Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments

Guillaume Thirel, Vazken Andréassian, Charles Perrin, Jean-Nicolas Audouy, Lionel Berthet, Pamela Edwards, Nathalie Folton, Carina Furusho-Percot, Anna Kuentz, Julien Lerat, et al.

Hydrological Sciences Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/thsj20

Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments

G. Thirel[a], V. Andréassian[a], C. Perrin[a], J.-N. Audouy[b], L. Berthet[b], P. Edwards[c], N. Folton[d], C. Furusho[a], A. Kuentz[aef], J. Lerat[g], G. Lindström[h], E. Martin[i], T. Mathevet[e], R. Merz[j], J. Parajka[k], D. Ruelland[l] & J. Vaze[m]

[a] Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), Antony, France
[b] Allier Basin Flood Forecasting Centre (DREAL d'Auvergne), Clermont-Ferrand, France
[c] Northern Research Station, USDA Forest Service, Parsons, West Virginia, 26287 USA
[d] OHAX Research Unit, Irstea, Aix-en-Provence, France
[e] EDF DTG, Grenoble, France
[f] LTHE, Grenoble, France
[g] Bureau of Meteorology, Canberra, Australia
[h] SMHI, Norrköping, Sweden
[i] CNRM-GAME, Météo-France, CNRS, Toulouse, France
[j] Department Catchment Hydrology, UFZ, Germany
[k] TU Vienna, Vienna, Austria
[l] CNRS, HydroSciences Montpellier, Montpellier, France
[m] Land and Water Flagship, CSIRO, Canberra, Australia

Accepted author version posted online: 17 Oct 2014.

To cite this article: G. Thirel, V. Andréassian, C. Perrin, J.-N. Audouy, L. Berthet, P. Edwards, N. Folton, C. Furusho, A. Kuentz, J. Lerat, G. Lindström, E. Martin, T. Mathevet, R. Merz, J. Parajka, D. Ruelland & J. Vaze (2014): Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments, Hydrological Sciences Journal, DOI: 10.1080/02626667.2014.967248

To link to this article: http://dx.doi.org/10.1080/02626667.2014.967248

Publisher: Taylor & Francis & IAHS Press

Journal: *Hydrological Sciences Journal*

DOI: 10.1080/02626667.2014.967248

Hydrology under change: An evaluation protocol to investigate how hydrological models deal with changing catchments

G. Thirel[1], V. Andréassian[1], C. Perrin[1], J.-N. Audouy[2], L. Berthet[2], P. Edwards[3], N. Folton[4], C. Furusho[1], A. Kuentz[1,5,6], J. Lerat[7], G. Lindström[8], E. Martin[9], T. Mathevet[5], R. Merz[10], J. Parajka[11], D. Ruelland[12] & J. Vaze[13]

1: Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), Antony, France

2: Allier Basin Flood Forecasting Centre (DREAL d'Auvergne), Clermont-Ferrand, France.

3: USDA Forest Service, Northern Research Station, Parsons, WV, 26287 USA

4: OHAX Research Unit, Irstea, Aix-en-Provence, France

5: EDF DTG, Grenoble, France

6: LTHE, Grenoble, France

7: Bureau of Meteorology, Canberra, Australia

8: SMHI, 60176 Norrköping, Sweden

9: CNRM-GAME, Météo-France, CNRS, Toulouse, France

10: Department Catchment Hydrology, UFZ, Germany

11: TU Vienna, Vienna, Austria

12: CNRS, HydroSciences Montpellier, Montpellier, France

13: Land and Water Flagship, CSIRO, Canberra, Australia

Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), Antony, France

guillaume.thirel@irstea.fr

Abstract: Testing hydrological models under changing conditions is essential to evaluate their ability to cope with changing catchments and their suitability for impact studies. With this perspective in mind, a workshop dedicated to this issue was held at the 2013 IAHS General Assembly in Göteborg, Sweden, in July 2013, in which the results of a common testing experiment were presented. Prior to the workshop, the participants had been invited to test their own models on a common set of basins showing varying conditions specifically set up for the workshop. All these

basins experienced changes, either in physical characteristics (e.g. changes in land cover) or climate conditions (e.g. gradual temperature increase). This article presents the motivations and organisation of this experiment – that is – the testing (calibration and evaluation) protocol and the common framework of statistical procedures and graphical tools used to assess the model performances. The basins datasets are also briefly introduced (a detailed description is provided within the associated supplementary material).

Key words non stationarity; IAHS workshop; calibration; evaluation; protocol; split-sample test; changing catchments dataset

Hydrologie sous changement : Un protocole d'évaluation pour examiner comment les modèles hydrologiques s'accommodent des bassins changeants

Résumé : Tester les modèles hydrologiques pour des conditions changeantes est essentiel pour évaluer leur capacité à faire face à des bassins changeants et leur pertinence pour les études d'impact. Un atelier, dédié à cette problématique et pendant lequel les résultats d'une expérimentation de tests conjoints ont été présentés, s'est tenu dans cette perspective en juillet 2013 à Göteborg lors de l'Assemblée Générale 2013 de l'AISH. Avant l'atelier, les participants avaient été invités à tester leurs propres modèles sur un jeu commun de bassins qui montraient des conditions changeantes et qui avait été spécifiquement préparé pour cet atelier. Tous ces bassins ont subi des changements, soit de leurs caractéristiques physiques (par exemple l'évolution de la couverture du sol), soit des conditions climatiques (par exemple une augmentation graduelle de la température). Cet article présente les motivations et l'organisation de cette expérimentation, c'est-à-dire le protocole de tests (calage et évaluation) et la manière dont les résultats ont été analysés dans un cadre commun utilisant des mesures statistiques d'efficacité et des outils graphiques. Les jeux de données des bassins sont également brièvement présentés (une description détaillée est fournie dans le document complémentaire associé à cet article).

Mots-clés non stationnarité; atelier AISH; calage; évaluation; protocole; test de calage-contrôle; jeu de données de bassins changeants

1 MOTIVATION AND SCOPE

1.1 On changing catchments

Better understanding how hydrological systems respond to changing conditions is a key question in the scientific community, which may help improving the prediction of the impacts of various future environmental changes (Peel and Blöschl, 2011).

However, the notion of change through time is relative, and depends on the time window that is considered (Koutsoyiannis, 2013). "Change" may have several meanings for catchments. Here we define a changing catchment as one that undergoes a significant change in land cover or climate conditions (a systematic deviation that is outside the range of the historic record). In addition, water management changes from the construction of a storage dam are also considered, but other human-induced changes that can have a significant impact on flow dynamics (e.g. water withdrawals, or dike construction, or streambed gravel mining) are not included in the analysis. Note that a change in land cover or climate does not necessarily imply a significant change in the hydrological behaviour of the catchment.

1.2 Should we fear changing catchments?

During the early days of hydrological modelling (in the 1960s and 1970s), there was a belief that computers eventually would be able to solve all arising problems, so there was little concern about changing catchments. Linsley (1982), who created one of the first modern hydrological models at Stanford University, considered that a good generic model should be able to adapt to the breadth of catchment conditions and events, provided that it would "represent the various processes with sufficient fidelity so that irrelevant processes can be 'shut off' or will simply not function." In those years, the idea was that calibration was a safe way to parameterise models so that they would adequately reproduce the catchment behaviour.

These happy days ended with increasing doubts on calibration. Already in the early 1970s, Johnston and Pilgrim (1973) had described the numerous disappointments caused by an extensive search for the optimum values of the parameters of Boughton's catchment model. They developed a comprehensive list of problems that have since been recognised as the major impediments to the calibration of hydrological models, including discontinuities of the response surface, multiplicity of equivalent optima, un-identifiability, and lack of robustness of calibrated parameter values. Eventually it became obvious that improvement in search algorithms could reduce the *numerical* miscalibration problems but the *hydrological* overcalibration problems remained to be addressed (Andréassian *et al.*, 2012). Obviously, model calibration could "work to accommodate reality, often in a subtle way" (Beven, 1977), but came to the cost of the model's predictive capacity.

When, on a perfectly steady catchment, parameters can change for (apparently) no reason from one calibration period to another, it seems unlikely that we can one day address the issue of fully representing responses in changing catchments. To be able to extrapolate to other climates or other land uses, it is essential to understand the causes of the dynamic behaviour of the catchment; otherwise the model will remain a "mathematical marionette," only able to "dance to a tune it has already heard" (Kirchner, 2006).

1.3 Towards solutions for dealing with change

The PUB (Prediction on Ungauged Basins) decade (2003-2012; Hrachowitz *et al.*, 2013), in combination with the initiation of the Panta Rhei decade (Montanari *et al.*, 2013), gave a new impetus to attempts to assess the extrapolation capacity of hydrological models in space and time. As Boughton (2005) put it, "the problem of estimating runoff from ungauged catchments […] is closely related to the problem of estimating the change in runoff that will occur when the land use of a catchment changes." Recent studies (Seibert, 2003; Coron *et al.*, 2011; Merz *et al.*, 2011; Coron *et al.*, 2012; Refsgaard *et al.*, 2013) have attempted to distinguish *apparent* changes, i.e. changes in observable properties like land cover and climate conditions (but that may not necessarily induce hydrological changes), from *intrinsic* changes, i.e. changes in the catchment hydrological behaviour (understood as an input-output relationship), which may be difficult to identify or attribute to specific causes on the sole basis of observations and that require the use of some kind of hydrological model.

However, we believe that hydrological modellers need to strive to respond collectively to the key growing questions in the challenging research field of hydrology under change (Peel and Blöschl, 2011). Towards this end, a workshop was held during the 2013 International Association of Hydrological Sciences (IAHS) General Assembly in Göteborg, Sweden, to foster the development of relevant hydrological parameterization methods for application to changing catchments. Before this workshop, the participants were asked to carry out calibrations of their models over successive and contrasted test periods. An evaluation protocol was defined, and a set of metrics was proposed as a means for analysing models' performance and parameter transferability. To promote comparability, we gathered

data sets from changing catchments and asked the modellers to use them preferentially as test beds. With the proposed framework, the participants were asked to contribute answers to the following questions during the Göteborg workshop:

- Can changes in model parameters calibrated over different periods tells us whether a catchment is changing, or are there too many numerical artefacts to answer this question (due to poor parameter identifiability, model overparameterization, etc.)?
- Are our models robust and/or flexible enough to be used under changing conditions?
- What approaches should be tried in the coming years to better handle hydrological modelling under change?

These questions guided the searches for improving model results on changing catchments, as reported in this special issue of the *Hydrological Sciences Journal*.

This article aims at presenting the protocol (Section 2) and the numerical graphical tools elaborated for analysing the model results (Section 3). The data sets of changing catchments are briefly introduced in this paper (Section 4) and described in more details in the supplementary material.

2 CALIBRATION AND EVALUATION PROTOCOL

2.1 On differential split-sample testing (DSST)

Due to the difficulty of exhaustively describing all the physical processes involved in the transformation of precipitation into streamflow, many hydrological models of a wide range of complexity have been proposed in the literature. These models require the calibration of several free parameters against discharge observations, because these parameters often cannot be linked to physical characteristics directly (Abebe *et al.*, 2010). However, calibration may be unable to reach a good representation of streamflow in all conditions, whatever the quantity of data used for calibration (Andréassian *et al.*, 2012), and may show unstable behaviour between different calibration periods. The modeller should try to solve this problem by seeking ways to obtain stable parameters.

The diagnosis of model stability is possible through testing procedures that have been proposed for assessing the dependence of model parameters to climate or

other potentially changing factors. The most famous example, which is also the most frequently used method, is the Split-Sample Test (SST) proposed by Klemeš (1986) within a hierarchical scheme for systematic testing of hydrological models. The temporal transposability of models is assessed through cross-calibration and validation tests of the models over two different periods. The even more demanding Differential Split-Sample Test (DSST) described by Klemeš (1986) seems particularly relevant for evaluating models under changing conditions. Within the DSST, the model should be calibrated and validated over contrasted periods: for example, calibrated over a dry period and validated during a wet period. The Klemes testing scheme also includes the evaluation of model spatial transposability by testing neighbouring catchments (proxy-basin test), possibly with contrasted conditions (differential split-sample proxy-basin test) (Klemeš, 1986; Klemeš, 2009).

In many cases, models fail when the most demanding test described above is applied, and for this reason, the full Klemes testing scheme is seldom applied (Andreassian *et al.*, 2009; Klemeš, 2009). Actually, only few studies applied it (Refsgaard and Knudsen, 1996; Donnelly-Makowecki and Moore, 1999; Stisen *et al.*, 2012). However, the less-demanding DSST still seems a suitable tool to evaluate models under changing conditions. For this reason, the protocol described in Section 2.2 is based on this test. The DSST is now more commonly used by hydrologists (Seibert, 2003; Vaze *et al.*, 2010; Brigode *et al.*, 2013; Refsgaard *et al.*, 2013) and has even been generalised by Coron *et al.* (2012) who proposed to use multiple test periods defined by a time-sliding window over the full available period. This testing approach also relates to other recent investigations on model parameter stability in time (see e.g., Gharari *et al.*, 2013; Razavi and Tolson, 2013).

2.2 Calibration and evaluation periods

For each data set six calibration periods were defined. The first period, called the "complete period", was set in such a way that it made use of as much data as possible. At least two years were kept at the beginning of the "complete period" for model warm-up. Calibration on the "complete period" represents Level 1 of the protocol. Since this test level does not include proper model validation, it should be considered as a preliminary step for modellers to check the general model suitability for the studied catchment. Then five sub-periods, nested within the "complete period" and

called P1, P2, P3, P4 and P5, were defined. These five periods have the same number of years and were chosen to cover the span of the "complete period" as fully as possible. Calibration on the five sub-periods represents Level 2 of the protocol. The details of the six test periods ("complete period" plus P1–P5) are defined in the supplementary material for each test catchment.

The modellers were asked to calibrate their hydrological models on each test period and to make simulations over the complete period using each of the six parameter sets successively. Simulations performance then was analysed over the six periods. The principle of this protocol is summarised in Fig. 1. Modellers were asked to calculate performance criteria on the "complete period" as well as on each of the five sub-periods to obtain common bases for the evaluation of models simulations.

2.3 Evaluation criteria

Prior to the workshop, a list of statistical scores was established, which are described in Table 1. We classified the scores into two groups:

- primary scores, which we considered to be the basis for a careful hydrological evaluation, since they evaluate simple expected properties in terms of water balance, representation of high- or low-flow, etc.;
- secondary scores, the examination of which was not mandatory, but which these can give valuable information on how well the models could handle the proposed changing catchments.

Primary scores

The Nash-Sutcliffe efficiency criterion (NSE; Nash and Sutcliffe, 1970) is a quadratic score commonly used to assess the quality of hydrological models. NSE ranges from $-\infty$ to 1, with negative values when the model performance falls below that of a simple model that would simulate a constant value equal to the average of the observed discharges. The maximum score of 1 means that the simulated discharges equal the observed discharges. Due to its quadratic nature, the NSE principally gives information about the capacity of the model to simulate high flows. For this reason, we also calculated a modified version of the NSE (symbolised NSE_{LF}) calculated with inverse transformed flows $1/(Q+\epsilon)$, where Q is the flow and ϵ is a small constant

introduced to avoid divisions by 0 in case of zero flows. Here ϵ was set to one-hundredth of the mean observed flows ($\epsilon = \mu(Q_{obs})/100$). This score has proven to be adequate for low flow estimation (Pushpalatha *et al.*, 2012).

Bias compares the mean simulated discharges and the mean observed discharges over the test period. Here the bias was computed as the ratio between mean simulated discharges and mean observed discharges. When model simulations are not biased, the bias value is 1. Values lower than 1 indicate underestimation of the mean discharge, while values larger than 1 indicate overestimation. The evolution of this criterion over contrasted periods often provides useful information on a model's capacity to reproduce the water balance over time.

Secondary scores

One model can perform well for a specific range of discharges (e.g., low flows), but not for the other magnitudes of discharges (e.g., high flows). NSE and NSE_{LF} criteria already give an indication of this kind of behaviour, but the distribution of the simulated discharges can give additional information. This is especially the case when an infrequent event occurs, because these strongly influence the NSE or NSE_{LF} values, due to their quadratic formulation. Visually comparing the simulated Flow Duration Curves (FDCs) with the observed FDC is one way to compare the distributions (see Section 3). Four specific quantiles, focusing on high and low flows, i.e. 95% ($Q_{0.95}$), 85% ($Q_{0.85}$), 15% ($Q_{0.15}$) and 5% ($Q_{0.05}$) quantiles, are also calculated. $Q_{0.95}$ represents the discharge value *not* exceeded 95% of the time (i.e. it is a high-flow indicator), while $Q_{0.05}$ represents the discharge value exceeded 95% of time (i.e. it is a low-flow indicator). Comparing simulated and observed quantiles over different periods gives an indication of how close the modelled flow distribution is to the observed distribution. However, in our data set some intermittent catchments are present. For them null discharges are quite common and the observed $Q_{0.05}$ and $Q_{0.15}$ are equal to 0; this also may be the case for the simulated quantiles. For this reason, the frequency of days with simulated discharges lower than 5% of the average observed discharge ($Freq_{LF}$) provides an alternative. The same formula can be applied with Q_{obs} instead of Q_{sim} so the frequency of low flows can be compared for observed and simulated discharges.

Although the NSE is a widely used criterion among hydrologists, due to representation of periodicity, NSE may take high values thus providing misleading conclusions about the actual model capacity, for example for snowmelt-driven catchments or for catchments facing strong seasonal meteorological graients.. As a result, some authors have suggested modified versions of NSE (Mathevet *et al.*, 2006; Criss and Winston, 2008). The NSE can be decomposed as a sum of different criteria (Murphy, 1988; Węglarczyk, 1998), which makes it easier to understand the origination of poor model performance. In this context, the three components (correlation coefficient r, ratio of standard deviations α_{NSE} and relative bias β_{NSE}) of the NSE decomposition proposed by Gupta *et al.* (2009) in their Equation 4 (see Table 1) were used. The ideal values for the three elements of the NSE decomposition are $r = 1$, $\alpha_{NSE} = 1$ and $\beta_{NSE} = 0$. Due to the double appearance of α_{NSE} in the NSE decomposition, Gupta *et al.* (2009) showed that the maximum NSE is obtained when $\alpha_{NSE} = r$. As a result, the α_{NSE} value selected for optimising the NSE leads to an underestimation of the variability of simulated flows. To overcome this problem, Gupta *et al.* (2009) proposed a new criterion, the Kling-Gupta efficiency (KGE), later slightly modified by Kling *et al.* (2012) (and called KGE′) to avoid interactions between the bias and variation coefficient ratios. In the proposed protocol, we decided to use this modified criterion (KGE′) and its decomposition. The NSE decomposition, the KGE′ and its decomposition (which is different from NSE's one) can be computed for low flows in the same way as NSE_{LF} is calculated from NSE.

3 GRAPHICAL ANALYSIS

Several types of plots were suggested by the workshop organisers to help modellers analyse their results through a common framework. Here we provide a detailed description of these graphical tools. Having a common graphical representation will indeed help readers to compare the results of different studies. In the following, the graphs describe the scores calculated on two time spans: a sub-period-based time span (Section 3.1) and an annual time span (Section 3.2). Presenting the scores described in Section 2.3 over these two time spans provide a complementary analysis of the results. While the plots given on sub-periods provide a means to analyse general trends over periods of several years, the annual scores highlight the effect of specific years on the scores. Most of the plots presented in this manuscript comprise not only

the scores over each evaluation period (i.e. the sub-periods or every year), but also the scores for each calibration made by the modellers. This two-dimensional analysis is necessary to analyse both the effect of the calibration period and the evaluation period on scores.

For both analyses, two ways of representing the results were proposed in most cases: classical curve-based plots and two-dimensional tables. Both representations can identify good or poor model performance. One difference is, however, the possibility of easily visualising the scores' stability across periods when using the curve-based plots (we remind that the stability of scores is an indicator of temporal transferability of hydrological parameters). The two-dimensional tables ease the identification of the time distance between the calibration and the evaluation period.

3.1 Sub-periods analysis

Recall, the "complete period", which is the longest period we could use for each catchment, was split into five sub-periods for both the calibration and the evaluation of the models (see Section 2.2). In other words, six simulations were made with each model (one simulation for each calibration) and each of them can be analysed over six periods. In this paper we present the plots obtained for an evaluation of the six simulations on the five sub-periods and on the "complete period" together.

The NSE, KGE′, their decompositions, and the bias could be represented using the type of graph shown in Fig. 2, where six coloured curves are drawn: each curve corresponds to the criterion values over the evaluation periods for a single model calibration (NSE is considered in Fig. 2). The x-axis presents six ticks: each of them is a NSE value for an evaluation on a given sub-period or on the "complete period". The y-axis actually represents the NSE values. The standard deviation of each curve is also given (see SD in the legend) as an additional indicator of the stability of the score value of the model simulations. However, providing a good simulation over a period is not sufficient to define a good model for simulating changing catchments: it is also necessary to reproduce this good performance for all evaluation periods. Please note that for the bias plot (not shown here), the "$y = 1$" line (i.e. the perfect value) is also drawn, represented by a dashed line.

Another representation of the same scores consisted in using two-dimensional table plots (see Fig. 3 for an example on NSE_{LF}). Here the score values are not given by

curves, but by coloured squares within two-dimensional tables. In these tables, a row corresponds to an evaluation period and a column to a calibration period. In other words, if one wishes to assess the impact of the period used for the model calibration on a score over one of the evaluation periods in Fig. 3, the row corresponding to this evaluation period will give that information. By contrast, for a single calibration, to check the change in the score versus the evaluation period, it is necessary to read the table in columns. To simplify interpretation of these tables, the squares for which the evaluation period is included in the calibration period are outlined in black (elements of the diagonal plus the first column). Figure 2 makes it easier to identify parameter sets that are stable over time, and Fig. 3 makes it easier to identify the impact of the distance between the calibration and evaluation periods considered.

The different quantiles ($Q_{0.95}$, $Q_{0.85}$, $Q_{0.15}$, and $Q_{0.05}$) and the $Freq_{LF}$ also can be represented through the graph of Fig. 2. The only information added to these plots is a curve for the observed discharge score. This score is represented by a dashed black line in Fig. 4. Since the objective of these criteria is not to be as stable and as good as possible, but to be as close as possible to the observations, the standard deviation metric is replaced here in the legend with the correlation between each simulated curve and the observed curve. This type of representation includes one graph for each quantile considered or for $Freq_{LF}$, and that each graph shows all the evaluation periods. The plots of monthly discharge regimes include one graph for each evaluation period.

The quantiles and the $Freq_{LF}$ can also be drawn with two-dimensional tables. The only difference compared to the NSE two-dimensional tables is the addition of a column for observations (e.g., see the first column "Obs." in Fig. 5). We did not draw the discharge regime curves (mean monthly values) in this way.

Simulated and observed FDCs are plotted as shown in Fig. 6. In this graph, each coloured line is obtained from a different model calibration, and the observed FDC is given with a dashed curve. The discharge values (m^3/s) are plotted against their frequency of exceedance. In total, six similar graphs are plotted, one for each evaluation period. There is no two-dimensional table version of these graphs.

3.2 Annual analysis

In addition to examining five contrasted sub-periods plus the "complete period", we decided to provide a more detailed graphical analysis computed on an annual basis to show how changes in scores can highlight sudden changes or outlier years. It should be remembered that model calibration was not different than for the plots presented in Section 3.1 (also based on the "complete period" plus the P1–P5 periods). No further calibration (e.g. on individual years) was performed by the modellers.

Each of the graphs presented on the sub-periods could be extended easily to the annual basis. An example of the extension of Fig. 2 to an annual basis is presented in Fig. 7: the six curves (i.e. the six different model calibrations) are plotted, but the x-axis, which gives the evaluation periods, is extended to each year of the "complete period". Note that the values of the score for the "complete period" are still indicated as the first element of the x-axis. Combining Fig. 7 and Fig. 2 shows that the poor performance of calibrations on P1 and P2 (blue and green curves) on the P3–P5 periods are not due to unusually poor event simulations biasing the NSE metrics, but rather to a systematic deficiency of the simulations (this is shown by the low NSE values for P1 and P2 calibrations from 1979 to 2008, which include the P3–P5 periods).

Figure 8 is the extension of Fig. 3 on an annual basis. Here there is one row for each evaluation year (plus one for the "complete period"). Since the extension on an annual basis of Fig. 4 is similar to the extension of Fig. 2, and the extension of Fig. 5 is similar to the extension of Fig. 3, they are not shown here. For Fig. 6, instead of having six different FDC plots, we provided as many plots as years in the "complete period" (plus one plot for the "complete period").

3.3 Comparing the models

From the plots presented in the previous sections, the modellers received a substantial amount of information available for analysing their simulations. However, we did not wish to limit this workshop to a diagnosis of hydrological models illnesses. We also wished to encourage the modellers to rid their models of the deficiencies identified and propose solutions (e.g. by explicitly accounting for the causes of change in their models, see e.g. Nalbantis *et al.*, 2011). Testing the possible solutions to remedy the models required having a simple visual way to check whether these solutions

improved or deteriorated model efficiency. Therefore, we also developed a way to represent the differences between the scores of several models or versions of models. The comparison graphs proposed to the modellers were of the two-dimensional table type. An example is given in Fig. 9 for a comparison of the NSE_{LF} values for two models, called "Model 1" and "Model 2". Since the structure of this representation is quite complex, a detailed description is provided:

1. First, note that on this graph, no NSE_{LF} values are actually given. Here the colour of the small squares corresponds to the difference between the NSE_{LF} values of two models. The upper right-hand block of the graph indicates NSE_{LF} for Model 2 minus NSE_{LF} for Model 1, and the lower left-hand block of the graph is NSE_{LF} for Model 1 minus NSE_{LF} for Model 2). In the colour scale, blue represents positive differences, white indicates that the two scores are very similar, and red denotes negative differences. Thus, the small blue squares drawn in the upper right-hand block of the graph indicate that Model 2 performs better than Model 1, and the small red squares indicate that Model 1 performs better than Model 2. For the lower left-hand block, the colours have the opposite meaning.
2. The blocks on the diagonal are filled with small grey squares, because these two blocks in fact compare one model with itself. Note that there is redundant information and that only one-fourth of these graphs would be sufficient. Nevertheless, as explained below, this graph could be extended, if needed, to compare more than two models or to include observations.

Now that the definition of the information content of this type of graph has been given, it is essential to explain how to read it correctly. The first way is of course to check the colour trends of the blocks: if one of the blocks is dominated by blue or red squares, then one of the models outperforms the other one.

Due to its two-dimensional structure, this table also plot can answer these two questions:

1. For a given calibration period, how do the two models compare when evaluated over different evaluation periods?
2. For a given evaluation period, how the two models compare when calibrated over different calibration periods?

Point (1) can be addressed by analysing the columns in Fig. 9: for example, let us imagine that we wish to compare how both models perform when they are

calibrated over P2. To do so, we have to look at the P2 column in Fig. 9 (either on the lower left-hand block or on the upper right-hand block). From the study of this column, it is clear that Model 1 outperforms Model 2 for all evaluation periods, because this column is globally blue (lower left-hand block) or red (upper right-hand block). Only when the models are evaluated over P2 do the models seem to have closer performance. This is an indication that both models have a similar performance over the period that is used for their calibration, but that the performance of Model 2 collapses when it is evaluated over another period (i.e. Model 2 is not robust when it is calibrated on P2).

Let us now focus on the second (2) type of comparison: for example, how the models behave on P3. For this, it is sufficient to study one of the P3 rows (let us take the one belonging to the upper right-hand block). From this row, it is clear that both models have a similar performance when the models are calibrated on P1 (white square), that Model 1 performs better when the models are calibrated on P2 (red square), and that Model 2 performs better when they are calibrated over the other periods (blue squares). This may indicate that the P2 calibration of Model 2 fails on P3 for some reason.

For some metrics, such as the quantiles and Freq_{LF}, a comparison with the observations is necessary. For these metrics, the 2×2 block table (e.g., Fig. 9) becomes a 3×3 block table (see Fig. 10 for $Q_{0.95}$). Here again in Fig. 10, only the differences between $Q_{0.95}$ values are given (and not $Q_{0.95}$ values). When using this type of table, direct comparisons between each model and the observations are allowed and are not limited to comparisons between models. The approach to interpreting Fig. 10 is similar to that for Fig. 9 (i.e. for evaluating a single calibration or evaluating several calibrations over all the periods), so it is not further detailed here. Again, it could be argued that instead of using a 3×3 block table, we could have limited it to a 2×2 block matrix (the four blocks of the upper right-hand part or the four blocks of the lower left-hand part) because of the reverse-symmetry of the complete table. However, this version proved to be much easier to explain to users.

4 THE BASINS DATASET

Assembling a representative dataset of changing catchments was necessary for the success of this experiment. The dataset provided to the workshop participants consisted of 14 catchments that experienced apparent changes. The locations of the catchments are shown in Fig. 11, while their main characteristics are given in Table 2. Each catchment and its corresponding dataset are described in details in the online supplementary material provided with this article. Summaries of mean monthly precipitation and streamflow data are given in Fig. 12. The hydrological regimes range from rain-fed (Axe Creek, Rivers Bani, Flinders, Gilbert, Rimbaud and Wimmera) to snow-fed or pluvio-nival (Rivers Allier, Durance, Garonne, Watershed 6 of the Fernow Experimental Forest, and Blackberry, Ferson and Obyån Creeks).

The variety of changes affecting the 14 catchments offers a wide test bed to evaluate the capacity of hydrological models to deal with these changes. Several types of changes are apparent. Temperature increase is the main factor of change in hilly or mountainous parts of Europe. This was the case for the Rivers Garonne, Allier, Durance and Kamp. In the River Allier, a secondtype of change is the construction of a dam for sustaining low flows. Afforestation, deforestation and reforestation also modify the response of catchments in a substantial way. This type of land cover modification affected the River Rimbaud (forest fire), Obyån Creek (Gudrun storm destroying the forest) and Watershed 6 of the Fernow Experimental Forest (deforestation of the native deciduous forest followed by planting and establishment of conifers). Another type of land cover change is urbanization. Two American catchments, Blackberry Creek and Ferson Creek, were subject to such changes. The case of prolonged droughts periods (extending over several years), which are known to affect the calibration of hydrological models, is illustrated by the "Millenium Drought" (River Wimmera and Axe Creek) and the West-Africa drought (River Bani). Finally, two semi-tropical basins (Rivers Gilbert and Flinders) were included in the dataset. The highly variable regimes of these catchments is challenging for hydrological modeling.

Meteorological data (precipitation, temperature, potential evapotranspiration) were provided to the participants as lumped values at the basin scale (no spatially-distributed data were available). These meteorological data were obtained from different sources, including point data from a single meteorological station, spatially

interpolated data from several meteorological stations, mix of meteorological stations and meteorological model reanalysis, and even climate reconstruction. Discharge data were provided for the outlet of each basin. All data were at the daily time step, except for. the Rimbaud River, for which hourly data were available (the hourly time step is more adequate to simulate this basin). The detailed descriptions of all the datasets are included in the online supplementary material.

5 CONCLUSIONS

Improving our capacity to work with changing catchments is a necessity for a variety of hydrological applications. Despite the efforts of modellers over the past few years, no joint action had been initiated previously to specifically address this issue on common basin data sets with different hydrological models applied to a common calibration and evaluation framework. The exercise conducted during the 2013 IAHS workshop was an attempt to address this deficiency. With the methodological framework described in this paper, we aimed to provide the basis for a stimulating and participative workshop, with a full set of criteria, including graphical metrics, specifically designed to easily analyse the stability properties of model simulation over multiple-period tests. This testing and analysis framework was applied to the set of catchments selected for the workshop. We hope that this attempt to rally the community of hydrological modellers around common questions of prediction under change will stimulate new initiatives during the start of the *Panta Rhei* decade.

Acknowledgements The authors would like to thank Francis Chiew from CSIRO for discussions to define the calibration and evaluation protocol and Laurent Coron (Irstea) for fruitful discussions on this article. We also thank IAHS and its STAHY (ICSH) and surface water (ICSW) commissions for their support in organising this workshop. Peter Krause and an anonymous reviewer, as well as the Associate Editor Andreas Efstratiadis, are thanked for comments that improved this manuscript and the supplementary material containing the catchments and datasets descriptions.

References

Abebe, N. A., Ogden, F. L. and Pradhan, N. R., 2010. Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: Implications for parameter estimation. *Journal of Hydrology,* 389(3–4): 301-310, DOI: 10.1016/j.jhydrol.2010.06.007.

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M. H., Oudin, L., Mathevet, T., Lerat, J. and Berthet, L., 2012. All that glitters is not gold: The case of calibrating hydrological models. *Hydrological Processes,* 26(14): 2206-2210, DOI: 10.1002/hyp.9264.

Andreassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M. H. and Valery, A., 2009. HESS Opinions 'Crash tests for a standardized evaluation of hydrological models'. *Hydrology and Earth System Sciences,* 13(10): 1757-1764, DOI: 10.5194/hess-13-1757-2009.

Beven, K., 1977. Response to General Reporter's comments on papers by Beven and Beven and Kirkby, Water Resources Publications: Littleton, CO. H. J. Morel-Seytoux, Salas, J.D., Sanders, T.G., Smith, R.E., 221-222.

Boughton, W., 2005. Catchment water balance modelling in Australia 1960-2004. 71, Agricultural Water Management, 91-116, DOI: 10.1016/j.agwat.2004.10.012.

Brigode, P., Oudin, L. and Perrin, C., 2013. Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology,* 476: 410-425, DOI: 10.1016/j.jhydrol.2012.11.012.

Coron, L., Andreassian, V., Bourqui, M., Perrin, C. and Hendrickx, F., 2011. Pathologies of hydrological models used in changing climatic conditions: a review. Hydro-Climatology: Variability and Change. S. W. Franks, E. Boegh, E. Blyth, D. M. Hannah and K. K. Yilmaz. 344: 39-44.

Coron, L., Andreassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M. and Hendrickx, F., 2012. Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resources Research,* 48(5), DOI: 10.1029/2011wr011721.

Criss, R. E. and Winston, W. E., 2008. Do Nash values have value? Discussion and alternate proposals. *Hydrological Processes,* 22(14): 2723-2725, DOI: 10.1002/hyp.7072.

Donnelly-Makowecki, L. M. and Moore, R. D., 1999. Hierarchical testing of three rainfall-runoff models in small forested catchments. *Journal of Hydrology,* 219(3-4): 136-152.

Gharari, S., Hrachowitz, M., Fenicia, F. and Savenije, H. H. G., 2013. An approach to identify time consistent model parameters: Sub-period calibration. *Hydrology and Earth System Sciences,* 17(1): 149-161.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology,* 377(1-2): 80-91, DOI: 10.1016/j.jhydrol.2009.08.003.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. and Cudennec, C., 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal,* 58(6): 1198-1255, DOI: 10.1080/02626667.2013.803183.

Johnston, P. R. and Pilgrim, D. H., 1973. A study of parameter optimisation for a rainfall-runoff model, Kensington, N.S.W. : School of Civil Engineering, University of New South Wales, 197p.

Kirchner, J. W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research,* 42(3), DOI: 10.1029/2005WR004362.

Klemeš, V., 1986. OPERATIONAL TESTING OF HYDROLOGICAL SIMULATION-MODELS. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques,* 31(1): 13-24, DOI: 10.1080/02626668609491024.

Klemeš, V., 2009 "Interactive comment on Hydrol. Earth Syst. Sci. Discuss.", http://www.hydrol-earth-syst-sci-discuss.net/6/C1069/2009.

Kling, H., Fuchs, M. and Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology,* 424-425: 264-277, DOI: 10.1016/j.jhydrol.2012.01.011.

Koutsoyiannis, D., 2013. Hydrology and change. *Hydrological Sciences Journal,* 58(6): 1177-1197.

Linsley, R. K., 1982. Hydrology for Engineers (Water Resources & Environmental Engineering), McGraw-Hill Education (ISE Editions) 512.

Mathevet, T., Michel, C., Andreassian, V. and Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. Large sample basin experiments for hydrological model parameterization: results of the model parameter experiment, IAHS Publication - Series Of Proceedings And Reports (Red Book Series)(307): 211-219.

Merz, R., Parajka, J. and Blöschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research,* 47(2), DOI: 10.1029/2010WR009505.

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S. J., Di Baldassarre, G., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D. A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z. and Belyaev, V., 2013. "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022. *Hydrological Sciences Journal,* 58(6): 1256-1275, DOI: 10.1080/02626667.2013.809088.

Murphy, A. H., 1988. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Monthly Weather Review,* 116(12): 2417-2424.

Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M. and Koutsoyiannis, D., 2011. Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems. *Hydrology and Earth System Sciences,* 15(3): 743-758.

Nash, J. E. and Sutcliffe, J. V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology,* 10(3): 282-290.

Peel, M. C. and Blöschl, G., 2011. Hydrological modelling in a changing world. *Progress in Physical Geography,* 35(2): 249-261.

Pushpalatha, R., Perrin, C., Le Moine, N. and Andreassian, V., 2012. A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology,* 420: 171-182, DOI: 10.1016/j.jhydrol.2011.11.055.

Razavi, S. and Tolson, B. A., 2013. An efficient framework for hydrologic model calibration on long data periods. *Water Resources Research,* 49(12): 8418-8431.

Refsgaard, J. C. and Knudsen, J., 1996. Operational validation and intercomparison of different types of hydrological models. *Water Resources Research,* 32(7): 2189-2202.

Refsgaard, J. C., Madsen, H., Andréassian, V., Arnbjerg-Nielsen, K., Davidson, T. A., Drews, M., Hamilton, D. P., Jeppesen, E., Kjellström, E., Olesen, J. E., Sonnenborg, T. O., Trolle, D., Willems, P. and Christensen, J. H., 2013. A framework for testing the ability of models to project climate change and its impacts. *Climatic Change*: 1-12, DOI: 10.1007/s10584-013-0990-2.

Seibert, J., 2003. Reliability of model predictions outside calibration conditions. *Nordic Hydrology,* 34(5): 477-492.

Stisen, S., Hojberg, A. L., Troldborg, L., Refsgaard, J. C., Christensen, B. S. B., Olsen, M. and Henriksen, H. J., 2012. On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes. *Hydrology and Earth System Sciences,* 16(11): 4157-4176.

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R. and Teng, J., 2010. Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology,* 394(3–4): 447-457, DOI: 10.1016/j.jhydrol.2010.09.018.

Węglarczyk, S., 1998. The interdependence and applicability of some statistical quality measures for hydrological models. *Journal of Hydrology,* 206(1-2): 98-103, DOI: 10.1016/S0022-1694(98)00094-8.

TABLES

Table 1 *Definition of the evaluation criteria. n is the number of days of the evaluation period, Q_{sim} and Q_{obs} stand for the simulated and observed discharges, μ is the arithmetic mean over the evaluation period, and σ is the standard deviation. r, α_{NSE}, β_{NSE},* KGE′, *$\beta_{KGE'}$, and $\gamma_{KGE'}$ can be adapted to low flows in the same way as NSE_{LF} is an adaptation of NSE to low flows*

Criterion	**Mathematical formulation**	**Name of the criterion**	**Best value**
NSE	$1 - \frac{\sum_{t=1}^{n}\left(Q_{obs}(t) - Q_{sim}(t)\right)^2}{\sum_{t=1}^{n}(Q_{obs}(t) - \mu(Q_{obs}))^2} = 2.\alpha_{NSE}.r - (\alpha_{NSE})^2 - (\beta_{NSE})^2$	Nash-Sutcliffe efficiency and decomposition	1
NSE_{LF}	$1 - \frac{\sum_{t=1}^{n}(\frac{1}{Q_{obs}(t)+\epsilon} - \frac{1}{Q_{sim}(t)+\epsilon})^2}{\sum_{t=1}^{n}(\frac{1}{Q_{obs}(t)+\epsilon} - \mu(\frac{1}{Q_{obs}+\epsilon}))^2}$	Nash-Sutcliffe efficiency on low flows	1
Bias (or $\beta_{KGE'}$)	$\frac{\mu(Q_{sim})}{\mu(Q_{obs})}$	Bias	1
$Freq_{LF}$	$freq(Q_{sim} < 0.05 * \mu(Q_{obs}))$	Frequency of low flows	Same as for Q_{obs}
r	$\frac{\sum_{t=1}^{n}(Q_{obs}(t) - \mu(Q_{obs})).(Q_{sim}(t) - \mu(Q_{sim}))}{\sigma(Q_{sim}).\sigma(Q_{obs})}$	Linear correlation coefficient	1
α_{NSE}	$\frac{\sigma(Q_{sim})}{\sigma(Q_{obs})}$	Relative variability in the simulated and observed discharges	1
β_{NSE}	$\frac{\mu(Q_{sim}) - \mu(Q_{obs})}{\sigma(Q_{obs})}$	Bias normalised by the standard deviation of the observed discharges	0
KGE'	$1 - \sqrt{(r-1)^2 + (\beta_{KGE'} - 1)^2 + (\gamma_{KGE'} - 1)^2}$	Modified Kling-Gupta efficiency	1
$\gamma_{KGE'}$	$\frac{\sigma(Q_{sim})/\mu(Q_{sim})}{\sigma(Q_{obs})/\mu(Q_{obs})}$	Variation coefficient ratio	1

Table 2 *Main characteristics of the 14 basins included in the database.*

#	River and gauging station	Country	Catchment area [km²]	Altitude range [m]	Annual precipitation [mm]	Available period	Type of change
1	River Allier at Vieille-Brioude	France	2 267	436-1551	900	1958–2008	Temperature increase, storage dam after 1983.
2	Axe Creek at Longlea	Australia	237	175-710	625	1970–2011	"Millennium Drought" during the 2000s.
3	River Bani at Douna	Mali, Ivory Coast and Burkina Faso	103 390	270-700	1075	1959–1990	West-Africa drought.
4	Blackberry Creek at Yorkville	USA	182	185-310	975	1980–2011	Growing urbanization.
5	Ferson Creek at St. Charles	USA	134	215-325	1000	1980–2011	Growing urbanization.
6	River Durance at La Clapière	France	2 170	787-4102	1275	1901–2010	Temperature increase.
7	River Flinders at Glendower	Australia	1 912	390-950	625	1967–2011	Arid catchment under cyclonic heavy rainfall influence. Major flood in 1974 and 2009.
8	River Gilbert at Gilberton	Australia	1 907	480-1070	750	1963–1988	Arid catchment under cyclonic heavy rainfall influence. Major floods in 1974.
9	River Garonne at Portet-sur-Garonne	France	9 980	140-3200	1125	1958–2008	Temperature increase.
10	River Kamp at Zwettl	Austria	622	500-1000	750	1976–2008	Increase in temperature, flood in August 2002.
11	Obyån Creek at Lissbro	Sweden	97	140-225	800	1981–2010	Loss of forest due to the Gudrun storm (2005).
12	River Rimbaud at Collobrières	France	1.4	470-622	1075	1966–2006	Forest fire in 1990.
13	Watershed 6 of the Fernow Experimental Forest	USA	0.2	730-860	1425	1956–2009	Deciduous forest cut, and converted to conifers.
14	River Wimmera at Glenorchy Concrete Weir Tail	Australia	2 000	170-950	550	1960–2009	"Millennium Drought" during the 2000s.

FIGURES

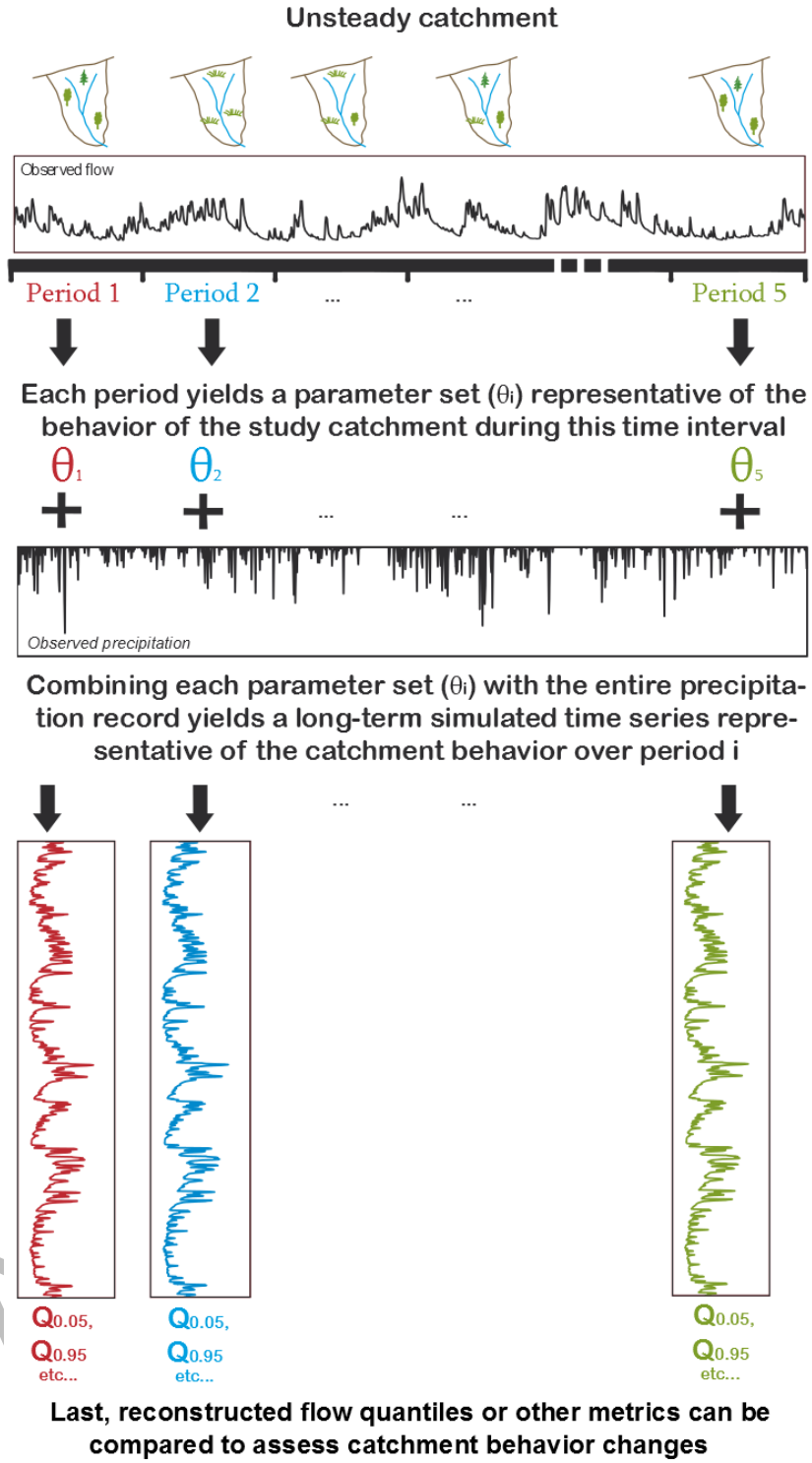

Fig. 1 *Principle of the calibration and evaluation protocol over the five sub-periods.*

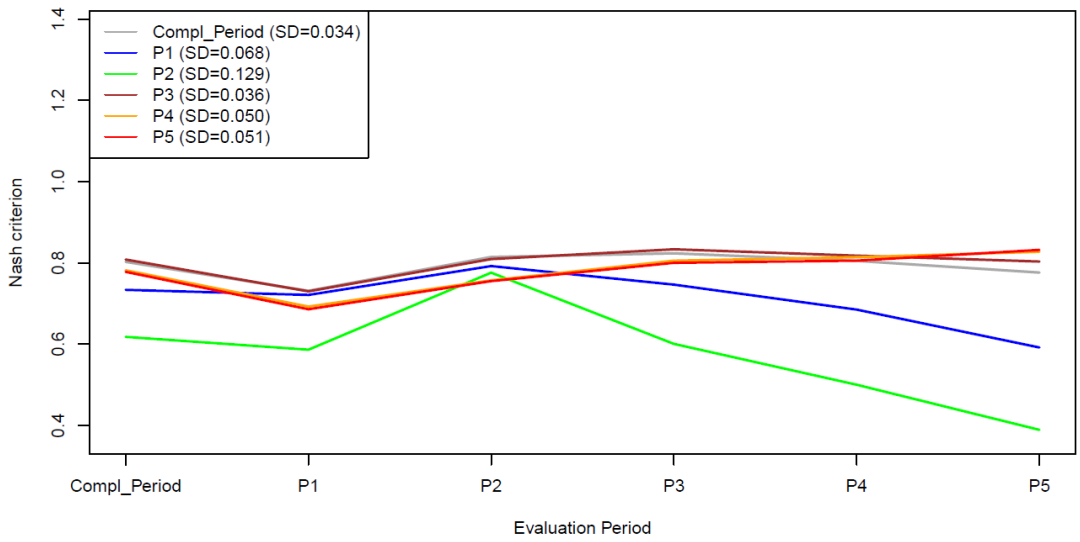

Fig. 2 *Example of a plot for the five sub-periods plus the "complete period". Here the NSE is represented. Each curve corresponds to the evolution of the score over the periods for a given calibration. SD stands for "Standard Deviation" of each curve from its mean.*

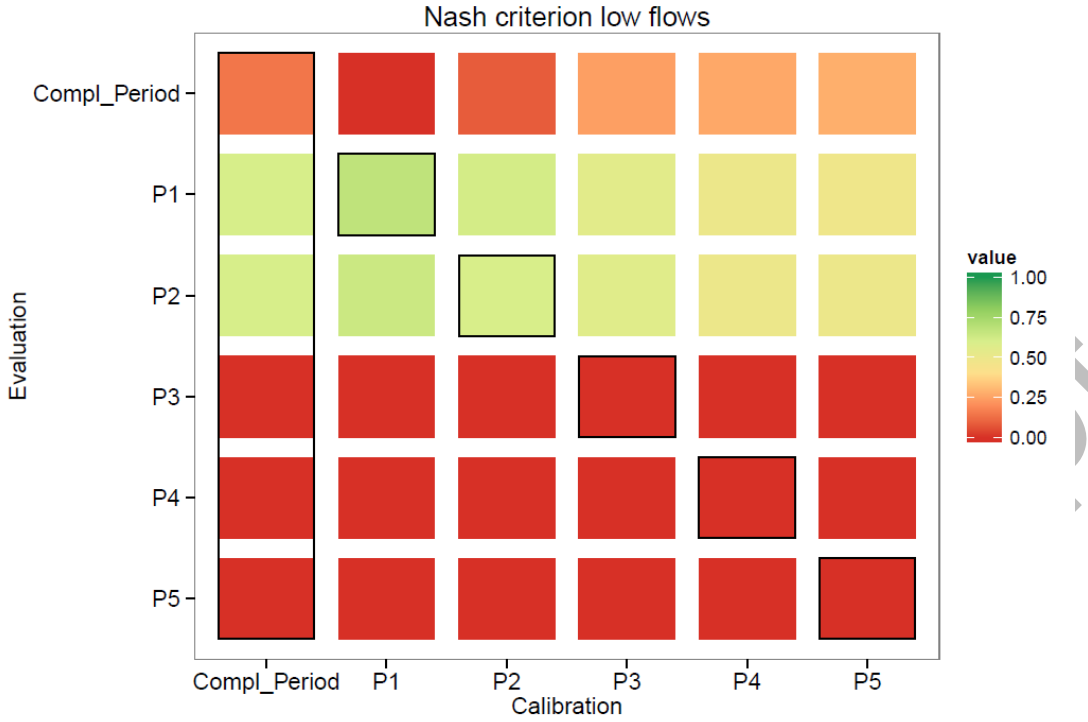

Fig. 3 *Example of a table plot for the five sub-periods plus the "complete period". Here the* NSE_{LF} *is represented. A row gives the scores of the six calibrations on a given evaluation period. A column gives the scores of one calibration over each of the six evaluation periods. The black outlines represent boxes for which the evaluation period is included in the calibration period.*

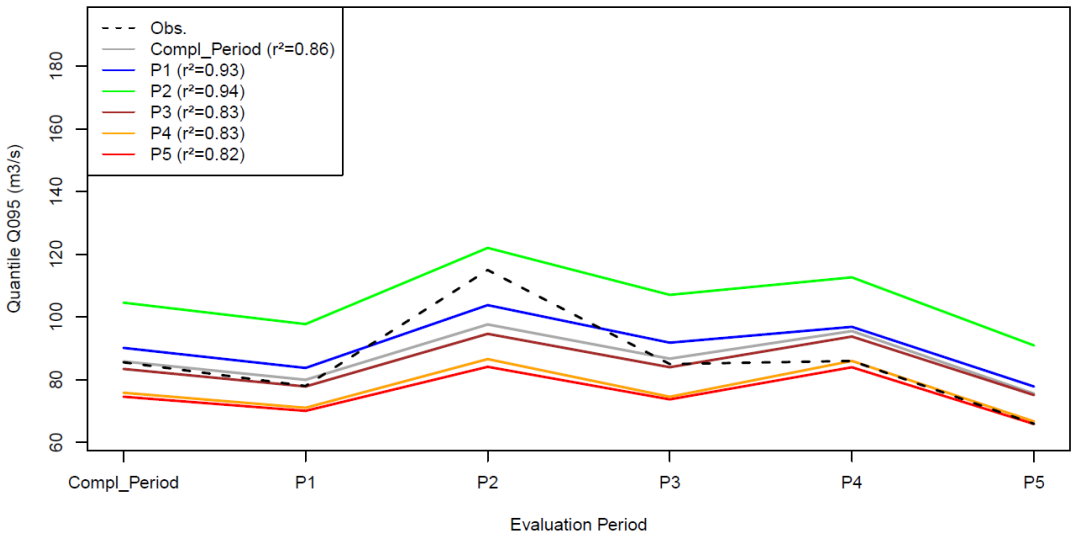

Fig. 4 *Example of a plot for the five sub-periods plus the "complete period", including the observed discharge score (dashed line) (here the* $Q_{0.95}$ *is represented).*

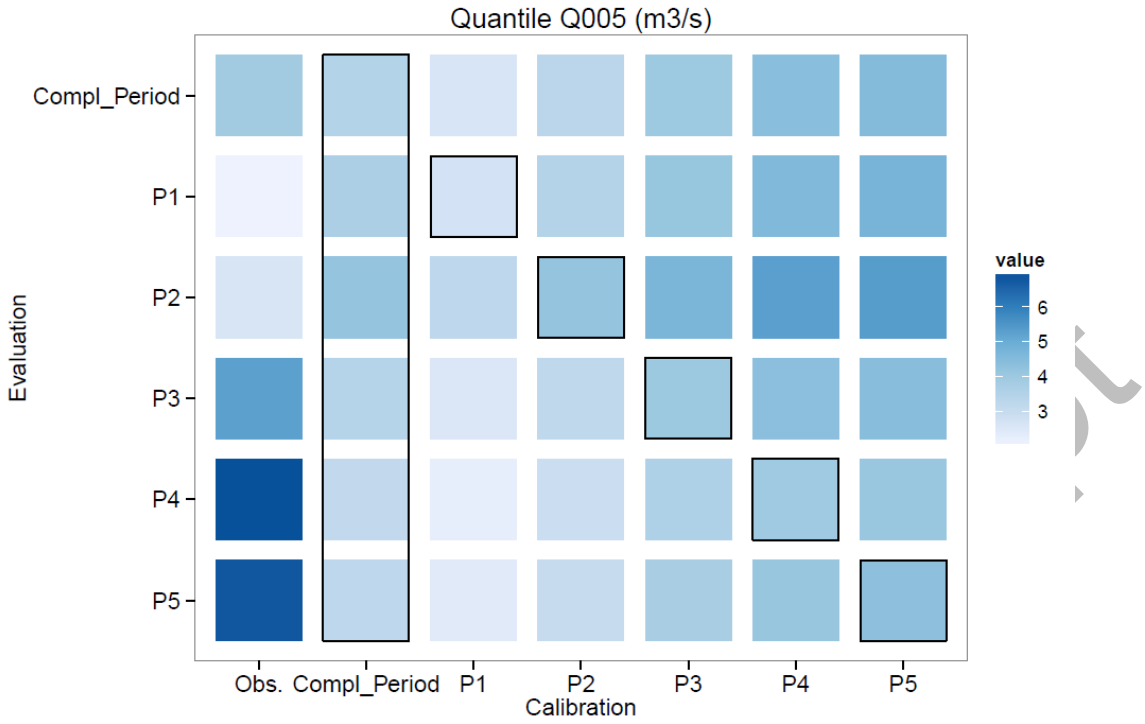

Fig. 5 *Example of a table plot for the five sub-periods plus the "complete period" and the observation (here the* $Q_{0.05}$ *is represented).*

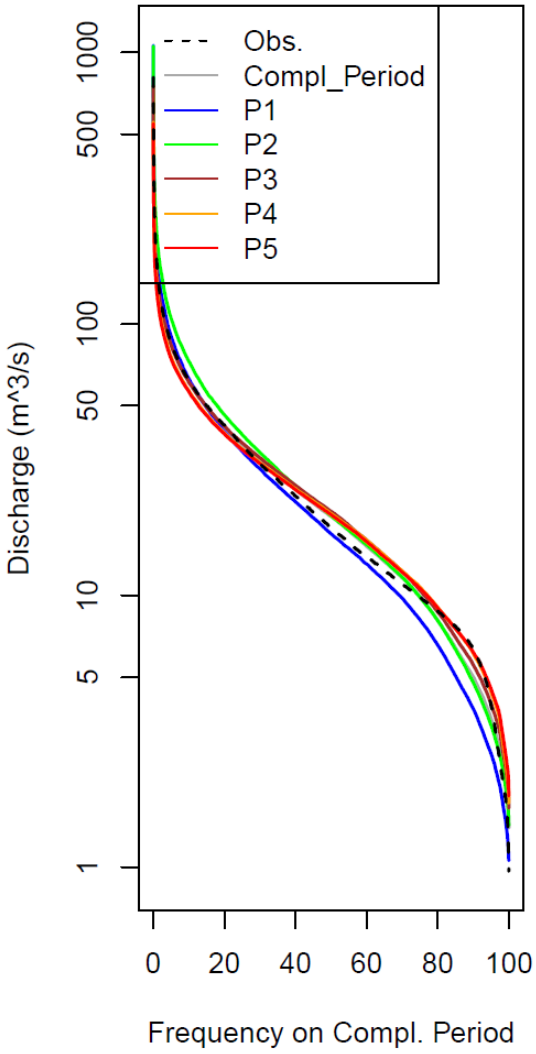

Fig. 6 *Example plot for the five sub-periods calibrations plus the "complete period" representing the Flow Duration Curve (FDC). The observed FDC is represented with a dashed line. On this plot the evaluation is done on the "complete period" only.*

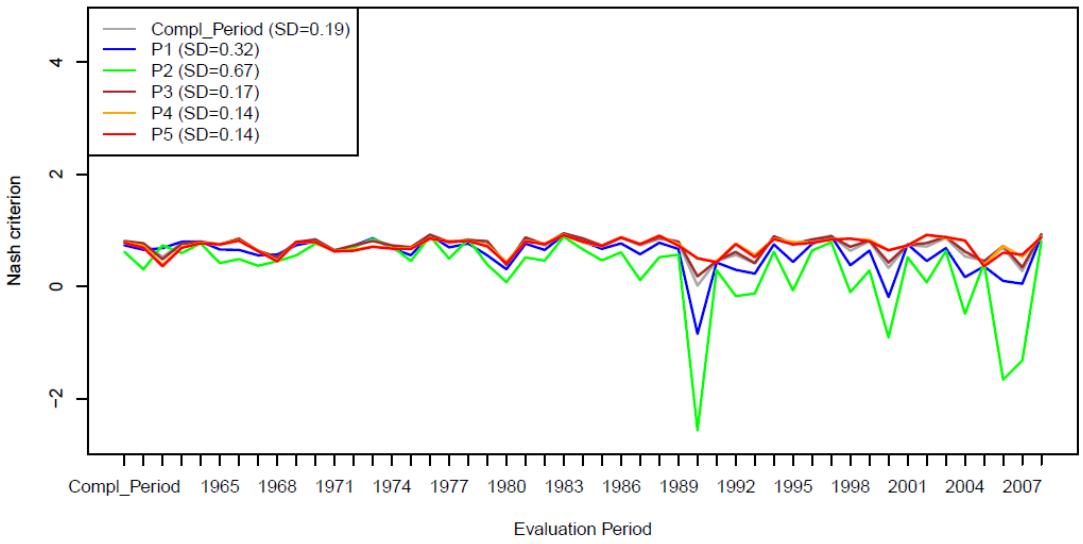

Fig. 7 *Example of a plot showing the year-to-year variations of an efficiency criterion (here the NSE is represented).*

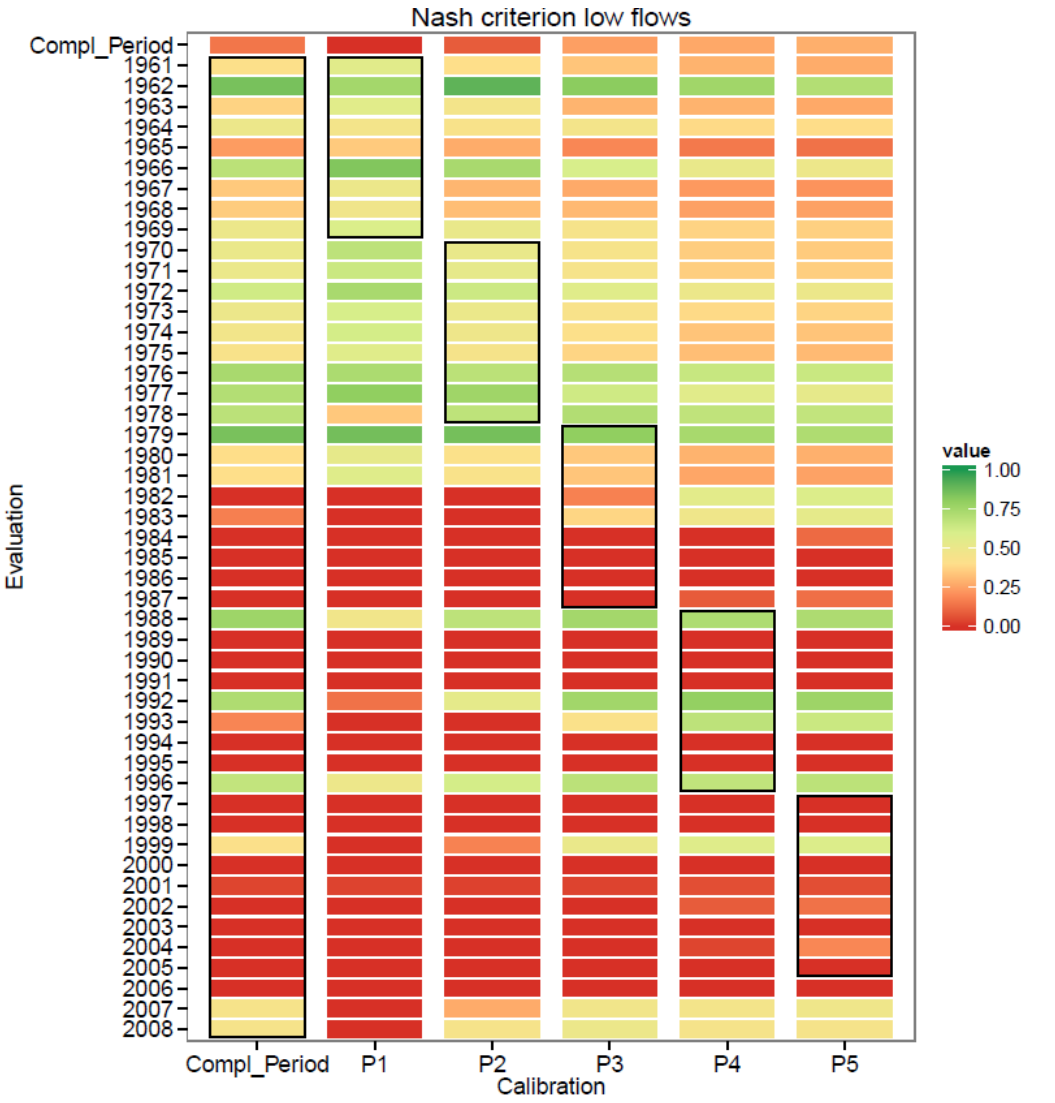

Fig. 8 *Example of a table plot on an annual basis (here the NSE_{LF} is represented).*

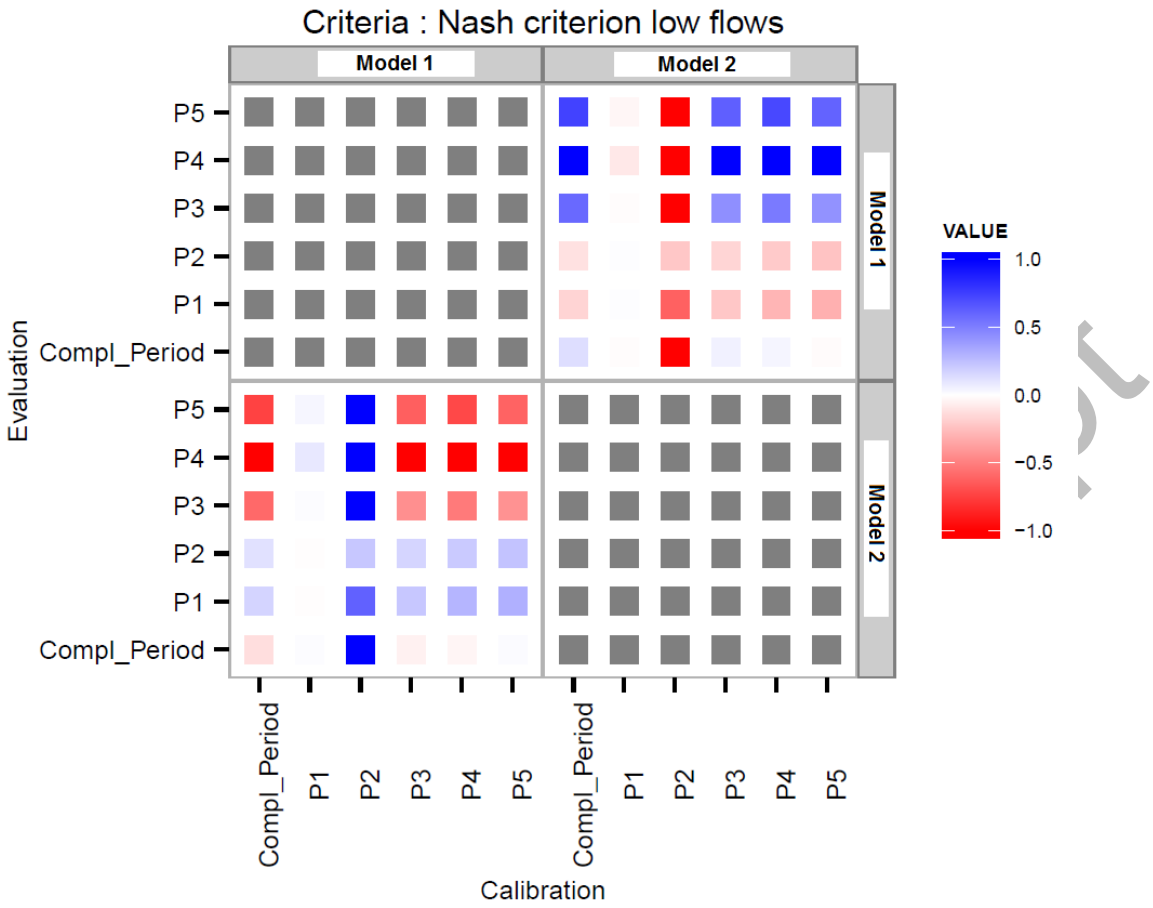

Fig. 9 *Comparison of NSE_{LF} between Model 1 and Model 2. The colours represent the difference between the NSE_{LF} of the two models.*

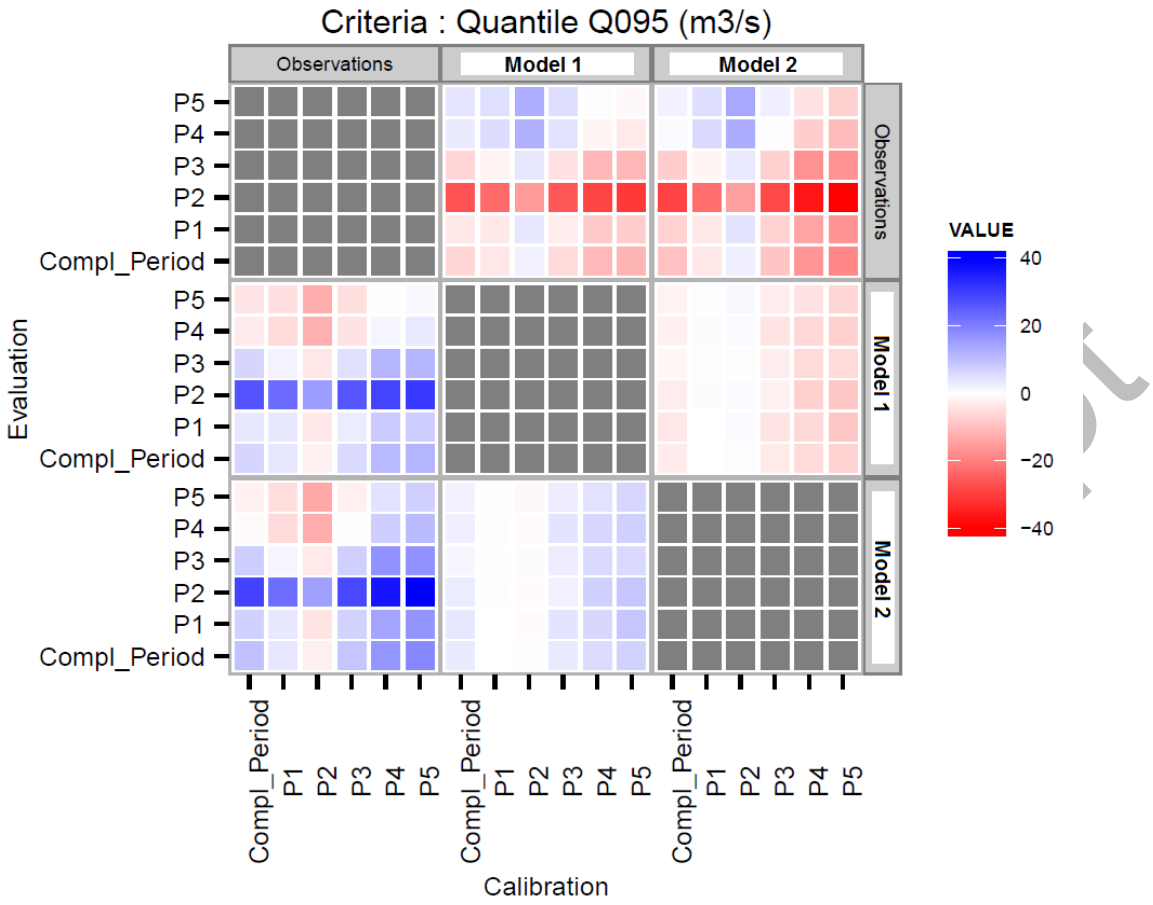

Fig. 10 *Comparison of $Q_{0.95}$ between models and observations. The colours represent the difference between the $Q_{0.95}$ of the two models.*

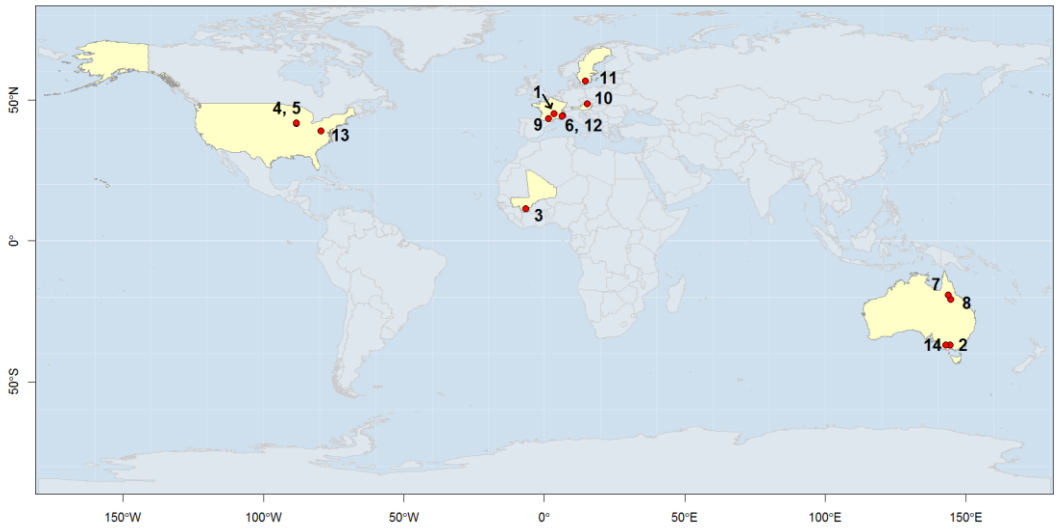

Fig. 11: *Location of the 14 catchments. The numbers correspond to the first column of Table 2.*

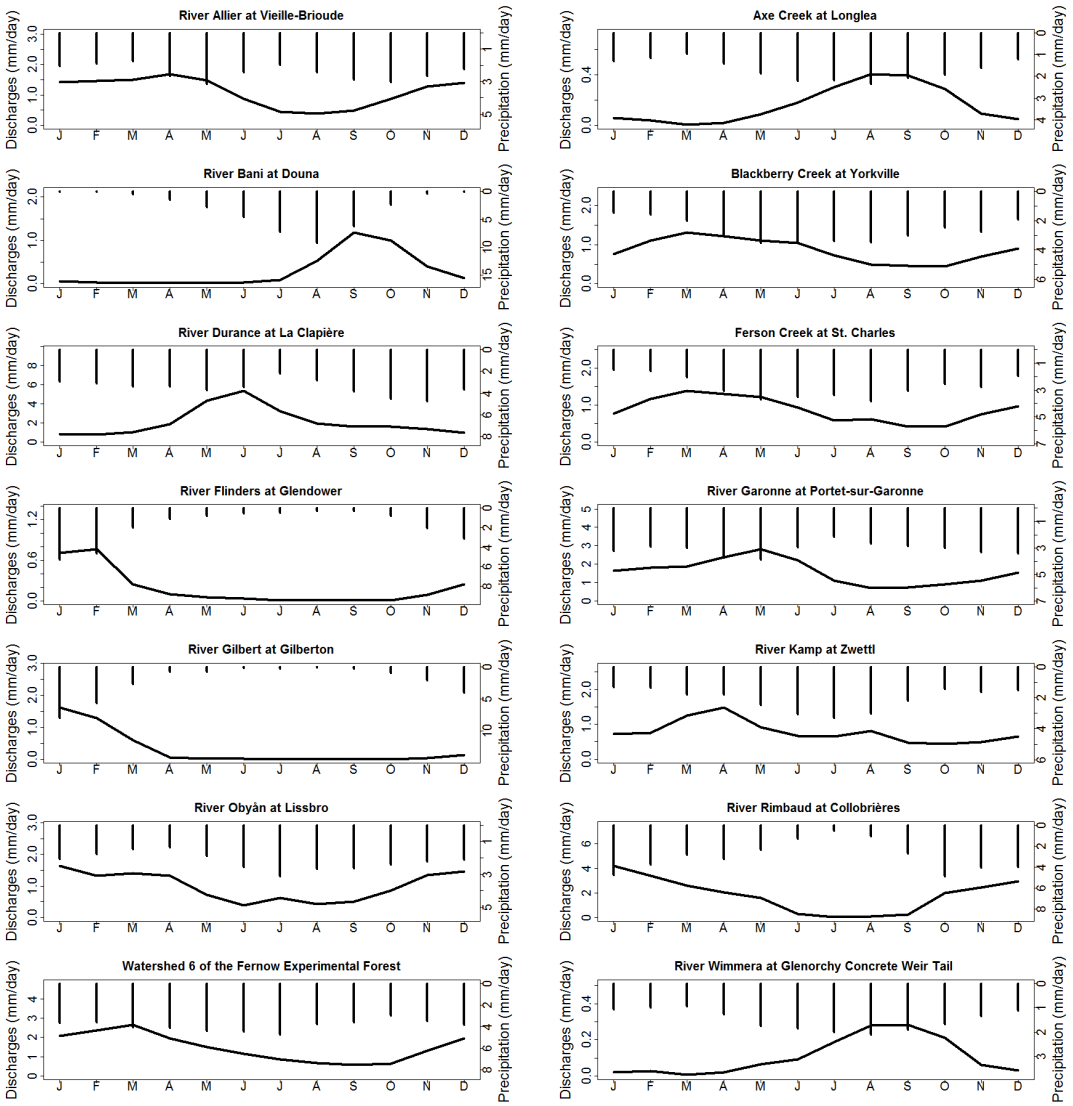

Fig. 12 *Mean monthly streamflow and precipitation for the 14 basins over the "complete period". Discharges (mm d^{-1}) are represented by the lines and precipitation (mm d^{-1}) by histograms.*

Supplementary material to article "Hydrology under change: An evaluation protocol to investigate how hydrological models deal with changing catchments"

Description of datasets and changes

G. Thirel[1], V. Andréassian[1], J.-N. Audouy[2], L. Berthet[2], P. Edwards[3], N. Folton[4], C. Furusho[1], A. Kuentz[1,5,6], J. Lerat[7], G. Lindström[8], E. Martin[9] ,T. Mathevet[5], R. Merz[10],J. Parajka[11], D. Ruelland[12] & J. Vaze[13]

1: Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), Antony, France

2: Allier Basin Flood Forecasting Centre (DREAL d'Auvergne), Clermont-Ferrand, France.

3: USDA Forest Service, Northern Research Station, Parsons, WV, 26287 USA

4: OHAX Research Unit, Irstea, Aix-en-Provence, France

5: EDF DTG, Grenoble, France

6: LTHE, Grenoble, France

7: Bureau of Meteorology, Canberra, Australia

8: SMHI, 60176 Norrköping, Sweden

9: CNRM-GAME, Météo-France, CNRS, Toulouse, France

10: Department Catchment Hydrology, UFZ, Germany

11: TU Vienna, Vienna, Austria

12: CNRS, HydroSciences Montpellier, Montpellier, France

13: Land and Water Flagship, CSIRO, Canberra, Australia

guillaume.thirel@irstea.fr

Abstract This supplementary material provides descriptions of the 14 catchments and corresponding datasets used during the workshop "Testing simulation and forecasting models in non-stationary conditions", which was held in Göteborg, Sweden, during the 2013 IAHS General Assembly.

Key words changing catchments; IAHS workshop; catchments dataset; contrasted calibration periods

Description des jeux de données et des changements

Résumé Ce matériel supplémentaire présente les descriptions des 14 bassins versants et de leurs jeux de données utilisés pendant l'atelier intitulé "Testing simulation and forecasting models in non-stationary conditions" (tester les modèles de simulation et de prévision dans des conditions non-stationnaires) qui s'est tenu à Göteborg, en Suède, durant l'Assemblée Générale 2013 de l'AISH.

Mots-clés bassins versants changeants; atelier AISH; jeu de données de bassins; périodes de calage contrastées

6 INTRODUCTION

This document provides a supplementary material to an article presenting a modelling experiment held in the framework of the 2013 IAHS conference in Göteborg, Sweden (Thirel *et al.*, 2014, this issue). The objective of this workshop entitled "Testing simulation and forecasting models in non-stationary conditions" was to discuss the issue of applying hydrological models under changing conditions. To this end, participating modellers were asked to follow a common calibration and evaluation protocol that had been defined months before the workshop and is described in Thirel *et al.* (2014, this issue). Meteorological and hydrological data for a worldwide ensemble of 14 catchments that presented changing conditions were made available to the modellers, and are described therein.

The objectives of this document are to present the 14 catchments and their main physiographical characteristics, their hydro-meteorological datasets, and their changes. For each dataset, we also give the calibration and evaluation periods chosen for the testing protocol. Summaries about previous studies for these catchments also are given. For all but one catchment, the data were made available to the participating modellers at the daily time step. Precipitation, temperature and potential evapotranspiration (PET) were averaged over the catchments, regardless of the raw data provided.

We will call the "complete period" the longest period used for both calibration and evaluation of the hydrological models. The "complete period" encompasses the longest time period possible for both the provided meteorological and hydrological

data, but excludes at least two years at the beginning of the time series, which were kept for warming up the models. The "complete period" was then split into five equal sub-periods, P1 through P5. These five periods were sequential, not overlapping, but not necessarily contiguous. While defining these five sub-periods, we tried to make use of as much data as possible.

7 DESCRIPTION OF THE CATCHMENTS

The dataset comprised 14 river basins. In the following sub-sections, each catchment and its dataset are described in alphabetical order, though we grouped Ferson Creek with Blackberry Creek and the River Gilbert with the River Flinders due to their proximities to one other and their common origin of data and cause of change. A summary of the monthly hydrological and precipitation data is given for each of the 14 catchments in Fig. 1 of the main article.

The variety of changes affecting these catchments offers wide perspectives to test the capacity of hydrological models to deal with them. Temperature increase is the main factor of unsteadiness in hilly or mountainous parts of Europe. Another factor of change was included: the construction of a dam for sustaining low flows for one river. Afforestation and deforestation modify the response of catchments. This type of land cover modification was represented by three catchments through forest fire, a storm destroying the forest and deforestation of the native deciduous forest followed by planting and establishment of conifers. Urbanization is another type of land cover change we included in the dataset. Large-scale droughts, which are known to affect the calibration of hydrological models, are represented by three catchments. Finally, two semi-tropical climate catchments were included in the dataset because they present an important variability.

7.1 River Allier at Vieille-Brioude

Physical description

The River Allier flows from the Massif Central highlands in central France (Fig. 1). The catchment area is 2269 km². Elevation ranges from 1551 to 436 m at Vieille-Brioude, with a mean slope of 0.7% (but with steeper slopes in the highest part of the catchment). The main tributaries are mountain torrents (the most important of them being the River Chapeauroux and the River Ance du Sud).

A flood scale has been installed under the old bridge of Vieille-Brioude for decades. At this location, the river flows in a V-shaped valley with steep slopes. The configuration is thought to be quite stable (except for some pebble banks in the low flow channel). The rating curve is checked and updated periodically, which allowed extracting daily data-series of water level and discharge from 1919 to the present.

Climatic conditions and hydrological regime

Mean annual precipitation on this catchment is approximately 900 mm. Spatially, precipitation is highly heterogeneous, ranging from approximately 2000 mm on the southern ridge to 500 mm in the driest zone, on a mean annual basis. In the highest mountainous part of the catchment, a significant fraction of the precipitation falls as snow: the hydrological regime is pluvio-nival, thus being influenced mainly by rainfall but also by snow melt. Floods occur most frequently during the late spring, due to a combination of the snow melt and large depressions coming from the Atlantic Ocean, and during fall, when violent storms originating from the Mediterranean Sea create very intense rains on the Southern part of the catchment. The most intense events occur when storm cells from the Mediterranean Sea initiate violent floods that are subsequently sustained downstream by stratiform rains coming off the Atlantic Ocean. Droughts can be severe on the River Allier. The longest and most severe low-flow period occurred during the 1940s.

Change

In order to mitigate low flows, a reservoir was built in 1982 on a small tributary (the River Donozau), just upstream of its confluence with the River Allier (upstream area 60 km²). The reservoir stores water from its own catchment and from the neighbour River Chapeauroux. It is also filled with water pumped from the River Allier when the discharge exceeds a threshold (mostly during the winter season). The reservoir is managed to sustain low flows downstream. Thus, the mean monthly discharge at Vieille-Brioude for August increased from 5.1 $m^3 s^{-1}$ (1919–1982) to 10.6 $m^3 s^{-1}$ (1983–2012). By contrast, the mean annual discharge fell from 28 $m^3 s^{-1}$ (1919–1982) to 22.5 $m^3 s^{-1}$ (1983–2012). The low flow distribution clearly was modified by this reservoir (Table 1). Temperature and PET also have increased while rainfall and more notably snowfall have decreased.

Data and calibration and evaluation periods

Daily precipitation, temperature and PET were available from August 1958 to July 2008 and discharge from 1919. We also provided information about the fraction of solid precipitation for each day. The meteorological data come from the Météo-France SAFRAN analysis (Quintana-Segui *et al.*, 2008; Vidal *et al.*, 2010) which makes use of both in situ observations and model outputs. PET was calculated using the Penman-Monteith formula (Monteith, 1965). Discharge data were obtained from the Banque HYDRO, the French hydrological database (www.hydro.eaufrance.fr). The "complete period" was defined as 1 January 1961 – 31 July 2008. The five sub-periods were defined as consecutive 9-year long periods, all of them starting on 1 January, and finishing on 31 December: 1961–1969 (P1), 1970–1978 (P2), 1979–1987 (P3), 1988–1996 (P4) and 1997–2005 (P5). Dam start-up is included in sub-period P3.

Past studies

The Allier basin has not been the subject of many publications in peer-reviewed journals. Some of the studies that can be found for this basin concern hydrological engineering studies about some of its tributaries (Dacharry, 1966) or floods (Onde, 1923). CETE (2009) described decreases in mean annual flows and summer flows, slight decreases in rainfall, and an increase in temperature and PET for four tributaries

of the River Allier, though none of these were impacted by the dam. The forest cover was not modified in these catchments. Recently, a nationwide project was initiated aimed at assessing the impact of climate change on hydrology of various French catchments including the River Allier catchment (Chauveau *et al.*, 2013). Lobligeois *et al.* (2014) showed that the use of a semi-distributed version of the normally lumped GR4J model instead of its lumped version could substantially improve the quality of the flow simulations in the upstream part of Allier basin.

7.2 Axe Creek at Longlea

Physical description

The Axe Creek catchment (Fig. 2) lies within the River Campaspe catchment in north-central Victoria, Australia. Elevation ranges from 175 to 710 m. The catchment area is 237 km^2 and it is located immediately south-east of Bendigo. Drainage is towards the northeast to the River Campaspe. Native vegetation was initially cleared in the 1850s during gold mining times and, subsequently, through the 1900s with agricultural activities. Land use today includes small-scale livestock grazing and agriculture. Native vegetation still covers 20% of the area, and there is some production forestry. The Axe Creek aquifer is an Intermediate Flow System, i.e. with groundwater recharge happening within about 100 to 200 km, in a fractured rock aquifer (fractured Palaeozoic bedrock). Hydraulic gradients for groundwater flow follow the topographic gradient from Big Hill Range in the south-west to the River Campaspe in the east. The defining feature of the catchment is the Whitelaw Fault that cuts the catchment into two distinct areas. Upslope of the fault the country is steep with incised streams, shallow soils and a native tree cover. Fractures mainly run in the dominant north-northwest direction with weak connection in the direction of the surface topography. Downsstream the landscape is much flatter with gently rolling terrain and extensive agricultural development. Recharge is mainly thought to occur in the high relief upslope areas of Axe Creek. Thin soils with direct connection to fractured bedrock provide an ideal conduit for recharge to the groundwater system from higher annual rainfall. Rates as high as 100 mm $year^{-1}$ are estimated although the average over the whole of Axe Creek is only 15 to 20 mm $year^{-1}$.

Climatic conditions and hydrological regime

The mean annual rainfall for the study catchment is about 600 mm, with winter and spring rainfall slightly dominant. Mean annual runoff is about 60 mm, thus the runoff coefficient is only 0.10. The highest discharges occur during the cold season, between June and October.

Change

A severe drought affected most part of Australia between 1997 and 2008. During this event, called the "Millenium Drought" (Petheram *et al.*, 2011), there was a 22% reduction in annual rainfall compared to the long-term mean but the corresponding reduction in runoff was almost 90% (thus resulting to runoff coefficients of about 0.01).

Data and calibration and evaluation periods

Precipitation, temperature and PET data were available from 1 January 1970 to 13 December 2011. Discharge data were available from 7 January 1972 to 13 December 2011. Precipitation and temperature were provided by the Bureau of Meteorology. PET was calculated by CSIRO with the Morton formula (Morton, 1983). Discharges were provided by Victorian Warehouse. The "complete period" was defined as the period from 1 January 1973 to 13 December 2011. The five sub-periods are 7-year long periods defined as follows: 1973–1979 (P1), 1980–1986 (P2), 1987–1993 (P3), 1994–2000 (P4) and 2001–2007 (P5). P1 to P3 are "wet" periods, P4 begins before the "Millenium Drought" and ends during the drought, and P5 is a "dry" period occurring completely within the "Millenium Drought" period.

Past studies

Since it belongs to a region that faces many water-related challenges, the Axe Creek catchment has been included in several studies that focused on large catchments

datasets in this part of Australia during the "Millenium Drought" (Potter *et al.*, 2010; Potter and Chiew, 2011; Potter *et al.*, 2011). For example, (Petheram *et al.*, 2011) undertook a detailed study using data from the Axe Creek catchment to investigate the changes in hydrological processes during the 'Millenium Drought' period. They concluded that the larger than normal reduction in runoff for a given reduction in rainfall during the drought was due partly to a loss in connectivity between the surface water and groundwater systems. Vaze *et al.* (2010a) examined a hydro-meteorological dataset for south-east Australia. Also, they applied six conceptual models to 232 catchments and showed that using parameters of nearby catchments resulted in reasonable performance. Vaze *et al.* (2010b) investigated whether the calibrated parameter values for rainfall–runoff models could be used to predict runoff responses to changes in future climate inputs. Chiew *et al.* (2013) studied "Millenium Drought"-related changes and showed that models calibrated before the drought did not represent the river responses post-drought. However, developing new models helped to better understanding the "Millenium Drought". Petheram *et al.* (2011) showed that the relation between outflow and aquifer storage was modified for the Axe Creek during this drought.

7.3 River Bani at Douna

Physical description

The River Bani runs primarily through southern Mali. Its catchment drains an area of around 100 000 km^2 at the Douna gauging station (Fig. 3). It was chosen for this study because of its large contribution to the flood of Inner Niger Delta (IND), the availability of data, and because its flows have not been disrupted by large-scale hydraulic works. It constitutes an ideal area for analysing the climate impacts on water resources in western Africa (Ruelland *et al.*, 2012).

The watershed's topography is gently sloping, with elevations between 270 and 700 m (Fig. 3). Soils are mostly ferralitic and leached with high sand and clay contents. Sandy hillwash often is found at the surface, while basal gravels are found in deeper parts of the profiles. The natural vegetation is savannah woodland. Agricultural areas are growing rapidly due to demographic pressure. Typical crops include millet, sorghum, cotton, manioc and peanuts. The aquifers are fissured formations of low permeability with a base layer of Birrimian mica-schist and metamorphic rocks in the southwest and Infracambrian sandstone in the northeast. The downstream area of the basin, therefore, has layers of higher permeability that can be expected to create more sustained low flows.

Climatic conditions and hydrological regime

The Bani catchment, located in a Sudano-Sahelian climatic regime, is characterized by a monsoon climate with a strong north-south rainfall gradient (from approximately 700 mm $year^{-1}$ in the North to 1500 mm $year^{-1}$ in the South) and a single rainy season between April and October (Ruelland *et al.*, 2012). As a result, the wet season is short and it is the only period during which the evapotranspiration demand can be satisfied. Consequently, the wet season is a crucial period for replenishing surface and subsurface water storages. The high flow period is between August and November. Highest discharge values (on 11-day moving average) are around 2700 $m^3 s^{-1}$ prior to 1970 and about 1000 $m^3 s^{-1}$ since 1970.

Change

The steady decline in rainfall since 1971 (-17%) had lasting effects on runoff. The flow observed at the Douna gauging station fell by 67% between the 1950–1970 and 1971–2000 periods (Ruelland *et al.*, 2012), with a decrease in deep water recharge and baseflow contribution to the annual flood (Ruelland *et al.*, 2009). Some of the low-water periods were so severe that the river flow stopped periodically at Douna during the 1980s. The only other Western African river system where a deficit of this magnitude has been observed is the Senegal catchment.

Moreover, historical remote sensing studies (Ruelland *et al.*, 2010b; Ruelland *et al.*, 2011) showed that the downstream Sahelian part of the catchment has

undergone drastic cropland expansion and deforestation since the 1960s. By contrast, in the most productive sub-basins (i.e., in the upstream part of the catchment), the observed land cover changes have been relatively limited due to lower demographic pressure and a better capacity of the natural vegetation to regenerate (Ruelland *et al.*, 2010a).

Data and calibration and evaluation periods

Daily rainfall series were derived from 72 rain gauges covering the area (Ruelland *et al.*, 2012). For the 1959–1990 period, these gauges were used to interpolate precipitation and develop rainfall maps by the inverse distance weighted method, which proved to be optimally accurate among the classic methods available for spatial integration of point data in the given context (Ruelland *et al.*, 2008). Since the only data available for estimating PET were temperature time series, a formula relying on solar radiation and on mean temperature was selected (Oudin *et al.*, 2005). This formula was used with a monthly temperature time step as provided by the CRU TS 2.1 World database on a 0.5° square grid (Mitchell and Jones, 2005). However, since extra-terrestrial radiation is a daily variable that depends on latitude and the Julian day of the year, PET was finally computed at a daily time step from monthly temperature data. Discharge data were from the Douna gauging station. This station is located upstream of the basin confluence with the Niger River, and it appeared to have a high-quality daily discharge series (less than 0.5% missing daily runoff values from 1961–1990). The "complete period" was defined as 1961–1990. The five sub-periods are 5-year long periods: 1961–1965 (P1), 1966–1970 (P2), 1973–1977 (P3), 1978–1982 (P4) and 1983–1987 (P5). P1 and P2 belong to the "wet" period, P3 to P5 belong to the "dry" period.

Past studies

For the past few decades, the Sudano-Sahelian regions have experienced a lasting drought, which started at the end of the 1960s and culminated in the 1980s with a rainfall deficit of 15–30% compared to the period 1950–1960 (see Nicholson *et al.*, 1998; L'Hôte *et al.*, 2002; Le Barbé *et al.*, 2002). Several studies have addressed the Bani basin's significant hydro-climatic variability over the last 50 years. Physically-

based hydrological modelling with the SWAT model was conducted to understand the factors controlling flow evolution over the last 50 years (Laurent and Ruelland, 2010). However, reservoir-based conceptual modelling was more appropriate for simulating long-term rainfall-runoff relationships in this large, poorly gauged, catchment. After analysing its sensitivity to various modes of rainfall interpolation (Ruelland *et al.*, 2008), the application of the HydroStrahler model allowed to reproduce runoff at the catchment outlet with high accuracy for 1950–2000 (Ruelland *et al.*, 2009; Ruelland *et al.*, 2012). These simulations showed a decrease in the simulated subsurface runoff and deep infiltration during the 1970s and 1980s, which can be attributed to the persistent rain deficit (Ruelland *et al.*, 2009). This has led to a drastic decrease in deep water recharge and the base runoff contribution of flood composition. These works were completed by a comparison of conceptual models (HydroStrahler *vs*. GR4J) which showed that moving from a lumped to a semi-distributed approach did not significantly improve the simulated hydrograph at the catchment outlet (Ruelland *et al.*, 2008; Ruelland *et al.*, 2010a). Attempts were made to account for land cover changes on the catchment and for their potential effects on runoff and infiltration attributable to alterations to surface features (Ruelland *et al.*, 2010b; Ruelland *et al.*, 2011). No conclusive results were obtained because the major land cover changes were located in the downstream areas where contributions to runoff were low. Attempts to link the water holding capacity of a monthly lumped model (GR2M) to satellite NOAA-AVHRR NDVI values at the end of the dry season were made (Dezetter and Ruelland, 2012). Finally, Ruelland *et al.* (2012) investigated future hydro-climatic conditions using climatic scenarios over the 21st century in the catchment. They showed that, based on a projected rainfall deficit and a continuing increase in PET, catchment discharge could decrease by the end of the 21st century to the same levels as those observed during the 1980s.

7.4 Blackberry Creek at Yorkville and Ferson Creek at St. Charles

Physical description

The Blackberry Creek watershed near Yorkville (Illinois, USA, US Geological Survey – USGS – station 05551700) drains an area of 169 km² (Fig. 4) and is located approximately 100 km west of metropolitan Chicago (Murphy *et al.*, 2007). Blackberry Creek, a 51 km-long stream, originates north of Elburn in central Kane County and drains to the Fox River near Yorkville in Kendall County. Nearly 54% of the Blackberry Creek watershed is planted in row crops, such as corn and soybeans. About 87% of the watershed has a slope of less than 4%, and 50% of the watershed has a slope of less than 1.2%. The topography varies from level or nearly level to rolling terrain with numerous small depressions and steeper slopes at headwater sections of the main stem and tributaries. The change in elevation from the headwaters to the mouth of Blackberry Creek is about 90 m. The watershed is located within the Bloomington Ridged Plain (Leighton *et al.*, 1948). The area is characterized by low, broad morainic ridges with intervening wide stretches of flat or gently undulating ground moraine. Parent soil materials are loess, glacial till, lacustrine, outwash alluvium, and organic deposits.

Ferson Creek near St. Charles is a USGS streamflow gauging station (USGS ID 05551200). The watershed drains 134 km² (Fig. 4) and is located on the urban fringe of Chicago. The Ferson Creek mainstream is 24-km long, originating north of Elburn in central Kane County and draining to the River Fox near St. Charles in Kane County. In 2007, row crops such as corn and soybeans covered nearly 36% of the watershed, and forest and rural grassland, respectively, covered approximately 13% and 37% of the watershed. Average land surface slope ranges from 0.5% to 2.8%. About 91% of the watershed has a slope of less than 4%, and 50 % of the watershed has a slope of less than 1.2% (Bartosova *et al.*, 2007b).

Climatic conditions and hydrological regime

The climate of northeastern Illinois is humid continental with warm to hot summers and moderate to fairly cold winters. The proximity of the watersheds to Lake Michigan has a moderating effect on climate (Federal Emergency Management

Agency, 2002). The long-term (132 years) average annual precipitation is 940 mm, with approximately 400 mm from thunderstorms and 75 to 90 mm from snowfall (Changnon *et al.*, 2004). The long-term (122 years) average temperature is approximately 9.4°C at Aurora (U.S. Department of Commerce, 2001). The largest streamflow values are observed from mid-winter to late spring when ground conditions (soil moisture and transpiration needs) are conducive to minimal infiltration rates and large runoff amounts. However, intense, short-duration storms during the summer can produce major floods in both watersheds.

Change

Urban development has increased in the Blackberry watershed during the past few decades, with appreciable residential and commercial lands spreading out within the jurisdictions of United City of Yorkville, Village of Montgomery, Kendall County, and in the eastern portion of the watershed near Aurora, as well as various other sections of the watershed. Population and urbanized land are expected to double by 2020. The Ferson watershed is located within Kane County, the fifth most populated county in Illinois, with 27.5% population growth from 2000-2010 (Chicago Metropolitan Agency for Planning, 2011). Annual data on the urbanization of both watersheds since 1980 were provided by Thomas Over (USGS) for the IAHS workshop (Table 2). Annual values of urbanized area were estimated using the threshold of 4 ha per housing unit and the interpolation of the decadal housing density data by Theobald (2005).

Data and calibration and evaluation periods

Precipitation and temperature were obtained from the meteorological DayMet analysis aggregated by the USGS Geo Data Portal (Blodgett *et al.*, 2011; Thornton *et al.*, 2012). PET was calculated using the Oudin formula (Oudin *et al.*, 2005). Daily precipitation, temperature, PET and discharge were available from 1980 to 2011 for both catchments. As a consequence, the "complete period" was defined as 1982–2011. The 5 sub-periods each were 6-years long: 1982–1987 (P1), 1988–1993 (P2), 1994–1999 (P3), 2000–2005 (P4) and 2006–2011 (P5). The urban fraction values

corresponding to each of these six periods and obtained from the yearly values are given in Table 2.

Past studies

The Ferson Creek and Blackberry Creek watersheds were included in the set of 78 drainage catchments in Illinois studied by O'Hearn and Gibb (1980) to estimate the groundwater contribution to baseflow using the graphical hydrograph separation technique applied by Walton (1965) in a precursory investigation. O'Hearn and Gibb (1980) identified numerous factors exerting influence on the regional distribution of baseflow, including but not limited to land use, point source discharges, surficial soil permeability, catchment topography, and climate. A year later, both watersheds were among the 131 catchments used by Singh (1981) to derive unit hydrograph parameters. Singh (1981) estimated unit hydrographs for determining 100-year flood and probable maximum flood hydrographs from catchment factors, such as drainage area, main channel length, and main channel slope for watersheds in a homogeneous region.

Later, the hydrologic behaviours of Ferson and Blackberry Creeks were studied by Knapp and Myers (1999) to update the hydrologic analysis used in the Illinois Streamflow Assessment Model. The model was adapted to better simulate flow frequencies that were influenced by population increases, overall water use, resulting effluent discharges, and general increases in streamflow caused by climatic variability and increases in average precipitation.

The River Fox Study Group selected the HSPF model (Bicknell *et al.*, 2001) to simulate watershed loading, and delivery and routing of nonpoint and point sources of pollution from the entire watershed. The specific development of the watershed loading model focused on two tributary watersheds (Blackberry and Poplar Creeks) in the River Fox Watershed (Bartosova *et al.*, 2007b). The HSPF model was calibrated to simulate daily streamflow and selected water quality constituents. Bartosova *et al.* (2007a) describes the estimation of model parameters using flow observations from five tributary watersheds not used in the calibration process, including Ferson Creek watershed.

Effective peak-flood discharges for Blackberry Creek (Federal Emergency Management Agency, 2002) for two locations in Kendall County were determined in 1976 using the USDA TR-20 hydrologic model (U.S. Department of Agriculture, 1992). The discharges were determined to be outdated in a comparison to flood frequencies, estimated using data from the Yorkville station (Soong *et al.*, 2004). A flood-hazard study of the Blackberry Creek watershed in Kane County has been completed by the USGS and Kane County Division of Environmental and Building Management (KCDEM; Soong *et al.*, 2005). The 100- and 500-year flood plain and 100-year floodway maps were generated for the determination of flood hazard areas in the Blackberry Creek watershed. In 2005, the USGS and KCDEM completed an addendum to the Soong *et al.* (2005) report that added the Aurora Chain-of-Lakes tributary to the analyses.

Before that, the USDA, Soil Conservation Service (now the Natural Resources Conservation Service; NRCS) conducted a watershed-wide flood-hazard analysis to estimate flood quantiles and flood stages along Blackberry Creek (U. S. Department of Agriculture, 1989). They used the TR-20 model for estimating peak discharges, and the Soil Conservation Service Water Surface Profile hydraulic model (U. S. Department of Agriculture, 1976) for estimating peak stages. Besides identifying the 100- and 500-year flood plains and the floodway, the study also identified developed areas that were prone to flooding, evaluated the importance of natural storage in the watershed, and suggested alternatives for flood-plain management.

Regional regression equations for Illinois were developed by Soong *et al.* (2004). The regional regression equation estimated the mean (logarithmic) value of flood quantiles obtained at different watersheds in a region with the same set of explanatory variables. The procedures used in developing hydrologic and hydraulic models and for estimating flood-peak magnitudes and recurrence intervals used for flood-hazard analysis have been described (Murphy *et al.*, 2007). To address the flood-hazard analysis on the watershed scale, the entire watershed (the main stem as well as seven tributaries of Blackberry Creek in Kane and Kendall Counties) was included in the hydrologic analyses. More recently, Soong *et al.* (2009) prepared a study on the effects of stormwater detention catchments with specified release rates on the watershed scale with the HSPF model in cooperation with the KCDEM and the Illinois Office of Water Resources.

7.5 River Durance at La Clapière

Physical description

The River Durance at La Clapière catchment stretches over approximately 2170 km² in the French south Alps (Fig. 5). La Clapière station is located at 787 m elevation. The watershed is mountainous, with 85% of its surface over 1500 m elevation and 20% over 2500 m. The highest parts of the watershed, which peak at 4102 m, are situated in the Ecrins massif on the western side of the watershed. Because of the mountains, the average bed slope between the spring from which the river originates and the outlet is nearly 25 m/km. Only a few small glaciers are present in the watershed. In 2009, these covered approximately 20 km² according to M. Gardent (personal communication with A. Kuentz, 14 March 2013).

Climatic conditions and hydrological regime

The mean annual precipitation over the watershed calculated from 1948–2010 was 1350 mm. The climate is Alpine but it also has Mediterranean influences. On average, during October, November, and December the catchment receives the most of precipitation, while July and August are the driest months. Local precipitation variations are strong due to orographic effects; the upper Durance valley is in the shadow of the surrounding mountains. Snowfall is common during winter. From December to March, a snowpack typically covers the majority of the watershed's surface, and snowmelt can last until July. The mean annual temperature over the catchment is 3.5°C (calculated from 1948–2010), with local means ranging from approximately between 8°C to -4°C, depending on elevation. These physical and climatic characteristics lead the hydrological regime of the Durance at La Clapière to be principally snow-influenced, with maximum monthly streamflows in late spring and early summer (May, June and July) and low flows in winter. The mean annual streamflow at La Clapière station is approximately 52 $m^3 s^{-1}$.

Change

The mean watershed temperature for River Durance at La Clapière has increased significantly over the last century. Using a linear model from 1883–2010, the mean annual temperatures show an increase of more than 0.01°C year^{-1}, resulting in a 1.2°C increase between 1900 and 2010. This trend is statistically significant according to a Mann-Kendall test (Mann, 1945). This increase of air temperature is one factor explaining the already mentioned 30% decrease in surface coverage by glaciers in the watershed from 1960 to 2009.

Data and calibration and evaluation periods

Time series of mean precipitations and air temperatures over the watershed have been reconstructed at a daily time-step for the 1883–2010 period based on local (observed series) and regional (climatic reanalysis) climatic data using the ANATEM method (Kuentz *et al.*, 2013). The PET was calculated using the Oudin formula (Oudin *et al.*, 2005). Daily precipitation, temperature and PET were available from 1 January 1901 to 30 December 2010. Discharges were available only from 1 January 1904. The "complete period" was defined as 1904–2010. The five sub-periods were 21-year-long periods: 1904–1924 (P1), 1925–1945 (P2), 1946–1966 (P3), 1967–1987 (P4) and 1988–2008 (P5).

Past studies

Due to its natural complexity and the variety of potential water uses, the Durance watershed has been studied for a long time. The first extensive study was published by Imbeaux (1892) focusing on floods and presented some of the first hypotheses of rainfall-runoff relations. Later, Wilhelm (1913) described the interest of building dams on this watershed. The River Durance hydrology was also extensively described in Pardé (1925). More contemporary studies were initiated in the mid-20th century with the development of dams (Serra, 1953; Morlat *et al.*, 1956). Today, the Durance watershed is often used as a study site for hydrological model developments (Garçon, 1996; Paquet and Garçon, 2002; Lafaysse *et al.*, 2011; François *et al.*, 2013; Kuentz *et al.*, 2013).

7.6 River Flinders at Glendower and River Gilbert at Gilberton

Physical description

The Flinders catchment is located in the northwestern part of the Queensland state in Australia (Fig. 6). The Flinders at Glendower covers an area of 1960 km². Elevation ranges from 390 to 950 m. The catchment is sparsely populated with approximately 6000 people; about two-thirds of the population reside in four towns: Cloncurry, Hughenden, Richmond and Julia Creek. The River Gilbert is located in the northwestern part of the Queensland state in Australia (Fig. 6). It covers an area of 1890 km² with an elevation ranging from 480 to 1070 m. The catchment is sparsely populated with approximately 1200 people, but has one urban centre in Georgetown. Both catchments are part of the headwaters of the larger area draining to the Gulf of Carpentaria. Vegetation cover in the two catchments is dominated by Eucalypt woodlands that alternate with grazing areas. The vast majority of the catchment is considered remote with little or no development.

Climatic conditions and hydrological regime

The two catchments have a semi-arid tropical climate. The mean and median annual rainfall spatially averaged across the Flinders catchment are 614 mm and 600 mm, respectively, and across the Gilbert catchment they are 750 mm and 700 mm, respectively. However, the historical annual rainfall series shows considerable variation among years. The highest annual rainfall (1110 mm in 2009 for Flinders and 1930 mm in 1974 for Gilbert) is about twice the median annual rainfall value. A

defining characteristic of the climate of the two catchments is the seasonality of rainfall, with more than 80% of rainfall occurring during the wet season (November to April, see Fig. 1 of the main article). The highest median monthly rainfall occurs during the months of January and February (~110 mm for Flinders, ~160 mm for Gilbert). The months with the lowest median rainfall are July and August (~1 mm for both catchments). The catchments have a mean annual potential evaporation of 1740 mm. Consequently, the majority of the Flinders and Gilbert catchments experiences a mean annual rainfall deficit of more than 1000 mm.

These characteristics cause the two rivers to be non-perennial with no flow during about 60% of the daily flow time series. There is a marked seasonality of the monthly flow patterns with high flows during the wet season (November to April) and low flows during the dry season (May to October). Due to high rainfall deficit, the runoff coefficient remains low, i.e. about 10%.

Change

The catchments did not exhibit important changes in hydrological flow regimes during the period of record. However, they show high interannual variability that is difficult to distinguish from long-term trends. In addition, the catchments are non-perennial, which remains a challenge for most hydrological models. As a result, these catchments can be seen as a benchmark test for handling of changing catchments.

Data and calibration and evaluation periods

Precipitation and PET data, estimated through the Morton's method (Morton, 1983), were obtained from the SILO climate data archive (http://www.longpaddock.qld.gov.au/silo; Jeffrey *et al.*, 2001). Temperature was not provided for these catchments. For the Flinders catchment, precipitation and PET were available from 1 January 1967 to 16 June 2011, and discharges were available from 2 September 1972 to 16 June 2011. The "complete period" was defined as 1973–2010. The five sub-periods were 7-years long: 1973–1979 (P1), 1980–1986 (P2), 1987–1993 (P3), 1994–2000 (P4) and 2001–2007 (P5). For the Gilbert catchment, precipitation and PET were available from 1 January 1963 to 30 September 1988, and discharges were available from 27 July 1968 to 30 September

1988. The "complete period" was defined as 1969–1987. The five sub-periods were 3-year long: 1969–1971 (P1), 1973–1975 (P2), 1977–1979 (P3), 1981–1983 (P4) and 1985–1987 (P5).

Past studies

The Gilbert and Flinders catchments have been modelled as part of the studies supporting the water resources planning of the Gulf of Carpentaria (Gulf Water Resources Plan; see http://www.nrm.qld.gov.au/wrp/gulf.html; DNRM, 2006a; DNRM, 2006b). Other studies focusing on the two catchments include the Northern Australia Sustainable Yield project (CSIRO, 2009a; CSIRO, 2009b), which provided a large-scale assessment of water resources across the northern part of Australia. The main findings were that the climate variability across northern Australia is extremely large and future projections from global climate models indicate an increase in the rainfall across the region. CSIRO is currently undertaking a major project, the Flinders and Gilbert Agricultural Resources Assessment (http://www.csiro.au/fgara) where the potential for irrigation development is explored due to recent interest in development of Northern Australia. The findings from this work were released to the public at the beginning of 2014 (see CSIRO, 2014).

7.7 River Garonne at Portet-sur-Garonne

Physical description

The River Garonne at Portet-sur-Garonne is located in the upper part of the basin, which lies in southwestern France. While the complete River Garonne basin (56 000 km^2) drains the northern slopes of the Pyrenean chain (along the French border with Spain) and the southern slopes of the Massif Central, the upper part drains only the Pyrenean chain (Fig. 7). The watershed area is 9980 km^2 and its elevations range from 140 m to 3200 m.

Climatic conditions and hydrological regime

The climate over the basin is influenced by oceanic conditions over its western part, and is characterized by heavy rainfall events during winter and relatively warm weather during summer. There is a significant precipitation gradient from the west to the east, ranging from approximately 1200 mm $year^{-1}$ in the Atlantic coastal region to about 600 mm $year^{-1}$ 300 km to the east. The upper River Garonne's hydrologic regime is marked by the spring snowmelt occurring in the Pyrenees (Caballero et al., 2007), while summer flows are very low due to relatively low precipitation.

Change

There is substantial human influence in the basin from irrigated agriculture. Irrigated area has been increased by a factor of five between the 1970s and the 1990s, and has stabilized at 160,000 ha (Sauquet *et al.*, 2010). The mean temperature increased by 1.1°C from 1901 to 2000, while there are no significant changes in precipitation (Moisselin *et al.*, 2002).

Data and calibration and evaluation periods

Daily precipitation, temperature, PET and discharge were available from 1 August 1958 to 31 July 2008. Information about snowfall for each day also was available. Meteorological data came from the Météo-France SAFRAN analysis (Quintana-Segui *et al.*, 2008; Vidal *et al.*, 2010), which makes use of both in situ observations and models outputs. PET was calculated using the Penman-Monteith formula (Monteith, 1965). Discharge data were obtained from the Banque HYDRO, the French hydrological database (www.hydro.eaufrance.fr). The "complete period" was defined as 1 January 1961–31 July 2008. The five sub-periods were defined as consecutive 9-year-long periods starting on 1 January, and finishing on 31 December: 1961–1969 (P1), 1970–1978 (P2), 1979–1987 (P3), 1988–1996 (P4) and 1997–2005 (P5). While precipitation showed no trend over the five sub-periods, temperature and PET increased, and discharge and the snowfall fraction decreased.

Past studies

Recent studies on the basin focused on the impact of future climate change on low flows (Caballero *et al.*, 2007) and on the evolution of water usage (Sauquet *et al.*, 2010; Hendrickx and Sauquet, 2013).

7.8 River Kamp at Zwettl

Physical description

The River Kamp is located in northern Austria, approximately 120 km northwest of Vienna. The catchment upstream of Zwettl covers 622 km² and its elevation ranges from 500 to 1000 m (Fig. 8). The higher elevations of the catchment, in the southwest, are hilly with deeply incised channels. Toward the outlet, in the northeast, the terrain is flatter and swampy areas exist along the streams. The geology of the catchment is primarily made of granite and gneiss. Weathering has produced sandy soils with a large storage capacity. Fifty percent of the catchment is forested.

Climatic conditions and hydrological regime

The mean annual precipitation is about 900 mm, of which about 300 mm becomes streamflow (Parajka *et al.*, 2005). Typical flow travel times in the river system range from 2 to 4 h. The average maximum annual peak discharge is about 65 $m^3 s^{-1}$. The largest flood volumes are produced by synoptic events, in which humid air is transported from the Mediterranean Sea. Other flood processes are flash floods driven by convective storms that occur at smaller spatial scales and can lead to a very rapid rise in the river stages. In addition, snow melt floods and rain-on-snow floods occur in winter or spring. These floods are typically characterized by gradual rises of stream water levels. During moderate flows events, only a small proportion of rainfall contributes to runoff and event runoff coefficients are 10% or less (Merz and Blöschl, 2005). As rainfall increases, the runoff response characteristics change fundamentally due to soil moisture changes in the catchment, and the runoff coefficient can exceed 50%. Therefore, the catchment is highly non-linear in its rainfall–runoff response.

Change

The Kamp catchment, like many Alpine catchments and most of Austria (Merz *et al.*, 2011), experienced air temperature increases during the last decades. The mean air temperature from 1976–1986 was approximately 6.0°C, compared to 7.3°C from 1998–2008. Precipitation and PET increased slightly, but discharge remained constant (Merz *et al.*, 2011). An extraordinary flood occurred in the Kamp catchment during August 2002, which is further described in the sub-section 'Past studies'.

Data and calibration and evaluation periods

Meteorological data were obtained from interpolated local weather stations. PET was estimated by the modified Blaney-Criddle method (Parajka *et al.*, 2003) using daily air temperature and potential sunshine duration. Daily precipitation, temperature, PET and discharge were available from 1976 to 2008.As a consequence, the "complete period" was defined as 1978–2008. The five sub-periods were 6-years long: 1978–1983 (P1), 1984–1989 (P2), 1990–1995 (P3), 1996–2001 (P4) and 2002–2007 (P5). Sub-periods P1 and P2 had average temperatures that were 1°C lower than the other sub-periods. The PET also was slightly lower for P1 and P2 than for the other sub-periods. The only notable characteristic for precipitation and discharge were that they were much higher for P5, primarily due to the 2002 flood.

Past studies

A number of floods have been recorded in this catchment. The flood record was in August 2002, which affected a large portion of Europe (Chorynski *et al.*, 2012). This caused significant damages to the Kamp catchment, which is the reason why this catchment was extensively studied in later years (Komma *et al.*, 2007; Blöschl *et al.*, 2008; Reszler *et al.*, 2008; Viglione *et al.*, 2010). The estimated peak flow was 460 m^3 s^{-1}, which is three times the second largest flood in the 55-year record. The generalized extreme value (GEV) distribution, fitted by the method of L-moments, gives a 100-year flood runoff (Q100) of 285 m^3 s^{-1}. Extrapolating this flood frequency

curve to large return periods results in a return period of 340 years for the 2002 event (Viglione *et al.*, 2013). Note that the study made by Viglione *et al.* (2013) used sub-daily data, while for this workshop daily data were used, which resulted in a daily averaged peak flow estimate of 272 m^3 s^{-1}.

7.9 Obyån Creek at Lissbro

Physical description

The Lissbro station measures the discharge in Obyån Creek, a tributary to the River Mörrumsån in south-eastern Sweden (Fig. 9). The catchment area is 97 km^2, most of which is covered by forested till soils. Lakes occupy only 1% of the catchment area. Elevations range from 140 to 225 m.

Climatic conditions and hydrological regime

The catchment is located in a one of the driest parts of Sweden. The hydrological regime can be described as pluvio-nival (i.e., equally influenced by snowmelt and rainfall), and is dominated by high winter runoff and dry summers. Discharge measurements began in 1984. For 1984–2010, the estimated annual averages of precipitation, temperature and runoff are 800 mm, 6.7ºC and 365 mm, respectively.

Change

On the 8 January 2005, a severe storm called Gudrun (a.k.a. Erwin) hit southern Sweden. The Gudrun storm was one of the three most severe storms in southern Sweden during the past 100 years, with maximum wind speeds around 30 m s^{-1} in the area. In total 18 people died in Sweden during the storm and its aftermath. In the worst hit areas about 8% of trees were blown down, and in the Lissbro catchment the loss of forest was about 40-50 m^3 ha^{-1}. This part of Sweden was also affected by flooding due to intense rainfall in July 2004. Therefore, there was considerable concern about increased flood risk in the area due to the loss of forest following the Gudrun storm.

Data and calibration and evaluation periods

Precipitation and temperature were obtained from SMHI, and they are estimated by interpolation of ground stations data (Johansson, 2000). PET was calculated from the Oudin's formula (Oudin *et al.*, 2005). Daily precipitation, temperature and PET were available from 1 January 1981 to 31 December 2010. Discharge was available only from 18 May 1983 through 31 December 2010. The "complete period" was defined as 1984–2010. The five sub-periods each were defined as 7-year periods: 1984–1988 (P1), 1989–1993 (P2), 1994–1998 (P3), 1999–2003 (P4) and 2006–2010 (P5). Since the Gudrun storm occurred in 2005, P1 to P4 are before the storm, and P5 is after the storm.

Past studies

The authors did not find previous studies of the effects of the Gudrun storm on the hydrology of the Lissbro catchment, but the storm provides a possibility for quantifying the effects of a large-scale change in land-use.

7.10 River Rimbaud at Collobrières

Physical description

The Rimbaud catchment is located in the Maures highlands of Southeastern France, close to the Mediterranean coast. It is part of the Réal Collobrier Research Catchment, and has been managed by Irstea since 1966. It drains an area of 1.4 km². The watershed elevations range from 470 to 622 m (Fig. 10). The average hillside slope is approximately 10% but it can reach 22% close to the talweg. The soil layer is thin (around 30 cm) and is mainly composed of small rock and sand. The vegetation is comprised of bushes, maritime pines and oaks (Cosandey *et al.*, 2005).

Climatic conditions and hydrological regime

Precipitation is abundant in this catchment, totalling around 1100 mm a year. This pattern is due to the humid Mediterranean climate due to the proximity of the sea and orographic effects. Runoff comprises 55% of total precipitation. The River Rimbaud has no discharge during the summer. While precipitation and runoff deficiencies have been rather stable through time (the coefficients of variation are 29% and 26%, respectively), runoff has been much more variable (coefficient of variation of 49%) due to the low capacity of the catchment to store water (Lavabre and Martin, 1997; Martin and Lavabre, 1997).

Change

A wildfire swept the western part of the Maures highlands in August 1990. Over 84% of the Rimbaud catchment was burnt, and only the eastern portion was spared (Puech *et al.*, 1993). In 1990 and 1991 (i.e., immediately after the fire), three floods with discharges higher than 5 m^3 s^{-1} were observed, compared to only three times during the 21 years before the fire. The rainfall events which caused these peak flows were not exceptional (Lavabre *et al.*, 1993). The recovery of the shrubland was rapid. In August 1993, more than 50% of the burnt area was covered by tree seedlings.

Data and calibration and evaluation periods

Precipitation data were obtained from a nearby rain gauge located outside the watershed, and temperature and Penman-Monteith (Monteith, 1965) PET were obtained from Météo-France's SAFRAN database (Quintana-Segui *et al.*, 2008; Vidal *et al.*, 2010). A rain gauge situated inside the catchment was excluded from the analysis of previous studies, because the stationarity of the data of this rain gauge had been questioned by Lavabre *et al.* (2000), due to the modification of its surrounding environment resulting from the fire. Daily precipitation, temperature and PET were available from 1 January 1966 to 31 December 2006. Discharge was available only from 24 August 1967 through 31 December 2006. The "complete period" was defined as 1968–2006. The five sub-periods each were 7-years long: 1968–1974 (P1), 1975–1981 (P2), 1982–1988 (P3), 1991–1997 (P4) and 1998–2004 (P5). Since the fire occurred in 1990, P1 to P3 are before the fire, and P4 to P5 are after the fire.

Additionally to the daily data, hourly precipitation and discharge data were provided for the same periods after the workshop.

Past studies

Several studies of the Rimbaud catchment investigated the effect of the 1990 fire on the forest cover, erosion, the water chemistry, and hydrology. Hydrological studies concluded that the destruction of the forest cover led to increased runoff and increased the frequency of flooding (Lavabre *et al.*, 1993; Lavabre and Martin, 1997; Cosandey *et al.*, 2005). The long-term impact of this forest fire is assessed in Folton *et al.* (2014, this issue).

7.11 Watershed 6 of the Fernow Experimental Forest

Physical description

Watershed 6 is a 0.22 km² catchment located on the U.S. Forest Service's Fernow Experimental Forest (FEF). This watershed is the smallest included in this analysis. The FEF is located in north central West Virginia in the unglaciated section of the Allegheny Plateau. Its elevation ranges from approximately 730 to 860 m (Fig. 11). Hillsides on the catchment are moderately steep, averaging 30 to 40% slope (Edwards and Wood, 1994). Soils are generally only about 1 to 1.5 m deep and are dominated by Calvin silt loams that overlay fractured sandstones and shales of the Hampshire formation (Losche and Beverage, 1967).

Climatic conditions and hydrological regime

Weather and precipitation measurements have been collected on the FEF since the early 1950s. Based on the past 30 years of data, the FEF receives approximately 1460 mm of precipitation annually. Precipitation is relatively evenly distributed throughout the year, though, May, June, and July typically receive the greatest precipitation, while February, September, and October are the driest months (Adams *et al.*, 2012).

Snow is common during the winter months, though it may occur as early as October and as late as May. The duration of snowpacks varies from year to year, depending upon winter weather patterns. In some years an uninterrupted snowpack exists throughout the winter, while during other years snowpacks are temporary due to rain-on-snow events. The mean annual temperature on the FEF for the past 30 years has been 9.3°C. The lowest mean monthly temperature occurs in January, at 2.8°C, and the highest, 20.4°C, occurs in July. Streamflow on watershed 6 is intermittent. Tipically, the flow disappears or becomes negligible in late summer or early fall, when evapotranspiration and soil storage demands exceed precipitation inputs, which remains a modelling challenge for most hydrological models.

Change

Like much of the eastern United States, Watershed 6 was harvested heavily around the turn of the twentieth century, after which it revegetated naturally to a mixed hardwood stand. From March through October 1964, the lower half of the catchment (0.11 km²) was clearcut with approximately 49% of the total watershed basal area removed (averaging 10.46 m^3 ha^{-1}). Following harvesting, it was herbicided annually with a variety of chemicals (Kochenderfer and Wendel, 1983) to retard vegetation regrowth through fall 1969. The upper 11.1 ha were subsequently clearcut from October 1967 to February 1968, then also herbicided annually through fall 1969. These annual herbicide applications were made manually using backpack sprayers. In spring 1973, the watershed was planted with 2-year-old Norway spruce seedlings. In August 1975, hardwoods that had naturally regenerated in the watershed were herbicided using 2, 4, 5-T and again in September 1980 with glyphosate (Edwards and Wood, 1994).

In the late 1980s, the spruce stand achieved canopy closure. After that time, the canopy was very dense, which created a microclimate that was quite different from adjacent hardwood watersheds – air temperature was typically several degrees cooler, relative humidity was substantially higher, and solar radiation inputs were visibly less than the adjacent catchments.

Data and calibration and evaluation periods

Precipitation and temperature were interpolated from two local weather stations. PET was calculated using the Oudin formula (Oudin *et al.*, 2005). Precipitation, temperature, and discharge data were obtained from the U.S. Forest Service research data archive (Edwards and Wood, 2011a; Edwards and Wood, 2011b; Edwards and Wood, 2011c). Daily precipitation, temperature, PET and discharge were available from 1 November 1956 to 31 December 2009. The "complete period" was defined as 1959–2009. The five sub-periods each were 10-years long: 1959–1968 (P1), 1969–1978 (P2), 1979–1988 (P3), 1989–1998 (P4) and 1999–2008 (P5). The harvesting, herbiciding, and stand conversion changes occurred during periods P1 to P3.

Past studies

Annual discharge responded to the combination of clearcutting and denudation to a greater degree and for a longer period than from clearcutting alone (Adams *et al.*, 2012). Annual streamflow increased by about 280 mm, which was approximately double that normally obtained only from clearcutting in the region. The duration of significant annual increases also lasted approximately 15 years, up to double the length without denudation. From 1987, the annual discharge fell below predicted levels, as the conifer stand became established. By the mid-2000s, the annual streamflow was approximately 200 mm below the levels predicted for the watershed if re-growing to hardwoods had been allowed (Adams *et al.*, 2012), due to the greater evapotranspiration demands of the conifers. In 2002, the peakflows and the stormflow volumes were approximately 0.2 m^3 s^{-1} km^{-2} and 2.9 m^3 km^{-2} below their respective pre-harvest levels (Edwards and Watson, 2002).

Comparisons of mean centroid lag times (i.e., the difference in time between the centroid mass of precipitation and the centroid mass of stormflow) were made among six time periods encompassing 34 years from 1957 through 1991 (Edwards and Wood, 1994). Statistical ranks were compared because the centroid lag times were not normally distributed. Overall, centroid lag time analysis was insensitive to detecting hydrologic responses to vegetative cover changes, perhaps due to the small size of the watershed. However, the mean rank for the period that the watershed was completely barren was statistically smaller than the ranks for the last two periods tested (spanning from 1979-1991), during which time the spruce stand had reached canopy closure and other hydrologic changes were becoming evident.

While the watershed supported hardwoods, stream morphology was characteristic of that throughout the FEF and most other streams in the area. It was a several feet-wide A channel (Rosgen, 1996), i.e. a small and steep headwater channel with low sinuosity, with a substrate dominated by medium and coarse gravels. After the Norway spruce became well established, the stream morphology changed dramatically in response to the hydrologic changes that occurred from the growth of the conifers. The channel throughout all but the lower approximately 30 m of length has become U-shaped and narrowed to an average of only about 24-cm wide, with the sides filling in with sediment covered by a thick mat of mosses (Edwards and Watson, 2002). Its sinuosity has increased within the very narrow valley segment that has developed from sediment accumulation in the channel. It appears to have transitioned to a G channel (Rosgen, 1996), i.e. a small and steep headwater channel with moderate sinuosity, with a much greater accumulation of sand and silt in the stream substrate (Edwards and Watson, 2002) than it had when the watershed supported a hardwood stand.

7.12 River Wimmera at Glenorchy Weir Tail

Physical description

The Wimmera region is based around the terminal Wimmera River, Avon River and Yarriambiack Creek. It includes the major centres of Horsham, Stawell and Ouyen. The region covers 3% of Murray-Darling basin (MDB) within western Victoria, Australia. The Wimmera River at Glenorchy Weir Tail lies in the extreme South of the Wimmera region and covers an area of about 2000 km^2 (Fig. 12). The dominant landuse is broad acre cropping of cereals, pulse crops and oilseeds in the central and northern areas and dry land livestock grazing in the south.

Climatic conditions and hydrological regime

The mean annual rainfall is about 460 mm. The rainfall and runoff across the Wimmera region varies substantially with a mean annual value of ~ 800 mm in the south and ~ 300 mm in the north. The rainfall varies considerably between years, but winter is typically the wettest season. The mean annual runoff is 25 mm, with a runoff coefficient of about 0.05. The highest discharges occur during the cold season, between June and October.

Change

The region's rainfall has been relatively consistent over the last 100 years but during the "Millenium Drought" the rainfall was ~ 13% lower than the long-term mean. A severe drought affected most of Australia between 1997 and 2008. As mentioned in the description of the Axe Creek basin, during the "Millenium Drought"(Petheram *et al.*, 2011; van Dijk *et al.*, 2013), there was a 22% reduction in annual rainfall compared to the long term mean but the corresponding reduction in runoff was almost 90% (runoff coefficient of about 0.01).

Data and calibration and evaluation periods

Precipitation, temperature and PET data were available from 1 January 1960 to 31 August 2009. Discharge data were available from 2 January 1965 to 31 August 2009. Precipitation and temperature were provided by the Bureau of Meteorology. PET was calculated by CSIRO, through the Morton formula (Morton, 1983). Discharges were provided by Victorian Warehouse. The "complete period" was defined as the period from 2 January 1965 to 31 August 2009. The 5 sub-periods each are 8-years long: 1966–1973 (P1), 1974–1981 (P2), 1982–1989 (P3), 1990–1997 (P4) and 1998–2005 (P5). P1 to P4 are wet periods and P5 is a dry period occurring completely within the "Millenium Drought" period.

Past studies

The reader is refered to the "Past studies" section of the Axe Creek.

Acknowledgments We would like to thank Thomas Over and Julie Kiang from the US Geological Survey who provided the data for Ferson and Blackberry Creeks. This manuscript is based on or contains data provided by the State of Queensland (Department of Natural Resources and Mines). We also thank IAHS and its STAHY (ICSH) and surface water (ICSW) commissions for their support in organizing this workshop.

References

Adams, M. B., Edwards, P. J., Ford, W. M., Schuler, T. M., Thomas-Van Gundy, M. and Wood, F., 2012. Fernow Experimental Forest: Research history and opportunities, EFR-2. Washington, DC: U.S. Department of Agriculture, Forest Service: 26.

Bartosova, A., Singh, J., Rahim, M. and McConkey, S., 2007a. Fox River Watershed Investigation: Stratton Dam to the Illinois River, Phase II: Hydrologic and Water Quality Simulation Models, Part 2: Validation Blackberry and Poplar Creek HSPF Models, Calibration and Initial Simulation Results. CR 2007-07, Illinois State Water Survey, Champaign, IL.

Bartosova, A., Singh, J., Rahim, M. and McConkey, S., 2007b. Fox River Watershed Investigation: Stratton Dam to the Illinois River, Phase II: Hydrologic and Water Quality Simulation Models, Part 3: Validation of Hydrologic Model Parameters, Brewster Creek, Ferson Creek, Flint Creek, Mill Creek, and Tyler Creek Watersheds. CR 2007-07, Illinois State Water Survey, Champaign, IL.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L. J., Jobes, T. H. and Donigian, A. S. J., 2001. Hydrological Simulation Program - Fortran (HSPF). User's Manual for Release 12, U.S. EPA National Exposure Research Laboratory, Athens, GA, in cooperation with U.S. Geological Survey, Water Resources Division, Reston, VA.

Blodgett, D. L., Booth, N. L., Kunicki, T. C., Walker, J. L. and Viger, R. J., 2011. Description and testing of the Geo Data Portal: Data integration framework and Web processing services for environmental science collaboration, U.S. Geological Survey Open-File Report 2011–1157 from http://pubs.usgs.gov/of/2011/1157/, 9.

Blöschl, G., Reszler, C. and Komma, J., 2008. A spatially distributed flash flood forecasting model. *Environmental Modelling and Software,* 23(4): 464-478, DOI: 10.1016/j.envsoft.2007.06.010.

Caballero, Y., Voirin-Morel, S., Habets, F., Noilhan, J., LeMoigne, P., Lehenaff, A. and Boone, A., 2007. Hydrological sensitivity of the Adour-Garonne river basin to climate change. *Water Resources Research,* 43(7): W07448, DOI: 10.1029/2005WR004192.

CETE, 2009. Dérive hydrologique en Auvergne, Laboratoire régional des ponts et chaussées de Clermont-Ferrand, 1-187.

Changnon, S. A., Angel, J. R., Kunkel, K. E. and Lehmann, C. M., 2004. *Climate Atlas of Illinois*. I. I. S. W. S. Champaign.

Chauveau, M., Chazot, S., Perrin, C., Bourgin, P.-Y., Sauquet, E., Vidal, J.-P., Rouchy, N., Martin, E., David, J., Norotte, T., Maugis, P. and De Lacaze, X.,

2013. Quels impacts des changements climatiques sur les eaux de surface en France à l'horizon 2070 ? *La Houille Blanche,*(4): 5-15, DOI: 10.1051/lhb/2013027.

Chicago Metropolitan Agency for Planning, 2011. Ferson-Otter Creek Watershed Plan.

Chiew, F. H. S., Potter, N. J., Vaze, J., Petheram, C., Zhang, L., Teng, J. and Post, D. A., 2013. Observed hydrologic non-stationarity in far south-eastern Australia: implications for modelling and prediction. *Stochastic Environmental Research and Risk Assessment*: 1-13, DOI: 10.1007/s00477-013-0755-5.

Chorynski, A., Pinskwar, I., Kron, W. and Brakenridge, G. R., 2012. Catalogue of Large Floods in Europe in the 20th Century. Changes in Flood Risk in Europe, CRC Press: 27-54.

Cosandey, C., Andréassian, V., Martin, C., Didon-Lescot, J. F., Lavabre, J., Folton, N., Mathys, N. and Richard, D., 2005. The hydrological impact of the mediterranean forest: A review of French research. *Journal of Hydrology,* 301(1-4): 235-249, DOI: 10.1016/j.jhydrol.2004.06.040.

CSIRO, 2009a. Water in the Flinders-Leichhardt region, Water in the Gulf of Carpentaria Drainage Division. A report to the Australian Government from the CSIRO Northern Australia Sustainable Yields Project. CSIRO Water for a Healthy Country Flagship, Australia, 187-274.

CSIRO, 2009b. Water in the South-East Gulf region, Water in the Gulf of Carpentaria Drainage Division. A report to the Australian Government from the CSIRO Northern Australia Sustainable Yields Project. CSIRO Water for a Healthy Country Flagship, Australia, 275-345.

CSIRO, 2014. Flinders and Gilbert Agricultural Resource Assessment from http://www.csiro.au/Organisation-Structure/Flagships/Water-for-a-Healthy-Country-Flagship/Sustainable-Yields-Projects/Flinders-and-Gilbert-Agricultural-Resource-Assessment-overview.aspx.

Dacharry, M., 1966. Précipitations et écoulements dans le bassin supérieur de l'Allier. Norois: 331-344.

Dezetter, A. and Ruelland, D., 2012. Parameterization based on NOAA-AVHRR NDVI to improve conceptual rainfall-runoff modelling in a large West African catchment, Remote Sensing and Hydrology, IAHS Publ. 352: 221-230.

DNRM, 2006a. Flinders River Basin IQQM Calibration Report, Surface Water Assessment Group, Department of Natural Resources and Mines, Brisbane, Australia.

DNRM, 2006b. Gilbert River Basin IQQM Calibration Report, Surface Water Assessment Group, Department of Natural Resources and Mines, Brisbane, Australia.

Edwards, P. J. and Watson, E. A., 2002. Converting a hardwood forest to spruce: Effects on channel morphology. Annual water resources conference abstract proceedings, Middleburg, Virginia, American Water Resources Association.

Edwards, P. J. and Wood, F., 1994. Centroid lag time changes resulting from harvesting, herbiciding, and stand conversion. Effects of human-induced changes on hydrologic systems: 727-734.

Edwards, P. J. and Wood, F., 2011a. Fernow Experimental Forest daily air temperature. Newtown Square, PA, U.S. Department of Agriculture, Forest Service, Northern Research Station, DOI: 10.2737/RDS-2011-0013.

Edwards, P. J. and Wood, F., 2011b. Fernow Experimental Forest daily precipitation. Newtown Square, PA, U.S. Department of Agriculture, Forest Service, Northern Research Station, DOI: 10.2737/RDS-2011-0014.

Edwards, P. J. and Wood, F., 2011c. Fernow Experimental Forest daily streamflow. Newtown Square, PA, U.S. Department of Agriculture, Forest Service, Northern Research Station, DOI: 10.2737/RDS-2011-0015.

Federal Emergency Management Agency, 2002. Flood insurance study, Kane County, Illinois (unincorporated areas): community number 170896.

Folton, N., Andréassian, V. and Duperray, R., 2014, this issue. Looking at the hydrological impact of forest-fire from two different perspectives: paired-catchment vs rainfall-runoff modelling. *Hydrological Sciences Journal*.

François, B., Hingray, B., Hendrickx, F. and Creutin, J. D., 2013. Storage water value as a signature of the climatological balance between resource and uses. *Hydrol. Earth Syst. Sci. Discuss.,* 10(7): 8993-9025, DOI: 10.5194/hessd-10-8993-2013.

Garçon, R., 1996. Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR. Bilan de l'année 1994-1995. *La Houille Blanche,*(5): 71-76, DOI: 10.1051/lhb/1996056.

Hendrickx, F. and Sauquet, E., 2013. Impact of warming climate on water management for the Ariège River basin (France). *Hydrological Sciences Journal,* 58(5): 976-993, DOI: 10.1080/02626667.2013.788790.

Imbeaux, E., 1892. La Durance : régime, crues et inondations, Paris : Vve Ch. Dunod, 200 pp.

Jeffrey, S. J., Carter, J. O., Moodie, K. B. and Beswick, A. R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling & Software,* 16(4): 309-330, DOI: 10.1016/S1364-8152(01)00008-1.

Johansson, B., 2000. Areal precipitation and temperature in the Swedish mountains : An evaluation from a hydrological perspective. Lyngby, DANEMARK, Nordic Association for Hydrology.

Knapp, H. V. and Myers, M. W., 1999. Fox River Streamflow Assessment Model: 1999 Update to the Hydrologic Analysis. CR-649, Illinois State Water Survey, Champaign, IL from http://www.isws.illinois.edu/pubs/abstract.asp?PID=2055.

Kochenderfer, J. N. and Wendel, G. W., 1983. Plant succession and hydrologic recovery on a deforested and herbicided watershed. *Forest Science,* 29(3): 545-558.

Komma, J., Reszler, C., Blöschl, G. and Haiden, T., 2007. Ensemble prediction of floods - Catchment non-linearity and forecast probabilities. *Natural Hazards and Earth System Science,* 7(4): 1-14, DOI: 10.5194/nhess-7-431-2007.

Kuentz, A., Mathevet, T., Gailhard, J., Perret, C. and Andréassian, V., 2013. Over 100 years of climatic and hydrologic variability of a Mediterranean and mountainous watershed: the Durance River. Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections Proceedings of H02, IAHS-IAPSO-IASPEI Assembly, Gothenburg, Sweden, July 2013, IAHS Publ. 360: 19-25.

L'Hôte, Y., Mahé, G. I. L., Somé, B. and Triboulet, J. P., 2002. Analysis of a Sahelian annual rainfall index from 1896 to 2000; the drought continues. *Hydrological Sciences Journal,* 47(4): 563-572, DOI: 10.1080/02626660209492960.

Lafaysse, M., Hingray, B., Etchevers, P., Martin, E. and Obled, C., 2011. Influence of spatial discretization, underground water storage and glacier melt on a physically-

based hydrological model of the Upper Durance River basin. *Journal of Hydrology,* 403(1–2): 116-129, DOI: 10.1016/j.jhydrol.2011.03.046.

Laurent, F. and Ruelland, D., 2010. Physically-based modelling of hydroclimatic variability across a large tropical watershed. Global Change: Facing Risks and Threats to Water Resources. *IAHS Publ*, 340: 474-484.

Lavabre, J. and Martin, C., 1997. Impact d'un incendie de forêt sur l'hydrologie et l'érosion hydrique d'un petit bassin versant méditerranéen. *IAHS Publ.,* 245: 39-47.

Lavabre, J., Martin, C. and Folton, N., 2000. Impact de l'incendie sur le comportement hydrologique du bassin versant. Chapitre III. In : Conséquences d'un incendie de forêt dans le bassin versant du Rimbaud (massif des Maures, Var, France) : destruction et régénération du couvert végétal, impacts sur l'hydrologie, l'hydrochimie et les phénomènes d'érosion mécanique., Cemagref Editions, Coll. Etudes - Gestion des milieux aquatiques. 16: 33-49.

Lavabre, J., Sempere Torres, D. and Cernesson, F., 1993. Changes in the hydrological response of a small Mediterranean basin a year after a wildfire. *Journal of Hydrology,* 142(1-4): 273-299, DOI: 10.1016/0022-1694(93)90014-Z.

Le Barbé, L., Lebel, T. and Tapsoba, D., 2002. Rainfall Variability in West Africa during the Years 1950–90. *Journal of Climate,* 15(2): 187-202, DOI: 10.1175/1520-0442(2002)015<0187:RVIWAD>2.0.CO;2.

Leighton, M. M., Ekblaw, G. E. and Horberg, L., 1948. Physiographic divisions of Illinois: State Geological Survey, Department of Registration and Education. *Reprinted from the Journal of Geology,* 56(1): 33.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P. and Loumagne, C., 2014. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrol. Earth Syst. Sci.,* 18(2): 575-594, DOI: 10.5194/hess-18-575-2014.

Losche, C. K. and Beverage, W. W., 1967. Soil survey of Tucker County and part of northern Randolph County, West Virginia, U.S. Department of Agriculture, Soil Conservation Service, Forest Service, and West Virginia Agriculture Station, 78.

Mann, H. B., 1945. Non parametric tests against trend. *Econometrica,* 13: 245-259.

Martin, C. and Lavabre, J., 1997. An estimation of the contribution of slope runoff to the Rimbaud stream floods (massif des Maures, Var, France) after the August 1990 forest fire. *Hydrological Sciences Journal,* 42(6): 893-907, DOI: 10.1080/02626669709492086.

Merz, R. and Blöschl, G., 2005. Flood frequency regionalisation - Spatial proximity vs. catchment attributes. *Journal of Hydrology,* 302(1-4): 283-306, DOI: 10.1016/j.jhydrol.2004.07.018.

Merz, R., Parajka, J. and Blöschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research,* 47(2), DOI: 10.1029/2010WR009505.

Mitchell, T. D. and Jones, P. D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology,* 25(6): 693-712, DOI: 10.1002/joc.1181.

Moisselin, J. M., Schneider, M., Canellas, C. and Mestre, O., 2002. Changements Climatiques en France au 20ème siècle - Étude des longues séries de données homogénéisées françaises de précipitations et températures. *La Météorologie*: 45-56.

Monteith, J. L., 1965. Evaporation and environment. Symposia of the Society for Experimental Biology, 19: 205-234.

Morlat, G., Billiet, A. and Bernier, J., 1956. Les crues de la haute Durance et la théorie des valeurs extrêmes. *IAHS Publications,* 42: 99-114.

Morton, F. I., 1983. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology,* 66(1–4): 1-76, DOI: 10.1016/0022-1694(83)90177-4.

Murphy, E. A., Straub, T. D., Soong, D. T. and Hamblen, C. S., 2007. Hydrologic, hydraulic, and flood analyses of the Blackberry Creek watershed, Kendall County, Illinois. Scientific Investigations Report 2007-5141, U.S. Geological Survey from http://pubs.usgs.gov/sir/2007/5141/, 49.

Nicholson, S. E., Tucker, C. J. and Ba, M. B., 1998. Desertification, Drought, and Surface Vegetation: An Example from the West African Sahel. *Bulletin of the American Meteorological Society,* 79(5): 815-829, DOI: 10.1175/1520-0477(1998)079<0815:DDASVA>2.0.CO;2.

O'Hearn, M. and Gibb, J. P., 1980. Groundwater Discharge to Illinois Streams. ISWS CR-246, Illinois State Water Survey, Champaign, IL from http://www.isws.illinois.edu/pubs/abstract.asp?PID=1278.

Onde, H., 1923. Les crues de l'Allier. *Revue de géographie alpine*: 301-372.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andreassian, V., Anctil, F. and 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, DOI: 10.1016/j.jhydrol.2004.08.026.

Paquet, E. and Garçon, R., 2002. Caprices du climat et de l'hydrologie en Haute-Durance : nos prévisions d'apports plurimensuels sont-elles encore fiables ? *La Houille Blanche,*(8): 62-68, DOI: 10.1051/lhb/2002109.

Parajka, J., Merz, R. and Blöschl, G., 2003. Estimation of daily potential evapotranspiration for regional water balance modeling in Austria. In: 11th. International Poster Day and Institute of Hydrology Open Day "Transport of Water, Chemicals and Energy in the Soil - Crop Canopy - Atmosphere System", Slovak Academy of Sciences, Bratislava, 299-306.

Parajka, J., Merz, R. and Blöschl, G., 2005. Regional water balance components in Austria on a daily basis. *Regionale Wasserbilanzkomponenten für Österreich auf Tagesbasis,* 57(3-4): 43-56.

Pardé, M., 1925. Le Régime du Rhône. Première partie, Etude générale, Lyon : Faculté des Lettres et P. Masson.

Petheram, C., Potter, N. J., Vaze, J., Chiew, F. H. S. and Zhang, L., 2011. Towards better understanding of changes in rainfall-runoff relationships during the recent drought in south-eastern Australia. 19th International Congress on Modelling and Simulation, Perth, Australia: 3622-3628.

Potter, N. J. and Chiew, F. H. S., 2011. An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray-Darling Basin using rainfall-runoff models. *Water Resources Research,* 47(12): W00G10, DOI: 10.1029/2010WR010333.

Potter, N. J., Chiew, F. H. S. and Frost, A. J., 2010. An assessment of the severity of recent reductions in rainfall and runoff in the Murray–Darling Basin. *Journal of Hydrology,* 381(1–2): 52-64, DOI: 10.1016/j.jhydrol.2009.11.025.

Potter, N. J., Petheram, C. and Zhang, L., 2011. Sensitivity of streamflow to rainfall and temperature in south-eastern Australia during the Millennium drought. 19th International Congress on Modelling and Simulation, Perth, Australia: 3636-3642.

Puech, C., Vine, P. and Lavabre, J., 1993. Remote sensing techniques in the mapping of vegetation and their application to runoff evolution in burnt areas. *Acta Geologica Hispanica,* 28(2-3): 103-110.

Quintana-Segui, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L. and Morel, S., 2008. Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France. *Journal of Applied Meteorology and Climatology,* 47(1): 92-107, DOI: 10.1175/2007jamc1636.1.

Reszler, C., Blöschl, G. and Komma, J., 2008. Identifying runoff routing parameters for operational flood forecasting in small to medium sized catchments. *Hydrological Sciences Journal,* 53(1): 112-129, DOI: 10.1623/hysj.53.1.11.

Rosgen, D., 1996. *Applied River Morphology*, Pagosa Springs, Colorado: Wildland Hydrology.

Ruelland, D., Ardoin-Bardin, S., Billen, G. and Servat, E., 2008. Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa. *Journal of Hydrology,* 361(1-2): 96-117, DOI: 10.1016/j.jhydrol.2008.07.049.

Ruelland, D., Ardoin-Bardin, S., Collet, L. and Roucou, P., 2012. Simulating future trends in hydrological regime of a large Sudano-Sahelian catchment under climate change. *Journal of Hydrology,* 424-425: 207-216, DOI: 10.1016/j.jhydrol.2012.01.002.

Ruelland, D., Guinot, V., Levavasseur, F. and Cappelaere, B., 2009. Modelling the long-term impact of climate change on rainfall-runoff processes over a large Sudano-Sahelian catchment, In: New approaches to hydrological prediction in data-sparse regions. IAHS Publ. 333: 59-68.

Ruelland, D., Larrat, V. and Guinot, V., 2010a. A comparison of two conceptual models for the simulation of hydro-climatic variability over 50 years in a large Sudano-Sahelian catchment, In: Global Change: Facing Risks and Threats to Water Resources. IAHS Publ. 340: 668-678.

Ruelland, D., Levavasseur, F. and Tribotté, A., 2010b. Patterns and dynamics of land-cover changes since the 1960s over three experimental areas in Mali. *International Journal of Applied Earth Observation and Geoinformation,* 12(SUPPL. 1): S11-S17, DOI: 10.1016/j.jag.2009.10.006.

Ruelland, D., Tribotté, A., Puech, C. and Dieulin, C., 2011. Comparison of methods for LUCC monitoring over 50 years from aerial photographs and satellite images in a Sahelian catchment. *International Journal of Remote Sensing,* 32(6): 1747-1777, DOI: 10.1080/01431161003623433.

Sauquet, E., Dupeyrat, A., Hendrickx, F., Perrin, C., Samie, R. and Vidal, J. P., 2010. IMAGINE 2030, climate and water management: uncertainties on water resources for the Garonne river basin in 2030? , Final report from http://cemadoc.irstea.fr/oa/PUB00028876-imagine-2030-climat-amenagements-garonne-quelles-i.html.

Serra, L., 1953. Les études hydrologiques sur la Haute-Durance. *La Houille Blanche*: 725-734, DOI: 10.1051/lhb/1953007.

Singh, K. P., 1981. Derivation and Regionalization of Unit Hydrograph Parameters for Illinois (Dam Safety Program), Illinois State Water Survey, Champaign, IL 84 p. from http://www.isws.illinois.edu/pubs/pubdetail.asp?CallNumber=ISWS+CR-258.

Soong, D. T., Ishii, A. L., Sharpe, J. B. and Avery, C. F., 2004. Estimating flood-peak discharge magnitudes and frequencies for rural streams in Illinois, U.S. Geological Survey Scientific Investigation Report 2004–5103, 147.

Soong, D. T., Murphy, E. A. and Straub, T. D., 2009. Effect of detention basin release rates on flood flows—Application of a model to the Blackberry Creek Watershed in Kane County, Illinois, U.S. Geological Survey Scientific Investigations Report 2009–5106 from http://pubs.usgs.gov/sir/2009/5106/, 33.

Soong, D. T., Straub, T. D. and Murphy, E. A., 2005. Continuous Hydrologic Simulation and Flood-Frequency, Hydraulic, and Flood-Hazard Analysis of the Blackberry Creek Watershed, Kane County, Illinois, Geological Survey Scientific Investigations Report 2005–5270 from http://pubs.usgs.gov/sir/2005/5270/, 65.

Theobald, D., 2005. Landscape patterns of exurban growth in the USA from 1980 to 2020. *Ecology and Society,* 10(32).

Thirel, G., Andréassian, V., Perrin, C., Audouy, J.-N., Berthet, L., Edwards, P. J., Folton, N., Furusho, C., Kuentz, A., Lerat, J., Lindström, G., Martin, E., Mathevet, T., Merz, R., Parajka, J., Ruelland, D. and Vaze, J., 2014, this issue. Hydrology under change. An evaluation protocol to investigate how hydrological models deal with changing catchments. *Hydrological Sciences Journal*.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wilhelmi, N., Wei, Y. and Cook, R. B., 2012. Daymet: Daily surface weather on a 1 km grid for North America, 1980 - 2011. Acquired online (http://daymet.ornl.gov/) on 23/12/2012 from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A., DOI: 10.3334/ORNLDAAC/Daymet_V2.

U. S. Department of Agriculture, 1976. Technical Release 61, WSP2 Computer Program, Soil Conservation Service.

U. S. Department of Agriculture, 1989. Floodplain management study, Blackberry Creek and Tributaries: Kane-Kendall Counties, Illinois, 24.

U.S. Department of Agriculture, 1992. TR-20 Computer Program for Project Formulation Hydrology, Soil Conservation Service from ftp://ftp.wcc.nrcs.usda.gov/wntsc/H&H/other/tr20userManual.pdf.

U.S. Department of Commerce, 2001. Climatological data annual summary, Illinois: Asheville, N. C., Environmental Data Information Service, National Oceanic and Atmospheric Administration.

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., Timbal, B. and Viney, N. R., 2013. The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research,* 49(2): 1040-1057, DOI: 10.1002/wrcr.20123.

Vaze, J., Chiew, F., Perraud, J.-M., Viney, N., Post, D., Teng, J., Wang, B., Lerat, J. and Goswami, M., 2010a. Rainfall-runoff Modelling across Southeast Australia: Datasets, Models and Results. *Australian Journal of Water Resources,* 14(2): 101-116.

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R. and Teng, J., 2010b. Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology,* 394(3–4): 447-457, DOI: 10.1016/j.jhydrol.2010.09.018.

Vidal, J. P., Martin, E., Franchisteguy, L., Baillon, M. and 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, DOI: 10.1002/joc.2003.

Viglione, A., Chirico, G. B., Komma, J., Woods, R., Borga, M. and Blöschl, G., 2010. Quantifying space-time dynamics of flood event types. *Journal of Hydrology,* 394(1-2): 213-229, DOI: 10.1016/j.jhydrol.2010.05.041.

Viglione, A., Merz, R., Salinas, J. L. and Blöschl, G., 2013. Flood frequency hydrology: 3. A Bayesian analysis. *Water Resources Research,* 49(2): 675-692, DOI: 10.1029/2011WR010782.

Walton, W. C., 1965. Ground-Water Recharge and Runoff in Illinois, Report of Investigation 48. Illinois Water Survey, Urbana, Illinois: 35-54.

Wilhelm, I., 1913. La Durance, étude de l'utilisation de ses eaux et de l'amélioration de son régime par la création de barrages, par Ivan Wilhelm, L. Laveur.

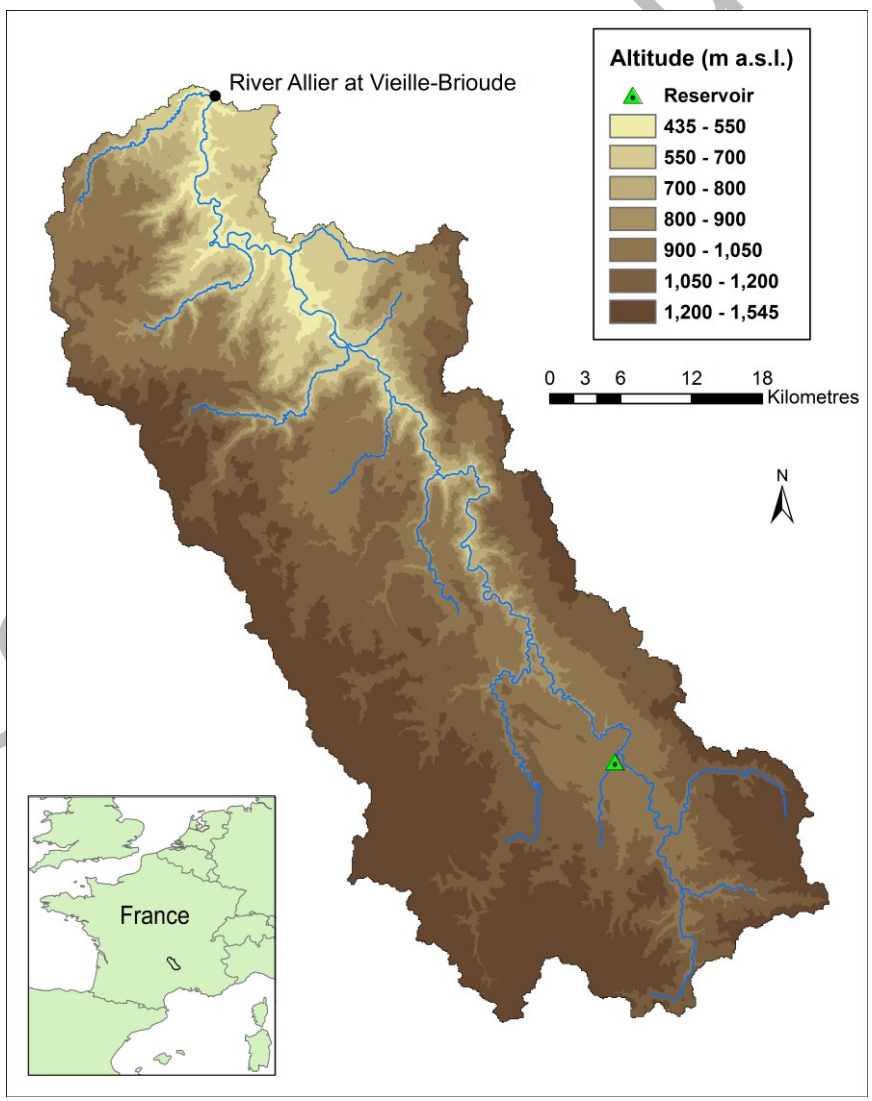

Fig. 1 *The River Allier catchment at Vieille-Brioude*

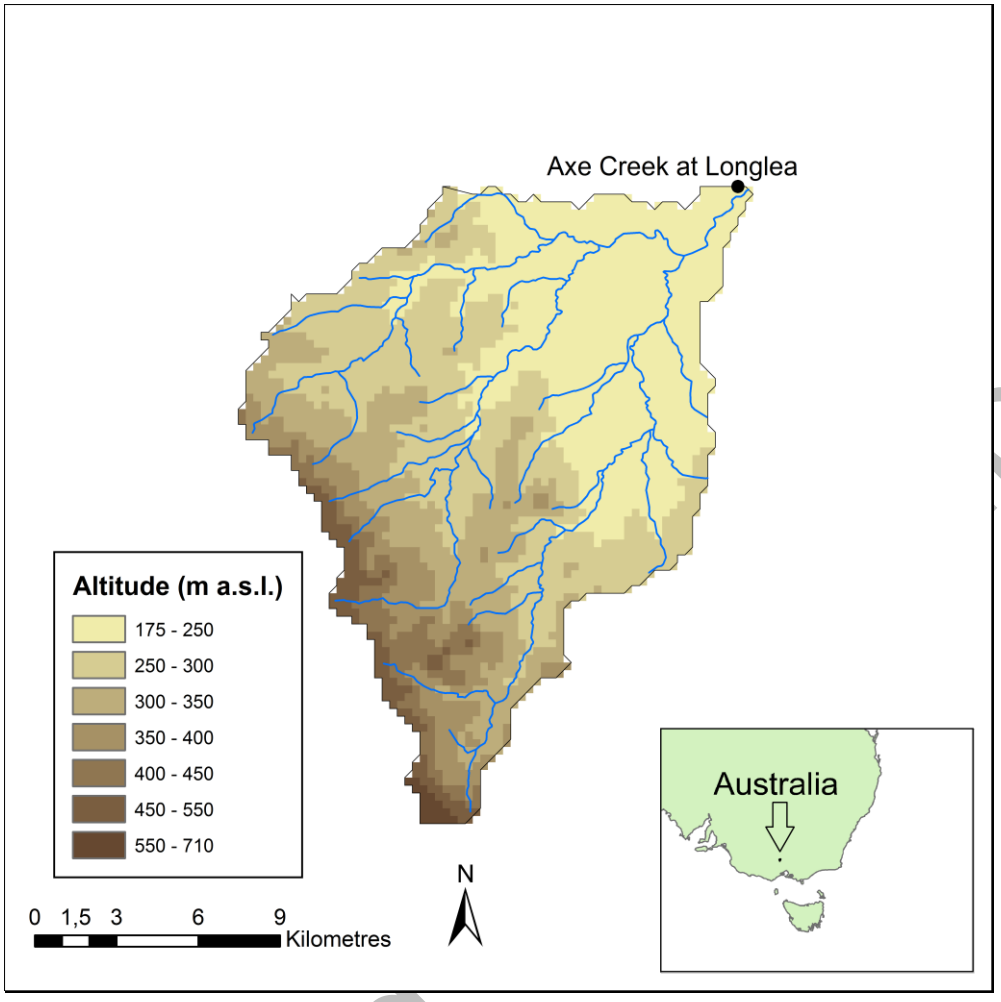

Fig. 2 *The Axe Creek catchment at Longlea*

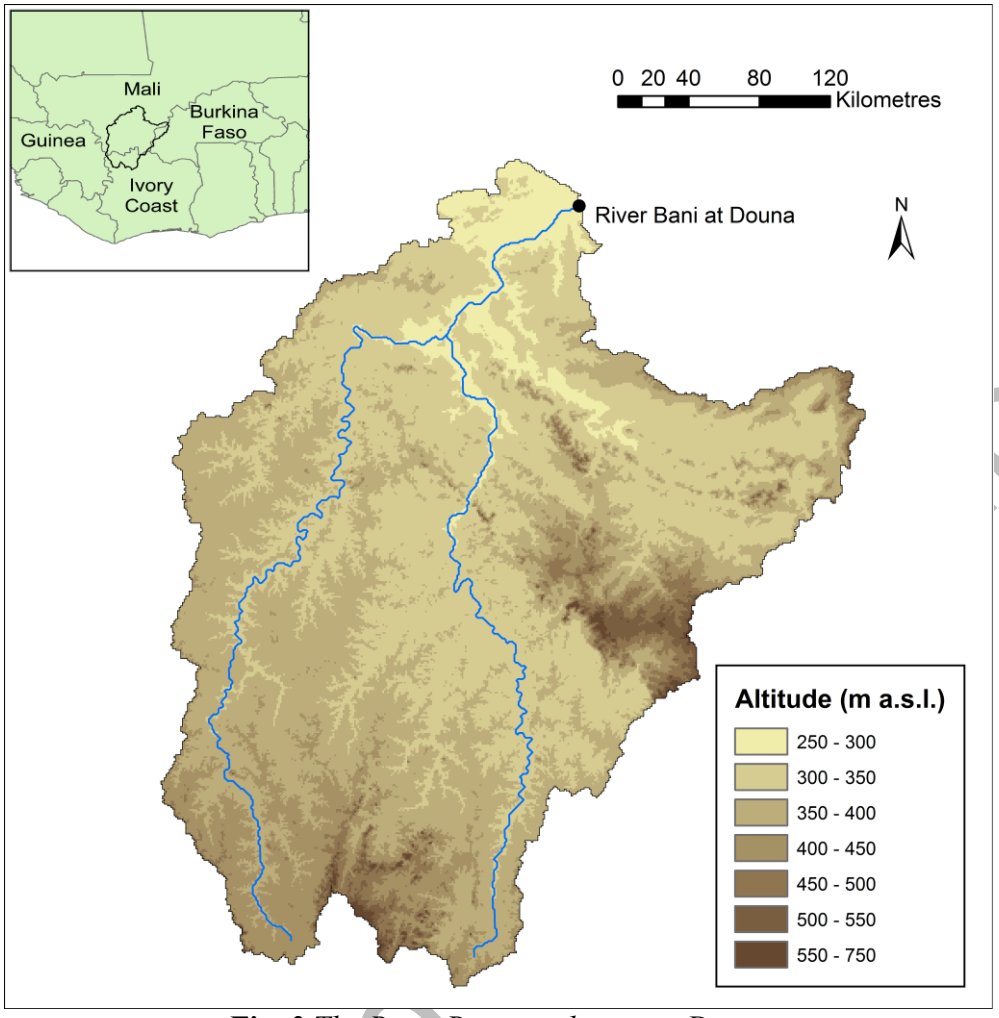

Fig. 3 *The River Bani catchment at Douna*

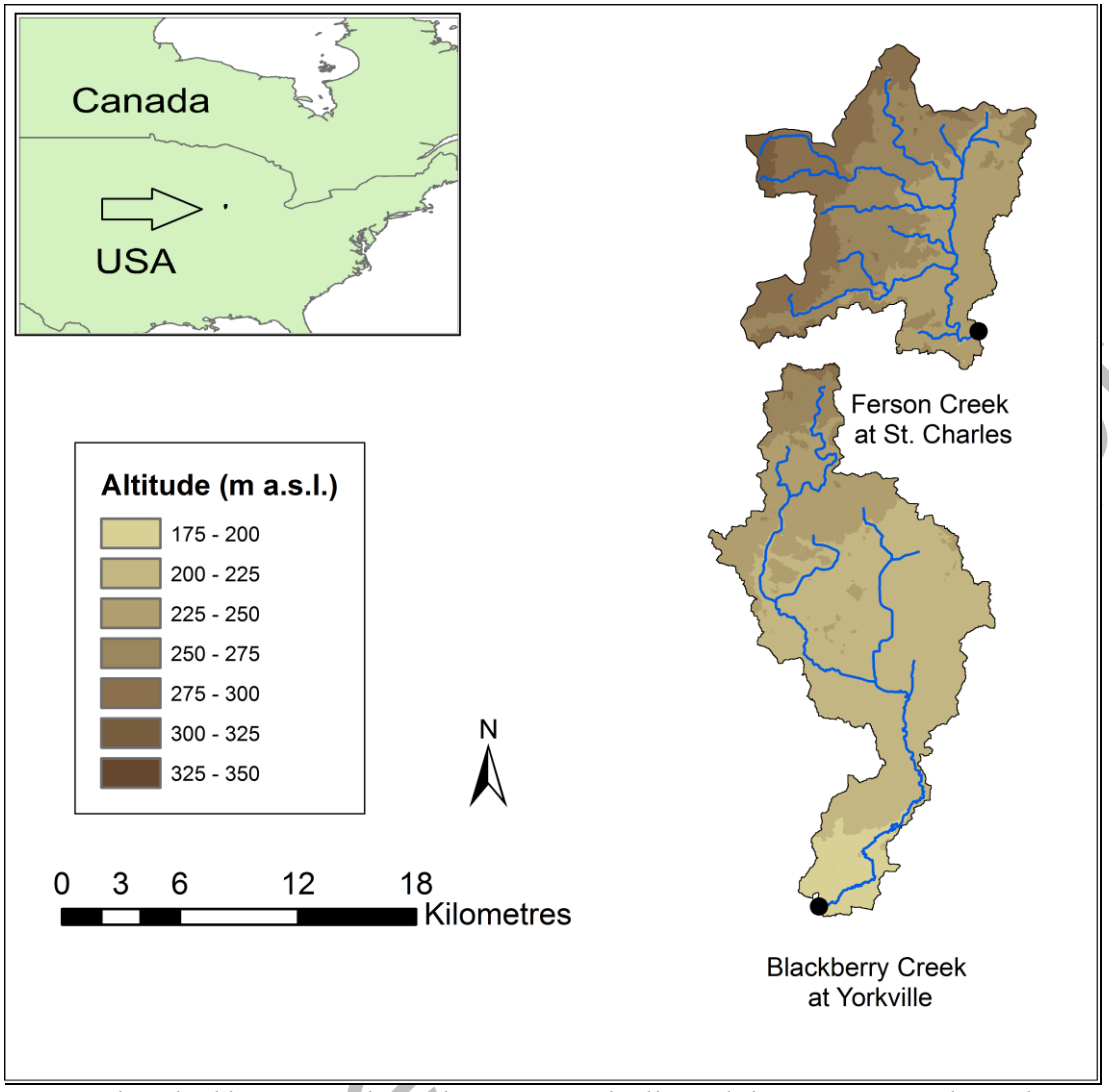

Fig. 4 *The Blackberry Creek catchment at Yorkville and the Ferson Creek catchment at St. Charles.*

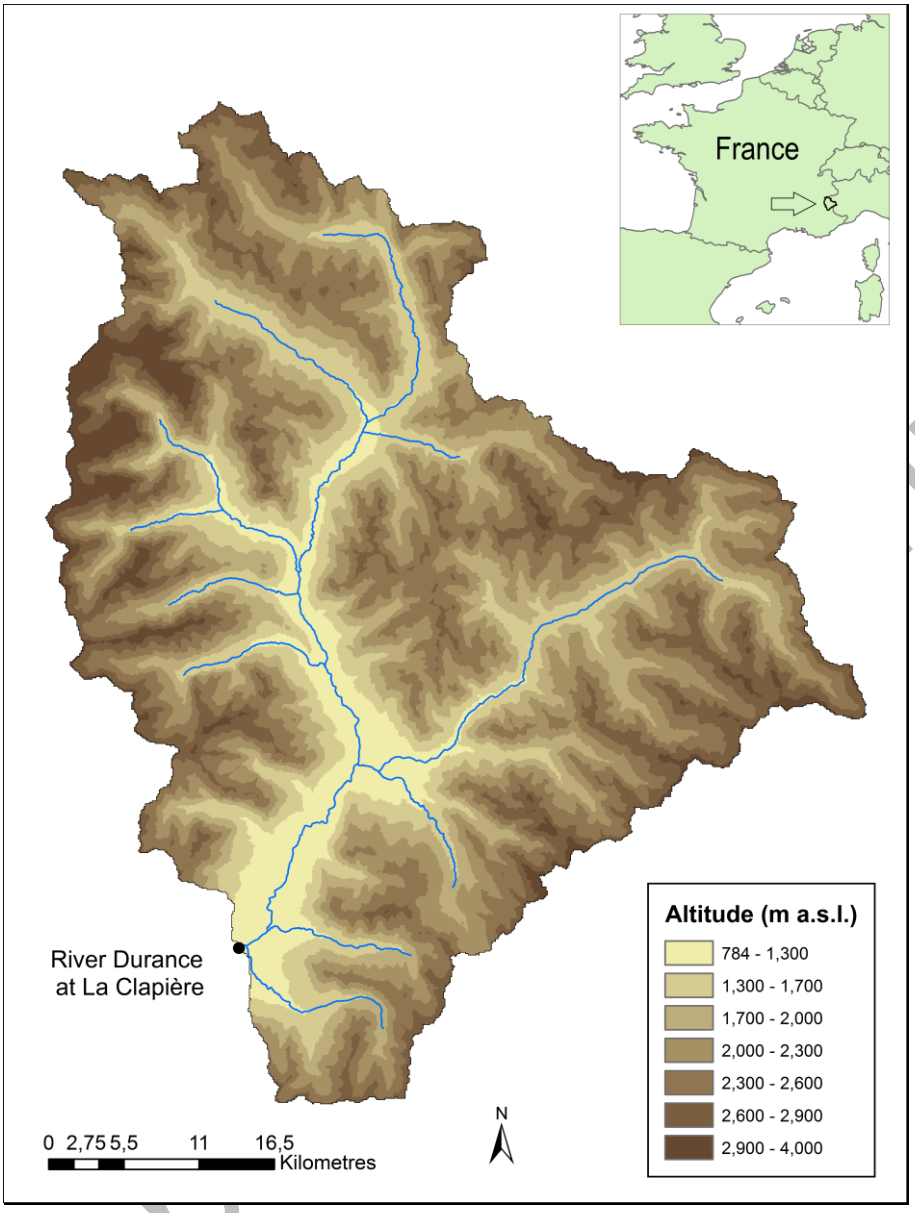

Fig.5 *The River Durance catchment at La Clapière*

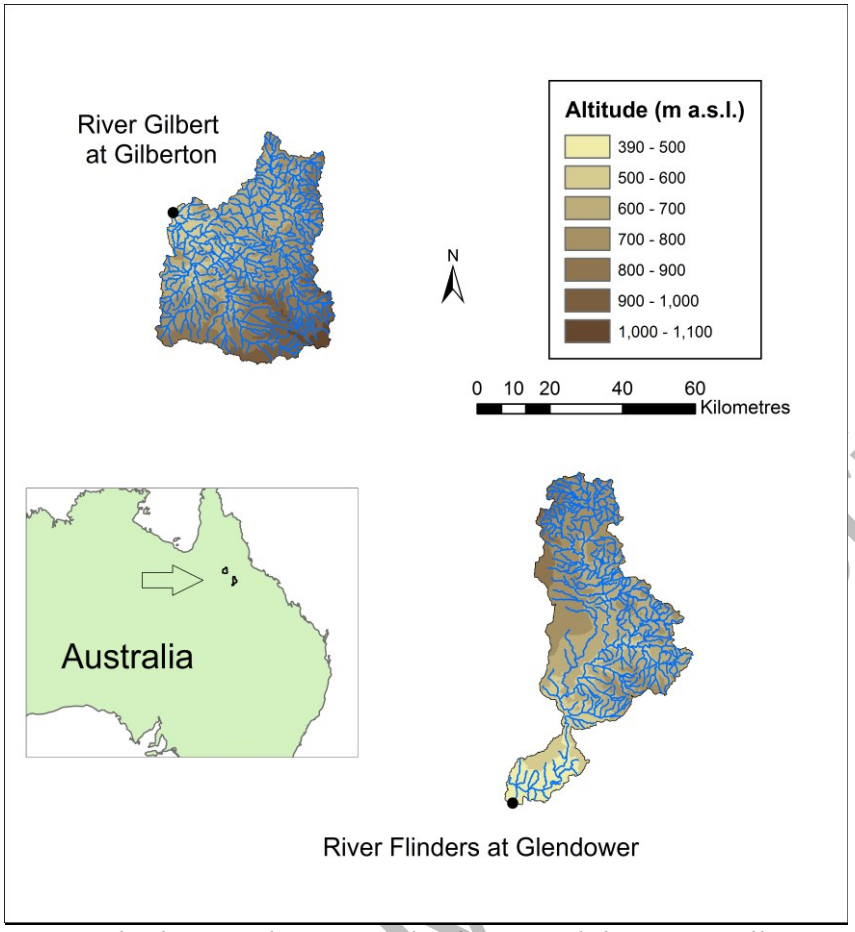

Fig. 6 *The River Flinders catchment at Glendower and the River Gilbert catchment at Gilberton*

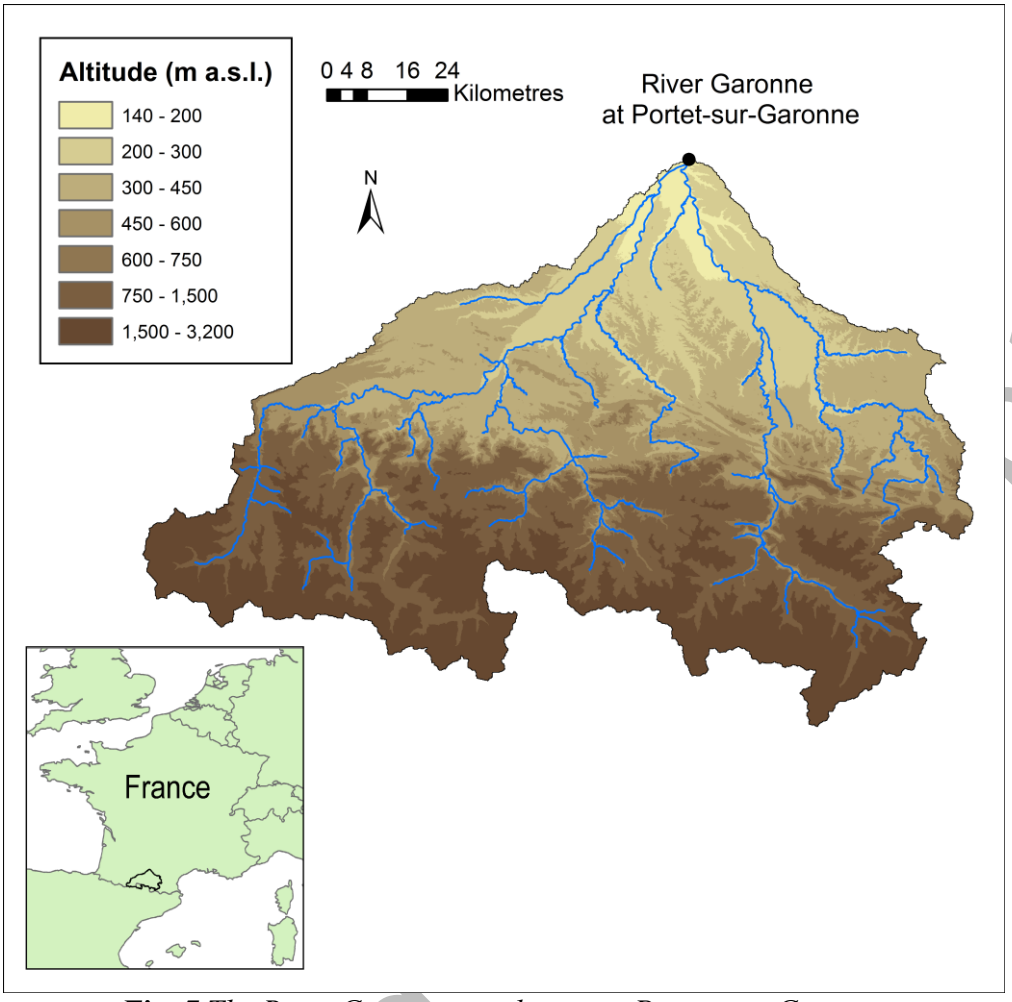

Fig. 7 *The River Garonne catchment at Portet-sur-Garonne*

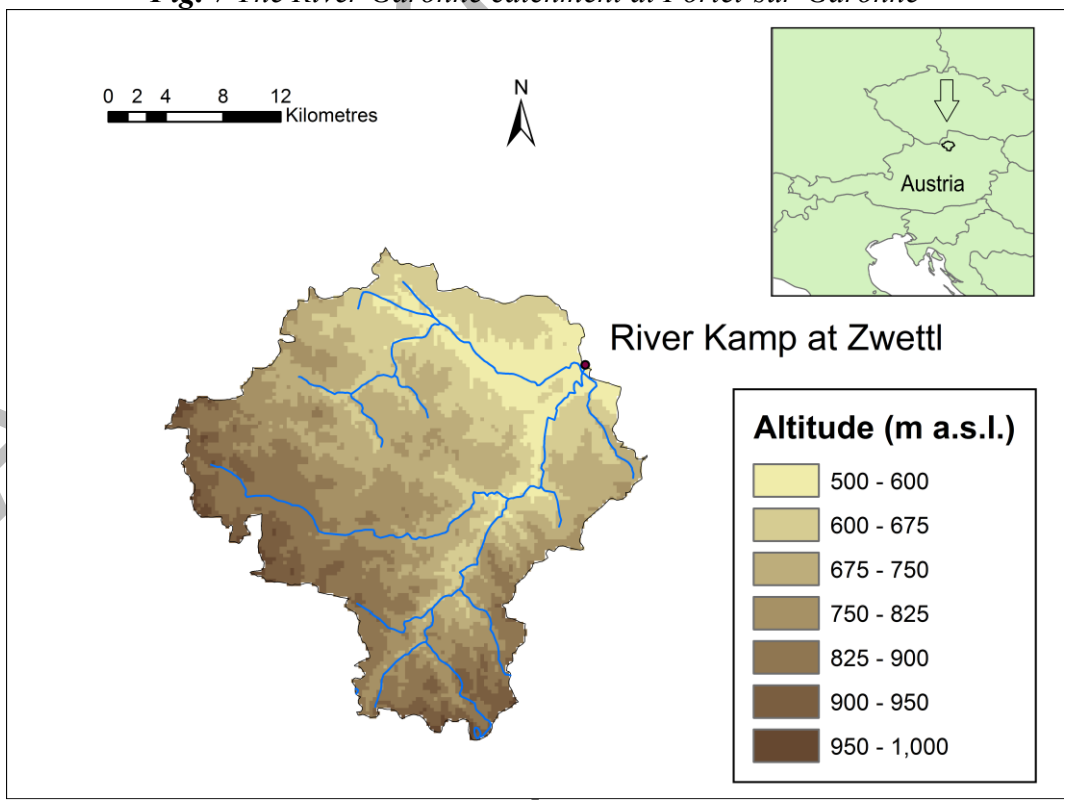

Fig. 8 *The River Kamp catchment at Zwettl*

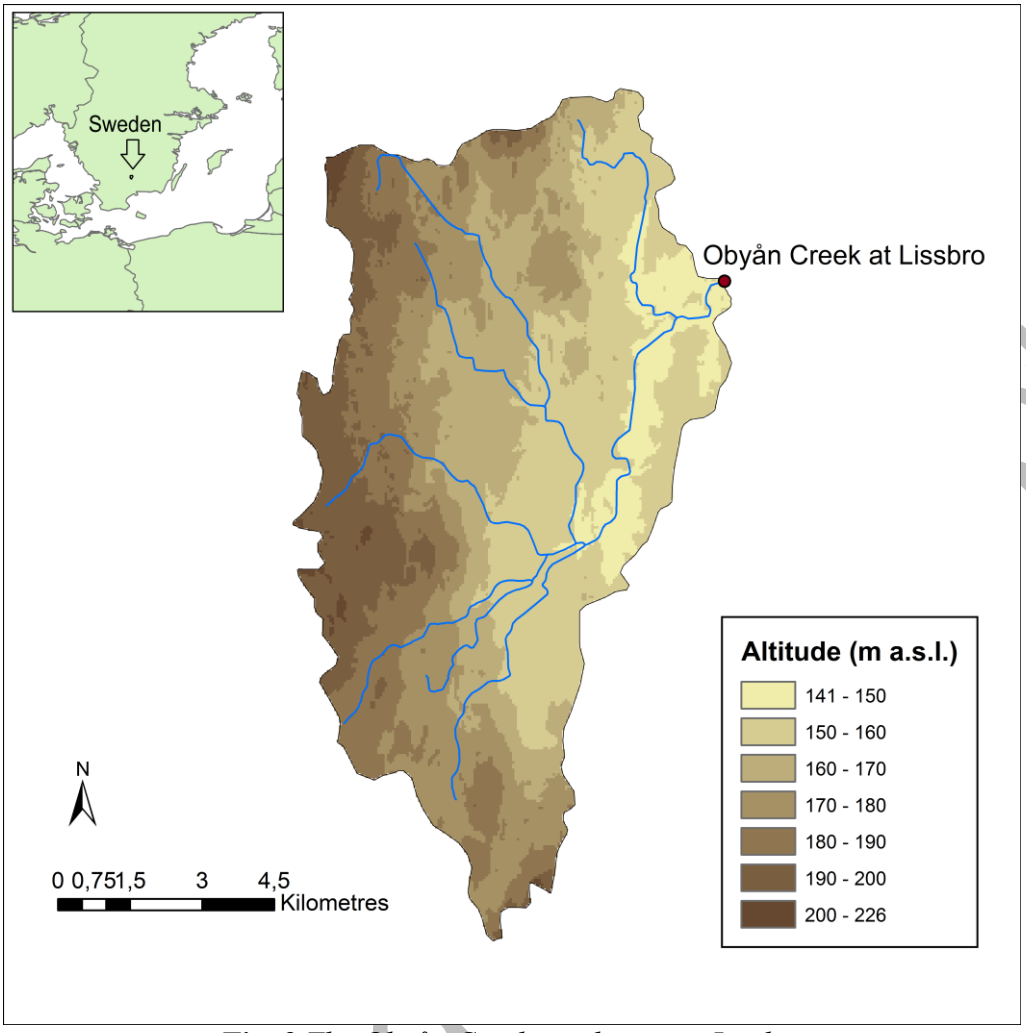

Fig. 9 *The Obyån Creek catchment at Lissbro*

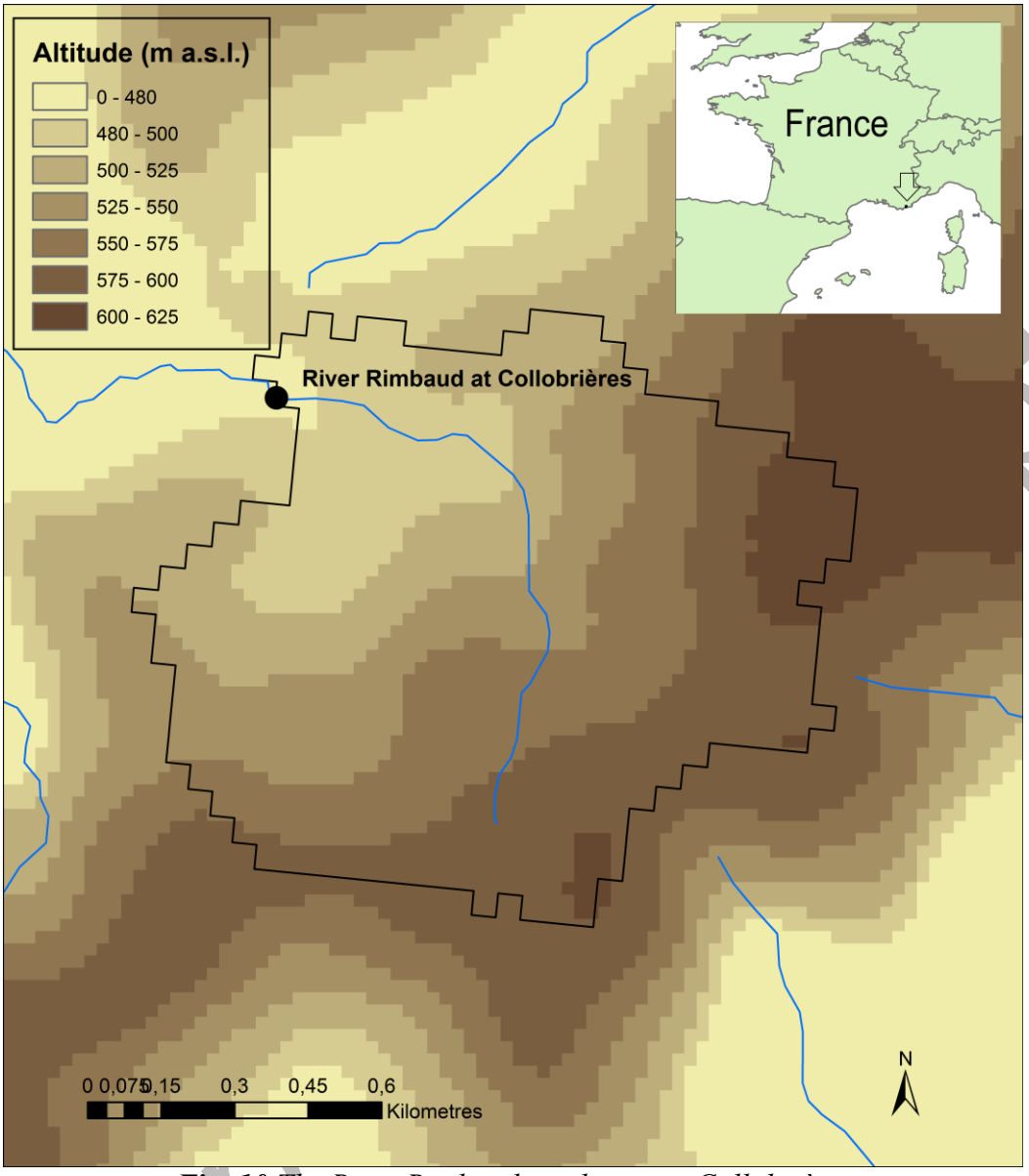

Fig. 10 *The River Rimbaud catchment at Collobrières*

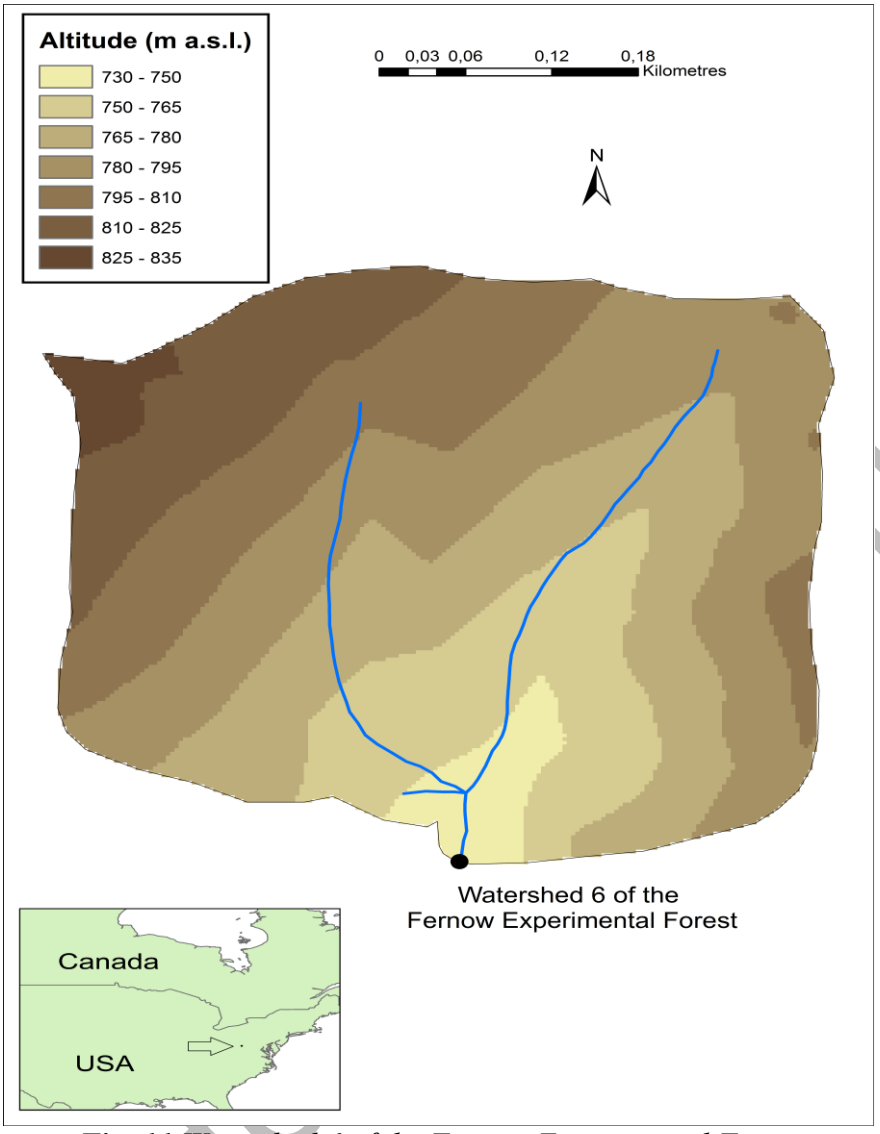

Fig. 11 *Watershed 6 of the Fernow Experimental Forest*

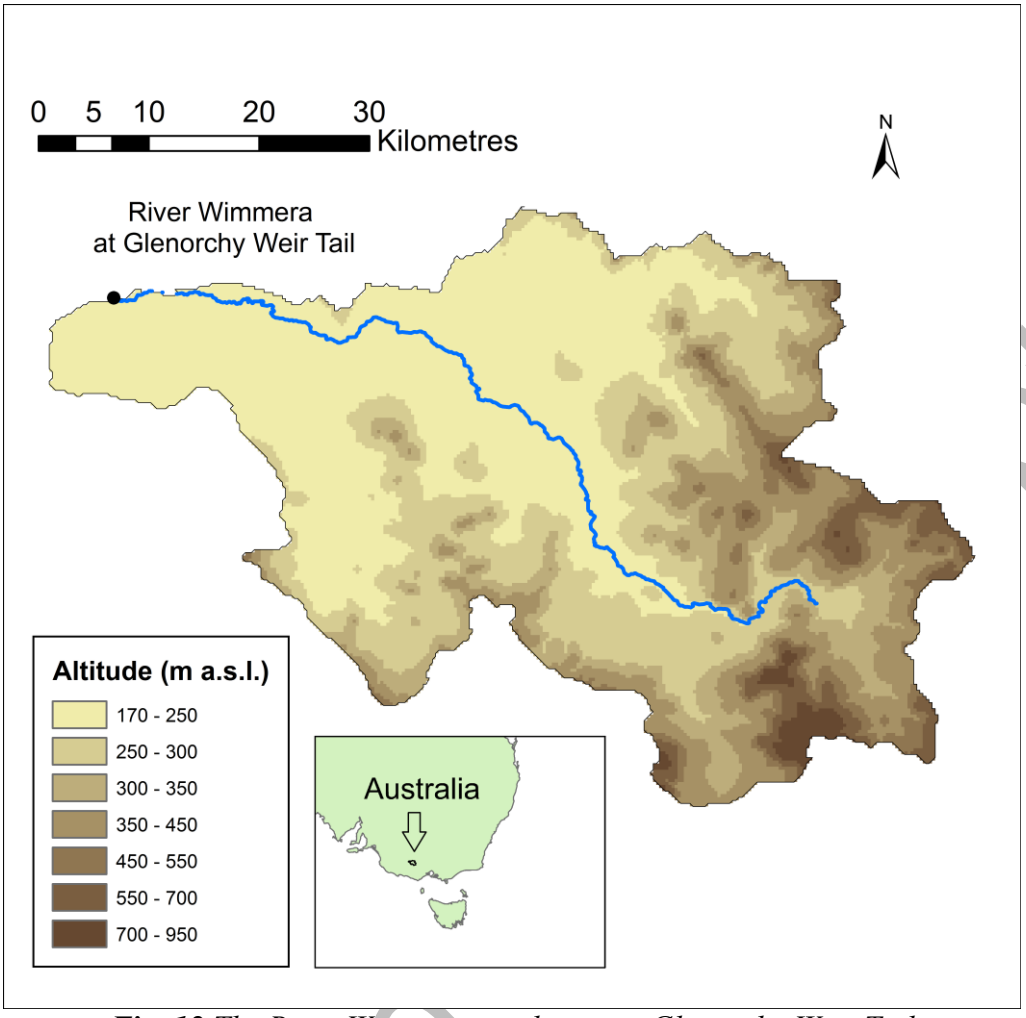

Fig. 12 *The River Wimmera catchment at Glenorchy Weir Tail*

Table 1 *Monthly low-flow values and 95% confidence intervals for different return periods for River Allier at Vieille-Brioude (source: HYDRO database).*

	Discharges in $m^3 s^{-1}$	
Return period of low flows (in years)	1919–1982	1983–2012
2	3.6 [3.1–4.1]	8.6 [8.1–9.1]
5	2.2 [1.8–2.6]	7.6 [7.1–8.1]
10	1.7 [1.4–2.0]	7.2 [6.6–7.6]
20	1.4 [1.1–1.7]	6.9 [6.2–7.3]
50	1.1 [0.8–1.4]	6.5 [5.8–7.0]

Table 2 *Urban area fraction over the Blackberry and Ferson Creeks watersheds for each period of the protocol, obtained from yearly values*

Basin	Complete period	P1	P2	P3	P4	P5
Blackberry Creek	23%	15%	17%	20%	27%	39%
Ferson Creek	39%	26%	32%	39%	48%	60%