



**HAL**  
open science

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

C. Catalogne, Eric Sauquet

► **To cite this version:**

C. Catalogne, Eric Sauquet. Valorisation des jaugeages épisodiques pour l'estimation du débit de référence d'étiage QMNA5. [Rapport de recherche] irstea. 2012, pp.49. hal-02599097

**HAL Id: hal-02599097**

**<https://hal.inrae.fr/hal-02599097v1>**

Submitted on 16 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Année de programmation 2012 – **Thème Hydro-morphologie** –

**Action 12** Prise en compte de chroniques incomplètes et de jaugeages volants dans la caractérisation des étiages



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

**Rapport final**

**Clotaire CATALOGNE (Irstea, Lyon)**  
**Eric SAUQUET (Irstea, Lyon)**

**Décembre 2012**

En partenariat avec :  
Laurence DURAND et Audrey FERNANDEZ (DREAL PACA, Le Tholonet)



- **AUTEURS**

**Clotaire CATALOGNE**, doctorant (Irstea)

**Eric SAUQUET**, chargé de recherche (Irstea), [eric.sauquet@irstea.fr](mailto:eric.sauquet@irstea.fr)

- **CORRESPONDANTS**

**Onema : Bénédicte AUGÉARD**, chargée de mission (Onema), [benedicte.augeard@onema.fr](mailto:benedicte.augeard@onema.fr)

- **AUTRES CONTRIBUTEURS**

**Droits d'usage** : accès libre

**Niveau géographique** : national

**Couverture géographique** : France

**Niveau de lecture** : professionnels, experts





- **RESUME**

Ce rapport s'attache à exposer 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. Cette approche vient ainsi en complément de méthodes mises en œuvre pour fournir une estimation en sites non jaugeés (Riffard *et al.*, 2012).

Le rapport contient tout d'abord des éléments techniques décrivant la méthode développée, les résultats obtenus et leur interprétation. La méthodologie retenue s'appuie sur l'ajustement d'une relation log-linéaire entre les débits jaugeés au site cible et les débits observés à la même date au droit d'une station hydrométrique voisine (un « site d'appui »). Cette relation peut alors être utilisée pour estimer le QMNA5 au site d'intérêt à partir du QMNA5 connu au site d'appui. L'interprétation des résultats, obtenus au travers d'une démarche de simulation de campagnes de jaugeage en validation croisée, s'est ensuite essentiellement orientée vers une analyse des performances de la méthode en fonction de la stratégie de jaugeage adoptée au site cible permettant ainsi de proposer quelques recommandations à destination des services opérationnels pour optimiser la programmation de futures campagnes de mesure.

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 s'atténue au-delà. Parallèlement il apparaît que la réalisation de plusieurs jaugeages au cours d'une même année (*i.e.* au cours d'une même saison d'étiage) conduit à un gain de performance moins rapide en raison, semble-t-il, de la dépendance intra-annuelle des débits jaugeés. Néanmoins les résultats indiquent qu'il est tout de même possible d'augmenter un peu la fréquence annuelle pour écourter la durée de suivi d'un site sans compromettre la performance de l'approche proposée. A cet égard la réalisation de trois jaugeages chaque année semble constituer le meilleur compromis.

L'application de la méthode est ensuite illustrée à l'aide d'un exemple concret d'application mené à partir de données de jaugeage disponibles sur le bassin de l'Asse, situé dans le département des Alpes de Haute-Provence en collaboration avec la Direction Régionale de l'Environnement, de l'Aménagement et du Logement (DREAL) Provence-Alpes-Cote d'Azur.

- **MOTS CLES (THEMATIQUE ET GEOGRAPHIQUE) : HYDROLOGIE, ETIAGE, JAUAGES VOLANTS, SITES PEU JAUAGES, PREDETERMINATION, MODELES D'INCERTITUDE**



- **USING SPOT GAUGING DATA TO ESTIMATE THE ANNUAL MINIMUM MONTHLY FLOW WITH A RETURN PERIOD OF 5 YEARS**

- **ABSTRACT**

A method using sporadic measurement data to estimate the annual minimum monthly flow with a return period of 5 years *QMNA5* is presented. This approach may complement the more traditional methods used to estimate low flow statistics at ungauged location (Riffard *et al.*, 2012). This report contains the suggested developments, results obtained through cross validation and a discussion of the main results.

The developed approach provides estimates of *QMNA5* at a poorly gauged site by multiplying *QMNA5* computed at a nearby gauging station by a scale coefficient *k*. The value of *k* is obtained through a log-linear regression fitted on synchronous flow measurements recorded at the poorly-gauged site and at the nearby station. A sensitivity analysis of the data collection strategy (including sensitivity to the number of spot gauging data) is carried out, allowing useful recommendations for operational service in charge of river flow monitoring.

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, performing several measurements the same year (*i.e.* the same low-flow period) does not significantly improve estimates, probably because of the intra-annual dependence of low-flow data. Even so, it remains possible to slightly increase the measurement frequency in order to shorten the duration of the gauging period without jeopardizing the model performance. Three measurements per year seem to be a good trade-off.

The method is illustrated by a case study in the French Département of Alpes de Haute-Provence. This application is performed in close collaboration with DREAL Provence-Alpes-Cote d'Azur.

- **KEY WORDS (THEMATIC AND GEOGRAPHICAL AREA): HYDROLOGY, LOW FLOW, SPOT GAUGING, POORLY GAUGED STATION, PREDETERMINATION, UNCERTAINTY ANALYSIS**



Valorisation des jaugeages  
épisodiques pour l'estimation du  
débit de référence d'étiage *QMNA5*  
Rapport final  
C. CATALOGNE, E. SAUQUET





• **SOMMAIRE**

1	Introduction .....	1
2	Présentation de la méthodologie .....	2
2.1	Présentation du jeu de données de validation croisée .....	2
2.2	Méthode d'estimation .....	3
2.3	Stratégies de jaugeage simulées .....	5
2.4	Choix d'un site de référence .....	8
2.5	Quantification des incertitudes .....	10
3	Résultats de la validation croisée et modélisation de l'incertitude .....	12
3.1	Evolution et modélisation de l'écart-type en fonction de $r$ , $N$ et $F$ .....	12
3.2	Evolution et modélisation du biais d'estimation en fonction de $r$ , $N$ , $F$ .....	19
3.3	Comparaison avec les performances d'estimation du <i>QMNA5</i> en sites non jaugés .....	21
4	Exemple d'application au bassin versant de l'Asse .....	23
4.1	Présentation du bassin versant de l'Asse .....	23
4.2	Analyse des données disponibles.....	25
4.3	Discussion des résultats .....	33
5	Conclusion .....	35
6	Bibliographie .....	36
7	Table des illustrations .....	38



## 1 Introduction

Cette note a pour objectif de présenter dans les grandes lignes la méthodologie développée au sein de l'Unité de Recherche Hydrologie-Hydraulique du centre Irstea de Lyon pour permettre l'estimation du débit de référence d'étiage *QMNA5* (débit mensuel minimum de période de retour 5 ans) en valorisant les données de débits issues de campagnes de jaugeage épisodique en un point du réseau hydrographique.

Inspirée de l'approche ébauchée par Chopart et Sauquet (2008) sur le secteur Seine-Normandie, la méthodologie proposée s'appuie sur une démarche de validation croisée menée sur le plan national à partir d'un échantillon de 632 stations hydrométriques pour définir les modalités optimales d'utilisation des jaugeages. Celle-ci a également permis, par la répétition d'un grand nombre de simulations, de quantifier et modéliser l'incertitude associée à chaque estimation, en fonction notamment de la stratégie de jaugeage mise en œuvre au site cible.

Les éléments présentés dans ce document doivent ainsi permettre de :

- fournir une estimation du *QMNA5* à partir des données de jaugeage disponibles au site cible en sélectionnant le site de référence le plus approprié ;
- quantifier l'incertitude associée à l'estimation obtenue ;
- déterminer éventuellement l'intérêt à poursuivre la campagne de jaugeage pour affiner l'estimation.

Pour ce faire les deux premières parties seront consacrées à l'exposé de la méthode et des principaux résultats obtenus lors de la validation croisée. Elles seront suivies d'une partie destinée à illustrer l'approche proposée au travers d'un exemple détaillé d'application au bassin versant de l'Asse. Cet exemple permettra notamment de pointer quelques problèmes susceptibles d'être rencontrés dans le choix des données de jaugeage mais aussi de pointer certaines limites de la méthode proposée.

## 2 Présentation de la méthodologie

Cette première partie est destinée à présenter les éléments de méthodologie mis en œuvre au cours de nos travaux. **Nous avons souhaité orienter ces divers éléments de manière à répondre à deux problématiques principales se rapportant (i) à l'influence de la stratégie de jaugeage mise en œuvre au site cible et (ii) au choix d'un site de référence pour réaliser l'estimation du *QMNA5*.**

Pour ce faire **la démarche employée repose sur une procédure de validation croisée** consistant à simuler différentes configurations de jaugeages à partir de tirages conditionnés dans les chroniques de débits journaliers d'un jeu étendu de 632 stations hydrométriques de référence identifiées sur le territoire français. A l'issue de la phase de validation croisée, appliquée tour à tour sur chacun des sites de référence, la comparaison des valeurs de *QMNA5* estimées et observées a permis de quantifier et confronter les incertitudes propres aux différentes simulations réalisées.

Finalement, ces résultats ont été résumés au travers de modèles analytiques simples pouvant être exploités dans un cadre opérationnel pour reconstituer l'incertitude associée à chaque estimation (sous la forme d'un intervalle de confiance).

### 2.1 Présentation du jeu de données de validation croisée

Le jeu de données employé pour réaliser cette étude comprend un total de 632 bassins jaugés répartis sur l'ensemble du territoire français et pour lesquels sont disponibles des séries chronologiques de débits moyens journaliers d'au moins 26 années sur la période 1970-2005, considérées comme peu ou pas influencées et de bonne qualité vis-à-vis de la mesure en basses eaux (Catalogne, 2012).

L'échantillon de bassins sélectionnés atteint une densité d'environ 1 pour 1000 km<sup>2</sup> et couvre un peu plus de 35 % du territoire (sans décompter les bassins emboîtés). La distribution spatiale des bassins n'est cependant pas uniforme, avec en particulier trois secteurs mal représentés dans le Nord, le bassin de la Charente et la Corse. L'échantillon demeure néanmoins suffisamment représentatif des grands contextes hydro-climatiques pouvant exister à l'échelle nationale. La gamme de surfaces drainées par les bassins de l'échantillon varie entre 4,5 et 8800 km<sup>2</sup>, néanmoins la majorité d'entre eux possède une superficie comprise entre 100 et 500 km<sup>2</sup>.

Les valeurs de *QMNA5* extraites des séries chronologiques, lorsqu'elles sont exprimées sous la forme de débits spécifiques, varient entre 0 et 14,7 l/s/km<sup>2</sup> (Figure 1). Les valeurs les plus faibles, incluant 22 bassins avec un *QMNA5* nul, sont principalement obtenues pour les bassins sur 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 souterraines de la part des aquifères sont importantes.

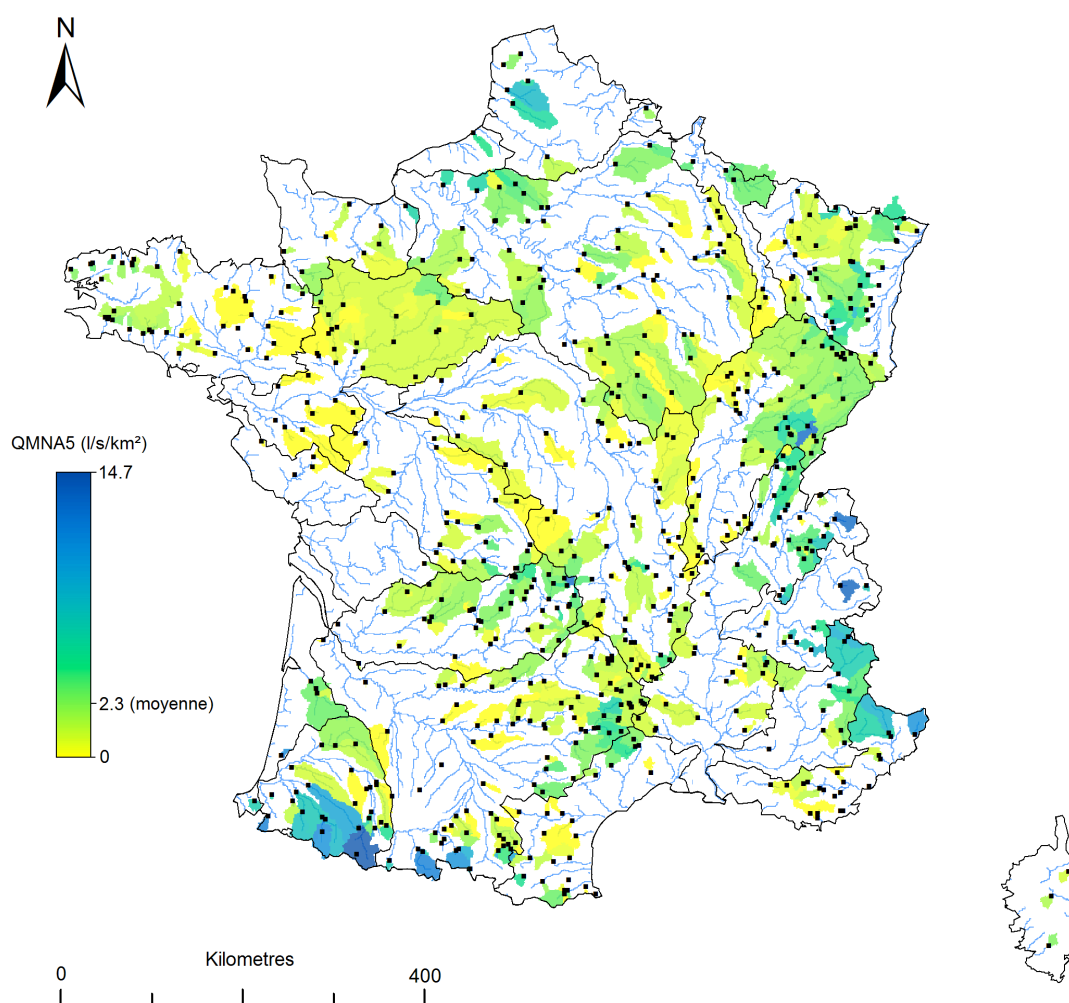


Figure 1 : Cartographie des QMNA5 spécifiques (en l/s/km<sup>2</sup>) aux stations de l'échantillon de référence © Clotaire Catalogne

## 2.2 Méthode d'estimation

La méthode d'estimation employée au cours de cette étude trouve en France son origine dans les travaux de Oberlin *et al.* (1973) qui les premiers ont reconnu l'intérêt des données de jaugeage pour l'estimation de variables d'étiage et jeté les bases d'une méthodologie permettant de les exploiter. La technique a depuis fait l'objet de plusieurs applications pratiques (e.g. DIREN Rhône-Alpes (2003), Galéa et Canali (2005)) parmi lesquelles on signalera les développements méthodologiques apportés par Chopart et Sauquet (2008) lors d'une application dédiée à l'estimation du QMNA5 dans le secteur Seine-Normandie. Bien que différents sous certains aspects, on trouve également des travaux comparables à l'étranger, en particulier aux Etats-Unis où la méthode visant la valorisation de jaugeages épisodiques a été initiée par Riggs (1972) puis perfectionnée par Stedinger et Thomas (1985), Reilly et Kroll (2003) et enfin Zhang et Kroll (2007) (méthode désignée sous l'appellation de « baseflow correlation method »).

D'une manière générale cette méthode s'appuie sur l'identification d'une **relation simple entre les débits jaugés au site cible et ceux observés à la même date à une station hydrométrique**

**voisine** sélectionnée à cet effet et que l'on désignera par la suite sous l'appellation de « site d'appui » ou « site de référence » (les modalités de sélection des sites d'appui sont exposées en détail dans la section 2.4). Il est alors fait l'hypothèse que la relation ajustée sur les débits concomitants constitue une bonne **approximation de la relation entre la valeur de QMNA5 recherchée au site cible et celle connue au site d'appui.**

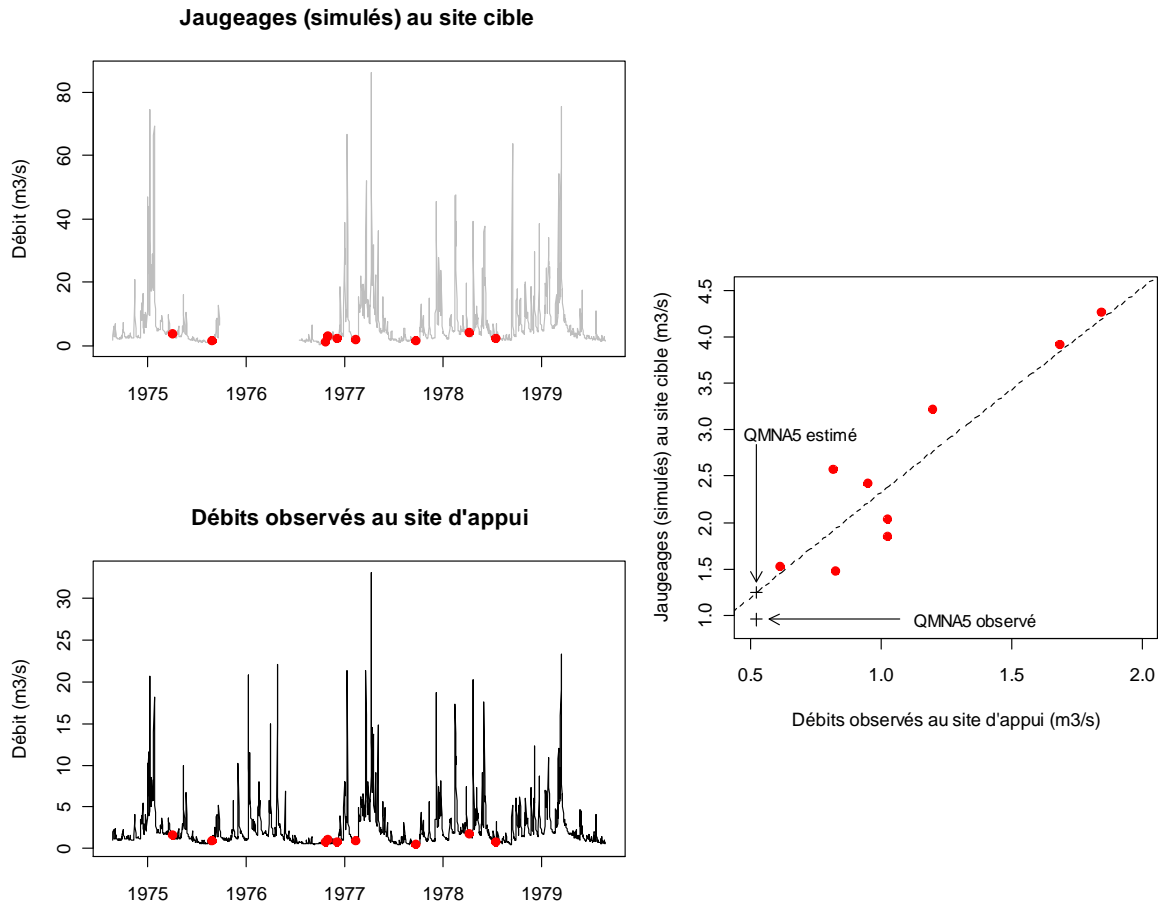


Figure 2 : Exemple de mise en relation entre débits jaugés au site cible (9 jaugeages simulés ici à partir d'une chronique) et débits observés à la même date sur un bassin de référence voisin © Clotaire Catalogne

Dans le cadre de nos travaux, trois types de formulation différents, inspirés des travaux relevés dans la littérature, ont été testés pour établir cette relation. La confrontation des résultats obtenus au terme de la validation croisée a permis de retenir une relation de forme log-linéaire comme suggéré dans les travaux américains :

$$q_{cib}(t) = L \cdot q_{réf}(t)^k \quad \text{ou de manière équivalente} \quad \ln(q_{cib}(t)) = \lambda + k \cdot \ln(q_{réf}(t))$$

où  $q_{cib}(t)$  sont les débits jaugés au site cible,  $q_{réf}(t)$  sont les débits observés à la même date au site d'appui,  $\lambda = \ln(L)$  et  $k$  sont les paramètres de la relation que l'on obtiendra par ajustement d'une régression linéaire entre les logarithmes des débits jaugés au site cible  $q_{cib}(t)$  et ceux observés

au site d'appui à la même date  $q_{réf}(t)$  (Figure 2)<sup>1</sup>.

Un des avantages de cette relation (par rapport à un simple modèle linéaire) est de donner intrinsèquement un poids plus important aux valeurs jaugées les plus basses. Elle **ne permet pas en revanche d'exploiter (ni d'estimer) des débits nuls**, valeurs que nous avons par défaut remplacées dans cette application par 1 l/s (seuil de précision des données de débits fournies par la Banque HYDRO). Enfin, il convient de préciser que si cet ajustement requiert un minimum théorique de deux mesures, **il nous paraît hasardeux d'employer moins de quatre couples de points**.

Les paramètres ainsi obtenus sont ensuite exploités pour estimer le  $QMNA5_{cib}$  (ou tout autre statistique de débit d'étiage) au site cible à partir de la valeur de  $QMNA5_{réf}$  connue au site de référence :

$$QMNA5_{cib} = L(QMNA5_{réf})^k$$

### 2.3 Stratégies de jaugeage simulées

D'un point de vue pratique, la définition d'une stratégie de jaugeage est un point essentiel qui doit permettre d'aboutir à une estimation fiable avec un minimum de données (de manière à minimiser le « coût opérationnel ») et/ou dans un délai raisonnable. Pour répondre à cette problématique nous nous sommes proposé d'examiner, au travers de différentes simulations :

- 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 ? sur combien de temps ?) ;
- l'influence de la fréquence annuelle des jaugeages (pour un même nombre final de jaugeages, peut-on écourter la période de suivi (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 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. Seuls Reilly et Kroll (2003) illustrent l'importance de travailler à l'aide de données indépendantes en testant des échantillonnages plus ou moins bien répartis dans le temps (jaugeages réalisés au cours d'une même période de récession, jaugeages répartis sur plusieurs périodes de récession successives, jaugeages réalisés au cours d'une unique période de récession chaque année).

A ces considérations viennent bien entendu s'ajouter des contraintes opérationnelles. Sur ce point, l'examen de bases de données mises à notre disposition sur les secteurs Seine-Normandie et Rhône-Alpes a permis de mieux connaître les pratiques de jaugeage adoptées dans ces deux régions. Il est ainsi apparu que, si les situations rencontrées d'un point de suivi à l'autre étaient très diversifiées (tant en termes de nombre de mesures réalisées qu'en termes de durée de la campagne de suivi), **la principale contrainte reposait sur la fréquence annuelle de jaugeage qui, sur le long terme, ne**

---

<sup>1</sup> Bien que cela n'ait pas été examiné dans le cadre de ces travaux, on signalera la possibilité d'introduire l'incertitude associée à la mesure de débit lors de l'ajustement en employant un critère de moindre carré pondéré.

**dépasse que rarement 4 à 5 mesures en moyenne chaque année.**

Parallèlement et étant donnés les objectifs propres à l'estimation de débits caractéristiques d'étiage, l'ensemble des travaux relevés dans la littérature s'accorde sur l'importance de réaliser des **mesures représentatives du fonctionnement de basses eaux**. Ainsi les études effectuées aux Etats-Unis se basent exclusivement sur l'échantillonnage au cours de périodes de récession (quelle que soit la période de l'année) tandis que les études réalisées en France préfèrent un échantillonnage réparti tout au long de la saison de basses eaux (Chopart et Sauquet (2008) démontrent en particulier le gain à procéder ainsi par rapport à des jaugeages réalisés tout au long de l'année). Dans cette application nous avons opté pour la seconde solution, sans doute plus simple à mettre en œuvre du point de vue opérationnel car moins contraignante, **en définissant la saison d'étiage comme la période correspondant au trimestre le plus sec dans l'année**. En pratique, s'il est possible de présumer d'une période favorable en fonction du contexte hydro-climatique du bassin cible (la majorité des cours d'eau français présentent une période d'étiage estivale entre les mois de juillet et de septembre), on pourra tout de même s'aider de la cartographie des régimes hydrologiques mensuels établie par Sauquet *et al.* (2008) en tout point du réseau hydrographique français (Figure 3).

**Enfin pour assurer un minimum d'indépendance entre les valeurs obtenues nous avons souhaité introduire une contrainte sur la durée minimale séparant deux mesures.** Celle-ci a été fixée par défaut à 15 jours, temps au bout duquel la décorrélation du signal de débit s'avère généralement suffisante pour obtenir deux mesures pas ou faiblement dépendantes (d'après les tests effectués 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 stations de référence). Ceci laisse alors la possibilité de réaliser jusqu'à 6 jaugeages au cours de la période d'étiage de trois mois.

On précisera que ces règles ont été en partie définies pour faciliter l'automatisation des simulations au cours de la validation croisée. Elles pourraient bien entendu être assouplies en pratique ; par exemple si la saison d'étiage se prolonge au-delà des trois mois suggérés ici ou si un épisode pluvieux vient interrompre une période d'étiage dans un délai inférieur aux 15 jours préconisés pour assurer l'indépendance des valeurs de débits jaugés.

Compte tenu de tous ces éléments et étant donné que la simulation des stratégies de jaugeage en validation croisée ne peut pas reproduire de manière exhaustive la grande diversité des situations rencontrées en pratique, il a été choisi de se limiter au cours de la validation croisée à un jeu de simulations faisant intervenir différentes combinaisons de durées de suivi en année (notées  $D$ ) et de fréquence de jaugeage annuelle (*i.e.* nombre de jaugeages réalisés durant la période d'étiage chaque année, noté  $F$ ) pour un nombre de jaugeages résultant donné par  $N = D \times F$  (Tableau 1).



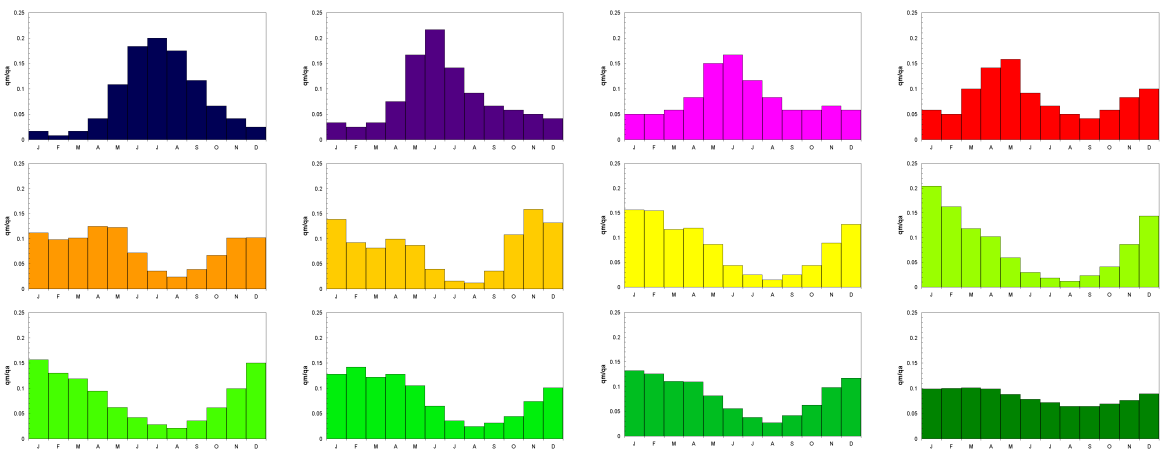
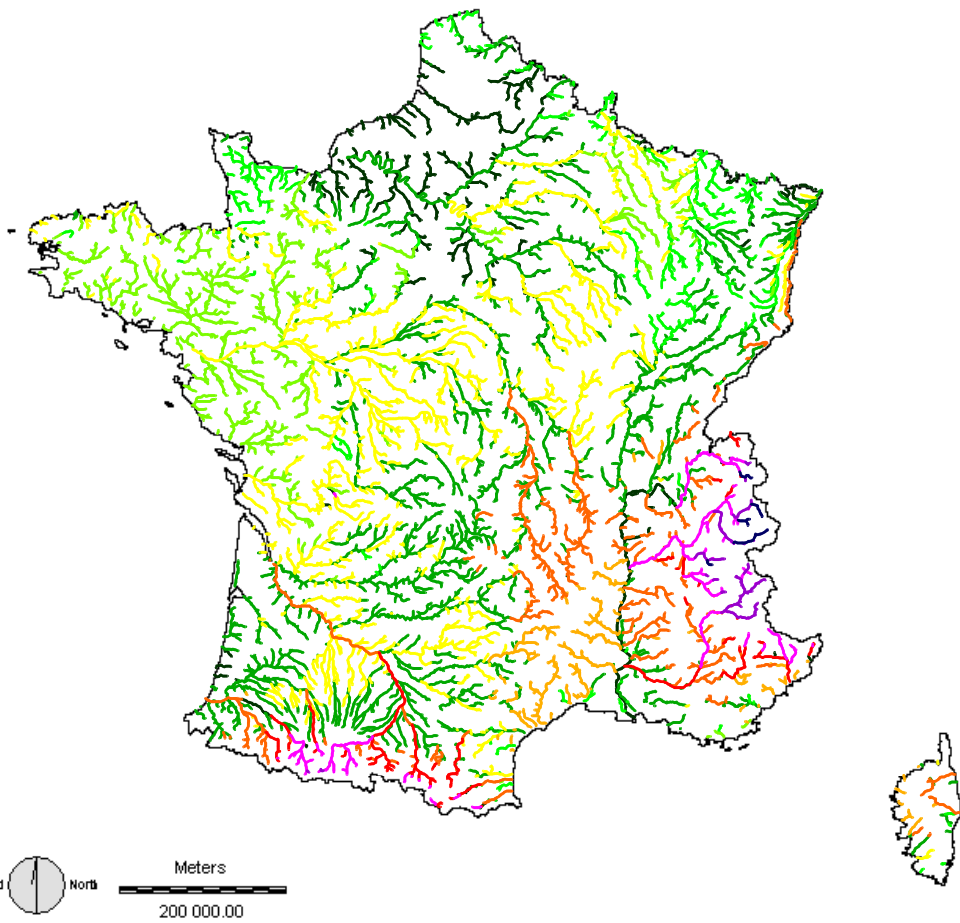


Figure 3 Carte des régimes hydrologiques le long du réseau hydrographique selon la typologie définie par les douze séries de coefficients mensuels de débit © Eric Sauquet

Tableau 1 : Combinaisons de durée de suivi (D en année) et de fréquence de jaugeage annuelle (F) simulées en validation croisée et nombre total de jaugeages résultant

(N) - Source des données : Catalogne (2012)

	D = 2	D = 4	D = 8	D = 16
F = 1	Non simulé	4	8	16
F = 2	4	8	16	32
F = 3	6	12	24	48
F = 4	8	16	32	64
F = 5	10	20	40	80

## 2.4 Choix d'un site de référence

Préalablement à la sélection d'un site d'appui pour réaliser l'estimation proprement dite, il convient de rappeler que les stations hydrométriques employées à cet effet doivent avant tout répondre à certaines contraintes élémentaires pour permettre la détermination fiable d'un *QMNA5* de référence. Ainsi elles doivent en premier lieu disposer de **données de débit fiables** (en particulier dans la gamme des débits d'étiage) et **peu ou pas influencés** de manière à fournir une valeur de *QMNA5* représentative d'un comportement naturel (on exclura toutes les stations soumises à l'influence de retenues, barrages, dérivations et/ou prélèvements ayant un impact significatif sur le régime hydrologique). Parallèlement la détermination d'un *QMNA5* statistiquement représentatif nécessite de disposer d'une **chronique suffisamment longue, comprenant au moins 26 années complètes (i.e. sans lacune) de données au pas de temps journalier** (cette recommandation est issue d'une étude de sensibilité menée par ailleurs). On remarquera que d'un point de vue pratique **les données au pas de temps journalier sont ici assimilées par défaut à des données de débits instantanés** (comme c'est le cas des données de jaugeage recueillies au site cible). Cette hypothèse, difficilement acceptable pour des périodes présentant de brusques variations du débit (e.g. en période de crue), a été jugée recevable en périodes de basses eaux au cours desquelles le débit varie peu au cours d'une même journée et durant lesquelles il est recommandé de pratiquer les jaugeages (cf. section 2.3).

Au delà de cette étape préliminaire de critique des données, la problématique de sélection d'un site d'appui constitue une étape clé qui vise, comme dans toute approche de régionalisation, à identifier le bassin de référence le plus approprié sur la base de similitudes comportementales avec le bassin cible. Ainsi dans le cadre de la valorisation des données de jaugeage, il est avant tout important de s'assurer que le bassin de référence sélectionné réagit aux mêmes événements météorologiques que ceux éprouvés par le bassin cible. C'est pourquoi la plupart des travaux relevés dans la littérature impose 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 que le site cible (Laaha et Blöschl, 2005). 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), toutefois la

solution la plus intéressante de ce point de vue consiste probablement à caractériser le degré de similitude entre bassin sur la base de la corrélation entre les couples de débits mesurés au site cible  $q_{cib}(t)$  et observés au site de référence  $q_{réf}(t)$  à la même date (Stedinger et Thomas (1985) préconisent par exemple de ne retenir que les sites de référence dont la corrélation dépasse 0.7). En effet l'emploi de la corrélation présente l'avantage de disposer d'un indicateur directement issu de l'information hydrologique disponible au site cible et pouvant se substituer de manière relativement « universelle » (*i.e.* quel que soit le contexte hydro-climatiques du site cible) à l'emploi de descripteurs de bassin dont la disponibilité et la pertinence ne sont pas toujours assurées.

Dans la continuité de ces différentes approches mais aussi dans un souci d'automatisation de la procédure, nous proposons dans cette application de **pré-sélectionner tout d'abord pour chaque site cible un ensemble de 50 sites de référence sur la base de la proximité géographique** définie par le calcul de la distance entre bassins suggérée par Ghosh (1951) :

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

où  $d$  est la distance entre bassins,  $S_{ref}$  et  $S_{cib}$  sont 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'emploi de cette distance possède l'avantage de permettre intrinsèquement une meilleure prise en compte de la topologie particulière du réseau hydrographique que la distance euclidienne entre centres de gravité des bassins traditionnellement utilisée (Gottschalk (1993)) et s'avère en particulier plus approprié pour quantifier la distance entre bassins emboîtés (Figure 4). A défaut, c'est sur la base de la distance entre centres de gravité que pourrait se faire la pré-sélection.

**Pour chacun des bassins de référence sélectionnés sont ensuite calculées les corrélations entre  $q_{cib}(t)$  et  $q_{réf}(t)$  (notées  $r$  dans la suite de ce document) pour finalement sélectionner le site d'appui présentant le plus fort degré de corrélation** (*i.e.* parmi les 50 sites de référence sélectionnés, on privilégiera celui pour lequel la corrélation entre  $q_{cib}(t)$  et  $q_{réf}(t)$  est élevée au dépens de sites géographiquement proches mais susceptibles de présenter un degré de corrélation plus faible).

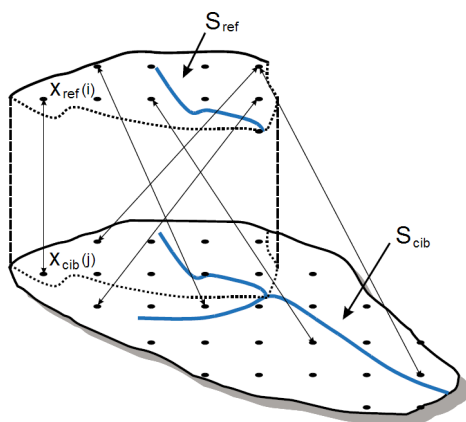


Figure 4 : Schéma illustrant la procédure de calcul de distance  $d$  pour deux bassins emboîtés. Modifié d'après Skoien et al. (2006) © Clotaire Catalogne

Selon ce protocole, le choix d'un site de référence est donc réalisé *a posteriori* à partir d'un échantillon préalable de sites géographiquement voisins<sup>2</sup> et pourra par conséquent évoluer avec l'augmentation du nombre de jaugeages réalisés. Cette évolution est à mettre en relation directe avec la significativité de la corrélation en fonction du nombre de jaugeages réalisés. En effet, la forte incertitude associée à la corrélation lorsque le nombre de jaugeages est faible se traduit par une incertitude sur le choix du site de référence le plus pertinent. S'il reste logique de privilégier par défaut le site de référence présentant le plus fort niveau de corrélation, ce choix ne sera plus forcément le plus pertinent avec l'ajout de nouveaux jaugeages et pourra donc évoluer vers la sélection d'un autre site plus approprié.

## 2.5 Quantification des incertitudes

Sur la base des éléments présentés précédemment la caractérisation des incertitudes obtenues en validation croisée a été réalisée de manière à répondre à deux objectifs principaux consistant à déterminer (et si possible modéliser) l'évolution de l'incertitude d'estimation du  $QMNA5$  en fonction :

- de la stratégie de jaugeage adoptée au site cible (*i.e.* en fonction de la durée de suivi ( $D$ ), de la fréquence annuelle de jaugeages ( $F$ ) et du nombre de jaugeages résultant ( $N$ );
- de la « proximité » du site de référence (quantifiée au travers de la corrélation  $r$  entre  $q_{cib}(t)$  et  $q_{réf}(t)$ ).

Pour ce faire le protocole retenu a consisté à considérer tour à tour chacune des 632 stations de notre échantillon de référence comme un site cible puis à fournir une estimation à l'aide de la relation présentée en section 2.1 en (*i*) simulant les différentes stratégies de jaugeages décrites en section 2.3 par échantillonnage dans la chronique de débit disponible et (*ii*) en s'appuyant successivement sur un ensemble de sites de référence voisins présentant un degré de corrélation plus ou moins élevé avec le site cible. Signalons que ces différentes opérations ont été répétées 100 fois pour chaque site cible afin d'obtenir une évaluation suffisamment robuste de l'incertitude (*i.e.* indépendante de l'échantillonnage) au terme de la validation croisée.

Pour chaque estimation fournie aux différents sites cibles, l'erreur commise a été quantifiée au moyen de l'écart entre les valeurs de  $QMNA5_{cib}$  prédites et  $QMNA5_{obs}$  observées, préalablement transformées en débit spécifique pour autoriser la mise en commun des résultats de tous les bassins examinés (les incertitudes sont ainsi exprimées en  $l/s/km^2$ ) :

$$QMNA5_{obs} = QMNA5_{cib} + err = L(QMNA5_{réf})^k + err$$

où les paramètres  $L$  et  $k$  sont estimés selon le protocole décrit en section 2.2 et l'erreur d'estimation  $err$  sera décrite dans la suite du document dans un sens statistique par sa moyenne (notée *bias*) et son écart-type (noté  $\sigma$ ).

L'ensemble des résultats obtenus (*i.e.* toutes stations confondues) a ensuite été rassemblé par classe de corrélation ( $r$ ) permettant d'obtenir, pour chacune des stratégies de jaugeages examinées, une représentation des distributions des erreurs en fonction de la « proximité » au site cible telle qu'illustrée

---

<sup>2</sup> Dans une application pratique, on pourra éventuellement affiner cet échantillon sur la base d'une expertise plus approfondie des contextes hydro-climatiques des bassins considérés.

sur la Figure 5 (sur cette figure, les distributions d'erreur sont représentées sous forme de « boîtes à moustache » dont les différents éléments indiquent la valeur médiane (trait gras), les premiers et troisièmes quartiles (boîte) et les bornes entre lesquelles sont incluses environ 95 % des valeurs (moustaches, correspondant à la médiane +/- 1.58 fois la distance interquartile)).

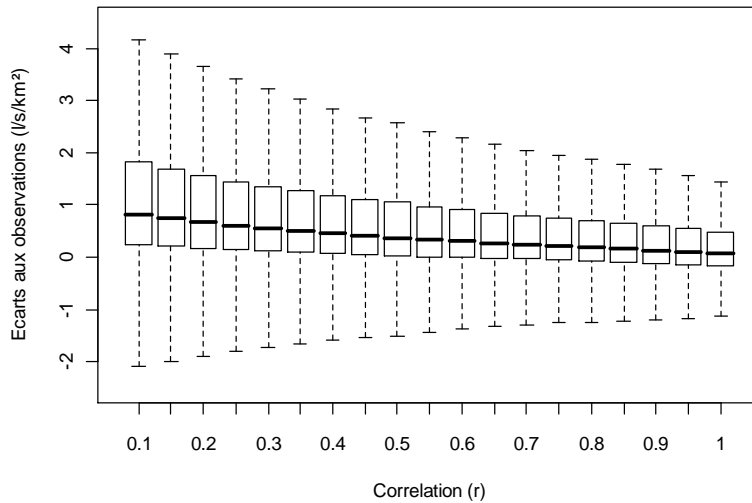


Figure 5 : Exemple de distributions d'erreur d'estimation du QMNA5 par classe de corrélation, obtenues pour une stratégie de jaugeage simulant une fréquence annuelle de jaugeage  $F = 2$  et une durée de suivi de  $D = 4$  années (soit au total  $N = 8$  jaugeages) © Clotaire Catalogne

Ces distributions, assez proches d'une distribution normale, ont enfin été résumées au moyen de leurs moyennes  $bias_{(r,N,F)}$  et de leur écart-type  $\sigma_{(r,N,F)}$  de manière à pouvoir construire à terme 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 ( $F$ ). Nous en donnons ici un exemple pour un niveau de confiance à 95 % :

$$P(Q_{est} + bias_{(r,N,F)} - 1.96 \cdot \sigma_{(r,N,F)} < Q_{obs} < Q_{est} + bias_{(r,N,F)} + 1.96 \cdot \sigma_{(r,N,F)}) \approx 0.95$$

L'expression des résultats en fonction du nombre de jaugeage (au détriment de la durée de suivi) est ici préférée de manière à bien faire ressortir en priorité le « coût opérationnel » (nombre de déplacements sur le terrain) impliqué par le choix d'une stratégie de jaugeage donnée. La transposition d'une grandeur à l'autre est cependant aisée (rappel :  $D = N/F$ ) et sera employée dans un second temps pour illustrer l'influence du couple fréquence de jaugeage / durée de suivi sur les performances d'estimation.

### 3 Résultats de la validation croisée et modélisation de l'incertitude

L'examen des résultats de la validation croisée en termes d'incertitude s'est avant tout concentré sur le critère d'écart type  $\sigma_{(r,N,F)}$  qui détermine seul la largeur de l'intervalle de confiance et, de ce fait, constitue le critère le plus pertinent pour évaluer ici l'évolution des performances en fonction de la stratégie de jaugeage adoptée au site cible et de la « proximité » du site de référence employé pour réaliser l'estimation. Le critère de  $bias_{(r,N,F)}$  sera traité dans un second temps de manière à permettre la correction des estimations et l'évaluation finale des bornes d'un intervalle de confiance de niveau de significativité voulu.

#### 3.1 Evolution et modélisation de l'écart-type en fonction de $r$ , $N$ et $F$

La Figure 6 illustre, indépendamment pour les cinq fréquences annuelles de jaugeage simulées, les valeurs de  $\sigma_{(r,N,F)}$  obtenues en validation croisée pour un nombre de jaugeages croissant et selon différentes classes de corrélation comprises ici entre 0,5 et 1 (gamme dans laquelle se trouvent plus de 99 % des valeurs de corrélation obtenues durant la validation croisée pour les sites de références sélectionnés selon le protocole décrit en section 2.4).

On constate sur cette figure que, pour une même fréquence de jaugeage, les courbes représentées s'avèrent globalement parallèles, indiquant que lorsque l'on double le nombre de jaugeages simulé (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 revanche, il apparaît visiblement que ce gain diminue progressivement avec l'augmentation du nombre de jaugeages. Quelques essais ont montré qu'il était possible de reproduire assez correctement cette évolution à l'aide d'une simple relation linéaire en fonction du logarithme du nombre de jaugeages simulé :

$$\sigma_{(r,N,F)} = a_{(r,F)} \cdot \ln(N) + b_{(r,F)} \quad (E1)$$

L'analyse des paramètres ajustés montre que  $a_{(r,F)}$  correspond comme prévu à une valeur négative (*i.e.* l'incertitude décroît avec l'augmentation du nombre de jaugeages) relativement constante quel que soit  $r$ . Ce résultat confirme indirectement que le gain de performance lorsque le nombre de jaugeage augmente (*i.e.* ici la différence entre les incertitudes lorsque le nombre de jaugeage double) ne dépend donc pas ou peu de la corrélation.

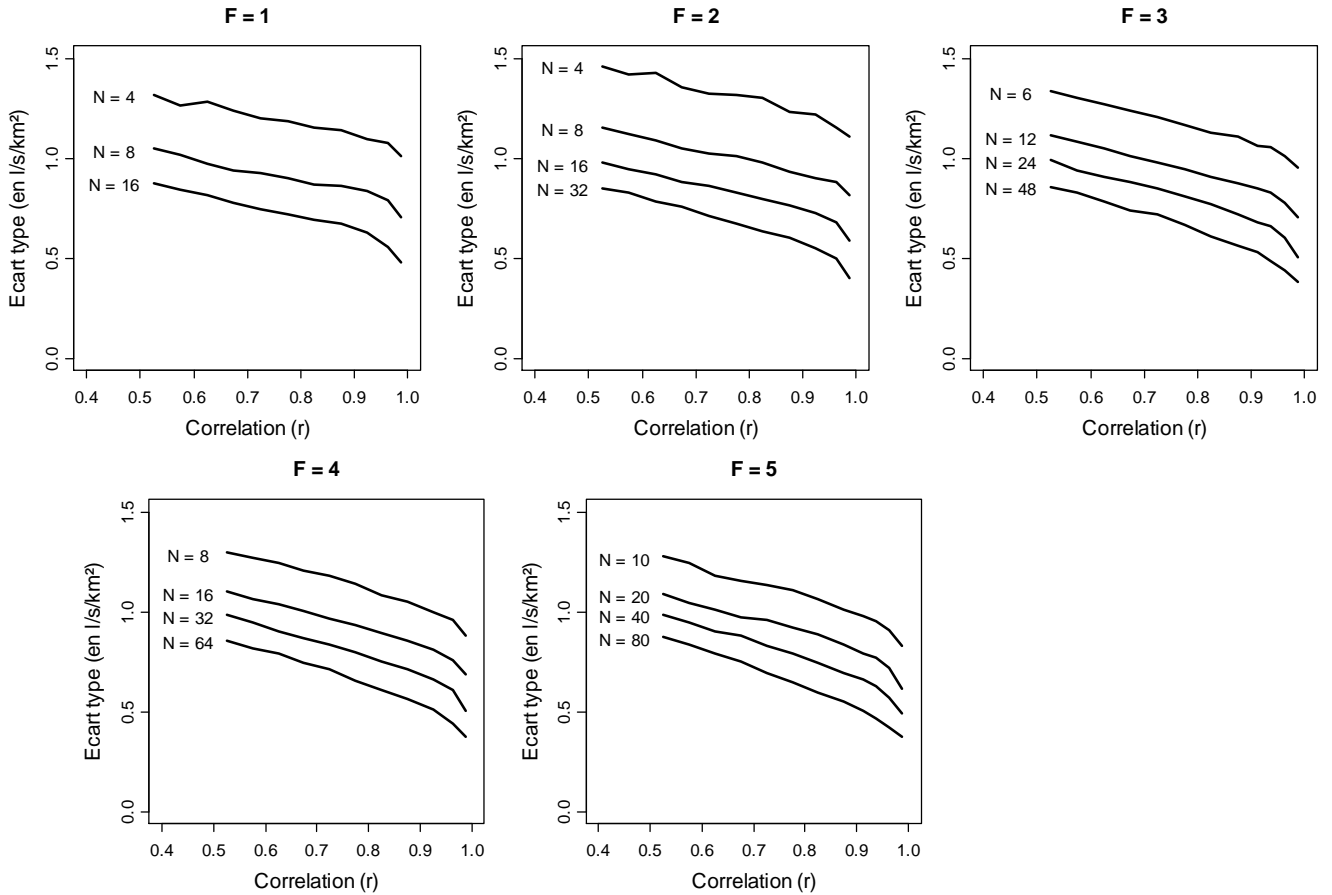


Figure 6 : Incertitudes d'estimation (écart type des erreurs d'estimation en l/s/km<sup>2</sup>) en fonction de la corrélation et du nombre de jaugeages pour les cinq fréquences de jaugeage simulées © Clotaire Catalogne

A l'inverse le paramètre  $b_{(r,F)}$  permet de rendre compte de l'évolution de l'incertitude en fonction du degré de corrélation et décroît lorsque  $r$  augmente. Il correspond en effet à la valeur d'incertitude théoriquement obtenue en extrapolation en ne réalisant qu'un unique jaugeage<sup>3</sup> et se réfère par conséquent à l'évolution de l'incertitude en fonction de la corrélation  $r$ , indépendamment du nombre de jaugeages réalisés ( $\ln(N=1) = 0$  d'où  $\sigma_{(r,N=1,F)} = b_{(r,F)}$ ). A l'examen des différents graphiques, il apparaît que cette évolution pourrait être à nouveau reproduite à l'aide d'une simple approximation linéaire, cependant pour pouvoir restituer correctement l'inflexion constatée entre  $r = 0.95$  et  $r = 1$ , une fonction puissance semble *a priori* plus appropriée. Les deux solutions ont été examinées et intégrées à l'équation *E1* pour donner deux nouvelles formulations analytiques de l'écart-type en fonction du nombre de jaugeages  $N$  et du degré de corrélation  $r$  :

$$\sigma_{(r,N,F)} = a_{(F)} \cdot \ln(N) + c_{(F)} \cdot (1-r) + d_{(F)} \quad (E2)$$

$$\sigma_{(r,N,F)} = a_{(F)} \cdot \ln(N) + c_{(F)} \cdot (1-r)^{d_{(F)}} \quad (E3)$$

Les résultats obtenus à partir de l'ajustement de ces deux modèles montrent que l'on parvient à des performances quasiment identiques (*i.e.* l'erreur absolue indique des écarts de l'ordre de  $4 \cdot 10^{-2}$  l/s/km<sup>2</sup>

3 Cette valeur reste théorique puisque l'ajustement de la relation entre  $q_{ref}(t)$  et  $q_{cib}(t)$  requiert un minimum théorique de 2 jaugeages.

en moyenne) et satisfaisantes avec des écarts aux valeurs observés pratiquement négligeables devant  $\sigma_{(r,N,F)}$ . La différence entre les deux modèles provient alors essentiellement de la manière dont sont répartis les écarts en fonction de la classe de corrélation qui s'avèrent plus équilibrés lorsque le modèle *E2* est employé (seules les incertitudes dans la gamme de corrélation  $r = 0.95$  à  $r = 1$  se distinguent par une sur-estimation plus importante et presque généralisée, ce qui reste néanmoins préférable à une sous-estimation). Ces observations nous ont amené à considérer le modèle *E2* comme le plus approprié bien qu'il soit peut-être analytiquement moins justifié que le modèle *E3*.

Les coefficients du modèle *E2*, ajustés indépendamment pour les cinq fréquences annuelles de jaugeages simulées, sont reportés dans le Tableau 2. Ils vont être exploités dans les paragraphes suivants pour mieux caractériser l'évolution des performances d'estimation du *QMNA5* en fonction de la stratégie de jaugeage.

Tableau 2 : Paramètres des cinq modèles d'incertitude ajustés pour chaque fréquence de jaugeage simulée - Source des données : Catalogne (2012)

$F$	$a_{(F)}$	$c_{(F)}$	$d_{(F)}$
1	-0.326	0.664	1.452
2	-0.290	0.770	1.468
3	-0.263	0.822	1.441
4	-0.248	0.849	1.446
5	-0.238	0.887	1.443

### 3.1.1 Influence du nombre de jaugeages

Dans un premier temps on se propose d'examiner ici, au travers des différents modèles d'incertitude, l'influence du nombre de jaugeages sur les performances d'estimation du *QMNA5* et, à partir de là, **déterminer l'intérêt à poursuivre une campagne de mesures pour une fréquence de jaugeage donnée**. Ainsi, il a été représenté sur la Figure 7, le gain théorique apporté par l'augmentation du nombre de jaugeages (volontairement poussé jusqu'à 100) indépendamment du niveau de corrélation :

$$gain = \sigma_{(r,N=i,F)} - \sigma_{(r,N=1,F)} = a_{(F)} \cdot \ln(N = i) - a_{(F)} \cdot \ln(N = 1)$$

On constate ici que **le gain de performance est assez sensible jusqu'à environ 20 ou 25 jaugeages quelle que soit la fréquence de jaugeage considérée** puis qu'il s'amortit au delà, d'autant plus rapidement que la fréquence de jaugeage est élevée. 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 être considéré comme une bonne indication sur la durée de suivi optimale d'un site selon la fréquence de jaugeage adoptée.

Ce résultat seul reste cependant insuffisant car, comme tend à le démontrer la Figure 7, **l'augmentation de la fréquence de jaugeage s'accompagne d'un gain de performance moins rapide pour un même nombre de jaugeage réalisé**. Il est donc 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.



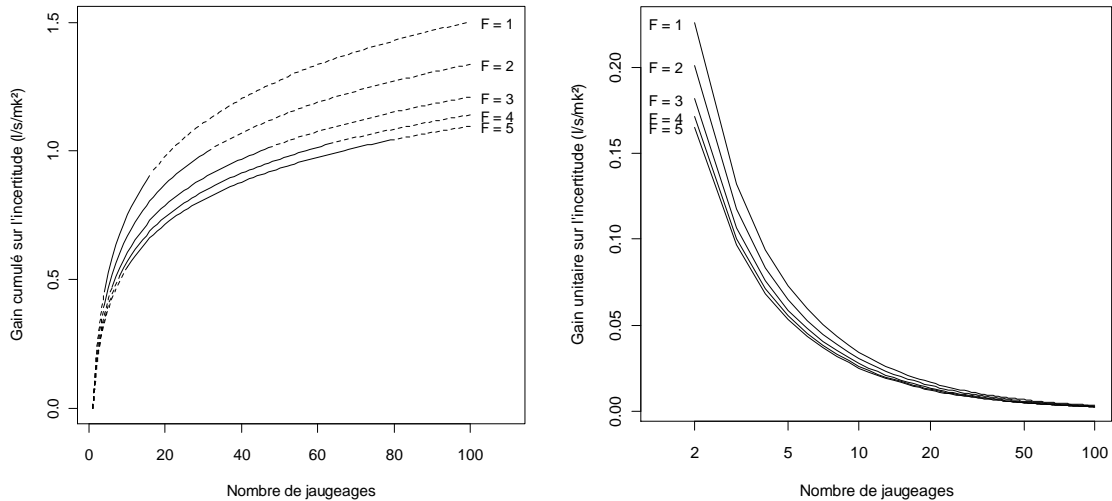


Figure 7 : Gains cumulés sur l'incertitude liés à l'augmentation du nombre de jaugeages pour les cinq fréquences de jaugeage simulées en supposant une corrélation constante. Les incertitudes fournies par les modèles pour des configurations non simulées (i.e. en extrapolation) sont représentées en pointillés (graphique de gauche). Gain par jaugeage complémentaire en fonction du nombre de jaugeages existant. © Clotilde Catalogne

Par ailleurs il n'est pas pris en compte ici la possible évolution de la corrélation (i.e. la mesure de proximité entre site cible et site d'appui) avec l'augmentation du nombre de jaugeages qui s'avère difficilement prévisible (cf. section 2.4). On peut néanmoins préciser, à partir de l'analyse des résultats de la validation croisée, que l'augmentation du nombre de jaugeages s'accompagne généralement d'une diminution du niveau de corrélation dans environ 75 % des simulations réalisées. Si cette perte s'avère le plus souvent largement insuffisante pour annuler le gain lié à l'augmentation du nombre de jaugeages, il en résulte tout de même que le gain « réel » (i.e. tenant compte du différentiel de corrélation) reste généralement inférieur au gain « théorique » tel que calculé ici en supposant la corrélation constante.

### 3.1.2 Influence de la fréquence de jaugeage

Comme évoqué précédemment, la Figure 7 permet à elle seule de montrer que l'augmentation de la fréquence de jaugeage s'accompagne d'un gain de performance moins rapide pour un même nombre de jaugeages réalisés, la différence étant d'autant plus marquée que le nombre de jaugeages final est élevé. Autrement dit l'intérêt à augmenter le nombre de mesures pour réduire l'incertitude est d'autant plus important que la fréquence de jaugeage est faible..

Une bonne illustration de ce résultat consiste à calculer, pour  $F > 1$ , le nombre de jaugeages équivalent (noté  $N_{eq}$ ) permettant de parvenir aux mêmes performances que celles obtenues pour  $F = 1$  et un nombre de jaugeages  $N$  donné. Ceci revient simplement à résoudre l'égalité suivante pour différentes valeurs de  $N$  :  $\sigma_{(r,N,F=1)} = \sigma_{(r,N_{eq},F>1)}$ .

La Figure 8 résume les résultats obtenus pour  $N$  allant jusqu'à 20 jaugeages et trois niveaux de corrélation distincts. Elle permet effectivement de constater que  $N_{eq}$  est toujours supérieur à  $N$  et croît (non linéairement) lorsque  $N$  augmente d'autant plus fortement que la corrélation est basse et que la fréquence de jaugeage est élevée. A titre d'exemple, il est ici obtenu un résultat équivalent en réalisant un total de 60 jaugeages pour  $F = 4$  et un total de 20 jaugeages pour  $F = 1$  pour un niveau de corrélation égal à  $r = 0.75$ . Il semble bien que l'hypothèse de l'existence d'une dépendance intra-

annuelle des valeurs jaugées (*i.e.* lorsque  $F > 1$ ) soit à l'origine de ce résultat et qu'elle soit d'autant plus fortement ressentie que le nombre de jaugages est élevé (autrement dit plus l'on jauge, plus le gain d'information par rapport à la réalisation d'un seul jaugage par année diminue).

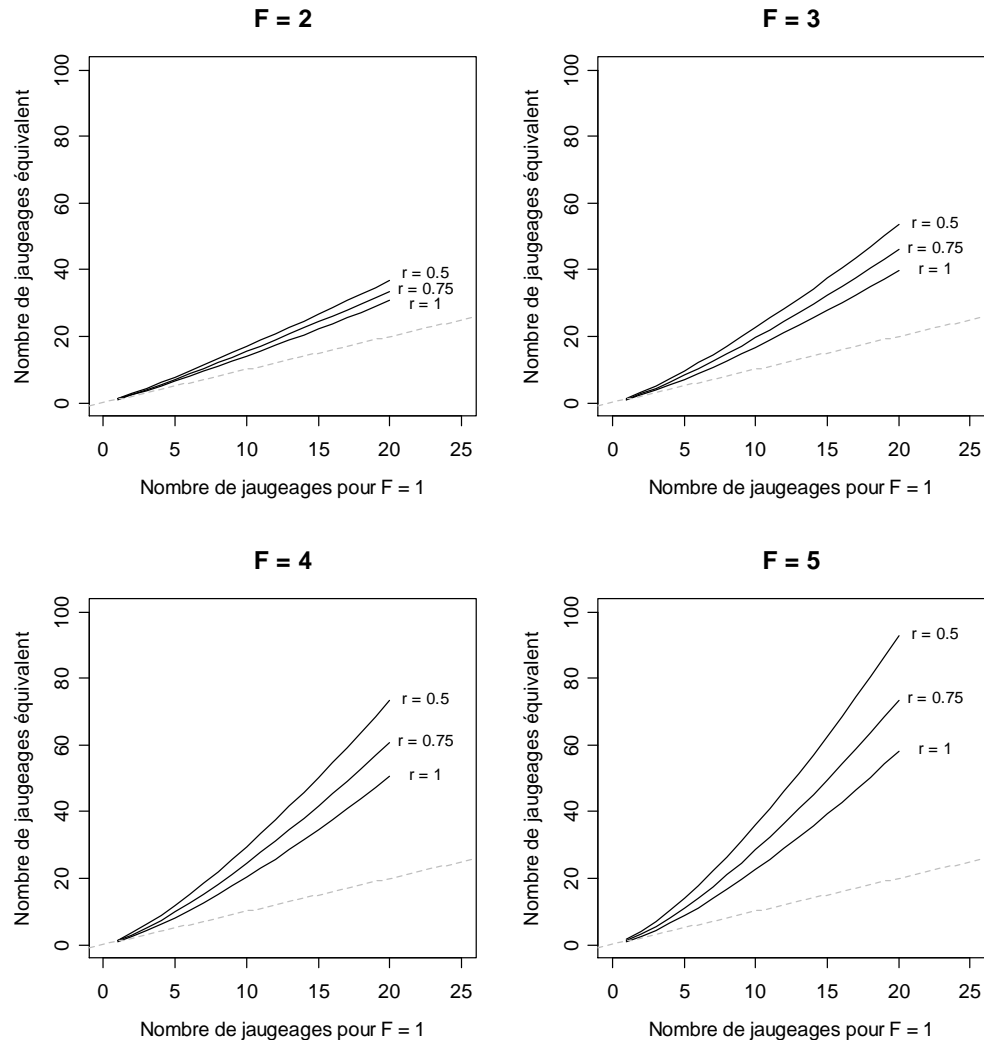


Figure 8 : Nombre de jaugages équivalent pour  $F > 1$  en fonction du nombre de jaugages réalisés pour  $F = 1$  © Clotilde Catalogne

Ainsi, si l'objectif est d'aboutir à la réalisation d'un total de 20 jaugages (*i.e.* si l'on donne priorité au moindre « coût opérationnel »), il sera préférable de répartir les mesures sur une durée de 20 ans (*i.e.*  $F = 1$ ) plutôt que sur 10 ( $F = 2$ ) ou 5 années ( $F = 4$ ). Toutefois cette conclusion cache une ambiguïté. En effet si l'on raisonne cette fois en termes de durées de suivi équivalentes ( $D_{eq} = N_{eq}/F$ ), il apparaît que des performances identiques peuvent être obtenues sur une durée plus courte en augmentant la fréquence de jaugage. Ce résultat est illustré sur la Figure 9. On constate bien que  $D_{eq}$  reste ici inférieure à  $D$ . Ainsi si l'objectif est d'aboutir à un résultat équivalent à celui obtenu pour 20 jaugages répartis sur 20 années, il est en réalité possible d'écourter la période de suivi à environ 17 ans pour une fréquence de jaugage  $F = 2$  (avec  $N_{eq} \approx 34$  jaugages) et à environ 15 années pour une fréquence de jaugage  $F = 4$  (avec  $N_{eq} \approx 60$  jaugages)<sup>4</sup>.

4 Il s'agit en réalité d'une fourchette de durées équivalentes dépendant du niveau de corrélation considéré.

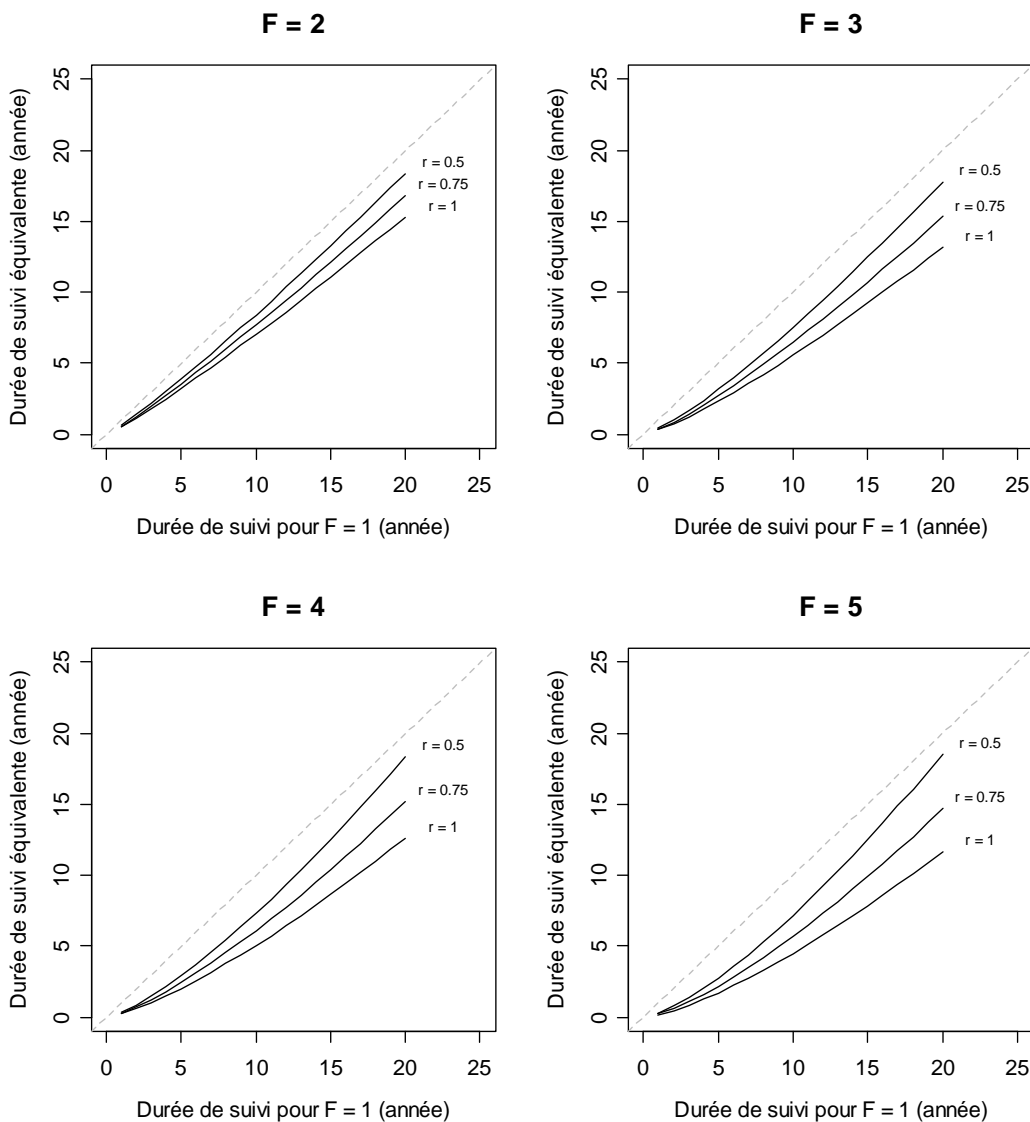


Figure 9 : Durée de suivi équivalente pour  $F > 1$  en fonction de la durée de suivi pour  $F = 1$  © Clotaire Catalogne

Compte tenu de ces observations **il apparaît donc avantageux d'augmenter la fréquence annuelle de jaugeage pour écourter (un peu) la durée de suivi d'un site (avec en contrepartie un accroissement du « coût opérationnel »)**. A cet égard, **la réalisation de trois jaugeages par année constitue sans doute le meilleur compromis**. Au delà (pour  $F = 4$  ou  $F = 5$ ) le gain supplémentaire en termes de durées de suivi s'avère pratiquement négligeable (inférieur à 1 an).

### 3.1.3 Cas d'une stratégie de jaugeage complexe

Jusqu'ici, l'analyse des modèles d'incertitude s'est limitée à des cas simples où la fréquence de jaugeages reste constante durant toute la campagne de suivi. **Une des principales difficultés tient à la possible irrégularité de la fréquence de jaugeage durant la période de suivi du site cible comme cela s'observe le plus souvent en pratique au cours des campagnes de jaugeage. Une solution consiste alors à traiter cette irrégularité comme une succession de périodes où la fréquence de jaugeage est constante puis à raisonner en termes de gain sur l'incertitude.**

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 au moyen de deux jaugeages par an sur les dix premières années ( $F_1 = 2$  et  $D_1 = 10$ ) puis au moyen de un jaugeage par an pendant les dix années restantes ( $F_2 = 1$  et  $D_2 = 10$ ). Le nombre total de jaugeages réalisés est donné par  $N = F_1 \cdot D_1 + F_2 \cdot D_2 = N_1 + N_2 = 30$  tandis que la fréquence de jaugeage moyenne est égale à  $F = N / D = 1,5$ .

Si le premier modèle d'incertitude permet de quantifier directement l'incertitude de l'estimation obtenue au bout des premières dix 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 :

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

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 constante ce gain serait théoriquement donné par :

$$\Delta\sigma_{(r_2, N_2, F_2)} = a_{(F_2=1)} \cdot (\ln(N_1 + N_2) - \ln(N_1)) + c_{(F_2=1)} \cdot (r_1 - r_2)$$

Or  $N_1$  correspond ici au nombre de jaugeages obtenu pour  $F_1 = 2$ , il est donc au préalable nécessaire de calculer le nombre de jaugeage équivalent à  $N_1$  en prenant  $F_1 = 1$ . Il en résulte que :

$$\Delta\sigma_{(r_2, N_2, F_2)} = a_{(F_2=1)} \cdot (\ln(N_{eq} + N_2) - \ln(N_{eq})) + c_{(F_2=1)} \cdot (r_1 - r_2)$$

En reprenant les caractéristiques de la campagne de jaugeage décrite plus haut, les résultats obtenus par cette méthode pour déterminer  $\sigma_{(r, N, F)}$  peuvent être illustrés à l'aide d'un graphique montrant l'évolution de l'incertitude en fonction du nombre de jaugeages (Figure 10, ici pour trois niveaux de corrélation distincts). Celui-ci permet bien de constater l'apparition d'une rupture dans l'évolution du gain de performance liée à la modification de la fréquence de jaugeage.

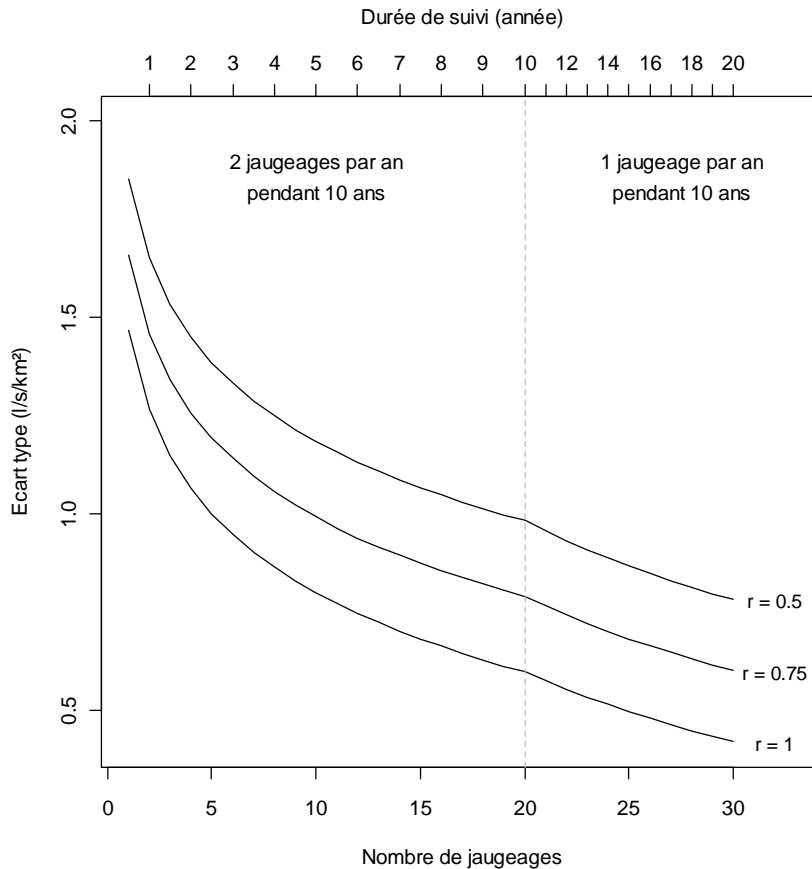


Figure 10 : Exemple d'évolution de l'incertitude en fonction du nombre de jaugeages dans le cas d'une modification de la fréquence de jaugeage au cours du suivi d'un site cible © Clotaire Catalogne

Bien entendu, cette démarche pourra être répétée autant de fois qu'il y a de modifications de la fréquence de jaugeage au cours de la période de suivi de manière à reproduire pratiquement n'importe quelle campagne de jaugeage (sous réserve que les fréquences de jaugeage mises en oeuvre correspondent à celles simulées).

### 3.2 Evolution et modélisation du biais d'estimation en fonction de $r$ , $N$ , $F$

Comme mentionné plus haut, l'objectif est de fournir une valeur approchée de l'intervalle de confiance de niveau de signification voulue pour chaque valeur de  $QMNA5$  estimée. Le modèle donnant l'écart type  $\sigma_{(r,N,F)}$  en fonction du degré de corrélation  $r$ , du nombre de jaugeages  $N$  et de la fréquence annuelle de jaugeage  $F$  doit donc être accompagné d'un modèle permettant de corriger préalablement le biais moyen affectant les estimations  $bias_{(r,N,F)}$ .

La Figure 11 illustre le biais d'estimation obtenu en fonction du niveau de corrélation pour les différentes fréquences annuelles de jaugeage simulées et pour un nombre de jaugeages croissant. On constate en premier lieu que ce biais est toujours positif et traduit une tendance à la sur-estimation du  $QMNA5$ .

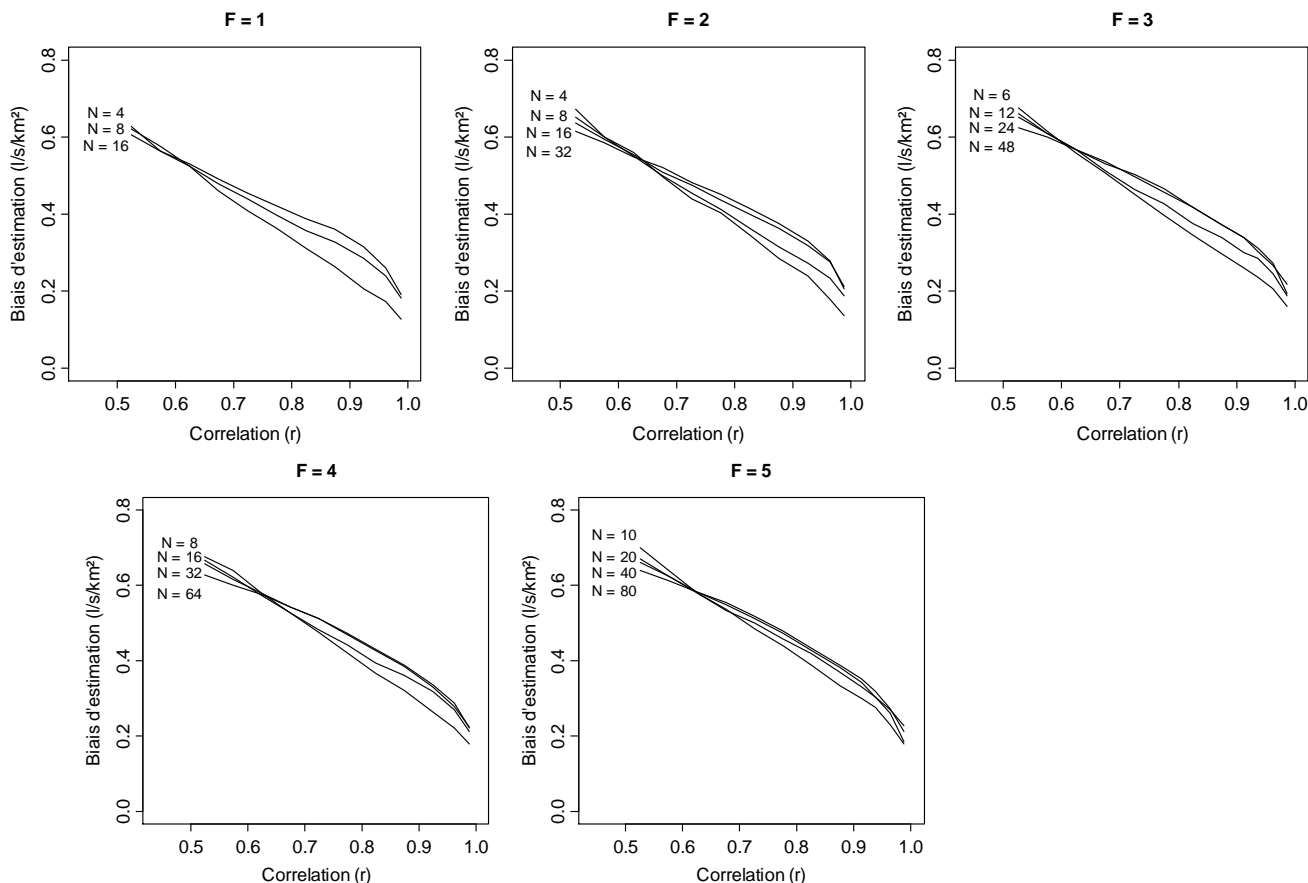


Figure 11 : Biais d'estimation en fonction de la corrélation et du nombre de jaugeages pour les cinq fréquences de jaugeage simulées © Clotaire Catalogne

Les graphiques montrent également que le biais est principalement conditionné par le niveau de corrélation, avec une évolution assez proche de celle observée précédemment pour  $\sigma_{(r,N,F)}$  (i.e. inflexion entre  $r=0.95$  et  $r=1$  suivie d'un accroissement assez constant jusqu'à  $r=0.5$ ), tandis que l'influence du nombre de jaugeages est assez faible. Compte tenu de ces observations il a été choisi, pour chaque fréquence de jaugeage simulée, de n'ajuster qu'une simple approximation linéaire ne dépendant que de la corrélation  $r$  :

$$bias_{(r,N,F)} = \alpha_{(F)} \cdot r + \beta_{(F)}$$

Les paramètres ajustés et les performances (quantifiées par la racine carrée de la moyenne des erreurs quadratiques, notée  $RMSE$ ) des cinq modèles en fonction de  $F$  sont reportés dans le Tableau 3. Les résultats obtenus permettent de constater que les ajustements réalisés sont très proches avec des pentes  $\alpha_{(F)}$  pratiquement identiques et des ordonnées à l'origine  $\beta_{(F)}$  croissantes mais comparables. Les performances s'avèrent, elles aussi, très voisines et montrent une reconstitution du biais satisfaisante en moyenne. Toutefois de forts écarts aux valeurs observées peuvent survenir (de l'ordre de 0,06 à 0,08 l/s/km<sup>2</sup> pour une corrélation autour de 0,95). Si ces écarts s'avèrent non négligeables devant les valeurs de biais associées, ils restent en revanche assez peu significatifs comparés aux valeurs d'un intervalle de confiance à 80% ou 95 % ce qui permet de considérer finalement ces modèles comme suffisamment satisfaisants.

De même que dans le cas de l'écart-type  $\sigma_{(r,N,F)}$ , on précisera que les modèles présentés ci-dessus ne

peuvent être appliqués directement en cas de modification de la fréquence de jaugeage au cours d'une campagne de suivi. Si le raisonnement présenté en section 3.1.3 reste identique, l'absence de termes liés au nombre de jaugeages  $N$  simplifie grandement ici la formulation du biais final. Ainsi en reprenant les caractéristiques de la campagne de jaugeage présentée en section 3.1.3 le biais final sera donné par :

$$bias_{(r,N,F)} = bias_{(r_1,N_1,F_1)} + \Delta bias_{(r_2,N_2,F_2)}$$

$$\text{avec : } bias_{(r_1,N_1,F_1)} = \alpha_{(F_1=1)} \cdot r_1 + \beta_{(F_1=1)} \text{ et } \Delta bias_{(r_2,N_2,F_2)} = \alpha_{(F_2=2)} \cdot (r_2 - r_1)$$

Tableau 3 : Paramètres et performances moyennes des modèles de biais ajustés sur les données en fonction de la corrélation pour les cinq fréquences annuelles de jaugeage simulées - Source des données : Catalogne (2012)

$F$	$\alpha(F)$	$\beta(F)$	$RMSE$
1	-0.908	1.095	0.023
2	-0.929	1.136	0.022
3	-0.938	1.154	0.022
4	-0.922	1.155	0.019
5	-0.942	1.177	0.019

### 3.3 Comparaison avec les performances d'estimation du *QMNA5* en sites non jaugés

On se propose ici d'illustrer brièvement l'intérêt de la valorisation des données de jaugeage dans le cadre de l'estimation du *QMNA5* par rapport à une technique d'estimation en site non jaugé. Pour ce faire nous avons repris les résultats de validation croisée obtenus sur le même jeu de stations hydrométriques au moyen de modèles régressifs régionaux établis entre le *QMNA5* exprimé en débit spécifique (en l/s/km<sup>2</sup>) et un ensemble de variables explicatives (*i.e.* descripteurs du bassin se rapportant au climat, à la géologie ou à l'occupation des sols)<sup>5</sup>. Ces résultats sont comparés à ceux fournis en simulation par la méthode de valorisation des jaugeages épisodiques dans diverses configurations de suivi (opération répétée 100 fois pour chaque site cible et chaque configuration).

Les résultats obtenus en termes d'erreur absolue (écart entre les valeurs de *QMNA5* prédites et observées en l/s/km<sup>2</sup>) et de largeur de l'intervalle de confiance à 95 % sont résumés sur la Figure 12 sous la forme de boxplots. La comparaison des erreurs absolues semble indiquer que l'emploi des données de jaugeage n'est réellement profitable qu'à partir de cinq années de suivi. En deçà, l'estimation en sites non jaugés peut s'avérer meilleure. Toutefois il apparaît que le recours aux données de jaugeage permet de réduire significativement la largeur des intervalles de confiance associés aux estimations, quelle que soit la durée de suivi. Autrement dit la robustesse des estimations obtenues s'avère globalement meilleure, ce qui constitue un réel avantage de l'exploitation des données de jaugeage disponibles au site cible.

<sup>5</sup> Travaux réalisés au centre Irstea de Lyon dans le cadre de la convention ONEMA-Irstea (action relative à la prédétermination des étiages). Les détails de la méthode utilisée et des résultats obtenus sont présentés dans le rapport de Catalogne et Sauquet (2011) et rappelées dans (Catalogne, 2012).

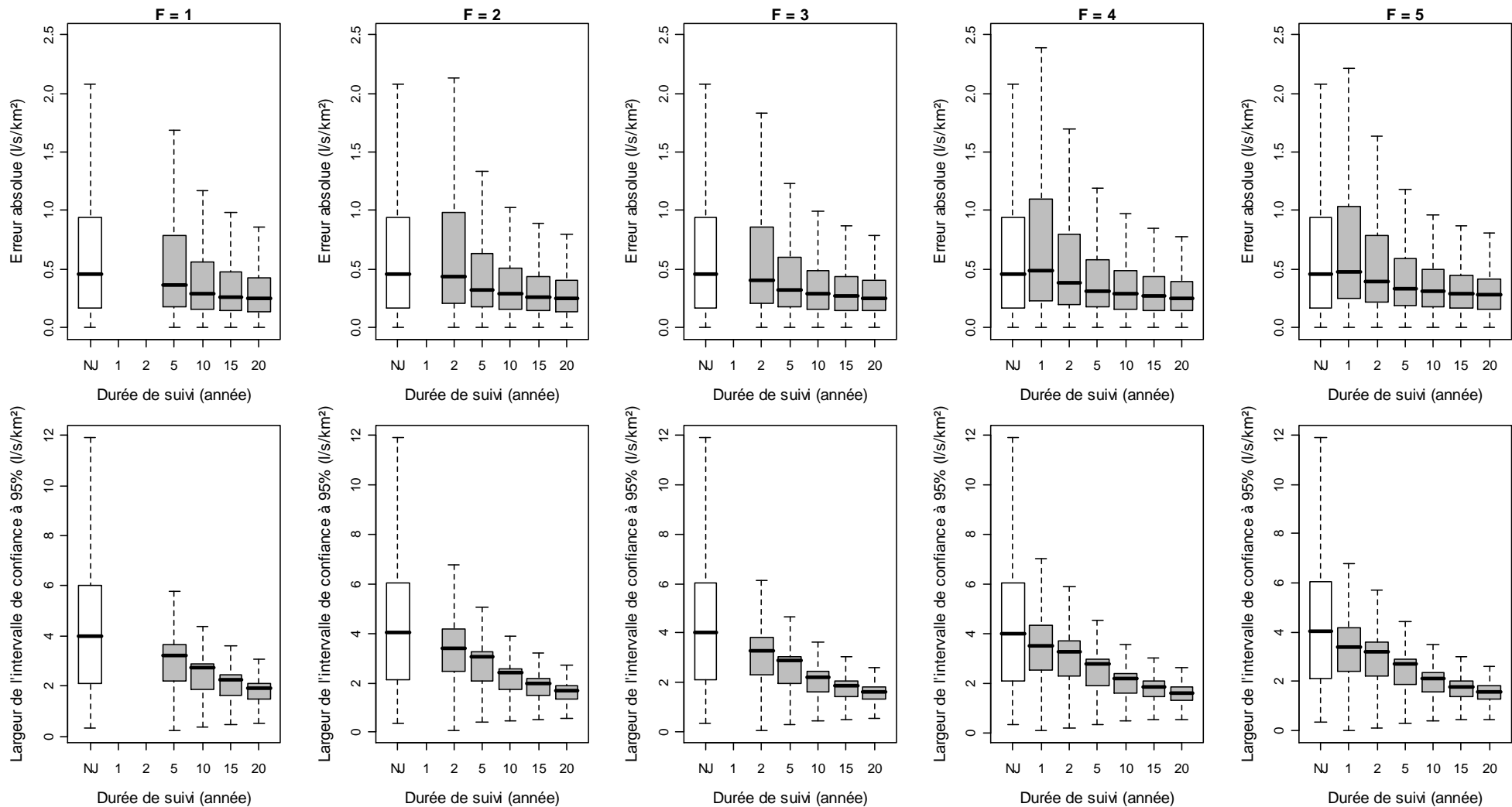


Figure 12 : Comparaison des erreurs d'estimation du QMNA5 et des intervalles de confiance à 95 % associés obtenus en sites non jaugés (en blanc) et par valorisation de jaugeages épisodiques (gris) en simulant différentes fréquences annuelles de jaugeage et pour une durée de suivi croissante © Clotaire Catalogne



## 4 Exemple d'application au bassin versant de l'Asse

Dans cette troisième partie nous nous proposons d'illustrer en détail l'application de la méthode présentée précédemment pour l'estimation du *QMNA5* dans un cadre opérationnel, c'est-à-dire en exploitant cette fois des données de jaugeage réelles. Pour ce faire nous avons choisi de présenter le cas d'une étude réalisée en collaboration avec la DREAL PACA sur le bassin de l'Asse, à hauteur de l'un des sept points stratégiques prévus par le Schéma Directeur d'Aménagement et de Gestion des Eaux (SDAGE) sur le bassin pour assurer le suivi des étiages (lieu dit « La Julienne »).

Ce choix tient avant tout à la disponibilité de données de jaugeage en nombre suffisant pour illustrer correctement les différents aspects de la méthode. Ainsi, il sera notamment question de la problématique de sélection des données de jaugeage à exploiter dans la procédure d'estimation. Nous verrons également ici quelques problèmes soulevés par l'application de la méthode, en particulier dans un contexte à la fois influencé par les prélèvements, l'existence d'échanges nappes-rivières et un régime hydrologique caractérisé par des étiages sévères.

### 4.1 Présentation du bassin versant de l'Asse

Le bassin de l'Asse correspond à un affluent rive gauche de la Durance dont l'exutoire se situe entre les communes d'Oraison à l'amont et de Manosque à l'aval dans le département des Alpes de Haute Provence (Figure 13). Au droit de la confluence avec la Durance, la surface drainée par le bassin atteint 692 km<sup>2</sup> pour un linéaire total de 130 km. L'altitude y varie entre 320 mètres à l'exutoire et 2280 mètres à sa source sur la montagne du Cheval Blanc pour une altitude médiane de 780 mètres. L'essentiel du bassin (80 %) se situe à une altitude inférieure à 1300 mètres.

Le bassin comprend dans sa partie amont trois tributaires principaux, appelés les « Trois Asses » (l'Asse de Clumanc, l'Asse de Moriez et l'Asse de Blieux), tandis que la partie aval ne compte qu'un affluent principal (l'Estoublaïsse). L'organisation du réseau hydrographique est essentiellement liée à une particularité remarquable du bassin qui comprend deux sous-ensembles géomorphologiques distincts (Figure 14) dont la limite physique est nettement marquée par la Clue de Chabrières. Celle-ci constitue en effet une frontière nette entre :

- un domaine que l'on pourra qualifier de montagneux à l'amont, au relief accentué avec une géologie constituée par les formations sédimentaires à dominante carbonatée du jurassique et du crétacé ;
- un domaine de plaine alluviale incisée dans les poudingues miocènes du plateau de Valensol où l'Asse prend la configuration d'un lit en tresse.

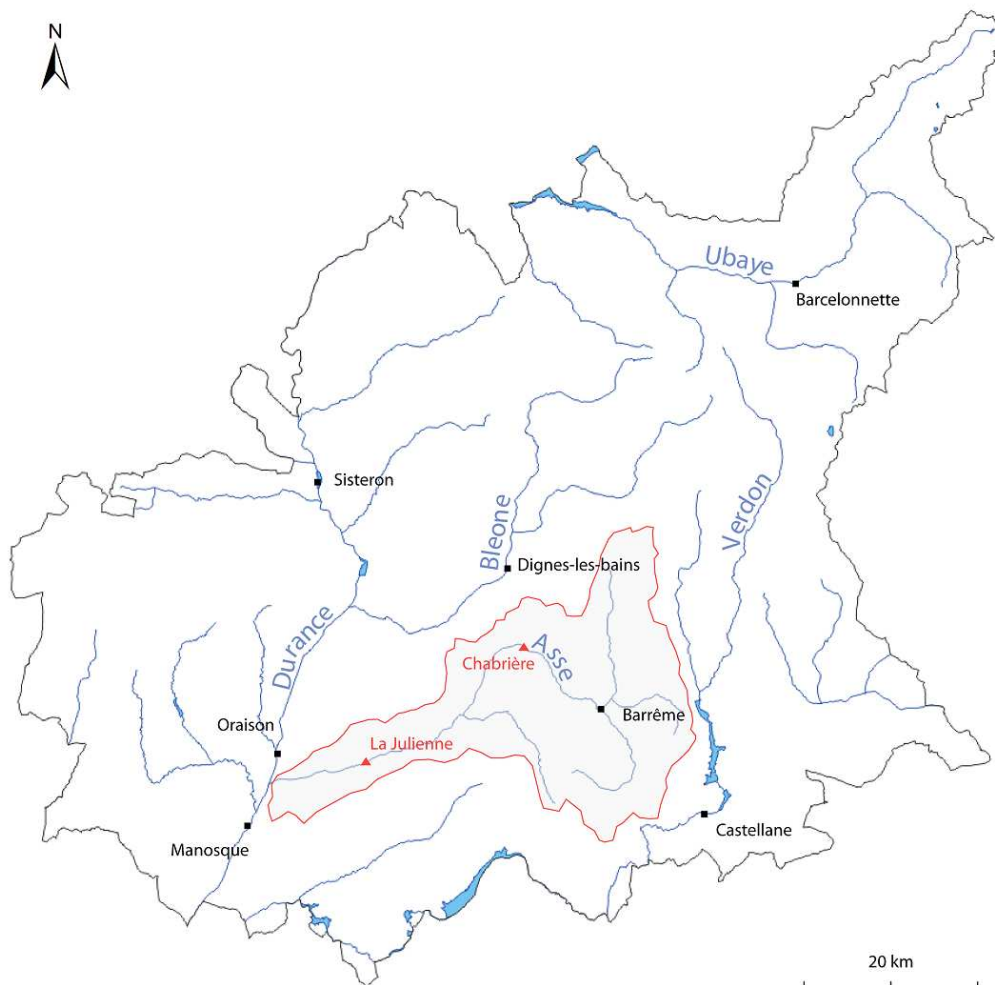


Figure 13 : Situation du bassin de l'Asse (délimité en rouge) dans le département des Alpes de Haute Provence © Daniel Dalet / d-maps.com

Bien entendu cette compartimentation du bassin implique des distinctions importantes en termes de (i) contexte climatique, plutôt humide à dominante montagnard en amont de la Clue de Chabrières avec des précipitations moyennes annuelles de l'ordre de 890 mm (dont une partie se produit sous forme solide) et de type plus méditerranéen en aval de celle-ci avec des précipitations moyennes annuelles de l'ordre de 650 mm, (ii) d'occupation des sols, avec une prédominance de la forêt dans les zones de relief et **une prédominance des zones cultivées dans la plaine alluviale, soumise à de fortes pressions agricoles durant la période estivale** et évidemment du point de vue de (iii) l'hydrologie avec **l'existence d'échanges importants entre l'Asse et sa nappe d'accompagnement dans la partie aval du bassin** (alors que ceux-ci peuvent être considérés comme pratiquement inexistant en amont de la Clue de Chabrières) où la géométrie de la plaine alluviale détermine alors le sens des échanges avec une tendance à l'infiltration des eaux de l'Asse dans la partie allant de la Clue de Chabrières jusqu'à Estoublon et une tendance à l'alimentation du cours d'eau par la nappe depuis Estoublon jusqu'à la confluence avec la Durance.

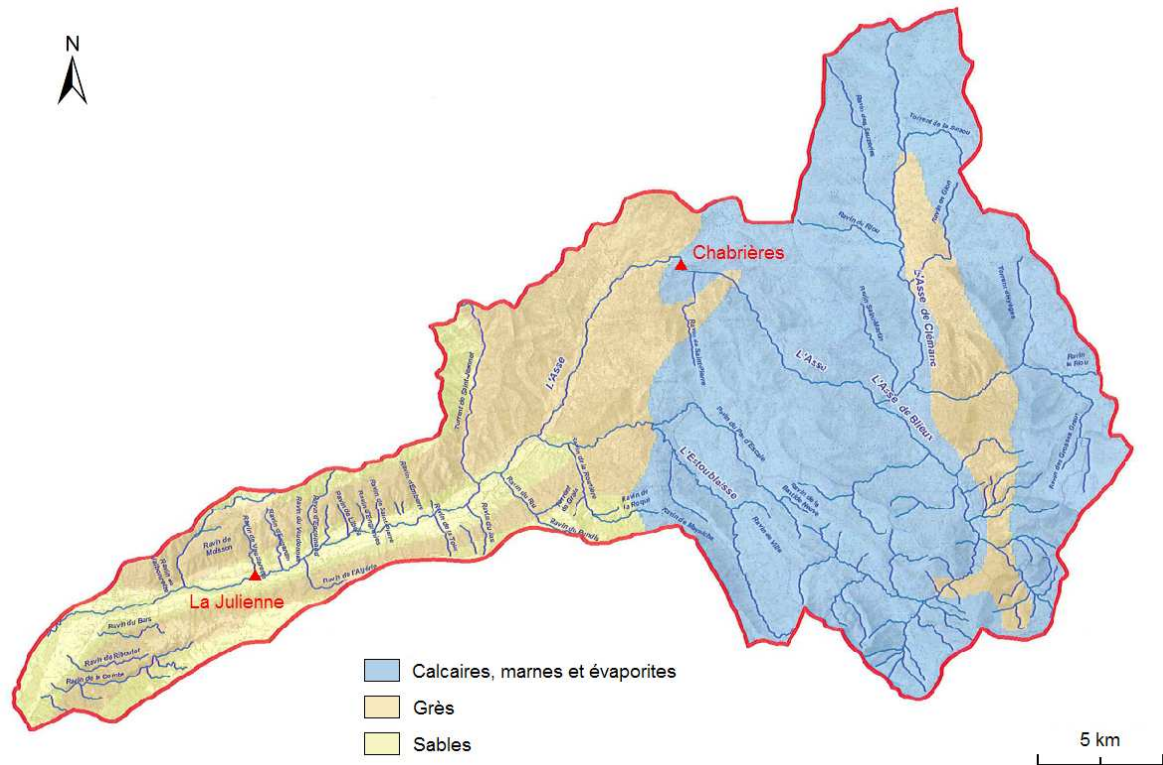


Figure 14 : Carte simplifiée des lithologies rencontrées sur le bassin de l'Asse. Figure modifiée d'après AERMC (2011) © AERMC

## 4.2 Analyse des données disponibles

L'analyse des données proposée dans ce paragraphe constitue une étape clé qu'il convient de bien détailler. Celle-ci aura tout d'abord pour objet la description préalable du régime hydrologique du bassin de l'Asse à partir des données d'une station hydrométrique. Il sera ensuite question de l'analyse des données de jaugeage disponibles au site cible de manière à illustrer le type de problèmes soulevé par l'emploi de données de jaugeage réelles et ne conserver que les données les plus pertinentes pour l'application de la méthode d'estimation du QMNA5.

### 4.2.1 Données de la station hydrométrique de l'Asse à Chabrières

Le bassin de l'Asse n'est équipé que d'une unique station hydrométrique pérenne situé à hauteur de la Clue de Chabrières : la station X1524010 de l'Asse à Beyne [Chabrières]. Cette station, gérée par EDF, comporte 54 années d'enregistrement mises à disposition par la banque HYDRO pour la période allant de 1953 à 2008 (incluant deux années incomplètes en 1978 et 1979). Les données des années 2009 à 2010 ont été directement fournies par EDF/DTG. La surface drainée à hauteur de la station atteint 375 km<sup>2</sup>.

L'analyse du régime hydrologique à l'échelle mensuelle à partir des données de cette station hydrométrique (Figure 15) montre que le bassin amont possède un régime pluvio-nival avec une saison de basses eaux principale durant la période estivale (du mois de juillet au mois de septembre) et un minimum secondaire au mois de février, lié à la rétention d'une partie des précipitations sous forme solide en tête de bassin.

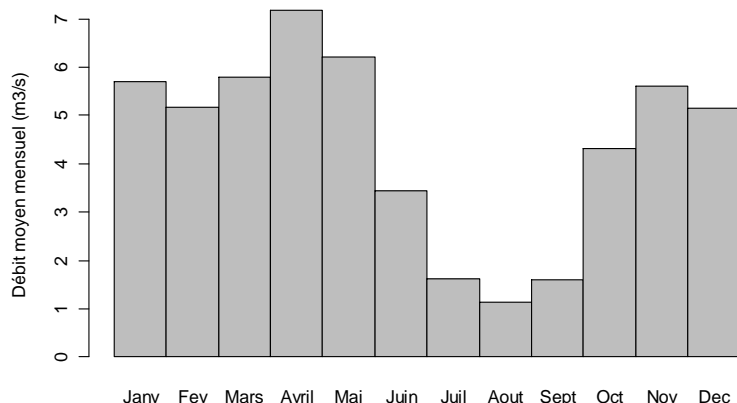


Figure 15 : Evolution des débits moyens mensuels interannuels © Clotaire Catalogne

Les débits mensuels minimums annuels (*QMNA*) s'observent préférentiellement au mois de septembre. Les écoulements se caractérisent par ailleurs par une assez faible composante de base comme en atteste la contribution des stocks au débit annuel (valeur du BFI « Base Flow Index » voisine de 0.56) et donc une forte sensibilité à l'absence de précipitations. Néanmoins, le régime pluviométrique du bassin semble suffisamment régulier pour limiter la durée des épisodes d'étiage : la durée moyenne de l'événement le plus sévère de l'année<sup>6</sup> est de 23 jours en moyenne (contre une moyenne nationale de 30 jours).

Sur le long terme, la valeur de *QMNA5* atteint 0,36 m<sup>3</sup>/s (période 1954-2010). Elle est plus élevée si l'on considère la période de référence 1970-2005 retenue au cours de nos propres travaux (Catalogne, 2012) avec une valeur de 0,46 m<sup>3</sup>/s (proche du débit de fréquence au dépassement 95% sur la même période). En tout état de cause, le bassin peut être qualifié d'assez peu productif avec un *QMNA5* spécifique de l'ordre 1 l/s/km<sup>2</sup>, inférieur aux valeurs observées pour les autres affluents de la Durance équipés d'une station hydrométrique en amont de la confluence avec le Verdon (*QMNA5* de l'ordre de 3,5 l/s/km<sup>2</sup> pour le Bes à la Javie (X1225010), 2 l/s/km<sup>2</sup> pour le Buech à Serre [Les Chambons] (X1034010) ou encore 4,5 l/s/km<sup>2</sup> pour l'Ubaye au Lauzet-Ubaye [Roche Rouse] (X0454010)).

Etant donné que la station de Chabrières se situe sur le même linéaire hydrologique que le site cible de la Julienne (en amont de celui-ci) et répond parfaitement aux critères de sélection évoqués en section 2.4 (*i.e.* au moins 26 années de chronique, régime peu ou pas perturbé par les activités humaines, données fiables), celle-ci sera considérée d'emblée comme le meilleur site de référence possible sans chercher au-delà un bassin instrumenté plus approprié. Le protocole de sélection proposé en section 2.4 dans le cadre de la validation croisée ne sera donc pas appliqué ici. Il restera cependant pertinent pour orienter le choix du site d'appui dans le cas de bassins cibles dépourvus d'une station hydrométrique située sur le même tronçon hydraulique et en l'absence de connaissances précises sur le contexte hydro-climatique du bassin étudié.

#### 4.2.2 Données de jaugeage à la Julienne

Le site de la Julienne est situé dans la partie aval du bassin étudié, à hauteur du hameau du même nom situé le long de le D907, en rive droite de la plaine alluviale de l'Asse ; il draine une surface de 624 km<sup>2</sup> soit 90% de la surface totale du bassin de l'Asse. A hauteur de ce site, le cours d'eau présente une configuration en tresse avec évolution latérale de la position des chenaux dans le temps. Si cette configuration exclut la possibilité d'implanter une station hydrométrique permanente, le site a en revanche fait l'objet d'un suivi par jaugeages sur la période allant de 2004 à 2010 par la DDT des Alpes de Haute Provence.

<sup>6</sup> Il s'agit de la plus longue séquence continue de débit sous le seuil Q90 (= débit dépassé 90% du temps sur toute la chronique).

Le nombre de jaugeages disponibles y est assez important avec au moins 9 mesures de débit réalisées chaque année (Tableau 4). La majorité d'entre elles se rapportent par ailleurs à la période de basses eaux identifiée précédemment à la station de Chabrières (*i.e.* trimestre le plus sec allant de juillet à septembre). Cette situation semble donc tout à fait favorable à l'application de la méthode destinée à estimer le QMNA5 au site de la Julienne (par ailleurs, la station de Chabrières dispose d'observations de débit pour l'ensemble des dates de jaugeage). Toutefois l'examen de ces données permet de montrer que l'ensemble des jaugeages réalisés n'est pas forcément exploitable dans une perspective d'estimation du QMNA5.

Tableau 4 : Dénombrement des jaugeages réalisés au site de la Julienne sur la période 2004-2010 - Source des données : DREAL PACA

Année	Nombre de jaugeages réalisés chaque année	Nombre de jaugeages réalisés durant le trimestre le plus sec
2004	9	6
2005	11	10
2006	21	13
2007	13	11
2008	13	10
2009	9	7
2010	9	6
<b>Total</b>	<b>85</b>	<b>63</b>

Pour illustrer cela, un exercice intéressant consiste probablement à confronter ces données de débit à celles observées à la station de Chabrières à la même date. Ainsi en considérant l'ensemble des données disponibles on constate que la concordance entre certains débits mesurés aux deux sites peut s'avérer assez médiocre malgré le fait que les deux bassins soient emboîtés (Figure 16).

La première explication repose sans doute sur la confrontation parfois inappropriée d'un débit instantané (mesuré au site cible) à un débit moyen journalier (observé à la station de référence), probablement accentué ici par l'existence d'un temps de transit non négligeable entre les deux sites. Le meilleur exemple en est sans doute le jaugeage du 6 juillet 2006 où le débit mesuré à la Julienne est nettement inférieur à celui observé à Chabrières. On constate en effet qu'à cette date correspond le début d'un petit épisode de crue estivale qui n'était probablement pas parvenu jusqu'au site de la Julienne au moment du jaugeage (Figure 17). Ceci illustre parfaitement la nécessité de s'assurer que les conditions de jaugeage correspondent à un régime suffisamment stabilisé au site cible comme au site de référence.

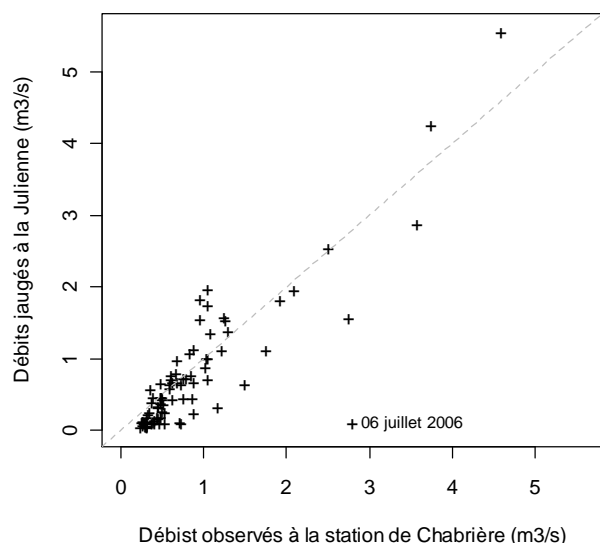


Figure 16 : Comparaison des débits jaugeés au site de la Julienne et des débits observés à la station de Chabrières à la même date © Clotaire Catalogne

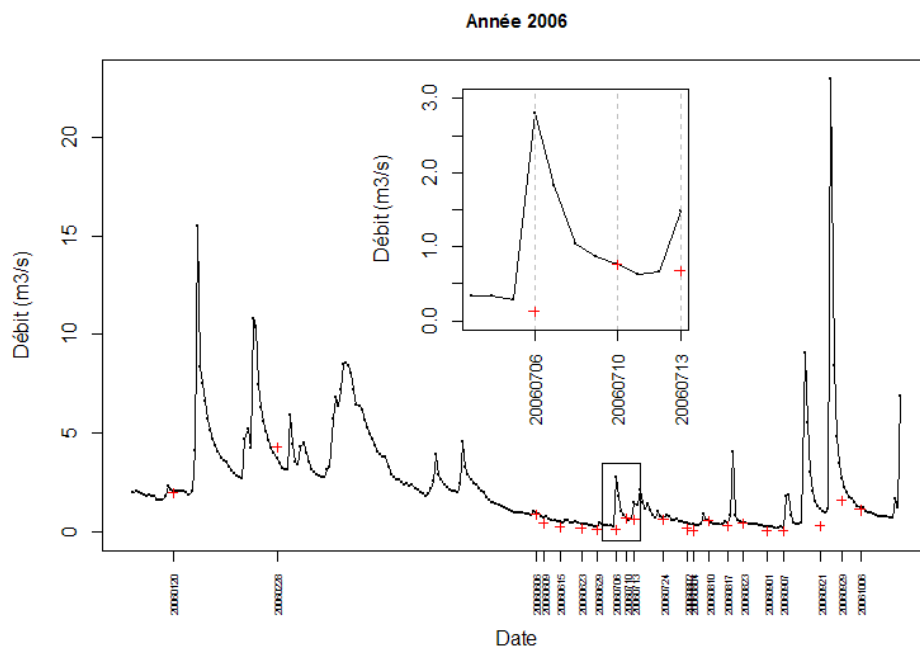


Figure 17 : Illustration de la non simultanéité des réponses hydrologiques aux sites de La Julienne et de Chabrières pour le jaugeage du 6 juillet 2006 © Clotaire Catalogne

On remarquera que dans le cas étudié, cette critique est évidemment largement facilitée par l'existence d'un site de référence situé sur le même linéaire que le site cible (et suffisamment proche de celui-ci) qui permet d'avoir une vision de l'historique des variations des débits sur le bassin. Dans le cas contraire, on pourra se référer à l'historique climatique du bassin considéré (on pourra par exemple exclure les données précédant un épisode pluvieux significatif sur le bassin et/ou dans un délai de  $d = S^{0.2}$  jours après celui-ci<sup>7</sup>).

A la problématique de simultanéité des réponses hydrologiques s'ajoute une seconde remarque d'importance, il est en effet surprenant de constater que les valeurs de débits mesurées au site de La Julienne s'avèrent souvent inférieures à celles observées à la station de Chabrières à la même date (alors que la surface drainée y est presque deux fois supérieure<sup>8</sup>). Considérant que les éventuelles imprécisions sur les données examinées ne sont pas suffisantes pour expliquer une telle anomalie, il est alors possible d'envisager que l'existence de prélèvements en amont du site de la Julienne durant la période estivale et/ou que l'existence de pertes vers la nappe d'accompagnement conduit à priver le cours d'eau d'une partie non négligeable de l'écoulement provenant de l'amont.

Comme en attestent les différentes études relatives aux usages de l'eau sur le bassin (AERMC (2011)), l'existence de prélèvements est en effet susceptible d'influencer significativement le régime hydrologique en période d'étiage, d'autant que la dispersion du nuage de point tendrait à démontrer l'existence d'une part de variabilité « aléatoire » pouvant être attribuée à la variation des volumes prélevés selon les dates de jaugeage (*i.e.* selon la saison et selon l'année). Cette situation constitue évidemment un réel obstacle à la détermination d'un débit de référence « naturel » au sens des études volumes prélevables (EVP). Ainsi la méthode d'estimation du QMNA5 ne pourra ici que restituer une valeur « influencée » par les éventuels prélèvements, inférieure à la valeur en régime naturel.

Malgré l'absence d'historique détaillé des prélèvements qui aurait pu permettre de ne retenir que les données les moins influencées, la critique des données s'est tout de même poursuivie de manière à

7 Il s'agit d'une manière simpliste mais pratique d'estimer le temps de ressuyage d'un bassin.

8 Lorsqu'ils sont exprimés en débit spécifique, les débits jaugés à La Julienne sont en moyenne 2.65 fois inférieurs à ceux observés à Chabrières !

retenir un total de 62 jaugeages (Figure 18). On remarquera qu'à ce stade, certaines mesures réalisées hors de la période correspondant au trimestre le plus sec sont encore présentes dans l'échantillon. Si la conservation des jaugeages réalisés aux mois de juin et octobre reste légitime, on pourra considérer que des débits tels que ceux mesurés au mois de janvier 2006 ou au mois d'avril 2008 (avec des valeurs de l'ordre de 2 m<sup>3</sup>/s) ne sont plus représentatifs du régime de basses eaux à proprement parler. Néanmoins, d'après la chronique observée à Chabrières, ces débits semblent tout de même se rapporter à la composante de débit de base et s'avèrent par ailleurs correspondrent aux données les moins susceptibles de se trouver influencées par d'éventuels prélèvements, c'est pourquoi il a été choisi de les conserver<sup>9</sup>.

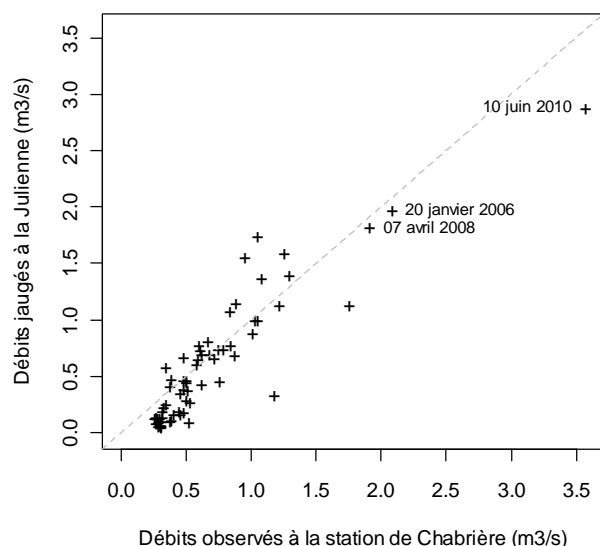


Figure 18 : Données de jaugeage retenues après exclusion des données se rapportant à des périodes de fortes variations du débit à la station de Chabrières © Clotaire Catalogne

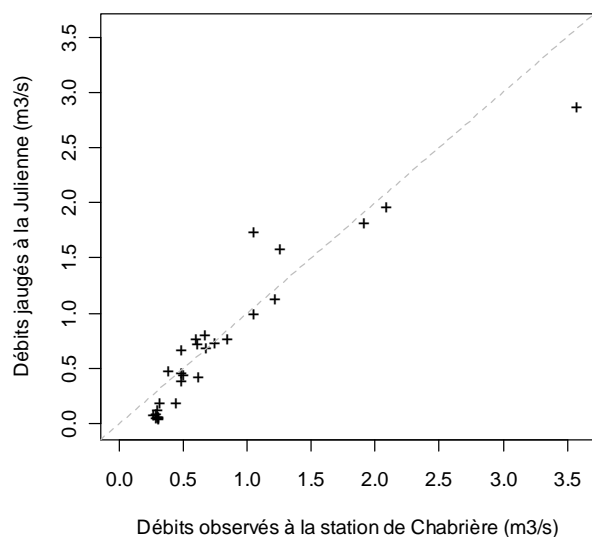


Figure 19 : Exemple de sélection de données de jaugeage relatives à des conditions hydrologiques contrastées © Clotaire Catalogne

A partir de cette présélection, une dernière étape consiste enfin à sélectionner les débits jaugés de manière à obtenir les valeurs les plus indépendantes possibles et ainsi ne conserver qu'un maximum de 5 mesures chaque année :

- soit en imposant une durée minimale de 15 jours entre deux jaugeages (comme proposé dans le

<sup>9</sup> On a là une bonne illustration de la tolérance pouvant être accordée dans la sélection des données de jaugeage par rapport aux contraintes fixées pour faciliter l'automatisation de la validation croisée

protocole de validation croisée),

- soit en considérant les éventuels épisodes de remontée de débits (suite par exemple à un événement pluvieux isolé au cours de la période de basses eaux) pour différencier les périodes d'étiages successives.

Evidemment, il n'existe pas de choix unique pour réaliser cette dernière étape de sélection. On pourra cependant recommander de privilégier des jaugeages représentatifs de situations contrastées. Un exemple de sélection réalisée selon cette recommandation est reporté sur la Figure 19, il sera utilisé par la suite pour illustrer l'application de la méthode.

#### 4.2.3 Mise en œuvre de la méthode pour l'estimation du QMNA5 et la détermination de l'incertitude associée

La critique des données ayant permis de retenir un jeu de dates de jaugeage pertinentes pour l'estimation du QMNA5, nous proposons ici d'appliquer la méthode d'estimation décrite dans la première partie de ce document.

Pour ce faire, et de manière à bien clarifier la mise en œuvre de la méthode, **nous proposons d'appliquer celle-ci en ajoutant progressivement les données de jaugeage obtenues au terme de chaque année**. A ce sujet il convient de préciser que nous nous trouvons dans le cas d'une stratégie de jaugeage complexe où la fréquence de jaugeage évolue d'une année sur l'autre. Il convient alors de se référer à la méthodologie exposée en section 3.1.3 pour permettre la détermination de l'incertitude autour de la valeur estimée. Pour illustrer cette application, nous considérerons l'échantillon de jaugeages retenu précédemment en guise d'exemple. Les caractéristiques de la campagne de jaugeage correspondantes sont reportées dans le Tableau 5.

Tableau 5 : Caractéristiques de la campagne de jaugeage retenue - Source des données : DREAL PACA

Année	Fréquence de jaugeage	Nombre total de jaugeage	Corrélation
2004	2	2	1
2005	4	6	0,95
2006	5	11	0,97
2007	5	16	0,97
2008	4	20	0,97
2009	4	24	0,97
2010	5	29	0,94

On remarque tout d'abord qu'au terme de l'année 2004, le nombre de jaugeages réalisés est encore insuffisant pour pouvoir effectuer l'ajustement ( $N < 4$ ). Il faudra alors attendre l'année 2005 pour obtenir une première valeur de QMNA5 estimée. A ce stade, il est en revanche d'ores et déjà possible d'initialiser le calcul du biais et de l'écart type d'estimation en s'appuyant sur les abaques issus de la validation croisée et en prenant  $F = 2$ ,  $N = 2$  et une corrélation  $r = 1$  (par défaut) :

$$bias_{2004} = \alpha_{(F=2)} \cdot r + \beta_{(F=2)} = \alpha_{(F=2)} + \beta_{(F=2)} (= 0.207 \text{ l/s/km}^2)$$

$$\sigma_{2004} = a_{(F=2)} \cdot \ln(N_{2004}) + c_{(F=2)} \cdot (1 - r) + d_{(F=2)} = a_{(F=2)} \cdot \ln(N_{2004}) + d_{(F=2)} (= 1.267 \text{ l/s/km}^2)$$

Au terme de l'année 2005, un premier ajustement à partir d'un total de  $N = 6$  valeurs est possible. La valeur de QMNA5 obtenue au site cible atteint 429 l/s/km<sup>2</sup> ou 0,707 l/s/km<sup>2</sup> selon que l'on considère la valeur sur la période 1954-2011 ou 1970-2005. Le calcul du biais et de l'écart type associé s'effectue ensuite en trois étapes :



- le calcul du nombre de jaugeages équivalent, tenant compte de la nouvelle fréquence de jaugeage  $F = 4$  :

$$N_{eq} = \exp\left(\frac{\sigma_{2004} - c_{(F=4)} \cdot (1 - r_{2004}) - d_{(F=4)}}{a_{(F=4)}}\right) (= 2,06)$$

- le calcul du gain apporté au biais et à l'écart type par les nouveaux jaugeages de l'année 2005 :

$$\Delta bias_{2005} = \alpha_{(F=4)} \cdot (r_{2005} - r_{2004}) (= 0,011 \text{ l/s/km}^2)$$

$$\Delta \sigma_{2005} = a_{(F=4)} \cdot (\ln(N_{eq} + N_{2005}) - \ln(N_{eq})) + c_{(F=4)} \cdot (r_{2004} - r_{2005}) (= 0,246 \text{ l/s/km}^2)$$

- le calcul du biais et de l'écart type final en ajoutant le gain apporté par les nouveaux jaugeages de l'année 2005 aux valeurs obtenues au terme de l'année 2004 :

$$bias_{2005} = bias_{2004} + \Delta bias_{2005} (= 0,196 \text{ l/s/km}^2)$$

$$\sigma_{2005} = \sigma_{2004} + \Delta \sigma_{2005} (= 1,021 \text{ l/s/km}^2)$$

Cette opération est répétée de manière identique pour chaque nouvelle année de mesures disponibles en employant les abaques correspondants à la fréquence de jaugeage de l'année en question. L'ensemble des résultats obtenus est synthétisé dans le Tableau 6 et sur la Figure 20 :

Tableau 6 : Détail des résultats d'estimation obtenus année par année - Source des données : DREAL PACA

Année	F	N	N <sub>eq</sub>	r	QMNA5 estimé (l/s/km <sup>2</sup> )		Biais (l/s/km <sup>2</sup> )	Ecart type (l/s/km <sup>2</sup> )
					Période de référence			
					1954-2010	1970-2005		
2004	2	2	2,06	1			0,207	1,267
2005	4	6	6,50	0,95	0,429	0,707	0,250	1,038
2006	5	11	11,50	0,98	0,463	0,608	0,228	0,881
2007	5	16	14,86	0,98	0,409	0,548	0,229	0,796
2008	4	20	18,86	0,98	0,462	0,608	0,228	0,736
2009	4	24	25,87	0,97	0,491	0,641	0,235	0,695
2010	5	29		0,94	0,624	0,781	0,260	0,677

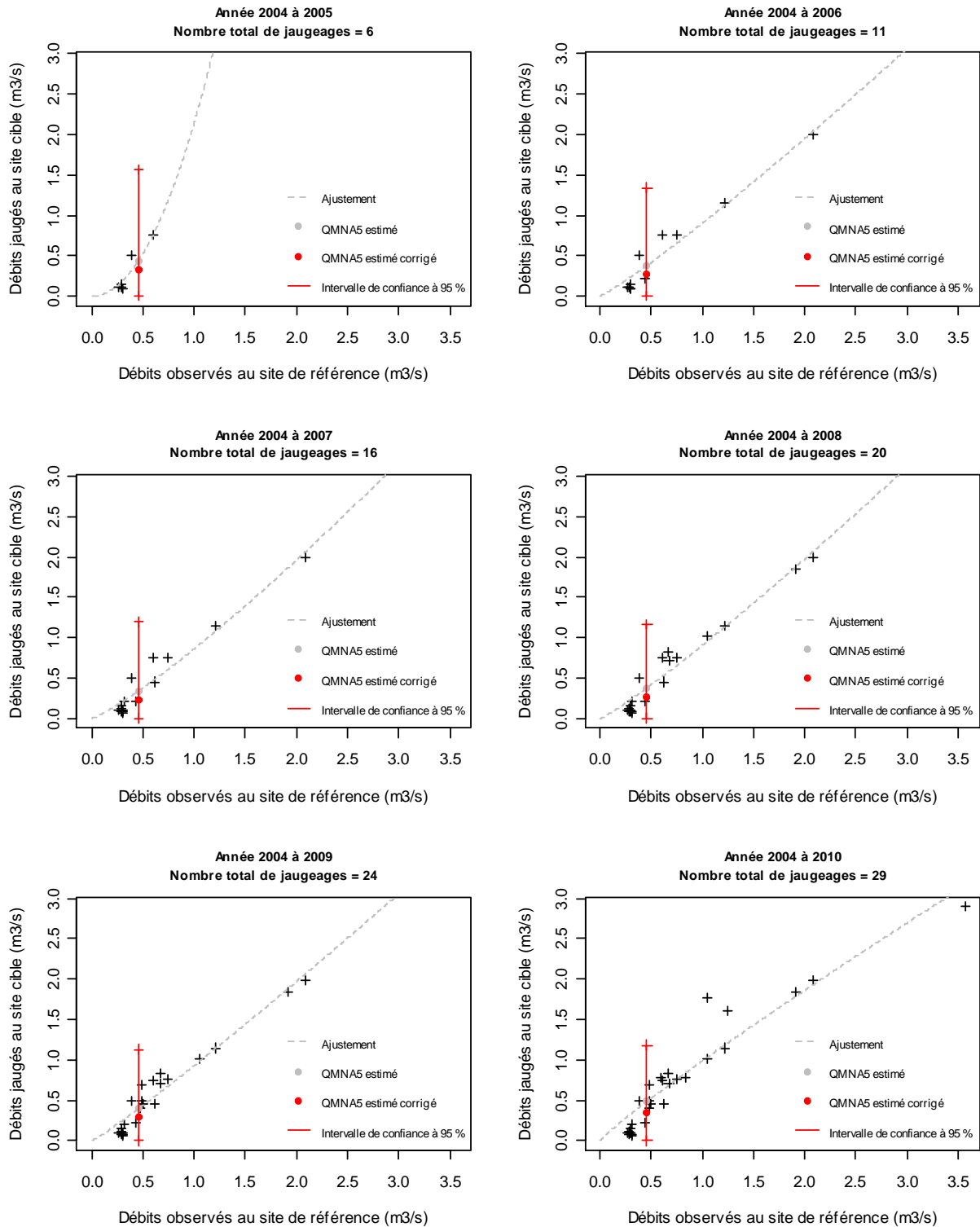


Figure 20 : Illustration des ajustements réalisés, des valeurs estimées (en retenant la période de référence 1970-2005) et des intervalles de confiance à 95 % associés au terme de chaque année de suivi © Clotaire Catalogne

Nous rappelons que finalement, l'intervalle de confiance associé à la valeur de  $QMNA5$  estimée au niveau de significativité voulu peut être déterminé au moyen de :

$$P(Q_{est} + bias_{(r,N,F)} - t_{\alpha/2} \cdot \sigma_{(r,N,F)} < Q_{obs} < Q_{est} + bias_{(r,N,F)} + t_{\alpha/2} \cdot \sigma_{(r,N,F)})$$

où  $t_{\alpha/2}$  correspond au quantile d'ordre  $\alpha/2$  de la loi normale ( $\alpha = 0.05$  et  $t_{\alpha/2} \approx 1.96$  pour un intervalle de confiance à 95%).

Finalement, il peut être également intéressant de résumer ces résultats à l'aide d'un graphique montrant

l'évolution des valeurs de *QMNA5* estimées au cours de la campagne de jaugeage et associées ici à leur intervalle de confiance à 95 % (par défaut borné à zéro, Figure 21).

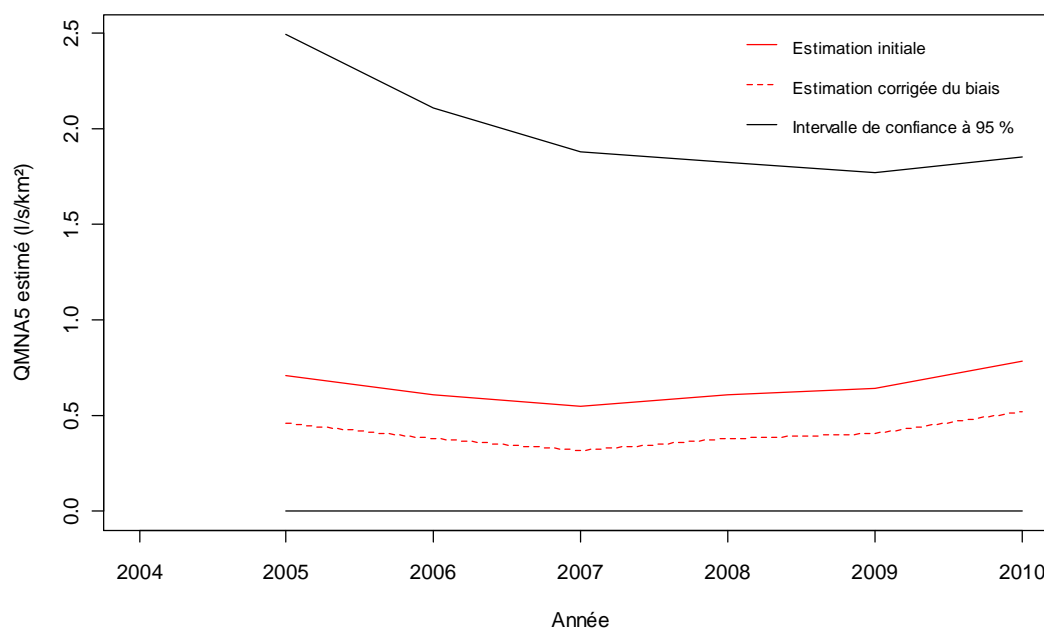


Figure 21 : Illustration de l'évolution du *QMNA5* estimé (en retenant la période de référence 1970-2005) et de l'intervalle de confiance à 95 % associé au cours de la période de suivi © Clotaire Catalogne

### 4.3 Discussion des résultats

Dans l'exemple présenté ci-dessus, on constate que les valeurs estimées s'avèrent relativement stables mais particulièrement faibles (autour de 0,65 l/s/km<sup>2</sup> pour la valeur brute et de 0,40 l/s/km<sup>2</sup> pour la valeur corrigée du biais) et nettement inférieures à la valeur observée à Chabrières (1,2 l/s/km<sup>2</sup> en retenant la période de référence 1970-2005). Malgré un niveau de corrélation élevé et un nombre de jaugeages déjà important, l'intervalle de confiance à 95 % reste particulièrement large à l'issue de la campagne de jaugeage puisque l'estimation finale corrigée du biais est ici susceptible de varier entre 0,521 l/s/km<sup>2</sup> +/- 230% (la valeur inférieure étant par défaut bornée à zéro).

Si ce résultat peut apparaître comme décevant, il montre surtout que la caractérisation de l'incertitude à l'aide des modèles ajustés sur les résultats de validation croisée reste peut-être trop grossière. En effet les modèles sont définis pour un jeu de bassins ayant des débits variant de 0 l/s/km<sup>2</sup> à 14,7 l/s/km<sup>2</sup>. La caractérisation de l'incertitude qui en résulte est donc plus adaptée aux valeurs de débit correspondant au cœur de la distribution des valeurs étudiées en validation croisée (de 2 à 4 l/s/km<sup>2</sup>), en contrepartie elle conduit probablement à surestimer l'incertitude pour les bassins peu productifs (*i.e.* bassin présentant de faibles débits spécifiques comme c'est le cas pour l'Asse) et, inversement, à sous-estimer l'incertitude pour les bassins les plus productifs. Cette remarque ne remet pas en question la méthode en tant que telle mais doit nous inciter à réaliser un travail de modélisation de l'incertitude plus fin par l'ajustement de modèles de biais et d'écart-type adaptés aux différents contextes hydro-climatiques. Le résultat obtenu ici à l'aide des jaugeages reste néanmoins plus fiable que l'estimation obtenue au moyen d'un modèle régressif régional établi entre le *QMNA5* exprimé en débit spécifique (en l/s/km<sup>2</sup>) et un ensemble de variables explicatives (*i.e.* descripteurs du bassin se rapportant au climat, à la géologie ou à l'occupation des sols). En effet selon cette méthode d'estimation, la valeur de *QMNA5* obtenue à La Julienne atteint 0,6 l/s/km<sup>2</sup> +/- 400%.

L'information sur l'incertitude reste par ailleurs particulièrement intéressante pour caractériser l'efficacité

de la campagne de jaugeage et déterminer s'il s'avère intéressant de poursuivre ou non celle-ci. On voit en effet ici que la largeur de l'intervalle de confiance diminue progressivement, l'écart type étant presque divisé par deux entre 2004 et 2010 et tendant à se stabiliser (Figure 22, en trait plein).

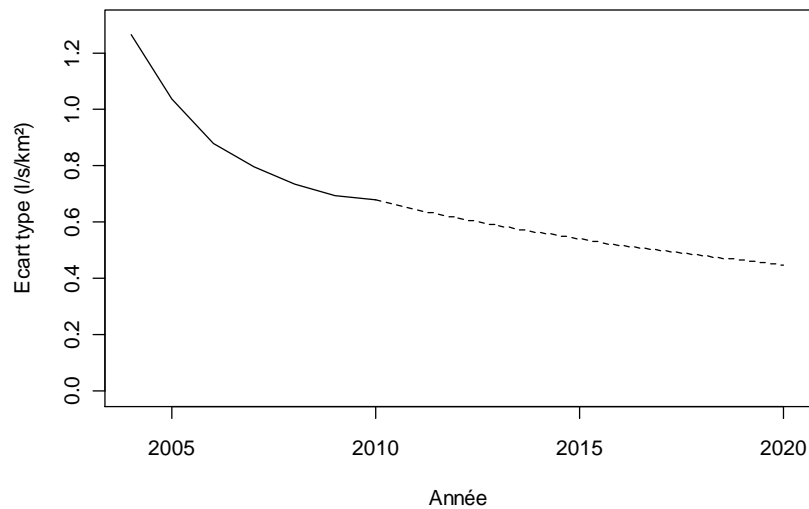


Figure 22 : Evolution de l'écart type au cours de la campagne de jaugeage (trait plein) et évolution prévue par les modèles pour une fréquence de 3 jaugeages par année en faisant l'hypothèse que le niveau de corrélation entre les débits jaugés à La Julienne et ceux observés à Chabrières reste constant (pointillés) © Clotaire Catalogne

En envisageant de poursuivre la campagne sur 10 années supplémentaires avec la réalisation de trois jaugeages chaque année (comme recommandé en 3.1.2) et en faisant l'hypothèse que le niveau de corrélation entre les débits jaugés à La Julienne et ceux observés à Chabrières reste constant on constate que le gain supplémentaire sur les dix années à venir est beaucoup moins soutenu qu'en début de campagne mais peut tout de même rester intéressant (gain supplémentaire de 17 % par rapport à la valeur de départ de 2004).

## 5 Conclusion

Ce document 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 s'appuie 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 jaugeages réalisés, 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.

Hormis l'importance de disposer d'un site d'appui présentant une dynamique hydrologique suffisamment voisine de celle du site cible (degré de corrélation entre débits observés à la même date élevé), 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 s'atténue assez rapidement au-delà. Parallèlement il apparaît que 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.

La confrontation des résultats obtenus avec ceux issus d'une méthode d'estimation en sites non jaugés montre dans la majorité des cas l'intérêt de la valorisation des données de jaugeage volant, notamment vis-à-vis de la fiabilité des valeurs estimées. La méthode s'appuyant sur les données de jaugeage pourra donc constituer une alternative intéressante (moins coûteuse) à l'implantation d'une station pérenne sur les bassins à enjeu ou sur les sites difficiles à instrumenter (*e.g.* lits en tresse, sections instables...).

Enfin ce document s'achève sur un exemple d'application de la méthode proposée au cas du bassin de l'Asse, situé dans les Alpes de Haute Provence. Cet exemple, mené à partir de données de jaugeage réelles réalisées sur l'un des sept points stratégiques définis sur le bassin pour le suivi de la situation hydrologique, permet d'illustrer le protocole d'estimation du  $QMNA5$  et de l'incertitude associée dans le cas d'une stratégie de jaugeage complexe (en réactualisant les résultats au terme de chaque nouvelle année de suivi du site cible). Il y est notamment question de l'indispensable étape d'expertise des données de jaugeage permettant de retenir les données les plus pertinentes pour l'estimation du  $QMNA5$ . Cette application a également permis de pointer les limites de la méthode, en particulier dans un contexte où l'écoulement se trouve perturbé par l'existence de prélèvements durant la saison d'étiage (la méthode ne peut alors que restituer des valeurs influencées) mais aussi dans un contexte hydro-climatique relativement sec pour lequel la caractérisation de l'incertitude au travers des modèles de biais et d'écart type fournit des valeurs d'intervalles de confiance probablement surestimées.

## 6 Bibliographie

AERMC (2011). Etude de détermination des volumes prélevables du bassin versant de l'Asse – Caractérisation du bassin versant et bilan des prélèvements (phases 1 et 2). 79 pages + annexes.

Catalogne C., Sauquet E. (2011). Interpolation des débits de référence d'étiage. Rapport d'étude, Irstea-Onema, 58 pages.

Catalogne C. (2012). Amélioration des méthodes de prédétermination des débits de référence d'étiage en sites peu ou pas jaugés. Thèse de l'Université de Grenoble, 171 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 Rhône Alpes (2003). Les débits d'étiage en Rhône-Alpes. Rapport d'étude, 14 pages.

Galéa 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*, Vol. 18(3), 331-352.

Ghosh B. (1951). Random distances within a rectangle and between two rectangles. *Bull. Calcutta Math. Soc.*, 43, 17-24.

Gottschalk L. (1993). Interpolation of runoff applying objective methods. *Stochastic Hydrology and Hydraulics*, 7: 269-281.

Laaha G., Blöschl, G. (2005). Low flow estimates from short stream flow records - a comparison of methods. *Journal of Hydrology*, 306(1-4): 264-286.

Oberlin G., Galéa 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.*, vol. X, No. 4, 349-368.

Reilly C.F., Kroll C.N. (2003). Estimation of low streamflow statistics using baseflow correlation. *Water Resources Research*, 39(9), 1236.

Riffard M., Andréassian V., Nicolle P., Peschard J. (2012). Combinaison multi-modèle et cartographie de consensus du débit de référence d'étiage et du débit moyen à l'échelle de la France. Rapport d'étude, Irstea-Onema, 40 pages.

Riggs H.C. (1972). *Techniques of Water-Resources Investigations of the United States Geological Survey - Chapter B1 : Low-flow estimation*. USGS publication.

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

Stedinger J.R., Thomas W.O. (1985). Low-flow frequency estimation using base-flow

measurements. USGS Scientific Investigations, Report 85-95.

Zhang Z., Kroll C. (2007). The baseflow correlation method with multiple gauged sites. *Journal of Hydrology*, 347: 371-380.

## 7 Table des illustrations

Figure 1 : Cartographie des QMNA5 spécifiques (en l/s/km <sup>2</sup> ) aux stations de l'échantillon de référence © Clotaire Catalogne.....	3
Figure 2 : Exemple de mise en relation entre débits jaugés au site cible (9 jaugeages simulés ici à partir d'une chronique) et débits observés à la même date sur un bassin de référence voisin © Clotaire Catalogne .....	4
Figure 3 Carte des régimes hydrologiques le long du réseau hydrographique selon la typologie définie par les douze séries de coefficients mensuels de débit © Eric Sauquet.....	7
Figure 4 : Schéma illustrant la procédure de calcul de distance $d$ pour deux bassins emboîtés. Modifié d'après Skoien et al. (2006) © Clotaire Catalogne.....	9
Figure 5 : Exemple de distributions d'erreur d'estimation du QMNA5 par classe de corrélation, obtenues pour une stratégie de jaugeage simulant une fréquence annuelle de jaugeage $F = 2$ et une durée de suivi de $D = 4$ années (soit au total $N = 8$ jaugeages) © Clotaire Catalogne .....	11
Figure 6 : Incertitudes d'estimation (écart type des erreurs d'estimation en l/s/km <sup>2</sup> ) en fonction de la corrélation et du nombre de jaugeages pour les cinq fréquences de jaugeage simulées © Clotaire Catalogne .....	13
Figure 7 : Gains cumulés sur l'incertitude liés à l'augmentation du nombre de jaugeages pour les cinq fréquences de jaugeage simulées en supposant une corrélation constante. Les incertitudes fournies par les modèles pour des configurations non simulées (i.e. en extrapolation) sont représentées en pointillés © Clotaire Catalogne.....	15
Figure 8 : Nombre de jaugeages équivalent pour $F > 1$ en fonction du nombre de jaugeages réalisés pour $F = 1$ © Clotaire Catalogne.....	16
Figure 9 : Durée de suivi équivalente pour $F > 1$ en fonction de la durée de suivi pour $F = 1$ © Clotaire Catalogne .....	17
Figure 10 : 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 © Clotaire Catalogne.....	19
Figure 11 : Biais d'estimation en fonction de la corrélation et du nombre de jaugeages pour les cinq fréquences de jaugeage simulées © Clotaire Catalogne .....	20
Figure 12 : Comparaison des erreurs d'estimation du QMNA5 et des intervalles de confiance à 95 % associés obtenus en sites non jaugés (en blanc) et par valorisation de jaugeages épisodiques (gris) en simulant différentes fréquences annuelles de jaugeage et pour une durée de suivi croissante © Clotaire Catalogne .....	22
Figure 13 : Situation du bassin de l'Asse (délimité en rouge) dans le département des Alpes de Haute Provence © Daniel Dalet / d-maps.com .....	24
Figure 14 : Carte simplifiée des lithologies rencontrées sur le bassin de l'Asse. Figure modifiée d'après AERMC (2011) © AERMC .....	25
Figure 15 : Evolution des débits moyens mensuels interannuels © Clotaire Catalogne .....	26
Figure 16 : Comparaison des débits jaugés au site de la Julienne et des débits observés à la station de Chabrières à la même date © Clotaire Catalogne .....	27
Figure 17 : Illustration de la non simultanété des réponses hydrologiques aux sites de La Julienne et de Chabrières pour le jaugeage du 6 juillet 2006 © Clotaire Catalogne .....	28
Figure 18 : Données de jaugeage retenues après exclusion des données se rapportant à des périodes de fortes variations du débit à la station de Chabrières © Clotaire Catalogne .....	29
Figure 19 : Exemple de sélection de données de jaugeage relatives à des conditions hydrologiques contrastées © Clotaire Catalogne .....	29



<i>Figure 20 : Illustration des ajustements réalisés, des valeurs estimées (en retenant la période de référence 1970-2005) et des intervalles de confiance à 95 % associés au terme de chaque année de suivi © Clotaire Catalogne</i> .....	32
<i>Figure 21 : Illustration de l'évolution du QMNA5 estimé (en retenant la période de référence 1970-2005) et de l'intervalle de confiance à 95 % associé au cours de la période de suivi © Clotaire Catalogne</i> .....	33
<i>Figure 22 : Evolution de l'écart type au cours de la campagne de jaugeage (trait plein) et évolution prévue par les modèles pour une fréquence de 3 jaugeages par année en faisant l'hypothèse que le niveau de corrélation entre les débits jaugés à La Julienne et ceux observés à Chabrières reste constant (pointillés) © Clotaire Catalogne</i> .....	34
<i>Tableau 1 : Combinaisons de durée de suivi (D en année) et de fréquence de jaugeage annuelle (F) simulées en validation croisée et nombre total de jaugeages résultant (N) - Source des données : Catalogne (2012)</i> .....	8
<i>Tableau 2 : Paramètres des cinq modèles d'incertitude ajustés pour chaque fréquence de jaugeage simulée - Source des données : Catalogne (2012)</i> .....	14
<i>Tableau 3 : Paramètres et performances moyennes des modèles de biais ajustés sur les données en fonction de la corrélation pour les cinq fréquences annuelles de jaugeage simulées - Source des données : Catalogne (2012)</i> .....	21
<i>Tableau 4 : Dénombrement des jaugeages réalisés au site de la Julienne sur la période 2004-2010 - Source des données : DREAL PACA</i> .....	27
<i>Tableau 5 : Caractéristiques de la campagne de jaugeage retenue - Source des données : DREAL PACA</i> .....	30
<i>Tableau 6 : Détail des résultats d'estimation obtenus année par année - Source des données : DREAL PACA</i> .....	31



Irstea  
1, rue Pierre Gilles de Gennes  
CS 10030  
92761 Antony  
01 40 96 61 21  
[www.irstea.fr](http://www.irstea.fr)

Onema  
Hall C – Le Nadar  
5, square Félix Nadar  
94300 Vincennes  
01 45 14 36 00  
[www.onema.fr](http://www.onema.fr)