Improving the structure of a hydrological model to forecast catchment response to intense rainfall

LHB: HYDROSCIENCE JOURNAL
https://doi.org/10.1080/27678490.2024.2341027

OPEN ACCESS Check for updates

Improving the structure of a hydrological model to forecast catchment response to intense rainfall

Paul C. Astagneau [a,b,c,d*], François Bourgin [a], Vazken Andréassian [a] and Charles Perrin [a]

[a]Université Paris-Saclay, INRAE, HYCAR Research Unit, Antony, France; [b]WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland; [c]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland; [d]Climate Change, Extremes and Natural Hazards in Alpine Regions Research Center CERC, Davos Dorf, Switzerland

ABSTRACT
We compared the flood forecasts issued by a model used by operational services in France (GRP) and by a model developed to improve the simulation of floods resulting from intense rainfall (GR5H_RI). We selected 10,652 flood events from 19 years of hourly data available for 229 French catchments. The models were combined with a state-updating procedure to produce forecasts at 3, 6, 12 and 24 h lead times. Results indicate that the GR5H_RI model performs better on average than the GRP model at all lead times, particularly for forecasting flash floods (rise time < 12 h), which occur mainly in summer and early autumn. The use of the last observed streamflow to update initial conditions does not compensate for GRP's structural errors in the case of fast catchment response to intense rainfall. The new structure therefore opens valuable operational perspectives.

KEYWORDS
Forecasting; hydrological modelling; rainfall intensity

MOTS CLÉS
Prévision; modélisation hydrologique; intensités de pluie

RÉSUMÉ
Nous présentons une évaluation comparée des prévisions de crue d'un modèle utilisé par les services opérationnels en France (GRP) et des prévisions d'un modèle développé pour améliorer la simulation des crues résultant de fortes intensités de pluie (GR5H_RI), réalisée sur un large échantillon de 229 bassins versants français répartis sur le territoire métropolitain. Nous avons sélectionné 10 652 événements, à partir de 19 ans de données au pas de temps horaire, pour effectuer cette évaluation. Les modèles sont couplés à une méthode de mise à jour des états initiaux pour effectuer des prévisions aux horizons 3, 6, 12 et 24 h. Les résultats indiquent que le modèle GR5H_RI a en moyenne de meilleures performances que le modèle GRP aux quatre horizons, en particulier pour prévoir les crues dont les temps de montée sont inférieurs à 12 h et qui ont majoritairement lieu en été et au début de l'automne. L'exploitation du dernier débit observé ne permet pas de rattraper les erreurs structurelles du modèle GRP lors de réactions rapides des bassins versants à de fortes intensités de pluie. La nouvelle structure offre ainsi des perspectives opérationnelles intéressantes.

1. Introduction

In order to anticipate floods, operational forecasting systems exist in many countries (Pappenberger et al., 2016). These systems use meteorological forecasts and observations to predict river flows at time scales ranging from a few hours to a few days (e.g. Pagano et al., 2014; Wu et al., 2020). This transformation is generally carried out using hydrological models, which are an important component of these forecasting systems. Although many improvements have been brought to flood forecasting systems over the last two decades (e.g. Jain et al., 2018; Zanchetta & Coulibaly, 2020), both in terms of meteorological inputs and in terms of development of more efficient models, the forecasts made by these models are still subject to considerable uncertainty (e.g. Berthet et al., 2020; Brunner et al., 2021; Troin et al., 2021).

In particular, hydrological models have lower predictive capacities in arid basins and in basins characterised by dry conditions in certain seasons (e.g. McMillan et al., 2016; Melsen et al., 2018), especially when flash floods occur (e.g. Hapuarachchi et al., 2011) and when rainfall intensities are high (Astagneau et al., 2021). To improve the predictive capabilities of models under these conditions, numerous studies have focused on the diagnosis of hydrological model simulations and on improving the simulation of flash floods. For example, De Boer-Euser et al. (2017) showed that some rainfall-runoff models face difficulties in reproducing hydrological signatures representing rapid summer dynamics in the Meuse basin. Knoben et al. (2020) showed that models incorporating a runoff process based on infiltration excess or a rapid routing component perform better in catchments experiencing floods

CONTACT Paul C. Astagneau paul.astagneau@slf.ch
*First author is now at b,c,d

under dry conditions. Roux et al. (2011) introduced a hydrological model dedicated to the simulation of flash floods in Mediterranean basins. Pang et al. (2020) tested a modification of the soil conservation service curve number (SCS-CN) to take into account rainfall intensities and slope in order to improve flood simulation in the Chao River catchment. To our knowledge, there are no studies specifically documenting the failure of flood forecasting models (i.e. a hydrological model combined with a data assimilation method) to simulate floods associated with high rainfall intensities.

The GRP flood forecasting model (Viatgé et al., 2019) has been used in France by most of the flood forecasting services in the Vigicrues network for more than 10 years. It is based on the work of Tangara (2005) and Berthet (2010), and has been applied to many French rivers. Its structure derives from the GR4J model (Perrin et al., 2003), which had been simplified to allow efficient use of the observed flow. GRP is a parsimonious deterministic model with three free parameters, which operates at time steps from sub-hourly to daily. It incorporates a procedure for assimilating observed streamflow in real time to update its initial states. Outputs are also corrected using the previous error to update future forecasts. The model can forecast flows for lead times ranging from less than 1 h to up to 5 days, the quality of forecasts at the various lead times being dependent on several factors such as model quality, basin dynamics, quality of rainfall forecasts, etc. Several limitations of the GRP model were identified based on end-users' feedbacks, in particular a tendency to underestimate flood volumes and peaks, and a delay in the flood rising limb. For example, in June 2016, the exceptional flood of the Seine River and its tributaries, and also the tributaries of the Loire River, was underestimated by the GRP model (e.g. Peredo, 2021). A number of aspects of the hydrological modelling chain may explain these difficulties, including the estimation of model parameters, its structure and the data assimilation method used.

Previous research work on the GRP model has sought an effective compromise in complexity between three elements: the structure of the model, the estimation of its parameters and the assimilation of the last observed discharge. The adopted strategy is to give a central place to the flow observed at the time of the forecast, which is used to update the level of the model store directly generating streamflow. The resulting structure is simple, with only three free parameters, which limits equifinality problems and increases the model's robustness. The updating procedures used by the GRP model have a major impact on the quality of its forecasts.

In this study, we seek an alternative modelling compromise for flood forecasting, particularly when catchments respond to high rainfall intensities under conditions of low antecedent wetness. In this situation, data assimilation is not always sufficient to compensate for the structural limitations of hydrological models. We are therefore seeking to answer the following question: Does improving the structure of a hydrological model lead to an improvement in the quality of deterministic forecasts, particularly when catchments react rapidly to high rainfall intensities?

To answer this question, we compared the forecasts of the GRP model with those of the GR5H_RI model developed specifically to improve the simulation of floods resulting from high rainfall intensities (Astagneau et al., 2022). The GR5H_RI model is combined with a state-updating method similar to that of GRP in order to produce forecasts up to 24 h ahead. The aim of this work is to quantify the differences between these two models as a function of flood type and forecast lead time. Comparisons were made over 19 years of available hourly data for 229 catchments in mainland France, from which 10,652 flood events were selected.

2. Data and methods

2.1. Data

This work is based on a large database of 229 French catchments (Figure 1) spread across mainland France. Anthropogenic activities and snow have limited influence on the hydrological behaviour of these catchments. Hourly time series were used over the period 2000–2018, representing a wide range of hydroclimatic conditions.

We used Météo-France's Comephore radar reanalysis (Tabary et al., 2012), available at a 1 km^2 resolution, to compile hourly time series of rainfall aggregated at the catchment scale. Hourly potential evapotranspiration (PET) series were calculated from a disaggregation of daily series derived from the formula of Oudin et al. (2005) using SAFRAN temperatures (Vidal et al., 2010). Temporal disaggregation of daily PET series was carried out using a parabolic distribution between 6 am and 7 pm. The temporal flow series were taken from an extraction of data from the Banque Hydro (Leleu et al., 2014). This extraction was carried out by Delaigue et al. (2020). We selected 10,652 flood events from this database using an automatic selection algorithm. The detailed characteristics of this database are presented by Astagneau (2022).

2.2. Models

Precipitation and PET time series were used as inputs for two forecasting models (Figure 2):

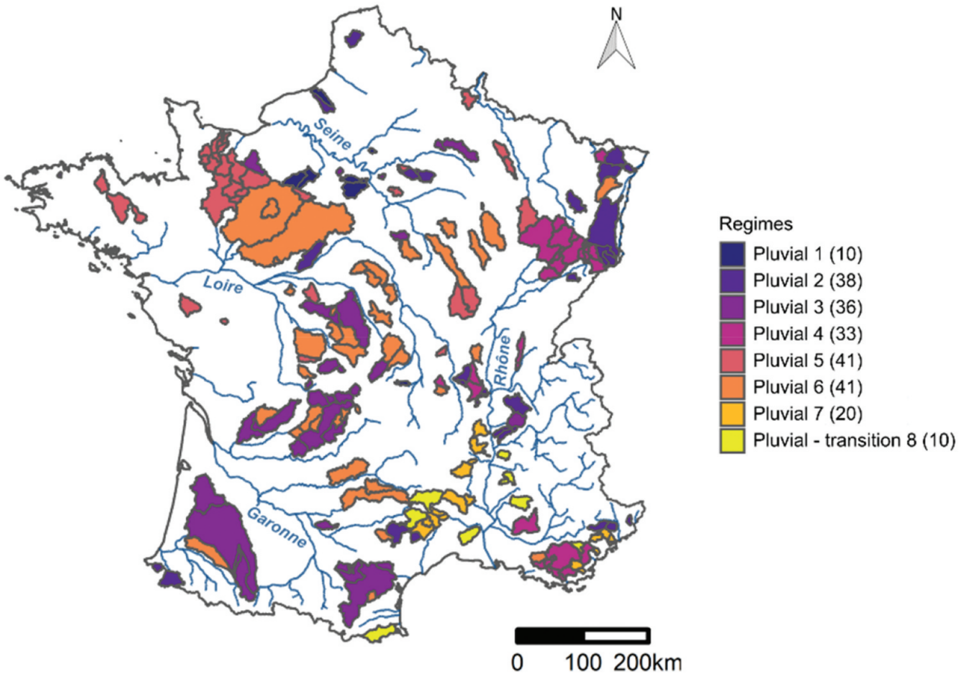

Figure 1. Location of the 229 catchments in mainland France. Classification into hydrological regimes as defined by Sauquet et al. (2008).

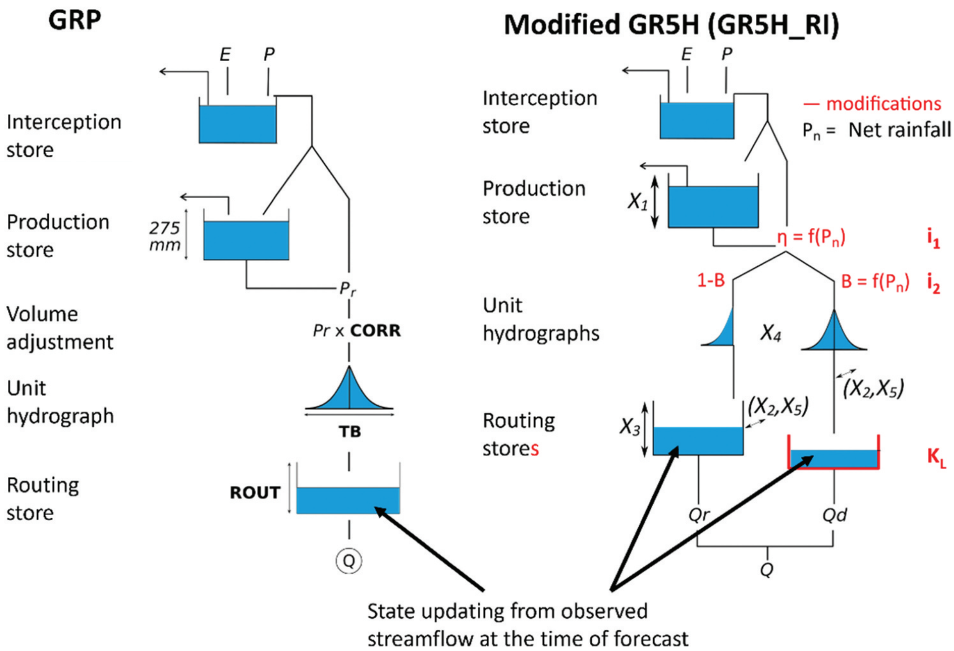

Figure 2. Schematic diagrams of the GRP model and the GR5H model integrating the rainfall intensity functions (GR5H_RI). CORR (effective rainfall correction coefficient; unitless), TB (unit hydrograph base time; h) and ROUT (routing store capacity; mm) are the free parameters of the GRP model. X_1 (production store capacity; mm), X_2 (exchange coefficient; mm/h), X_3 (indirect branch routing store capacity; mm), X_4 (unit hydrograph base half-time; h) and X_5 (exchange threshold; unitless) are free GR5H model parameters. i_1 (rainfall intensity coefficient; h/mm), i_2 (rainfall intensity coefficient; h/mm) and K_L (direct branch linear store emptying coefficient; unitless) are free parameters added to the GR5H model to improve the simulation of catchment response to intense rainfall. E, P and Q are potential evaporation, precipitation and streamflow, respectively. Other symbols are internal model state variables.

- The GRP model (Berthet, 2010; Tangara, 2005; Viatgé et al., 2019) has three free parameters and incorporates a procedure for updating the routing store using the flow observed at the time of forecast. The procedure for correcting forecast flows using the previous error at 1 h is not activated for the forecasts evaluated in this study.
- The GR5H_RI model, developed by Astagneau (2022), is a modified version of the GR5H model (Ficchì et al., 2019; Le Moine, 2008), which aims

to improve the simulation of catchment response to intense rainfall. GR5H_RI has eight free parameters and incorporates a procedure for updating its two routing stores (see the description of this procedure in section 2.3).

Two modelling functions have been introduced into the structure of the GR5H model in order to improve its performance when intense rainfall events occur in the dry season (Figure 3; for more details on these modelling functions, see Astagneau et al. (2022) and Peredo et al. (2022)). The first hypothesis aims to modify the net rainfall production rate (provided by the production store) as a function of the net rainfall intensity (volume hypothesis). The production rate calculated by GR5H depends solely on the level of the production store (high rate when the store level is high, i.e. when soil moisture is high, and a low rate when the store level is low). In the GR5H_RI model, when rainfall intensity is high and the level of the production store is low, the production rate increases (i.e. it becomes different from the rate initially calculated by GR5H). This change is controlled by parameter i_1. The second function aims to increase the partitioning of effective rainfall with fast dynamics when rainfall intensities are high (temporal distribution hypothesis). In the GR5H structure, 10% of the effective rainfall systematically passes through the direct branch of the routing function. In the GR5H_RI model, this fraction depends on the intensity of the net rainfall. This change is controlled by the i_2 parameter. A linear store (parameter K_L) has been added to the direct branch to improve the simulation of recessions when a large proportion of the effective rainfall passes through this branch.

2.3. Calibration, updating and evaluation

2.3.1. Hydrological forecasts

To test the models on past series, the models are run successively at each time step in the series, as if they were used in real time. At each forecast time t_0, specific model states are updated on the basis of the observed flow. Once the initial states have been updated, the forecasts are run between $t_0 + 1$ and $t_0 + H$, where H is the forecast horizon. This reproduces a close to real-time forecast situation: all the information available up to the time of the forecast is used to produce a forecast for the future. However, in real time rainfall forecasts are used between $t_0 + 1$ and $t_0 + H$, while in this study we used so called "perfect" rainfall scenarios, corresponding to the catchment rainfall records observed a posteriori. This allows us to compare the forecasting models independently of the uncertainties arising from the rainfall forecasts, and therefore with less uncertainty than in real time.

2.3.2. State updating

The initial conditions of the GRP model are updated on the basis of the last observed flow. Only the level of

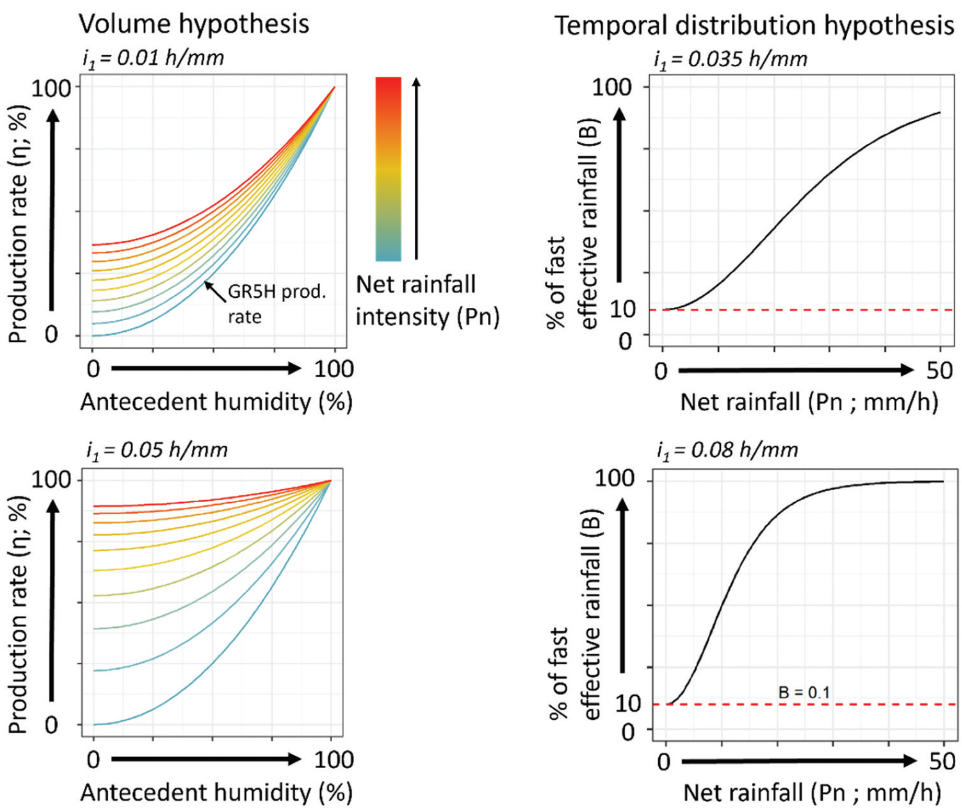

Figure 3. Simplified diagram of the rainfall intensity functions proposed by Astagneau et al. (2022).

the routing store is updated. At each forecast time, the level of the routing store is adjusted to reproduce exactly the last observed flow (direct insertion updating method).

The GR5H_RI model has two routing branches. This means that the direct insertion updating procedure is necessarily different from the GRP updating procedure. When testing the structure of the GRP model, Berthet (2010) evaluated a version with two routing branches (with a single routing store on one of the two branches). To update the routing store, he suggests to find the level whose outflow Q^{*}_{R,t_0} is equal to $Q_{obs,t_0} - Q_{D,t_0}$, with Q_{obs,t_0} the observed flow at the forecast time and Q_{D,t_0} the flow coming from the pseudo-direct branch. However, this rule implies that when $Q_{D,t_0} \geq Q_{obs,t_0}$, the routing store on the indirect branch empties almost completely to adjust to the observed flow. In this case, the update does not fully adjust to the observed flow so as not to obtain negative store level values. To limit this effect, we assume that the information provided by the model error at a given time does not allow to estimate an "observed" flow partitioning. We therefore prefer to use the partitioning initially simulated by the model.

We proposed to determine a target outflow Q^{*}_{R,t_0} as follows:

$$Q^{*}_{R,t_0} = \frac{Q_{R,t_0}}{Q_{R,t_0}+Q_{D,t_0}} \times Q_{obs,t_0} = \alpha \times Q_{obs,t_0} \tag{1}$$

$Q_{R,t0}$ is the simulated flow that comes from the indirect branch (before updating); α is the fraction of the simulated flow that comes from the indirect branch at the forecast time (before updating).

As the GR5H_RI model has a linear store on the pseudo-direct branch, a target outflow Q^{*}_{D,t_0} can also be calculated:

$$Q^{*}_{D,t_0} = (1-\alpha) \times Q_{obs,t_0} \tag{2}$$

The equation of the outflow of the GR5H routing store (instantaneous function to the power of 5 of the level) is not analytically inversible. There is therefore no analytical solution giving the level of the store whose outflow is equal to Q^{*}_{R,t_0}. To find an approximate level at each time t_0, 1000 values between $X3/1000$ and $X3$ are tested, $X3$ being the routing store capacity in mm (see Pelletier (2021) for a similar procedure with the GR6J model).

The linear store of the pseudo-direct branch of GR5H_RI, on the other hand, is inversible. The level of the updated linear store $L^{*}_{t_0}$ at time t_0 is calculated as follows:

$$L^{*}_{t_0} = Q^{*}_{D,t_0} \times \frac{1-K_L}{K_L} \tag{3}$$

where K_L [-] is the linear store coefficient. K_L takes values between 0 and 1.

2.3.3. Calibration of model parameters

The model parameters are estimated from a calibration procedure carried out per catchment and over two independent sub-periods (P1: 1 January 2000–30 June 2009; P2: 1 July 2009–31 December 2018). As we wish to analyse the performance of the models independently of the parametric compensations induced by a calibration of the models with the update procedures activated, the two models are calibrated in simulation mode. This means that we take the calibrated parameters for each catchment and each sub-period independently of the initial state updating procedures. A period of 2 years preceding the start of each sub-period is applied to initialise the model states (the performance criteria are not calculated over these periods). The calibration algorithm used is based on the exhaustive gridding discretisation (EGD) algorithm of Perrin et al. (2008). This algorithm is implemented in the airGR package (Coron et al., 2017, 2020). The two models were calibrated using the NSE criterion (Nash & Sutcliffe, 1970) as the objective function.

2.3.4. Evaluation of forecast quality

Forecasts are evaluated for a fixed lead time. In other words, all the forecasts made successively for each time step in the data series for a given lead time are put together to create a series of forecast flows for that lead time. This time series is then compared with the observations according to the chosen criterion. For this study, we use three event-based error criteria:

- All the forecasts within the flood event temporal window (from start to end of an event) were considered to calculate a bounded version of the NSE (Mathevet et al., 2006). For an event j at lead time H, this criterion is expressed as follows:

$$bounded\_NSE_{j,H} = \frac{NSE_{j,H}}{2-NSE_{j,H}} \tag{4}$$

This bounded version of the NSE takes values between −1 and 1. Negative values of bounded NSE indicate that the catchment mean flow is a better predictor than the model in forecasting a flood event. The optimal value of this criterion is 1. The use of the bounded version avoids a too-large impact of strongly negative values of NSE when calculating the average performance on a large set of catchments and flood events.

- The root mean square error (RMSE) calculated on flood rises. To calculate this criterion, the fixed-lead time forecasts used are those that fall within the rising window of the observed event (from the start of the event to the observed peak). The start of the event is defined as the time at which the flow exceeds 20% of the maximum flow for the event. Then, for each event and each lead time, we calculate the ratio between the flood rise RMSE of GRH_RI forecasts and the flood rise RMSE of GRP forecasts.

- The peak flow relative error criterion. It was adjusted to be relevant for fixed-lead-time time series analyses for the following reason: when the lag time between the observed peak flow and the predicted peak flow equals the lead time, it means that the peak flow was predicted at the forecast time of the observed peak flow. Therefore, it should not be taken into account in the calculation of the peak flow relative error criterion. To address this issue, the peak flow relative error of an event j at lead time H was calculated as follows:

$$Q_{max_{j,H}} = \frac{max\left[Q_{P,j,H}\right]}{max\left[Q_{obs,j,H}\right]} - 1 \tag{5}$$

where $max\left[Q_{P,j,H}\right]$ is the maximum predicted flow between $t_{max} - H/2$ and $t_{max} + H/2$, where t_{max} is the time of the observed peak flow. This criterion was then bounded:

$$bounded\_Q_{max_{j,H}} = \frac{Q_{max_{j,H}}}{Q_{max_{j,H}}+2} \tag{6}$$

In order to simplify the presentation of the results, the values of the criteria are presented in cross-evaluation (i.e. on P1 with the parameters estimated on P2 and vice versa) without distinguishing the temporal sub-periods. The results are presented for lead times of 3, 6, 12 and 24 h. First, the 10,652 events are divided into two categories: events with peak flows between November and April (winter) and events with peak flows between May and October (summer). In a second step, the events are categorised according to the time to rise to the observed peak. Thirdly, the events are categorised according to the average intensity of the associated rainfall events.

3. Results

We first look at the performance of the forecasts according to the season of peak flood occurrence. Figure 4 shows the distribution of event errors according to the bounded NSE (optimum value = 1) on flood events for the two models. The bounded-NSE distributions show higher values for the GR5H_RI model than for the GRP model, for all forecast horizons and for both seasons. The greatest differences in performance are observed for summer events. In winter and for the 3 h lead time, the differences in performance are less apparent.

Comparing the performance of GR5H_RI in terms of flood rise forecast as a function of the observed rise time (Figure 5), it can be seen that the shorter the rise time, the better the performance of GR5H_RI compared with GRP, for all forecast horizons. These differences increase as the lead time is reduced. Between 65% and 70% of events with a rise time of less than 12 h are better predicted by GR5H_RI than by GRP at 3 and 6 h lead times. In other words, the greatest differences in performance are observed for rise times of less than 12 h and for short lead times. The 50% quantile of the performance ratio is equal to 1 or even less than 1 for events with a rise time greater than 24 h for the 6, 12 and 24 h lead times, which means that GR5H-RI performs at least as well as GRP in these situations. Conversely, at the 3 h lead time, the 50% quantile of the performance ratio is slightly greater than 1, and 53% of these events are better predicted by GRP. These results indicate that the rise time, and therefore the flood dynamics, has a strong impact on the differences in performance between GR5H_RI and GRP, even in the very short term. The error at the forecast time, which is taken into account through the assimilation process, contains less information when large variations in streamflow occur in a very short

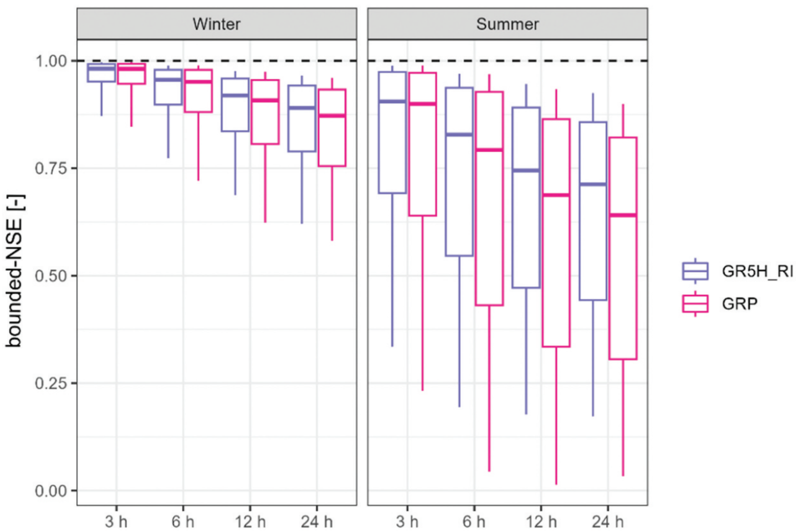

Figure 4. Distribution of the event performances (bounded NSE; cross-validation; one criterion value per event) of the GRP and GR5H_RI models for four lead times; 8,290 winter events and 2,362 summer events are considered here. Box plots are plotted from the 5% quantile to the 95% quantile.

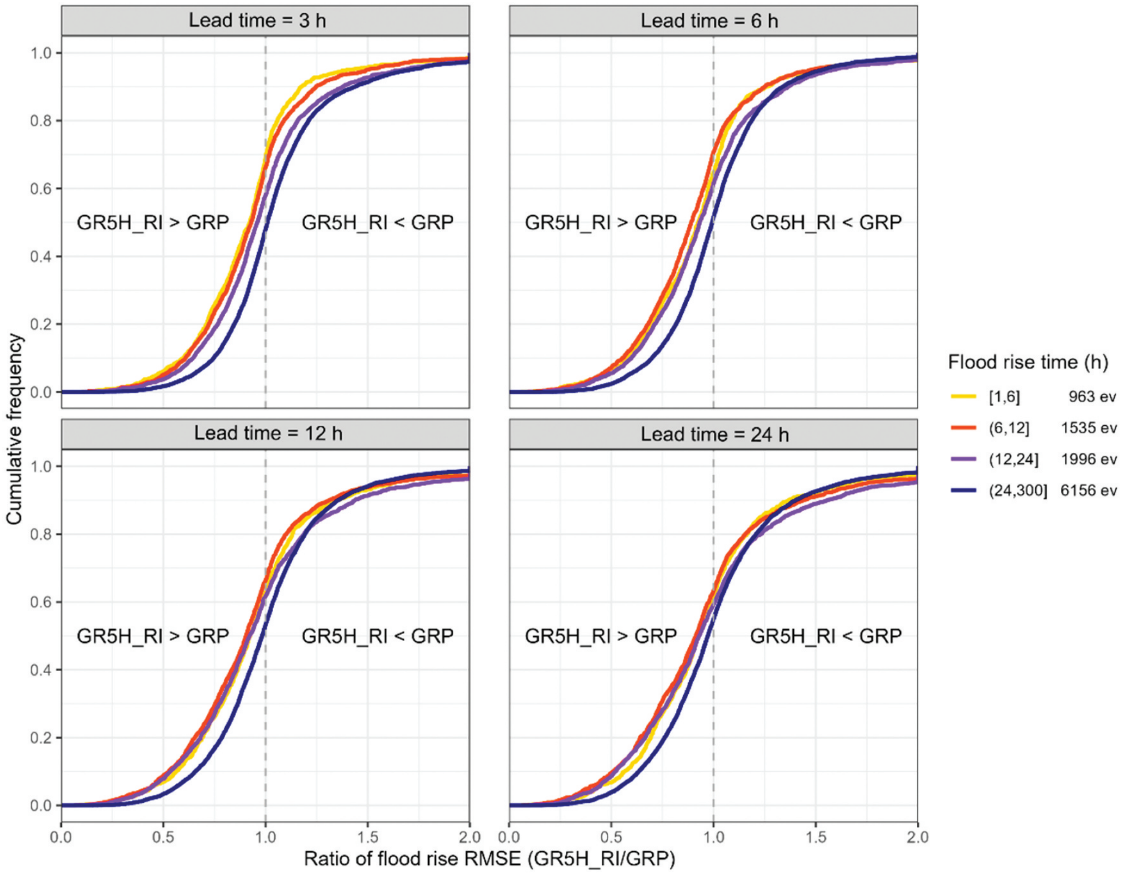

Figure 5. Cumulative distributions of the RMSE ratio on flood rises between GR5H_RI and GRP as a function of the rise time of the events. The results are presented in cross-evaluation for four lead times. The division of the rise times into four groups follows the values of the forecast times (3, 6, 12, 24 h). "> and <" "indicate better" than and "worse than", respectively.

time. In this case, the only way to improve forecasts is to improve the structure of the model.

The GR5H_RI model has been developed to take better account of rainfall intensity compared with GR5H and GRP, particularly in conditions of low antecedent humidity. Furthermore, the GRP model is known to frequently underestimate flood peaks. We now look at the impact of the change in structure on the ability of GR5H_RI and GRP forecasts to simulate flood peaks when rainfall intensities are high (Figure 6). The error of both models increases with increasing average rainfall intensity for all four lead times. At 3 h lead time, there is very little difference between the two models. At 6, 12 and 24 h lead time, the error on flood peaks is on average lower for GR5H_RI than for GRP, particularly for the highest rainfall intensities.

Figure 7 shows four hydrographs forecasted by the GRP and GR5H_RI models at 3, 12 and 24 h lead times for four catchments in our dataset. The hydrographs presented here are only illustrations of the previous analyses and do not constitute a representative sample of the results.

The Ardèche at Ucel (478 km^2) is a tributary of the Rhône downstream of Lyon. Numerous flood events take place in this basin in early autumn. The event that took place in November 2014 was characterised by rainfall intensities of up to 20 mm/h. The GR5H_RI model forecasts the flood peak better than GRP for all three lead times. At the 3 h lead time, the peak predicted by GRP at a fixed lead time is shifted by 3 h, which means that the associated time of forecast (t_0) was the time at which the peak was observed and the time at which the peak error was assimilated by the state updating procedure. Although this peak is of the same order of magnitude as the peak forecasted by GR5H_RI, the forecast is of poorer quality because the flood peak forecasted by GRP is offset from the observed peak by the same order of magnitude as the lead time considered.

The Ill at Osthouse (3296 km^2) is a tributary of the Rhine. It is characterised by slow dynamics and periods of high water, most of which occur in winter. The flood of December 2010 is a typical example of this basin. The rainfall event associated with this flood is spread over 5 days. The forecasts of the two models are satisfactory for the three lead times, but the flood peak is underestimated at 12 and 24 h. The GR5H_RI model overestimates part of the flood rise and underestimates the peak more strongly than GRP does.

The Estéron at Broc (443 km^2) is a Mediterranean coastal catchment whose high flows occur in winter and early autumn. It can be seen that both models underestimate the volume and peak of a flood that occurred in November 2014, particularly at 12 and

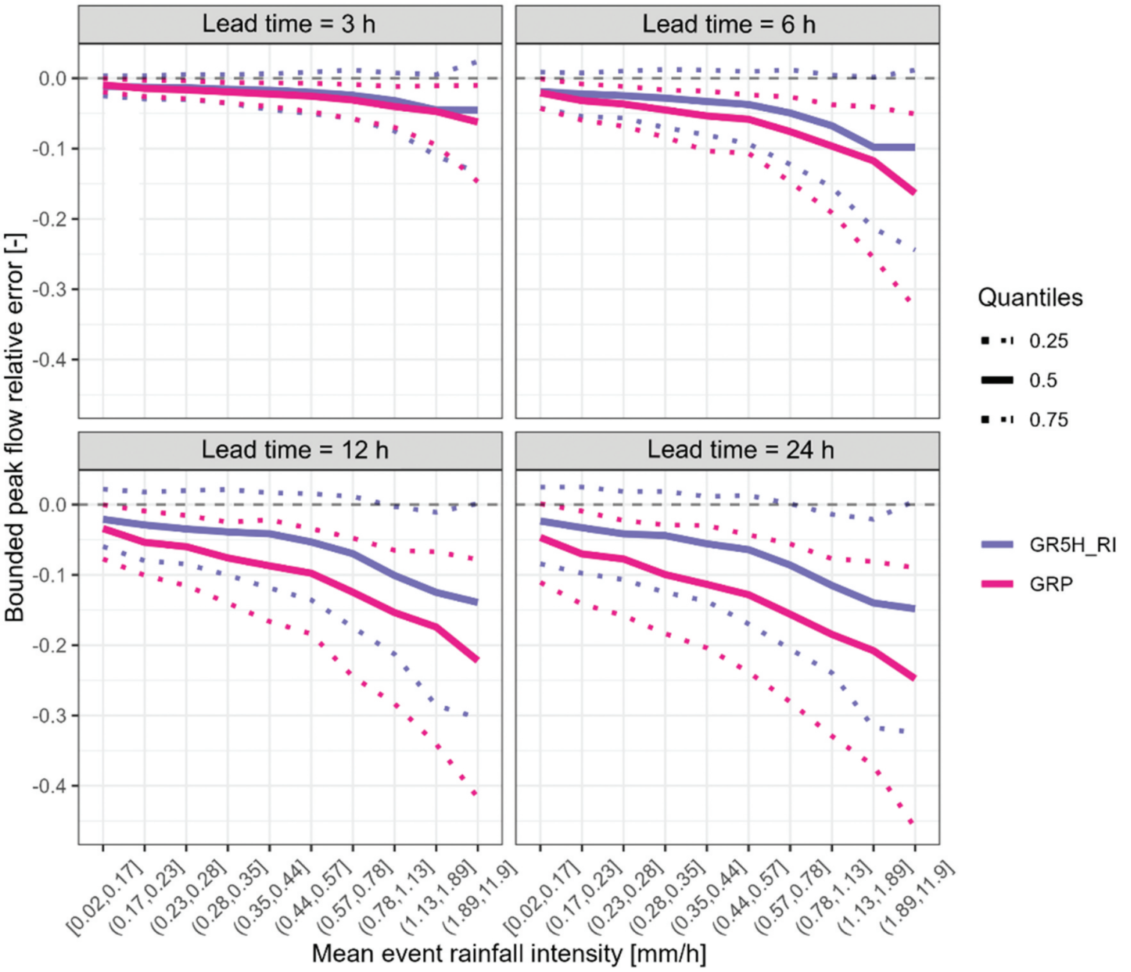

Figure 6. Distribution of the performance of GRP and GR5H_RI on flood peaks as a function of the average rainfall intensity of the events (cross-evaluation). The results are presented for four lead times. The events were divided into 10 quantile classes of mean rainfall intensity.

24 h. The GR5H_RI model forecasts the flood peak and rise better than GRP for these lead times. The difference is more pronounced at 24 h. At 3 h, the forecasts of the two models are close, but the peak and the rise in flood level are simulated slightly better by GR5H_RI. The Cèze at Tharaux (665 km^2) is a tributary of the Rhône. High-flow periods occur in autumn in this catchment. The September 2015 flood was characterised by rainfall intensities of up to 28 mm/h and a flood rise lasting 5 h. GR5H_RI forecasts the rise, volume and peak of the flood better than GRP for all three lead times. For this event, GRP's significant underestimation of flood volume is not compensated for by its data assimilation procedure, even in the short term.

4. Conclusion and perspectives

The aim of this study was to determine whether improving the structure of a hydrological model could lead to an improvement in the quality of deterministic forecasts, particularly during rapid catchment responses to high rainfall intensities. To meet this objective, we compared the forecasts of the GRP model, which has been used by flood forecasting services in France for several years, with the forecasts of a hydrological model developed to simulate floods resulting from high rainfall intensities. In order to compare the two models fairly, we constructed a method for updating the initial states of GR5H_RI similar to that used by GRP. The two models were used to produce forecasts up to 24 h ahead for 229 catchments from which 10,652 flood events were selected. The forecasts were evaluated using three event error criteria. Performance was compared as a function of the season of peak flood occurrence, event rise times and rainfall intensities.

The results showed that the GR5H_RI model, combined with a method to update its two routing stores, performed better on average than the GRP model on the flood events selected for the four lead times evaluated. Forecasts of flood events occurring in summer and early autumn are particularly improved. The differences in performance between the two models are highly dependent on the rise time of the events, even in the short term. Events associated with slower basin dynamics are forecasted equivalently by both models. Flood events associated with higher rainfall intensities

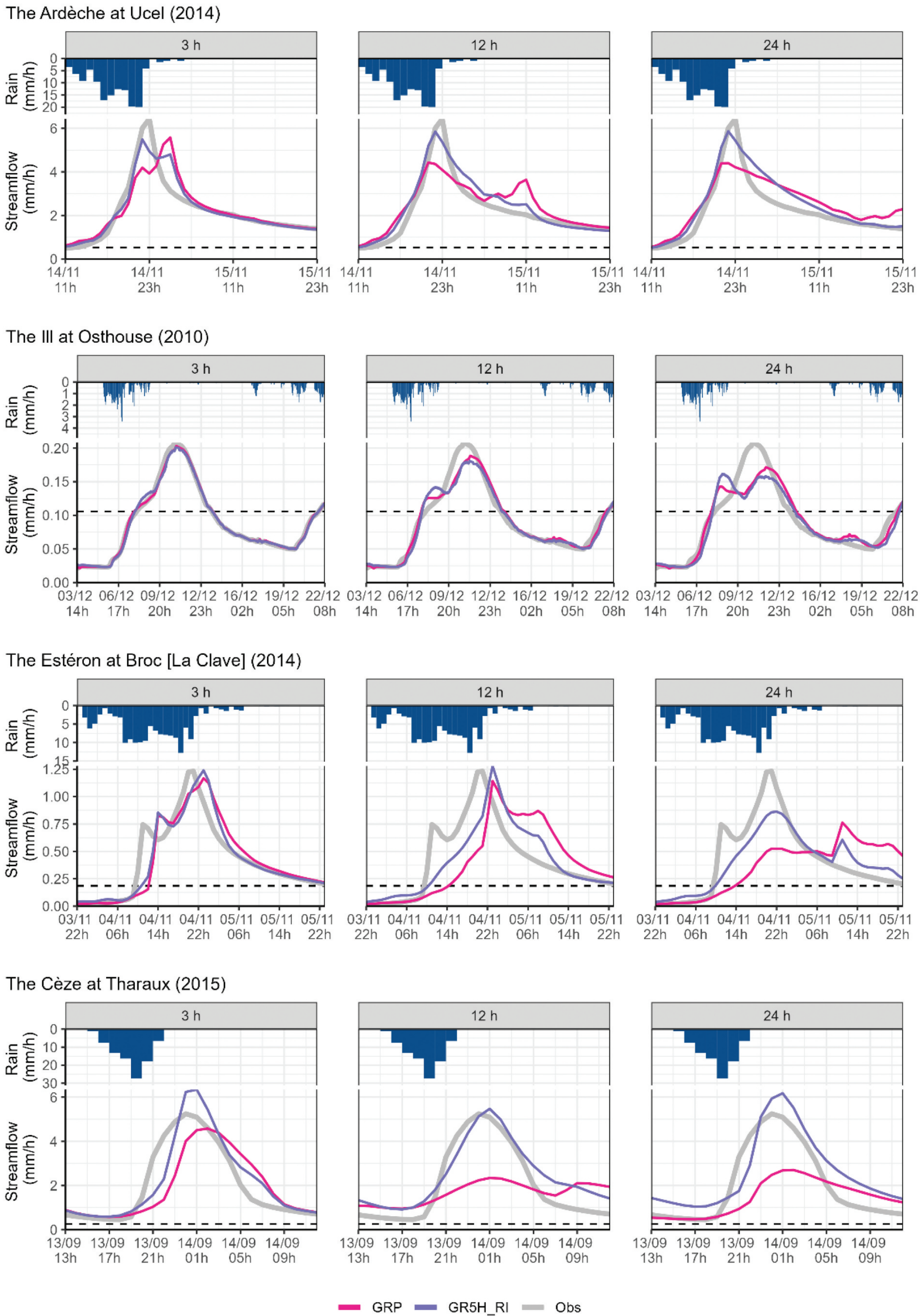

Figure 7. Examples of flood hydrographs (fixed-lead time) forecasted by the GRP model (pink line) and the GR5H_RI model (purple line) over four catchments for three lead times (evaluation). The black dotted line is the 95% quantile of flows over the entire series.

are better forecasted by GR5H_RI. These results indicate that the choice of a structure in which the assimilation of the last observed discharge has a very high impact on the forecasts is not appropriate when major variations in streamflow occur in a very short time.

The new version of the forecasting model proposed as a result of this work (GR5H_RI) should exceed the overall efficiency of GRP under operational conditions, and offer an applicability to a wider range of events and hydroclimatic contexts. However, the proposed improvements make the estimation of the parameters and the assimilation method more complex.

The conclusions of this work depend on the updating method used and the way the models were parameterised. In particular, the choice of estimating model parameters in simulation could have an influence on very short-term performance. More generally, the results presented in this work mean that we need to think again about the trade-offs between increasing the complexity of the structure, performance gains and parametric uncertainty. If the new version of the forecasting model is to be incorporated into future GRP operational developments, it will be necessary to determine whether the improvement in deterministic forecasts for some of the events associated with rapid catchment dynamics justifies increasing the complexity of the structure, parameter estimation and updating. Work is underway to find a parameter estimation method that is better suited to the GR5H_RI intensity functions.

Acknowledgments

This work was supported by the French Ministry of Environment (DGPR/SNRH/SCHAPI). The authors thank Météo-France and Banque HYDRO for providing the climatic and hydrological data, respectively. Finally, we thank the Associate Editor, Michel Lang, and two anonymous reviewers for their constructive comments that helped us improve the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the French Ministry of Environment.

ORCID

Paul C. Astagneau http://orcid.org/0000-0002-6688-5783
François Bourgin http://orcid.org/0000-0002-2820-7260
Vazken Andréassian http://orcid.org/0000-0001-7124-9303
Charles Perrin http://orcid.org/0000-0001-8552-1881

Data availability statement

Streamflow data were obtained from https://hydro.eaufrance.fr/. Precipitation and temperature data were provided by Méteo-France and are subject to third-party restrictions. All data used during the study are available from the corresponding author by request.

References

Astagneau, P. C. (2022). *Pistes d'amélioration de la généralité et de l'efficacité d'un modèle opérationnel de prévision des crues* [Doctoral dissertation, Sorbonne Université, INRAE Antony]. https://theses.hal.science/tel-04028903

Astagneau, P. C., Bourgin, F., Andréassian, V., & Perrin, C. (2021). When does a parsimonious model fail to simulate floods? Learning from the seasonality of model bias. *Hydrological Sciences Journal*, *66*(8), 1288–1305. https://doi.org/10.1080/02626667.2021.1923720

Astagneau, P. C., Bourgin, F., Andréassian, V., & Perrin, C. (2022). Catchment response to intense rainfall: Evaluating modeling hypotheses. *Hydrological Processes*, *36*(8), e14676. https://doi.org/10.1002/hyp.14676

Berthet, L. (2010). *Prévision des crues au pas de temps horaire: pour une meilleure assimilation de l'information de débit dans un modèle hydrologique* [Doctoral dissertation, Agro-ParisTech]. https://tel.archives-ouvertes.fr/pastel-00529652

Berthet, L., Bourgin, F., Perrin, C., Viatgé, J., Marty, R., & Piotte, O. (2020). A crash testing framework for predictive uncertainty assessment when forecasting high flows in an extrapolation context. *Hydrology and Earth System Sciences*, *24*(4), 2017–2041. https://doi.org/10.5194/hess-24-2017-2020

Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *WIREs Water*, *8*(3). https://doi.org/10.1002/wat2.1520

Coron, L., Delaigue, O., Thirel, G., Dorchies, D., Perrin, C., & Michel, C. (2020). *airGR: Suite of GR Hydrological models for precipitation-runoff modelling. R package version 1.6.9.27*. https://cran.r-project.org/web/packages/airGR/index.html

Coron, L., Thirel, G., Delaigue, O., Perrin, C., & Andréassian, V. (2017). The suite of lumped GR hydrological models in an R package. *Environmental Modelling and Software*, *94*, 166–171. https://doi.org/10.1016/j.envsoft.2017.05.002

de Boer-Euser, T., Bouaziz, L., De Niel, J., Brauer, C., Dewals, B., Drogue, G., Fenicia, F., Grelier, B., Nossent, J., Pereira, F., Savenije, H., Thirel, G., & Willems, P. (2017). Looking beyond general metrics for model comparison – Lessons from an international model intercomparison study. *Hydrology and Earth System Sciences*, *21*(1), 423–440. https://doi.org/10.5194/hess-21-423-2017

Delaigue, O., Génot, B., Lebecherel, L., Brigode, P., & Bourgin, P. (2020). *Database of watershed-scale hydroclimatic observations in France*. https://webgr.inrae.fr/base-de-donnees

Ficchì, A., Perrin, C., & Andréassian, V. (2019). Hydrological modelling at multiple sub-daily time steps: Model improvement via flux-matching. *Journal of Hydrology*, *575*, 1308–1327. https://doi.org/10.1016/j.jhydrol.2019.05.084

Hapuarachchi, H. A. P., Wang, Q. J., & Pagano, T. C. (2011). A review of advances in flash flood forecasting. *Hydrological Processes*, *25*(18), 2771–2784. https://doi.org/10.1002/hyp.8040

Jain, S. K., Mani, P., Jain, S. K., Prakash, P., Singh, V. P., Tullos, D., Kumar, S., Agarwal, S. P., & Dimri, A. P. (2018). A brief review of flood forecasting techniques and their applications. *International Journal of River Basin Management*, *16*(3), 329–344. https://doi.org/10.1080/15715124.2017.1411920

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., & Woods, R. A. (2020). A brief analysis of conceptual model structure uncertainty using 36 models and 559 catchments. *Water Resources Research*, *56*(9), e2019WR025975. https://doi.org/10.1029/2019wr025975

Le Moine, N. (2008). *Le bassin versant de surface vu par le souterrain: une voie d'amélioration des performances et du réalisme des modèles pluie-débit ?* [Doctoral dissertation, UPMC, Cemagref]. https://hal.inrae.fr/tel-02591478

Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., & Ndima, P. O. (2014). La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche*, *100*(1), 25–32. https://doi.org/10.1051/lhb/2014004

Mathevet, T., Michel, C., Andréassian, V., & Perrin, C. (2006). A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. *IAHS Publication*, *307*, 211–219. https://hal.inrae.fr/hal-02588691

McMillan, H., Booker, D., & Cattoën, C. (2016). Validation of a national hydrological model. *Journal of Hydrology*, *541*, 800–815. https://doi.org/10.1016/j.jhydrol.2016.07.043

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., & Teuling, A. J. (2018). Mapping (dis)agreement in hydrologic projections. *Hydrology and Earth System Sciences*, *22*(3), 1775–1791. https://doi.org/10.5194/hess-22-1775-2018

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

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2: Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, *303*(1–4), 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Pagano, T. C., Wood, A. W., Ramos, M. H., Cloke, H. L., Pappenberger, F., Clark, M. P., Cranston, M., Kavetski, D., Mathevet, T., Sorooshian, S., & Verkade, J. S. (2014). Challenges of operational river forecasting. *Journal of Hydrometeorology*, *15*(4), 1692–1707. https://doi.org/10.1175/JHM-D-13-0188.1

Pang, S., Wang, X., Melching, C. S., & Feger, K. H. (2020). Development and testing of a modified SWAT model based on slope condition and precipitation intensity. *Journal of Hydrology*, *588*, 125098. https://doi.org/10.1016/j.jhydrol.2020.125098

Pappenberger, F., Pagano, T. C., Brown, J. D., Alfieri, L., & Lavers, D. A. (2016). Hydrological Ensemble Prediction Systems Around the Globe. In Q. Duan, F. Pappenberger, Thielen, A. Wood, H. L. Cloke, J. C. Schaake, et al. (Eds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 1–35). Springer Berlin Heidelberg.

Pelletier, A. (2021). *Nappes et rivières: la piézométrie peut-elle améliorer la prévision des étiages des cours d'eau ?* [Doctoral dissertation, INRAE Antony, Ecole des Ponts, Sorbonne Université]. https://tel.archives-ouvertes.fr/tel-03783485

Peredo, D. (2021). *Quels gains d'une modélisation hydrologique adaptée et d'une approche d'ensemble pour la prévision des crues rapides ?* [Doctoral dissertation, INRAE Antony, Sorbonne Université]. https://hal.sorbonne-universite.fr/tel-03469456v1

Peredo, D., Ramos, M. H., Andréassian, V., & Oudin, L. (2022). Investigating hydrological model versatility to simulate extreme flood events. *Hydrological Sciences Journal*, *67*(4), 628–645. https://doi.org/10.1080/02626667.2022.2030864

Perrin, C., Andréassian, V., Serna, C. R., Mathevet, T., & Le Moine, N. (2008). Discrete parameterization of hydrological models: Evaluating the use of parameter sets libraries over 900 catchments. *Water Resources Research*, *44*(8), W08447. https://doi.org/10.1029/2007WR006579

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, *279*(1–4), 275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

Roux, H., Labat, D., Garambois, P. A., Maubourguet, M. M., Chorda, J., & Dartus, D. (2011). A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments. *Natural Hazards and Earth System Sciences*, *11*(9), 2567–2582. https://doi.org/10.5194/nhess-11-2567-2011

Sauquet, E., Gottschalk, L., & Krasovskaia, I. (2008). Estimating mean monthly runoff at ungauged locations: An application to France. *Hydrology Research*, *39*(5–6), 403–423. https://doi.org/10.2166/nh.2008.331

Tabary, P., Dupuy, P., L'Henaff, G., Gueguen, C., Moulin, L., Laurantin, O., Merlier, C., & Soubeyroux, J. M. (2012). A 10-year (1997–2006) reanalysis of quantitative precipitation estimation over France: Methodology and first results. *IAHS Publication*, *351*, 255–260. http://hydrologie.org/redbooks/a351/abs_351_0255.pdf

Tangara, M. (2005). *Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global* [Doctoral dissertation, Ecole pratique des hautes études]. https://webgr.inrae.fr/publications/theses

Troin, M., Arsenault, R., Wood, A. W., Brissette, F., & Martel, J. L. (2021). Generating ensemble streamflow forecasts: A review of methods and approaches over the past 40 years. *Water Resources Research*, *57*(7), e2020WR028392. https://doi.org/10.1029/2020wr028392

Viatgé, J., Pinna, T., Ficchì, A., Perrin, C., Dorchies, D., Garandeau, L., & Tilmant, F. (2019). Vers une plus grande flexibilité temporelle du modèle opérationnel de prévision des crues GRP. *La Houille Blanche*, *105*(2), 72–80. https://doi.org/10.1051/lhb/2019017

Vidal, J. P., Martin, E., Franchistéguy, L., Baillon, M., & Soubeyroux, J. M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, *30*(11), 1627–1644. https://doi.org/10.1002/joc.2003

Wu, W., Emerton, R., Duan, Q., Wood, A. W., Wetterhall, F., & Robertson, D. E. (2020). Ensemble flood forecasting: Current status and future opportunities. *WIREs Water*, *7*(3). https://doi.org/10.1002/wat2.1432

Zanchetta, A., & Coulibaly, P. (2020). Recent advances in real-time pluvial flash flood forecasting. *Water*, *12*(2), 570. https://doi.org/10.3390/w12020570