Response to reviewer hess-2019-166
The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: evaluation over a 60 year period.

The authors wish to thank the two reviewers for all their corrections, comments and suggestions of improvement for this manuscript. The following text proposes a detailed response for each of them. Bold texts correspond to the author's response.

## Anonymous Referee #1

- P2L13: includes > include. **Done**
- P6L4: Error in reference. **Done**
- P10L10: newfluxes > new fluxes. **Done**

## Anonymous Referee #2

1. P1, L32: "Nash-Sutcliffe" **Done**

2. P3, L13: "State scale" **Done**

3. P4, L9-25: I found this part not well placed in the introduction, since it starts describing in details the AquiFR platform, which is the objective of the subsequent section (section 2), with some overlap between these two parts. I suggest moving this part at the beginning of section 2 and removing the redundant information with this section. **The authors agree with this comment. We merged this part with section 2 it in the new revised manuscript.**

4. P4, L13: "and contain" **Done**

5. P5, L32: Remove "Erreur…introuvable" **Done**

6. P6, L4: Remove "Erreur…introuvable" **Done**

7. P6, 26: "assigned to" **Done**

8. P7, L13: "is accounted for" **Done**

9. P7, L26: "dimensional" **Done**

10. P9, L13: "99 km²" **Done**

11. P9, L21-22: Is this use of average pumping estimates realistic? Was this mean cycle calculated over all pumping data available? Would not it be more realistic to take only the first years of the available records and to extend this over the past years? For aquifers where water abstractions have much increased, taking a mean over all the available pumping data may overestimate the abstractions over the most remote parts of the evaluation period. Could the author add a few words on this? **It is not that obvious that the pumping has increased over time. Indeed, pumping from industry was very important in some places in the past, and has decreased significantly. Irrigation varies according to the climate and the equipment. For this reason, it sounds better to impose a mean. The revised manuscript includes a few words about this: "This choice is linked to the lack of knowledge about past pumping. However, we do know that there have been antagonistic**

developments between irrigation and industrial pumping. Irrigation has increased in accordance with the irrigated areas while it varied greatly depending on the climate. Industrial pumping was dominant in the past but has considerably decreased during the past decades (Service de l'observation et des statistiques, 2016)".

Service de l'observation et des statistiques: Repères. L'eau et les milieux aquatiques. Chiffres clés., Ministère de l'Environnement, de l'Energie et de la Mer, 2016.

12. P9, L26: "do not cover" **Done**

13. P10, L3: "new fluxes" **Done**

14. P10, L4: "will be soon updated": any reference to this ongoing work? **It is expected to do it in the next phase of the project, therefore, in the next 2 coming years.**

15. P10, L5: I found it difficult to check all the cited references to get the information on the periods used for calibration. Could the authors add a column in Table 1 to give the period used for calibration in each case? (for each line of the table, if different calibration periods are considered, maybe show the longest period used in each case). This would help the reader assess the level of overlap between the simulation period used in this study and the calibration period. **Thank you for this comment. The revised manuscript includes the calibration period shown directly in the column named Recalibration if necessary.**

16. P10, L10: Similarly to the previous comment, it would be useful for the reader if the authors could add an extra column in Table 1 giving the type of criteria used for calibration (and possibly the various variables – piezometric level and/or flows). This would show the various strategies used in these past works. **An extra column is added to specify the type of variables used for recalibrated the models. The following text is added to the legend: "Periods of calibration are given in the Recalibration column and the type of variables used for recalibration in the Variables column. GW means groundwater level and RF river flow. GW levels were evaluated using RMSE and bias criteria. River flows were evaluated using NSE and ratio criteria."**

17. P10, L29: "relative": the formulation is not expressed in relative but absolute terms **Done**

18. P10, L31: In the equations, the summation symbol refers to i index but then the variable X is referring to t. Maybe use instead Q(i) to be more consistent in notations. Change notations accordingly for the other equations. Maybe remind that BIAS has the same unit as X and that the perfect value is 0, with negative values corresponding to overestimation and positive values corresponding to underestimation by the model (or maybe there is an error in Eq.1, see also comment # 25 on this issue). **Done**

19. P11, top: Please give a few words on how the values of the NRMSE_BE criterion should be interpreted (what could be considered a good NMRSE_BE value?). **A comment is added to the text: "The NRMSE_BE criterion is always positive and starts from 0 for a perfect simulation of the observed amplitudes. A NRMSE_BE lower than 0.8 can be considered as a reasonable estimation of the temporal evolution of the observed water table."**

20. P11, L11: I found this sentence unclear. What do you mean? **The authors agree. NSE is strongly affected by biases between observation and simulation. Therefore, a bias value that can be considered as reasonable (i.e. < 1 m) can lead to poor NSE scores. If the reviewer agree, we replace this sentence in the revised manuscript by "Its use for comparing groundwater levels is less obvious regarding its strong sensitivity to the biases between observation and simulation.**

21. P12, L8-21: This part is not well placed in the Results section, since it presents the data used. I suggest moving it to a new subsection on data at the end of section 3. **The authors agree. This part is moved in a new section « 3.4 Dataset and model setup » in the revised manuscript.**

22. P12, L8-21: I did not find information on the way model warm-up is made at the beginning of the simulation period. I guess this can be an issue and a potential source of error. Could the authors give some information on this? **Initial conditions correspond to a proxy year of the 1958 year that was defined using time series of long-term observed groundwater levels starting from 1958. A sentence about this is included in the new subsection 3.4 Data : "State variables from August 1, 2013 were chosen from a first simulation over the 1958-2018 period in order to initialize the simulation in August 1, 1958. The 2013 year was chosen as the best proxy year of the 1958 year by analysing long-term time series of observed groundwater levels having data since 1958."**

23. P12, L12: "a few measurements" **Done**

24. P12, L20: "not yet fully available" **Done**

25. P12, L23-24: From equation (1), a positive value would mean an underestimation. **Thank you. This sentence was right but there was a mistake in the description of Equation 1. It is now corrected in the revised manuscript.**

26. P12, L26: instead of "accumulated distribution", use "cumulative distribution". Change accordingly elsewhere in the text and figures (graphs and captions in Fig. 6, 7, 15) **Done**

27. P12, L29-30: How these values can be interpreted? (see comment #19) **Section 3.3 includes now a few word about this (see answer to comment #19).**

28. P13, L6-7: The observations show annual cycles starting from 2008. Is there any change in the measurement device? **Yes, observed values were monitored monthly until 1995, then weekly until 2008, and then daily after this date.**

29. P13, L8 and 11: "Helloin" is written with uppercase elsewhere in the article. **Done**

30. P13, L12: This period seems also badly simulated for the Omiécourt and Farceaux case studies, which are all in the same region. Is it just a coincidence or were there specific climatic conditions on this period, which are not well captured by the models? **Thank you for this remark. The authors agree about this comment. The fact that these biases occur in different regions with different hydrogeological models suggests indeed an underestimation of recharge from SURFEX. However, errors in the hydrogeological models cannot be excluded. A specific analysis of these fluxes that is beyond the scope of this paper could be useful to explain these biases.**

31. P13, L24-26: I am not sure these two figures are really useful. Very little is said on these figures and they do not seem to convey different information. I would remove the figure on correlation. **The authors agree to the fact that the Figure on Correlation brings little information. We keep only the NSE Figure in the revised manuscript, as suggested by the reviewer.**

32. P14, L29: "Kling-Gupta efficiency" **Done**

33. P15, L13: "negative" **Done**

34. P15, L19: Not sure this comment is fully consistent with what is said on top of page 13. Please clarify this. **Thank you for this remark. This was an error. We meant poor bias scores instead of poor NRMSE-BE scores for the Loire River. This comment is now consistent with top of page 13.**

35. P16, L18: "in section 3.2" **Done**

36. P16, L17-19: Generally, testing a model outside calibration range is considered a good way to evaluate its extrapolation capacity and robustness. Here it is difficult to evaluate this since the calibration and evaluation periods overlap. The authors could discuss this point. Is there some perspective to more thoroughly evaluate the robustness of model proposed in AquiFR? **According to the period of calibration of each model now described in the new Table 1 of the revised manuscript, no calibration was undertaken before 1986 and after 2015 whatever the model used. Thus, in a few years, we will have several independent years after 2015 with observations to consider a better evaluation of the robustness of the model.**

37. Fig. 6: Could the cumulative distribution give also the value for an absolute bias lower than 1, to make a better correspondence with the scale of the map above (in fig. a). **Done**

38. Fig. 12, caption: "for June 2016 for" **Done**

39. Fig. 13: It seems that the observations are not available on the whole period in each case. But it is a bit difficult to distinguish between the case where observed and simulated values perfectly overlap from the case where there are gaps in the observations. I suggest adding the information on periods of gaps in the observations, e.g. by putting a grey background on the graphs for the corresponding periods. **Done**

40. Fig. 13: in the caption: "dashed blue" **Done**

41. Fig. 14: Same comments as #39 and 40 **Done**

42. Table 4: The information contained in the table could be directly introduced in Fig. 14 (in brackets after the name of the catchment, as done in Fig. 13). This would reduce the number of tables. **Done**

## List of all relevant changes

All the relevant changes made to the manuscript correspond to the technical corrections previously stated by the reviewers as well as complements that were arise by the reviewers, including:

- A paragraph of the Introduction that was moved and merged to Section 2: The AquiFR Hydrometeorological Modelling Platform
- The introduction of the Section 4 Results that is moved to a new subsection Section 3.4 Dataset and model setup
- Figure 10 and Table 4 were deleted according to the suggestions of reviewer #2.

## Marked-up manuscript

# The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: evaluation over a 60-year period.

Jean-Pierre Vergnes[1], Nicolas Roux[2], Florence Habets[3,4], Philippe Ackerer[5], Nadia Amraoui[1], François
Besson[6], Yvan Caballero[1], Quentin Courtois[7], Jean-Raynald de Dreuzy[7], Pierre Etchevers[6], Nicolas
Gallois[8], Delphine J. Leroux[2], Laurent Longuevergne[7], Patrick Le Moigne[2], Thierry Morel[9], Simon
Munier[2], Fabienne Regimbeau[6], Dominique Thiéry[1], Pascal Viennot[8]

[1]BRGM French Geological Survey, Water Environment and Ecotechonologies Division, Orléans Cedex 2, France
[2]GAME/CNRM, Météo France, CNRS, Toulouse, France
[3]CNRS/Sorbonne University, UMR 7619 Métis, Paris Cedex 5, France
[4]Laboratoire de Géologie, Ecole Normale Supérieure, PSL Research University, CNRS UMR 8538, 24 rue Lhomond, 75005
Paris, France
[5]LHYGES, UMR 7517 CNRS, EOST / Université de Strasbourg, France
[6]METEO-FRANCE/DCSC, Toulouse, France
[7]Géosciences Rennes, UMR 6118, CNRS, Université de Rennes I, Rennes, France
[8]Centre de Géosciences, MINES ParisTech, Fontainebleau, France
[9]CERFACS, Toulouse, France

*Correspondence to*: Jean-Pierre Vergnes (jp.vergnes@brgm.fr)

> **Code de champ modifié**

**Abstract.** The new AquiFR hydrometeorological modelling platform was developed to provide short to long-term forecasts
for groundwater resource management in France. This study aims to describe and assess this new tool over a long period of 60
years. This platform gathers in a single numerical tool several hydrogeological models covering much of the French
metropolitan area. Eleven aquifer systems are simulated through spatially distributed models using either the MARTHE
groundwater modelling software or the EauDyssée hydrogeological platform. Twenty-three karstic systems are simulated by
a lumped reservoir approach using the EROS software. AquiFR computes the groundwater level, the groundwater surface
water exchanges, and the river flows. A simulation covering a 60-year period from 1958 to 2018 is achieved in order to evaluate
the performance of this platform. The 8 km resolution SAFRAN meteorological analysis provides the atmospheric variables
needed by the SURFEX land surface model in order to compute surface runoff and groundwater recharge used by the
hydrogeological models. The assessment is based on more than 600 piezometers and more than 300 gauging stations
corresponding to simulated rivers and outlets of karstic systems. For the simulated piezometric heads, 42% and 60% of the
absolute biases are lower than 2 m and 4 m respectively. The Standardized Piezometric Level Index (SPLI) was computed to
assess the ability of AquiFR to identify extreme events such as groundwater floods or droughts in the long-term simulation
over a set of piezometers used for groundwater resource management. 56% of the Nash-SutcliffSutcliffe coefficient calculated
between the observed and simulated SPLI time series are greater than 0.5. The quality of the results makes it possible to

consider using the platform for real-time monitoring and seasonal forecasts of groundwater resources as well as for climate change impact ~~assessements~~assessments.

## 1 Introduction

Groundwater is the most important freshwater resources on Earth. It is widely used for drinking water, agricultural, and
5    industrial use. Knowing the spatial and temporal evolutions of groundwater and being able to predict its future evolution over short to long-term periods are essential to manage water resources management and anticipate climate change impacts. However, a strong spatial heterogeneity characterizing groundwater makes its monitoring difficult. Thus, it is mostly monitored through well networks that can give information only at specific locations (Aeschbach-Hertig and Gleeson, 2012; Fan et al., 2013). Remote sensing gravimeters can provide large scale estimates of groundwater storage changes (Long et al., 2015) but
10   it is not suited for regional scale studies (Longuevergne et al., 2010). Therefore, modeling can be a useful tool to provide meaningful information on the groundwater resources (Aeschbach-Hertig and Gleeson, 2012) at different spatial scales and different temporal periods in the past or in the future.

An increasing number of numerical weather prediction models include~~s~~ a representation of groundwater (Barlage et al., 2015; Sulis et al., 2018). Nevertheless, such representations are not detailed enough to be used to monitor or to forecast the
15   groundwater resources. This is the reason why some dedicated approaches aim at providing groundwater level forecasts at the well scale with lumped models (Prudhomme et al., 2017) or neural networks (Amaranto et al., 2018; Dudley et al., 2017; Guzman et al., 2017).

At the regional scale, only a few modelling approaches use spatially distributed models to monitor and forecast the groundwater resource. Henriksen et al. (2003) presented the development of national hydrogeological models in Denmark aiming at
20   gathering competencies from research organizations and water agencies and establishing a national overview of the present state and future trends of groundwater resources. An integrated groundwater/surface water hydrological model covering a spatial extension of about 43 000 km$^2$ with a 1 km grid resolution was then developed by making full use of the available data (Henriksen et al., 2003). The modelling system is composed of 11 regional sub-models. This model has been regularly updated, as reported by Højberg et al. (2013) who used local studies in relation with active stakeholders to include local data to improve
25   the national model. The Danish model is planned to be used for real time monitoring (He et al., 2016) and climate change studies (Højberg et al., 2013).

In the Netherlands, national and regional water authorities decided to build the Netherlands Hydrological Instrument (NHI) which couples various physical models for all parts of the water system in order to support long-term plans for sustainable water use and safety under changing climate conditions (De Lange et al., 2014). The model was developed by research
30   institutions but local knowledge has been adopted in cooperation with the national water boards (Højberg et al., 2013). It aims to be a model for long-term national policymaking and real-time forecasting for daily water management.

In the United Kingdom (UK), Pachocka et al. (2015) used a numerical model to compute the piezometric head evolution of the three most important UK unconfined aquifers using a finite difference scheme. These three unconfined aquifer basins were discretized into a 5 km resolution grid and connected to a river network. The model was tested against 37 gauging stations distributed across the country. Good fit to the observations was obtained in a steady state run. This study seems to be the first

5  step toward a system that will be used for water management studies and climate impact studies.

Another study covering a wide domain corresponding to the major part of the United States (6.3 billion of km$^2$) was carried out by Maxwell et al. (2015). A 3 dimensional hydrogeological model (ParFlow) was used at a 1 km grid resolution in steady state run. This model has four layers over the first meter of soil, and then a fifth layer from 1 to 100 meter depths. The computation time was one week on high-performance computer for a steady-state simulation. Thus, while this study confirms

10  the possibility of running a 3D groundwater model at fine resolution over a very large territory, it is still difficult to consider its application for operational water management purpose.

Other examples include the Texas Water Development Board that has implemented several sub-models to help monitor groundwater resources at the State scale (more than 500 000 km$^2$ (Texas Water Development Board, 2018)) or New Zealand where a nationwide groundwater recharge model is currently under development (Westerhoff et al., 2018).

15  In France, the hydrometeorological model SAFRAN-ISBA-MODCOU (Habets et al., 2008) that is used for long-term reanalyses (Vidal et al., 2010) as well as real time monitoring (Coustau et al., 2015) and forecast (Singla et al., 2012; Thirel et al., 2010) includes an explicit representation of two aquifer systems. However, the representation of these aquifer systems is rather coarse and is mostly used to have a realistic representation of the river base flow (Rousset et al., 2004) rather than providing consistent information on groundwater resources. Vergnes et al. (2012) developed a hydrogeological model

20  dedicated to climate modelling that was first applied over France and on a global scale (Vergnes and Decharme, 2012). However, only a single layer at the resolution of approximately 10 km over France was considered. This approach is still too coarse to be used for groundwater management over France.

The need to have a national scale consistent representation of groundwater resources in France clearly appeared during the project EXPLORE2070 led by the French environment ministry that aimed at providing climate projections of the evolution

25  of the water resource in France including groundwater (Stollsteiner, 2012). Several regional hydrogeological models were used in this project, together with downscaled climate change projections. The results were difficult to analyse due to the differences in the way the surface water balance was calculated (either lumped-parameters model or soil-vegetation-atmosphere scheme), in the initialization methods, and in the way the evolutions were estimated. Moreover, in the meantime, several regional groundwater models were developed independently by research institutions in close relationship with the stakeholders for

30  regional water management purposes or climate impact studies (Amraoui et al., 2014; Croiset et al., 2013; Douez, 2015; Habets et al., 2010; Monteil et al., 2010; Vergnes and Habets, 2018).

In such context, the AquiFR project was initiated to capitalize these developments in order to provide real time monitoring (Coustau et al., 2015) and forecasts (Singla et al., 2012; Thirel et al., 2010) of groundwater resources in France, as well as long-term reanalyses and future projections. The project associates research teams in hydrogeological, numerical modelling,

3

and atmospheric fields. ~~It is funded by a national stakeholder in charge of the water resource, the French Agency for Biodiversity (AFB)~~A national stakeholder in charge of the water resource, the French Agency for Biodiversity (AFB), funds it. The main idea of AquiFR is to include existing hydrogeological models developed with different groundwater modelling software and to connect them with real time atmospheric analysis and weather forecasts for producing relevant information

5      for water resource management through a single numerical tool. This project also encourages new developments over areas where no groundwater models currently exist. To achieve these objectives the AquiFR hydrogeological modelling platform was developed. The main objectives of this paper are to describe this platform, to evaluate its performance against observations, and to prove its suitability and robustness for operational and research purposes.

~~In its present form, the AquiFR system includes 3 hydrogeological modelling software covering 11 sedimentary aquifers and~~
10     ~~23 karstic systems: the EauDyssée hydro(geo)logical numerical platform (Saleh et al., 2013), the MARTHE groundwater flow software (Thiéry, 2015a) and the EROS lumped model software used for karstic systems (Thiéry, 2018a). These software are embedded in an application developed with the OpenPALM coupling system (Duchaine et al., 2015). All these models cover an area of about 149 000 km² and contains up to 10 overlaid aquifer layers.~~ Prior to real time monitoring and forecast, AquiFR need to be assessed on a long-time period, which is presented in the present study. The evaluation is carried out over a 60-year

15     period from 1958 to 2018 at a daily time step. This long-term simulation provides a unique insight on the long-term evolution of groundwater in France, as most of the groundwater data are available over about 30 years. This long-term simulation can then be used to characterize the daily situation compared to past events. ~~The SAFRAN meteorological analysis (Vidal et al., 2010), available over the French metropolitan area at an 8 km resolution, allow to supply the meteorological variables to the SURFEX land surface model (Masson et al., 2013) which evaluates the water balance over the French metropolitan area. It~~
20     ~~must be stressed that the hydrogeological models could have been classically fed with the SAFRAN analysis precipitation, potential evapotranspiration, and temperature data using their own water balance calculation. However, the combined use of SURFEX and SAFRAN provides a consistent set of hydro-meteorological data over an 8 km resolution grid over France, including groundwater recharge and surface runoff from SURFEX, as well as potential evapotranspiration, precipitation, and temperature from SAFRAN. The use of these SURFEX 8 km resolution fluxes made necessary the recalibration of the~~
25     ~~hydrogeological models included in the platform.~~A wide range of gauging stations and piezometers were selected in order to perform the evaluation of the simulated piezometric heads, river flows and karstic spring flows. This evaluation allows to identify extreme events such as groundwater floods or droughts over a long-term period. In this paper, a detailed description of the AquiFR platform and its components is presented in section 2. Section 3 provides information on the regional models, their calibration and the statistical criteria used to evaluate their performance. Section 4 presents the assessment of the long-
30     term simulation based on comparison with observations of river flows, karstic spring flows and piezometric heads. The results are then discussed in section 5, prior to the conclusions.

4

## 2 The AquiFR Hydrometeorological Modelling Platform

The AquiFR hydrometeorological modelling platform represents the main hydrological processes occurring within the watersheds from precipitations to groundwater flows as shown in Figure 1. In its present form, the AquiFR system includes 3 hydrogeological modelling software covering 11 sedimentary aquifers and 23 karstic systems: the EauDyssée hydro(geo)logical numerical platform (Saleh et al., 2013), the MARTHE groundwater flow software (Thiéry, 2015a) and the EROS lumped model software used for karstic systems (Thiéry, 2018a). These software are embedded in an application developed with the OpenPALM coupling system (Duchaine et al., 2015). All these models cover an area of about 149 000 km$^2$ and contain up to 10 overlaid aquifer layers.

AquiFR accounts for spatial heterogeneity by using different spatial scales. The SAFRAN meteorological analysis (Quintana-Seguí et al., 2008; Vidal et al., 2010), available over the French metropolitan area at an 8 km resolution, allow to supply the meteorological variables to the SURFEX land surface model (Masson et al., 2013) which evaluates the water balance over the French metropolitan area. ~~The atmospheric fields from SAFRAN and the estimation of the surface water budget fluxes by SURFEX are provided on an 8 km resolution grid. The~~ SAFRAN ~~meteorological analysis (Quintana Segui et al., 2008)~~ provides hourly precipitation (rainfall and snowfall), temperature, relative air humidity, wind speed and downward radiations. ~~The~~ SURFEX ~~land surface model (Masson et al., 2013)~~ uses these atmospheric variables to solve the energy and surface water budget at the land-atmosphere interface at a 5-minute time step. SURFEX estimates the spatial partition of the flow between surface runoff and groundwater recharge on the SAFRAN 8 km resolution grid. It accounts for different soil and vegetation types and uses a diffusion scheme to represent the transfer of heat and water through the soil. The soil in SURFEX is represented by a multilayer approach. Its depth varies according to the vegetation (in France from 0.2 to 3 m) and is partly accessible to plant roots. Deep soil infiltration constitutes groundwater recharge flux. Surface runoff can occur according to saturation excess or infiltration excess.

The simulation of the watersheds depends on its hydrogeologic characteristics. For sedimentary basins, these two fluxes are transferred to the MARTHE (Thiéry, 2015a) or EauDyssée (Saleh et al., 2013) groundwater models. These models simulate the transfer to the unsaturated zone, groundwater flows within and between the aquifer layers, the routing of surface runoff to and within rivers, and river-aquifer exchanges. They also account for the numerous groundwater abstractions within the river basins. The temporal resolution is daily and the spatial resolution varies from 100 m to a maximum of 8000 m. The depth of the deepest aquifer layer can reach locally about 1000 meters. It must be stressed that the hydrogeological models could have been classically fed with the SAFRAN analysis precipitation, potential evapotranspiration, and temperature data using their own water balance calculation. However, the combined use of SURFEX and SAFRAN provides a consistent set of hydro-meteorological data over an 8 km resolution grid over France, including groundwater recharge and surface runoff from SURFEX, as well as potential evapotranspiration, precipitation, and temperature from SAFRAN. The use of these SURFEX 8 km resolution fluxes made necessary the recalibration of the hydrogeological models included in the platform.

Code de champ modifié

Code de champ modifié

5

Karstic aquifer systems are simulated through a conceptual reservoir modelling approach using the EROS software (Thiéry, 2018a). Each karstic system is represented by a lumped reservoir model solved at a daily time scale. Conceptual approaches are preferred for simulating karstic systems. Indeed, their heterogeneities make it difficult to use a physically based approach. EROS uses the daily precipitation, snow, temperature and potential evapotranspiration provided by SAFRAN to compute
5   karstic spring flows.

Technically, the AquiFR hydrogeological modelling platform was developed using the OpenPALM coupling system (Buis et al., 2005; Duchaine et al., 2015). OpenPALM allows the easy integration of high-performance computing applications in a flexible and scalable way. It was originally designed for oceanographic data assimilation algorithms, but its application domain extends to multiple scientific applications. In the framework of OpenPALM, applications are split into elementary components
10   that can exchange data. The AquiFR platform is an OpenPALM application that currently gathers 5 components. Figure 2 shows the linkage between these components and the workflow of an AquiFR run. In this version 1.2 of AquiFR, no feedback from groundwater to the soil of SURFEX is taken into account. Therefore, a preliminary step illustrated by Figure 2a is carried out in order to estimate groundwater recharge and surface runoff with SURFEX accounting for the atmospheric forcing from SAFRAN prior to an OpenPALM run. This preliminary step gives access to 60 years of daily groundwater recharge and surface
15   runoff on a regular 8 km resolution over all the French metropolitan area.

These water fluxes are then accessible by the OpenPALM application that includes the three hydrogeological modelling components, the pre-processing component, and the post-processing component as shown in Figure 2b. b. All these components exchange data during the parallel execution of a single OpenPALM run. At each daily time step, a first pre-processing component retrieves both the atmospheric forcing and the SURFEX groundwater recharge and surface runoff at the beginning
20   of the time step. Then, the EauDyssée, MARTHE and EROS modelling software compute the evolution of the simulated hydrogeological variables for the current time step for each groundwater model independently. A last post-processing component synchronizes the simulation (it waits until all the models have ended their computations for the current time step) and collects the individual outputs of each model to write to comprehensive outputs for the entire domain. At last, a signal is sent by the post-processing component in order to allow the platform to compute the next time step. The use of OpenPALM
25   allows running each instance of the models in parallel on several processors. The 60-year simulation presented in this study needs approximately 1.5 days of computation time on a high-performance computer. The following subsections present a brief description of the components integrated within the OpenPALM application in AquiFR.

## 2.1 The SAFRAN meteorological analysis

The SAFRAN meteorological analysis is a mesoscale atmospheric analysis system for surface variables. It provides
30   meteorological forcing data over France on an 8 by 8 km grid at the hourly time step using observed data and atmospheric simulations. Originally intended for mountainous areas, it was later extended to cover France (Quintana-Seguí et al., 2008). SAFRAN estimates eight variables: rainfall, snowfall, incoming solar and atmospheric radiations, cloudiness, air temperature and relative humidity 2 m above ground, and wind speed 10 m. Potential evapotranspiration can also be computed from these

6

atmospheric variables. SAFRAN is based on climatic zones where the atmospheric variables only vary according to the topography. More than 600 homogeneous climate zones are defined over France. The average area for each zone is about 1000 $km^2$ so that each one contains one surface meteorological station and at least two rain gauges. SAFRAN uses all the observations available to estimate each atmospheric variable except for radiations. For each variable, values are ~~assign~~assigned

5 to given altitudes using an optimal interpolation method. The analyses are computed every 6 hours and an interpolation is made to an hourly time step. Radiation fluxes are computed using a radiative transfer scheme. The daily precipitation rates are estimated using a wide range of daily rain gauges and converted to hourly data using the evolution of the air relative humidity. The vertical profiles of the atmospheric parameters are then computed in each climatic zone and the values are spatially interpolated over the 8 km grid as a function of the altitude within each climatic zone. Further details on the SAFRAN analysis

10 system can be found in Quintana-Seguí et al. (2008) and Vidal et al. (2010).

**2.2 The SURFEX modelling platform**

SURFEX is a modelling platform aiming at simulating the water and energy fluxes at the interface between the surface and the atmosphere (Masson et al., 2013). SURFEX is built to be coupled to forecast and climate models. It includes databases, interpolation schemes and several physical options that allow its use over different spatial and temporal scales. SURFEX

15 gathers several physical schemes in a single platform, allowing the simulation of the urban surfaces and the main components of the water cycle: sea and ocean, lake, vegetation and soil.

Land surface processes are taken into account using the Interaction between Soil Biosphere and Atmosphere (ISBA) (Noilhan and Planton, 1989) land surface scheme. ISBA uses a short list of parameters depending on vegetation and soil types. The temporal evolution of the soil water and energy budget is computed using a multilayer soil scheme based on the explicit

20 resolution of the one-dimension Fourier law as well as the mixed form of the Richards equation (Boone et al., 2000; Decharme et al., 2013). Surface/groundwater capillary exchanges can be explicitly taken into account (Vergnes et al., 2014) as well as the vertical root profile in the soil (Braud et al., 2005).

In the present study, no bidirectional coupling between the soil of SURFEX and the aquifers is ~~accounting~~accounted for. Thus, a one-way coupling from the soil of SURFEX to the aquifer is taken into account in order to provide groundwater recharge

25 and surface runoff to the AquiFR platform. The soil column thickness represented in each 8 km resolution grid cell varies from 0.20 to 3 meters according to the land cover. It corresponds mostly to the root zone layer (Decharme et al., 2013). Thus, the recharge provided by SURFEX is the vertical flux leaving the bottom of the soil column of each grid cell. Further details on ISBA can be found in Decharme et al. (2013).

**2.3 The EauDyssée groundwater modelling software**

30 The EauDyssée modelling platform gathers numerical modules representing several hydrological processes, the most important being the aquifer module based on the Simulation des Aquifères Multicouches SAM (multilayer aquifer system) regional

7

groundwater modelling software (Ledoux et al., 1989) and the river routing scheme based on the Routing Application for Parallel computatIon of Discharge (RAPID) model (David et al., 2011).

SAM computes the evolution of the piezometric heads of multilayer aquifers using a finite difference numerical scheme to solve the groundwater diffusivity equation with a square grid discretization. Groundwater horizontal flows are two-dimensionals and vertical flows through aquitards are taken into account. Therefore, unconfined and confined aquifers can be represented. SAM was successfully used to predict groundwater and surface water flows in different basins of various scales and hydrogeological contexts: the Seine basin (Viennot, 2009), the Somme basin (Habets et al., 2010), the Loire basin (Monteil et al., 2010) or the Rhine basin (Thierion et al., 2012; Vergnes and Habets, 2018).

The RAPID software is a river routing model based on the Muskingum routing scheme (David et al., 2011). It can be coupled to groundwater and land surface models. Volumes and river flows are computed along a river network discretized into square grid-cells to ease the simulation of the exchanges with groundwater. River-groundwater exchanges are taken into account in both directions.

**2.4 The MARTHE groundwater modelling software**

The Modélisation d'Aquifères par un maillage Rectangulaire en régime Transitoire pour le calcul Hydrodynamique des Ecoulements MARTHE (Modelling Aquifers with Rectangular cells, Transport and Hydrodynamics) computer code is the hydrogeological modelling software from the French Geological Survey (BRGM) (Thiéry, 2015a, 2015c, 2015b). MARTHE embeds single to multilayer aquifers and hydrographic networks. It is designed for 2D or 3D modelling of flows and mass transfers in aquifer systems, including climatic, human influences and possible geochemical reactions. Groundwater flow is computed by a 3D finite volume approach to solve the hydrodynamic equation based on the Darcy's law and mass conservation, using irregular rectangular grids, with the possibility of nested grids. River flows are simulated based on a kinematic wave approach that is fully coupled to groundwater flow. Groundwater-river exchanges are taken into account in both directions.

Other options are available and can be integrated to the simulation: mass transfer for pollutants in water, temperature effects, impact of salinity, degradation of pollutants, transfers in the unsaturated zone and geochemical reactions.

This software is widely used for groundwater resources management in France: for example in the Somme River basin (Amraoui et al., 2014), in the Poitou-Charentes region (Douez, 2015), in the Basse-Normandie region (Croiset et al., 2013) or in the Aquitaine sedimentary basin (Saltel et al., 2016). It is also used in other environmental fields such as pollutant infiltration in unsaturated zones (Herbst et al., 2005; Thiéry et al., 2018) or for the simulation of pollution plume coming from a contaminated area.

**2.5 The EROS software**

The Ensemble de Rivières Organisés en Sous-bassins EROS (set of rivers organized in sub-basins) numerical code is a distributed reservoir modelling software dedicated to large river systems (Thiéry, 2018a; Thiéry and Moutzopoulos, 1992). It

8

allows the simulation of river flow or karstic spring flow and piezometric head 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 local river discharge 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

5 simplified physical laws (Thiéry, 2015d). 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 spatially physically based model. In the framework of AquiFR, this software was adapted in order to simulate in a single instance 23 karstic systems as independent sub-models (Thiéry, 2018b). It is not connected to SURFEX but directly to SAFRAN as described in Figure 1.

10 **3 Methodology**

**3.1 The regional models implemented in the AquiFR platform**

AquiFR aims at covering all groundwater resources in France. Figure 3 shows the main aquifers covering France classified by geological type as defined in the French hydrogeological reference system BDLISA (https://bdlisa.eaufrance.fr/). The current version of AquiFR gathers 13 spatially distributed models corresponding to regional single or multilayer aquifers (Table 1 and

15 Figure 4).

Some regions are simulated by two spatialized models (Figure 4): the Somme and the Basse-Normandie basins are covered by MARTHE and EauDyssée models, and the chalk aquifer of the Seine basin is covered by both the EauDyssée Seine model and four EauDyssée sub-models (Marne-Loing, Marne-Oise, Seine-Eure, and Seine-Oise regional models, see Figure 4). This allows a multi-model approach, which can be useful for forecast and climate change impact studies. For these regions, the

20 results presented in this paper correspond to the models that were considered as the best calibrated with the SURFEX fluxes. It corresponds to the four EauDyssée sub-models over the Seine basin and the Somme and Basse Normandie MARTHE models. Figure 4 also shows the 23 karstic systems (median catchment area of 99 km$^2$) simulated by EROS (Thiéry, 2018b) as well as the hard rock aquifer in Britany that will be simulated using a hillslope model (Courtois, 2018; Marçais et al., 2017) and integrated in the near future.

25 Groundwater withdrawals are integrated as time-dependent boudary conditions in the spatially distributed models. On annual average and with respect to the total surface area of the simulated domain, it corresponds to about 16 mm/year (2.4 billion of cubic meters per year) distributed in more than 16 000 grid cells. Data on groundwater pumping are provided by the regional water agencies based on tax reporting. Pumping concerns drinking water, irrigation, and industrial use. The quality of the dataset as well as its temporal extension varied for each regional modelling, although the later does not exceed 20 years. Further

30 details on regional models can be found in the references listed in Table 1. To extend the pumping estimation to the 1958-2018 period, a monthly mean annual cycle is used for the years without data.defined for the years without data. This choice is linked to the lack of knowledge about past pumping. However, we do know that there have been antagonistic developments between

9

irrigation and industrial pumping. Irrigation has increased in accordance with the irrigated areas while it varied greatly depending on the climate. Industrial pumping was dominant in the past but has considerably decreased during the past decades (Service de l'observation et des statistiques, 2016).

Each regional model uses its own river network at its own resolution. Most of the simulated domains encompass the entire
5 river basins corresponding to the simulated rivers. Only the Alsace and the Poitou-Charentes basins are partially represented. Therefore, they need to prescribe time-dependent boundary conditions at the upstream of some rivers based on river flow observations. If the observed data ~~don't~~do not cover the full period, the missing values are filled by the daily mean annual observed river flow. In the near future, the advantage to have the atmospheric forcing and surface fluxes over the entire domain will be used to estimate the upstream flow based either on a lumped-parameter rainfall-runoff model integrated in the
10 MARTHE computer code or by the RAPID river routing model using a fine scale river network covering all France.

**3.2 Calibration of the hydrogeological models**

The original hydrogeological regional models were developed independently most often based on stakeholder requests. The water budgets were usually computed using less physical methods and atmospheric local data (precipitation, temperature and potential evapotranspiration) that differ from the physically based approach using SURFEX and the SAFRAN analysis. As a
15 result, in order to be consistent with the estimation of the groundwater recharge estimated by SAFRAN-SURFEX, most of the regional models were recalibrated based on ~~the newfluxes~~new fluxes (Habets et al., 2017). This recalibration effort was not undertaken for the Alsace and Loire models since both of them will be soon updated and then recalibrated.

Periods of recalibration were the same as those initially used to develop and calibrate each model (see references in Table 1) in order to facilitate the comparison between the recalibrated models and the initial models. Hydrodynamic parameters,
20 including hydraulic conductivities and specific yields, were modified based on hydrogeological expertise in order to obtain the best fit between observations and simulations. The calibration was made only on the piezometric heads, except for the MARTHE Somme model for which piezometric heads and riverflows were accounted for, and for the karstic systems with karst spring flows only. All the models were recalibrated using the same statistical criteria. A comparison between the initial water budget of the models and the SURFEX fluxes was performed as a first step to estimate the need for recalibration of each
25 model.

Some models, such as the Seine EauDyssée model, were not recalibrated since they perform equally well with the use of the SURFEX fluxes (see Table 1). In contrast, the MARTHE Somme river basin model was characterized by an excess of surface runoff in the north and a deficit in the south. In order to compensate for this imbalance, the total runoff provided by SURFEX was split into surface runoff and groundwater recharge using the original water balance scheme of MARTHE. This water
30 balance scheme is based on a reservoir for which parameters are calibrated in order to compute the main components of the surface water budget (Thiéry, 2014). Only one reservoir was used, enabling to modify the partition of the total runoff and to account for a delay on the groundwater recharge in order to mimic the impact of the deep unsaturated zone. It improved the simulation of the river flows using the SURFEX total runoff. Once the new partition was estimated, the aquifer permeability

10

was recalibrated. The Somme basin is the only one for which only the total runoff from SURFEX was used. For the other basins, the estimation by SURFEX of the partition of the water fluxes between surface runoff and groundwater recharge was used. Overall, the performance of the models are similar with the original water balance fluxes and the ones simulated by SAFRAN-SURFEX, although locally, they may be better or otherwise degraded.

5  For the karst system software EROS, the models were calibrated based on the SAFRAN atmospheric analysis by using an optimization of the statistical comparison between observed and simulated daily river flows.

More information about the calibration is given in Habets et al. (2017).

### 3.3 Evaluation criteria of the 60-year simulation

Statistical criteria are used to evaluate the long-term simulation. The bias allows evaluating the ~~relative~~absolute mean deviation

10  between the observation and the simulation. It is calculated as follows:

$$BIAS = \frac{1}{n} \sum_{i=1}^{n} \left( X_{obs} X_{sim}(t) - X_{sim obs}(t) \right),$$  (1)

with $n$ the number of observed values, $X_{obs}(t)$ and $X_{sim}(t)$ the observed and simulated values respectively at time t. BIAS has the same unit as $X_{obs}(t)$ and $X_{sim}(t)$. The perfect value is 0 while negative values correspond to underestimation and positive values correspond to overestimation.

15  The Root Mean Square Error (RMSE) allows estimating the differences between the observed and simulated values. It is often used to compare observed and simulated piezometric heads. However, the computation of the RMSE is strongly affected by the biases. Therefore, we computed a RMSE bias-excluded in order to better assess the simulation in terms of amplitude and synchronization. Moreover, this RMSE bias-excluded is normed with respect to the observed standard deviation for each observation in order to account for the differences of variability between the numerous wells to help spatial comparison or

20  aggregation. This normed RMSE bias-excluded (NRMSE_BE) is ~~expresed~~expressed as follows:

$$NRMSE\_BE = \frac{1}{\sigma_{obs}} \sqrt{\frac{\sum_{t=1}^{n}[(X_{sim}(t) - \overline{X_{sim}}) - (X_{obs}(t) - \overline{X_{obs}})]^2}{n}}$$  (2)

with $\overline{X_{sim}}$ the temporal mean of simulated values over the considered period and $\sigma_{obs}$ the observed standard deviation. The NRMSE_BE criterion is always positive and starts from 0 for a perfect simulation of the observed amplitudes. A NRMSE_BE

25  lower than 0.8 can be considered as a reasonable estimation of the temporal evolution of the observed water table.

The Nash-Sutcliffe model Efficiency coefficient NSE (Nash and Sutcliffe, 1970) measures the variance between the observed and simulated values. It is often applied to compare observed and simulated river flows but can be used for other variables. Its ~~sensitivity to high-frequency fluctuations makes its~~ use for comparing groundwater levels is less obvious regarding its strong sensitivity to the biases between observation and simulation. It is equal to 1 when the model fits perfectly the observations. A

11

NSE above 0.7 is generally accepted as a good estimate of the signal dynamic, however depending on the hydrogeological and climate context of the basin. A negative NSE means that the mean observed signal is a better predictor than the model. The NSE is calculated as follows:

$$NSE = 1 - \frac{\sum_{t=1}^{n}(X_{obs}(t)-X_{sim}(t))^2}{\sum_{t=1}^{n}(X_{obs}(t)-\overline{X_{obs}})^2},$$

(3)

with $\overline{X_{obs}}$ the temporal mean of observed values over the considered period.

The annual discharge ratio criterion helps to compare the mean simulated and observed river flows as follows:

$$Ratio = \frac{\overline{Q_{sim}}}{\overline{Q_{obs}}},$$

(4)

with $\overline{Q_{sim}}$ and $\overline{Q_{obs}}$ the mean siumlated and observed river flows respectively.

One way to evaluate the ability of the simulation to capture extreme events is to use the Standardized Piezometric Level Index (SPLI). The SPLI is an indicator used to compare groundwater level time series and to characterize the severity of extreme events such as long dry period or groundwater overflows (Seguin, 2015). It is currently used in France for the Monthly Hydrological Survey (MHS) (Office International de l'Eau, 2019). This MHS provides monthly information to the policymakers and the public on the hydrological state of groundwater. Assessing the ability of the AquiFR modelling platform to reproduce this indicator is important since the main objective of this platform is to predict such extreme events in short-to-long terms hydrogeological forecasts for groundwater management. The SPLI indicator is based on the same principles as the Standardised Precipitation Index (SPI) defined by McKee et al. (1993) to characterize meteorological drought at several time scales. First, monthly mean time series are computed from time series of piezometric heads. Then, twelve monthly time series (January to December) are constituted over the N years of the time series period. For each time series of N monthly values, a non-parametric kernel density estimator allows estimating the best probability density function fitting the histogram of monthly values. At last, for each month from January to December, a projection over the standardised normal distribution using a quantile-quantile projection allows deducing the SPLI for each value of the monthly mean time series of piezometric heads. The SPLI values most often range from -3 (extremely low groundwater levels corresponding to a return period of 740 years) to +3 (extremely high groundwater levels). The SPLI allows representing wetter and drier periods in a similar way all over the simulated domain.

### 3.4 Dataset and model setup

The long-term simulation was carried out over a 60-year period from the August 1, 1958 to July 31, 2018 at a daily time step using the SAFRAN meteorological analyses. State variables from August 1, 2013 were chosen from a first simulation over the
5 1958-2018 period in order to initialize the simulation in August 1, 1958. The 2013 year was chosen as the best proxy of the 1958 year by analysing time series of long-term observed groundwater levels having data since 1958. The mean precipitations corresponding to the simulated domain of Figure 4 and averaged over the 60-year period is equal to 743 mm/year. The SURFEX computes then the surface water budget is then computed by SURFEX from the SAFRAN outputs. The mean simulated total runoff is partitioned between 163 mm/year of groundwater recharge and 60.5 mm/year of surface runoff. Thus,
10 the groundwater abstractions represent about 25% of the groundwater recharge.

The evaluation of this simulation is made using the numerous in situ datasets available in France. Observed piezometric heads are available in the "Accès aux Données sur les Eaux Souterraines" (ADES) database (http://www.ades.eaufrance.fr/). 639 observation boreholes covering the AquiFR domain corresponding to both confined and unconfined aquifers, and with at least 10 years of continuous time series, were selected. Figure 5 shows the temporal evolution of the number of daily measurements
15 along the 60-year period. Starting in 1958, only a few measures measurements are available. Starting from 1970, the number of wells increases slowly to reach about 100 in 1990. Then the number of daily measurements quickly increases to reach more than 450 in 2010. This number remains stable then, except for the last year (2018) with a decrease because the datasets were not yet fully transmitted available. In situ daily river flow observations at 362 gauging stations were also selected for evaluating the daily simulated river flows from the Hydro database (http://hydro.eaufrance.fr/).

20 ### 4 Results

#### 4.1 Piezometric head

Figure 6a shows the spatial distribution of the bias for the 639 observed piezometers. A positive value means that the simulation overestimates the mean observed piezometric head while a negative value means the opposite. The north of the Loire river basin, corresponding to the Beauce region, shows a significant underestimation of the mean observed groundwater level.
25 Elsewhere, no significant patterns appear. Figure 6b summarizes these results with the accumulated cumulative distribution of the absolute biases for all the piezometers. 42% and 60% of the absolute biases are lower than 2 m and 4 m, respectively.

Figure 7 shows the spatial and accumulated cumulative distribution of the NRMSE_BE. 16%, 39 and 62% of the wells obtain a value lower than 0.6, 0.8 and 1 respectively, while 88% have a value lower than 2. Some piezometers that were affected by important biases in Figure 6a exhibit however good NRMSE_BE, in particular over the Loire river basin and in the northern
30 Poitou-Charentes region, meaning that the temporal evolution is well simulated.

13

**Code de champ modifié**

Five examples of simulated and observed daily evolution of piezometric heads are shown in Figure 8. These piezometers are encircled in Figure 3 and statistical scores are available in Table 2. They were chosen to characterize different hydrogeological contexts. The first piezometer named Omiécourt is located in the Chalk aquifer of the Somme River basin. The temporal evolution of the groundwater level is characterized by multiyear cycles well captured by the model. However, the simulation

5  displays annual cycles that are not observed. It explains why the NRMSE_BE is equal to 0.93 while the bias is equal to -0.86 m. The two piezometers named Ruffec and Le Bec ~~hellouin~~Hellouin correspond to limestone aquifers and are located in the Poitou-Charente region and near the coast of the English Channel respectively. The first one is characterized by large annual cycles with wide amplitudes. The model is able to reproduce these annual cycles (correlation of 0.82) but with an underestimation of the peaks leading to a negative bias of -1.44 m. The Le Bec ~~hellouin~~Hellouin piezometer is characterized

10  by both multiyear and annual cycles that are captured by the model, although between 2005 and 2015 the simulated groundwater level is underestimated with respect to the observation. The piezometer named Farceaux is located in a chalk aquifer in the Seine River basin. It is characterized by a systematic bias of about -8.3 m. Otherwise, the multiyear and annual cycles are well reproduced by the model, which is confirmed by the NRMSE_BE equal to 0.52. The last example corresponds to a piezometer for which the model cannot reproduce the strong seasonal decrease of level occurring each year. Such

15  behaviours in the observation are likely due to groundwater withdrawals that are not well prescribed in the model near this well.

**4.2 The Standardised Piezometric Level Index**

The SPLI is categorized into seven classes summarized in Table 3 from the driest to the wettest conditions. According to Seguin and Klinka (2016), a set of piezometers were chosen in order to compute the SPLI indicator in this MHS with the

20  following characteristics: a continuous time series with at least 15 years and no impact of pumping wells. Among the 639 selected observation wells in Figure 6 and Figure 7, 103 contribute to the MHS.

Figure 9 ~~and Figure 10~~ show~~s~~ the spatial distribution of the NSE ~~and the correlation respectively~~ computed between the observed and simulated SPLI indicator for the 103 selected piezometers. It assesses the ability of the model to reproduce the SPLI indicator in different locations. 20% of the NSE are greater than 0.7 and 56% greater than 0.5, while 12% are lower than

25  zero. ~~In parallel, 65% of the correlation.~~ Figure 10~~Figure 11~~ focuses on five examples of observed and simulated temporal evolutions of the SPLI indicator. These piezometers correspond to the ones shown in Figure 8 and are part of the selected piezometers used for the MSH. Table 2 presents the related NSE and correlation scores. The NSE computed for these SPLI time series are all greater than or equal to 0.6 except for the Bourdet piezometer characterized by a NSE equal to -0.51. This lower score may be due to a lack in the model input, such as the underestimation of withdrawal data in its vicinity.

30  The SPLI, as a frequency indicator, does not account for the potential biases between the observed and simulated groundwater levels. This is the reason why the systematic biases found in Figure 8 do not appear in the monthly SPLI comparisons in Figure 10~~Figure 11~~, in particular for the Farceaux piezometer. The right part of Figure 10~~Figure 11~~ shows the histograms of the simulated (in red) and observed (in blue) monthly SPLI values for each classes of Table 3. The histograms are similar for both

14

the observed and simulated SPLI at the Ruffec and Le Bec Hellouin piezometers. The occurrences of the wetter conditions are well reproduced for the Omiécourt piezometer but the model tends to underestimate the number of moderately dry conditions (26% and 18% events for the simulation and the observation respectively). For the Farceaux piezometer, the model underestimates the occurrences of the driest events and overestimates the occurrences of the wetter events. Despite the poor scores obtained for the Bourdet piezometer, in particular for the correlations, the distribution of all the monthly SPLI values with respect to the classes of Table 3 is similar for both the observation and the simulation.

The MHS published every month in France for water resources management includes the calculation of the SPLI. As an example, Figure 11Figure 12a shows the observed SPLI values calculated for the 103 selected piezometers for June 2016. We chose this specific month since it follows large precipitation events that leads to floods in the Seine and Loire basins (Philip et al., 2018). Figure 11Figure 12b shows the simulated SPLI values computed for this specific month. The model reproduces the overall pattern of normal and wet conditions but tends to overestimate the importance of the moderately wet conditions: 19% (29%) of the simulated (observed) piezometers are in normal conditions, 46% (31%) are in moderately wet conditions, 16% (13%) are in wet conditions, and 16% (18%) are in extremely wet conditions. The background map of Figure 11Figure 12b shows the SPLI computed in the cells of the whole outcropping domain. These SPLI values were computed with respect to a 30 years reference period from 1981 to 2010, which might lead to differences between the simulated SPLI map and observed values. This map shows a large area of extremely wet conditions located in the south of the Loire River, which refers to the extreme event episode of rainfall from the end of May 2016.

### 4.3 River flow and karstic spring flow

The 23 karstic systems simulated by the EROS model are evaluated against gauging stations located at the outlet of the corresponding karstic systems. All these gauging stations were also used to calibrate the model (Thiéry, 2018b). Figure 12Figure 13 shows the comparison on the 60-year period of the observed and simulated monthly river flows for four examples of karstic systems located in Figure 3. There is a tight agreement between the observation and the simulation. The NSE of the square root of the daily karstic spring flows are given for each example. Using the square root of the daily karstic spring flow allows attenuating the importance of the flood peaks characterizing these small karstic systems and enables to better evaluate the karstic spring flow simulation. Such transformation is necessary because of the excessive sensitivity of the NSE criteria to extreme values in a river flow time series (Legates and McCabe Jr, 1999). Other statistical scores using less assumptions on the underlying data distribution, such as the non-parametric variant of the KlungeKling-Gupta efficiency (Pool et al., 2018), could be used to reduce the sensitivity to the extremes. For these four examples, all the NSE are greater than 0.7.

The distributed models included in the AquiFR modelling platform integrate river networks and the simulation of river flows on each river grid cell. 362 gauging stations were selected to evaluate the simulated river flows. Figure 13Figure 14 shows the comparison between the observed and simulated daily river discharges for four of them: the Charente River, the Somme River, the Seine River and the Loire River. The statistical scores are given for each of them. The locations of these gauging stations areis shown in Figure 3. The Charente River and the Somme River correspond to watersheds with areas lower than 10 000 km$^2$

15

and the Loire River and the Seine Rivers to regional watersheds with areas greater than 80 000 km$^2$. ~~The statistical scores are~~ ~~summarized in Table 4.~~ The simulated river flow of the Charente River underestimates the peak floods, which leads to a ratio of 0.81. The river flows of the Seine River and the Loire River are well reproduced with daily NSE of 0.86 and 0.9 respectively. The Somme river flow is also well reproduced with a NSE equal to 0.69 due to a lower ratio of 0.92.

5    Figure 14~~Figure 15~~a shows the spatial distribution of NSE for the 362 gauging stations (circles) and the 23 karstic systems (stars). The NSE calculated for the karstic systems correspond to the square root of the daily karstic spring flows in order to attenuate the importance of the flood peaks characterizing these small karstic systems, as explained above. The corresponding ~~accumulated~~cumulative distributions are shown for the 23 karstic systems and for the 362 gauging stations in Figure 14~~Figure~~ ~~15~~b and Figure 14~~Figure 15~~c respectively. 96% of the NSE using the square root of the daily karstic spring flows are greater

10   than 0.7. Regarding the results of Figure 14~~Figure 15~~c, for rivers in continuous aquifers, 34% of the NSE are greater than 0.7. Moreover 63% of these NSE are greater than 0.5 while 18% are ~~negatives~~negative.


## 5 Discussion

The results obtained on the 1958-2018 period demonstrates the feasibility and the utility to gather several regional models developed separately in different research institutes into a single numerical tool to provide simulations of the water resource

15   at a daily time step at the national scale. It was shown that the AquiFR platform is able to reproduce the evolution of the observed hydrological variables, including piezometric levels and river flows, with reasonable statistical scores. Some regions are nevertheless better reproduced than others. For example, the Loire region exhibit poor ~~NRMSE_BE~~bias scores in Figure 6. Part of the error is linked to the estimation of the groundwater recharge by SURFEX. Indeed, the regional groundwater models were developed and calibrated independently using various methods and data to compute groundwater recharge. As a

20   consequence, the development of AquiFR reinforced the need to calibrate these models based on the SURFEX forcing fields. This work was accomplished for most of the models included in AquiFR except for some of them, including the Loire River basin (Habets et al., 2017). The use of an inverse model as the one proposed by Hassane Maina et al. (2017) should help to improve such calibration.

Starting from scratch with an integrated method could have prevented the burden of maintaining each model separately and

25   handling the different outputs of the models. Such method was applied by Kollet et al. (2018) over the North Rhine-Westphalia domain using Parflow-CLM by integrating all the physical processes related to groundwater/surface water into a single numerical tool. De Lange et al. (2014) used an approach closer to AquiFR by coupling five physically different models for different water domains with different concepts, different temporal and spatial scales, and different national and regional databases altogether embedded into the National Hydrological Instrument NHI. These two models are used for integrated water

30   management and policymaking issues. The areas covered by these models are 22 500 km$^2$ and 41 500 km$^2$ respectively. According to Figure 3, the BDLISA database references regional sedimentary aquifer systems (in green) and alluvial aquifers (in blue) in France both covering an area of about 355 000 km$^2$. Reaching the complexity of a fine-tuned regional model,

16

including the geometrical, geological and physical contexts, in a single integrated numerical tool covering such large territory would be time consuming to build, to calibrate, and to evaluate. It would also demand big resources of computational power. An attempt was made by Vergnes et al. (2012) to simulate a single integrated model groundwater over the French metropolitan territory. Even though the results obtained were good enough to be used for large-scale climate applications, it was not fitted
5 for operational water management: only one layer was defined at the coarse resolution of about 10 km, no pumping were defined, no calibration was achieved and very simplified parameterizations were used.

To overcome this difficulty, the choice was made for AquiFR to bring together different models developed independently. Currently, the area covered by the platform is equal to around 133 000 km$^2$ with a number of layers that can reach 10 layers for some models, and these numbers will increase in the future with the addition of new models. This multi-model approach
10 allows to promote the share of knowledge in hydrogeology and to gather the competencies accumulated in the different research institutes involved in AquiFR in water resource modelling. Moreover, thanks to the evolutive approach of the OpenPALM coupling software, the platform facilitates the addition of new software and new models.

Results from AquiFR show a global view of the performance of the AquiFR platform but are characterized by uncertainties. These uncertainties are mainly related to the calibration of the models and to the lack of some input data like groundwater
15 abstractions. Indeed, the regional models have been calibrated over a shorter period compared to the long-term simulation, as stated in the section 3.2. Consequently, the calibration process may have not included the extreme climatic conditions encountered on the 1958-2018 period.

Other uncertainties may be related to the choices of the resolution, the databases used, the geometry of the models, and more generally the representation of the physical processes in the hydrogeological software. Some regions are better monitored than
20 others and the global view of the performance of the AquiFR platform is certainly affected by this. Moreover, the chosen method of evaluation based on statistical scores such as NSE could also be improved. Indeed, some authors report that the use of more realistic upper and lower benchmarks for each simulated basins could improve the judgement of model performances with respect to the climate and hydrogeological context of the basins (Pappenberger et al., 2015; Seibert et al., 2018). At last, in order to diminish uncertainties, a long-term calibration effort using a denser observation network could be undertaken to
25 improve the AquiFR performance.

The AquiFR platform can be seen as an improvement of the SIM hydrometeorological tool for operational water management purpose (Habets et al., 2008). These two systems share common points as the SURFEX land surface model or the groundwater component of the EauDyssée platform. However, the SIM tool uses coarse hydrogeological modelling with less aquifer layers or no river loss to the aquifer. It mainly focuses on operational forecasts of river flows and monitoring of soil humidity. The
30 AquiFR platform is intended to focus also on the forecast of groundwater levels for the multilayer aquifers and karstic systems described in Figure 4. To achieve this goal, the SPLI indicator will be calculated to provide forecasts of extreme events. For this purpose, AquiFR is able to produce different representation of this indicator and to compare it with other variables depending on the need of the stakeholders. For example, Figure 15Figure 16 compares the simulated time series of daily mean groundwater recharge, stream-groundwater exchanges budget, monthly mean piezometric head and SPLI averaged over the

17

chalk aquifer of the Somme model. It gives a global view of the past states of groundwater related to climatic extreme events, such as the severe flood of the 2001 year, characterized by groundwater flooding and sustained stream-to-groundwater exchanges (Amraoui and Seguin, 2012).

**6 Conclusions**

5   This study introduces the AquiFR hydrogeological modelling platform aiming to provide at the French national scale short-term to seasonal hydrological forecasts as well as real-time monitoring for daily water management and long-term simulations for climate impact studies. It was developed using a coupling software in order to gather different software and several models covering a set of French multilayer aquifers. Daily surface runoff and groundwater recharge provided by the SURFEX land surface model using the SAFRAN meteorological analysis were used for simulating the daily evolution of groundwater levels

10  and river flows of French regional multilayer aquifers and karstic systems on the 1958-2018 period.

The results confirm the feasibility of gathering independent hydrogeological model developed in different research institutes into the same coupling platform. All these models were initially developed and calibrated on shorter periods with heterogeneous geological and meteorological databases. Some of these models were recalibrated against the SAFRAN-SURFEX fluxes. The evaluation of the 1958-2018 long-term simulation shows a good comparison with the observations

15  available on the same period. It confirms the relevance of using AquiFR as a tool for long-term impact studies. The evaluation of the SPLI indicator also shows that AquiFR could be used in an operational context to monitor future events and be part of the Monthly Hydrological Survey, provided the necessary caution in terms of communication of model uncertainties and performances.

The advantage of this platform lies in its modularity. AquiFR encourages the development of groundwater modelling where it

20  is missing and, more generally, it has the potential to be a valuable tool for many applications in water resource management and research studies, as for instance climate change studies and seasonal forecasts. In the future, more regional models developed with MARTHE or EauDyssée will be included in order to extend the cover of AquiFR, as the Tarn and Garonne aquifer (Figure 4). A new software will be included in order to simulate bedrock aquifers located in Britany (Courtois, 2018). A new modelling method based on a lumped-parameter rainfall-runoff model will be used to provide upstream river flows as

25  boundary conditions for the MARTHE models that required it. Assessment of the seasonal forecast of the groundwater resource is now in progress (Roux, 2018). Since errors in the initial conditions can alter significantly the skill of the forecast, dedicated studies on data assimilation to improve initial state conditions are also done in parallel (Hassane Maina et al., 2017).

18

data, models and software for making this project feasible (BRGM, CERFACS, Météo France, Mines ParisTech, Geosciences Rennes). Special thanks come to Claire Magand, Thimotée Leurent and Bénédicte Augeard for supporting this project with AFB, and Nathalie Dörfliger who supports this project with BRGM. Finally, the authors wish also to thank Eric Martin and Jean-Michel Soubeyroux that were involved at the beginning of this project.

5 **References**

Aeschbach-Hertig, W. and Gleeson, T.: Regional strategies for the accelerating global problem of groundwater depletion, Nat. Geosci., 5(12), 853–861, doi:10.1038/ngeo1617, 2012.

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P. and Meyer, G.: Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland, J. Hydroinformatics, 20(6), 1227–1246, doi:10.2166/hydro.2018.002, 10  2018.

Amraoui, N. and Seguin, J.-J.: Simulation par modèle maillé de l'impact d'épisodes pluvieux millénaux sur les niveaux de la nappe de la craie et les débits du fleuve Somme. Rapport final, BRGM/RP-61864-FR, Orléans., 2012.

Amraoui, N., Castillo, C. and Seguin, J.-J.: Évaluation de l'exploitabilité des ressources en eau souterraine de la nappe de la craie du bassin de la Somme. Rapport final, BRGM/RP-63408-FR, Orléans., 2014.

15 Barlage, M., Tewari, M., Chen, F., Miguez-Macho, G., Yang, Z.-L. and Niu, G.-Y.: The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP, Clim. Change, 129(3), 485–498, doi:10.1007/s10584-014-1308-8, 2015.

Bessière, H., PIcot, J., Picot, G. and Parmentier, M.: Affinement du modèle hydrogéologique de la Craie du Nord-Pas-de-Calais autour des champs captants de la métropole Lilloise. Rapport final, BRGM/RP-63689-FR, Orléans., 2015.

20 Boone, A., Masson, V., Meyers, T. and Noilhan, J.: The Influence of the Inclusion of Soil Freezing on Simulations by a Soil–Vegetation–Atmosphere Transfer Scheme, J. Appl. Meteorol., 39(9), 1544–1569, doi:10.1175/1520-0450(2000)039<1544:TIOTIO>2.0.CO;2, 2000.

Braud, I., Varado, N. and Olioso, A.: Comparison of root water uptake modules using either the surface energy balance or potential transpiration, J. Hydrol., 301(1–4), 267–286, doi:10.1016/j.jhydrol.2004.06.033, 2005.

25 Buis, S., Piacentini, A. and Déclat, D.: PALM: a computational framework for assembling high-performance computing applications, Concurr. Comput. Pract. Exp., 18(2), 231–245, doi:10.1002/cpe.914, 2005.

Courtois, Q.: A hillslope-based aquifer model of free-surface flows in crystalline regions, p. 15, Saint-Malo. [online] Available from: https://www.irisa.fr/sage/jocelyne/CMWR2018/pdf/CMWR2018_paper_429.pdf, 2018.

Coustau, M., Rousset-Regimbeau, F., Thirel, G., Habets, F., Janet, B., Martin, E., de Saint-Aubin, C. and Soubeyroux, J.-M.: 30  Impact of improved meteorological forcing, profile of soil hydraulic conductivity and data assimilation on an operational Hydrological Ensemble Forecast System over France, J. Hydrol., 525, 781–792, doi:10.1016/j.jhydrol.2015.04.022, 2015.

Croiset, N., Wuilleumier, A., Bessière, H., Gresselin, F. and Seguin, J.-J.: Modélisation des aquifères de la plaine de Caen et du bassin de la Dives. Phase 2 : construction et calage du modèle hydrogéologique, BRGM/RP-62648-FR, Orléans., 2013.

19

David, C. H., Habets, F., Maidment, D. R. and Yang, Z.-L.: RAPID applied to the SIM-France model, Hydrol. Process., 25(22), 3412–3425, doi:10.1002/hyp.8070, 2011.

De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Oude Essink, G. H. P., van Walsum, P. E. V., Delsman, J. R., Hunink, J. C., Massop, H. Th. L. and Kroon, T.: An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument, Environ. Model. Softw., 59, 98–108, doi:10.1016/j.envsoft.2014.05.009, 2014.

Decharme, B., Martin, E. and Faroux, S.: Reconciling soil thermal and hydrological lower boundary conditions in land surface models, J. Geophys. Res. Atmospheres, 118(14), 7819–7834, doi:10.1002/jgrd.50631, 2013.

Douez, O.: Actualisation 2008-2011 du modèle maillé des aquifères du Jurassique en Poitou-Charentes. Rapport final, BRGM/RP-64816-FR, Orléans., 2015.

Duchaine, F., Jauré, S., Poitou, D., Quémerais, E., Staffelbach, G., Morel, T. and Gicquel, L.: Analysis of high performance conjugate heat transfer with the OpenPALM coupler, Comput. Sci. Discov., 8(1), 015003, doi:10.1088/1749-4699/8/1/015003, 2015.

Dudley, R. W., Hodgkins, G. A. and Dickinson, J. E.: Forecasting the Probability of Future Groundwater Levels Declining Below Specified Low Thresholds in the Conterminous U.S., JAWRA J. Am. Water Resour. Assoc., 53(6), 1424–1436, doi:10.1111/1752-1688.12582, 2017.

Fan, Y., Li, H. and Miguez-Macho, G.: Global Patterns of Groundwater Table Depth, Science, 339(6122), 940–943, doi:10.1126/science.1229881, 2013.

Guzman, S. M., Paz, J. O. and Tagert, M. L. M.: The Use of NARX Neural Networks to Forecast Daily Groundwater Levels, Water Resour. Manag., 31(5), 1591–1603, doi:10.1007/s11269-017-1598-5, 2017.

Habets, F., Boone, A., Champeaux, J. L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Moigne, P. L., Martin, E., Morel, S., Noilhan, J., Quintana Seguí, P., Rousset-Regimbeau, F. and Viennot, P.: The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, J. Geophys. Res., 113(D6), D06113, doi:10.1029/2007JD008548, 2008.

Habets, F., Gascoin, S., Korkmaz, S., Thiéry, D., Zribi, M., Amraoui, N., Carli, M., Ducharne, A., Leblois, E., Ledoux, E., Martin, E., Noilhan, J., Ottlé, C. and Viennot, P.: Multi-model comparison of a major flood in the groundwater-fed basin of the Somme River (France), Hydrol Earth Syst Sci, 14(1), 99–117, doi:10.5194/hess-14-99-2010, 2010.

Habets, F., Amraoui, N., Caballero, Y., Thiéry, D., Vergnes, J.-P., Morel, T., Le Moigne, P., Roux, N., Dreuzy, J.-R. D., Ackerer, P., Maina, F., Besson, F., Etchevers, P., Regimbeau, F. and Viennot, P.: Plateforme de modélisation hydrogeologique nationale AQUI-FR : Rapport final de 1ère phase. (The national hydrogeological plateform Aqui-FR. Report of the first phase), ONEMA. [online] Available from: https://www.metis.upmc.fr/~aqui-fr/Rapport_fin_phase1_Aqui-FR_VF.pdf, 2017.

Hassane Maina, F., Delay, F. and Ackerer, P.: Estimating initial conditions for groundwater flow modeling using an adaptive inverse method, J. Hydrol., 552, 52–61, doi:10.1016/j.jhydrol.2017.06.041, 2017.

He, X., Stisen, S., Wiese, M. B. and Henriksen, H. J.: Designing a Hydrological Real-Time System for Surface Water and Groundwater in Denmark with Engagement of Stakeholders, Water Resour. Manag., 30(5), 1785–1802, doi:10.1007/s11269-016-1251-8, 2016.

Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C. and Madsen, B.: Methodology for construction, calibration and validation of a national hydrological model for Denmark, J. Hydrol., 280(1–4), 52–71, doi:10.1016/S0022-1694(03)00186-0, 2003.

Herbst, M., Fialkiewicz, W., Chen, T., Pütz, T., Thiéry, D., Mouvet, C., Vachaud, G. and Vereecken, H.: Intercomparison of Flow and Transport Models Applied to Vertical Drainage in Cropped Lysimeters, Vadose Zone J., 4(2), 240–254, doi:10.2136/vzj2004.0070, 2005.

Højberg, A. L., Troldborg, L., Stisen, S., Christensen, B. B. S. and Henriksen, H. J.: Stakeholder driven update and improvement of a national water resources model, Environ. Model. Softw., 40, 202–213, doi:10.1016/j.envsoft.2012.09.010, 2013.

Kollet, S., Gasper, F., Brdar, S., Goergen, K., Hendricks-Franssen, H.-J., Keune, J., Kurtz, W., Küll, V., Pappenberger, F., Poll, S., Trömel, S., Shrestha, P., Simmer, C. and Sulis, M.: Introduction of an Experimental Terrestrial Forecasting/Monitoring System at Regional to Continental Scales Based on the Terrestrial Systems Modeling Platform (v1.1.0), Water, 10(11), 1697, doi:10.3390/w10111697, 2018.

Korkmaz, S.: Modélisation des régimes de crue des systèmes couplés aquifères-rivières, thesis, Paris, ENMP, 1 January. [online] Available from: http://www.theses.fr/2007ENMP1495 (Accessed 30 November 2018), 2007.

Ledoux, E., Girard, G. and De Marsily, G.: Spatially distributed modelling: conceptual approach, coupling surface water and groundwater, in Unsaturated Flow in Hydrologic Modeling: Theory and Practice., vol. 275, pp. 435–454, H.J. Morel-Seytoux., 1989.

Legates, D. R. and McCabe Jr, G. J. M.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, doi:10.1029/1998WR900018, 1999.

Long, D., Longuevergne, L. and Scanlon, B. R.: Global analysis of approaches for deriving total water storage changes from GRACE satellites, Water Resour. Res., 51(4), 2574–2594, doi:10.1002/2014WR016853, 2015.

Longuevergne, L., Scanlon, B. R. and Wilson, C. R.: GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA, Water Resour. Res., 46(11), doi:10.1029/2009WR008564, 2010.

Marçais, J., de Dreuzy, J.-R. and Erhel, J.: Dynamic coupling of subsurface and seepage flows solved within a regularized partition formulation, Adv. Water Resour., 109, 94–105, doi:10.1016/j.advwatres.2017.09.008, 2017.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V. and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6(4), 929–960, doi:10.5194/gmd-6-929-2013, 2013.

Maxwell, R. M., Condon, L. E. and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, Geosci Model Dev, 8(3), 923–937, doi:10.5194/gmd-8-923-2015, 2015.

McKee, T. B., Doesken, N. J. and Kleist, J.: The relationship of drought frequency and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology, vol. 17, pp. 179–183, American Meteorological Society Boston, MA., 1993.

21

Monteil, C., Flipo, N., Poulin, M., Habets, F., Krimissa, M. and Ledoux, E.: Assessing the contribution of the main aquifer units of the Loire basin to river discharge during low flow. [online] Available from: https://hal-mines-paristech.archives-ouvertes.fr/hal-00800812 (Accessed 8 January 2018), 2010.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J.
5    Hydrol., 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

Noilhan, J. and Planton, S.: A Simple Parameterization of Land Surface Processes for Meteorological Models, Mon. Weather Rev., 117(3), 536–549, doi:10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2, 1989.

Noyer, M. L. and Elsass, P.: Modélisation de la salure profonde au droit et en aval du Bassin Potassique - Rapport final, BRGM/RP-54389-FR, Orléans., 2006.

10  Office International de l'Eau: Bulletin de situation hydrologique de janvier 2019, Office International de l'Eau, France. [online] Available from: https://www.eaufrance.fr/publications/bsh/2019-01 (Accessed 7 February 2019), 2019.

Pachocka, M., Mansour, M., Hughes, A. and Ward, R.: Challenges of modelling a complex multi-aquifer groundwater system at a national scale: case study from the UK, pp. 216–220, Colorado School of Mines, Colorado, USA. [online] Available from: http://igwmc.mines.edu/conference/Mod2015/MM15_Proceedings.pdf (Accessed 24 July 2018), 2015.

15  Pappenberger, F., Ramos, M. H., Cloke, H. L., Wetterhall, F., Alfieri, L., Bogner, K., Mueller, A. and Salamon, P.: How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction, J. Hydrol., 522, 697–713, doi:10.1016/j.jhydrol.2015.01.024, 2015.

Philip, S., Kew, S. F., Jan van Oldenborgh, G., Aalbers, E., Vautard, R., Otto, F., Haustein, K., Habets, F. and Singh, R.: Validation of a Rapid Attribution of the May/June 2016 Flood-Inducing Precipitation in France to Climate Change, J.
20  Hydrometeorol., 19(11), 1881–1898, doi:10.1175/JHM-D-18-0074.1, 2018.

Prudhomme, C., Hannaford, J., Harrigan, S., Boorman, D., Knight, J., Bell, V., Jackson, C., Svensson, C., Parry, S., Bachiller-Jareno, N., Davies, H., Davis, R., Mackay, J., McKenzie, A., Rudd, A., Smith, K., Bloomfield, J., Ward, R. and Jenkins, A.: Hydrological Outlook UK: an operational streamflow and groundwater level forecasting system at monthly to seasonal time scales, Hydrol. Sci. J., 62(16), 2753–2768, doi:10.1080/02626667.2017.1395032, 2017.

25  Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L. and Morel, S.: Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France, J. Appl. Meteorol. Climatol., 47(1), 92–107, doi:10.1175/2007JAMC1636.1, 2008.

Rousset, F., Habets, F., Gomez, E., Moigne, P. L., Morel, S., Noilhan, J. and Ledoux, E.: Hydrometeorological modeling of the Seine basin using the SAFRAN-ISBA-MODCOU system, J. Geophys. Res., 109(D14), D14105,
30  doi:10.1029/2003JD004403, 2004.

Roux, N.: The Aqui-FR project : the future operational modeling platform for the main regional multilayer aquifers in France, Saint-Malo. [online] Available from: https://www.irisa.fr/sage/jocelyne/CMWR2018/pdf/CMWR2018_paper_429.pdf, 2018.

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,
35  doi:10.1016/j.jhydrol.2012.10.027, 2013.

22

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 – Module 1 – Année 5 - Convention 2008-2013. Rapport final, BRGM/RP-65039-FR, Orléans., 2016.

5    Seguin, J.-J.: Proposition d'un indicateur piézométrique standardisé pour le Bulletin de Situation Hydrologique "Nappes". Rapport final, BRGM/RP-64147-FR, Orléans., 2015.

Seguin, J.-J. and Klinka, T.: Index Piézométrique Standardisé (IPS) - Bilans et comparaisons avec l'indicateur du BSH - Note technique, BRGM/RP-67251-FR, Orléans., 2016.

Seibert, J., Vis, M. J. P., Lewis, E. and Meerveld, H. J. van: Upper and lower benchmarks in hydrological modelling, Hydrol. Process., 32(8), 1120–1125, doi:10.1002/hyp.11476, 2018.

10   Service de l'observation et des statistiques: Repères. L'eau et les milieux aquatiques. Chiffres clés., Ministère de l'Environnement, de l'Energie et de la Mer., 2016.

Singla, S., Céron, J.-P., Martin, E., Regimbeau, F., Déqué, M., Habets, F. and Vidal, J.-P.: Predictability of soil moisture and river flows over France for the spring season, Hydrol Earth Syst Sci, 16(1), 201–216, doi:10.5194/hess-16-201-2012, 2012.

Stollsteiner, P.: PROJET Explore 2070 - Evaluation de l'impact du changement climatique - Rapport final, BRGM/RP-61483-
15   FR, Orléans., 2012.

Sulis, M., Keune, J., Shrestha, P., Simmer, C. and Kollet, S. J.: Quantifying the Impact of Subsurface-Land Surface Physical Processes on the Predictive Skill of Subseasonal Mesoscale Atmospheric Simulations, J. Geophys. Res. Atmospheres, 123(17), 9131–9151, doi:10.1029/2017JD028187, 2018.

Texas Water Development Board: Groundwater Division | Texas Water Development Board, [online] Available from:
20   https://www.twdb.texas.gov/groundwater/ (Accessed 6 November 2018), 2018.

Thierion, C.: Modélisation hydrogéologique des aquifères du Bajocien - Bathonien dans le secteur des Bocages Normands, Master 2, Université Pierre et Marie Curie, Ecole des Mines de Paris & Ecole Nationale du Génie Rural des Eaux et des Forêts., 2007.

Thierion, C., Longuevergne, L., Habets, F., Ledoux, E., Ackerer, P., Majdalani, S., Leblois, E., Lecluse, S., Martin, E.,
25   Queguiner, S. and Viennot, P.: Assessing the water balance of the Upper Rhine Graben hydrosystem, J. Hydrol., 424–425, 68–83, doi:10.1016/j.jhydrol.2011.12.028, 2012.

Thiéry, D.: Logiciel GARDÉNIA, version 8.2. Guide d'utilisation. Rapport final, BRGM/RP-62797-FR, Orléans., 2014.

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), BRGM/RP-64554-FR, Orléans., 2015a.

30   Thiéry, D.: Modélisation 3D des écoulements en Zone Non Saturée avec le code de calcul MARTHE - version 7.5, BRGM/RP-64495-FR, Orléans., 2015b.

Thiéry, D.: Modélisation 3D du transport de masse et d'énergie avec le code de calcul MARTHE - version 7.5, BRGM/RP-64765-FR, Orléans., 2015c.

Thiéry, D.: Validation du code de calcul GARDÉNIA par modélisations physiques comparatives, BRGM/RP-64500-FR,
35   Orléans., 2015d.

23

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

Thiéry, D.: Modélisation hydrologique globale des débits de 23 sources karstiques avec le logiciel ÉROS. Rapport final, BRGM/RP-67723-FR, Orléans., 2018b.

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, pp. 285–295, ORSTOM Editions. [online] Available from: https://hal-brgm.archives-ouvertes.fr/hal-01061971/document (Accessed 10 January 2018), 1992.

Thiéry, D., Amraoui, N. and Noyer, M.-L.: Modelling flow and heat transfer through unsaturated chalk – Validation with experimental data from the ground surface to the aquifer, J. Hydrol., 556, 660–673, doi:10.1016/j.jhydrol.2017.11.041, 2018.

Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S. and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France – Part 1: Description and validation of the assimilation system, Hydrol Earth Syst Sci, 14(8), 1623–1637, doi:10.5194/hess-14-1623-2010, 2010.

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(10), 3889–3908, doi:10.5194/hess-16-3889-2012, 2012.

Vergnes, J.-P. and Habets, F.: Impact of river water levels on the simulation of stream–aquifer exchanges over the Upper Rhine alluvial aquifer (France/Germany), Hydrogeol. J., 1–15, doi:10.1007/s10040-018-1788-0, 2018.

Vergnes, J.-P., Decharme, B., Alkama, R., Martin, E., Habets, F. and Douville, H.: A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France, J. Hydrometeorol., 13(4), 1149–1171, doi:10.1175/JHM-D-11-0149.1, 2012.

Vergnes, J.-P., Decharme, B. and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res. Atmospheres, 119(19), 2014JD021573, doi:10.1002/2014JD021573, 2014.

Vidal, J., Martin, E., Franchistéguy, L., Baillon, M. and Soubeyroux, J.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, Int. J. Climatol., 30(11), 1627–1644, doi:10.1002/joc.2003, 2010.

Viennot, P.: Modélisation mathématique du fonctionnement hydrogéologique du bassin de la Seine. Représentation différentiée des aquifères du Tertiaire/Intégration des prélèvements anthropiques, R071128PVIE, Centre de Géosciences, Fontainebleau., 2009.

Viennot, P. and Abasq, L.: Modélisation de la pollution nitrique des grands aquifères du bassin de Seine-Normandie à l'échelle des masses d'eau. Développement des sous modèles hydrogéologiques "Craie" du bassin de la Seine. Rapport final, Centre de Géosciences, Fontainebleau., 2013.

Westerhoff, R., White, P. and Rawlinson, Z.: Incorporation of Satellite Data and Uncertainty in a Nationwide Groundwater Recharge Model in New Zealand, Remote Sens., 10(1), 58, doi:10.3390/rs10010058, 2018.

~~Aeschbach-Hertig, W. and Gleeson, T.: Regional strategies for the accelerating global problem of groundwater depletion, Nat. Geosci., 5(12), 853–861, doi:10.1038/ngeo1617, 2012.~~

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P. and Meyer, G.: Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland, J. Hydroinformatics, 20(6), 1227–1246, doi:10.2166/hydro.2018.002, 2018.

Amraoui, N. and Seguin, J.-J.: Simulation par modèle maillé de l'impact d'épisodes pluvieux millénaux sur les niveaux de la nappe de la craie et les débits du fleuve Somme. Rapport final, BRGM/RP-61864-FR, Orléans., 2012.

Amraoui, N., Castillo, C. and Seguin, J. J.: Évaluation de l'exploitabilité des ressources en eau souterraine de la nappe de la craie du bassin de la Somme. Rapport final, BRGM/RP-63408-FR, Orléans., 2014.

Barlage, M., Tewari, M., Chen, F., Miguez-Macho, G., Yang, Z. L. and Niu, G. Y.: The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP, Clim. Change, 129(3), 485–498, doi:10.1007/s10584-014-1308-8, 2015.

Bessière, H., PIcot, J., Picot, G. and Parmentier, M.: Affinement du modèle hydrogéologique de la Craie du Nord-Pas-de-Calais autour des champs captants de la métropole Lilloise. Rapport final, BRGM/RP-63689-FR, Orléans., 2015.

Boone, A., Masson, V., Meyers, T. and Noilhan, J.: The Influence of the Inclusion of Soil Freezing on Simulations by a Soil–Vegetation–Atmosphere Transfer Scheme, J. Appl. Meteorol., 39(9), 1544–1569, doi:10.1175/1520-0450(2000)039<1544:TIOTIO>2.0.CO;2, 2000.

Braud, I., Varado, N. and Olioso, A.: Comparison of root water uptake modules using either the surface energy balance or potential transpiration, J. Hydrol., 301(1–4), 267–286, doi:10.1016/j.jhydrol.2004.06.033, 2005.

Buis, S., Piacentini, A. and Déclat, D.: PALM: a computational framework for assembling high-performance computing applications, Concurr. Comput. Pract. Exp., 18(2), 231–245, doi:10.1002/cpe.914, 2005.

Courtois, Q.: A hillslope-based aquifer model of free-surface flows in crystalline regions, p. 15, Saint-Malo. [online] Available from: https://www.irisa.fr/sage/jocelyne/CMWR2018/pdf/CMWR2018_paper_429.pdf, 2018.

Coustau, M., Rousset-Regimbeau, F., Thirel, G., Habets, F., Janet, B., Martin, E., de Saint-Aubin, C. and Soubeyroux, J.-M.: Impact of improved meteorological forcing, profile of soil hydraulic conductivity and data assimilation on an operational Hydrological Ensemble Forecast System over France, J. Hydrol., 525, 781–792, doi:10.1016/j.jhydrol.2015.04.022, 2015.

Croiset, N., Wuilleumier, A., Bessière, H., Gresselin, F. and Seguin, J. J.: Modélisation des aquifères de la plaine de Caen et du bassin de la Dives. Phase 2 : construction et calage du modèle hydrogéologique, BRGM/RP-62648-FR, Orléans., 2013.

David, C. H., Habets, F., Maidment, D. R. and Yang, Z. L.: RAPID applied to the SIM-France model, Hydrol. Process., 25(22), 3412–3425, doi:10.1002/hyp.8070, 2011.

De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Oude Essink, G. H. P., van Walsum, P. E. V., Delsman, J. R., Hunink, J. C., Massop, H. Th. L. and Kroon, T.: An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument, Environ. Model. Softw., 59, 98–108, doi:10.1016/j.envsoft.2014.05.009, 2014.

Decharme, B., Martin, E. and Faroux, S.: Reconciling soil thermal and hydrological lower boundary conditions in land surface models, J. Geophys. Res. Atmospheres, 118(14), 7819–7834, doi:10.1002/jgrd.50631, 2013.

Douez, O.: Actualisation 2008-2011 du modèle maillé des aquifères du Jurassique en Poitou-Charentes. Rapport final, BRGM/RP-64816-FR, Orléans., 2015.

25

Duchaine, F., Jauré, S., Poitou, D., Quémerais, E., Staffelbach, G., Morel, T. and Gicquel, L.: Analysis of high performance conjugate heat transfer with the OpenPALM coupler, Comput. Sci. Discov., 8(1), 015003, doi:10.1088/1749-4699/8/1/015003, 2015.

Dudley, R. W., Hodgkins, G. A. and Dickinson, J. E.: Forecasting the Probability of Future Groundwater Levels Declining Below Specified Low Thresholds in the Conterminous U.S., JAWRA J. Am. Water Resour. Assoc., 53(6), 1424–1436, doi:10.1111/1752-1688.12582, 2017.

Fan, Y., Li, H. and Miguez-Macho, G.: Global Patterns of Groundwater Table Depth, Science, 339(6122), 940–943, doi:10.1126/science.1229881, 2013.

Guzman, S. M., Paz, J. O. and Tagert, M. L. M.: The Use of NARX Neural Networks to Forecast Daily Groundwater Levels, Water Resour. Manag., 31(5), 1591–1603, doi:10.1007/s11269-017-1598-5, 2017.

Habets, F., Boone, A., Champeaux, J. L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Moigne, P. L., Martin, E., Morel, S., Noilhan, J., Quintana Seguí, P., Rousset Regimbeau, F. and Viennot, P.: The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, J. Geophys. Res., 113(D6), D06113, doi:10.1029/2007JD008548, 2008.

Habets, F., Gascoin, S., Korkmaz, S., Thiéry, D., Zribi, M., Amraoui, N., Carli, M., Ducharne, A., Leblois, E., Ledoux, E., Martin, E., Noilhan, J., Ottlé, C. and Viennot, P.: Multi-model comparison of a major flood in the groundwater-fed basin of the Somme River (France), Hydrol Earth Syst Sci, 14(1), 99–117, doi:10.5194/hess-14-99-2010, 2010.

Habets, F., Amraoui, N., Caballero, Y., Thiéry, D., Vergnes, J. P., Morel, T., Le Moigne, P., Roux, N., Dreuzy, J. R. D., Ackerer, P., Maina, F., Besson, F., Etchevers, P., Regimbeau, F. and Viennot, P.: Plateforme de modélisation hydrogeologique nationale AQUI-FR : Rapport final de 1ère phase. (The national hydrogeological plateform Aqui-FR. Report of the first phase), ONEMA. [online] Available from: https://www.metis.upmc.fr/~aqui-fr/Rapport_fin_phase1_Aqui FR_VF.pdf, 2017.

Hassane Maina, F., Delay, F. and Ackerer, P.: Estimating initial conditions for groundwater flow modeling using an adaptive inverse method, J. Hydrol., 552, 52–61, doi:10.1016/j.jhydrol.2017.06.041, 2017.

He, X., Stisen, S., Wiese, M. B. and Henriksen, H. J.: Designing a Hydrological Real-Time System for Surface Water and Groundwater in Denmark with Engagement of Stakeholders, Water Resour. Manag., 30(5), 1785–1802, doi:10.1007/s11269-016-1251-8, 2016.

Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C. and Madsen, B.: Methodology for construction, calibration and validation of a national hydrological model for Denmark, J. Hydrol., 280(1–4), 52–71, doi:10.1016/S0022-1694(03)00186-0, 2003.

Herbst, M., Fialkiewicz, W., Chen, T., Pütz, T., Thiéry, D., Mouvet, C., Vachaud, G. and Vereecken, H.: Intercomparison of Flow and Transport Models Applied to Vertical Drainage in Cropped Lysimeters, Vadose Zone J., 4(2), 240–254, doi:10.2136/vzj2004.0070, 2005.

Højberg, A. L., Troldborg, L., Stisen, S., Christensen, B. B. S. and Henriksen, H. J.: Stakeholder driven update and improvement of a national water resources model, Environ. Model. Softw., 40, 202–213, doi:10.1016/j.envsoft.2012.09.010, 2013.

Kollet, S., Gasper, F., Brdar, S., Goergen, K., Hendricks Franssen, H. J., Keune, J., Kurtz, W., Küll, V., Pappenberger, F., Poll, S., Trömel, S., Shrestha, P., Simmer, C. and Sulis, M.: Introduction of an Experimental Terrestrial Forecasting/Monitoring System at Regional to Continental Scales Based on the Terrestrial Systems Modeling Platform (v1.1.0), Water, 10(11), 1697, doi:10.3390/w10111697, 2018.

26

Korkmaz, S.: Modélisation des régimes de crue des systèmes couplés aquifères rivières, thesis, Paris, ENMP, 1 January. [online] Available from: http://www.theses.fr/2007ENMP1495 (Accessed 30 November 2018), 2007.

Ledoux, E., Girard, G. and De Marsily, G.: Spatially distributed modelling: conceptual approach, coupling surface water and groundwater, in Unsaturated Flow in Hydrologic Modeling: Theory and Practice., vol. 275, pp. 435–454, H.J. Morel-Seytoux.,
5   1989.

Legates, D. R. and McCabe Jr, G. J. M.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, doi:10.1029/1998WR900018, 1999.

Long, D., Longuevergne, L. and Scanlon, B. R.: Global analysis of approaches for deriving total water storage changes from GRACE satellites, Water Resour. Res., 51(4), 2574–2594, doi:10.1002/2014WR016853, 2015.

10   Longuevergne, L., Scanlon, B. R. and Wilson, C. R.: GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA, Water Resour. Res., 46(11), doi:10.1029/2009WR008564, 2010.

Marçais, J., de Dreuzy, J. R. and Erhel, J.: Dynamic coupling of subsurface and seepage flows solved within a regularized partition formulation, Adv. Water Resour., 109, 94–105, doi:10.1016/j.advwatres.2017.09.008, 2017.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F.,
15   Brousseau, P., Brun, E., Calvet, J. C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A. L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J. F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V. and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6(4), 929–960, doi:10.5194/gmd-6-929-2013, 2013.

20   Maxwell, R. M., Condon, L. E. and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, Geosci Model Dev, 8(3), 923–937, doi:10.5194/gmd-8-923-2015, 2015.

McKee, T. B., Doesken, N. J. and Kleist, J.: The relationship of drought frequency and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology, vol. 17, pp. 179–183, American Meteorological Society Boston, MA., 1993.

25   Monteil, C., Flipo, N., Poulin, M., Habets, F., Krimissa, M. and Ledoux, E.: Assessing the contribution of the main aquifer units of the Loire basin to river discharge during low flow. [online] Available from: https://hal-mines-paristech.archives-ouvertes.fr/hal-00800812 (Accessed 8 January 2018), 2010.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

30   Noilhan, J. and Planton, S.: A Simple Parameterization of Land Surface Processes for Meteorological Models, Mon. Weather Rev., 117(3), 536–549, doi:10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2, 1989.

Noyer, M. L. and Elsass, P.: Modélisation de la salure profonde au droit et en aval du Bassin Potassique - Rapport final, BRGM/RP-54389-FR, Orléans., 2006.

Office International de l'Eau: Bulletin de situation hydrologique de janvier 2019, Office International de l'Eau, France. [online]
35   Available from: https://www.eaufrance.fr/publications/bsh/2019-01 (Accessed 7 February 2019), 2019.

27

Pachocka, M., Mansour, M., Hughes, A. and Ward, R.: Challenges of modelling a complex multi-aquifer groundwater system at a national scale: case study from the UK, pp. 216–220, Colorado School of Mines, Colorado, USA. [online] Available from: http://igwmc.mines.edu/conference/Mod2015/MM15_Proceedings.pdf (Accessed 24 July 2018), 2015.

5 Pappenberger, F., Ramos, M. H., Cloke, H. L., Wetterhall, F., Alfieri, L., Bogner, K., Mueller, A. and Salamon, P.: How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction, J. Hydrol., 522, 697–713, doi:10.1016/j.jhydrol.2015.01.024, 2015.

Philip, S., Kew, S. F., Jan van Oldenborgh, G., Aalbers, E., Vautard, R., Otto, F., Haustein, K., Habets, F. and Singh, R.: Validation of a Rapid Attribution of the May/June 2016 Flood-Inducing Precipitation in France to Climate Change, J. Hydrometeorol., 19(11), 1881–1898, doi:10.1175/JHM-D-18-0074.1, 2018.

10 Prudhomme, C., Hannaford, J., Harrigan, S., Boorman, D., Knight, J., Bell, V., Jackson, C., Svensson, C., Parry, S., Bachiller-Jareno, N., Davies, H., Davis, R., Mackay, J., McKenzie, A., Rudd, A., Smith, K., Bloomfield, J., Ward, R. and Jenkins, A.: Hydrological Outlook UK: an operational streamflow and groundwater level forecasting system at monthly to seasonal time scales, Hydrol. Sci. J., 62(16), 2753–2768, doi:10.1080/02626667.2017.1395032, 2017.

Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L. and Morel, 15 S.: Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France, J. Appl. Meteorol. Climatol., 47(1), 92–107, doi:10.1175/2007JAMC1636.1, 2008.

Rousset, F., Habets, F., Gomez, E., Moigne, P. L., Morel, S., Noilhan, J. and Ledoux, E.: Hydrometeorological modeling of the Seine basin using the SAFRAN-ISBA-MODCOU system, J. Geophys. Res., 109(D14), D14105, doi:10.1029/2003JD004403, 2004.

20 Roux, N.: The Aqui-FR project : the future operational modeling platform for the main regional multilayer aquifers in France, Saint-Malo. [online] Available from: https://www.irisa.fr/sage/jocelyne/CMWR2018/pdf/CMWR2018_paper_429.pdf, 2018.

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, doi:10.1016/j.jhydrol.2012.10.027, 2013.

25 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 – Module 1 – Année 5 - Convention 2008-2013. Rapport final, BRGM/RP-65039-FR, Orléans., 2016.

Seguin, J. J.: Proposition d'un indicateur piézométrique standardisé pour le Bulletin de Situation Hydrologique "Nappes". Rapport final, BRGM/RP-64147-FR, Orléans., 2015.

30 Seguin, J. J. and Klinka, T.: Index Piézométrique Standardisé (IPS) - Bilans et comparaisons avec l'indicateur du BSH - Note technique, BRGM/RP-67251-FR, Orléans., 2016.

Seibert, J., Vis, M. J. P., Lewis, E. and Meerveld, H. J. van: Upper and lower benchmarks in hydrological modelling, Hydrol. Process., 32(8), 1120–1125, doi:10.1002/hyp.11476, 2018.

Service de l'observation et des statistiques: Repères. L'eau et les milieux aquatiques. Chiffres clés., Ministère de 35 l'Environnement, de l'Energie et de la Mer., 2016.

Singla, S., Céron, J.-P., Martin, E., Regimbeau, F., Déqué, M., Habets, F. and Vidal, J.-P.: Predictability of soil moisture and river flows over France for the spring season, Hydrol Earth Syst Sci, 16(1), 201–216, doi:10.5194/hess-16-201-2012, 2012.

28

Stollsteiner, P.: PROJET Explore 2070 - Evaluation de l'impact du changement climatique - Rapport final, BRGM/RP-61483-FR, Orléans., 2012.

Sulis, M., Keune, J., Shrestha, P., Simmer, C. and Kollet, S. J.: Quantifying the Impact of Subsurface-Land Surface Physical Processes on the Predictive Skill of Subseasonal Mesoscale Atmospheric Simulations, J. Geophys. Res. Atmospheres, 123(17), 9131–9151, doi:10.1029/2017JD028187, 2018.

Texas Water Development Board: Groundwater Division | Texas Water Development Board, [online] Available from: https://www.twdb.texas.gov/groundwater/ (Accessed 6 November 2018), 2018.

Thierion, C.: Modélisation hydrogéologique des aquifères du Bajocien – Bathonien dans le secteur des Bocages Normands, Master 2, Université Pierre et Marie Curie, Ecole des Mines de Paris & Ecole Nationale du Génie Rural des Eaux et des Forêts., 2007.

Thierion, C., Longuevergne, L., Habets, F., Ledoux, E., Ackerer, P., Majdalani, S., Leblois, E., Lecluse, S., Martin, E., Queguiner, S. and Viennot, P.: Assessing the water balance of the Upper Rhine Graben hydrosystem, J. Hydrol., 424–425, 68–83, doi:10.1016/j.jhydrol.2011.12.028, 2012.

Thiéry, D.: Logiciel GARDÉNIA, version 8.2. Guide d'utilisation. Rapport final, BRGM/RP-62797-FR, Orléans., 2014.

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), BRGM/RP-64554-FR, Orléans., 2015a.

Thiéry, D.: Modélisation 3D des écoulements en Zone Non Saturée avec le code de calcul MARTHE - version 7.5, BRGM/RP-64495-FR, Orléans., 2015b.

Thiéry, D.: Modélisation 3D du transport de masse et d'énergie avec le code de calcul MARTHE - version 7.5, BRGM/RP-64765-FR, Orléans., 2015c.

Thiéry, D.: Validation du code de calcul GARDÉNIA par modélisations physiques comparatives, BRGM/RP-64500-FR, Orléans., 2015d.

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

Thiéry, D.: Modélisation hydrologique globale des débits de 23 sources karstiques avec le logiciel ÉROS. Rapport final, BRGM/RP-67723-FR, Orléans., 2018b.

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, pp. 285–295, ORSTOM Editions. [online] Available from: https://hal-brgm.archives-ouvertes.fr/hal-01061971/document (Accessed 10 January 2018), 1992.

Thiéry, D., Amraoui, N. and Noyer, M. L.: Modelling flow and heat transfer through unsaturated chalk – Validation with experimental data from the ground surface to the aquifer, J. Hydrol., 556, 660–673, doi:10.1016/j.jhydrol.2017.11.041, 2018.

Thirel, G., Martin, E., Mahfouf, J. F., Massart, S., Ricci, S. and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France – Part 1: Description and validation of the assimilation system, Hydrol Earth Syst Sci, 14(8), 1623–1637, doi:10.5194/hess-14-1623-2010, 2010.

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(10), 3889–3908, doi:10.5194/hess-16-3889-2012, 2012.

Vergnes, J.-P. and Habets, F.: Impact of river water levels on the simulation of stream–aquifer exchanges over the Upper Rhine alluvial aquifer (France/Germany), Hydrogeol. J., 1–15, doi:10.1007/s10040-018-1788-0, 2018.

Vergnes, J. P., Decharme, B., Alkama, R., Martin, E., Habets, F. and Douville, H.: A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France, J. Hydrometeorol., 13(4), 1149–1171, doi:10.1175/JHM-D-11-0149.1, 2012.

Vergnes, J. P., Decharme, B. and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res. Atmospheres, 119(19), 2014JD021573, doi:10.1002/2014JD021573, 2014.

Vidal, J., Martin, E., Franchistéguy, L., Baillon, M. and Soubeyroux, J.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, Int. J. Climatol., 30(11), 1627–1644, doi:10.1002/joc.2003, 2010.

Viennot, P.: Modélisation mathématique du fonctionnement hydrogéologique du bassin de la Seine. Représentation différentiée des aquifères du Tertiaire/Intégration des prélèvements anthropiques, R071128PVIE, Centre de Géosciences, Fontainebleau., 2009.

Viennot, P. and Abasq, L.: Modélisation de la pollution nitrique des grands aquifères du bassin de Seine-Normandie à l'échelle des masses d'eau. Développement des sous modèles hydrogéologiques "Craie" du bassin de la Seine. Rapport final, Centre de Géosciences, Fontainebleau., 2013.

Westerhoff, R., White, P. and Rawlinson, Z.: Incorporation of Satellite Data and Uncertainty in a Nationwide Groundwater Recharge Model in New Zealand, Remote Sens., 10(1), 58, doi:10.3390/rs10010058, 2018.

30

**Figure 1: Scheme of the AquiFR physical system. The simulation of the watersheds depends on its hydrogeologic characteristics. For sedimentary basins, the transfer of water within the watersheds is estimated by MARTHE or EauDyssée. It accounts for flows in the unsaturated zones, to (red thin arrow) and in the rivers, in (black arrows) and between (blue arrows) aquifer layers, as well as the exchange between the river and the aquifer (purple arrow). The temporal resolution is daily and the spatial resolution varies from 100 m to a maximum of 8000 m. The depth of the deepest aquifer layer can reach locally about 1000 meters. The 8 km spatial partition of the flow between surface runoff and groundwater recharge (red thick arrows) is estimated by the SURFEX land surface scheme. It solves the water and energy budget at a 5 minutes time step. It accounts for the local type of vegetation and soil, the presence of snow, and a multilayer soil that can reach a depth of 3 meters. The atmospheric forcing is provided by SAFRAN. For the karstic systems, the EROS conceptual model is used. It represents each karstic system as lumped basins based on a reservoir approach at a daily time scale. The incoming atmospheric forcing is provided by SAFRAN.**

**Figure 2: Scheme of the numerical implementation of AquiFR. (a) SAFRAN and SURFEX are run separately, as well as the processes that extract the daily surface runoff and groundwater recharge at 8 km resolution on a daily time step over the full 60-year period. (b) The components implemented within the coupling system O-Palm are presented. Pre-processing in blue gives access to the surface runoff and groundwater recharge as well as atmospheric forcing to the 3 groundwater models for the current time steps. Then, each hydrogeologic software runs all of their models for the current time step. The fluxes and state variables are then transferred daily to the post-processing, that writes the model outputs and manages the following time step.**

**Figure 3: Main aquifers of France classified by geological type from the BDLISA database (https://bdlisa.eaufrance.fr/). The names of the gauging stations and piezometers shown in Figure 8, Figure 12~~Figure 13~~ and Figure 13~~Figure 14~~ are written.**

**Figure 4: Map of the regional multilayer aquifers and the karstic systems simulated in AquiFR. The outlines of the models are also shown with colours corresponding to the outcropping aquifers with respect to their geological contexts. Grey areas correspond to models that will be integrated in a near future.**

**Figure 5: Temporal evolution of the number of piezometric head measurements per day among the 639 selected piezometers over the 1958-2018 simulated period.**

**Figure 6: (a) Spatial distribution of the biases calculated between the simulated and observed piezometric heads for the 639 selected piezometers. The grey background colour corresponds to the simulated aquifer domain. (b) ~~Accumulated~~Cumulative distribution of absolute biases for all the piezometers.**

**Figure 7: (a) Spatial distribution of NRMSE_BE calculated between the simulated and observed piezometric heads for the 639 selected piezometers. The grey background colour corresponds to the simulated aquifer domain. (b) ~~Accumulated~~Cumulative distribution of NRMSE_BE for all the piezometers.**

**Figure 8: Daily observed (dotted blue) and simulated (red) piezometric head variations for the five piezometers encircled in green in Figure 3.**

**Figure 9: NSE criterion calculated between the observed and simulated SPLI for the 103 selected piezometers.**

**~~Figure 10: Correlations calculated between the observed and simulated SPLI for the 103 selected piezometers.~~**

**~~Figure 11:~~ (left) Monthly observed (dotted blue) and simulated (red) SPLI indicator variations for the five piezometers encircled in green in Figure 3. Font colours correspond to the classes of Table 3 from the driest (red) to the wettest (blue) intervals. (Right) Histograms in percentage of the SPLI values distributed against the classes of Table 3.**

**Figure 11~~12~~: Standardised Piezometric Level Index for ~~the august~~June 2016 ~~month~~ for the (a) observed and (b) simulated piezometers. The SPLI are computed on the period for which the observations are available. (b) The map of the SPLI indicators calculated in each grid cell of the AquiFR domain with a common reference period is shown in the background.**

**Figure 12~~13~~: Monthly observed (~~dotted~~dashed blue) and simulated (red) river flows of the gauging stations monitoring the four karstic systems encircled in red in Figure 3. NSE of the square root of the daily karstic spring flows and drainage area are given in parenthesis for each gauging station. Grey backgrounds correspond to periods of gaps in the observations.**

31

**Figure 13~~14~~: Daily observed (~~dotted~~dashed blue) and simulated (red) river flows for the four gauging stations encircled in yellow in Figure 3. NSE, correlation and ratio scores are given in parenthesis for each gauging station. Grey backgrounds correspond to periods of gaps in the observations.**

**Figure 14~~15~~: (a) Spatial distribution of the NSE calculated between the observed and simulated karstic spring flows and river flows for the gauging stations monitoring the 23 simulated karstic systems (stars) and the 362 selected gauging stations located in the distributed hydrogeological models (circles). ~~Accumulated~~Cumulative distribution of NSE for (a) the 23 karstic systems and (b) the 362 gauging stations of the distributed models are also shown. The NSE for karstic systems are computed using the square root of the daily karstic spring flows. The simulated river network is shown in the background.**

**Figure 15~~16~~ : (a) Spatial average over the Somme model of daily mean groundwater recharge, (b) daily mean river-groundwater exchanges budget over the Somme river network, (c) spatial average over the Somme model of the monthly mean piezometric head and (d) monthly SPLI. Red colors in (b) indicate groundwater to river flows and blue colors stream to groundwater flows. The background colors in (d) correspond to the SPLI classes from Table 3.**

**Table 1: Short description of the regional multilayer aquifer models available in AquiFR. Periods of calibration are given in the Recalibration column and the type of variables used for recalibration in the Variables column. GW means groundwater level and RF river flow. GW levels were evaluated using RMSE and bias criteria. River flows were evaluated using NSE and ratio criteria.**

| Software | Model | Number of layers | Number of cells | References | Recalibration | Variables |
|---|---|---|---|---|---|---|
| EauDyssée | Basse-Normandie | 4 | 37 667 | (Thierion, 2007) | ~~Yes~~1986 – 2013 | GW |
| | Loire | 3 | 37 620 | (Monteil et al., 2010) | No | |
| | Marne-Loing | 4 | 66 235 | (Viennot and Abasq, 2013) | ~~Yes~~1996 – 2015 | GW |
| | Marne-Oise | 2 | 45 904 | (Viennot and Abasq, 2013) | ~~Yes~~1986 – 2015 | GW |
| | Seine | 6 | 41 609 | (Viennot, 2009) | Not necessary | GW |
| | Seine-Eure | 1 | 57 306 | (Viennot and Abasq, 2013) | In progress | |
| | Seine-Oise | 4 | 87 178 | (Viennot and Abasq, 2013) | ~~Yes~~1996 – 2015 | GW |
| | Somme | 1 | 63 226 | (Korkmaz, 2007) | No | |
| MARTHE | Alsace | 3 | 40 947 | (Noyer and Elsass, 2006) | No | |
| | Basse-Normandie | 10 | 93 800 | (Croiset et al., 2013) | No | |
| | Nord Pas-de-Calais | 10 | 226 077 | (Bessière et al., 2015) | ~~Yes~~1995 – 2009 | GW, RF |
| | Poitou-Charentes | 8 | 90 084 | (Douez, 2015) | Not necessary | GW, RF |
| | Somme | 1 | 66 924 | (Amraoui et al., 2014) | ~~Yes~~1989 – 2012 | GW, RF |

5

33

**Table 2: Statistical scores of the comparison between the simulated and observed daily evolution of the piezometers shown in Figure 8.**

| Piezometer | Model | Time series | | | SPLI | |
|---|---|---|---|---|---|---|
| | | NRMSE-BE | Correlation | Biases (m) | NSE | Correlation |
| Omiécourt | Somme | 0.93 | 0.85 | -0.86 | 0.73 | 0.87 |
| Ruffec | Poitou-Charentes | 0.58 | 0.82 | -1.44 | 0.6 | 0.79 |
| Le Bec Hellouin | Basse-Normandie | 0.57 | 0.84 | -2.76 | 0.73 | 0.86 |
| Farceaux | Seine-Oise | 0.52 | 0.86 | -8.34 | 0.67 | 0.84 |
| Bourdet | Poitou-Charentes | 1.02 | 0.18 | -0.72 | -0.51 | 0.24 |

**Table 3: Classification of water table level classes related to the values of the SPLI corresponding to the MSH limits**

| Classification | SPLI values | Return periods |
|---|---|---|
| Very low groundwater level | < -1.28 | > 10 dry years |
| Low groundwater level | between -1.28 and -0.84 | Between 10 dry years and 5 dry years |
| Moderately low groundwater level | between -0.84 and -0.25 | Between 5 dry years and 2.5 dry years |
| Normal groundwater level | between -0.25 and 0.25 | Between 2.5 dry years and 2.5 wet years |
| Moderately high groundwater level | between 0.25 and 0.84 | Between 2.5 wet years and 5 wet years |
| High groundwater level | between 0.84 and 1.28 | Between 5 wet years and 10 wet years |
| Very high groundwater level | > 1.28 | > 10 wet years |

5

**Table 4: Statistical scores of the comparison between the simulated and observed daily evolution of river discharges shown in Figure 14.**

| Gauging Station | NSE | Correlation | Ratio |
|---|---|---|---|
| Charente at Jarnac | 0.45 | 0.72 | 0.81 |
| Somme at Abbeville | 0.69 | 0.86 | 0.92 |
| Seine at Poses | 0.86 | 0.93 | 1.01 |
| Loire at Montjean | 0.94 | 0.97 | 1.05 |