
UNIVERSITÉ PARIS SUD

Master 2 Hydrologie, Hydrogéologie et sols

Quelles perspectives de l'intégration de l'expertise dans le calage de modèle hydrologique ?

Benjamin CONTE

Institut National de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

Équipe HBAN-HYDRO

Encadrants : Pierre NICOLLE et François TILMANT



Table des matières

Résumé	3
Contexte de l'étude	4
Introduction	5
I. Objet de l'étude	6
<i>I.1 La prévision des étiages</i>	6
I.1.1 L'étiage	6
I.1.2 La modélisation hydrologique.....	6
I.1.3 Calage d'un modèle hydrologique.....	7
I.1.4 L'expertise : une étape indispensable.....	8
<i>I.2 PREMHYCE : un outil de prévision des débits d'étiage</i>	9
<i>I.3 Données</i>	10
<i>I.4 Bassins versants</i>	10
I.4.1 Régimes hydrologiques des bassins versants de l'étude.....	11
<i>I.5 Modèles utilisés par PREMHYCE</i>	13
I.5.1 Calage dans la plateforme PREMHYCE.....	15
I.5.2 Calage expert.....	17
II. Résultats	17
<i>II.1 Calage automatique de la plateforme PREMHYCE</i>	17
<i>II.2 Calage expertisé des modèles</i>	19
II.2.1 Synthèse des réponses pour le modèle PRESAGES	20
II.2.2 Synthèse des réponses pour le modèle Gardenia.....	22
II.2.3 Synthèse des réponses pour le modèle Mordor SD	23
II.2.4 Synthèse globale.....	25
<i>II.3 Comparaison des méthodes automatiques et expertes</i>	27
II.3.1 Méthode de calage d'EDF pour le modèle Mordor	27
II.3.2 Méthode de calage du BRGM pour le modèle Gardénia.....	30
III. Méthode d'utilisation des modèles de la plateforme PREMHYCE	32
III.1.1 Utilisation des différents modèles.....	32
<i>III.2 Analyse des données d'entrée</i>	35
III.2.1 La pluviométrie	35
III.2.2 Les débits	36
IV. Discussion	37
<i>IV.1 Les limites des résultats</i>	37
<i>IV.2 Ouverture</i>	38
V. Conclusion générale	38
VI. Bibliographie	40
VII. Annexes	42

Figure 1: Hydrogrammes des régimes hydrologiques du jeu de bassins versants proposés par la classification	12
Figure 2: Grahique de Turc-Budyko.....	13
Figure 3: Schéma du modèle hydrologique.....	14
Figure 4: Diagramme schématique de l'expertise appliquée sur le modèle PRESAGES	20
Figure 5: Diagramme schématique de l'expertise appliquée sur le modèle GARDENIA.....	22
Figure 6: Diagramme schématique de l'expertise appliquée au modèle MordorSD	23
Figure 7: Graphique présentant les valeurs de la méthode expertisée de MordorSD (vert) comparée aux valeurs de la méthode automatisée (marron) concernant les critères NSE_Q, NSE_sqrt, NSE_log, NSE_inv.....	27
Figure 8: La fiche graphique (hydrogramme, courbe de débits classés, régime, sim = f(obs)) de comparaison de la méthode experte et automatique sur le bassin « K6402520 » avec la fonction objectif NSE_sqrt_Q	28
Figure 9: La fiche graphique (hydrogramme, courbe de débits classés, régime, sim = f(obs)) de comparaison de la méthode experte et automatique sur le bassin « H4262010 » avec la fonction objectif NSE_log_Q.....	29
Figure 10: La fiche graphique (hydrogramme, courbe de débits classés, régime, sim = f(obs)) de comparaison de la méthode experte et automatique sur le bassin « H4262010 » avec la fonction objectif NSE_log	30
Figure 11: Graphique présentant les valeurs de la méthode expertisée de Gardénia comparée aux valeurs de la méthode automatisée concernant les critères NSE_Q, NSE_sqrt, NSE_log, NSE_inv.	30
Figure 12: La fiche graphique (hydrogramme, courbe de débits classés, régime, sim = f(obs)) de comparaison de la méthode experte et automatique sur le bassin « K6402520 » avec la fonction-objectif NSE_log.	31
Figure 13: Hydrogramme de comparaison de la méthode experte et automatique sur le bassin « O3121010 » avec la fonction objectif NSE_log.	32
Figure 14: Diagramme de l'expertise apportée au modèle Gardénia après calage.....	33
Figure 15: Diagramme de l'expertise apportée au modèle PRESAGES après calage.....	34
Figure 16: Diagramme de l'expertise apportée au modèle MordorSD après calage.....	35
Figure 17: Ellipses de bois	36
Figure 18: Fiche de débits caractéristiques ajoutée à PREMHYCE	36

Résumé

Les modèles hydrologiques sont des outils indispensables dans la prévision des étiages afin de mieux anticiper les périodes de pénurie pour améliorer la gestion des ressources en eau. Ces modèles hydrologiques nécessitent, pour certains, d'être paramétrés afin de reproduire le comportement hydrologique des bassins versants, avant d'émettre une prévision de débits. Les procédures de calage de modèles hydrologiques sont nombreuses et peuvent être sujettes à de nombreux problèmes. Ainsi, une part du calage d'un modèle hydrologique peut faire intervenir une forme d'expertise du modélisateur ou de l'utilisateur du modèle pour améliorer la cohérence de la modélisation.

L'objectif de cette étude est de comprendre l'expertise associée au calage de différents modèles hydrologiques, et de tenter d'en intégrer certains aspects de manière automatisée. Cette étude explore les réponses fournies par différents modélisateurs impliqués dans le projet PREMHYCE (Prévision des Étiages par des Modèles Hydrologiques, Comparaison et Évaluation) au travers d'un questionnaire, et d'une étude comparative entre une méthode de calage expertisée pour les modèles MordorSD de EDF-DTG et Gardénia du BRGM avec une méthode automatique comprenant un système de rang. Une investigation a donc été réalisée afin d'intégrer l'expertise au sein du calage. Les premiers résultats montrent que certains aspects de l'expertise sont communs à différents modélisateurs (calage automatique), que l'analyse de la donnée d'entrée est l'approche de l'expertise la plus importante. D'autre part, les grandes différences de qualité de la modélisation proviennent de l'utilisation d'algorithmes et de fonctions-objectifs différentes.

Contexte de l'étude

En période estivale, les débits des cours d'eau diminuent, naturellement sous l'effet des prélèvements accru par évaporation, et plus particulièrement lors d'épisodes climatiques de sécheresse. Ces phénomènes peuvent avoir des conséquences sur les différents usages de l'eau. En effet, près de 80% des volumes prélevés pour divers usages (production énergétique, irrigation, approvisionnement en eau potable, navigation) proviennent des eaux de surface. Ainsi, l'anticipation des périodes d'étiage est nécessaire afin d'améliorer la gestion des ressources en eau pour atténuer les impacts socio-économiques et écologiques des restrictions des usages de l'eau. La perspective d'avoir des étiages plus marqués dans le contexte de changement climatique renforce la nécessité de disposer d'outils de gestion appropriés.

Les modèles hydrologiques sont des outils adaptés pour évaluer des volumes d'eau déficitaire en période sèche. Ces outils permettent de réaliser des prévisions hydrologiques pour anticiper le phénomène d'étiage en vue d'une meilleure gestion en eau.

Cette étude s'inscrit dans le cadre du projet PREMHYCE qui a été lancé en 2010 par l'Agence française pour la biodiversité (AFB, ex Office national des eaux des milieux aquatique – ONEMA) et la Direction de l'Eau et de la Biodiversité du Ministère de la Transition Hydrologique et Solidaire (MTES, ex. Ministère de l'Environnement, de l'Energie et de la Mer – MEEM). Ce projet est piloté par l'équipe Hydrologie d'IRSTEA Antony et réunit plusieurs partenaires : le BRGM, EDF, Météo-France et l'Université de Lorraine. La première partie du projet a eu pour objectif d'évaluer les forces et les faiblesses de cinq modèles de prévision d'étiage (GR6J, GARDENIA, MORDOR, SIM et PRESAGES) respectivement développés par ces différents acteurs. Le projet a ensuite donné lieu au développement d'une plateforme opérationnelle de prévision des étiages (PREMHYCE) qui intègre l'ensemble de ces modèles hydrologiques.

Cet outil est mis à disposition d'utilisateurs extérieurs (DREAL, Syndicats de rivière,...). La prise en main et l'utilisation de plusieurs modèles hydrologiques requiert des compétences en modélisation hydrologique, nécessitant une formation approfondie de ces utilisateurs. Le modélisateur peut mettre en place des procédures semi-automatique ou automatique, notamment concernant le calage des paramètres du modèle.

Le premier objectif de l'étude sera de comprendre l'expertise (jugement de l'expert) des acteurs vis-à-vis du calage du modèle, les critères qu'ils utilisent, pourquoi ils les utilisent, et sous quelles conditions?

Le second objectif sera d'intégrer ces résultats au sein d'une méthode de calage de manière à automatiser la procédure en accord avec une expertise définie sur la base des réponses des différents modélisateurs.

Ce rapport s'inscrit dans la continuité de travaux visant à analyser les relations entre les jugements faits par les experts et les critères numériques au travers d'évaluations des performances de modèles hydrologiques en comparant des hydrogrammes observés et simulés (Crochemore et al., 2014).

Introduction

La sémantique du mot étiage est complexe car elle fait intervenir plusieurs significations. Il n'existe pas de consensus sur la définition, ce qui va nous permettre d'introduire le sujet en prenant l'étiage comme une période correspondant aux faibles débits des cours d'eau.

Les étiages sont caractérisés par leurs dynamiques lentes résultantes de phénomènes hydrométéorologiques (augmentation des températures et diminution des précipitations) qui s'étendent de plusieurs semaines à plusieurs mois. Les débits faibles sont de plus en plus fréquents sur le territoire Français métropolitain ce qui impacte directement les activités industrielles et agricoles.

Depuis le début du 20^{ème} siècle, un grand nombre d'étiages exceptionnels ont pu être recensés avec notamment celui de 2011. Cet étiage est caractérisé en raison de sa forte sévérité (période de retour > 50 ans) et de sa longue durée le classant comme le plus extrême de la dernière décennie. Ce fut un événement de grande ampleur car il s'est étendu du Jura à la Bretagne en passant par la région Centre impliquant donc un manque en eau sur une grande surface du territoire.

La prévision des étiages est de ce fait une question majeure à laquelle de nombreux scientifiques tentent de répondre par le biais de modèles hydrologiques. Ce sont des questions de biens communs tant l'Homme est dépendant de la ressource en eau qui se fait de plus en plus rare et met en danger grand nombre de nos activités. Les prévisions peuvent être réalisées au travers de modèles hydrologiques, comme le fait la plateforme PREMHYCE. Afin de pouvoir reproduire le bon fonctionnement hydrologique des bassins versants, une paramétrisation des modèles (appelée calage) de cette plateforme doit être réalisée. Cette étape de calage est essentielle pour reproduire le comportement hydrologique d'un bassin versant, et ainsi simuler ou prévoir le débit à son exutoire. Il existe une multitude de moyens de la réaliser, fonction des objectifs de modélisation et ou du modèle lui-même.

De plus, des méthodes automatiques permettent de réaliser des calages et bien qu'elles ne soient pas parfaites, elles peuvent faire l'objet de modification et d'amélioration par l'expertise du modélisateur. Un certain nombre de modélisateurs du projet réalise le calage de cette manière (calage puis expertise), ce qui laisse place à une certaine subjectivité liée au modélisateur lui-même et à son affinité au modèle. L'objet de l'étude réside donc dans cet aspect de l'expertise, et donc intrinsèquement de le caractériser pour l'intégrer dans l'automatisation de la procédure de calage d'un modèle.

Une première étape serait de comparer et d'évaluer les avantages et inconvénients de chacune des approches (automatique et expertisée) et de pouvoir ainsi caractériser l'aspect subjectif de l'expertise.

I. Objet de l'étude

I.1 La prévision des étiages

I.1.1 L'étiage

L'étymologie du mot étiage serait dérivée du terme « étier » qui était affecté aux canaux reliant les marais salants à la mer (Dacharry, 1996). D'une origine commune à estuaire provenant du mot latin « aestus » se traduisant par marée, l'étiage désignait l'état d'un étier après le retrait de l'eau.

Pourtant dans le monde scientifique, le terme d'étiage est sujet à différentes définitions et donc à différentes interprétations. Cette acceptation large du terme induit un amalgame entre étiage et basses eaux. Cette confusion est due à la période de basses eaux qui se définit par le niveau des cours d'eau en dessous de leur module. Ainsi, la définition de M. Dacharry (1996) en explicite les différences : « Les étiages correspondent à un « débit exceptionnellement faible d'un cours d'eau, qu'il ne faut pas confondre avec les basses eaux saisonnières, même s'il en est l'exacerbation » (M. Dacharry, 1996).

Bien que cette définition soit partiellement correcte, elle ne manque pas moins d'être clarifiée sur ses tenants et ses aboutissants. Pour comprendre le caractère exceptionnel des étiages, il faut alors faire des analyses statistiques qui permettent de calculer des périodes de retour afin de juger de leur sévérité. La sévérité est donc étroitement liée à la durée de l'évènement et à son intensité.

Les étiages sont des évènements hydrologiques pouvant s'étaler sur de longues durées causés pour les plus sévères par une sécheresse météorologique et dans le cas le plus général par un appauvrissement des nappes explicitées par la définition de P. George et F. Verger : « Période pendant laquelle les débits sont très bas et peu variables parce que issus des seules nappes souterraines en voie d'épuisement, de tarissement ».

Les étiages peuvent aussi se définir de manière statistique grâce à leurs débits qui sont calculés à différent pas de temps (journaliers, mensuels, moyennes mobiles, débit seuil).

Pour revenir dans un cadre plus général, le niveau d'un cours d'eau est une réponse du bassin versant lié à l'apport météorique, et que l'on peut modéliser grâce à des outils de prévisions.

I.1.2 La modélisation hydrologique

L'objet d'étude en hydrologie est le bassin versant topographique. La ligne de partage des eaux se fait donc en fonction des côtes altimétriques et non en fonction du pendage des couches géologiques sous-jacentes. Ce dernier représente donc toutes les surfaces qui contribuent à l'écoulement de surface des eaux météoriques qui alimenteront la section d'une rivière (exutoire). Cet objet admet donc des échanges avec l'atmosphère (pluviométrie, ETR) mais aussi avec des bassins versants adjacents par des écoulements souterrains.

La modélisation est une simplification mathématique, physique de processus (dans le cas présent) hydrologique, permettant la simulation des états du système étudié en fonction d'un état initial et de forçage connus. Cette réduction du système permet donc d'obtenir des résultats probants notamment en termes de prédiction de débits à l'exutoire en s'affranchissant de la complexité du terrain. Plusieurs modèles sont alors utilisés en hydrologie et se différencient par leurs méthodes de

résolutions et par leurs données d'entrées disponibles. Cinq types de modèle sont alors distingués selon Perrin (2000) :

- Les modèles linéaires (statistiques),
- les modèles stochastiques (probabilistes),
- les modèles physiques (mécanistes),
- les modèles empiriques,
- les modèles conceptuels.

En revanche, pour réaliser une modélisation correcte, il faut répondre à certaines normes présentées ci-dessous.

Lors d'une modélisation hydrologique, la première étape consiste à choisir un modèle, selon la finalité escomptée et son utilisation :

- La finalité du modèle correspond au choix des variables d'étude, et est fonction du niveau de détail, du pas de temps et d'espace et des niveaux d'incertitudes.
- Les données d'entrée disponibles qui vont sélectionner le type du modèle listé dans la partie modélisation.
- La réalité du terrain qui permettra de se rendre compte de la possible difficulté de modéliser les débits si le modèle défini est trop minimaliste ou bien trop complexe qui fournira donc beaucoup d'incertitude.

Dans notre cas d'étude nous nous restreindrons aux modèles conceptuels présents dans la plateforme PREMHYCE, dont l'identification des valeurs des paramètres constituera la première étape de la modélisation. « Les modèles conceptuels sont utilisés pour simuler les comportements d'un bassin » (Glossaire international d'hydrologie, 1992). Ils fonctionnent grâce à une représentation du cycle de l'eau dans plusieurs réservoirs distincts sujets à l'évaporation, aux écoulements souterrains et surfaciques.

I.1.3 Calage d'un modèle hydrologique

S'ensuit la phase de calage qui consiste à identifier les valeurs des paramètres du modèle hydrologique, pour simuler une chronique de débits qui soit la plus proche de celle des débits observés, à partir de données d'entrée et de sortie connues. La suite de l'étude s'attardera sur des modèles pluie-débit de type conceptuels et de spatialisation globale.

Le calage permet donc de d'évaluer les paramètres du modèle pour un bassin donné. Les différents types de calage qui peuvent être réalisés sont :

- de manière automatique, par un algorithme d'optimisation itératif qui cherche à optimiser une fonction objectif, généralement définie comme une distance entre les séries observées et simulées de débit. Il existe de nombreux algorithmes d'optimisation, que l'on peut grossièrement classer en deux catégories : les méthodes globales, qui explorent très largement l'espace des paramètres à partir d'une population de points, et les méthodes locales qui partent d'un point de l'espace des paramètres et explorent son voisinage. Des approches mixtes associant approches globales et locales existent également, de même qu'il

existe des algorithmes multi-objectifs qui peuvent considérer simultanément plusieurs fonctions objectif (avec construction de fronts de Pareto) ;

- de manière manuelle, par le modélisateur qui cherche à faire coïncider les hydrogrammes observés et simulés, sur la base d'un ensemble de critères, qui peuvent être tant numériques que graphiques, et d'une approche heuristique (méthode de calcul rapide qui fournit une solution réalisable rapidement), souvent mise au point sur une base d'expérience(s) individuelle(s) ou collective(s) et fondée sur la compréhension du fonctionnement du modèle. Une telle démarche peut revêtir différentes modalités, avec par exemple des approches optimisant les paramètres de manière séquentielle en considérant des critères différents suivant le rôle de chaque paramètre considéré. Elle peut également faire intervenir des analyses sur les données, la structure du modèle, etc. ;
- de manière mixte, en exploitant les approches automatiques et manuelles de manière combinées, par exemple en utilisant l'approche automatique pour générer un jeu de paramètres initial, et en affinant ensuite les résultats du calage automatique de manière manuelle.

Les modèles globaux sont couramment utilisés, ils permettent de simuler une variable de sortie telle que le débit à l'exutoire d'un bassin versant en fonction de variables d'entrées comme les pluies et l'ETP. Ils permettent donc de simuler des chroniques de débits à des pas de temps variables allant communément de l'horaire au mensuel. Ce modèle demande très peu d'informations sur les bassins versants, ce qui induit un calage plus facile.

Il s'affranchit donc de la spatialisation des variables d'entrées et il transforme les pluies homogènes d'un bassin versant en un débit à son exutoire. Or les résultats de la modélisation peuvent confronter les modélisateurs à plusieurs problèmes :

- Des erreurs aberrantes de valeurs de paramètres,
- Des valeurs de critères de validation trop faibles
- Des aspects graphiques insatisfaisants
- Le sur-calage du modèle (très grande spécialisation des paramètres selon la période étudiée)
- L'équifinalité : des jeux de valeurs de paramètres différents menant à un résultat similaire mais dont les significations des physiques du terrain sont différentes.

Le problème commun à toutes les méthodes locales est le risque d'être piégé sur un optimum local (Perrin, 2000). Ceci est donc dû à des problèmes d'observabilité de l'optimum globale en raison de la complexité du modèle, de la fonction-objectif choisie, et des erreurs sur les données d'entrée. Ceci induit des jeux de paramètres qui peuvent être différents, et qui possiblement n'ont pas de parenté conceptuelle (Gan et Biftu (1996)).

Cette liste exhaustive met donc le modélisateur dans l'obligation de revoir son calage et les données avec lesquelles il fut réalisé. Cet aspect-là se nomme l'expertise et c'est ce qui va nous intéresser par la suite du rapport.

I.1.4 L'expertise : une étape indispensable

L'expertise est « la traduction de l'énoncé d'une connaissance dans une communauté scientifique dans un énoncé apporté en réponse à une question posée dans un objectif décisionnel (Roqueplo,

1991) ». En modélisation hydrologique, l'expertise prend tout son sens concernant la fiabilité des modèles réalisés. Elle s'appuie sur la légitimité des opérateurs à pouvoir réaliser une modélisation juste selon un statut « d'expert » qu'ils auront acquis vis-à-vis de la science. Cette vision suit le modèle standard positiviste qui considère « les acquis de la science comme indiscutables, et reconnaît aux scientifiques le pouvoir de définir à eux-seuls (sur la base des études produites) l'intérêt général (Soyer, 2015) ».

Cette définition de l'expertise reste bornée par le fait que les experts eux-mêmes ne sont pas tous des humanoïdes dotés d'une pensée commune, et de prises de décisions similaires. L'expertise conserve donc une part importante de subjectivité. Dans ce cas, l'expertise perd toute neutralité car la donnée produite sera étroitement corrélée avec la réflexion de l'expert quand bien même aucune pression politique ou financière ne viendrait interférer.

En modélisation hydrologique, les travaux de Crochemore et al. (2014) se sont axés sur l'étude de la similitude des jugements d'experts et une recherche des liens entre les jugements et les critères numériques. Les conclusions de ses travaux ont montrées que le jugement ne peut pas se résumer à un critère numérique et que l'expertise multiple devrait être le moteur des évaluations relatives comme absolues. Dans le cas de l'étude, l'expertise peut concerner plusieurs aspects tels que l'expertise sur la donnée, sur la paramétrisation et sur les sorties du modèle hydrologique.

Afin de mieux comprendre cet aspect subjectif, plusieurs experts participant au projet PREMHYCE se sont prêtés au jeu afin que l'étude puisse mieux cerner l'expertise effectuée en matière de modélisation hydrologique visant évidemment la simulation des étiages.

I.2 PREMHYCE : un outil de prévision des débits d'étiage

Les objectifs principaux du projet PREMHYCE (Prévision des Etiages par des Modèles Hydrologiques, Comparaison et Evaluation) sont d'évaluer, de comparer et in fine d'améliorer les modèles de prévision des étiages. Le projet a été réalisé en collaboration avec cinq partenaires scientifiques précédemment présentés dans la partie « contexte de l'étude », et a évalué les forces et faiblesses de ces modèles sur trente-cinq bassins versants subdivisés en deux échantillons (de vingt-quatre et de onze) représentatifs des conditions d'écoulements naturelles et influencées, situés en France métropolitaine et sur l'île de la Réunion.

Ce projet a donné lieu au développement d'une plateforme de prévision des étiages intégrant les cinq modèles hydrologiques évalués, et qui permet de fournir des prévisions de débit sur environ 120 bassins versants en France métropolitaine depuis 2017.

Cette plateforme permet de réaliser des prévisions d'ensemble de débit jusqu'à un horizon de 90 jours avec cinq modèles hydrologiques (Gardénia, GR6J, Mordor, PRESAGES et SIM) (PREMHYCE : un outil de prévision des débits d'étiage), en utilisant différents scénarios en entrée (rejeu de la climatologie, précipitation nulle, prévisions du centre européen à 10 jours). Cette plateforme est composée de deux utilitaires qui permettent :

- le calage des modèles hydrologiques
- l'émission en temps réel de prévisions de débit

Dans cette étude, nous nous restreindrons à l'aspect calage des modèles.

Les cinq modèles utilisés sont donc associés aux cinq organismes de recherche et d'expertise (Irstea, BRGM, EDF-DTG, Université de Lorraine et Météo-France) et répondent aux noms de GR6J (Pushpalatha et al, 2011), GARDENIA (Thiery, 2014), MORDOR (Garçon, 1999), PRESAGES (Lang et al., 2006) et SIM (HABETS et al., 2008).

Ces modèles seront présentés en partie « Modèles utilisés par PREMHYCE » Données et méthode

I.3 Données

Les données utilisées sont des séries chronologiques de débits (L/s) au pas de temps journalier (sur la période 1974-2010) issues de la banque HYDRO. Ces données sont assorties de codes de validité qui permettent de qualifier la qualité de la donnée. Les données de précipitations (mm/jour), d'évapotranspiration potentielle (mm/jour) et de températures (°C) journalières s'étalent entre 1959 et 2010 et sont issues de la réanalyse SAFRAN (Quintana-Ségui et al, 2008, Vidal et al, 2009) dont les données sont disponibles sur l'ensemble du territoire sur une grille de 8 kilomètres de côté, et ont été agrégées à l'échelle du bassin versant.

I.4 Bassins versants

Cette étude présente des résultats sur vingt-et-un bassins versants en France métropolitaine, peu influencés par les activités humaines. Ces bassins sont issus de la base de données de la phase de recherche du projet PREMHYCE (Nicolle et al, 2014). Cette sélection a été réalisée avec l'aide des services opérationnels de l'Etat (Direction Régionales de l'Environnement, de l'Aménagement et de Logement (DREAL)) grâce à des données dont la qualité est considérée comme satisfaisante. Les critères déterminants ont été la qualité des données, les longueurs des séries disponibles, les enjeux qui incombent aux bassins versants. De plus, ils ont été choisis pour représenter la diversité des caractéristiques physiques et hydro-climatiques de l'ensemble du territoire, que l'on peut notamment constater grâce aux classifications des régimes hydrologiques de Sauquet et de Turk Budyko dans la partie « **Régimes hydrologiques des bassins versants de l'étude.** ».

Tableau 1: Caractéristiques hydrologiques, climatiques, géographiques des bassins versants

Bassin versant	Débit moyen (l/s)	Précipitation moyenne (mm)	ETP moyenne (mm)	Température moyenne (°C)	Altitude moyenne (m)	Surface (km ²)
A1080330	4442,4	949	657	9,4	390	669
B2220010	21831,8	949	657	9,2	350	2543
H2342020	5850,8	839,5	693,5	10,1	308	1119
H4252010	2752,6	657	693,5	10,6	133	934
H7401010	24526,7	803	657	9,8	137	4320
H8212010	2800,1	912,5	620,5	9,7	159	377
I5221010	7379,9	985,5	620,5	10,1	159	883
J7483010	3427,4	730	693,5	11,2	70	809
K1321810	15969,3	985,5	693,5	9,7	431	1792
K6402520	6692,9	803	693,5	10,4	220	1220
L0563010	5795,7	1095	730	10,7	386	605
L4411710	6131	912,5	730	10,3	392	853
M0243010	1812,8	730	693,5	10,9	103	501
M7112410	6676,7	912,5	730	11,2	170	872
O0592510	29880,9	1350,5	693,5	9,4	985	1579
O3121010	11997,8	1387	693,5	7,9	1019	588
Q5501010	58855,9	1350,5	693,5	9,4	915	2592
S2242510	11469,6	1022	730	12,5	78	1678
U4644010	5020,1	912,5	730	10	516	798
V4264010	11471,1	1168	730	7,9	935	1127
Y4624010	2570,7	949	1022	13,5	316	536

Le **Tableau 1** présente des caractéristiques de chaque bassin versant mettant en évidence la diversité climatique et physiographique du territoire Français. Ces bassins ont effectivement des caractéristiques différentes avec des surfaces comprises entre 337 km² à 2592 km² et des altitudes moyennes variant entre 70 m et 1019 m.

I.4.1 Régimes hydrologiques des bassins versants de l'étude.

En ce qui concerne les bassins versants constituant les objets de l'étude, leur régimes ont été acquis selon la méthode de Sauquet. Ils ont été générés sur la période 1974 à 2010 et présentés sur la **figure 1**.

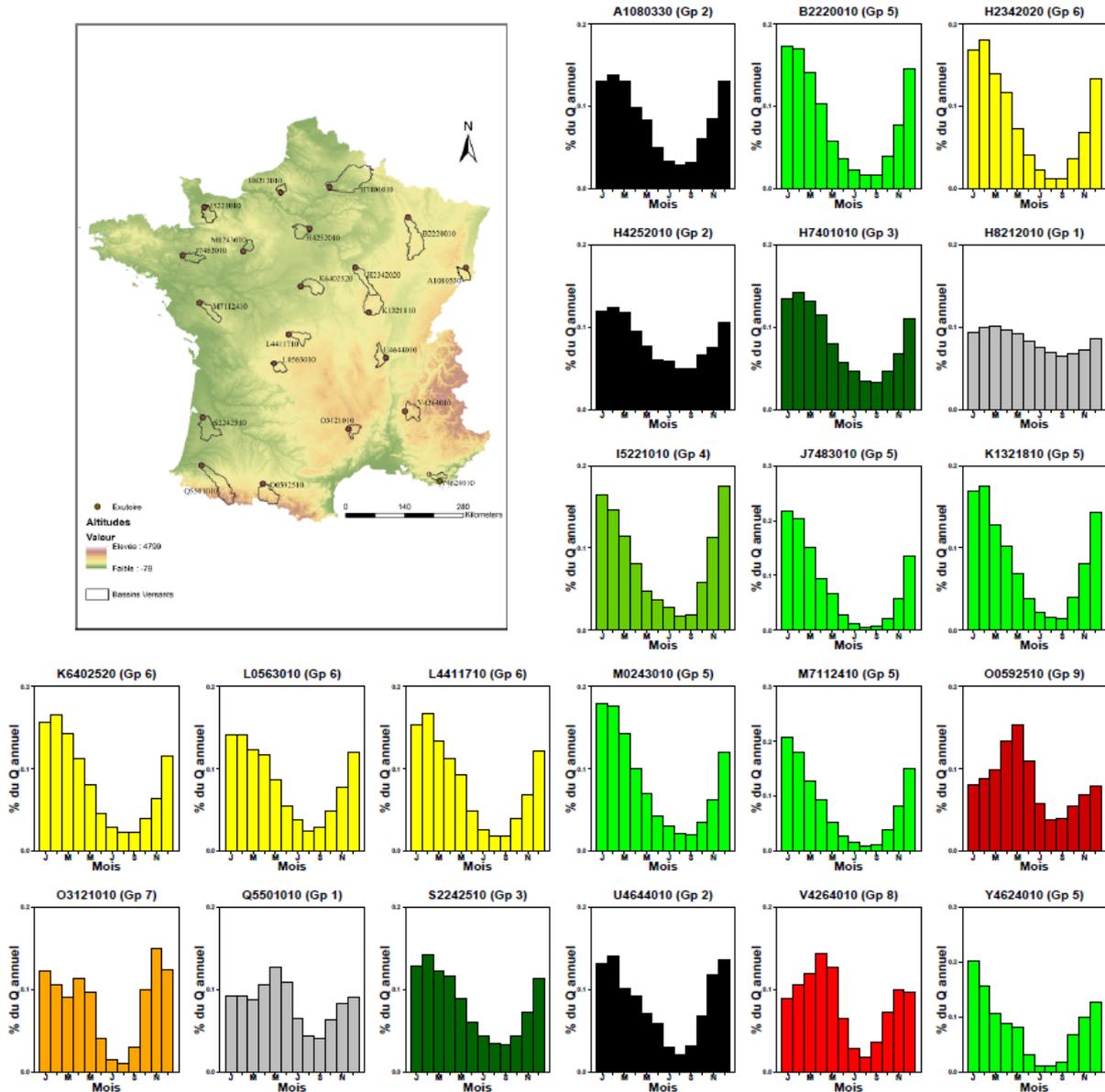


Figure 1: Hydrogrammes des régimes hydrologiques du jeu de bassins versants proposés par la classification de Sauquet(2006) et de la carte de la France

Cette méthode a été développée en langage de programmation R, ce qui permet de classer les régimes de manière simple et automatique. Elle permet de calculer le critère de ressemblance comme procédure de classification. Ce critère est une distance euclidienne dont l'équation 1 est donnée ci-dessous :

Équation 1

$$d(ref, A) = \sqrt{\sum_{t=1}^{12} \left(\frac{qm(t, A)}{qa(A)} - qref(t) \right)^2}$$

- Un schéma de surface Interaction sol-biosphère-atmosphère (Isba) qui est un modèle de surface qui fait intervenir les données physiographiques pour le sol et la végétation, photosynthèse, humidité, évaporation et les infiltrations gravitaires.
- Un modèle hydrogéologique (Modcou) qui permet de modéliser les débits journalier à l'exutoire des bassins versants.

Le modèle SIM ne nécessite pas d'être calibré comme les autres modèles conceptuels. En effet, les paramètres de SIM sont définis par rapport à la réalité physique du terrain. Ce modèle ne sera pas abordé dans cette étude

Tableau 2 : Tableau récapitulatif des modèles utilisés dans le projet PREMHYCE

Modèle	Institution	Type	Résolution spatiale	Nombre de paramètres à caler	Méthode d'assimilation	Méthode de post-traitement
GARDENIA	BRGM	Conceptuel	Global	4 à 13	Oui	Non
GR6J	Irstéa	Conceptuel	Global	6 à 8	Oui	Non
MORDOR	EDF	Conceptuel	Global	11 à 15	Oui	Non
PRESAGES	Univ. Lorraine	Conceptuel	Global	7 à 10	Oui	Oui
SIM	Météo-France	Base physique	Distribué	0	Non	Oui

I.5.1 Calage dans la plateforme PREMHYCE

Dans le projet PREMHYCE, la méthode de calage « pas-à-pas » (Michel 1989 ; Nascimento, 1995) est utilisée et fut développée à la Division Hydrologie du Cemagref Antony. Cette méthode suit un processus d'optimisation de la fonction-objectif itératif selon un déplacement le long des axes des paramètres. C'est un algorithme fonctionnant en deux étapes qui combine l'approche globale et locale :

- Une première étape consistant en un préfiltrage, qui parcourt l'espace des paramètres afin de localiser la zone de l'optimum qui sert de point de départ pour la seconde étape.
- La seconde étape est une recherche locale « pas-à-pas » qui explore les environs du point de départ pour trouver l'optimum.

Cette méthode permet d'éviter de tomber sur un optimum secondaire grâce notamment à la première étape.

De plus, le calage effectué sur les modèles ne suit pas la procédure de calage validation (Split sample test) habituellement utilisée car les modèles furent préalablement validés dans la phase de recherche, ce qui permet d'utiliser toutes les données pour se placer dans des conditions de prévision en temps réel.

1.5.1.1 L'importance de la fonction-objectif

La fonction-objectif est un critère numérique de performance qui permet de mesurer les écarts entre les débits observés et les débits simulés. L'algorithme d'optimisation permet de faire varier les paramètres du/des modèles pour minimiser ces écarts. La plateforme PREMHYCE intègre deux

fonctions-objectifs : les critères de Nash-Sutcliffe (NSE) (Nash et Sutcliffe, 1970) selon l'Équation 2 et de Kling Gupta Efficiency (KGE) (Gupta et al. 2009) selon l'Équation 3. Les formules sont dérivées de celle des moindres carrés. Plus la valeur est proche de 1, plus la simulation est corrélée avec les observations.

Équation 2

$$NSE = 1 - \frac{\sum_{i=1}^n (Qsim - Qobs)^2}{\sum_{i=1}^n (Qobs - \bar{Qobs})^2}$$

Équation 3

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

L'absence de biais est la qualité première d'un modèle hydrologique. En effet, lorsque les données d'observation tendent vers l'infini, la somme des erreurs liée à la modélisation tend vers 0. Sur un nombre de données d'observation fini, il existe et il existera toujours une part d'erreur entre les valeurs observées et simulées, il faut toutefois que sa valeur tende au maximum vers 0.

1.5.1.2 Transformation des débits

La transformation préalable des débits lors du calage est utilisée afin de donner une pondération plus importante à certaines classes de débits dans le but de mieux les reproduire. En effet, « certains critères accordent, par la nature de leur formulation, une importance relative plus grande à certaines classes de débits » (Perrin, 2000). Sans transformation, les erreurs relatives liées au modèle sont toujours plus élevées lorsque les débits sont élevés car les volumes transportés sont plus conséquents. Ceci induit de fait une erreur relative de la modélisation plus élevée causée par les forts débits et donc une mauvaise modélisation des faibles débits. « Ceci provient du fait que les résidus du modèle ne sont généralement pas homoscedastiques, c'est-à-dire que leur variance n'est pas indépendante de la valeur du débit » (Perrin, 2000). Ainsi, il est possible de réduire le poids des erreurs liées aux forts débits, en utilisant certaines transformations de débits, parmi lesquelles :

- La transformation sur la racine carrée, pour représenter la globalité des classes de débits.
- La transformation logarithmique, pour représenter de la même manière les étiages.
- La transformation sur l'inverse des débits, qui est la plus « virulente » et qui a pour but de caractériser les faibles débits.

1.5.1.3 Méthode des rangs

Si notre choix se porte plus naturellement vers des transformations de débits favorisant la reproduction des bas débits, nous avons choisi de réaliser le calage des modèles hydrologiques sur l'ensemble des fonctions objectifs et des transformations de débits disponibles. Ceci présente plusieurs avantages :

- vérifier que l'algorithme d'optimisation n'a pas abouti à la définition d'un jeu de paramètre correspondant à un optimum local (identifié si les performances du modèle calé sur une fonction objectif et une transformation de débit ne sont pas les plus élevées)

- assurer une cohérence de la modélisation, en vérifiant que les performances du modèle sont globalement satisfaisantes sur l'ensemble de la gamme des débits. En effet, un utilisateur qui verrait un modèle performant pour reproduire les bas-débit mais incapable de reproduire les crues pourrait douter de la qualité du modèle.

Ainsi, une méthode simple a été mise en place afin de sélectionner la fonction-objectif pour chacun des bassins, qui permettra de fournir de bonnes performances sur toute la gamme des débits, avec un poids plus forts sur les bas débits, et d'éviter les problèmes d'optimums locaux.

Le calage des modèles fut opéré avec les fonctions-objectif KGE et NSE et les transformations des débits log, l'inverse, la racine carré pour le NSE et l'inverse ainsi que la racine carré pour le KGE. Ainsi, sept calages ont été réalisés avec chaque fonction-objectif pour chaque bassin. Pour chaque calage, il est possible de calculer des critères de performance (qui sont les mêmes que les fonctions objectifs). Pour chaque bassin versant et pour chaque fonction-objectif, le rang de chaque critère de performance est déterminé, et la moyenne de ces rangs est calculée. De plus, un classement fut également opéré suivant les rangs du critère de performance focalisé sur les bas débits (NSE_log Q). Un rang global est ensuite calculé pour conserver le calage qui donne le jeu de paramètres le plus représentatif de l'ensemble des critères de performance, en mettant un poids plus important sur les bas débits (représentés par le critère NSE sur les logs).

I.5.2 Calage expert

Afin de comparer l'expertise du calage du modèle hydrologique, il a été proposé à chaque modélisateur de caler son modèle hydrologique selon sa propre procédure de calage et sa propre expertise. Ainsi, nous disposons d'un jeu de paramètre calé de manière automatisée au sein de la plateforme PREMHYCE, et d'un jeu de paramètre expert.

De plus, afin de comprendre la vision d'expertise de chacun des scientifiques et partenaires du projet, un questionnaire leur a été envoyé pour détailler les procédures de calage et l'expertise. Les réponses au questionnaire sont disponibles en annexes et les résultats seront présentés dans la partie « **Résultats** » du rapport.

II. Résultats

Les résultats se concentrent sur les jeux de données liés au modèle MordorSD.

II.1 Calage automatique de la plateforme PREMHYCE

Le **Tableau 3** présente le résultat de la sélection des critères de calage pour la méthode automatique correspondant au modèle de MordorSD.

Tableau 3: Tableau représentatif des jeux de paramètres pour la méthode automatique pour MordorSD

Nom de la station	Code	Crit_cal	KGEQ	KGE1Q	KGERCQ	NSEQ	NSE1Q	NSERCQ	NSEIQ	CRIT_MOY	RANG_NSERCQ	Cmoyen
L'Ill à Didenheim	A1080330	KGE_sqrt_Q	0.93	0.81	0.95	0.87	0.78	0.90	0.90	0.86	2	2.29
La Meuse à Saint-Mihiel	B2220010	NSE_log_Q	0.93	0.95	0.95	0.90	0.92	0.94	0.96	0.93	3.5	3.36
Le Serein à Chablis	H2342020	NSE_log_Q	0.92	0.83	0.94	0.89	0.72	0.92	0.92	0.86	4	3.43
L'Orge à Morsang-sur-Orge	H4252010	NSE_log_Q	0.87	0.79	0.88	0.73	0.58	0.76	0.73	0.70	4	3.29
L'Oise à Sempigny	H7401010	NSE_log_Q	0.92	0.88	0.95	0.86	0.87	0.91	0.92	0.89	3	3.14
L'Andelle à Vascoeuil	H8212010	NSE_log_Q	0.91	0.84	0.90	0.86	0.78	0.86	0.85	0.84	2	3.14
La Vire à Saint-Lô	I5221010	NSE_sqrt_Q	0.92	0.89	0.94	0.89	0.85	0.93	0.94	0.91	2	2.86
La Seiche à Bruz	J7483010	NSE_log_Q	0.77	0.44	0.86	0.84	0.35	0.90	0.89	0.74	5	3.71
L'Arroux à Étang-Sur-Arroux	K1321810	NSE_sqrt_Q	0.94	0.85	0.95	0.93	0.70	0.95	0.94	0.88	2	3
La Sauldre à Salbris	K6402520	NSE_sqrt_Q	0.94	0.83	0.94	0.90	0.75	0.92	0.91	0.87	1	2.86
La Briançonnais à Condat-Sur-Vienne	L0563010	NSE_log_Q	0.85	0.81	0.94	0.79	0.74	0.89	0.91	0.83	2.5	3.5
La petite Creuse à Fresselines	L4411710	NSE_sqrt_Q	0.90	0.76	0.93	0.89	0.57	0.92	0.90	0.82	1	3
L'Orne Saosnoise à Montbizot	M0243010	NSE_log_Q	0.82	0.63	0.88	0.81	0.58	0.88	0.90	0.79	4	4.43
La Sèvre Nantaise à Tiffauges	M7112410	NSE_log_Q	0.87	0.49	0.95	0.88	0.45	0.93	0.90	0.79	3	3.29
La Salat à Roquefort-sur-Garonne	O0592510	NSE_log_Q	0.83	0.92	0.88	0.85	0.84	0.89	0.90	0.87	2	3.71
Le Tarn à Montbrun	O3121010	NSE_log_Q	0.82	0.86	0.89	0.83	0.74	0.91	0.93	0.85	2	3.577
Le Gave de Pau à Bérenx	Q5501010	NSE_log_Q	0.68	0.81	0.75	0.66	0.69	0.71	0.73	0.70	2	3.42
L'Eyre à Salles	S2242510	KGE_sqrt_Q	0.97	0.83	0.97	0.93	0.79	0.94	0.93	0.90	2	2.71
L'Azergues à Lozanne	U4644010	NSE_log_Q	0.84	0.69	0.90	0.83	0.67	0.90	0.91	0.83	3	3.14
La Drôme à Saillans	V4264010	NSE_log_Q	0.91	0.92	0.93	0.86	0.87	0.89	0.91	0.88	3	3.36
Le Gapeau à Hyères	Y4624010	NSE_sqrt_Q	0.93	0.69	0.95	0.88	0.28	0.91	0.86	0.74	1	3

Les critères choisis sont donc des fonctions-objectif qui permettent de caler le modèle selon un objectif favorisant la modélisation des bas débits avec notamment l'utilisation des fonctions-objectif NSE_Log_Q, NSE_sqrt_Q et KGE_sqrt_Q. Ces critères fournissent respectivement les meilleures performances 14, 5 et 2 fois avec la méthode des rangs. Ces résultats sont cohérents puisque la méthode de calage accentue le poids sur le NSE_log_Q. Pourtant, sept bassins n'ont pas été calés suivant ce critère mais les résultats de l'ensemble des valeurs de validation sont moins bons pour ces bassins.

D'autre part, on remarque que les bassins versants « H4252010 » et « Q5501010 » ont des performances médiocres. Leurs régimes hydrologiques sont différents (respectivement pluvial et pluvio-nival), ce qui permet d'écarter l'hypothèse d'une mauvaise représentation éventuelle de l'influence de la neige sur le second bassin. De plus, que ce soit par la méthode experte ou la méthode automatique, la modélisation des ces deux bassins versants reste laborieuse et ne permet pas d'acquiescer des résultats significativement bons, par comparaison avec les autres bassins.

La suite des résultats s'axe sur la compréhension d'expertise des différents acteurs qui ont participé au sondage (**Calage expert**).

Le **Tableau 4** présente le résultat de la sélection des critères de calage pour la méthode automatique correspondant au modèle de Gardénia.

Tableau 4; Tableau représentatif des jeux de paramètres pour la méthode automatique pour Gardénia

stations	Code	Crit_cal	KGEQ	KGE1Q	KGERCQ	NSEQ	NSE1Q	NSERCQ	NSEIQ	CRIT_MOY	RANG_NSEIQ	Cmoyen
L'Ill à Didenheim	A1080330	NSE_log_Q	0,89	0,74	0,94	0,82	0,72	0,89	0,88	0,84	1,00	3,43
La Meuse à Saint-Mihiel	B2220010	NSE_log_Q	0,89	0,94	0,96	0,83	0,91	0,93	0,95	0,91	1,00	3,71
Le Serein à Chablis	H2342020	NSE_log_Q	0,92	0,83	0,96	0,90	0,79	0,93	0,93	0,89	1,00	2,86
L'Orge à Morsang-sur-Orge	H4252010	NSE_log_Q	0,86	0,75	0,87	0,74	0,68	0,78	0,78	0,78	1,00	2,86
L'Oise à Sempigny	H7401010	NSE_log_Q	0,94	0,90	0,95	0,91	0,87	0,92	0,92	0,92	2,00	2,57
L'Andelle à Vascoeuil	H8212010	KGE_sqrt	0,89	0,88	0,92	0,85	0,79	0,86	0,85	0,86	2,00	3,43
La Vire à Saint-Lô	I5221010	NSE_log_Q	0,94	0,82	0,96	0,89	0,79	0,93	0,93	0,90	1,00	3,43
La Seiche à Bruz	J7483010	NSE_log_Q	0,83	0,32	0,93	0,85	0,30	0,90	0,86	0,71	1,00	3,29
L'Arroux à Étang-Sur-Arroux	K1321810	NSE_log_Q	0,93	0,77	0,97	0,91	0,75	0,94	0,94	0,89	1,00	3,71
La Sauldre à Salbris	K6402520	NSE_log_Q	0,89	0,77	0,94	0,82	0,68	0,88	0,88	0,84	1,00	3,00
La Briance à Condat-Sur-Vienne	L0563010	NSE_log_Q	0,88	0,79	0,93	0,80	0,73	0,89	0,90	0,84	1,00	2,57
La petite Creuse à Fresselines	L4411710	NSE_log_Q	0,91	0,70	0,95	0,86	0,66	0,90	0,90	0,84	1,00	3,29
L'Orne Saosnoise à Montbizot	M0243010	NSE_log_Q	0,87	0,61	0,92	0,76	0,57	0,87	0,88	0,78	1,00	3,29
La Sèbre Nantaise à Tiffauges	M7112410	NSE_log_Q	0,92	0,44	0,95	0,88	0,42	0,91	0,88	0,77	1,00	3,14
La Salat à Roquefort-sur-Garonne	O0592510	NSE_log_Q	0,86	0,83	0,86	0,76	0,75	0,79	0,80	0,81	1,00	2,57
Le Tarn à Montbrun	O3121010	NSE_log_Q	0,87	0,81	0,95	0,82	0,75	0,90	0,91	0,86	1,00	3,00
Le Gave de Pau à Bérenx	Q5501010	NSE_log_Q	0,87	0,81	0,86	0,76	0,73	0,78	0,78	0,80	1,00	2,57
L'Eyre à Salles	S2242510	KGE_sqrt	0,96	0,96	0,96	0,93	0,90	0,94	0,94	0,94	3,00	3,07
L'Azergues à Lozanne	U4644010	KGE_sqrt	0,89	0,67	0,93	0,86	0,62	0,91	0,89	0,83	2,00	3,43
La Drôme à Saillans	V4264010	NSE_log_Q	0,87	0,88	0,89	0,77	0,82	0,82	0,85	0,84	1,00	2,71
Le Gapeau à Hyères	Y4624010	NSE_log_Q	0,76	0,77	0,93	0,75	0,69	0,88	0,87	0,81	1,00	3,14

Pour le modèle Gardénia, les fonctions-objectifs choisies sont celles qui représentent le mieux les bas débits avec notamment le NSE_log, KGE_sqrt (KGE avec transformation de type racine), retenues respectivement 18 et 3 fois avec la méthode des rangs. Ainsi, trois bassins n'ont pas été calés avec la fonction choisie (NSE_log), ce qui est également normal car dans l'ensemble, les résultats des valeurs des critères de validations sont moins bons.

II.2 Calage expertisé des modèles

Les réponses fournies par les différents modélisateurs ont permis d'identifier de grands axes d'expertise réalisés lors du calage des modèles hydrologiques. Ils ont été schématisés selon des arbres décisionnels (**Figure 4, Figure 5, Figure 6**) réalisés par mes soins sur la base du questionnaire (**fourni en « Annexes »**) afin de mieux illustrer les différentes démarches de calage des modélisateurs, et pour en faire ressortir les points communs et les différences.

II.2.1 Synthèse des réponses pour le modèle PRESAGES

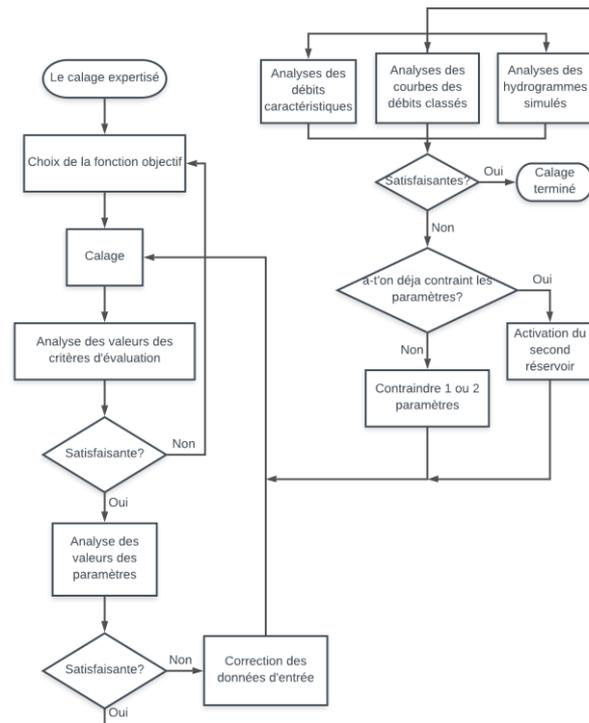


Figure 4: Diagramme schématisant de l'expertise appliquée sur le modèle PRESAGES

PRESAGES a été développé pour simuler les débits de basses eaux. La procédure de calage expertisée (Figure 4) est réalisée en quatre étapes :

- 1) **Choix de la fonction objectif** : Cette étape est de loin l'une des plus importantes car c'est sur cette base que s'appuie toute la procédure de paramétrisation du modèle pour la simulation des débits et en particulier les bas débits. Cette procédure comporte de nombreux intérêts notamment pour traiter ou réduire les erreurs liées aux classes de débits (débits observés utilisés pour le calage ou l'évaluation du modèle), aux estimations des paramètres et aux variables d'entrée (pluie, ETP, température). Dans un premier temps, le modélisateur choisit une fonction-objectif qui permet de mieux représenter la classe de débits à modéliser. Ainsi, un critère de performance qui permet de mettre l'accent sur les bas débits sera utilisé afin de caler le modèle de manière automatique.
- 2) **Analyse de critères numériques** : Les premiers résultats permettront d'évaluer les valeurs des critères d'évaluation en fonction du critère de performance. Autrement dit, parmi tous les critères d'évaluations calculés, celui qui a la valeur la plus proche de 1,0 devra être celui correspondant à la fonction-objectif. Le rapport entre le débit moyen simulé et observé est calculé pour trouver un optimum dont la valeur est de 1, au-dessus le débit moyen observé est surestimé et en-dessous il est sous-estimé. Une valeur de critère satisfaisant pourrait être entre 0,8 et 1,2. Dans le cas contraire, on considère que le modèle ne permet pas de simuler correctement les débits observés, ce qui conduira le modélisateur à changer de fonction-objectif pour caler de nouveau le modèle.

- 3) **Analyse des valeurs des paramètres** : Cette analyse permet d'identifier des valeurs « inhabituelles », en lien avec leurs rôles au sein du modèle : par exemple un paramètre correcteur de la pluie qui doublerait les valeurs de la pluie en entrée du modèle pourrait signifier des problèmes dans les données d'entrée. Ces données pourraient ensuite être expertisées et corrigées avant de réitérer le calage. Par ailleurs, si les valeurs ne présentent pas d'erreur alors l'analyse des données de sortie sera effectuée.
- 4) **Analyse graphique** : Elle comporte l'examen des débits caractéristiques, des courbes de débits classés, et des hydrogrammes observés et simulés. Elles apportent un critère de satisfaction qui permet à l'utilisateur d'apprécier visuellement la qualité de la modélisation. Les critères graphiques peuvent dans certains cas aider à identifier certains paramètres à modifier pour améliorer la simulation des débits, grâce notamment à l'inspection de phases particulières de l'hydrogramme (crue, tarissement). Le cas le plus parlant est celui de coefficient d'écoulement de crue qui peut amener à produire des hydrogrammes très « nerveux » ou « apathiques ». Cette inspection visuelle permet de borner des paramètres entre des valeurs seuils.

Si ces différentes étapes ne permettent pas une reproduction satisfaisante des débits, l'activation d'un réservoir supplémentaire de routage peut être orchestrée en dernier recours bien que cette option ne soit que très rarement utilisée.

II.2.2 Synthèse des réponses pour le modèle Gardenia

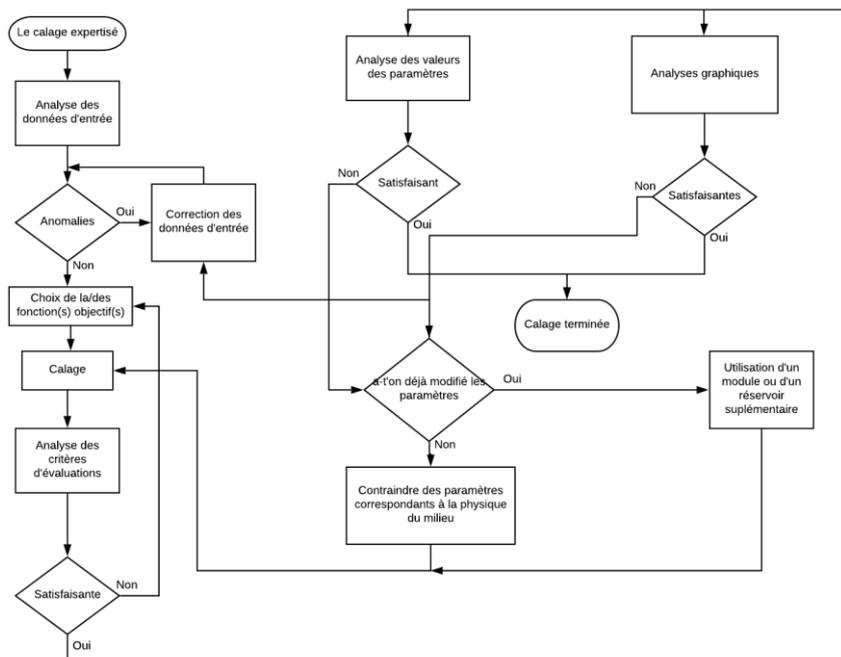


Figure 5: Diagramme schématisé de l'expertise appliquée sur le modèle GARDENIA

Pour le cas de GARDENIA, l'expertise (**Figure 5**) présente des similitudes avec celle réalisée pour le modèle PRESAGES, avec néanmoins quelques différences.

- 1) **Analyse des données d'entrée** : une analyse des séries chronologiques des données d'entrée se fait préalablement au choix de la fonction-objectif. Cette analyse permet de corriger ou filtrer les fichiers d'entrées selon des critères qui leur sont propres comme les lacunes sur des périodes plus ou moins longues.
- 2) **Choix d'une fonction objectif** : une fois cette étape terminée, un critère de performance est sélectionné et le calage automatique lancé.
- 3) **Analyse de critères numériques** : De la même manière que pour PRESAGES, l'analyse des critères d'évaluations succède au calage. Cette étape est également cruciale car elle permet de valider ou non la fonction-objectif. Si les valeurs ne sont pas satisfaisantes, un changement de critère est alors opéré.
- 4) **Analyse des paramètres et analyse graphique** : ces analyses sont réalisées conjointement. Si et seulement si les deux analyses sont satisfaisantes, alors le calage est terminé. L'analyse graphique permet d'identifier deux sources potentielles d'erreurs provenant soit des données d'entrée soit des paramètres, alors que l'analyse des valeurs des paramètres permettra d'affirmer ou non si les valeurs des paramètres sont cohérentes avec les valeurs de la physique du milieu.

Si ces différentes étapes ne permettent pas une reproduction satisfaisante des débits, l'utilisation d'un module ou d'un réservoir supplémentaire peut être implémentée. La stratégie de calage est basée sur le schéma du modèle le plus simple afin de vérifier qu'il arrive à modéliser correctement les débits observés, qui peut être complexifié si nécessaire.

II.2.3 Synthèse des réponses pour le modèle Mordor SD

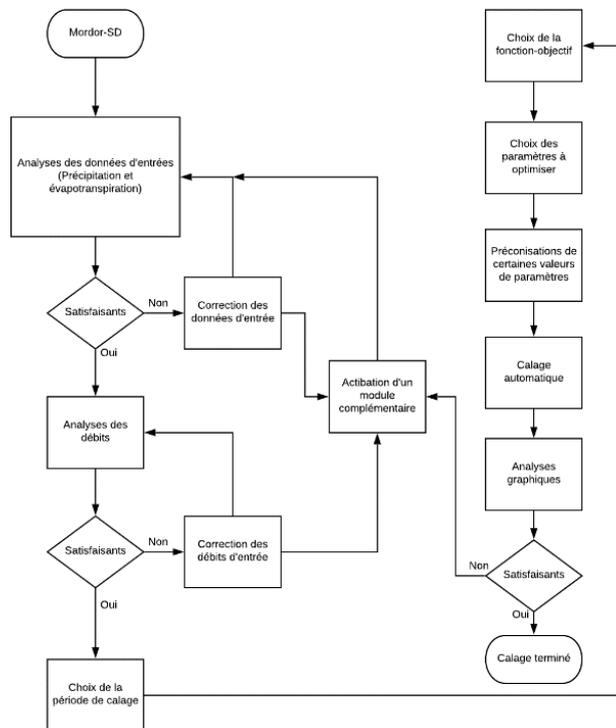


Figure 6: Diagramme schématisé de l'expertise appliquée au modèle MordorSD

En ce qui concerne MordorSD, la stratégie de calage (**Figure 6**) suit elle aussi une approche mixte mêlant le calage automatique et manuel.

Tout d'abord, elle s'organise via le choix et les critiques des données d'entrée s'effectuant en deux étapes :

1. - Choix et critique des forçages météo avec les entrées de stations pluviométriques et d'évapotranspiration. Un choix de stations pluviométriques les mieux situées pour représenter au mieux la pluviométrie sur le bassin versant est réalisé. De plus, ces données sont critiquées et homogénéisées (ellipse de Bois)
 - Les débits subissent une correction via une naturalisation éventuelle s'ils sont impactés par les activités humaines et peuvent être dans un même temps eux aussi sujet à une homogénéisation.
2. Dès lors que la critique des données est effectuée, une période de calage ad hoc est choisie. Elle permet d'utiliser des périodes qui corroborent les bonnes données chronologiques d'entrée.

3. Le choix de la fonction-objectif est alors effectué comprenant une agrégation multicritères. Cela permet notamment de simuler au mieux l'ensemble des classes de débits.
4. Un choix de paramètres à optimiser est alors effectué selon la typologie du bassin versant avec notamment une préconisation de certaines valeurs de paramètres avec leurs gammes de variations (descripteur bassin versant, typologie des régimes interannuels).
5. Le calage automatique est alors effectué selon un algorithme maison qui, en fonction des résultats obtenus par analyse visuelle de graphiques, fera office de test afin de réitérer ou non toutes les étapes précédentes.

L'expertise du processus de calage d'un modèle hydrologique est un exercice routinier qui mêle de nombreux aspects. Cette synthèse des questionnaires a permis de dégager quatre grands axes de l'expertise du calage de ces modèles hydrologiques : l'analyse sur les données d'entrée, le calage en lui-même (i.e. choix d'une fonction objectif), la structure du modèle (i.e paramètres), et les aspects graphiques. L'analyse visuelle est de loin le critère le plus important de par plusieurs points :

- Elle est issue de l'analyse graphique des courbes de débits simulés et observés. En un regard succinct, le modélisateur est capable de dire si le modèle calé donne des résultats satisfaisants.
- Elle permet de trouver des valeurs aberrantes qui pourront être corrigées.
- Elle permet d'affirmer que le calage est terminé et réussi.

Le choix de la fonction objectif est également important car il permet de cibler la classe de débits qu'on souhaite modéliser.

Les valeurs des paramètres sont déterminées à l'aide d'un algorithme d'optimisation. Cet outil mathématique présente des limites et peut générer des valeurs erronées pour certains paramètres. En ce sens, le modélisateur se doit de vérifier que les valeurs des paramètres sont cohérentes et d'en déduire des améliorations.

De plus, aller du schéma le plus simple au plus compliqué constitue en soi une première expertise. Ceci permet de conserver un modèle parcimonieux (avec un nombre limité de paramètres à caler), qui permet d'assurer une certaine robustesse du modèle (capacité du modèle à représenter les débits sur une période sur laquelle il n'a pas été calé). L'ajout de modules ou de réservoirs souterrains peut permettre de représenter des processus plus complexes que le modèle de base.

II.2.4 Synthèse globale

Nous présentons ici (**Tableau 5**) la synthèse de l'ensemble de ces questionnaires. La lecture des différents questionnaires permet d'identifier de grands axes dans l'expertise liée au calage du modèle hydrologique.

Tableau 5: Tableau récapitulatif de l'automatisation de l'expertise

Axes	Objets	Automatisation	Procédure
Données	Précipitation, ETP, Température	Non/Oui	Homogénéisation
	Débit	Oui/Oui	Homogénéisation Naturalisation
Structure	Valeurs des paramètres	Oui	Borner les paramètres
	Algorithme d'optimisation	Non	Optimisation séquentielle (un groupe de paramètres d'abord, puis un autre)
Calage	Fonction(s)-objectif(s)	Oui	Intégrer plusieurs fonctions-objectifs différentes
	Fonction(s)-objectif(s)	Oui	Combiner des fonctions-objectifs
Graphiques	hydrogrammes (obs/sim)	Non	
	Courbe débit classé	Non	

Ces grandes étapes sont décrites ci-dessous :

1. Les données d'entrée :

Ce sont des données brutes qui correspondent à la précipitation, évapotranspiration, température, débits et qui sont susceptibles d'être erronées tant par leurs valeurs que par les lacunes dans les chroniques. Dans un premier temps, leur analyse est donc essentielle à une bonne appropriation du modèle et des données par les modélisateurs. Dans un second temps, la modification peut alors s'effectuer de manière manuelle ou automatique selon des critères de validation qui existent déjà dans les données fournies dans la banque Hydro ou par homogénéisation.

De plus, lorsque les débits des cours d'eau sont fortement influencés par les activités humaines, une naturalisation des débits dans la mesure du possible peut être réalisée. En effet, il est possible de naturaliser un débit s'il est influencé seulement par un barrage. A l'inverse il est plus difficile de naturaliser le débit du bassin versant de la Seine à son exutoire.

Afin de valider les différentes étapes d'analyse et de correction effectuées, il est préconisé de réaliser une analyse de bilan hydrologique ainsi qu'une analyse de sensibilité des données qui permettront d'identifier la robustesse du modèle vis-à-vis de la perturbation de données saines.

2. La structure du modèle :

Ceci correspond aux valeurs de paramètres, à l'algorithme d'optimisation et aux modules complémentaires. Seuls les paramètres et les modules seront sujets à discussion. D'une part, les paramètres ne sont pas les mêmes, leur nombre diffère selon les modèles. De plus, l'importance attachée à leurs valeurs est aussi différente selon les méthodes expertisées. Néanmoins, certains comportements de paramètres permettent d'identifier des valeurs d'entrée défectueuses et peuvent être optimisées et préconisées selon la typologie du bassin versant afin de réduire les équifinalités. D'autre part, des procédures pour prémunir le sur-calage peuvent être mises en place afin d'éviter une trop grande spécialisation des paramètres due notamment à leur grand nombre.

Dans un second temps, l'activation de modules complémentaires peut être alors ajoutée au processus, soit comme dernier recours à une modélisation peu convaincante ou alors en réponse aux valeurs de débits observés, aux échanges souterrains, ou encore à la prise en compte du cycle végétal. L'analyse des données d'entrée sera encore un fois déterminante dans l'activation ou non des différents modules.

3. Le calage du modèle :

La première étape dans le calage du modèle est donc le choix d'une ou plusieurs fonction(s)-objectif(s). Elles sont choisies selon les classes de débits qu'on veut modéliser ou bien peuvent être combinées afin de modéliser tous les classes de débits. Le calage est alors effectué de manière automatique pour tous les modélisateurs ce qui leur permette d'en extraire les données de sorties.

4. Les graphiques :

La comparaison des débits simulés et observés sur différents graphiques est notamment un critère de performance dans l'acceptation de la modélisation. Elle permet très souvent d'affirmer que la modélisation est réussie ou non.

A l'inverse, elle peut aiguiller le modélisateur à réévaluer son expertise au travers de grandes différences qu'elle peut mettre en avant. Auquel cas, il pourra notamment se pencher sur d'éventuelles erreurs qui proviennent soit des données d'entrée, soit des valeurs de paramètres ou bien de l'oubli d'un module complémentaire.

De manière intéressante, si les modèles sont différents et les expériences des modélisateurs également, nous pouvons identifier des étapes communes dans la réalisation du calage des modèles. Ce sont ces étapes communes que nous pourrions tenter d'intégrer à la plateforme. Par contre, toutes ces étapes ne sont pas automatisables.

L'automatisation de l'analyse et de la correction des données d'entrée et des valeurs de paramètres semble être le secteur sur lequel l'automatisation doit se faire. En effet, ces étapes sont primordiales dans la construction d'une modélisation robuste et performante. L'utilisation de données saines pourra de ce fait permettre d'acquérir de meilleures données de sorties et d'en réduire les erreurs qui subsistaient jusqu'alors. D'autant plus qu'elle peut influencer sur de meilleures valeurs de paramètres. Elle pourra se concentrer sur la critique et l'homogénéisation des données, le calcul d'un bilan hydrologique de type « Turc-Budiko » ou autres afin de corroborer les corrections faites. L'analyse de sensibilité peut être elle aussi automatisée et généralisable à tous les modèles.

De plus, l'avantage de l'automatisation de l'homogénéisation des données d'entrée est universelle ce qui est un grand pas vers une harmonisation de la comparaison des modèles hydrologiques.

II.3 Comparaison des méthodes automatiques et expertes

II.3.1 Méthode de calage d'EDF pour le modèle Mordor SD

EDF utilise une approche multicritère basée sur le KGE et ses transformées. Le calage est réalisé sur quatre critères qui sont le KGE Q, le KGE Régime, le KGE classé mensuels et le KGE sur les séquences d'étiage. Ils correspondent respectivement aux débits, débit moyen journalier interannuel, débit classé mensuel et aux périodes de récessions. Cette méthode permet un rendu globalement bon mais avec quelques déficiences au niveau des étiages. La **Figure 7** compare les résultats du modèle MordorSD calé de manière automatique (en marron) et expertisé (en vert), pour chaque bassin versant et selon des critères de validation représentatifs de différentes gammes de débit (NSE Q à gauche pour les crues, NSE sur les racines des débits le 3^{ème} en partant de la gauche pour toutes les gammes de débit, et NSE sur les log à droite et sur les inverses à 2^{ème} en partant de la gauche pour les bas débits).

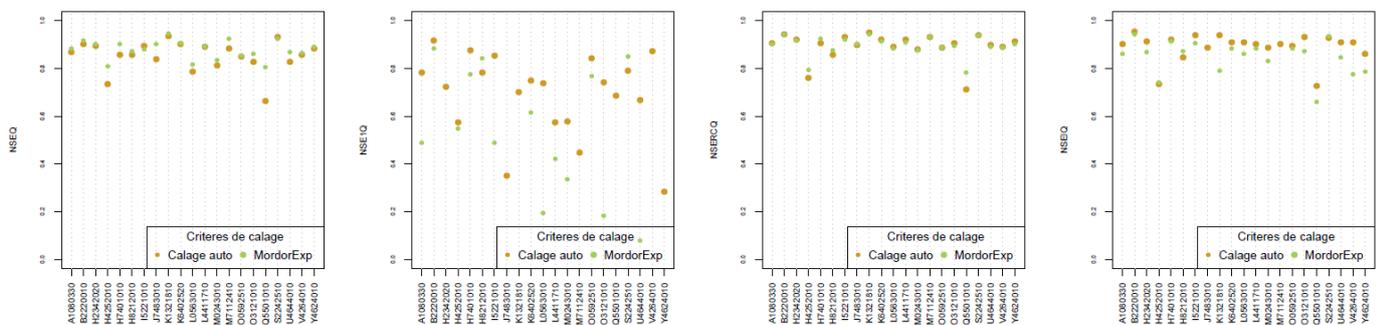


Figure 7: Graphique présentant les valeurs de la méthode expertisée de MordorSD (vert) comparée aux valeurs de la méthode automatisée (marron) concernant les critères NSE_Q, NSE_sqrt, NSE_log, NSE_inv.

Cette méthode permet d'une part de constater que les meilleures valeurs de critères de validation obtenues avec la méthode automatisée sont celles calculées à partir de la même fonction-objectif.

De plus, en regardant plus spécifiquement des résultats de la **Figure 7**, il est possible de s'apercevoir que, pour le critère de validation sur les hauts débits (NSE_Q), la méthode experte a systématiquement de meilleurs résultats. Bien que les différences ne soient pas conséquentes sur l'ensemble des bassins versants, deux valeurs de critère de validation sont bien inférieures concernant la méthode automatique par rapport à la méthode experte. Elles concernent les bassins « H4252010 » et « Q5501010 ». Néanmoins, ce constat ne s'applique pas sur le critère représentant

l'ensemble des débits et est même inversé quand il s'agit des critères favorisant les faibles débits. En effet, les différences sont d'autant plus accentuées sur les graphiques avec les critères de validation NSE_inv_Q et NSE_log_Q .

Ce constat en somme toute logique est lié au fait que l'expertise ne se focalise pas sur les faibles débits mais sur l'ensemble des classes de débits que peuvent enregistrer les cours d'eau à leur exutoire. Ainsi le poids mis par la méthode experte ne privilégie pas les faibles débits mais d'autres aspects de l'hydrogramme.

Ainsi, la méthode experte faite par les experts d'EDF-DTG montre une bonne représentativité des différentes classes de débits et de leurs variabilités saisonnières sur l'ensemble des bassins versants. La méthode automatisée conduit à une sur-estimation systématique des forts débits et une meilleure modélisation des faibles débits dans une faible mesure comme il est possible de le constater sur la « **Figure 8** » représentant les modélisations du bassin « K6402520 ».

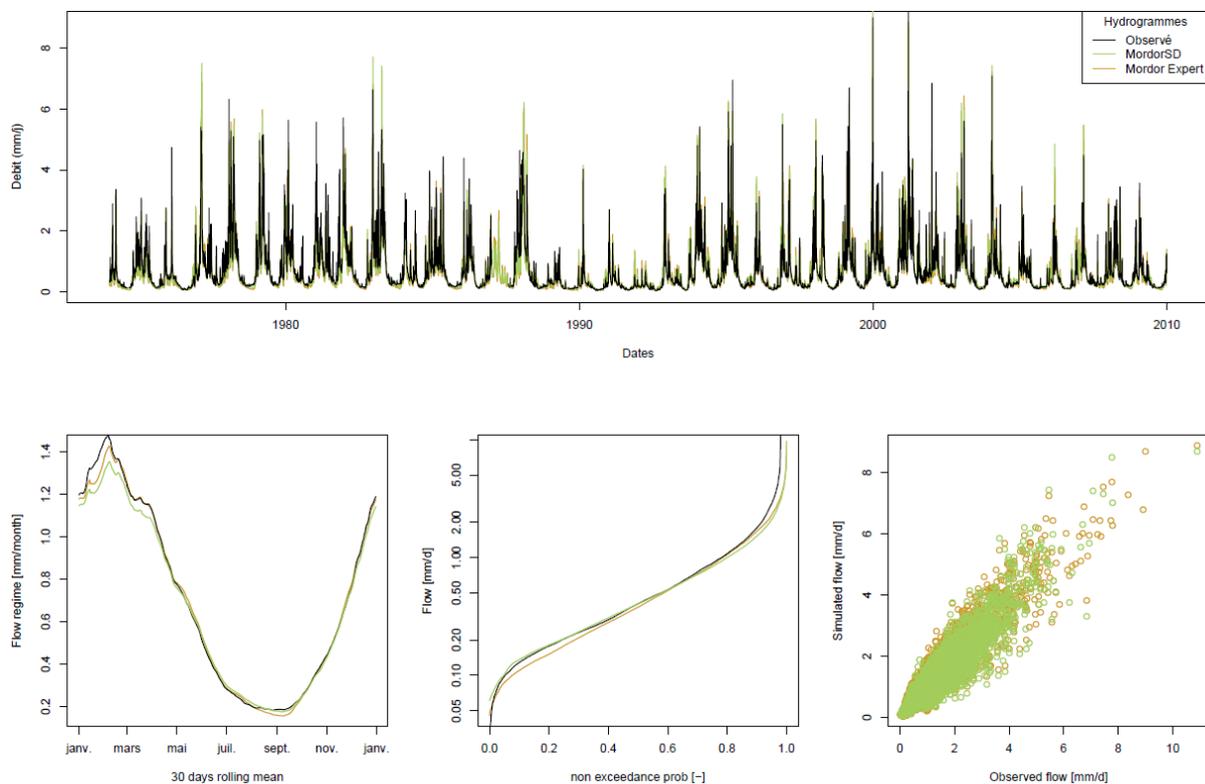


Figure 8: La fiche graphique (hydrogramme, courbe de débits classés, régime, $sim = f(obs)$) de comparaison de la méthode experte et automatique sur le bassin « K6402520 » avec la fonction objectif NSE_sqrt_Q

Bien que ce constat soit généralisable, certains bassins sont moins bien modélisés avec la méthode automatique comme par exemple les bassins versants « H4262010 » et « Q5501010 » présentés en « **Figure 9, Figure 10** ».

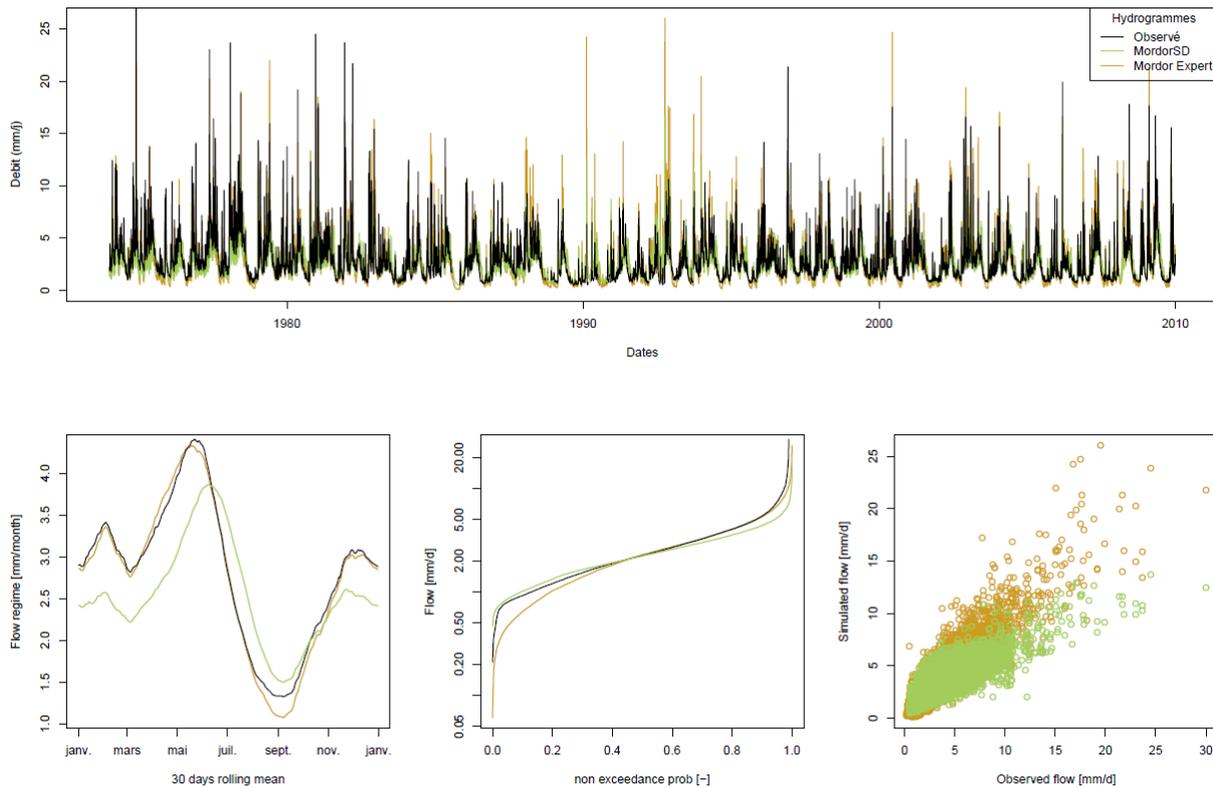


Figure 9: La fiche graphique (hydrogramme, courbe de débits classés, régime, $sim = f(obs)$) de comparaison de la méthode experte et automatique sur le bassin « H4262010 » avec la fonction objectif NSE_log_Q

Il se trouve que les débits du bassin « Q5501010 » sont les plus difficiles à modéliser à l'aide de la méthode experte et surtout de la méthode automatique. Le calage de MordorSD réalisé par la méthode des rangs montre effectivement une sous-estimation des forts débits et une surestimation des faibles débits. La variabilité saisonnière des débits présentée par la courbe des moyennes des débits glissants sur 30 jours et par le nuage de points corrélant les débits observés et simulés montre la difficulté du modèle à reproduire le régime hydrologique du bassin versant.

En ce qui concerne le bassin versant « H4252010 », la méthode automatique qui prend en compte le calage sur le NSE avec la transformation sur les logarithmes montre les mêmes travers que pour le bassin « Q5501010 » et une surestimation de la pente allant vers les faibles débits.

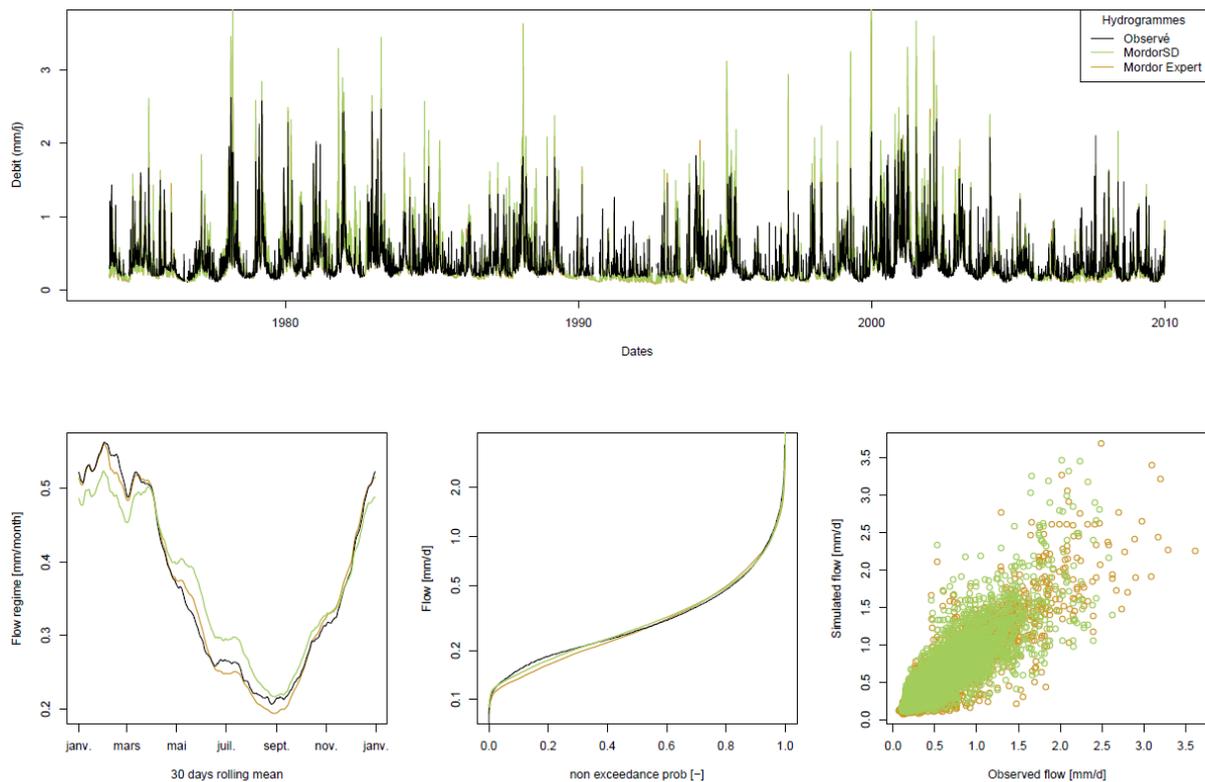


Figure 10: La fiche graphique (hydrogramme, courbe de débits classés, régime, $sim = f(obs)$) de comparaison de la méthode experte et automatique sur le bassin « H4262010 » avec la fonction objectif NSE_{log}

Ce constat est visible grâce au graphique représentant le régime qui montre un décalage des débits simulés par rapport aux débits observés notamment sur les moyens et hauts débits.

→ Pour ces deux bassins, il semblerait qu’une analyse plus approfondie du calage du modèle telle que l’apporte l’expertise soit essentielle pour une bonne modélisation des débits.

II.3.2 Méthode de calage du BRGM pour le modèle Gardénia

Gardénia est un modèle utilisé pour se focaliser sur l’aspect des hauteurs d’eau dans la nappe. Ainsi, ils ont un intérêt certain pour une meilleure représentation des étiages plutôt que des crues car la relation nappe/rivière sera plus impactée par une absence relative comme absolue d’eau dans la rivière. Ainsi comme pour MordorSD, les comparaisons de la méthode expertisée (vert) et automatique (marron) sont représentées par la « Figure 11 » pour chacun des critères.

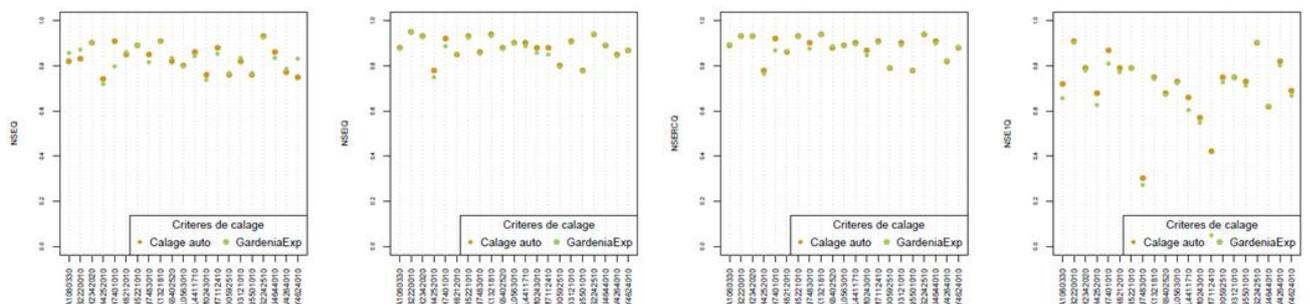


Figure 11: Graphique présentant les valeurs de la méthode expertisée de Gardénia comparée aux valeurs de la méthode automatisée concernant les critères NSE_Q , NSE_{sqrt} , NSE_{log} , NSE_{inv} .

Dans l'ensemble, les graphiques montrent de meilleurs résultats de la part de la méthode automatique tant sur les représentations des hauts débits que sur les bas débits. Bien que ce constat soit inhérent au modèle de Gardénia, une étude graphique des différents bassins a été réalisée afin de mieux représenter les différences. Dans un premier temps, six bassins versants dont les modélisations sont jugées correctes (A1080330, B2220010, H7401010, K6402520, M7112410 et S2242510) ont été identifiés, parmi lesquels seulement trois sont vraiment satisfaisants (B2220010, K6402520 et le M7112410) par une succincte analyse graphique comme l'illustre la **Figure 12**.

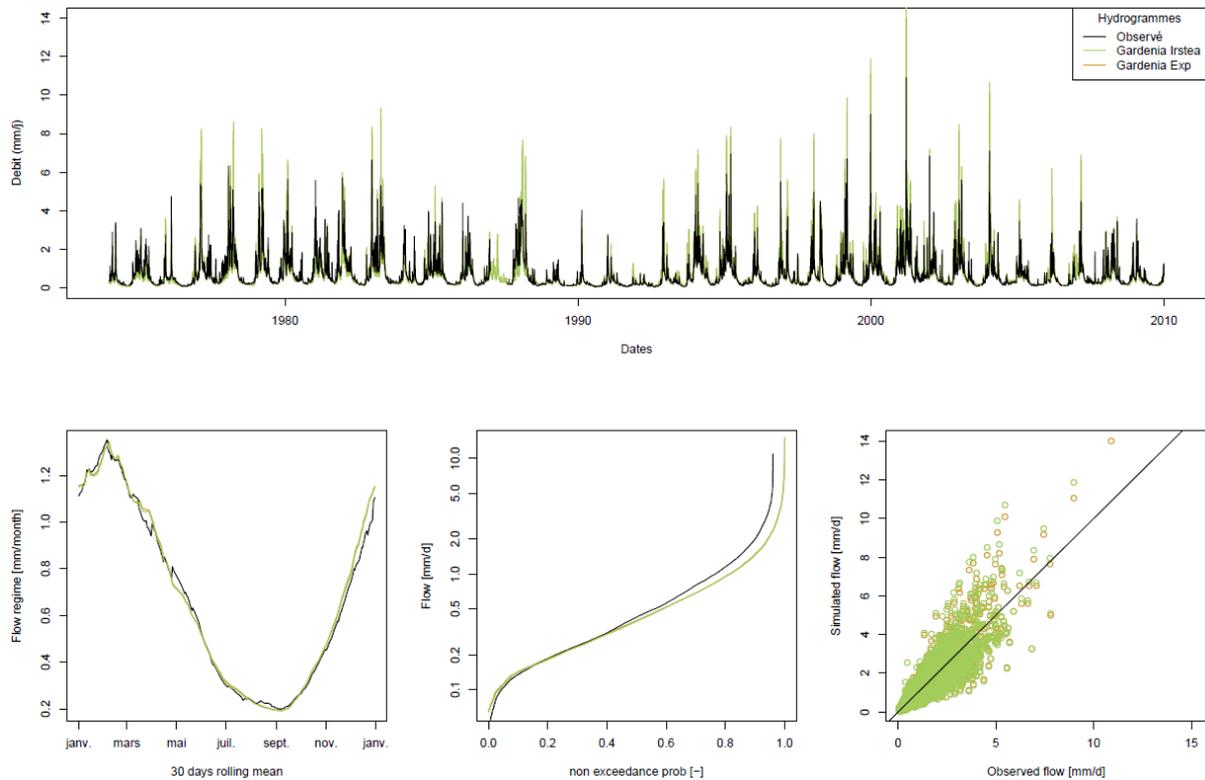


Figure 12: La fiche graphique (hydrogramme, courbe de débits classés, régime, $sim = f(obs)$) de comparaison de la méthode experte et automatique sur le bassin « K6402520 » avec la fonction-objectif NSE_log.

Ainsi, la **Figure 12** présente une modélisation satisfaisante de la part des deux méthodes avec une légère surestimation des hauts débits qu'on peut apercevoir sur l'hydrogramme.

Pour les autres bassins, les simulations sont médiocres, avec des mauvaises représentativités de l'ensemble de la classe des débits comme l'illustre la **Figure 13**.

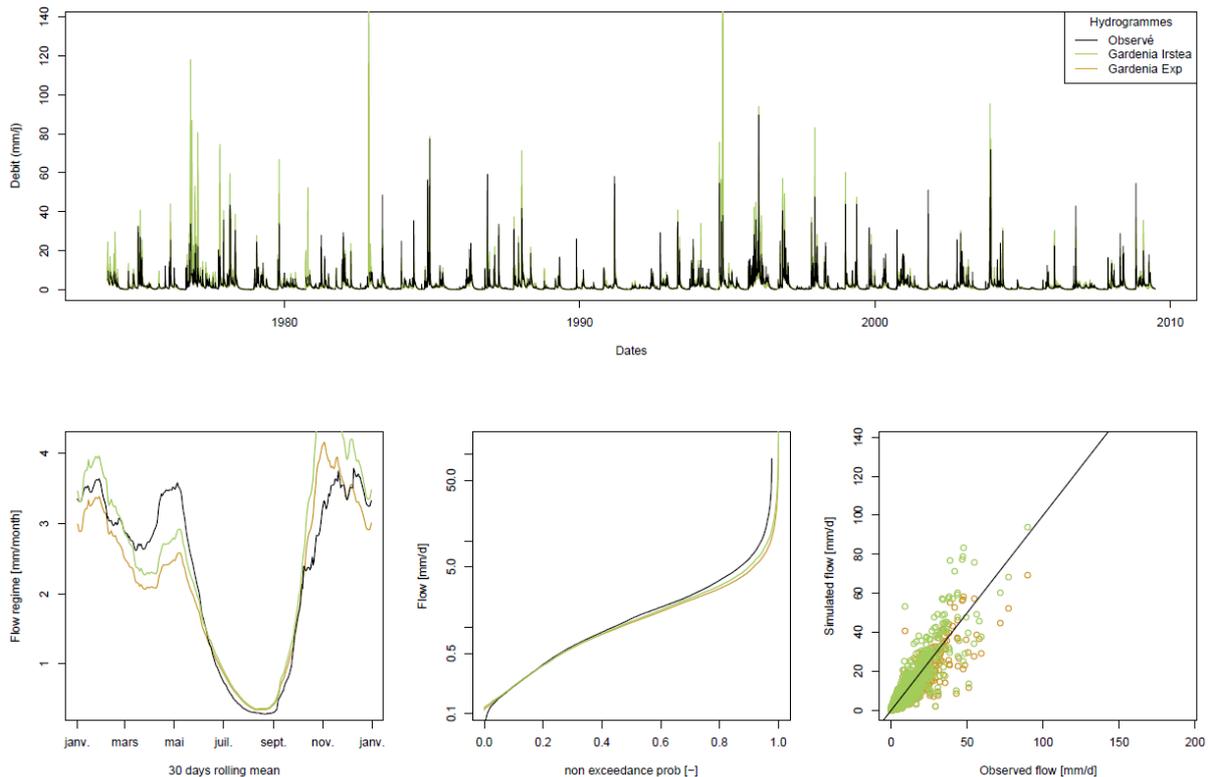


Figure 13: Hydrogramme de comparaison de la méthode experte et automatique sur le bassin « O3121010 » avec la fonction objectif NSE_log.

En effet, la « **Figure 13** » montre des difficultés dans la simulation des hauts débits avec une surestimation parfois très importantes et des surestimations dans les bas débits.

De plus, pour corrélérer ces résultats avec le « **Tableau 4** », on peut s’apercevoir que dix-huit bassins versants sont calés grâce à la fonction-objectif NSE avec la fonction de transformation logarithme.

III. Méthode d’utilisation des modèles de la plateforme PREMHYCE

III.1.1 Utilisation des différents modèles

Les différents partenaires de la plateforme PREMHYCE ont insisté sur le fait que leur calage se fait de manière automatique. Ainsi chacun vis-à-vis de leur modèle respectif à un calage automatique maison. En d’autres termes, la différence de calage entre tous les modèles conceptuels utilisés se fait sur la différence des réponses que les fonctions-objectifs restituent en fonction de chaque modèle. Cette différence est accentuée par l’objectif final que les modélisateurs veulent cibler. Cela marque la première divergence dans le choix des fonctions-objectifs.

Ayant défini cet objectif, une expertise peut alors commencer et se définir en plusieurs points : pour le cas de Gardénia :

- Le premier point est de réaliser un calage automatique et d’en regarder les valeurs des critères de validation. Si cette valeur est au-dessus de la valeur de 0,70 pour le KGE par

exemple et que les analyses des hydrogrammes sont correctes, alors on considère que le calage est réussi (mais pas idéal).

- S'ajoute à cela un regard sur les valeurs des paramètres extrêmes. On s'assure que les paramètres ont des valeurs réalistes pour favoriser la robustesse du modèle et intrinsèquement sa fiabilité.

Sur un jeu de bassins versants, seuls ceux qui auront de mauvais résultats sont sujet à une analyse plus approfondie et qui feront l'objet d'une modification soit par suppression de certaines données d'entrée, soit par l'ajout d'un module complémentaire comme l'illustre la « **Figure 14** ».

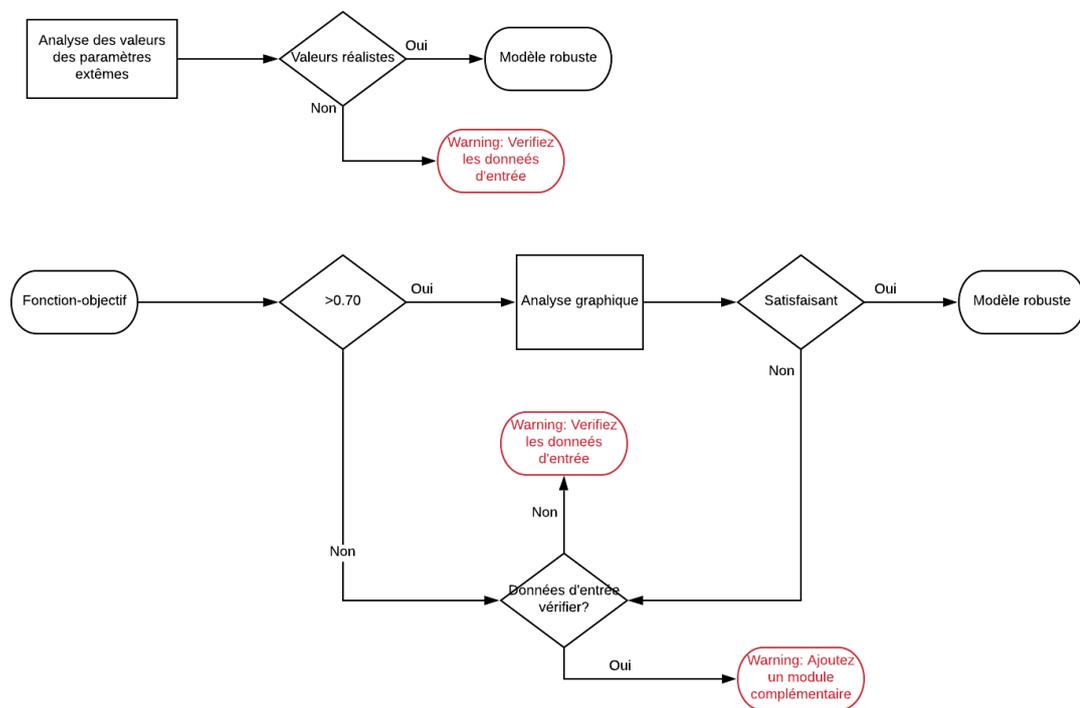


Figure 14: Diagramme de l'expertise apportée au modèle Gardénia après calage.

En ce qui concerne le modèle hydrologique PRESAGES, un paramètre a été introduit afin de corriger les valeurs de précipitation. Cela fait suite à un constat qui met en avant des erreurs liées à de mauvaises estimations de la pluviométrie moyenne causée notamment par les zones de montagnes et/ou karstiques.

Ce paramètre évolue donc entre la valeur 0,75 et 1,25. Les valeurs comprises entre ces deux bornes sont situées dans un intervalle de confiance. Au-delà de cet intervalle, le modélisateur se doit de vérifier les valeurs d'entrée liées aux précipitations ainsi que le contexte géographique et géologique du bassin versant concerné.

En effet, la différence entre la pluie et la lame d'eau écoulée est comprise entre 450 et 650 mm au cours d'une année hydrologique en France métropolitaine. Au-delà de ces valeurs, une erreur est

présente signifiant avec pour les valeurs plus faibles un climat froid qui cause peu de perte (peu d'ETP et d'infiltration).

De plus, grâce à l'analyse des débits caractéristiques et des débits classés, il est possible de s'apercevoir si le calage respecte les débits extrêmes ou non. Ces observations permettent d'orienter la fonction-objectif en fonction des objectifs de modélisation. Cela comprend l'ajout d'un Nash ou d'un KGE avec une transformée sur les bas débits au critère déjà présent afin d'avoir une fonction-objectif composite.

La « **Figure 15** » présente l'expertise effectuée après le calage du modèle PRESAGES.

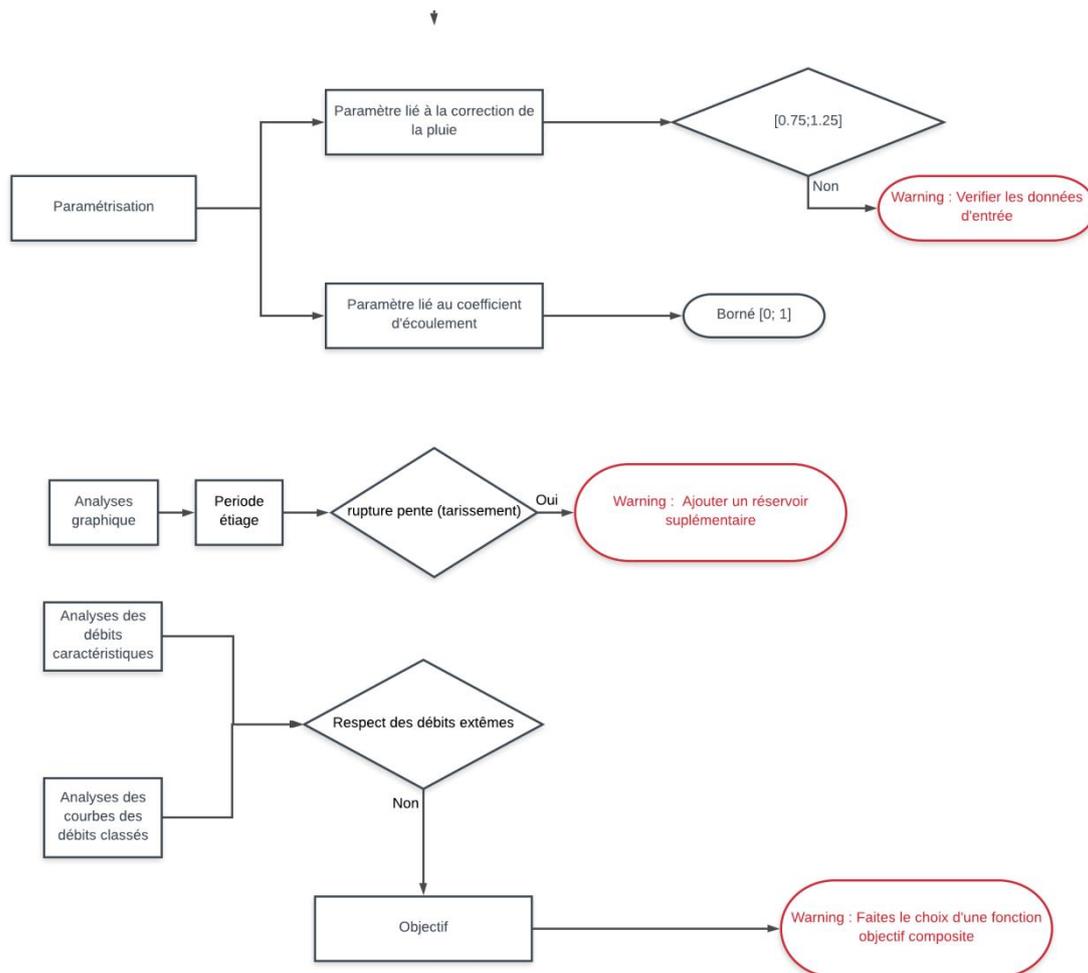


Figure 15: Diagramme de l'expertise apportée au modèle PRESAGES après calage.

Pour ce qui est du modèle MordorSD, la stratégie de calage est automatique selon plusieurs fonctions-objectif voulant allier une fiabilité sur la simulation de l'ensemble de la classe de débits. Ainsi leur expertise ne se trouve pas sur la stratégie de calage mais sur les données d'entrée grâce notamment à l'analyse graphique des débits observés. En d'autre terme, la stationnarité des données d'entrée est contrôlée afin d'identifier des problèmes sur certaines périodes.

Si les stratégies de calage des différents modèles présents dans la plateforme PREMHYCE sont automatiques, alors le rôle des différents modélisateurs sera de prévenir d'éventuelles erreurs liées à

leurs interprétations sous la forme de messages d'avertissements. Un petit guide d'interprétation peut être mis à l'œuvre illustré par la « **Figure 16** » pour le modèle de MORDOR SD.

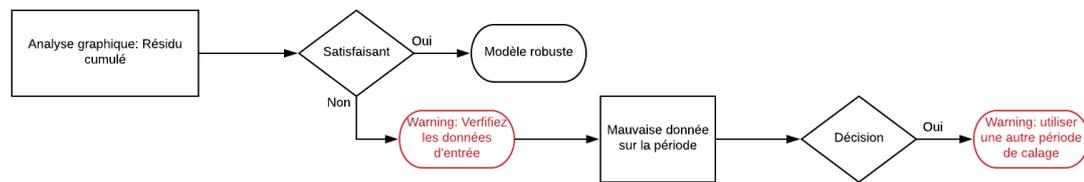


Figure 16: Diagramme de l'expertise apportée au modèle MordorSD après calage.

III.2 Analyse des données d'entrée

L'analyse des données d'entrée constitue une étape importante dans le déroulement d'une modélisation. En effet, grâce à l'identification des erreurs, les modélisateurs prétendent avoir des meilleures simulations en sortie de modélisation. Pour ce faire, plusieurs étapes sont à préconiser en particulier pour les données de pluviométrie.

Précédemment, nous avons pu remarquer que les simulations fournis par EDF grâce au modèle MordorSD constituaient la modélisation fournissant les meilleurs résultats. A partir de cette constatation, nous avons donc choisi de calquer leurs méthodes d'analyse des données d'entrée qui correspondent en majeure partie à la réussite de leur modélisation.

III.2.1 La pluviométrie

EDF utilise les ellipses de Bois (Bois, 1986) pour l'étude statistique des résidus liés aux chroniques de pluviométrie. Afin de mettre en œuvre une telle méthode, il faut avoir à disposition les chroniques de données pluviométriques observées de 2 stations proches.

Dans notre cas, la réanalyse SAFRAN est utilisée pour définir la pluviométrie du bassin. Il sera donc possible d'effectuer ce test (mono-station) en comparant la chronique de pluie avec le modèle représentant la moyenne. Pour ce faire, il faudra calculer les résidus à chaque pas de temps et d'en faire le cumul.

Cette méthode permet graphiquement d'identifier plusieurs types d'hétérogénéités liées aux :

- Erreur multiplicative
- Erreur additive
- A la dérive dans le temps
- Au changement de corrélation

Elle a pour but de déceler des périodes d'hétérogénéité issues d'erreurs de mesure, de changement d'appareillage ou de phénomènes naturels. La « **Figure 17** » présente graphiquement les ellipses de bois avec un cumul de résidus de séries correctes à gauche et un cumul de résidus de séries erronées à droite.

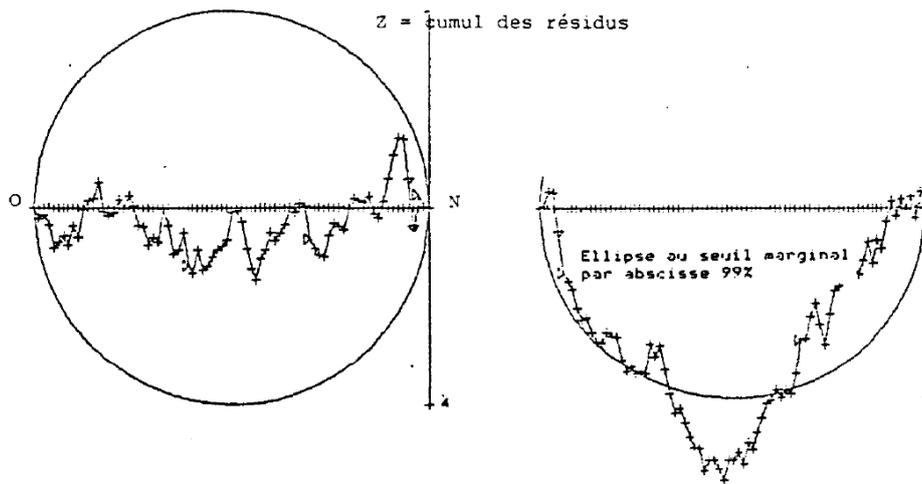


Figure 17: Ellipses de bois

III.2.2 Les débits

Nous avons ajouté à la plateforme PREMHYCE un certain nombre de critères graphiques liés au débit caractéristique des étiages dans le but de réaliser une fiche rétrospective présentée sur la « **Figure 18** » ci-dessous.

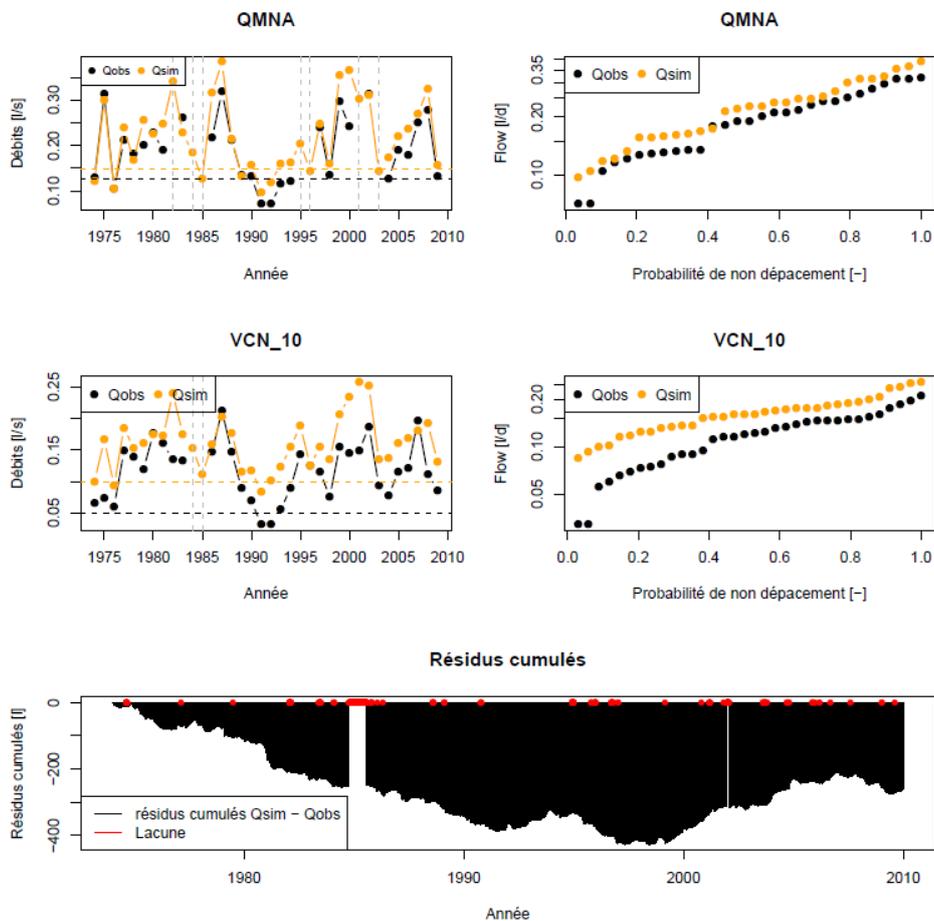


Figure 18: Fiche de débits caractéristiques ajoutée à PREMHYCE

Cette fiche permet d'analyser les QMNA (débit moyen mensuel le plus bas de l'année), ainsi que la représentation des « QMNA 5 » (débit d'étiage de référence correspond au débit moyen mensuel sec de récurrence 5 ans) en pointillés horizontales de la série observée et simulée.

S'ajoute à cela les VCN de durée de 10 jours qui correspondent à des valeurs extraites annuellement en fonction d'une durée fixée (en l'occurrence 10 jours), grâce aux moyennes mobiles calculées à partir des débits moyens journaliers consécutifs.

De plus, le graphique des résidus cumulés $Q_{sim} - Q_{obs}$ permet d'identifier les hétérogénéités dans et les tendances dans les chroniques de débits.

IV. Discussion

IV.1 Les limites des résultats

Plusieurs aspects sont à aborder vis-à-vis des différentes méthodes utilisées, qui sont sujet à discussion vis-à-vis des procédés utilisés.

D'une part, le choix du questionnaire fût de prime abord la solution qui m'a paru légitime quant au fait d'interroger un petit nombre de modélisateurs sur la question de l'expertise. Il permet de récupérer des réponses qui peuvent être plus ou moins précises en fonction des questions posées. En effet, le fait d'introduire le questionnaire et de ce fait plonger l'interlocuteur au cœur du sujet, d'en délimiter le cadre et le fait d'organiser le questionnaire de façon à avoir des questions tantôt ouvertes, tantôt fermés a permis d'acquérir des réponses variées et/ou précises. Malgré cela, il subsiste une certaine orientation des réponses liées à la spécificité des questions. Le fait d'interroger un nombre limité de participants pose un problème de représentativité qui ne permet pas pour l'instant d'affirmer que ces avis pluriels soient généralisables au plus grand nombre.

De plus, seuls les modélisateurs qui utilisent les modèles conceptuels ont pu répondre à ce dernier bien que le nombre de paramètres des modèles diffère.

Ainsi, se pose la deuxième problématique suggérant que le nombre de paramètres et que leurs rôles différents sont susceptibles d'amener des sources d'erreurs différenciées. Cela peut mettre en opposition la valeur conceptuelle des paramètres qui se veulent plus représentatifs de la physique du terrain et de leurs efficacités sur la qualité du modèle. Ceci rend l'automatisation de l'expertise plus abstraite et plus difficile à réaliser.

La méthode des rangs utilisée pour le calage automatique soulève elle aussi un certain nombre de questions liées aux critères de calage. Elle permet de choisir parmi les calages réalisés une fonction-objectif et une transformation basée sur la représentativité des bas débits. Ce critère de calage fait partie de la fonction-objectif parmi tant d'autres, d'autant plus que la méthode ne prend pas en compte le calage multicritères.

La sélection d'une fonction-objectif par la méthode automatique des rangs nous a semblé présenter certaines limites. Ainsi, de manière visuelle (comparaison des débits observés et simulés), nous

aurions pu sélectionner une autre fonction-objectif que celle retenue par la méthode automatique pour certains bassins versants.

De ce fait, cette méthode reste à améliorer et à perfectionner afin d'obtenir des résultats plus proches des modèles expertisés.

IV.2 Ouverture

L'intégration de l'expertise au sein de la méthode automatique constituera les prochaines étapes de travail à réaliser. Plusieurs pistes ont été identifiées et devraient être testées :

- L'automatisation de la filtration des données d'entrée pour permettre d'ores et déjà d'obtenir des résultats qui devraient être meilleurs. Pour ce faire, l'homogénéisation des données d'entrée grâce à l'ellipse de bois permettrait de diminuer les variations non naturelles causées par les réseaux d'observations. Ce dispositif aura pour but de commencer un calage avec des données plus saines. Peut s'ajouter à cela la comparaison des séries hydrologiques des stations voisines qui retranscrivent généralement des variations climatiques similaires.
- L'utilisation des fonctions-objectifs propres aux différents modèles sera donc préconisée en ce qui concerne les modèles MordorSD et PRESAGES. En ce qui concerne le modèle Gardénia, la méthode automatique fournie de meilleurs résultats sur l'ensemble des bassins versants et par conséquent, son utilisation prévaudra sur la fonction-objectif choisie par le BRGM.

V. Conclusion générale

Au cours de cette étude, la réalisation d'une analyse de réponses à un questionnaire par différents experts couplée à une méthode de calage automatique a été effectuée afin d'étudier l'impact de l'expertise au sein d'un protocole de modélisation hydrologique.

Grâce à ce travail, le rôle de trois experts a pu être caractérisé. Il a été démontré d'une part que les différences de performance entre les calages effectués par chacun des experts, sont dues à l'utilisation de multiples modèles présentant des jeux de paramètres et des algorithmes d'optimisation différents. Partant donc du fait d'une certaine subjectivité de l'expertise, les résultats ont montré que les modélisateurs agissaient sur quatre grands axes : l'analyse sur les données d'entrée, le calage en lui-même (i.e. choix d'une fonction objectif), la structure du modèle (i.e. paramètres), et les aspects graphiques. Bien qu'ils eussent été identifiés, l'automatisation de l'expertise ne s'applique pas d'une manière générique à tous.

Ainsi, l'automatisation réalisable et applicable à tous les modèles se définit par la filtration des données d'entrée et à l'utilisation d'un calage multicritères également utilisé par les modélisateurs de MordorSD au sein d'EDF-DTG.

De plus, la comparaison de la méthode automatique et expertisée du modèle MordorSD, donne d'emblée des résultats divergeants :

- Pour la méthode experte, les résultats montrent une meilleure représentativité de l'ensemble des classes des débits.
- Pour la méthode automatique, les résultats montrent une meilleure représentativité des bas débits mais en revanche une reproductibilité moins franche des hauts et des moyens débits. Pour deux des bassins versants, un décalage est mis en évidence dans la saisonnalité.

En ce qui concerne le modèle Gardénia, la méthode automatique permet une meilleure représentativité de l'ensemble des différentes classes de débits en comparaison avec la méthode maison du BRGM.

Le calage expertisé du modèle hydrologique MordorSD est plus performant que la méthode automatique sur les critères calculés. Cette valeur ajoutée de l'expertise s'explique notamment par un filtrage des données d'entrée selon un protocole d'homogénéisation par l'ellipse de Bois et l'utilisation de plusieurs fonctions-objectifs capables de reproduire de manière plus précise les débits observés impliquant une meilleure reproductibilité de l'ensemble des classes de débits. Toutefois, cette observation ne s'applique pas si on s'intéresse uniquement aux faibles débits. En effet, la méthode expertisée surestime les étiages de manière fréquente et est moins performante que la méthode automatique axée volontairement sur les bas débits.

VI. Bibliographie

Crochemore, L., Perrin, C., Andréassian, V., Ehret, U., Seibert, S. P., Grimaldi, S., Gupta, H. and Paturel, J.-E.: Comparing expert judgement and numerical criteria for hydrograph evaluation, *Hydrol. Sci. J.*, 0(ja), null, doi:10.1080/02626667.2014.903331, 2014.

Dacharry M., dictionnaire français d'hydrologie, 1996.

Lang, C.: Étiages et tarissements : vers quelles modélisations ? L'approche conceptuelle et l'analyse statistique en réponse à la diversité spatiale des écoulements en étiage des cours d'eau de l'Est français., phdthesis, Université de Metz, 13 December. [online] Available from: <https://tel.archives-ouvertes.fr/tel-00534656/document> (Accessed 21 July 2015), 2007.

Lang, C., Freyermuth, A., Gille, E. and François, D., Le dispositif PRESAGES (PREvisions et Simulations pour l'Annonce et la Gestion des Etiages Sévères) : des outils pour évaluer et prévoir les étiages. *Géocarrefour*, 81(1): 1-15, 2006a.

Gan, T.Y. et Biftu, G.F. Automatic calibration of conceptual rainfall-runoff Models: optimization algorithms, catchment conditions, and model structure. *Water Resources Research*, 32(12), 3513-3524, 1996.

Garçon, R., Carre, C., and Lyaudet, P.: An example of forecasting and operating simulation of low water flows, *Houille BlancheRevue Internationale De L Eau*, 54, 37–42, 1999.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, 2009.

Habets, F., Boone, A., Champeaux, J. L., Etchevers, P., Franchisteguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Quintana Seguí, P., Rousset-Regimbeau, F., and Viennot, P.: The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *J. Geophys. Res.-Atmos.*, 113, D06113, doi:10.1029/2007jd008548, 2008.

Mathilde Soyer. Solidité de l'expertise, prudence de l'innovation: chercheurs et praticiens dans les observatoires d'hydrologie urbaine. *Sociologie*. Université Paris-Est, 2014. Français. NNT: 2014PEST1048. tel-01226225

Michel, C. Hydrologie appliquée aux petits bassins versants ruraux, Cemagref, Antony, 1989.

Nascimento, N.O. Appréciation à l'aide d'un modèle empirique des effets d'action anthropiques sur la relation pluie-débit à l'échelle du bassin versant. Thèse de Doctorat, CERGRENE/ENPC, Paris, 550 p, 1995.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models, Part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.

Nicolle, P., Pushpalatha, R., Perrin, C., François, D., Thiéry, D., Mathevet, T., Le Lay, M., Besson, F., Soubeyroux, J.-M., Viel, C., Regimbeau, F., Andréassian, V., Maugis, P., Augeard, B., and Morice, E.: Benchmarking hydrological models for low-flow simulation and forecasting on French catchments, *Hydrology and Earth System Sciences*, 18, 2829–2857, doi:10.5194/hess-18-2829-2014, 2014.

Perrin C. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Thèse de doctorat, INPG, Grenoble, Paris, 530 p, 2000.

Bois Ph.: Contrôle des séries hydrologiques corrélées par étude du cumul des résidus. Deuxièmes journées hydrologiques de l'ORSTOM p 89-100, 1986

Pushpalatha, R.: Low-flow simulation and forecasting on French river basins: a hydrological modelling approach, AgroParisTech, 2013.

Roqueplo, P. «L'expertise scientifique: convergence ou conflit des rationalités.». *Environnement, science et politique: les experts sont formels*. JParis, GERMES. 13, pp. 43-80, 1991.

Sauquet E., Gottschalk L., Krasovskaia I. Estimating mean monthly runoff at ungauged locations: an application to France, *Hydrology Research*, vol 39 n°5-6, p 403-423, 2008.

Thiéry, D.: Logiciel GARDENIA, version 8.2, Guide d'utilisation (GARDENIA software, version 8.2, User Guide), BRGM/RP62797-FR 102, BRGM, 2013.

Valéry A. Modélisation précipitations-débit sous influence nivale, laboration d'un module neige et évaluation sur 380 bassins versants, thèse de doctorat, ENGREF, Cemagref, Paris, 405 p, 2010.

Vidal J.-P., Martin, É., Franchistéguy, L., Baillon, M. et Soubeyroux, J.-M. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, 2009.

VII. Annexes

Questionnaire sur les stratégies de calages expertisés des modèles hydrologiques

Stage de Benjamin Conte, Projet PREMHYCE

Avril 2019

Contexte et objectifs du questionnaire :

Le calage d'un modèle hydrologique permet d'évaluer les paramètres du modèle pour un bassin donné. Cette étape de calage peut être réalisée :

- de manière automatique, par un algorithme d'optimisation qui cherche à optimiser une fonction objectif, généralement définie comme une distance entre les séries observées et simulées de débit. Il existe de nombreux algorithmes d'optimisation, que l'on peut grossièrement classer en deux catégories : les méthodes globales, qui explorent très largement l'espace des paramètres à partir d'une population de points, et les méthodes locales qui partent d'un point de l'espace des paramètres et explorent son voisinage. Des approches mixtes associant approches globales et locales existent également, de même qu'il existe des algorithmes multi-objectifs qui peuvent considérer simultanément plusieurs fonctions objectif (avec construction de fronts de Pareto) ;
- de manière manuelle, par le modélisateur qui cherche à faire coïncider les hydrogrammes observés et simulés, sur la base d'un ensemble de critères, qui peuvent être tant numériques que graphiques, et d'une approche heuristique, souvent mise au point sur une base d'expérience(s) individuelle(s) ou collective(s) et fondée sur la compréhension du fonctionnement du modèle. Une telle démarche peut revêtir différentes modalités, avec par exemple des approches optimisant les paramètres de manière séquentielle en considérant des critères différents suivant le rôle de chaque paramètre considéré. Elle peut également faire intervenir des analyses sur les données, la structure du modèle, etc. ;
- de manière mixte, en exploitant les approches automatiques et manuelles de manière combinées, par exemple en utilisant l'approche automatique pour générer un jeu de paramètres initial, et en affinant ensuite les résultats du calage automatique de manière manuelle.

Dans ce qui suit, seront désignées par calage expertisé les approches manuelles ou mixtes, c'est-à-dire les approches faisant intervenir de manière explicite l'expertise du modélisateur dans le processus d'estimation des paramètres.

L'objectif du stage est de mieux comprendre la démarche de calage expertisé et d'essayer d'en formaliser certaines étapes pour en faire bénéficier la plateforme PREMHYCE. Cette plateforme regroupe un ensemble de modèles correspondant à des approches de modélisation et d'estimation des différents paramètres, chaque modèle ayant montré son intérêt dans les phases antérieures du projet. Il est probable que les approches de calage expertisé pourraient donc bénéficier à l'ensemble des modèles, et donc in fine aux utilisateurs de la plateforme.

L'objectif du questionnaire est de mieux cerner les différentes facettes de l'expertise impliquées dans l'estimation des paramètres des modèles de la plateforme PREMHYCE, pour identifier celles qui pourraient faire l'objet d'un traitement plus détaillé dans le cadre de ce stage.

La réponse au questionnaire ne doit prendre qu'une vingtaine de minutes. Les questionnaires sont à renvoyer par mail à Benjamin Conte (benjamin.conte@irstea.fr) jusqu'au 10 mai 2019.

Vous et le modèle que vous utilisez

Prénom Nom	
Institution	
Années d'expérience en modélisation hydrologique	
Modèle sur la base duquel vous répondrez au questionnaire	
Publication de référence décrivant ce modèle	
Publications éventuelles sur les procédures de calage de ce modèle	

Question 1

	Oui	Non
La définition d'un calage expertisé d'un modèle hydrologique donnée précédemment vous convient-elle ?	<input type="checkbox"/>	<input type="checkbox"/>
Si non, pouvez-vous donner la définition correspondant davantage à votre perception du calage expertisé ? <u>Définition</u> :		

Question 2

	Oui	Non
Mettez-vous en œuvre une démarche de calage expertisé telle que définie précédemment (définition proposée ou votre propre définition) pour l'estimation des paramètres de votre modèle ?	<input type="checkbox"/>	<input type="checkbox"/> Allez directement à la Question 18

Question 3

Pour quelle(s) raison(s) une approche de calage expertisée a été mise en œuvre pour votre modèle ?
<input type="checkbox"/> Une approche purement automatique donnait rarement de bons résultats
<input type="checkbox"/> Pour que l'utilisateur soit plus familier du fonctionnement du modèle
<input type="checkbox"/> Pour éviter certains problèmes de calage revenant de manière récurrente dans l'application du modèle. Précisez lesquels :
<input type="checkbox"/> Pour capitaliser l'expérience sur le modèle
<input type="checkbox"/> Autres (précisez) :

Question 4

L'approche de calage expertisée mise en œuvre pour votre modèle est basée sur ?
<input type="checkbox"/> Votre propre expérience avec le modèle
<input type="checkbox"/> L'expérience d'un petit groupe de modélisateurs travaillant sur le développement du modèle
<input type="checkbox"/> L'expérience d'une communauté d'utilisateurs
<input type="checkbox"/> Autres (précisez) :

Question 5

Diriez-vous que l'approche de calage expertisée mise en œuvre pour votre modèle permet de mieux traiter (ou éventuellement réduire) préférentiellement les sources d'incertitudes liées aux aspects suivants ? Cochez les sources concernées et donnez-leur éventuellement un ordre de priorité dans la deuxième colonne (1 : le plus prioritaire dans la démarche)	
<input type="checkbox"/>	Variables utilisées en entrée du modèle (typiquement pluie, ETP, température, etc.)
<input type="checkbox"/>	Sorties du modèle (typiquement le débit observé utilisé pour le calage ou l'évaluation du modèle)

<input type="checkbox"/>		Estimation des paramètres du modèle
<input type="checkbox"/>		Structure du modèle
<input type="checkbox"/>		Autres (précisez) :

Question 6

	Oui	Non
L'approche de calage expertisé que vous mettez en œuvre pour votre modèle a-t-elle fait l'objet d'une formalisation sous forme d'un guide, de consignes ou d'autre manière ?	<input type="checkbox"/>	<input type="checkbox"/>
Si oui : quelle est la référence du document concerné et est-il possible d'y avoir accès ? (dans le cas où le document peut-être fourni avec la réponse au questionnaire, la réponse aux questions 7 à 13 peut simplement faire référence à des parties de ce document)		
Si non : pour quelles raisons une telle formalisation n'a pas été réalisée ?		
<input type="checkbox"/> Elle n'est pas simplement formalisable car très dépendante de l'expérience du modélisateur		
<input type="checkbox"/> Manque de moyen et/ou de temps		
<input type="checkbox"/> Autres (précisez) :		

Question 7

Comment qualifieriez-vous l'approche de calage expertisée que vous mettez en œuvre pour votre modèle ?
<input type="checkbox"/> Une approche de calage purement manuelle, incluant d'autres aspects d'expertise (cf. question 5)
<input type="checkbox"/> Une approche de calage automatique suivie d'une expertise
<input type="checkbox"/> Une approche mixte mêlant calage automatique et manuel et d'autres étapes d'expertise
<input type="checkbox"/> Autres (précisez) :

Question 8

Décrivez les principes généraux et grandes étapes de la démarche de calage expertisée que vous mettez en œuvre (10 lignes max)

Question 9

Dans le cas où votre approche de calage expertisée comprend un algorithme de calage automatique, précisez le nom de l'algorithme utilisé :
Une référence bibliographique décrivant cet algorithme

Question 10

Si votre approche de calage expertisée comprend des étapes d'analyse des données d'entrée et/ou de sortie, détaillez :		
La nature des analyses, traitements et corrections effectués pour chaque variable		
Leur éventuellement séquençement et/ou leur organisation par rapport à d'autres étapes d'expertise		
	Oui	Non
Utilisez-vous des données externes (hors entrées/sorties du modèle) lors de cette étape ?	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, précisez lesquelles		

Question 11

Si votre approche de calage expertisé comprend des étapes sur la structure du modèle elle-même, détaillez :		
La nature des analyses et traitements effectués		
Leur éventuellement séquençement et/ou leur organisation par rapport à d'autres étapes d'expertise		

Question 12

Concernant le calage expertisé des paramètres du modèle proprement dit, pouvez-vous préciser :		
	Oui	Non
S'il y a une séquence définie dans l'estimation des paramètres	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, précisez lequel		
Quels sont les critères numériques ou graphiques généraux utilisés :		
Si des critères (numériques ou graphiques) différents sont utilisés pour estimer différents paramètres	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, précisez les critères pour les paramètres concernés		
Si vous avez une procédure pour vous prémunir d'un sur-calage (trop grande spécialisation des paramètres à la période étudiées)	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, précisez laquelle		

Question 13

	Oui	Non
Au final, avez-vous des critères ou procédures permettant de juger que votre calage expertisé est terminé (par ex. qualité générale du modèle,	<input type="checkbox"/>	<input type="checkbox"/>

impossibilité de trouver de nouvelles sources d'amélioration, validation sur une période indépendante) ?		
Si oui, détaillez les critères ou procédures concernés		

Question 14

	Oui	Non
Parmi les étapes d'expertises décrites précédemment (questions 10-12), certaines sont-elles déjà automatisées ou semi-automatisées, ou pensez-vous que certaines seraient potentiellement automatisables ?	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, détaillez les étapes concernées		

Question 15

	Oui	Non
Parmi les étapes d'expertises décrites précédemment, pensez-vous que certaines revêtent un caractère générique applicable à d'autres modèles ?	<input type="checkbox"/>	<input type="checkbox"/>
Si oui, détaillez les étapes concernées		

Question 15

Quelle sont les principaux avantages et les inconvénients que vous pourriez identifier sur votre approche de calage expertisé ?
<u>Avantage(s) :</u> -
<u>Inconvénient(s) :</u> -

Question 16

Quel temps pensez-vous nécessaire pour réaliser un calage expertisé de votre modèle sur un bassin ? Donnez un temps moyen et/ou une gamme de temps nécessaire en fonction des difficultés rencontrées

Question 17

Avez-vous des informations complémentaires à apporter sur l'approche de calage expertisé mise en œuvre pour votre modèle ?

--

La réponse au questionnaire est désormais terminée. Nous vous remercions de votre participation.

Question 18

Pour quelle(s) raison(s) ne mettez-vous pas en œuvre l'approche de calage expertisée ?

- Elle n'est pas aisée à mettre en œuvre pour mon modèle
- Elle ne me paraît pas utile pour mon modèle
- Elle ne peut pas être systématisée
- Je n'y avais pas pensé
- Autres (précisez) :

Question 19

Pensez-vous que l'approche de calage expertisée ?

- Elle n'est pas aisée à mettre en œuvre pour mon modèle
- Elle ne me paraît pas utile pour mon modèle
- Elle ne peut pas être systématisée
- Je n'y avais pas pensé
- Autres (précisez) :

La réponse au questionnaire est désormais terminée. Nous vous remercions de votre participation.