RU cannot simulate irrigation, and its spatialization would require plot-scale irrigation data that are rarely available.

In some cases, SAMIR shows a relatively low KGE ($Q^{0.5}$) and difficulties reproducing the right amount of drainage (see Fig. 4.d, 4.h, and 5.e). Modifying the SAMIR formalism by taking inspiration from RU and adding a stage

related to the soil water availability in which recharge is limited by a factor β could help improve the precision of
545 SAMIR-SIDRA and SAMIR-Reservoir.

3.2 Accuracy evaluation

Figures 6 and 7 show for each of the 16 model/sub-basin/period combinations the drainage simulated by using the mean values of the default parameters as input (red line), and the 2000 model runs from randomly generated input
550 parameters sets within pre-defined distributions provided by the literature (gray lines). Table 7 summarizes the KGE ($Q^{0.5}$) obtained with the average default parameters. It indicates that all the 16 cases present "unsatisfactory" KGE ($Q^{0.5}$) values (below 0.5). Moreover, nine cases show KGE ($Q^{0.5}$) lower than -0.41, the value corresponding to the KGE obtained with the temporal average of the observed data.

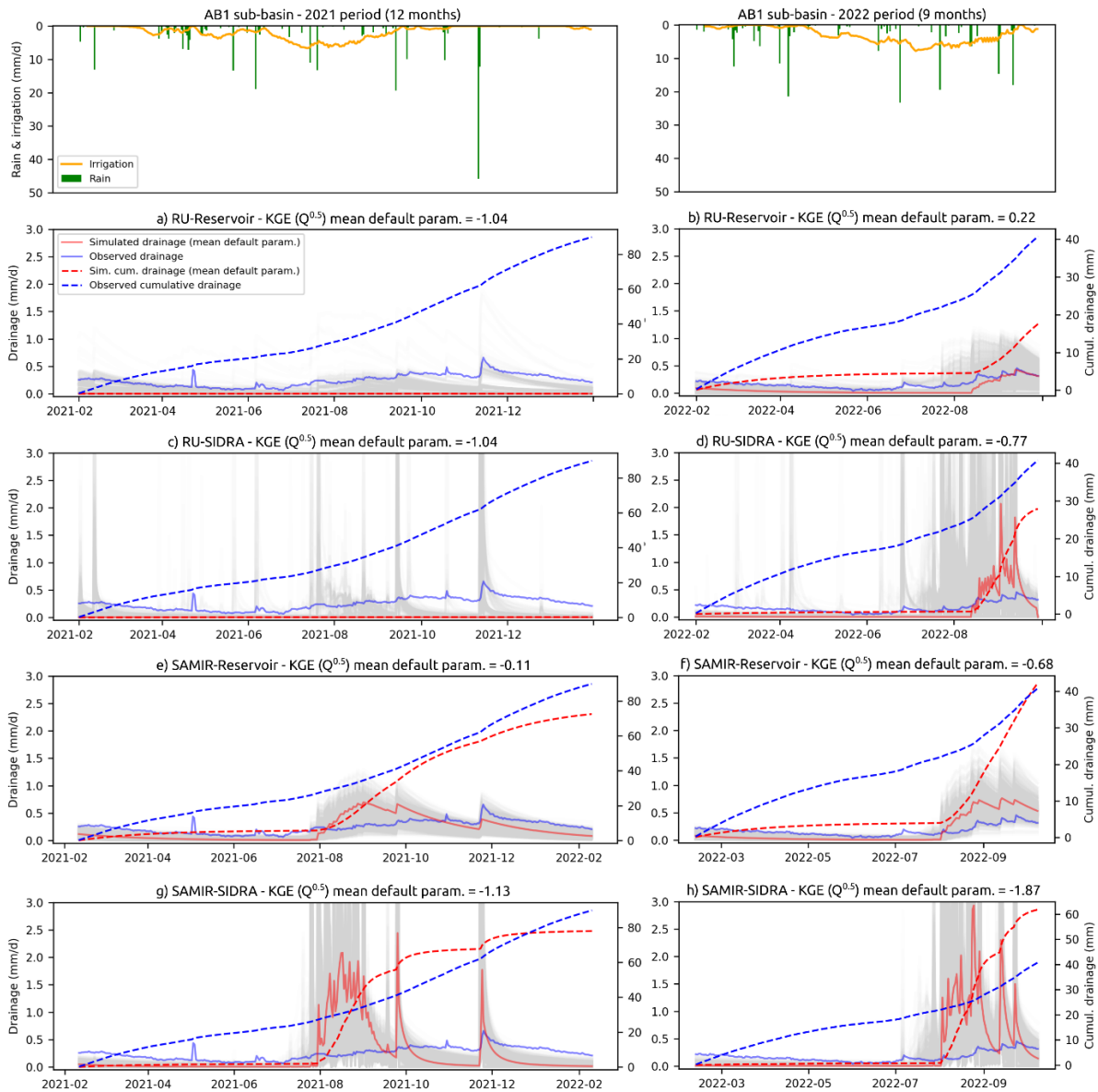
RU-Reservoir shows relatively satisfactory KGE ($Q^{0.5}$) for 2022 (0.22 for AB1 and 0.29 for AB2). However, when
555 looking at the drainage dynamics and amounts illustrated in Fig. 6.b and 7.b, it appears that these performances are due to the nature of the objective function KGE ($Q^{0.5}$), giving significant importance to low flows. SAMIR-Reservoir shows a relatively good KGE ($Q^{0.5}$) for AB2 for 2022 (0.29). Furthermore, the timing and quantities that SAMIR-Reservoir simulated for AB1 for both periods 2021 and 2022, as well as for AB2 for the 2021 period, are more consistent than those simulated by the other three models.

560

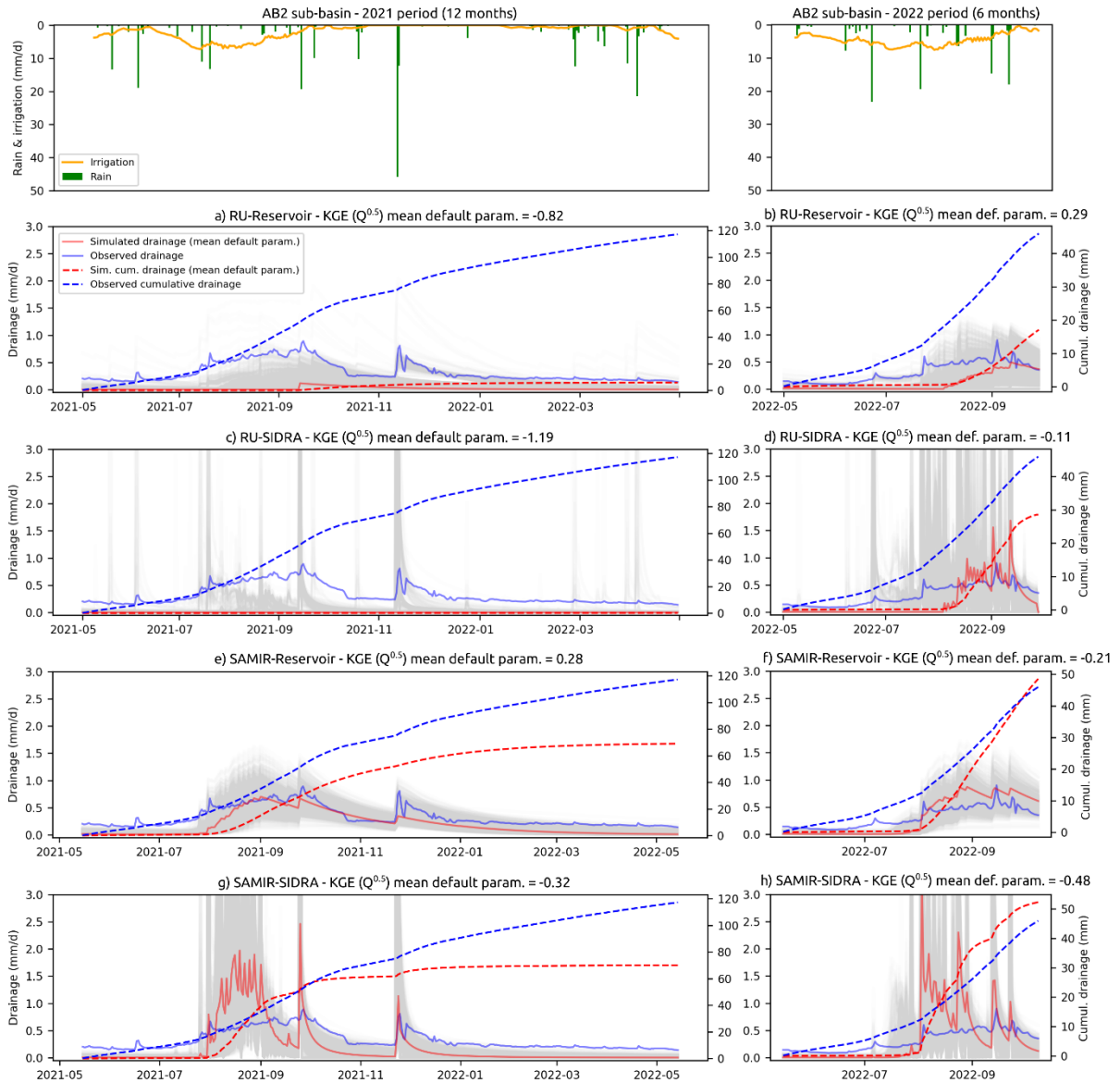
Figure 6. Daily drainage of AB1 sub-basin simulated using average default parameters of the four models separately (red line), and the drainage values ensemble obtained by running each model 2000 times using randomly generated input parameters sets within pre-defined distributions (gray lines).

565

Figure 7. Same as Fig. 6 but for the AB2 sub-basin.

Table 7. KGE ($Q^{0.5}$) values obtained with average default values of model parameters for the AB1 and AB2 sub-basins, the four models, and the two study periods separately.

	KGE ($Q^{0.5}$) for AB1		KGE ($Q^{0.5}$) for AB2	
	2021 period	2022 period	2021 period	2022 period
RU-Reservoir	-1.04	0.22	-0.82	0.29
RU-SIDRA	-1.04	-0.77	-1.19	-0.11
SAMIR-Reservoir	-0.11	-0.68	0.28	-0.21
SAMIR-SIDRA	-1.13	-1.87	-0.32	-0.48

570

3.2.1 Accuracy of RU-Reservoir and RU-SIDRA

Figures 6.a, 6.c, 7.a, and 7.c show that the RU-based models do not simulate any discharge for the 2021 period with the average default parameters (red lines). This is related to the fact that, in a context where RU is not spatialized while irrigation is spatially heterogeneous, the optimal S_{inter} value to generate sufficient recharge is lower than in the literature (optimal S_{inter} values are shown in Table 5). The S_{inter} values taken from the literature for the accuracy evaluation are then too high to simulate enough recharge. This implies that without calibration with drainage data, models based on RU are only effective in contexts where irrigation practices are homogeneous, e.g., on sites under monoculture. Note that for the 2022 period, the RU-based models simulate more drainage than in 2021 because the irrigation amounts applied by farmers in 2022 (587 mm between May and October 2022 in average over the AB district) are significantly larger than in 2021 (509 mm between May and October 2021).

3.2.2 Accuracy of SAMIR-SIDRA

Figures 6 and 7 show that with average default parameter values, the SAMIR-based models simulate the drainage dynamics with some consistency with the irrigation season and the rain events for both the 2021 and 2022 periods. However, SAMIR-SIDRA shows lower performance than SAMIR-Reservoir. Several factors may explain this result. First, although they are physical parameters, the values of the SIDRA parameters K and μ found in the literature do not necessarily correspond to their optimal values for a given site. Second, SIDRA is more appropriate for more responsive hydrosystems than AB1 and AB2. Finally, the lack of information on the half of the drain spacing L on the AB district led us to set it at 6 m, whereas when calibrated on AB1 and AB2, this parameter is on average 37 m (average value considering the plots that are not equipped with field drains). This lower L value results in a high reactivity of the simulated drainage.

3.2.3 Accuracy of SAMIR-Reservoir

Figures 6.e, 6.f, 7.e, and 7.f. show that the drainage simulated by SAMIR-Reservoir with average default parameter values is more consistent with the observed drainage than the other three models. It also presents less variability within the 2000 simulations ensemble. Fig.7.e shows that SAMIR-Reservoir simulates particularly well the drainage dynamics and quantities for the AB2 sub-basin during the 2021 period. SAMIR-Reservoir tends to overestimate the drainage when the discharge is lower for AB1 during the 2021 and 2022 periods and AB2 during

the 2022 periods. The fact that SAMIR-Reservoir shows higher accuracy than the more complex models using SIDRA is an interesting result since it shows that the descriptive complexity of the models may not be useful for predictive purposes. One reason for this is the difficulty of linking SIDRA soil parameters to physically measurable soil properties.

600

The SAMIR-Reservoir accuracy varies spatially between AB1 and AB2 and temporally between 2021 and 2022. These differences reflect the semi-empirical nature of SAMIR and Reservoir models. Indeed, they can be attributed to the impact of unrepresented processes (e.g., lateral flows) or misrepresentation of *ET* and recharge processes in unusual situations (e.g., 2022 drought and heat waves). SAMIR especially fully neglects the lateral flow. Such

605 subsurface flows may come from outside sub-basin boundaries and contribute significantly to the sub-basin discharge measured at the outlet, depending on the hydrometeorological conditions encountered in a given year. Note that representing subsurface lateral flows would be challenging, especially to estimate them accurately across system boundaries. Furthermore, this would require more complex models than those tested in this study and additional data (e.g., piezometric), which are currently not available in the study area.

610 **3.2.4 Recommendations and perspectives**

Due to the semi-empirical nature of the four models investigated in this study, it is difficult to reproduce the drainage discharge with default parameters from the literature. A limitation of RU-based models is their lack of spatialization, leading to the use of an average irrigation in forcing, while the irrigation of AB1 and AB2 sub-basins is spatially heterogeneous. This results in RU not having enough irrigation to generate a correct recharge

615 with the S_{inter} values suggested in the literature (being too high). One way to overcome this would be to spatialize RU with irrigation data at each plot, but these data are rarely available.

SAMIR-Reservoir offers a certain consistency with the observed data and could provide an approximate idea of the drainage dynamics and amount occurring in an ungauged irrigated sub-catchment. Figures 6.e, 6.f, 7.e, and 7.f show that SAMIR-Reservoir 2000 simulations are less dispersed than for the other three models, suggesting greater

620 robustness. For decision support, we therefore recommend the use of SAMIR-Reservoir. However, its accuracy should be further evaluated on other sites. Furthermore, its accuracy could be improved i) by modeling lacking physical processes and ii) by investigating a link between the depletion coefficient parameter ω and the soil texture or the characteristics of the drainage network.

625 **4 Summary and conclusion**

Estimating the drainage in semi-arid irrigation conditions is essential to prevent soil salinization issues and assess the water productivity and the irrigation impact on downstream ecosystems. A few studies have used physically-based models to simulate drainage in an irrigated context. However, these models have many parameters requiring extensive data that are rarely available. In this paper, we thus assessed the capacity of four parsimonious semi-

630 empirical models to simulate drainage at the scale of two sub-basins of the AB district in northeastern Spain. The four models are built from the combination of two surface models (RU and SAMIR) and two subsurface models (Reservoir and SIDRA) with varying complexity levels: RU-Reservoir (two main parameters), RU-SIDRA (three

main parameters), SAMIR-Reservoir (three main parameters) and SAMIR-SIDRA (four main parameters). SAMIR is based on the FAO-56 *ET* and crop water balance formulations relying on two main sensitive parameters, while RU is a simplified version of the FAO-56 relying on a single sensitive parameter only. SIDRA solves the Boussinesq equation from two main sensitive parameters, while Reservoir is an empirical drainage model based only on a single depletion parameter.

The precision of the four models, i.e., their ability to reproduce observed drainage data with a site-specific calibration, was first evaluated. An optimal calibration approach was implemented for each model and each sub-basin using the multi-objective genetic algorithm NSGA-II. The comparison between the drainage simulated by site-specific calibrated models and observations indicates that RU-Reservoir presents a better precision, followed closely by SAMIR-Reservoir. This is explained by the fact that the Reservoir model is well suited to represent the low responsiveness of both studied sub-basins and that the RU model manages to artificially better spread out the recharge events during the irrigation period than the SAMIR model. In addition, the calibrated parameter values vary between the two periods analyzed for a given sub-basin. This indicates that the models have limited predictive capacities (robustness). It is therefore necessary to identify the processes poorly simulated or not simulated at all, such as lateral subsurface flows, root growth, and evapotranspiration, and to better take them into account in the models.

Complex models, with many processes simulated and parameters to calibrate, are likely to simulate drainage more precisely than simple models when calibrated on observed data. However, when no data is available for calibration, the most complex models are also the most prone to uncertainty (Puy et al., 2022). Moreover, drainage observations required for site-specific calibration are rarely available. Therefore, the accuracy of the four models was also evaluated. By accuracy, we mean the ability of the four models to reproduce the observed drainage using default parameter values provided by the literature. The comparison between the drainage simulated by default-calibrated models and observations indicates that SAMIR-Reservoir is the only model among the four tested capable of giving a rough estimate of the drainage dynamics and amounts from default parameters.

However, it was found that the accuracy of SAMIR-Reservoir is quite variable from one sub-basin to another and from one hydrological year to another. Therefore, calibration strategies are still needed to reduce uncertainties in SAMIR-Reservoir drainage estimates in sub-basins with contrasted conditions. In addition, better constraining the value of the Reservoir's depletion coefficient *k*, especially by seeking a link with soil texture or with the characteristics of the drainage network, should be investigated in future studies to improve the accuracy of SAMIR-Reservoir.

Furthermore, our study took place in an irrigation district where the water use is known and accurately monitored through the pumping data. Such irrigation data are rarely available in practice, and no model can predict drainage accurately based on inaccurate irrigation forcing, regardless of the model calibration issue. Hence, it is crucial to develop tools to retrieve the irrigation practices, notably at the integrated spatial scales of sub-basin or irrigation districts. To this end, many recent works seek to assimilate satellite products of soil moisture or *ET* in crop water balance models at a range of scales (e.g., Ouaadi et al., 2021; Massari et al., 2021; Dari et al., 2022; Olivera-Guerra et al., 2020). The coupling of such remote sensing approaches with surface and subsurface models is likely to improve the predictive capabilities of drainage in irrigated areas.

Code availability. All Python codes will be provided by the corresponding author upon request. The Python package spotpy (Houska et al., 2015) used for the NSGA-II multi-objective optimization is available at https://github.com/thouska/spotpy.

Data availability. Drainage data are not publicly available. Daily precipitation and ET_0 data were provided by the Meteorological Service of Catalonia and are available at https://ruralcat.gencat.cat/agrometeo.estacions (accessed 2 March 2023). The 2 m resolution DEM was provided by the Institute of Cartography of Catalonia and is available at https://www.icgc.cat/ca/Descarregues/Elevacions/Model-d-elevacions-del-terreny-de-2x2-m (accessed 2 March 2023). Crop type information was provided by the Department of Climate Action, Food, and Rural Agenda of the Region of Catalonia and is available at https://agricultura.gencat.cat/ca/ambits/desenvolupament-rural/sigpac/descarregues (accessed 2 March 2023). SoilGrid products are available at https://maps.isric.org (accessed 2 March 2023).

Author contributions. PL and OM conceptualized the work. PL, OM, LOG, and PQS provided the methodological guidelines. VA and JMV collected data. PL, VR, and LOG developed the code. PL drafted the paper. PL, OM, LOG, VR, VA, JMV, AJ, JT, PQS, ABO, and OCC all revised the paper and contributed to its analyses and discussions.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgment. The authors would like to thank the Comunitat de Regants Canal Algerri Balaguer and the Automatic Hydrological Information System of the Ebro Basin for providing the irrigation observation data used in this study. We also want to thank Eric Chavanon (CESBIO), who helped optimize the SAMIR code.

Financial support. This study was supported by the IDEWA project (ANR-19-P026-003) of the Partnership for Research and Innovation in the Mediterranean Area (PRIMA) program and by the Horizon 2020 ACCWA project (grant agreement # 823965) in the context of the Marie Sklodowska-Curie research and innovation staff exchange (RISE) program.

References

Abdi, D. E., Owen, J. S., Wilson, P. C., Hinz, F. O., Cregg, B., and Fernandez, R. T.: Reducing pesticide transport in surface and subsurface irrigation return flow in specialty crop production, Agricultural Water Management, 256, 107124, https://doi.org/10.1016/j.agwat.2021.107124, 2021.

Ale, S., Gowda, P. H., Mulla, D. J., Moriasi, D. N., and Youssef, M. A.: Comparison of the performances of DRAINMOD-NII and ADAPT models in simulating nitrate losses from subsurface drainage systems, Agricultural Water Management, 129, 21–30, https://doi.org/10.1016/j.agwat.2013.07.008, 2013.

Allen, R., Pereira, L., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 1998.

Altés, V., Bellvert, J., Pascual, M., and Villar, J. M.: Understanding Drainage Dynamics and Irrigation Management in a Semi-Arid Mediterranean Basin, Water, 15, https://doi.org/10.3390/w15010016, 2022.

Amazirh, A., Merlin, O., Er-Raki, S., Bouras, E., and Chehbouni, A.: Implementing a new texture-based soil evaporation reduction coefficient in the FAO dual crop coefficient method, Agricultural Water Management, 250, 106827, https://doi.org/10.1016/j.agwat.2021.106827, 2021.

Bekele, E. G. and Nicklow, J. W.: Multi-objective automatic calibration of SWAT using NSGA-II, Journal of Hydrology, 341, 165–176, https://doi.org/10.1016/j.jhydrol.2007.05.014, 2007.

Blann, K., Anderson, J., Sands, G., and Vondracek, B.: Effects of Agricultural Drainage on Aquatic Ecosystems: A Review, Critical Reviews in Environmental Science and Technology - CRIT REV ENVIRON SCI TECHNOL, 39, 909–1001, https://doi.org/10.1080/10643380801977966, 2009.

Bouarfa, S. and Zimmer, D.: Water-table shapes and drain flow rates in shallow drainage systems, Journal of Hydrology, 235, 264–275, https://doi.org/10.1016/S0022-1694(00)00280-8, 2000.

Boussinesq, J.: Recherches théoriques sur l'écoulement des nappes d'eau infiltrées dans le sol et sur le débit des sources, Journal de Mathématiques Pures et Appliquées, 10, 5–78, 1904.

Cavero, J., Barros, R., Sellam, F., Topcu, S., Isidoro, D., Hartani, T., Lounis, A., Ibrikci, H., Cetin, M., Williams, J. R., and Aragüés, R.: APEX simulation of best irrigation and N management strategies for off-site N pollution control in three Mediterranean irrigated watersheds, Agricultural Water Management, 103, 88–99, https://doi.org/10.1016/j.agwat.2011.10.021, 2012.

Cenobio-Cruz, O., Quintana-Seguí, P., Barella-Ortiz, A., Zabaleta, A., Garrote, L., Clavera-Gispert, R., Habets, F., and Beguería, S.: Improvement of low flows simulation in the SASER hydrological modeling chain, Journal of Hydrology X, 18, 100147, https://doi.org/10.1016/j.hydroa.2022.100147, 2023.

Chang, X., Wang, S., Gao, Z., Chen, H., and Guan, X.: Simulation of Water and Salt Dynamics under Different Water-Saving Degrees Using the SAHYSMOD Model, Water, 13, 1939, https://doi.org/10.3390/w13141939, 2021.

Chelil, S., Oubanas, H., Henine, H., Gejadze, I., Malaterre, P. O., and Tournebize, J.: Variational data assimilation to improve subsurface drainage model parameters, Journal of Hydrology, 610, 128006, https://doi.org/10.1016/j.jhydrol.2022.128006, 2022.

van Dam, J. C., Groenendijk, P., Hendriks, R. F. A., and Kroes, J. G.: Advances of Modeling Water Flow in Variably Saturated Soils with SWAP, Vadose Zone Journal, 7, 640–653, https://doi.org/10.2136/vzj2007.0060, 2008.

Dari, J., Brocca, L., Modanesi, S., Massari, C., Tarpanelli, A., Barbetta, S., Quast, R., Vreugdenhil, M., Freeman, V., Barella-Ortiz, A., Quintana-Seguí, P., Bretreger, D., and Volden, E.: Regional data sets of high-resolution (1 and 6 km) irrigation estimates from space, Earth System Science Data Discussions, 1–34, https://doi.org/10.5194/essd-2022-403, 2022.

David, C. H., Maidment, D. R., Niu, G.-Y., Yang, Z.-L., Habets, F., and Eijkhout, V.: River Network Routing on the NHDPlus Dataset, Journal of Hydrometeorology, 12, 913–934, https://doi.org/10.1175/2011JHM1345.1, 2011.

De Schepper, G., Therrien, R., Refsgaard, J. C., and Hansen, A. L.: Simulating coupled surface and subsurface water flow in a tile-drained agricultural catchment, Journal of Hydrology, 521, 374–388, https://doi.org/10.1016/j.jhydrol.2014.12.035, 2015.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation, 6, 182–197, https://doi.org/10.1109/4235.996017, 2002.

FAO: The state of the world's land and water resources for food and agriculture, available at: https://www.fao.org/documents/card/fr/c/cb7654en (last access: 20 April 2023), 2021

Feng, G., Zhu, C., Wu, Q., Wang, C., Zhang, Z., Mwiya, R. M., and Zhang, L.: Evaluating the impacts of saline water irrigation on soil water-salt and summer maize yield in subsurface drainage condition using coupled HYDRUS and EPIC model, Agricultural Water Management, 258, 107175, https://doi.org/10.1016/j.agwat.2021.107175, 2021.

Fritsch, F. N. and Carlson, R. E.: Monotone Piecewise Cubic Interpolation, SIAM J. Numer. Anal., 17, 238–246, https://doi.org/10.1137/0717021, 1980.

García-Garizábal, I. and Causapé, J.: Influence of irrigation water management on the quantity and quality of irrigation return flows, Journal of Hydrology, 385, 36–43, https://doi.org/10.1016/j.jhydrol.2010.02.002, 2010.

Gassman, P., Williams, J., Wang, X. (Susan), Saleh, A., Osei, E., Hauck, L., Izaurralde, R., and Flowers, J.: Agricultural Policy Environmental EXtender (APEX) Model: An Emerging Tool for Landscape and Watershed Environmental Analyses, The, Center for Agricultural and Rural Development (CARD) at Iowa State University, Center for Agricultural and Rural Development (CARD) Publications, 53, 2009.

Golmohammadi, G., Rudra, R. P., Parkin, G. W., Kulasekera, P. B., Macrae, M., and Goel, P. K.: Assessment of Impacts of Climate Change on Tile Discharge and Nitrogen Yield Using the DRAINMOD Model, Hydrology, 8, 1, https://doi.org/10.3390/hydrology8010001, 2020.

Gowda, P., Mulla, D., Desmond, E., Ward, A., and Moriasi, D.: ADAPT: Model use, calibration and validation, Transactions of the ASABE (American Society of Agricultural and Biological Engineers), 55, 1345–1352, https://doi.org/10.13031/2013.42246, 2012.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hengl, T., Jesus, J. M. de, Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil

information based on machine learning, PLOS ONE, 12, e0169748, https://doi.org/10.1371/journal.pone.0169748, 2017.

Henine, H., Jeantet, A., Chaumont, C., Chelil, S., Lauvernet, C., and Tournebize, J.: Coupling of a subsurface drainage model with a soil reservoir model to simulate drainage discharge and drain flow start, Agricultural Water Management, 262, 107318, https://doi.org/10.1016/j.agwat.2021.107318, 2022.

Houska, T., Kraft, P., Chamorro-Chavez, A., and Breuer, L.: SPOTting Model Parameters Using a Ready-Made Python Package, PLOS ONE, 10, e0145180, https://doi.org/10.1371/journal.pone.0145180, 2015.

Jahn, R., Blume, H. P., Asio, V., Spaargaren, O., and Schád, P.: FAO Guidelines for Soil Description, 2006.

Jarvis, N. and Larsbo, M.: MACRO (v5.2): Model Use, Calibration, and Validation, Transactions of the ASABE, 55, 1413–1423, https://doi.org/10.13031/2013.42251, 2012.

Jeantet, A., Henine, H., Chaumont, C., Collet, L., Thirel, G., and Tournebize, J.: Robustness of a parsimonious subsurface drainage model at the French national scale, Hydrology and Earth System Sciences, 25, 5447–5471, https://doi.org/10.5194/hess-25-5447-2021, 2021.

Jeantet, A., Thirel, G., Jeliazkov, A., Martin, P., and Tournebize, J.: Effects of Climate Change on Hydrological Indicators of Subsurface Drainage for a Representative French Drainage Site, Frontiers in Environmental Science, 10, 899226, https://doi.org/10.3389/fenvs.2022.899226, 2022.

Jiang, Q., Qi, Z., Lu, C., Tan, C. S., Zhang, T., and Prasher, S. O.: Evaluating RZ-SHAW model for simulating surface runoff and subsurface tile drainage under regular and controlled drainage with subirrigation in southern Ontario, Agricultural Water Management, 237, 106179, https://doi.org/10.1016/j.agwat.2020.106179, 2020.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrology and Earth System Sciences, 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Laluet, P., Olivera-Guerra, L., Rivalland, V., Simonneaux, V., Inglada, J., Bellvert, J., Er-raki, S., and Merlin, O.: A sensitivity analysis of a FAO-56 dual crop coefficient-based model under various field conditions, Environmental Modelling & Software, 160, 105608, https://doi.org/10.1016/j.envsoft.2022.105608, 2023.

Larsbo, M., Roulier, S., Stenemo, F., Kasteel, R., and Jarvis, N.: An Improved Dual-Permeability Model of Water Flow and Solute Transport in the Vadose Zone, Vadose Zone Journal, 4, 398–406, https://doi.org/10.2136/vzj2004.0137, 2005.

Lehmann, P., Merlin, O., Gentine, P., and Or, D.: Soil Texture Effects on Surface Resistance to Bare-Soil Evaporation, Geophysical Research Letters, 45, 10,398-10,405, https://doi.org/10.1029/2018GL078803, 2018.

Lesaffre, B. and Zimmer, D.: Field evaluation of a subsurface drainage simulation model predicting peak flows, in: 3ème séminaire national de l'ASAE drainage design and management, Chicago, USA, 1987, 8, 1987.

Ma, L., Ahuja, L., Nolan, B. T., Malone, R., Trout, T., and Qi, Z.: Root zone water quality model (RZWQM2): Model use, calibration and validation, Transactions of the ASABE, 55, 1425–1446, https://doi.org/10.13031/2013.42252, 2012.

Massari, C., Modanesi, S., Dari, J., Gruber, A., De Lannoy, G. J. M., Girotto, M., Quintana-Seguí, P., Le Page, M., Jarlan, L., Zribi, M., Ouaadi, N., Vreugdenhil, M., Zappa, L., Dorigo, W., Wagner, W., Brombacher, J., Pelgrum, H., Jaquot, P., Freeman, V., Volden, E., Fernandez Prieto, D., Tarpanelli, A., Barbetta, S., and Brocca, L.: A Review of Irrigation Information Retrievals from Space and Their Utility for Users, Remote Sensing, 13, 4112, https://doi.org/10.3390/rs13204112, 2021.

Merlin, O., Stefan, V. G., Amazirh, A., Chanzy, A., Ceschia, E., Er-Raki, S., Gentine, P., Tallec, T., Ezzahar, J., Bircher, S., Beringer, J., and Khabba, S.: Modeling soil evaporation efficiency in a range of soil and atmospheric conditions using a meta-analysis approach, Water Resources Research, 52, 3663–3684, https://doi.org/10.1002/2015WR018233, 2016.

Moursi, H., Youssef, M. A., and Chescheir, G. M.: Development and application of DRAINMOD model for simulating crop yield and water conservation benefits of drainage water recycling, Agricultural Water Management, 266, 107592, https://doi.org/10.1016/j.agwat.2022.107592, 2022.

Muma, M., Rousseau, A. N., and Gumiere, S. J.: Modeling of subsurface agricultural drainage using two hydrological models with different conceptual approaches as well as dimensions and spatial scales, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 42, 38–53, https://doi.org/10.1080/07011784.2016.1231014, 2017.

Negm, L. M., Youssef, M. A., and Jaynes, D. B.: Evaluation of DRAINMOD-DSSAT simulated effects of controlled drainage on crop yield, water balance, and water quality for a corn-soybean cropping system in central Iowa, Agricultural Water Management, 187, 57–68, https://doi.org/10.1016/j.agwat.2017.03.010, 2017.

Nousiainen, R., Warsta, L., Turunen, M., Huitu, H., Koivusalo, H., and Pesonen, L.: Analyzing subsurface drain network performance in an agricultural monitoring site with a three-dimensional hydrological model, Journal of Hydrology, 529, 82–93, https://doi.org/10.1016/j.jhydrol.2015.07.018, 2015.

Olivera-Guerra, L., Merlin, O., and Er-Raki, S.: Irrigation retrieval from Landsat optical/thermal data integrated into a crop water balance model: A case study over winter wheat fields in a semi-arid region, Remote Sensing of Environment, 239, https://doi.org/10.1016/j.rse.2019.111627, 2020.

Olivera-Guerra, L.-E., Laluet, P., Altés, V., Ollivier, C., Pageot, Y., Paolini, G., Chavanon, E., Rivalland, V., Boulet, G., Villar, J.-M., and Merlin, O.: Modeling actual water use under different irrigation regimes at district scale: Application to the FAO-56 dual crop coefficient method, Agricultural Water Management, 278, 108119, https://doi.org/10.1016/j.agwat.2022.108119, 2023.

Oosterbaan, R. J.: SAHYSMOD (version 1.7 a), Description of principles, user manual and case studies, International Institute for Land Reclamation and Improvement, Wageningen, Netherlands, 140, 2005.

Ouaadi, N., Jarlan, L., Khabba, S., Ezzahar, J., Le Page, M., and Merlin, O.: Irrigation Amounts and Timing Retrieval through Data Assimilation of Surface Soil Moisture into the FAO-56 Approach in the South Mediterranean Region, Remote Sensing, 13, 2667, https://doi.org/10.3390/rs13142667, 2021.

Pereira, L. S., Paredes, P., Hunsaker, D. J., López-Urrea, R., and Jovanovic, N.: Updates and advances to the FAO-56 crop water requirements method, Agricultural Water Management, 248, 106697, https://doi.org/10.1016/j.agwat.2020.106697, 2021.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, SOIL, 7, 217–240, https://doi.org/10.5194/soil-7-217-2021, 2021.

Puy, A., Beneventano, P., Levin, S. A., Lo Piano, S., Portaluri, T., and Saltelli, A.: Models with higher effective dimensions tend to produce more uncertain estimates, Science Advances, 8, eabn9450, https://doi.org/10.1126/sciadv.abn9450, 2022.

Quintana-Seguí, P., Turco, M., Herrera, S., and Miguez-Macho, G.: Validation of a new SAFRAN-based gridded precipitation product for Spain and comparisons to Spain02 and ERA-Interim, Hydrology and Earth System Sciences, 21, 2187–2201, https://doi.org/10.5194/hess-21-2187-2017, 2017.

Román Dobarco, M., Cousin, I., Le Bas, C., and Martin, M. P.: Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty, Geoderma, 336, 81–95, https://doi.org/10.1016/j.geoderma.2018.08.022, 2019.

Santos, L., Thirel, G., and Perrin, C.: Technical note: Pitfalls in using log-transformed flows within the KGE criterion, Hydrology and Earth System Sciences, 22, 4583–4591, https://doi.org/10.5194/hess-22-4583-2018, 2018.

Schultz, B., Zimmer, D., and Vlotman, W. F.: Drainage under increasing and changing requirements, Irrigation and Drainage, 56, S3–S22, https://doi.org/10.1002/ird.372, 2007.

Shafii, M. and De Smedt, F.: Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm, Hydrology and Earth System Sciences, 13, 2137–2149, https://doi.org/10.5194/hess-13-2137-2009, 2009.

Simonneaux, V., Lepage, M., Helson, D., Metral, J., Thomas, S., Duchemin, B., Cherkaoui, M., Kharrou, H., Berjami, B., and Chehbouni, A.: Estimation spatialisée de l'évapotranspiration des cultures irriguées par télédétection : application à la gestion de l'irrigation dans la plaine du Haouz (Marrakech, Maroc), Science et changements planétaires / Sécheresse, 20, 123–130, https://doi.org/10.1684/sec.2009.0177, 2009.

Singh, A.: Environmental problems of salinization and poor drainage in irrigated areas: Management through the mathematical models, Journal of Cleaner Production, 206, 572–579, https://doi.org/10.1016/j.jclepro.2018.09.211, 2019.

Skaggs, R. W., Youssef, M., and Chescheir, G. M.: DRAINMOD: model use, calibration, and validation, Transactions of the ASABE, 55, 1509–1522, https://doi.org/10.13031/2013.42259, 2012.

Tournebize, J., Kao, C., Nikolic, N., and Zimmer, D.: Adaptation of the STICS model to subsurface drained soils, Agronomie, 24, 305–313, https://doi.org/10.1051/agro:2004030, 2004.

875 Turunen, M., Warsta, L., Paasonen-Kivekäs, M., Nurminen, J., Myllys, M., Alakukku, L., Äijö, H., Puustinen, M., and Koivusalo, H.: Modeling water balance and effects of different subsurface drainage methods on water outflow components in a clayey agricultural field in boreal conditions, Agricultural Water Management, 121, 135–148, https://doi.org/10.1016/j.agwat.2013.01.012, 2013.

Vergnes, J.-P. and Habets, F.: Impact of river water levels on the simulation of stream–aquifer exchanges over the
880 Upper Rhine alluvial aquifer (France/Germany), Hydrogeol J, 26, 2443–2457, https://doi.org/10.1007/s10040-018-1788-0, 2018.

Wen, Y., Shang, S., Rahman, K. U., Xia, Y., and Ren, D.: A semi-distributed drainage model for monthly drainage water and salinity simulation in a large irrigation district in arid region, Agricultural Water Management, 230, 105962, https://doi.org/10.1016/j.agwat.2019.105962, 2020.

885 Xian, C., Qi, Z., Tan, C. S., and Zhang, T.-Q.: Modeling hourly subsurface drainage using steady-state and transient methods, Journal of Hydrology, 550, 516–526, https://doi.org/10.1016/j.jhydrol.2017.05.016, 2017.

Zimmer, D., Tournebize, J., Bouarfa, S., Kao, C., and Lesaffre, B.: Land drainage functioning and hydrological impacts in rural catchments: model development and field experiments, Comptes Rendus. Géoscience, 355, 1–19, https://doi.org/10.5802/crgeos.194, 2023.