

Valorisation des données de jaugeages épisodiques pour l'estimation du débit de référence d'étiage QMNA5

Clotaire CATALOGNE¹, Eric SAUQUET², Michel LANG³

¹ Irstea, Lyon – e-mail : clotaire.catalogne@irstea.fr

² Irstea, Lyon – e-mail : eric.sauquet@irstea.fr

³ Irstea, Lyon – e-mail : michel.lang@irstea.fr

RÉSUMÉ. – Cet article propose une méthode permettant la valorisation des données de jaugeage disponibles localement, en un point du réseau hydrographique (un « site cible »), en vue de l'estimation du débit mensuel minimum de période de retour 5 ans QMNA5. La méthodologie mise en place au travers de validations croisées s'appuie sur l'ajustement d'une relation linéaire entre couples des logarithmes des débits jaugeés au site cible et observés simultanément au droit d'une station hydrométrique voisine (un « site d'appui ») pour laquelle une valeur du QMNA5 est connue et fiable. La relation est alors appliquée pour estimer le QMNA5 au site d'intérêt à partir du QMNA5 connu au site d'appui. Les erreurs d'estimation associées sont quantifiées au travers de formulations empiriques intégrant les caractéristiques des campagnes de mesure (nombre et fréquence des jaugeages) et la qualité de la relation entre les deux sites (coefficient de corrélation). Les résultats obtenus en termes de performance montrent qu'au delà de 20 jaugeages, les gains n'évoluent plus significativement et qu'un jaugeage est plus informatif si les campagnes sont espacées dans le temps (une fréquence de trois jaugeages par saison d'étiage est recommandée). Le protocole enfin peut s'accommoder de suivis irréguliers et être étendu à d'autres caractéristiques d'étiage. Cette méthode peut conduire à des estimations plus précises qu'une estimation basée sur des techniques d'interpolation – même sophistiquées – dès lors qu'un nombre de jaugeages est atteint.

Mots-clés : campagne de jaugeage, prédétermination de débits de basses eaux

Using spot gauging data to estimate the annual minimum monthly flow with a return period of 5 years

ABSTRACT. – The annual minimum monthly flow with a return period of 5 years QMNA5 at a poorly gauged site is commonly used for water quality and quantity management in France. A method using spot gauging data to estimate this low flow statistic at poorly gauged sites is presented. The estimate for QMNA5 at the partial record site is derived from the value of QMNA5 at one nearby-gauged site with long-term and high quality records. The relationship between the logarithm of these values is supposed to be well approximated by a linear regression fitted to the logarithm of concurrent flows observed at the two sites. A delete-one cross-validation analysis was performed (i) to assess sensitivity of the data collection strategy, allowing useful recommendations for operational service in charge of river flow monitoring and (ii) to derive bias and standard error models as function of the correlation coefficient between synchronous flows, the total number and the frequency of spot gauging data. The results show that increasing the number of spot gauging data leads to a significant increase in the model performance until approximately 20 gauging data; the gain becomes limited afterwards. Moreover, gauging several times the same year does not significantly improve estimates, probably because of the intra-annual dependence of low-flow data. Three measurements per year seem to be a good trade-off. The method may adjust to non-stationary strategies for data collection, be applied to other low flow statistics and may yield more accurate predictions of QMNA5 than advanced interpolation methods.

Key-words: flow gauging strategy, low flow statistics

I. INTRODUCTION

Le débit moyen mensuel minimum annuel de fréquence quinquennale sèche (*i.e.* de période de retour 5 ans, couramment noté QMNA5) est - avec le dixième du module - l'un des seuls indicateurs statistiques relatif aux étiages à être officiellement défini dans la réglementation sur l'eau à l'échelle nationale. Il constitue, selon le décret nomenclature n°93-743 de la loi de 1992, le débit d'étiage de référence pour l'application de la police de l'eau de manière à fixer les volumes des prélèvements et rejets en eaux superficielles

en fonction de la sensibilité des milieux naturels concernés. A ce titre, le QMNA5 intervient parfois dans la définition des seuils de franchissement adoptés par les « arrêtés cadres sécheresse » pour fixer les mesures de limitation ou de suspension des usages de l'eau en période de crise. Il est également une statistique de référence pour les études d'estimation des « Volumes prélevables » qui visent à répartir de manière concertée les ressources disponibles entre les différents usages, y compris les services écologiques.

La détermination d'un tel indicateur de débit est donc incontournable dans le cadre d'une gestion raisonnée de la

ressource en eau en période d'étiage. Cette connaissance n'est évidemment pas garantie en dehors des sites instrumentés disposant d'une chronique de débit suffisamment longue pour constituer une archive statistiquement exploitable et fait encore défaut en de nombreux points du réseau hydrographique, en particulier au droit des points dits « stratégiques » définis par les Schémas Directeurs d'Aménagement et de Gestion des Eaux (SDAGEs).

Pour combler ces lacunes, l'implantation de nouvelles stations hydrométriques pérennes, coûteuses ou techniquement problématique de par la configuration du site d'intérêt, ne constitue pas forcément une solution avantageuse. Des techniques d'estimation peuvent alors être envisagées. Parmi elles, deux approches sont d'usage fréquent. La première à base de schémas d'interpolation adaptés ou de régressions multiples ajustées à l'échelle régionale fournit directement une estimation de la variable d'intérêt [Laaha *et al.*, 2013]. La seconde s'appuie sur une modélisation hydrologique et reconstitue les chroniques de débits à partir de forçages météorologiques, avant d'en extraire la statistique recherchée [Engeland et Hisdal, 2009]. En ingénierie, ce sont des modèles conceptuels qui sont le plus souvent mis en œuvre ; la démarche de spatialisation des paramètres internes de transformation pluie-débit est alors requise pour une application en sites non jaugés [Folton et Lavabre, 2007]. Néanmoins, malgré les constantes améliorations apportées ces dernières années (décennie PUB), force est de constater qu'en site non jaugé, ces méthodes restent encore bien souvent approximatives, en particulier dans le cas des statistiques d'étiage ou lorsque le secteur étudié présente une certaine complexité hydrogéologique (par exemple en milieu karstique). Une solution intermédiaire, souvent négligée, consiste alors à s'appuyer sur les données de jaugeages ponctuelles, dits « jaugeages volants » ou encore « jaugeages épisodiques ».

En France, une telle démarche de valorisation des données de jaugeage à des fins d'estimation de variables d'étiage trouve son origine dans les travaux de Oberlin *et al.* [1973] qui, les premiers, ont jeté les bases d'une méthodologie permettant de les exploiter (*cf.* section III). La technique a depuis fait l'objet de quelques applications dans différents contextes climatiques [DIREN Rhône-Alpes, 2003 ; <http://www.haute-normandie.developpement-durable.gouv.fr/estimation-des-debits-de-reference-al26.html>] et de travaux de recherche complémentaires [Galéa et Canali, 2005 ; Chopart et Sauquet, 2008]. On trouve également des travaux comparables à l'étranger, en particulier aux Etats-Unis, qui ont abouti à la conception de la méthode désignée sous l'appellation de « baseflow correlation method » [Stedinger et Thomas, 1985 ; Reilly et Kroll, 2003 ; Zhang et Kroll, 2007]. Une autre approche consiste à dresser des profils en long à l'aide de jaugeages en différents points d'un même linéaire hydrologique, puis à établir une relation entre les débits jaugés et la surface drainée [Riggs, 1972 ; Carter *et al.*, 1988]. La relation peut alors être employée pour déterminer le $QMNA5$ en tout point du tronçon jaugé à partir d'une valeur de référence connue au droit d'un site instrumenté [Arts et Sarry, 2000]. Enfin, l'usage des données de jaugeages volants a également été examiné dans un contexte de modélisation hydrologique conceptuelle dans le but de mieux caler les paramètres et d'améliorer ainsi la qualité des chroniques reconstituées en site non jaugés [Rojas-Serna *et al.*, 2006 ; Seibert et Beven, 2009].

Dans la continuité des travaux ébauchés par Chopart et Sauquet [2008] sur l'estimation du $QMNA5$ sur le secteur Seine-Normandie, Catalogne [2012] a cherché à progresser sur certains aspects restés inexplorés, en particulier concernant

l'évaluation de l'incertitude affectant les estimations en fonction de la stratégie de jaugeage adoptée au site d'intérêt. Les principaux résultats sont présentés ici. La démarche engagée s'appuie sur un protocole de simulation de campagnes de jaugeage en validation croisée, à partir d'un échantillon de chroniques hydrométriques disponibles sur le plan national. Elle se place dans une perspective opérationnelle en tentant de fournir des recommandations sur la stratégie de jaugeage optimale pour parvenir à un résultat fiable dans un délai raisonnable.

Après une première partie consacrée à la présentation de données exploitées, il sera question de la méthodologie mise en œuvre, du protocole de simulation des campagnes de jaugeage et de la détermination des incertitudes propres à l'estimation du $QMNA5$. Une troisième partie sera ensuite dédiée à l'exposé des résultats obtenus avant de terminer par les principales conclusions issues de ces travaux.

II. DONNEES

Le jeu de données employé pour réaliser cette étude comprend un total de 632 bassins jaugés répartis sur l'ensemble du territoire français. De manière à permettre la détermination d'un $QMNA5$ fiable et représentatif d'un régime naturel, les chroniques correspondantes, extraites de la banque HYDRO (<http://hydro.eaufrance.fr>), comprennent au moins 26 années complètes de débits moyens journaliers sur la période 1970-2005, considérés comme peu ou pas influencés et de bonne qualité vis-à-vis de la mesure en basses eaux.

L'échantillon de bassins ainsi constitué assure une couverture spatiale d'environ 35 % du territoire et une bonne représentativité des principaux contextes hydro-climatiques rencontrés à l'échelle nationale. Les surfaces drainées se situent majoritairement dans la gamme allant de 100 à 500 km² tandis que les petits (< 20 km², peu instrumentés) et grands bassins (> 1000 km², de fait fortement anthropisés) s'avèrent assez peu représentés.

Les valeurs de $QMNA5$ ont été obtenues par ajustement d'une loi gamma à deux paramètres sur les échantillons de débits mensuels minimaux annuels extraits des chroniques. Lorsqu'ils sont exprimés sous la forme de débits spécifiques, les $QMNA5$ calculés varient entre 0 et 14,7 l/s/km² (Figure 1) pour une valeur moyenne autour de 2 l/s/km². Les valeurs les plus faibles, incluant 22 bassins avec un $QMNA5$ nul, sont principalement obtenues pour les bassins à substratum imperméable et/ou soumis à des étés particulièrement secs. A l'inverse les valeurs les plus élevées sont observées en majorité dans les secteurs montagneux mais aussi dans les secteurs où les contributions d'origine souterraine aux débits des cours d'eau sont importantes (principalement dans le Bassin Parisien).

III. METHODE

Dans les travaux de référence ébauchés par Oberlin *et al.* [1973], la démarche de valorisation des données de jaugeage passe par l'identification d'une relation entre les débits jaugés au site cible $q_{cib}(t)$ et ceux observés $q_{ref}(t)$ à la même date t au droit d'une station hydrométrique voisine que l'on désignera par la suite sous l'appellation de « site d'appui » (Figure 2). Il est alors fait l'hypothèse que cette relation constitue une bonne approximation de la relation entre la valeur de $QMNA5$ recherchée au site cible $QMNA5_{cib}$ et celle connue au site d'appui $QMNA5_{ref}$.

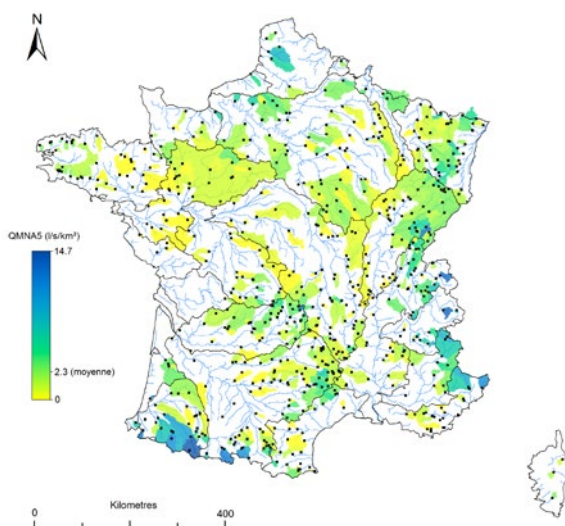


Figure 1 : Valeurs des $QMNA5$ spécifiques (en $l/s/km^2$) aux 632 stations de l'échantillon de référence

Il est question dans les prochains paragraphes de décrire le protocole mis en œuvre pour tester cette approche en validation croisée (modalités de simulation de stratégies de jaugeage, sélection d'un site d'appui et choix de la relation destinée à estimer le $QMNA5$) de manière à quantifier les incertitudes propres à l'estimation du $QMNA5$ selon la stratégie de jaugeage adoptée au site cible et la proximité du site d'appui.

III.1. Simulation des stratégies de jaugeage

L'objectif principal étant d'examiner l'influence de la stratégie de jaugeage sur les performances d'estimation, des campagnes de jaugeage ont été simulées à partir de débits tirés aléatoirement dans les chroniques des stations hydrométriques de l'échantillon de 632 bassins de référence.

Le protocole de simulation explore plusieurs couples de durées de suivi D (en année) et de fréquences annuelles de jaugeage F permettant d'obtenir un nombre total de jaugeages $N = D \times F$. Deux contraintes sont cependant imposées :

- la réalisation des F jaugeages annuels pendant la période correspondant aux trois mois le plus secs dans l'année de manière à obtenir des mesures représentatives du fonctionnement hydrologique en basses eaux,
- un espacement de 15 jours entre deux mesures consécutives de manière à assurer un minimum d'indépendance entre les débits jaugés (sur les données en saison d'étiage, l'auto-corrélation à 15 jours est inférieure à 0,3 pour environ 90 % de l'échantillon de 632 stations de référence).

Enfin, pour conférer un cadre réaliste à nos simulations vis-à-vis des contraintes opérationnelles, des bases de données de jaugeage mises à disposition par l'Agence de l'Eau Seine-Normandie et les DREALs Rhône-Alpes et Provence-Alpes-Côte d'Azur ont été examinées. Les situations rencontrées d'un point à l'autre sont très diverses, tant en termes de nombre de mesures réalisées qu'en termes de durées. Certains points de suivi présentent en effet un nombre élevé de jaugeages (N supérieur à 100 voire proche de 500) tandis que d'autres ne disposent que d'une unique mesure. La majorité des campagnes comptent moins de 30 mesures pour une durée de suivi généralement inférieure à 10 années. Il apparaît toutefois de manière presque

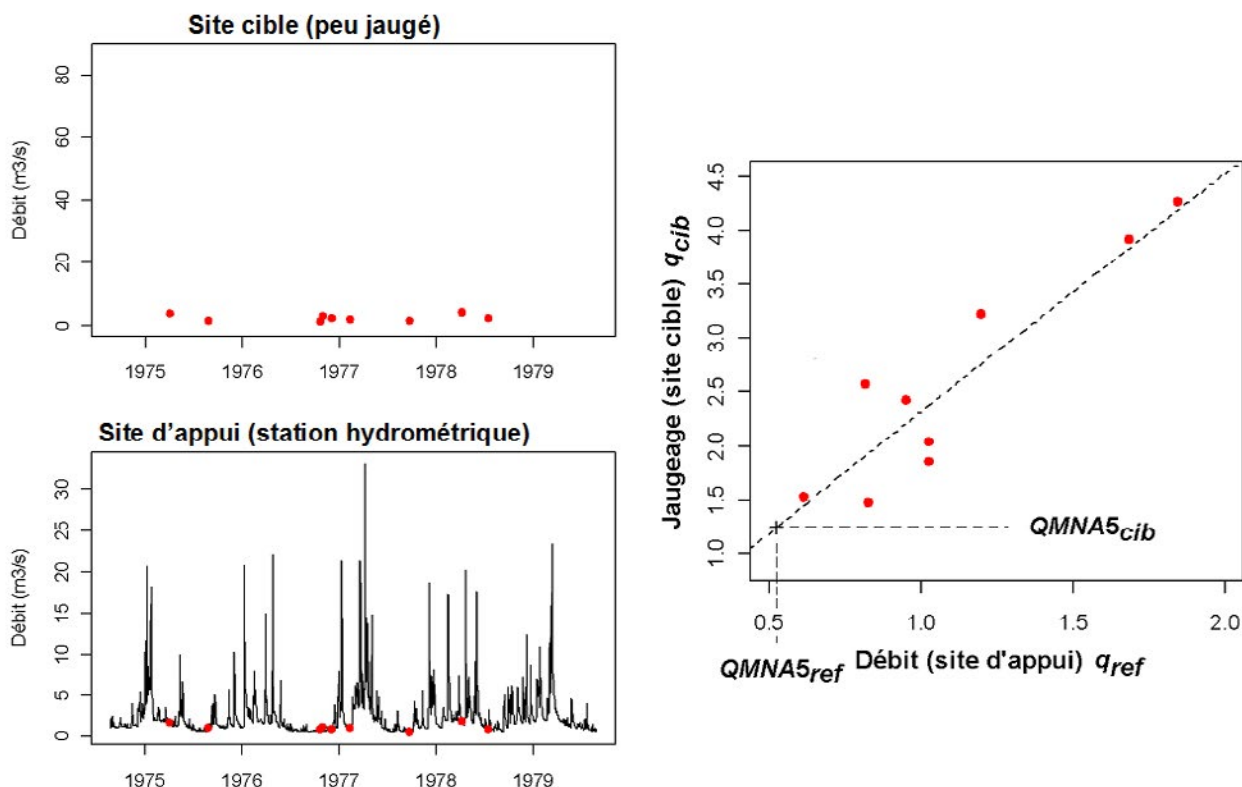


Figure 2 : Principe de l'exploitation des jaugeages épisodiques sur la base d'une relation ajustée sur neuf couples jaugeages au site cible et débits observés aux mêmes dates sur le bassin d'appui

généralisée que la principale contrainte opérationnelle repose sur la fréquence F qui, sur le long terme, ne dépasse que rarement 4 à 5 mesures en moyenne chaque année.

Compte tenu de ces éléments et étant donné que la simulation des stratégies de jaugeage ne peut pas reproduire de manière exhaustive l'ensemble des situations rencontrées en pratique, il a été choisi de se limiter ici à un jeu de simulations reposant sur l'ensemble des combinaisons de durées $D = 2, 4, 8, 16$ et de fréquences de jaugeage $F = 1, 2, 3, 4, 5$. Seule la configuration faisant intervenir $D = 2$ et $F = 1$ n'a pas été mise en œuvre de manière à disposer d'au moins quatre jaugeages simulés, nombre de mesures en deçà duquel il serait hasardeux d'ajuster une relation entre $q_{cib}(t)$ et $q_{réf}(t)$. Le nombre de jaugeages simulés varie au final entre 4 et 80.

Dans cette application, il est admis que les débits moyens journaliers échantillonnés peuvent être assimilés à des débits instantanés. Cette hypothèse, imposée par la nature des données disponibles aux stations de référence, reste cependant à discuter en raison de l'existence de variations de débits au cours de la journée. Pour aborder la question, les débits instantanés disponibles pour 455 stations de l'échantillon de référence ont été comparés aux débits moyens journaliers correspondants pour chaque date de la période retenue dans le protocole d'échantillonnage (durant le trimestre le plus sec). Les valeurs de débit instantané examinées s'écartent de moins de 10% de la moyenne journalière dans environ la moitié des cas, quelle que soit la gamme de débit considérée. Ces écarts ne dépassent 50% que pour 15 à 10% des dates échantillonnées. Ces chiffres montrent qu'il peut exister des variations de débit importantes au cours d'une même journée en réponse à des épisodes de pluie qui peuvent altérer *in fine* la qualité de la relation (cf. section III.3) établie entre les débits des sites cible et de référence. Le protocole de simulation adopté ici occulte ces problèmes de stabilisation du régime d'écoulement, qui participent certainement aux incertitudes sur l'estimation finale du $QMNA5$, sans que leur contribution relative ait pu être quantifiée. En pratique, les services opérationnels pourront éliminer certaines valeurs « aberrantes » après expertise des données recueillies (en se référant par exemple à l'historique climatique du bassin étudié) ou restreindre leurs mesures aux périodes jugées favorables, sur la base des conditions hydrométéorologiques observées.

III.2. Sélection d'un site d'appui

La problématique de sélection d'un site d'appui est une étape clé qui vise ici à identifier le bassin instrumenté le plus approprié pour parvenir à une estimation correcte (sous-entendu qu'il réponde déjà aux critères permettant la détermination d'un $QMNA5$ de référence fiable, cf. section II). Ceci renvoie à la question de l'identification de similitudes entre bassins et la construction de voisinages hydrologiques autour du site cible. La littérature abonde sur le sujet ; les approches sont multiples et, de ce fait, les inter-comparaisons le sont également (par exemple : Laaha et Blöschl [2006] sur les caractéristiques d'étiage, Sauquet et Catalogne [2011] sur les courbes des débits classés ou encore Yuan [2013] sur une série complète de débits journaliers).

Dans le cadre de la valorisation des données de jaugeage, il semble important de s'assurer que le bassin de référence sélectionné réagit aux mêmes événements météorologiques que ceux éprouvés par le bassin cible. Pour ce faire, la plupart des travaux antérieurs imposent en premier lieu une contrainte sur la proximité géographique entre le bassin cible et le bassin de référence sélectionné en privilégiant

notamment les sites instrumentés situés sur le même linéaire hydrologique que le site cible [Chopart et Sauquet, 2008]. Toutefois, dans de nombreux cas, l'absence de station hydrométrique sur le tronçon étudié justifiera d'étendre la recherche d'un site de référence approprié à d'autres bassins voisins du bassin cible. A cet effet, certains auteurs proposent d'orienter la sélection des sites de référence en considérant d'autres critères de similitude entre bassins. Reilly et Kroll [2003] ont par exemple privilégié l'emploi de variables auxiliaires descriptives du bassin (*i.e.* descripteurs se rapportant au climat, à la morphologie ou à la géologie des bassins). D'autres recommandent de caractériser indirectement le degré de similitude entre bassins sur la base de la corrélation entre débits mesurés à la même date [Archfield et Vogel 2010]. L'emploi de la corrélation présente alors l'avantage de disposer d'un indicateur directement issu de l'information hydrologique disponible et pouvant se substituer de manière relativement « universelle » et objective à l'emploi de descripteurs de bassin (*i.e.* quel que soit le contexte hydro-climatique du site cible). Dans le cadre de l'usage de jaugeages volants, Stedinger et Thomas [1985] préconisent par exemple de ne retenir que les sites de référence dont le coefficient de corrélation dépasse 0,7.

Dans la continuité de cette dernière approche, nous proposons tout d'abord de présélectionner pour chaque site cible un ensemble de 50 sites de référence sur la base de la proximité géographique mesurée par la distance d entre bassins suggérée par Ghosh [Gottschalk *et al.*, 2011] :

$$d = \frac{1}{S_{ref} \cdot S_{cib}} \cdot \sum_{i=1}^n \sum_{j=1}^{n'} |x_{ref}(i) - x_{cib}(j)| \quad (1)$$

où S_{cib} et S_{ref} sont respectivement les surfaces drainées par le site cible et le site de référence, $x_{ref}(i)$, $i = 1, \dots, n$ et $x_{cib}(j)$, $j = 1, \dots, n'$ sont les coordonnées des points d'un maillage régulier superposé sur les bassins considérés. L'avantage de cette distance est la prise en compte de la topologie particulière du réseau hydrographique ; elle s'avère notamment plus appropriée pour quantifier la distance entre bassins emboîtés. Cette distance n'étant pas aisée à calculer, la distance entre centres de gravité pourra toutefois être envisagée par défaut pour réaliser la présélection.

Pour chacun de ces 50 bassins de référence est ensuite calculé le coefficient de corrélation entre $q_{cib}(t)$ et $q_{réf}(t)$ pour finalement sélectionner le site d'appui présentant le plus fort degré de corrélation r . Ce coefficient sera employé par la suite pour quantifier la « proximité » entre bassin cible et bassin d'appui puis déterminer dans quelle mesure celle-ci influe sur les performances d'estimation.

Remarquons que, selon ce protocole, le choix d'un site d'appui pourra évoluer avec l'augmentation du nombre de jaugeages réalisés. En effet, la forte incertitude associée à la corrélation calculée lorsque le nombre de jaugeages est faible se traduit par une incertitude sur le choix du site de référence le plus pertinent. L'ajout de nouveaux jaugeages doit alors permettre de converger progressivement vers le site de référence le plus approprié. Cet aspect doit évidemment s'inscrire dans une démarche opérationnelle de mise à jour des résultats au cours de la campagne de jaugeage.

III.3. Quelle relation entre les débits de deux bassins à la même date ?

Contrairement aux travaux français antérieurs qui préconisaient l'ajustement de relations linéaires entre $q_{cib}(t)$ et $q_{réf}(t)$,

après différents tests, la formulation employée dans les travaux américains a été privilégiée ici :

$$\ln(q_{cib}(t)) = \lambda + k \cdot \ln(q_{réf}(t)) \quad (2)$$

où k et λ sont les paramètres ajustés par régression linéaire sur l'ensemble des couples de débits $q_{cib}(t)$ et $q_{réf}(t)$ après transformation logarithmique.

L'avantage est alors de donner un poids plus important aux valeurs jaugées les plus basses.

De plus, par opposition aux relations linéaires, la forme log-linéaire semble plus justifiée pour décrire l'évolution conjointe des débits en phase de récession pour deux bassins dotés de propriétés de vidange différentes. En effet, en supposant la loi de Maillet appropriée pour décrire la dynamique de récession ($Q(t) = Q_0 \exp(-\alpha t)$), la formulation (2) correspond effectivement à la solution analytique permettant de décrire la relation entre les débits de deux bassins caractérisés par des coefficients de tarissement différents. La relation retenue ici permet ainsi de s'affranchir d'une trop stricte conformité des dynamiques de récession de deux bassins voisins sous-entendue par les relations linéaires.

Elle ne permet pas en revanche d'exploiter (ni d'estimer) des débits nuls. Pour traiter ce cas de figure, les valeurs nulles sont par défaut remplacées par 1 l/s, seuil de précision des données extraites de la banque HYDRO.

Les paramètres obtenus suite à l'ajustement sont ensuite exploités pour estimer le $QMNA5_{cib}$ au site cible à partir de la valeur de $QMNA5_{réf}$ connue au site d'appui :

$$QMNA5_{cib} = \exp(\lambda) (QMNA5_{réf})^k \quad (3)$$

III.4. Evolution et modélisation des incertitudes obtenues en validation croisée

Le protocole d'estimation décrit précédemment (*i.e.* choix d'un site d'appui puis ajustement de la relation entre $q_{cib}(t)$ et

$q_{réf}(t)$) a été appliqué en validation croisée pour chaque stratégie de jaugeage simulée en considérant tour à tour chaque site instrumenté disponible comme un site cible. L'erreur commise est ensuite quantifiée au moyen de l'écart entre la valeur de $QMNA5$ prédite par la relation ($QMNA5_{cib}$) et la valeur de référence calculée au site cible $QMNA5_{obs}$ suite à l'ajustement de la loi gamma sur les échantillons de débits mensuels minimaux annuels extraits de la totalité de la chronique disponible :

$$\varepsilon = QMNA5_{cib} - QMNA5_{obs} \quad (4)$$

Elle est exprimée en l/s/km² pour autoriser la comparaison entre sites.

Pour chaque stratégie simulée, l'ensemble des résultats obtenus (*i.e.* toutes stations confondues) a ensuite été rassemblé par classe de corrélation (r) de manière à obtenir une représentation des distributions d'erreurs en fonction de la « proximité » au site cible.

Après vérification, ces distributions se sont avérées suffisamment proches de distributions normales pour être finalement résumées au travers de leur moyenne $\mu_{(r,N,F)}$ (ou biais) et de leur écart-type $\sigma_{(r,N,F)}$ et ainsi permettre de construire un intervalle de confiance autour de chaque estimation en fonction de la corrélation (r), du nombre de jaugeages (N) et de la fréquence annuelle de jaugeage (F) :

$$P\left(\mu_{(r,N,F)} - u_{1-\alpha/2} \cdot \sigma_{(r,N,F)} < \varepsilon < \mu_{(r,N,F)} + u_{1-\alpha/2} \cdot \sigma_{(r,N,F)}\right) = 1 - \alpha \quad (5)$$

où $u_{1-\alpha/2}$ correspond au quantile normal centré réduit de fréquence au non-dépassement $1-\alpha/2$.

Un exemple d'évolution du biais et de l'écart-type en fonction de ces différents paramètres est donné sur la Figure 3 pour une fréquence de cinq jaugeages par an. Dans les deux cas, on constate une nette diminution des valeurs obtenues μ et σ avec l'augmentation de la corrélation (autrement dit l'incertitude diminue logiquement avec le renforcement

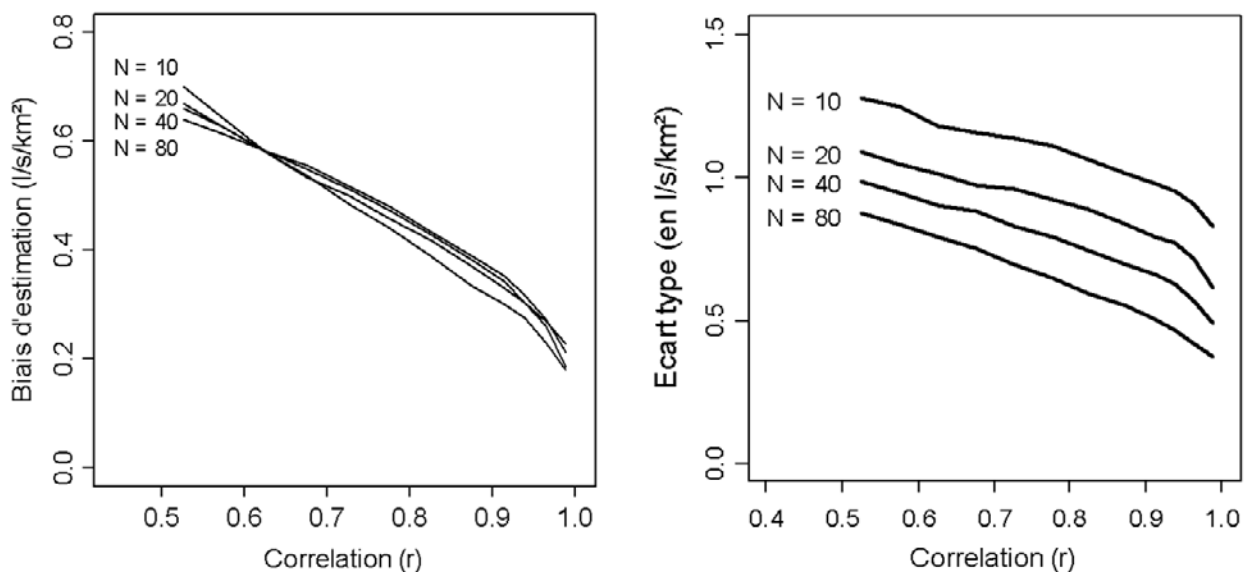


Figure 3 : Exemple d'évolution du biais μ et d'écart-type σ en fonction de la corrélation et du nombre de jaugeages pour une fréquence annuelle de cinq jaugeages

de la « proximité » entre le bassin cible et le bassin de référence). En revanche, l'évolution en fonction du nombre de jaugeages réalisés diffère sensiblement. En effet, alors que le biais semble peu sensible à ce paramètre, on constate une diminution significative de l'écart-type lorsque N augmente. Par ailleurs, les courbes représentées s'avèrent globalement parallèles, indiquant que lorsque l'on double le nombre de jaugeages (ou, par équivalence, la durée de suivi), le gain de performance peut être considéré comme relativement constant quel que soit le degré de corrélation entre les débits mesurés au site cible et les débits observés aux sites de référence.

En s'appuyant sur ces observations, l'étape suivante a consisté à synthétiser ces résultats au travers de modèles analytiques simples de manière à pouvoir estimer μ et σ pour des situations de jaugeage non simulées et faciliter l'interprétation de l'influence de chaque aspect de la stratégie de jaugeage sur l'incertitude d'estimation. Un modèle décrivant le biais μ et l'écart-type σ a donc été ajusté aux données pour chaque fréquence de jaugeage simulée en fonction de la corrélation (r) et du nombre de jaugeages (N) :

$$\mu_{(r,N,F)} = \alpha_{(F)} \cdot r + \beta_{(F)} \quad (6)$$

$$\sigma_{(r,N,F)} = a_{(F)} \cdot \ln(N) + b_{(F)} \cdot (1-r) + c_{(F)} \quad (7)$$

Quoique reposant sur des hypothèses très simplificatrices (évolutions en fonction de r et N considérées comme

linéaires et indépendantes), ces modèles semblent satisfaisants avec des écarts aux valeurs observées suffisamment faibles (exemples donnés en Figure 4 pour l'écart-type σ et pour des fréquences de jaugeage $F = 1, 2$ et 4) pour apprécier les bornes d'un intervalle de confiance à 80 % ou 95 %. Les coefficients de ces modèles sont reportés dans le Tableau 1. Ils sont exploités dans les paragraphes suivants pour mieux caractériser l'évolution des performances d'estimation du *QMNA5* en fonction de la stratégie de jaugeage.

IV. RECOMMANDATIONS OPÉRATIONNELLES

Au-delà des considérations liées à la « proximité » entre site cible et bassin de référence, nécessaire à l'obtention de résultats performants, l'analyse menée ici vise à définir les modalités d'une campagne de jaugeage permettant de parvenir à une estimation fiable avec un minimum de données (de manière à minimiser le « coût opérationnel ») et/ou dans un délai raisonnable. Cette analyse repose ainsi sur deux interrogations principales concernant (i) l'influence du nombre de jaugeages réalisés (pour une même fréquence annuelle de jaugeage, y a-t-il un gain à poursuivre ou non une campagne de mesure et sur quelle durée ?) et (ii) l'influence de la fréquence annuelle des jaugeages (pour un même nombre final de jaugeage, peut-on écarter la période de suivi – c'est-à-dire augmenter la fréquence de jaugeage – sans perdre d'information compte tenu de la possible dépendance intra-annuelle des valeurs jaugées ?).

L'examen de ces deux aspects permet d'aller plus loin que la plupart des travaux relevés dans la littérature qui se limitent

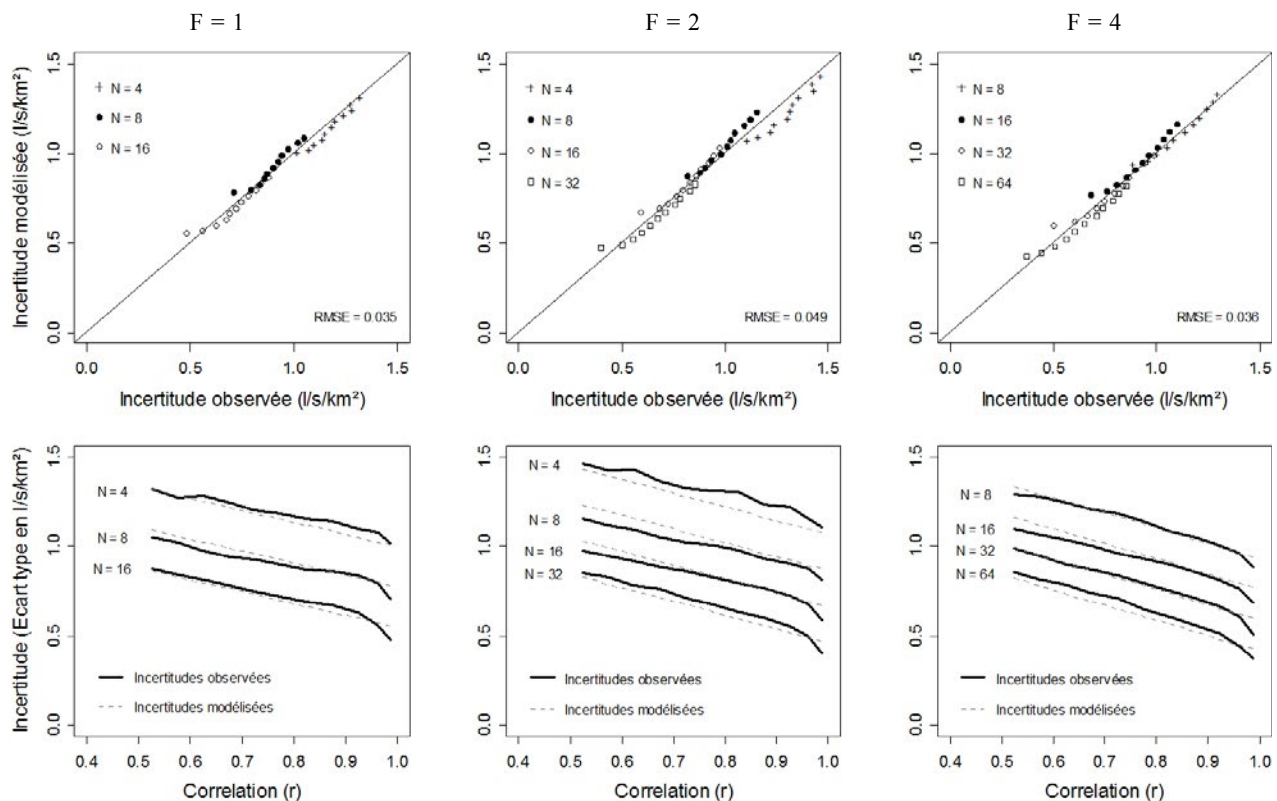


Figure 4 : Exemple de modèles d'écart-type ajustés en fonction de la corrélation et du nombre de jaugeages pour trois fréquences de jaugeage par saison d'été

Tableau 1 : Paramètres des modèles de biais et d'écart-type des erreurs ajustés pour chaque fréquence de jaugeage simulée

F	$\alpha_{(F)}$	$\beta_{(F)}$	$a_{(F)}$	$b_{(F)}$	$c_{(F)}$
1	-0.908	1.095	-0.326	0.664	1.452
2	-0.929	1.136	-0.290	0.770	1.468
3	-0.938	1.154	-0.263	0.822	1.441
4	-0.922	1.155	-0.248	0.849	1.446
5	-0.942	1.177	-0.238	0.887	1.443

généralement à examiner l'influence de l'accroissement du nombre de jaugeages sans réelle prise en considération de la durée de suivi et surtout de la fréquence de jaugeage.

Pour traiter ces questions il sera uniquement fait appel aux modèles d'écart-type $\sigma_{(r,N,F)}$ présentés précédemment. En effet, l'écart-type détermine l'étendue de l'intervalle de confiance et, de ce fait, constitue le critère le plus important pour évaluer ici l'évolution des incertitudes en fonction de la stratégie de jaugeage adoptée au site cible ; le biais moyen $\mu_{(r,N,F)}$ n'intervient, lui, que dans un second temps pour permettre la correction des estimations et l'évaluation finale des bornes d'un intervalle de confiance de niveau de significativité voulu.

IV.1. Nombre de jaugeages

La Figure 5 illustre le gain cumulé théorique apporté par l'augmentation du nombre de jaugeages (volontairement poussé jusqu'à 100) pour chaque fréquence de jaugeage à corrélation constante :

$$gain_{(r,N,F)} = -\sigma_{(r,N,F)} + \sigma_{(r,1,F)} = -a_{(F)} \cdot \ln(N) \quad (8)$$

ou exprimé au travers de la dérivée :

$$\partial gain_{(r,N,F)} / \partial N = -a_{(F)} / N \quad (9)$$

On constate ici que ce gain est assez sensible jusqu'à environ 20 ou 25 jaugeages quelle que soit la fréquence de jaugeage F considérée, puis qu'il s'amortit au delà d'autant plus rapidement que $a_{(F)}$ est faible et donc que F est élevée (équation (8)).

Ce résultat s'accorde relativement bien avec les conclusions des différents travaux relevés dans la littérature [Stedinger, 1985 ; Reilly et Kroll, 2003] et peut donc permettre d'anticiper sur le nombre de jaugeages à réaliser lors de la programmation d'une campagne de mesures (et ainsi chiffrer le coût prévisionnel de celle-ci) et, au-delà, sur la durée optimale de suivi d'un site selon la fréquence de jaugeage adoptée. Notons que le résultat est également du même ordre de grandeur que l'effectif minimal conseillé d'environ 30 jaugeages répartis sur une période de 3 à 5 ans et de 32 valeurs sur 10 ans suggéré respectivement par Rojas-Serna *et al.* [2006] et par Seibert et Beven [2009] pour une exploitation des jaugeages volants dans un contexte de modélisation hydrologique.

Cette conclusion seule reste cependant insuffisante car, comme tend à le démontrer la Figure 5, pour un même nombre de jaugeages réalisés, le gain est d'autant plus faible que la fréquence F est forte. Concrètement, à nombre de jaugeage égal, cela se traduit par un intervalle de confiance

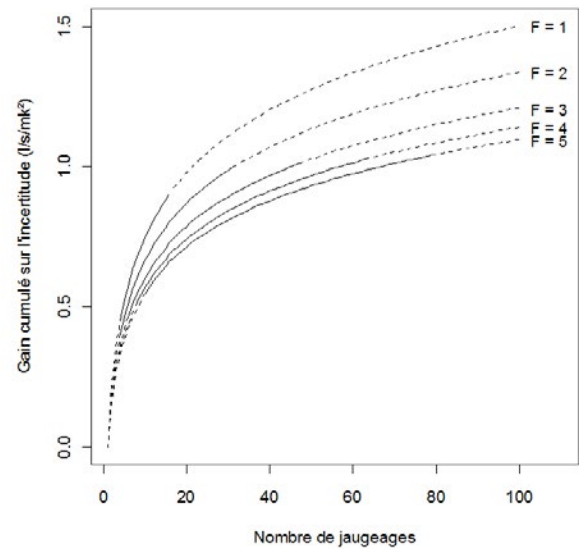


Figure 5 : Gain cumulé sur l'incertitude lié à l'augmentation du nombre de jaugeages pour cinq fréquences de jaugeage en supposant une corrélation constante (en trait plein sur la gamme des configurations simulées, en pointillé en dehors)

plus large lorsque la fréquence de jaugeage augmente. Par exemple à l'aide du Tableau 1, on peut déterminer que : pour un niveau de corrélation « moyen » $r = 0,75$ et un total de $N = 15$ jaugeages réalisés, l'intervalle de confiance à 95 % ($\alpha = 0,05$) est de $[-1,03 ; 1,86]$ l/s/km² lorsque $F = 1$ et atteint $[-1,53 ; 2,47]$ l/s/km² lorsque $F = 5$. Il semble bien que l'hypothèse de l'existence d'une dépendance intra-annuelle résiduelle des débits jaugés (*i.e.* lorsque $F > 1$), malgré la contrainte d'espacement minimal de 15 jours imposée ici entre deux jaugeages simulés, soit à l'origine de ce résultat et qu'elle soit d'autant plus fortement ressentie que le nombre de jaugeages est élevé (autrement dit plus on jauge, plus le gain d'information par rapport à la réalisation d'un seul jaugeage par année diminue). Pour affiner cette conclusion, il est indispensable d'examiner avec plus d'attention l'influence de la fréquence de jaugeage avant de déterminer le couple durée de suivi / fréquence de jaugeage optimal.

IV.2. Fréquence de jaugeage

Pour illustrer l'influence de la fréquence de jaugeage sur l'incertitude des estimations il a été choisi de raisonner au travers d'une variable dénommée « durée de suivi équivalente » (D_{eq}). Pour une fréquence de jaugeage donnée F , D_{eq} est définie comme étant la durée requise pour parvenir aux mêmes performances en termes d'écart-type que celles obtenues avec un autre couple de fréquence de jaugeage $F_{réf}$ et de durée de référence $D_{réf}$. Ainsi, pour une corrélation identique r et en se fondant uniquement sur l'équation (7), l'équation à résoudre est la suivante :

$$\begin{aligned} a_{(F_{réf})} \cdot \ln(D_{réf} \cdot F_{réf}) + b_{(F_{réf})} \cdot (1-r) + c_{(F_{réf})} \\ = a_{(F)} \cdot \ln(D_{eq} \cdot F) + b_{(F)} \cdot (1-r) + c_{(F)} \end{aligned} \quad (10)$$

La Figure 6 représente les valeurs de D_{eq} en fonction de $D_{réf}$ obtenus pour $F = 2, 3, 4$ et 5 , en prenant comme fréquence

de référence la réalisation d'un jaugeage chaque année ($F_{\text{réf}} = 1$) et en considérant trois coefficients de corrélation $r = 0,5, 0,75$ ou 1 .

Les points se placent sous la première bissectrice (D_{eq} reste ici inférieure à $D_{(F=1)}$). Ainsi si l'objectif est d'aboutir à un résultat équivalent à celui obtenu pour 20 jaugeages répartis sur 20 années ($F = 1$), il est en réalité possible (pour $r = 0,75$) d'écourter la période de suivi à environ 17 ans pour une fréquence de jaugeage $F = 2$ (soit un total de 34 jaugeages) et à environ 15 années pour une fréquence de jaugeage $F = 4$ (soit un total de 60 jaugeages).

En généralisant au-delà de cet exemple particulier, il apparaît donc avantageux d'accroître la fréquence annuelle de jaugeage pour écourter la durée de suivi d'un site, sous réserve que le « coût opérationnel » lié à l'augmentation du nombre de jaugeages nécessaire ne soit pas une contrainte. A cet égard, la réalisation de trois jaugeages par année constitue sans doute le meilleur compromis. Au delà (pour $F = 4$ ou $F = 5$) la réduction des durées de suivi s'avère pratiquement négligeable (inférieure à 1 an).

IV.3. Exploitation de jeux de données existant dans le cas de configurations complexes

Jusqu'à présent, l'analyse des modèles d'incertitude a porté sur des configurations simples, de fréquence de jaugeage constante. Il s'agit ici de préciser comment exploiter des campagnes existantes présentant des suivis irréguliers, cas rencontré le plus souvent en pratique. La solution proposée consiste à traiter ces séries comme une succession de périodes pour lesquelles les fréquences de jaugeage sont constantes.

A titre d'exemple, considérons une campagne de jaugeage d'une durée totale de $D = 20$ années au cours de laquelle le suivi a été effectué avec une fréquence de deux jaugeages par an sur les dix premières années ($F_1 = 2$ et $D_1 = 10$) puis d'un jaugeage par an les dix années suivantes ($F_2 = 1$ et $D_2 = 10$). Le nombre total de jaugeages réalisés atteint $N = 30$ et la fréquence de jaugeage moyenne est égale à $F = N / D = 1,5$.

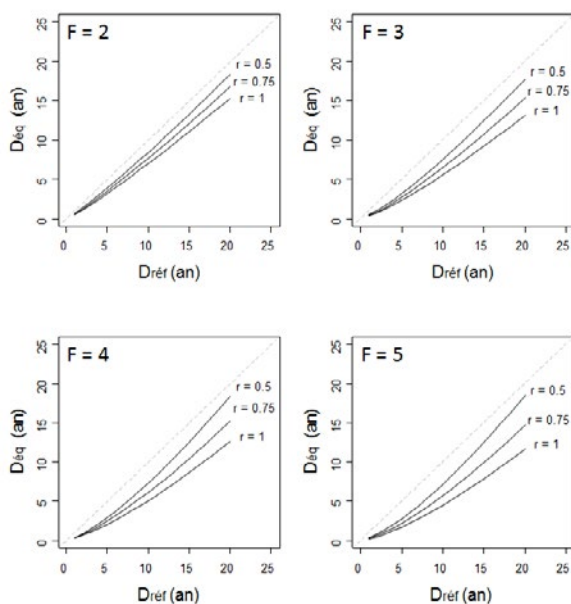


Figure 6 : Durée de suivi équivalente D_{eq} pour $F > 1$ en fonction de la durée de suivi $D_{\text{réf}}$ pour $F_{\text{réf}} = 1$ pour différents niveaux de corrélation

Si le premier modèle d'incertitude permet de quantifier directement l'incertitude de l'estimation obtenue au bout des dix premières années, notée $\sigma_{(r_1, F_1, D_1)}$, le second modèle servira en revanche à quantifier le gain apporté par les nouvelles mesures réalisées au cours de la période D_2 , noté $\Delta\sigma_{(r_2, F_2, D_2)}$. Ainsi pour l'exemple proposé, l'incertitude finale sur la totalité de la période de suivi sera simplement quantifiée par :

$$\sigma_{(r, N, F)} = \sigma_{(r_1, N_1, F_1)} + \Delta\sigma_{(r_2, F_2, D_2)} \quad (11)$$

Pour parvenir à ce résultat il est cependant nécessaire de quantifier correctement le gain $\Delta\sigma_{(r_2, F_2, D_2)}$. En effet, dans le cas d'une fréquence de jaugeage F et d'une corrélation r constantes, ce gain (équations (7) et (11)) serait théoriquement donné par :

$$\Delta\sigma_{(r, N_2, F)} = a_{(F)} \left(\ln(N_1 + N_2) - \ln(N_1) \right) \quad (12)$$

Dans le cas de campagnes complexes, on se ramènera à une fréquence constante en calculant préalablement le nombre de jaugeages équivalent N_{eq} à N_1 (en prenant donc $F_1 = 1$), ce qui permet de prolonger le modèle sur la période la plus récente. Il en résulte que :

$$\Delta\sigma_{(r, N_2, F_2)} = a_{(F_2=1)} \cdot \left(\ln(N_{\text{eq}} + N_2) - \ln(N_{\text{eq}}) \right) \quad (13)$$

Les résultats obtenus par cette méthode pour déterminer $\sigma_{(r, N, F)}$ avec les valeurs numériques prises en exemple sont illustrés en Figure 7 pour trois niveaux de corrélation distincts. Celle-ci permet bien de constater l'apparition d'une rupture dans l'évolution de l'écart-type liée au changement de la fréquence de jaugeage. En pratique, cette démarche pourra être répétée autant de fois qu'il y a de modifications de la fréquence de jaugeage au cours de la campagne de suivi.

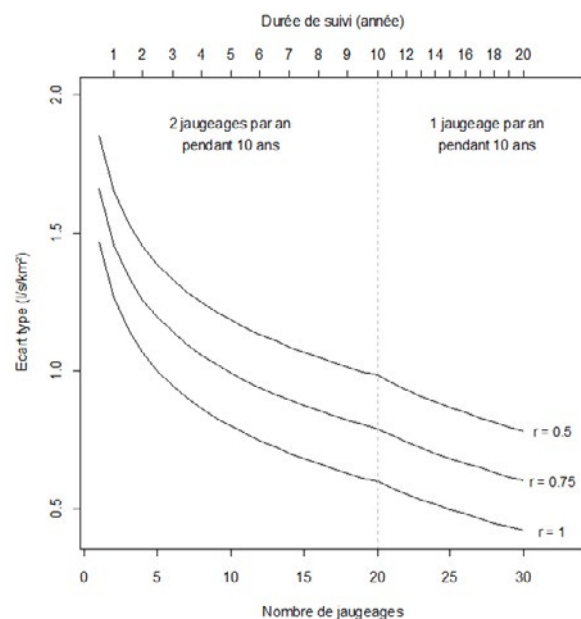


Figure 7 : Exemple d'évolution de l'incertitude en fonction du nombre de jaugeages dans le cas d'une modification de la fréquence de jaugeage au cours du suivi d'un site cible

V. CONCLUSION

Cet article présente une méthode destinée à valoriser des données de jaugeage volant dans le cadre de l'estimation du débit de référence d'étiage *QMNA5*. Celle-ci repose sur la mise en relation entre débits jaugés au site cible et débits observés à la même date au droit d'une station hydrométrique voisine (un « site d'appui »). Il est alors fait l'hypothèse que cette relation constitue une bonne approximation de la relation entre la valeur de *QMNA5* recherchée au site cible et celle connue au site d'appui.

Dans le cadre de nos travaux, la démarche employée pour déterminer les performances de cette méthode repose sur une procédure de validation croisée, menée à l'échelle nationale sur un jeu de 632 bassins de référence représentatifs de différents contextes hydro-climatiques. Celle-ci a permis de quantifier, au travers de différentes simulations, les incertitudes d'estimation du *QMNA5* en fonction (i) de la stratégie de jaugeage adoptée au site cible (nombre de jauges réalisées, fréquence annuelle de jaugeage) et (ii) du choix d'un site référence (réalisé sur la base de la corrélation entre débits jaugés et débits observés à la même date au droit d'une station hydrométrique géographiquement voisine). L'ensemble des résultats obtenus a été résumé à l'aide de modèles analytiques simples, permettant de déterminer un intervalle de confiance autour de chaque estimation puis de fournir quelques recommandations opérationnelles pour optimiser la programmation de futures campagnes de mesure.

Ainsi, outre l'importance de disposer d'un site d'appui présentant un comportement hydrologique suffisamment voisin de celui du site cible (degré élevé de corrélation entre débits observés à la même date), les résultats obtenus montrent que le gain de performance apporté par l'augmentation du nombre de jaugeages est sensible jusqu'à 20 mesures environ puis qu'il est moindre au-delà. Parallèlement il apparaît que, pour un même nombre de jaugeages, la réalisation de plusieurs jaugeages au cours d'une même année (*i.e.* au cours d'une même saison d'étiage chaque année) conduit à un gain de performance moins rapide en raison, semble-t-il, de la dépendance intra-annuelle des débits en basses eaux. Néanmoins les résultats indiquent qu'il est tout de même possible d'écourter un peu la durée de suivi d'un site en augmentant la fréquence annuelle de jaugeage sans compromettre la qualité de l'estimation. A cet égard, la réalisation de trois jaugeages chaque année semble constituer le meilleur compromis.

Compte tenu des besoins en la matière, la méthode qui est présentée ici constitue sans doute une alternative intéressante (moins coûteuse) à l'implantation d'une station pérenne sur les bassins non jaugés à enjeu ou sur les sites difficiles à instrumenter (ex. lits en tresse, sections instables...). Les travaux menés dans la thèse de Catalogne [2012] ont par ailleurs confirmé que la valorisation d'une information hydrologique locale sous forme de jaugeages ponctuels, même très fragmentaire (dès deux ans de suivis), peut conduire à des estimations plus fiables que celles obtenues par rapport à une estimation en sites non jaugés, bien que celle-ci s'appuie sur des méthodes de régionalisation relativement sophistiquées. Ainsi, la démarche proposée pourra sans doute être avantageusement étendue à d'autres statistiques d'étiage (par exemple les quantiles $VCNd(T)$ de débits moyens sur d jours minimaux dans l'année et de période de retour T), sous réserve de reproduire le protocole de validation croisée pour permettre de quantifier les incertitudes associées. Pour une application sur une région spécifique, il pourrait enfin être pertinent de réajuster les modèles de biais et d'écart-type des erreurs d'estimation pour une appréciation plus fine des

incertitudes dans un contexte hydro-climatique plus homogène. Ceci nécessite d'identifier dans le secteur d'intérêt un jeu de stations de référence aux enregistrements longs et fiables et de reprendre le protocole de validation croisée.

VI. REMERCIEMENTS

Cette étude a été réalisée grâce au soutien financier de l'Onema et d'une collaboration entre les différents centres régionaux d'Irstea. Nous souhaitons également adresser nos remerciements à L. Durand de la DREAL PACA pour l'échange de données et de résultats qui ont permis de tester la méthode en conditions réelles et à l'Agence de l'Eau Seine Normandie et à la DREAL Rhône-Alpes pour la mise à disposition des données de jaugeage.

VII. BIBLIOGRAPHIE

- ARCHFIELD S. A., VOGEL R. M. (2010) — Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments. *Water Resour. Res.* **46** W10513
- ARTS I., SARY M. (2000) — Cartographie des débits d'étiage : application au bassin de la Meurthe. mis en ligne le 05 août 2013, consulté le 27 octobre 2013. *Revue Géographique de l'Est*. **40(1-2)** URL :
- CATALOGNE C. (2012) — *Amélioration des méthodes de prédétermination des débits de référence d'étiage en sites peu ou pas jaugés*. Thèse de Doctorat, Université Joseph Fourier, Grenoble. 285 pages
- CHOPART S., SAUQUET E. (2008) — Usage des jaugeages volants en régionalisation des débits d'étiage. *Revue des sciences de l'eau / Journal of Water Science*. **21(3)** 267-281
- DIREN RHONE ALPES (2003) — *Les débits d'étiage en Rhône-Alpes. Rapport d'étude*
- ENGELAND K., HUSDAL H. (2009) — A comparison of low flow estimates in ungaged catchments using regional regression and the HBV-model. *Water Resources Management*. **23(12)** 2567-2586
- GALEA G., CANALI S. (2005) — Régionalisation des modules annuels et des régimes d'étiage du bassin hydrographique de la Moselle française : lien entre modèles régionaux. *Revue des Sciences de l'Eau / Journal of Water Science*. **18(3)** 331-352
- GOTTSCALK L., LEBLOIS E., SKØIEN J.O. (2011) — Distance measures for hydrological data having a support. doi 10.1016/j.jhydrol.2011.03.020. *Journal of Hydrology*. **402(3-4)** 415-421
- LAAHA G., BLÖSCHL G. (2006) — A comparison of low flow regionalisation methods - catchment grouping. *Journal of Hydrology*. **323(1-4)** 193-214
- LAAHA G., DEMUTH S., HUSDAL H., KROLL C.N., VAN LANEN H.A.J., NESTER T., ROGER M., SAUQUET E., TALLAKSEN L.M., WOODS R., YOUNG A. (2013) — Prediction of low flows in ungauged basins. In « *Runoff Prediction in Ungauged Basins - Synthesis across Processes, Places and Scales* », (Blöschl G., Sivapalan M., Wagener T., Viglione A. & Savenije H., eds.), Chap 8. Cambridge University Press, Cambridge, UK. 163-188
- FOLTON N., LAVABRE J. (2007) — Approche par modélisation pluie-débit pour la connaissance régionale de la ressource en eau : application à la moitié du territoire français. *La Houille Blanche*. **3** 64-70
- OBERLIN G., GALEA G.C., TONI J.T. (1973) — Intérêt des jaugeages épisodiques pour l'estimation des étiages de petits bassins non équipés. *Cahiers de l'ORSTOM, sér. Hydrol.* **X(4)** 349-368
- REILLY C.F., KROLL C.N. (2003) — Estimation of low streamflow statistics using baseflow correlation. *Water Resources Research*. **39(9)** : 1236, doi10.1029/2002WR001740

- RIGGS H.C. (1972) — Low-flow estimation. *Techniques of Water-Resources Investigations of the United States Geological Survey. USGS publication*
- ROJAS-SERNA C., MICHEL C., PERRIN C. ANDRÉASSIAN V. (2006) — Ungauged catchments: How to make the most of a few streamflow measurements? *IAHS Publication*. **307** 230-236
- SAUQUET E., CATALOGNE C. (2011) — Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France. doi10.5194/hess-15-2421-2011. *Hydrology and Earth System Sciences*. **15** 2421–2435
- SEIBERT J., BEVEN K.J. (2009) — Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci.* **13** 883-892
- STEDINGER J.R., THOMAS W.O. (1985) — Low-flow frequency estimation using base-flow measurements. *USGS Scientific Investigations, Report*. 85-95
- YUAN L. L. (2013) — Using correlation of daily flows to identify index gauges for ungauged streams. *Water Resour. Res.* 49
- ZHANG Z., KROLL C. (2007) — The baseflow correlation method with multiple gauged sites. *Journal of Hydrology*. **347** 371-380