Supplement to:

A large transient multi-scenario multi-model ensemble of future streamflow and groundwater projections in France

Eric Sauquet[1], Guillaume Evin[2], Sonia Siauve[3], Ryma Aissat[4], Patrick Arnaud[5], Maud Bérel[6], Jérémie Bonneau[1,7], Flora Branger[1], Yvan Caballero[8], François Colléoni[5], Agnès Ducharne[9], Joël Gailhard[10], Florence Habets[11], Frédéric Hendrickx[12], Louis Héraut[1], Benoît Hingray[2], Peng Huang[9], Tristan Jaouen[1], Alexis Jeantet[13], Sandra Lanini[8], Matthieu Le Lay[10], Claire Magand[14], Louise Mimeau[1], Céline Monteil[12], Simon Munier[13], Charles Perrin[15], Olivier Robelin[1,8], Fabienne Rousset[17], Jean-Michel Soubeyroux[17], Laurent Strohmenger[15], Guillaume Thirel[15,16], Flore Tocquer[17], Yves Tramblay[18], Jean-Pierre Vergnes[4], Jean-Philippe Vidal[1]

[1]UR RiverLy, INRAE, Villeurbanne, France
[2]Univ. Grenoble Alpes, INRAE, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]OiEau, 87100 Limoges, France
[4]BRGM – French Geological Survey, Orléans, France
[5]UMR RECOVER, INRAE, Aix-Marseille University, Le Tholonet, France
[6]MTEECPR, La Défense, France
[7]INSA Lyon, DEEP, UR 7429, Villeurbanne, France
[8]UMR 183 G-Eau, INRAE, CIRAD, IRD, AgroParisTech, Institut Agro, BRGM, Montpellier
[9]Sorbonne Université/CNRS/EPHE, METIS-IPSL, Paris, France
[10]Département Eau Environnement, EDF-DTG, Saint Martin le Vinoux, France
[11]Geology Laboratory of Ecole Normale Supérieure, Pierre Simon Laplace Research University, CNRS UMR 8538, Paris, France
[12]Département LNHE, EDF-R&D, 78401 Chatou, France
[13]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[14]OFB, Direction de la recherche et de l'appui scientifique, Nantes, France
[15]Université Paris-Saclay, INRAE, UR HYCAR, Antony, France
[16]Univ Toulouse, CNES/IRD/CNRS/INRAE, CESBIO, Toulouse, France
[17]Météo-France, Direction de la Climatologie et des Services Climatiques, Toulouse, France
[18]UMR Espace Dev (Univ. Montpellier, IRD), Montpellier, France

Correspondence to: Eric Sauquet (eric.sauquet@inrae.fr)

Abstract. A large transient multi-scenario and multi-model ensemble of future streamflow and groundwater projections in France developed in a national project named Explore2 was recently made available. The main objective of Explore2 is to provide rich and spatially-consistent information for the future evolution of hydrological (surface and groundwater) resources and extremes in France to support adaptation strategies. The Explore2 dataset was obtained using a nested multi-scenario multi-model approach to estimate future uncertainty and to assess local climate at the catchment scale: three greenhouse gas

(GHG) emission scenarios, a set of 17 combinations of Global Climate Models and Regional Climate Models (GCM-RCM), and two bias correction methods provide the meteorological forcing for nine surface hydrology models and four groundwater hydrology models (one to simulate groundwater recharge and three to simulate groundwater level). In this paper, we present the methodology underlying the dataset, the evaluation of the hydrological models against daily streamflow and groundwater level observations, the assessment of the future streamflow and recharge projections, the data availability and the ways of accessing the data and understanding the results (mainly through visualisation tools).

This large set of hydrological projections shows a high model agreement on the decrease in seasonal flows in the South of France under the RCP8.5 high-emission scenario, confirming its hotspot status. The surface HMs agree on the decrease in summer flows across France under the RCP8.5 scenario, with the exception of northern part France. This area may indeed benefit from more active winter recharge that may counterbalance decrease in summer precipitation and increase in evapotranspiration. In the mountainous areas, winter flows will increase as a result of higher air temperature and the high degree of agreement between the models holds regardless of the RCP considered. Unsurprisingly, the higher the GHG emission scenario, the higher the median changes. Most of these changes are organised in France along a north-south gradient, regardless of the RCP considered.

S1. Description of hydrological models

S1.1 AquiFR

The AquiFR groundwater modelling platform covers around one-third of the French mainland area and half of the French aquifers. AquiFR aims at monitoring and providing seasonal forecasts of the groundwater as well as to project its long-term evolution in order to support decision-making at regional scale. It consists in a modelling chain gathering meteorological inputs (SAFRAN, Vidal *et al*., 2010), a land surface model (SURFEX, Le Moigne *et al*., 2020) and regional groundwater models (Vergnes *et al*., 2020). Within the modelling chain, the SURFEX land surface model simulates groundwater recharge and surface runoff. Then, hydrogeological models use these variables as inputs (Vergnes *et al*., 2020). SURFEX uses atmospheric variables provided by SAFRAN to solve the energy and water budgets at the land–atmosphere interface. Within the platform, MARTHE (Thiéry, 2015) and EAUDYSSÉE (Saleh *et al*., 2013) simulate unsaturated zone transfer, groundwater flows, and groundwater-rivers exchanges over five and seven regions, respectively. Each model is spatially distributed and solves a 2D diffusivity equation (Darcy law in combination with the conservation mass principle) in multi-layer aquifers. Groundwater withdrawals are included in the model based on incomplete reported values (Vergnes *et al*., 2020). The finest spatial resolution of each regional model varies from 100 m to 1 km. In Vergnes *et al*. (2020), if necessary, new calibrations were eventually performed using the SURFEX groundwater recharge and surface runoff. Indeed, before being implemented in AquiFR, these models were previously calibrated independently based on stakeholder requests and used their own soil water balance estimates that differ from the ones provided by SURFEX (Vergnes *et al.*, 2020). Groundwater parameters (aquifers thickness, specific yield, hydraulic conductivity and their spatial variability) were first defined based on available local data (geological maps, geological logs, and pumping tests). Calibration relied on groundwater levels and on river flows. The AquiFR platform was assessed by comparison of the simulation with observed piezometric levels and river flows in Vergnes *et al.* (2020).

S1.2 CTRIP

The CTRIP model is a physically-based river routing model coupled to the ISBA land surface model. While ISBA represents the vertical exchanges of water and energy at the soil-atmosphere interface, CTRIP simulates river flows over an entire hydrographic network (Decharme *et al*., 2019). CTRIP is based on a regular grid with a resolution of 1/12° (i.e. around 6-8 km over France) obtained by the upscaling of the MERIT-Hydro global hydrographic network (Yamazaki *et al*., 2019), available at a 90-m resolution and currently considered to be the most accurate on a global scale. A number of hydro-geomorphological parameters, such as the lengths and slopes of river sections, are obtained from high-resolution data from MERIT-Hydro, while other parameters, such as widths, depths and roughness, are obtained from empirical formulae (Munier and Decharme, 2022). It is assumed that each grid cell contains one and only one river reach, represented as a reservoir flowing into the downstream grid cell. The Manning equation is used to calculate the flow velocity as a function of the volume of water in the section, itself updated by the inflows from the upstream reaches and the runoff from the ISBA model (in the same configuration as in the SIM2 model). The CTRIP model also benefits from a two-dimensional representation of aquifer dynamics and groundwater-river exchanges (Vergnes and Decharme, 2012). Finally, surface processes linked to vegetation (including actual evapotranspiration) and snow cover (including sublimation and melting) are taken into account in the ISBA model.

CTRIP does not benefit from a parameter calibration stage. This choice ensures spatial consistency when the model is used in other regions of the world – or even globally – where few observations are available.

S1.3 EROS

EROS is a distributed reservoir modelling framework dedicated to large river systems (Thiéry, 2018; Thiéry and Moutzopoulos, 1992). It allows the simulation of river flow or karstic spring flow and piezometric-head measurements in heterogeneous river basins. These river basins are represented in EROS as a cluster of elementary lumped-parameter hydrological models connected with each other. For each sub-model, a hydroclimatic lumped model computes the streamflow at the outlet of the sub-model and the piezometric head in the underlying water table. Each sub-model simulates the main mechanisms of the water cycle through simplified physical laws. Snow accumulation, snow melting and pumping are taken into account. The total river flow at the outlet of each sub-basin is computed from the upstream tree of sub-basins. EROS was initially developed to simulate regional watersheds avoiding the complexity of a spatially- and physically-based model.

Two versions of the EROS model were used in Explore2. The first one simulates daily time series of streamflowss at the outlet of 96 catchments located in Brittany as well as daily time series of groundwater levels at 41 piezometers. The second one subdivides the Loire River basin into 368 catchments for which the model simulates daily time series of streamflows

(Seyedhashemi *et al*., 2022). Calibration was done against observed streamflows and groundwater levels for each simulated time series using a semi-automatic method based on a modified version of the Rosenbrock algorithm (Rosenbrock, 1960).

S1.4 GRSD

The GRSD model is a semi-distributed rainfall-runoff model developed at INRAE (de Lavenne *et al*. 2019), based on GR4J (Perrin *et al*., 2003), and CemaNeige, a snow accumulation and melt model (Valéry *et al*., 2014). The GRSD model is used for flood simulations or forecasts (Royer-Gaspard *et al*., 2024), climate change impact assessment (Thirel *et al*., 2019), or the assessment of climate change water use adaptation strategies (Lemaitre-Basset *et al*., 2024). The GRSD model calibration and simulation were applied within the airGR and airGRiwrm open-source R packages (Coron *et al*., 2017, 2021; Dorchies *et al*., 2024). The core of the GRSD model, GR4J, is a lumped process-based model using a production store (for partitioning rainfall into actual evapotranspiration and net rainfall), and a routing store and a unit hydrograph to simulate the hydrological transfer to the basin outlet. An additional component represents the inter-catchment subsurface fluxes. The streamflows simulated by GR4J are routed downstream thanks to a lag function. The GRSD model uses as inputs rainfall and potential evapotranspiration aggregated at the sub-basin scale. The CemaNeige model, using air temperature and solid precipitation, simulates snow accumulation and melt thanks to a two-parameter degree-day approach applied on five zones of equal area in each sub-basin. The snow melt is added to the rainfall that feeds GRSD. Outputs from the model are daily streamflow at simulation points but also a range of hydrological variables for the related sub-catchments (daily actual evapotranspiration, snow water equivalent, soil moisture, etc.) calculated at the sub-catchment scale.

The GRSD model was calibrated thanks to the KGE (Gupta *et al*. 2009) criterion, applied with a square root streamflow transformation to equilibrate the weight put on high and low streamflows. A sequential approach (from upstream catchments to downstream catchments) was used.

S1.5 J2000

The J2000 model (Krause *et al*., 2006) is a process-oriented, distributed and modular hydrological model developed by the University of Jena and INRAE. It was used in many studies with a variety of applications for catchments ranging 10 – 10^6 km²: groundwater recharge (Watson *et al*., 2018), nutrient transport and erosion (Steudel *et al*., 2015), river flow intermittence (Mimeau *et al*., 2024), land use change and urban stormwater management (Branger *et al*., 2013), climate change impacts (Gao *et al*., 2012), etc. J2000 is open source and available through the JAMS modelling framework (Kralisch and Krause, 2006). The model was designed to simulate key hydrological processes at the catchment scale, on an irregular grid composed of Hydrological Response Units (HRUs) considered hydrologically homogeneous. The delineation of HRUs came from a meshing algorithm which is also open source (https://forgemia.inra.fr/michael.rabotin/hru-delin). This algorithm required a Digital Elevation Model, and catchment maps of land use, soils and geology in order to provide a calculated drainage network, and the connections between HRUs and river reaches. Within the Explore2 project, the model was run at a daily time step. It required precipitation, potential evapotranspiration and temperature as inputs. J2000 simulates critical processes such as interception of rainfall by vegetation, rainfall and snow partitioning, snow accumulation and melt, surface (hortonian) runoff, soil infiltration into two conceptual reservoirs, one drained by evapotranspiration and the other one drained by gravity through percolation to groundwater, baseflow from a conceptual deeper reservoir (see Horner (2020) for more details). The model also takes into account hydraulic routing within the river network. Outputs from the model are daily streamflows at relevant simulation points and river reaches but also a range of hydrological variables for relevant sub-catchments (daily actual evapotranspiration, snow water equivalent, soil water index, etc.) calculated at the sub-catchment scale.

Two applications of the J2000 model were used in the Explore2 project: one existing (Morel *et al*., 2022) for the Rhône catchment (31,426 HRUs) and one was set up for the Loire catchment (15,349 HRUs). For both catchments, calibration was minimal as most parameters can be related to physical features of the catchment. Calibration was performed manually and iteratively until an "acceptable" parametrization was found.

S1.6 MONA

The MOdel of North Aquitaine (MONA) is a regional pseudo-3D hydrodynamic model developed by the French Geological Survey (BRGM) to simulate groundwater flow and investigate the impact of pumping in the extensive unconfined aquifers supplying the city of Bordeaux (Aissat *et al*., 2023). The model covers the northern part (46,032 km²) of the French south-west sedimentary basin. It has 15 aquifers interbedded by aquitards discretized with a regular grid of 2 × 2 km from Plio-Quaternary down to Jurassic units. The model does not explicitly simulate flow in the aquitards but accounts for vertical flows adjusted by a conductance parameter (pseudo-3-D assumption). The domain is bounded by Cretaceous and Jurassic outcrops to the east and north, the Atlantic Ocean and the Gironde Estuary to the west. Hydraulic heads are imposed on the western boundary, accounting for seawater level along the Atlantic Coast. Heads are also prescribed along the Garonne River and its estuary, and no-flow boundaries are assumed at the southern limit, which corresponds with the separation from the southern part of the Aquitaine basin. Recharge is estimated annually by an empirical formula using climatic data (precipitation and reference evapotranspiration) from a series of weather stations (Pédron and Platel, 2005). The pumping database includes 6,235 wells distributed within the 15 geological formations (Saltel *et al*., 2016). The diffusivity equation is solved at the annual time step in a finite volume scheme with MARTHE (Thiery, 2015). The model calibration is based on observed groundwater levels, which the model must reproduce as closely as possible. A total of 430 groundwater level time series were used for manual calibration and evaluation (Saltel *et al.*, 2016).

S1.7 MORDOR-SD

The MORDOR-SD hydrological model (Garavaglia *et al*. 2017) is the operational rainfall-runoff model of Électricité De France (EDF) and is applied in different contexts (real-time forecasting, flood frequency analysis, continuous monitoring of water resources and climate change impact assessment) with a time step ranging from hourly to daily. Like many conceptual rainfall-runoff models, its structure is composed of different interconnected storages. MORDOR-SD is a lumped rainfall-runoff model, but it uses an elevation zone approach to spatialize the main meteorological forcings and hydrological processes.

Here, spatial variability of meteorological forcings is summarised with two orographic gradients, one for precipitation and one for temperature. Based on precipitation, air temperature and potential evapotranspiration data, six conceptual interconnected storage components evolve and supply the hydrographic network. These six storage components are: a superficial, an intermediate, a deep, an evaporating (capillary trapping), a snow and a glacier storage component. The transfer function is based on the diffusive wave equation (Hayami, 1951). The snow module is based on an enhanced degree-day approach applied on each elevation zone. The glacier module was used for seven catchments with a significant glacier area. The formulation of this module, which couples changes in glacier mass and area, enables the effects of glacial retreat observed and projected in the simulations to be taken into account.

The parameterization of the daily version applied in this study contains 8 to 12 free parameters, depending on whether or not the snow and a glacier storage component is applied.. Calibration was carried out over the maximum period of data availability, using a genetic algorithm optimizer described by Garavaglia *et al.* (2017) and based on a multi-criteria objective function incorporating several hydrological signatures (flows, interannual regimes, classified flows, low-water sequences). For lowland and mid-mountain basins, a snow-covered surface criterion was also used (via Fractional Snow Cover MODIS data).

S1.8 MORDOR-TS

The MORDOR-TS model (Rouhier *et al*., 2017) is a spatially-distributed version of the rainfall-runoff conceptual MORDOR-SD model (Garavaglia *et al.*, 2017). The catchment is divided into several hydrological meshes, based on topography, connected according to the hydrographic network. At each daily time step, the model calculates the water production of each mesh independently and routes all productions to the simulation points, which can be any mesh outlet. The production module aims at quantifying the exchanges between different components of the hydrologic cycle. Based on precipitation, air temperature and potential evapotranspiration data, six conceptual interconnected storage components evolve and supply the hydrographic network. The production module is applied to every hydrological mesh to estimate the runoff contributions. The routing module is designed to propagate the runoff contributions of all the meshes to the simulation points through the hydrographic network. It combines the intra-mesh and inter-mesh transfers by means of a formulation based on the 1D diffusive wave model, with celerity and diffusion independent of runoff (Hayami, 1951; Litrico and Georges, 1999). The parameterization of the version applied in this study contains 6 to 10 free parameters, depending on whether or not the snow module is applied.

An implementation of the model on the Loire catchment at Saumur is used for the Explore2 project (81,200 km², 1102 topographic meshes). The calibration methodology comes from Rouhier (2018). Depending on the parameters, there are different scenarios of spatialisation: some parameters are prescribed uniformly, others are derived from spatialized data, and others are calibrated according to an "upstream-downstream" pattern according to the available control flows. The calibration is carried out using the CaRamel multi-objective optimizer (Monteil *et al*., 2020) over the period 1987-2017. The multi-objective function is composed of three KGE criteria applied to: (1) daily streamflows, (2) daily interannual average, and (3) monthly classified streamflows.

S1.9 ORCHIDEE

The ORCHIDEE model is the land component of the IPSL (Institut Pierre Simon Laplace) climate model (Boucher *et al*., 2020). ORCHIDEE couples physically-based descriptions of water, energy, and carbon budgets, calculated on a 30-minute time step at atmospheric forcing resolution, here 8x8 km². Evapotranspiration does not depend on potential evapotranspiration, but is calculated as the sum of plant transpiration, soil evaporation, interception loss, and snow sublimation, all deduced from bulk aerodynamic formulations. These fluxes depend on vegetation properties (albedo, roughness, resistances), which vary spatially at a subgrid scale following high-resolution maps of vegetation types, and over time following vegetation phenology (Krinner *et al*., 2005). Soil evaporation and transpiration also depend on soil moisture.

Snowpack and its dynamics are described by a 3-layer model (Wang *et al*., 2013), and the resulting snowmelt combines with throughfall to feed soil water infiltration. The soil is 2-m deep, discretized in 22 layers to calculate vertical water fluxes from the Richards equation, assuming free drainage at the bottom, and infiltration-excess surface runoff. In each grid cell, the hydraulic parameters depend on soil texture, further modified for the hydraulic conductivity by the root profile, which also controls transpiration (Tafasca *et al*., 2020). Streamflow is calculated by aggregating surface runoff and drainage along the river network, based on a high-resolution topography (around 1.3 km over France). Each grid cell is decomposed into a graph of hydrological transfer units, each including three linear reservoirs corresponding to streams, hillslopes, and groundwater, with increasing transit times (Polcher *et al*., 2023). Because of interactions between the water, energy, and carbon budgets, the simulated evapotranspiration and streamflow depend on atmospheric CO_2 concentration.

A simple trial-and-error calibration (Huang *et al*., 2024) was carried out by modifying parameter sets in order to minimise the biases of the simulated water budget in relation to evapotranspiration products (GLEAM, Martens *et al*., 2017; FLUXCOM, Jung *et al*., 2019) and observed river flow at 1785 gauging stations located across France available in the Hydroportail database.

S1.10 RECHARGE

The RECHARGE model estimates groundwater potential recharge over a given territory including the influence of snowmelt and land cover. “Potential recharge” refers to the total infiltration flow, including both the part that actually reaches and recharges the aquifer, and the part that returns to the river downstream through sub-surface flows. The RECHARGE model has been applied to estimate groundwater potential recharge at various scales, ranging from catchments to entire countries or continents (Lanini *et al*., 2021), with historical data to assess past and current recharge or with climate projections to explore the future groundwater recharge evolution (Caballero *et al*., 2021).

The RECHARGE model computes the effective precipitation based on simple soil water balance models to which have been added a degree-day equation for snow melting and a generalised version of the FAO methodology to account for the land cover on potential evapotranspiration (Allen *et al*., 1998). For the latter, a seasonal crop coefficient is associated to each category of vegetation cover and land use including bare and urbanised soils, as defined in the Corine Land Cover map (2018). Effective rainfall is split between runoff and infiltration using a semi-distributed coefficient known as the effective rainfall/infiltration ratio (EPIR). The EPIR is defined at the scale of groundwater bodies with the strong assumption that infiltration is controlled by the local properties of the soil and subsurface and is independent of climate conditions. EPIR is estimated in gauged catchments using the base flow index (BFI). Empirical relationships between the cartographic index (IDPR) combining the influence of drainage density and hydrological connectivity (Mardhel *et al.*, 2021) and the baseflow index (BFI) taking into account the geological context are calibrated to estimate EPIR at the scale of groundwater bodies. The groundwater potential recharge is calculated using EPIR and the effective precipitation at the scale of groundwater bodies on a daily time step.

The EPIR coefficient is only one parameter to be calibrated. The set of 611 reference gauging stations was used to estimate observed BFI values, and to derive thereafter the empirical relationships between BFI and IDPR. For evaluation purposes, results from the RECHARGE model have been compared to the outputs of the land surface model SURFEX (Le Moigne *et al.*, 2020, see 4.1) over the period 1959-2019.

S1.11 SIM2

The SIM2 model is the operational physically-based hydro-meteorological model of Meteo-France (Le Moigne *et al.*, 2020). It has been in use for a long time for climatology (from 1959 to the present), for real-time monitoring, and for forecasting of water resources over France (Besson *et al.*, 2020), as well as for studies on the impact of climate change on water resources: the Explore2070 project (Chauveau *et al.*, 2013).

SIM2 is made up of three elements. The first one is the atmospheric reanalysis SAFRAN providing meteorological inputs. The second element of SIM2 is the SURFEX module, which includes the ISBA surface scheme. ISBA relies on the ECOCLIMAP2 database to describe soil and vegetation, and uses a multi-layer diffusion scheme (Decharme *et al.*, 2011) to compute the water and energy budget of the soil up to a 12-m depth. ISBA describes vegetation-linked processes such as photosynthesis, evaporation, transpiration, using prescribed parameters (some of them coming from the ECOCLIMAP2 database): the annual cycle of the leaf area index, roughness length, albedo of soil and vegetation. ISBA also features a snow scheme (ISBA-ES) that simulates a 10-layer snow pack and computes its evolution, including melting and sublimation. Moreover, in mountainous areas, a sub-grid orography representation using elevation bands is adopted, and reservoirs are added to represent the effect of aquifers that are not explicitly described, with different parameterizations for mountain and lowland areas. Water information outputs from ISBA, such as runoff and gravitational drainage, are then an input for the third element of SIM2, the hydrogeological MODCOU model. MODCOU computes the streamflows at a 3-hour time step, and the groundwater levels and the exchanges between groundwater and rivers at a daily time step.

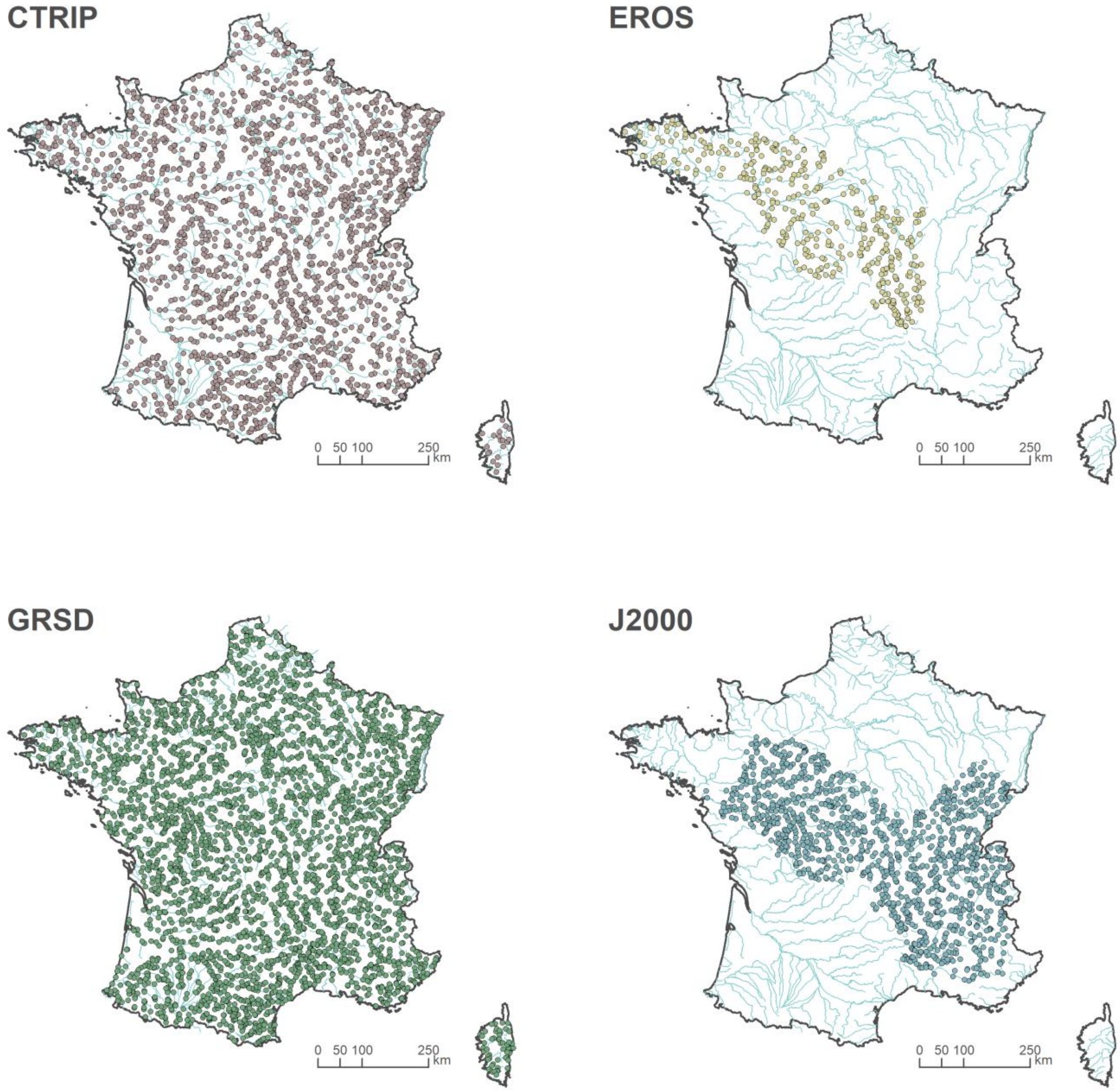
SIM2 operates on a regular 8-km grid that covers the whole of France. In addition, MODCOU describes river grid cells at a resolution from 1 to 8 km. No automatic calibration process is required for the SIM2 model.

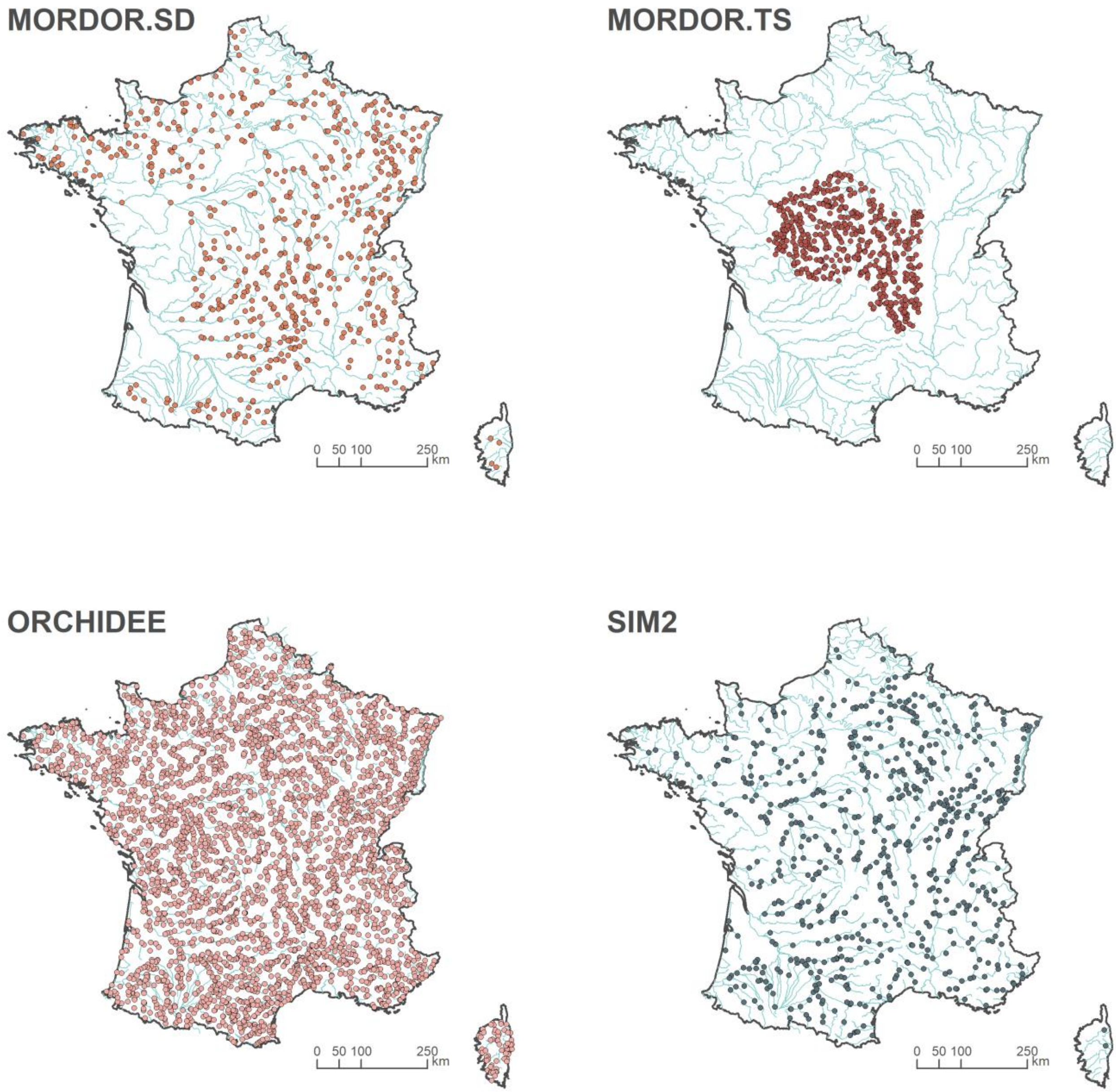
S1.12 SMASH

SMASH is an open-source computing platform developed at INRAE for continuous, spatially distributed hydrological modelling and multi-source data assimilation (https://smash.recover.inrae.fr/). The SMASH platform is based on a conceptual representation of the hydrological processes (several modules are provided) including several optimizers and methods (Jay-Allemand *et al*. 2020). Here, a variational calibration method suited to large-scale issues was adopted to perform a spatially-distributed calibration of all parameters. Furthermore, this spatially-distributed calibration can include multiple gauging stations and physiographic descriptors (Hyunh *et al.*, 2023). As part of the Explore2 project, the SMASH model was implemented with a daily time step, 1-km spatial resolution and a 5-parameter hydrological operator structure. This structure includes a degree-day snow module, the production and transfer reservoirs of the GR4J model (Perrin *et al.*, 2003), and a reservoir for routing hydrographs along the hydrograph network.

A spatially-uniform calibration (a single value for each parameter across all catchment cells) was applied locally for each catchment by maximising the KGE criterion calculated on streamflows. The parameter transfer method used in the regionalization was a transfer that considered the maximum overlap rate and spatial proximity between catchments.

S2. Surface hydrographic network and actual simulated points




MORDOR.SD
0 50 100 250 km
MORDOR.TS
0 50 100 250 km
ORCHIDEE
0 50 100 250 km
SIM2
0 50 100 250 km

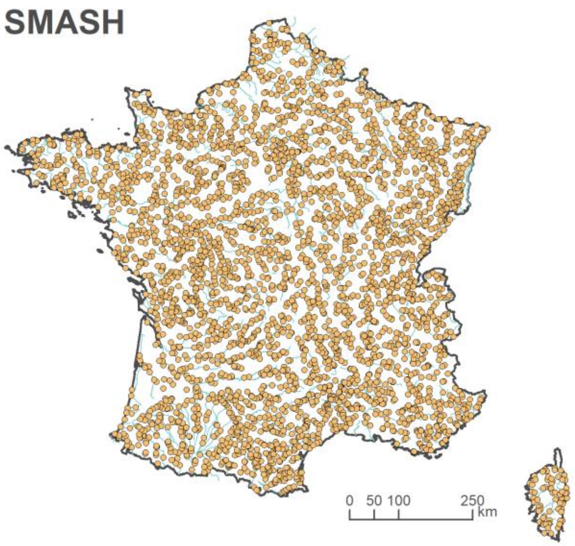


Figure S1: Location of the simulation points for each surface hydrological model.

S3. Local hydrological assessment summary sheets

The header of the local summary sheet (Fig. C1) reports the main basin characteristics (name and coordinates of the gauging station, start, end and length of the observed streamflow time series, and the river flow regime). Two panels in upper right display the hydrograph of the mean monthly streamflow and a map of France with the delineation of the gauged catchment and the location of the selected gauging station. Graphs (a, b, c, d) draw comparisons between observed streamflows and those simulated with SAFRAN as input. The y-axis of graphs c (flow regime) and d (flow duration curve) displays a square-root scale for an easier assessment of low flows. In graph d, exceedance probability values are displayed as the corresponding z value from the standard normal distribution in order to better analyse extreme values. The twelve evaluation scores computed at the gauging station are given in graph e for each available HM with respect to the statistical distribution of the scores at the regional scale. These graphs allow comparisons between HMs at the station, and show how the performance at this point differs from that of the model at other reference gauging stations within the region. The non-shaded area represents the acceptable range for each score. The score on the long-term trend (αQA) is computed and displayed only if the Mann-Kendall test is significant (10 %) at this gauging station. The RAT robustness tests are not represented graphically, their results appear in the warnings at the bottom of the page. The caption at the bottom of the page details the colour associated with each HM (common to all sheets), and the drainage area estimated at the simulation point by the available HMs. The bottom right panel contains pre-constructed sentences that are automatically generated according to a decision tree (if 'scores for one specific HM out of the range' then 'warning for this HM'). Warnings are ordered by scores and limited to seven. The first is an overall assessment based on KGE$\sqrt{}$ and Bias. For each model, when at least one of the two values of these scores is outside the prescribed range, the model is likely to have difficulty in reproducing the hydrological regime. The following warnings are then established: the results of the RAT robustness tests, criteria on the quantiles of flow duration curves, the occurrence of extreme high flows (medtQJXA) and low flows (medtVCN10), the contrast between low-flow and high-flow regimes (aCDC), the long-term trend in annual flow (αQA) and the ϵ elasticities.

Groundwater projections were obtained by running HMs over separate areas, so there is no possible pre-selection there. The reference piezometers have not been the subject of summary sheets. However, the diagnosis is presented on the scale of the national hydrogeological units, as defined by the level-1 of the French hydrogeological database BDLISA (http://bdlisa.eaufrance.fr). Finally, 27 sheets have been devised, 7 of which include only one reference piezometer. The regional summary sheets for groundwater HMs (Fig. C2) are organised as follows:

- The header indicates the name and three-digit code of the hydrogeological unit, the number of reference piezometers included, and the minimum and maximum values of the elevation at which the piezometers are located (metadata extracted from the national ADES database).
- Then, below, four identical graphs are displayed for up to four reference piezometers of the hydrological unit. They compare the observed and simulated interannual median piezometric levels. To ensure the representativeness of the

results, the selected piezometers are those associated with the maximum, 75 % quantile, 25 % quantile and minimum of NSE_{GSI}.

- The set of graphs (e) is the counterpart of the set (e) of the local summary sheets, with the median identified by a coloured thick horizontal line.
- The panel bottom right “COMMENTS” explains how the four reference piezometers were selected.

The content of these sheets is detailed in French in a technical note (Sauquet and Héraut, 2024).

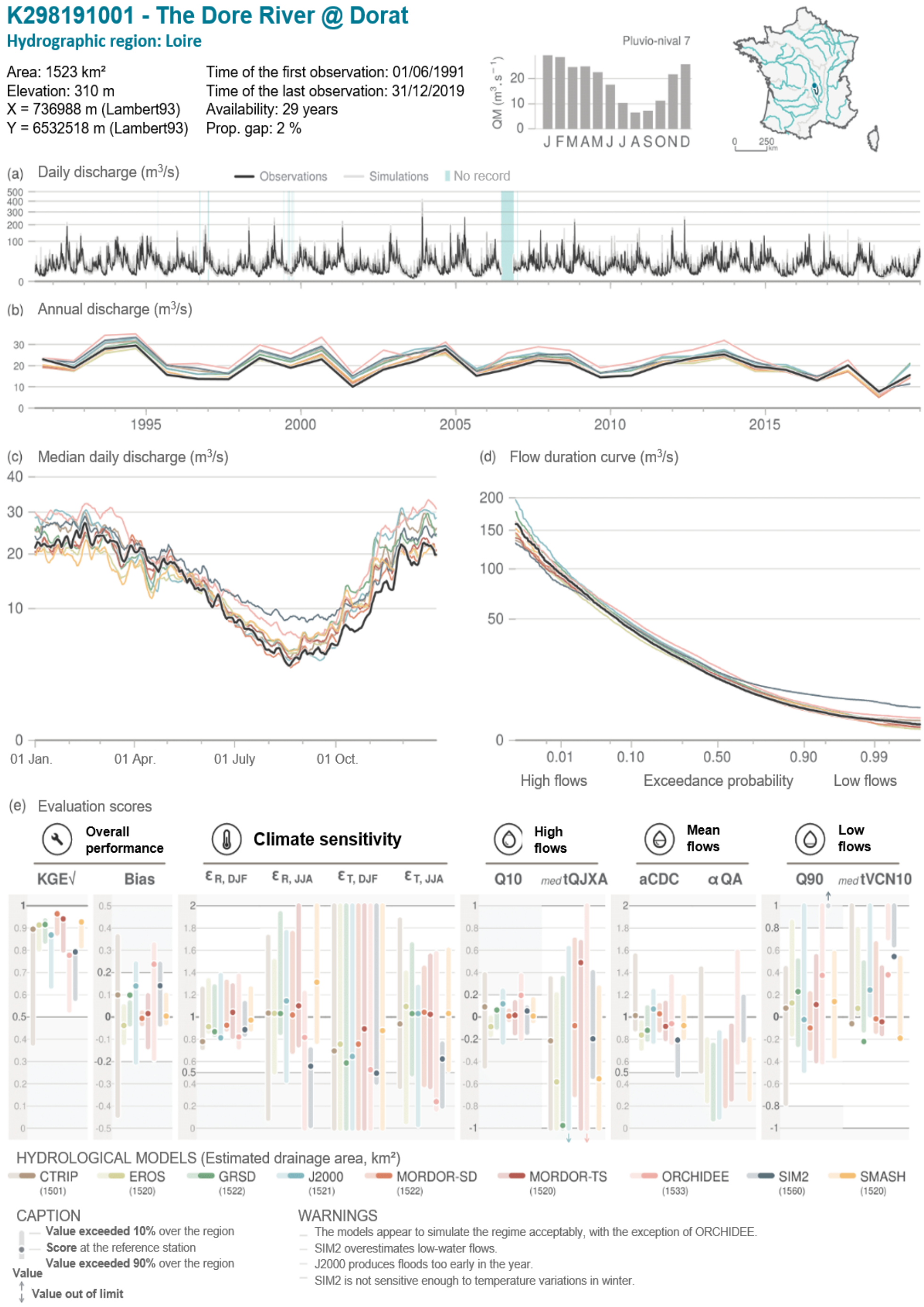


Figure S2: Example of summary sheet produced at one reference gauging station for the evaluation of surface HMs (translated from French).

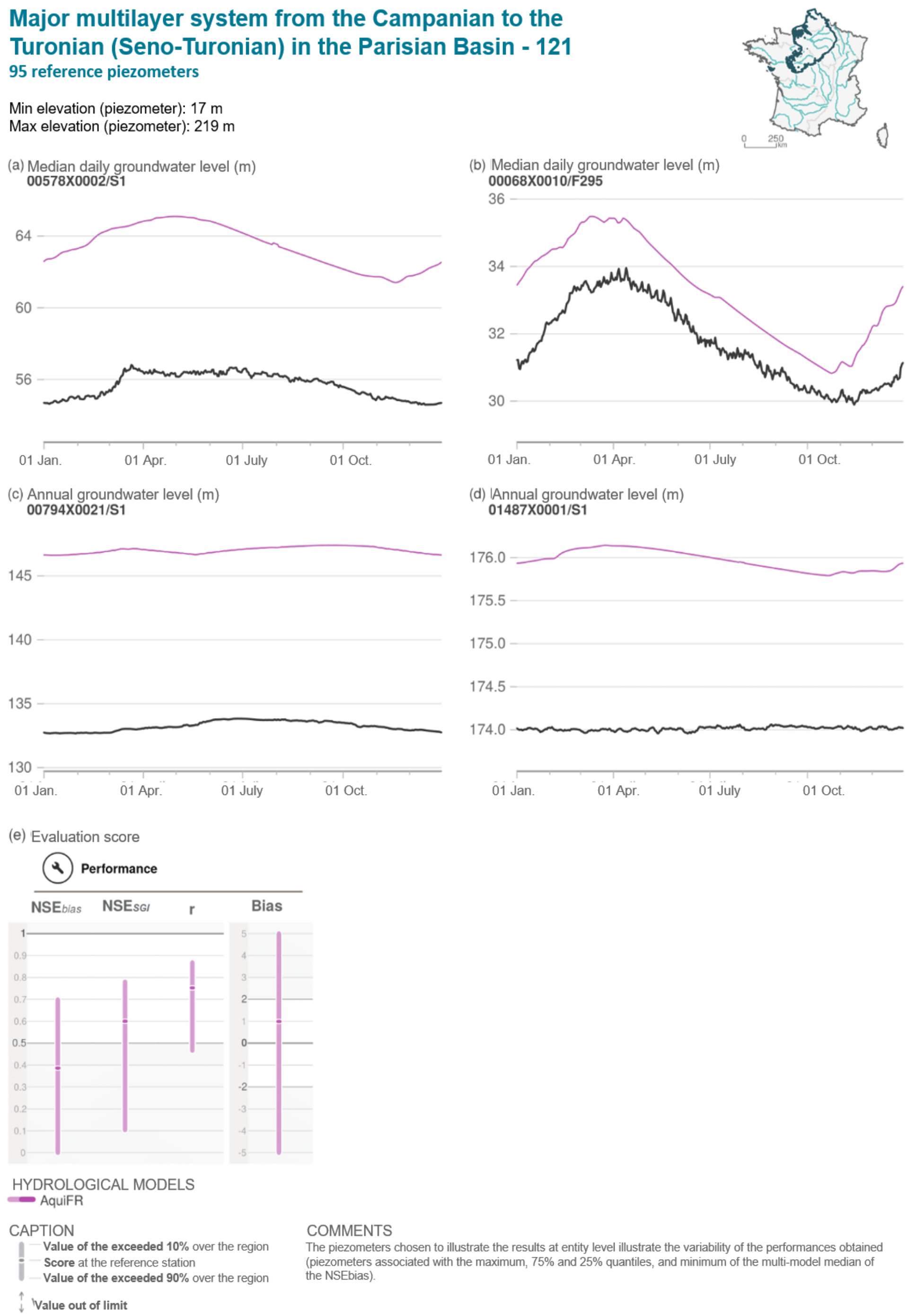


Figure S3: Example of summary sheet produced at one hydrogeological BDLISA unit for the evaluation of groundwater HMs (translated from French).

S4. References

Aissat, R., Pryet, A., Saltel, M., and Dupuy, A.: Comparison of Different Pilot Point Parameterization Strategies When Measurements Are Unevenly Distributed in Space, Hydrogeol. J., 31, 2381-2400. https://doi.org/10.1007/s10040-023-02737-z, 2023.

Amraoui, N., Sbai, M. A., and Stollsteiner, P.: Assessment of Climate Change Impacts on Water Resources in the Somme River Basin (France), Water Resour. Manage., 33, 2073-2092. https://doi.org/10.1007/s11269-019-02230-x, 2019.

Besson F., Etchevers P., Habets F., Le Moigne P., Rousset-Regimbeau F., Soubeyroux J.-M., Viel C., and Vincendon B.: Suivi en temps réel des sécheresses: de l'analyse à la prévision saisonnière, La Houille Blanche, 4, 82–92, https://doi.org/10.1051/lhb/2020042, 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., Ethe, 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., Levy, 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., Ottle, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thieblemont, R., Traore, A. K., Vancoppenolle,M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J. Adv. Model. Earth Syst., 12, e2019MS002 010, https://doi.org/10.1029/2019MS002010, 2020.

Branger, F., Kermadi, S., Jacqueminet, C., Michel, K., Labbas, M., Krause, P., Kralisch, S, and Braud, I.: Assessment of the influence of land use data on the water balance components of a peri-urban catchment using a distributed modelling approach, J. Hydrol., 505, 312-325, 2013.

Caballero, Y., Lanini, S., Le Cointe, P., Pinson, S., Hevin, G., Jódar, J., Lambán, J., Zabaleta, A., Antigüedad, I., and Beguería, S.: Groundwater recharge and groundwater water resources under present and future climate over the Pyrenees (France, Spain, Andorre), EGU General Assembly 2021, Online, 19–30 Apr 2021, EGU21-16471, https://doi.org/10.5194/egusphere-egu21-16471, 2021.

Corine Land Cover: https://doi.org/10.2909/71c95a07-e296-44fc-b22b-415f42acfdf0, 2018.

Coron, L., Delaigue, O., Thirel, G., Dorchies, D., Perrin, C., and Michel, C.: airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling. R package version 1.6.12, 2021.

Coron, L., Thirel, G., Delaigue, O., Perrin, C., and Andréassian, V.: The Suite of Lumped GR Hydrological Models in an R package, Environmental Modelling & Software, 94, 166-171, https://doi.org/10.1016/j.envsoft.2017.05.002, 2017.

de Lavenne, A., Andréassian, V., Thirel, G., Ramos, M.-H., and Perrin, C.: A regularization approach to improve the sequential calibration of a semi-distributed hydrological model, Water Resources Research, 55, 8821-8839. https://doi.org/10.1029/2018WR024266, 2019.

Decharme, B., Boone A., Delire C., and Noilhan J.: Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions, J. Geophys. Res., 116, D20126, https://doi.org/10.1029/2011JD016002, 2011.

Decharme, B., Delire, C., Minvielle, M., and Colin, J.: Recent changes in the ISBA-CTRIP land surface system for using in the CNRM-CM6 climate model and in global off-line hydrological applications, Journal of Advances in Modeling Earth Systems, 1, 1–92, https://doi.org/10.1029/2018MS001545, 2019.

Dorchies, D., Delaigue, O. and Thirel, G.: airGRiwrm: Modeling of Integrated Water Resources Management based on airGR. R package version 0.7.0, doi: 10.57745/XKN6NC, URL: https://cran.r-project.org/package=airGRiwrm, 2024

Gao, T., Kang, S., Krause, P., Cuo, L., and Nepal, S.: A test of J2000 model in a glacierized catchment in the central Tibetan Plateau, Environmental Earth Sciences, 65, 1651-1659, 2012.

Garavaglia, F., Le Lay, M., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E., and Mathevet, T.: Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi distributed approach, Hydrology and Earth System Sciences, 21, 3937–3952, https://doi.org/10.5194/hess-21-3937-2017, 2017.

Hayami, S.: On the propagation of flood waves. Disaster Prevention Res. Inst. Bull. 1, 1951.

Héraut, L., Vidal, J.-P., Évin, G., and Sauquet, E.: Notice de lecture des fiches de résultats des modèles hydrologiques de surface, Recherche Data Gouv, https://doi.org/10.57745/6YNIUF, 2024.

Horner, I.: Design and evaluation of hydrological signatures for the diagnostic and improvement of a process-based distributed hydrological model, Ph.D. thesis, Université Grenoble Alpes, France, 2020.

Huang, P., Ducharne, A., Rinchiuso, L., Polcher, J., Baratgin, L., Bastrikov, V., and Sauquet, E.: Multi-objective calibration and evaluation of the ORCHIDEE land surface model over France at high resolution, Hydrol. Earth Syst. Sci., 28, 4455–4476, https://doi.org/10.5194/hess-28-4455-2024, 2024.

Hyunh, N. N. T., Garambois, P.-A., Colleoni, F., Renard, B., and Roux, H.: Multi-gauge hydrological variational data assimilation: Regionalization learning with spatial gradients using multilayer perceptron and bayesian guided multivariate regression. Proceeding, SHF conference "Prévision des crues et inondations, Avancées, valorisations et perspectives", Toulouse, https://hal.inrae.fr/hal-04149040v1, 2023.

Jay-Allemand, M., Javelle, P., Gejadze, I., Arnaud, P., Malaterre, P.-O., Fine, J.-A., and Organde, D.: On the potential of variational calibration for a fully distributed hydrological model: application on a Mediterranean catchment, Hydrology and Earth System Sciences, 5519-5538, https://doi.org/10.5194/hess-24-5519-2020, 2020.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Scientific Data, 6, 74, https://doi.org/10.1038/s41597-019-0076-8, 2019.

Kralisch, S. and Krause, P.: JAMS–A framework for natural resource model development and application, 2006.

Krause, P., Bäse, F., Bende-Michl, U., Fink, M., Flügel, W., and Pfennig, B.: Multiscale investigations in a mesoscale catchment–hydrological modelling in the Gera catchment, Advances in Geosciences, 9, 53-61, https://doi.org/10.5194/adgeo-9-53-2006, 2006.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, https://doi.org/10.1029/2003GB002199, 2005.

Lanini, S., Caballero, Y., Hevin, G., Le Cointe, P., Pinson, S., and Desprats, J.-F.: Groundwater Recharge Indicators at the European scale. EGU General Assembly 2021, Online, 19–30 Apr 2021, EGU21-2327, https://doi.org/10.5194/egusphere-egu21-2327, 2021.

Litrico, X. and Georges, D.: Robust continuous-time and discrete-time flow control of a dam-river system. (I) Modelling, Appl. Math. Model., 23, 809–827, https://doi.org/10.1016/S0307-904X(99)00014-1, 1999.

Mardhel, V., Pinson, S., and Allier, D.: Description of an indirect method (IDPR) to determine spatial distribution of infiltration and runoff and its hydrogeological applications to the French territory, J. Hydrol., 592, 125609. https://doi.org/10.1016/j.jhydrol.2020.125609, 2021.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Mimeau, L., Künne, A., Branger, F., Kralisch, S., Devers, A., and Vidal, J.-P.: Flow intermittence prediction using a hybrid hydrological modelling approach: influence of observed intermittence data on the training of a random forest model, Hydrol. Earth Syst. Sci., 28, 851–871, https://doi.org/10.5194/hess-28-851-2024, 2024.

Monteil, C., Zaoui, F., Le Moine, N., and Hendrickx, F.: Multi-objective calibration by combination of stochastic and gradient-like parameter generation rules – the caRamel algorithm, Hydrol. Earth Syst. Sci., 24, 3189-3209, https://doi.org/10.5194/hess-24-3189-2020, 2020.

Morel, M., Pella, H., Branger, F., Sauquet, E., Grenouillet, G., Côte, J., Braud, I., and Lamouroux, N.: Catchment-scale applications of hydraulic habitat models: Climate change effects on fish, Ecohydrology, e2513, https://doi.org/10.1002/eco.2513, 2022.

Munier, S. and Decharme, B.: River network and hydro-geomorphological parameters at 1⁄12° resolution for global hydrological and climate studies, Earth Syst. Sci. Data, 14, 2239–2258, https://doi.org/10.5194/essd-14-2239-2022, 2022.

Pédron N. and Platel, J.-P.: Gestion des eaux souterraines en Région Aquitaine. Développements et maintenance du Modèle Nord Aquitain de gestion des nappes. Rapport, Orléans, BRGM/RP -53659- FR, 2005.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, J. Hydrol., 279, 275-289, https://doi.org/10.1016/S0022-1694(03)00225-7, 2003.

Polcher, J., Schrapffer, A., Dupont, E., Rinchiuso, L., Zhou, X., Boucher, O., Mouche, E., Ottlé, C., and Servonnat, J.: Hydrological modelling on atmospheric grids: using graphs of sub-grid elements to transport energy and water, Geosci. Model Dev., 16, 2583–2606, https://doi.org/10.5194/gmd-16-2583-2023, 2023.

Rosenbrock, H. H.: An Automatic Method for Finding the Greatest or Least Value of a Function, The Computer Journal, 3, 175-184, https://doi.org/10.1093/comjnl/3.3.175, 1960.

Rouhier, L., Le Lay, M., Garavaglia, F., Le Moine, N., Hendrickx, F., Monteil, C., and Ribstein, P.: Impact of mesoscale spatial variability of climatic inputs and parameters on the hydrological response, J. Hydrol., 553, 13–25, https://doi.org/10.1016/j.jhydrol.2017.07.037, 2017.

Rouhier, L.: Régionalisation d'un modèle hydrologique distribué pour la modélisation de bassins non jaugés. Application aux vallées de la Loire et de la Durance, Ph.D. thesis, Sorbonne Université, France, 2018.

Royer-Gaspard P., Bourgin F., Perrin C., Andréassian V., De Lavenne A., Thirel G., and Tilmant F.: Benefits of upstream data for downstream streamflow forecasting: data assimilation in a semi-distributed flood forecasting model, LHB, 2374081, https://doi.org/10.1080/27678490.2024.2374081, 2024.

Saleh, F., Ducharne, A., Flipo, N., Oudin, L., and Ledoux, E.: Impact of river bed morphology on discharge and water levels simulated by a 1D Saint–Venant hydraulic model at regional scale, J. Hydrol., 476, 169-177. https://doi.org/10.1016/j.jhydrol.2012.10.027, 2012.

Saltel, M., Wuilleumier, A., and Cabaret, O.: Gestion des eaux souterraines en Région Aquitaine - Développements et maintenance du Modèle Nord-Aquitain de gestion des nappes. Rapport, Orléans, BRGM/RP-65039-FR, 2016.

Sauquet, E. and Héraut, L.: Notice de lecture des fiches « diagnostic » des modèles hydrologiques, Recherche Data Gouv, https://doi.org/10.57745/MDHS0D, 2024.

Seyedhashemi, H., Vidal, J.-P., Diamond, J. S., Thiéry, D., Monteil, C., Hendrickx, F., Maire, A., and Moatar, F.: Regional, multi-decadal analysis on the Loire River basin reveals that stream temperature increases faster than air temperature, Hydrol. Earth Syst. Sci., 26, 2583–2603, https://doi.org/10.5194/hess-26-2583-2022, 2022.

Steudel, T., Bugan, R., Kipka, H., Pfennig, B., Fink, M., de Clercq, W., Flügel, W.-A., and Helmschrot, J.: Implementing contour bank farming practices into the J2000 model to improve hydrological and erosion modelling in semi-arid Western Cape Province of South Africa, Hydrol. Res., 46, 192-211, 2015.

Tafasca, S., Ducharne, A., and Valentin, C.: Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model, Hydrol. Earth Syst. Sci., 24, 3753–3774, https://doi.org/10.5194/hess-24-3753-2020, 2020.

Thiéry, D. and Moutzopoulos, C.: Un modèle hydrologique spatialisé pour la simulation de très grands bassins : le modèle EROS formé de grappes de modèles globaux élémentaires, VIIIèmes journées hydrologiques de l'ORSTOM : Régionalisation en hydrologie, application au développement, ORSTOM Editions, 285-295, https://hal-brgm.archives-ouvertes.fr/hal-01061971/document, 1992.

Thiéry, D.: Code de calcul MARTHE - Modélisation 3D des écoulements dans les hydrosystèmes - Notice d'utilisation de la version 7.5 (MARTHE: Modeling software for groundwater flows), Rapport, Orléans, BRGM/RP-64554-FR, 2015.

Thiéry, D.: Logiciel ÉROS version 7.1 - Guide d'utilisation, Rapport final, Orléans, BRGM/RP-67704-FR, 2018.

Valéry, A., Andréassian, V., and Perrin, C.:'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments, J. Hydrol., 517, 1176-1187, https://doi.org/10.1016/j.jhydrol.2014.04.058, 2014.

Vergnes, J. P. and Decharme, B.: A simple groundwater scheme in the TRIP river routing model: Global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges, Hydrol. Earth Syst. Sci., 16, 3889–3908. https://doi.org/10.5194/hess-16-3889-2012, 2012.

Vergnes, J.-P., Roux, N., Habets, F., Ackerer, P., Amraoui, N., Besson, F., Caballero, Y., Courtois, Q., de Dreuzy, J.-R., Etchevers, P., Gallois, N., Leroux, D. J., Longuevergne, L., Le Moigne, P., Morel, T., Munier, S., Regimbeau, F., Thiéry, D., and Viennot, P.: The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: evaluation over a 60-year period, Hydrol. Earth Syst. Sci., 24, 633–654, https://doi.org/10.5194/hess-24-633-2020, 2020.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, J. Geophys. Res.: Atmos., 118, 6064–6079, https://doi.org/10.1002/jgrd.50395, 2013.

Watson, A., Miller, J., Fleischer, M., and De Clercq, W.: Estimation of groundwater recharge via percolation outputs from a rainfall/runoff model for the Verlorenvlei estuarine system, west coast, South Africa, J. Hydrol., 558, 238-254, https://doi.org/10.1016/j.jhydrol.2018.01.028, 2018.

Yamazaki, D., Sosa, J., Bates, P. D., Allen, G., and Pavelsky, T.: MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets, Water Resour. Res., 2019WR024873, https://doi.org/10.1029/2019WR024873, 2019.