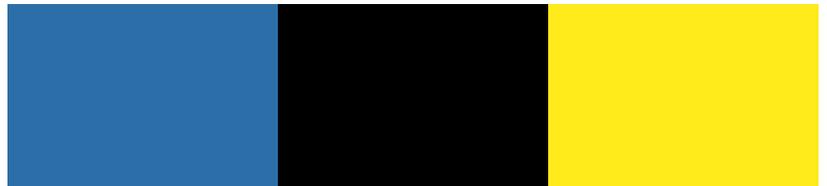


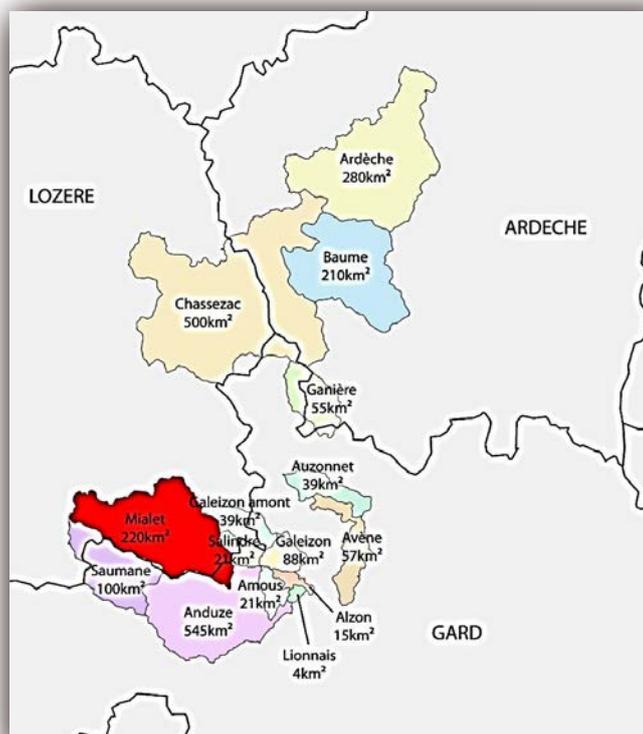


31 décembre 2013

Contributeurs : Anne Johannet,  
Guillaume Artigue, Pierre-Alain  
Ayrat, Délizar Ben Nessib, Audrey  
Bornancin-Plantier, Georges-Marie  
Saulnier, Yann Visserot.



Tâche 06 : Changement d'échelles  
spatiales et temporelles



# Sommaire

<b>1. INTRODUCTION</b> .....	<b>2</b>
<b>2. GENERALISATION AUX BASSINS DE LA CEZE ET DE L'ARDECHE</b> .....	<b>2</b>
<b>2.1 LE PRINCIPE DE LA GENERALISATION</b> .....	<b>2</b>
2.1.1 Bassin versant pilote et bassin versant cible.....	2
2.1.2 Le bassin versant pilote : le Gardon de Mialet .....	3
<b>2.2 CHANGEMENT D'ECHELLE TEMPORELLE</b> .....	<b>7</b>
<b>2.3 CHANGEMENT D'ECHELLE SPATIALE</b> .....	<b>7</b>
<b>2.4 LES BASSINS VERSANT CIBLES</b> .....	<b>8</b>
<b>2.5 PERFORMANCE DES GENERALISATIONS</b> .....	<b>10</b>
<b>2.6 CONCLUSION</b> .....	<b>11</b>
<b>3. GENERALISATION AU BASSIN VERSANT DE LA SOMME</b> .....	<b>12</b>
<b>3.1 PRESENTATION DU BASSIN VERSANT DE LA SOMME</b> .....	<b>12</b>
3.1.1 Géologie.....	12
3.1.2 Hydrologie .....	13
<b>3.2 DESCRIPTION D'UNE CRUE DE NAPPE</b> .....	<b>14</b>
<b>3.3 DYNAMIQUE D'UNE CRUE DE NAPPE</b> .....	<b>15</b>
3.3.1 Comparaison crue éclair/ crue de nappe .....	15
<b>3.4 CHOIX DES VARIABLES POUR LE MODELE</b> .....	<b>15</b>
3.4.1 Sélection des variables .....	15
3.4.2 Démarche suivie.....	17
<b>3.5 PREVISION PAR RESEAU DE NEURONES</b> .....	<b>19</b>
<b>3.6 CONCLUSION</b> .....	<b>21</b>
<b>4. GENERALISATION AU BASSIN VERSANT DE LA LIANE</b> .....	<b>21</b>
<b>4.1 PRESENTATION DU BASSIN VERSANT</b> .....	<b>21</b>
<b>4.2 BASE DE DONNEES HYDROMETEOROLOGIQUE</b> .....	<b>22</b>
<b>4.3 SYNTHESE DU PREDICTEUR</b> .....	<b>22</b>
<b>4.4 PERFORMANCES SUR L'ANNEE 2012</b> .....	<b>24</b>
<b>4.5 CONCLUSION</b> .....	<b>26</b>
<b>5. CONCLUSION GENERALE</b> .....	<b>26</b>
<b>6. BIBLIOGRAPHIE</b> .....	<b>27</b>

# 1. INTRODUCTION

Le projet ANR FLASH (Flood forecasting with machine Learning, data Assimilation and Semi-physical modeling) a pour objectif de compléter les modèles hydrologiques à base physique pour effectuer la prévision des crues ; il s'intéresse dans un premier temps à la modélisation de la chaîne hydrométéorologique (tâche T<sub>2</sub>) appliquée aux bassins rapides du Gardon d'Anduze, sujet à des crues dévastatrices. Les méthodes ainsi définies sont ensuite évaluées, pour mieux appréhender leur pertinence et leur caractère générique, sur d'autres bassins de comportement semblables, mais de dimension très différentes (10 fois plus petits à 6 fois plus grands). Il est aussi intéressant d'étudier la capacité des méthodes définies à être transposées à des bassins très différents, hydrologiquement parlant, en particulier les bassins sujets à des crues de nappe très lentes.

C'est cette généralisation à des bassins de caractéristiques temporelles et spatiales très différentes qui est l'objet de la tâche 6 du projet FLASH, et le sujet du présent rapport. Dans la proposition du projet FLASH, la tâche 6 comportait 2 sous-tâches :

- T<sub>6.1</sub> Généralisation aux bassins de la Cèze et de l'Ardèche, de même dynamique temporelle et mêmes types de précipitations, mais de superficies très différentes
- T<sub>6.2</sub> Généralisation au bassin de la Somme de superficie et de comportement hydrologique sans relation avec les bassins cévenols

Le déroulement du projet nous a amenés à nous intéresser à un autre bassin, qui peut ainsi constituer une troisième sous-tâche de la tâche 6 :

- T<sub>6.3</sub> Généralisation au bassin de la Liane, bassin de dimension moitié de celui d'Anduze, mais de dynamique bien plus lente (6 fois plus lente)

Les études et développements effectués pour chacune des sous-tâches sont présentés de manière synthétique ci-dessous en conservant la proposition initiale du projet. Néanmoins des renvois sont effectués à d'autres rapports si la présentation détaillée de l'étude peut être envisagée avec plus de cohérence par ce moyen.

## 2. GENERALISATION AUX BASSINS DE LA CEZE ET DE L'ARDECHE

### 2.1 LE PRINCIPE DE LA GENERALISATION

#### 2.1.1 Bassin versant pilote et bassin versant cible

Dans le rapport de la tâche T<sub>2</sub>, (Johannet, 2013) la généralisation de méthodes de conception de modèles de prévision fondés sur les réseaux de neurones, d'un bassin plus petit à un plus grand bassin le contenant, a été présentée. Si cette méthode développée pour le bassin d'Anduze (543 km<sup>2</sup>) et généralisée au bassin du Gardon à Remoulins (1848 km<sup>2</sup>) (Bornancin-Plantier, 2013) a présenté des résultats positifs, elle ne correspondait pas à ce que l'on entend traditionnellement par un « changement d'échelle » ; c'est donc ce concept très novateur pour les réseaux de neurones que Guillaume Artigue (Artigue, 2012) a entrepris d'explorer sur un ensemble de bassins choisis dans la bordure cévenole.

L'idée est de modéliser le comportement du bassin versant représentatif d'un comportement particulier ou bassin versant pilote, pour notre étude, les crues éclair

cévenoles, de normaliser le comportement de ce bassin, tant en temporel qu'en spatial, puis d'appliquer ce modèle à un autre bassin cible grâce à une dénormalisation (Figure 1).

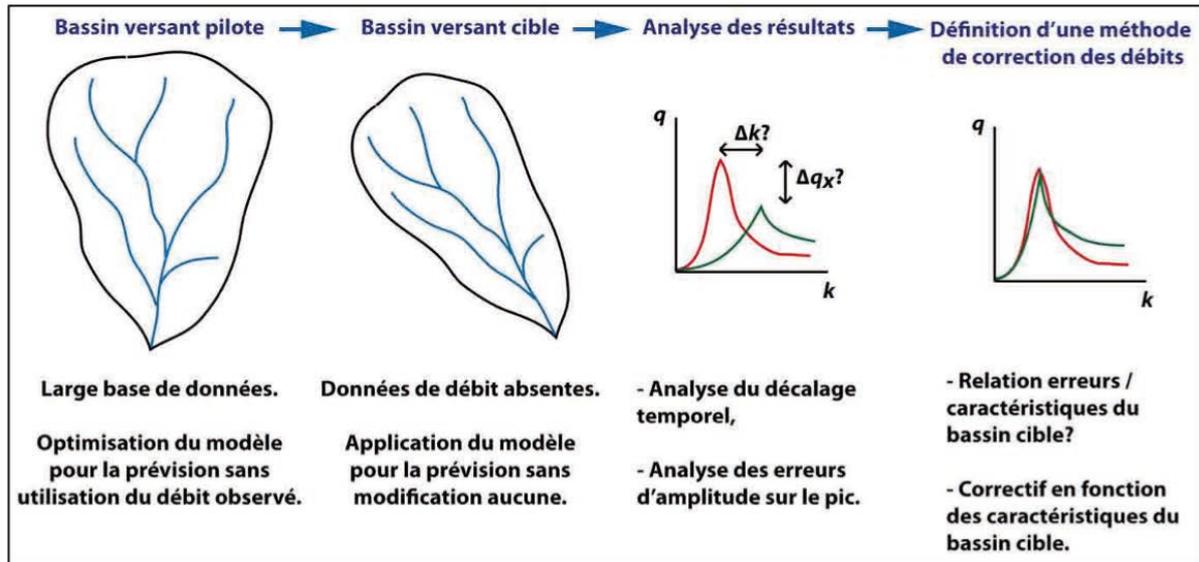


Figure 1 : Méthode de généralisation (Artigue, 2012)

Cette méthode comporte donc 3 étapes : (i) le choix et la modélisation du bassin versant pilote, (ii) la définition des méthodes de normalisation, (iii) la définition des méthodes de dénormalisation.

### 2.1.2 Le bassin versant pilote : le Gardon de Mialet

Guillaume Artigue a choisi comme bassin versant pilote le Gardon de Mialet à Mialet, bassin compris dans le bassin versant d'Anduze, de superficie  $220 \text{ km}^2$ , de pente moyenne du drain 3% et de pente moyenne des versants 33%. Il s'agit donc d'un bassin versant montagneux très pentu (Figure 2, Tableau 4).

Le parti pris de la modélisation a été de réaliser un modèle de type perceptron multicouches récurrent, ainsi il effectue son apprentissage et ses prévisions sans références externes. Ce type de modèle a de plus l'avantage de pouvoir être appliqué aux bassins versants dits « non-jaugés », c'est à dire pour lesquels on ne dispose pas de station de mesure. Le modèle intègre des liens linéaires comme présenté dans le rapport de la tâche T<sub>2</sub> (Johannet, 2013).

Le modèle ainsi défini est représenté en Figure 3. Chaque fenêtre temporelle est définie indépendamment des autres par la validation croisée partielle (Johannet, 2013).

Les notations sont les suivantes :

- $k$  est le temps discrétisé,
- $p_1$  correspond à la pluie à Mialet,  $p_2$  à Saint-Roman-de-Tousque,  $p_3$  à Barre-des-Cévennes,  $p_4$  au cumul moyen depuis le début de l'épisode (depuis son début jusqu'à l'instant de prévision),
- $w_i$  correspond à la dimension de la fenêtre temporelle d'entrée (*window*),
- $h$  correspond au nombre de neurones présents sur la couche cachée (*hidden layer*),
- $r$  correspond à l'ordre du modèle récurrent,
- $l$  correspond à l'horizon de prévision visé (*lead-time*),

- $s(k+l)$  est la valeur de la fonction de sortie estimée par le modèle.

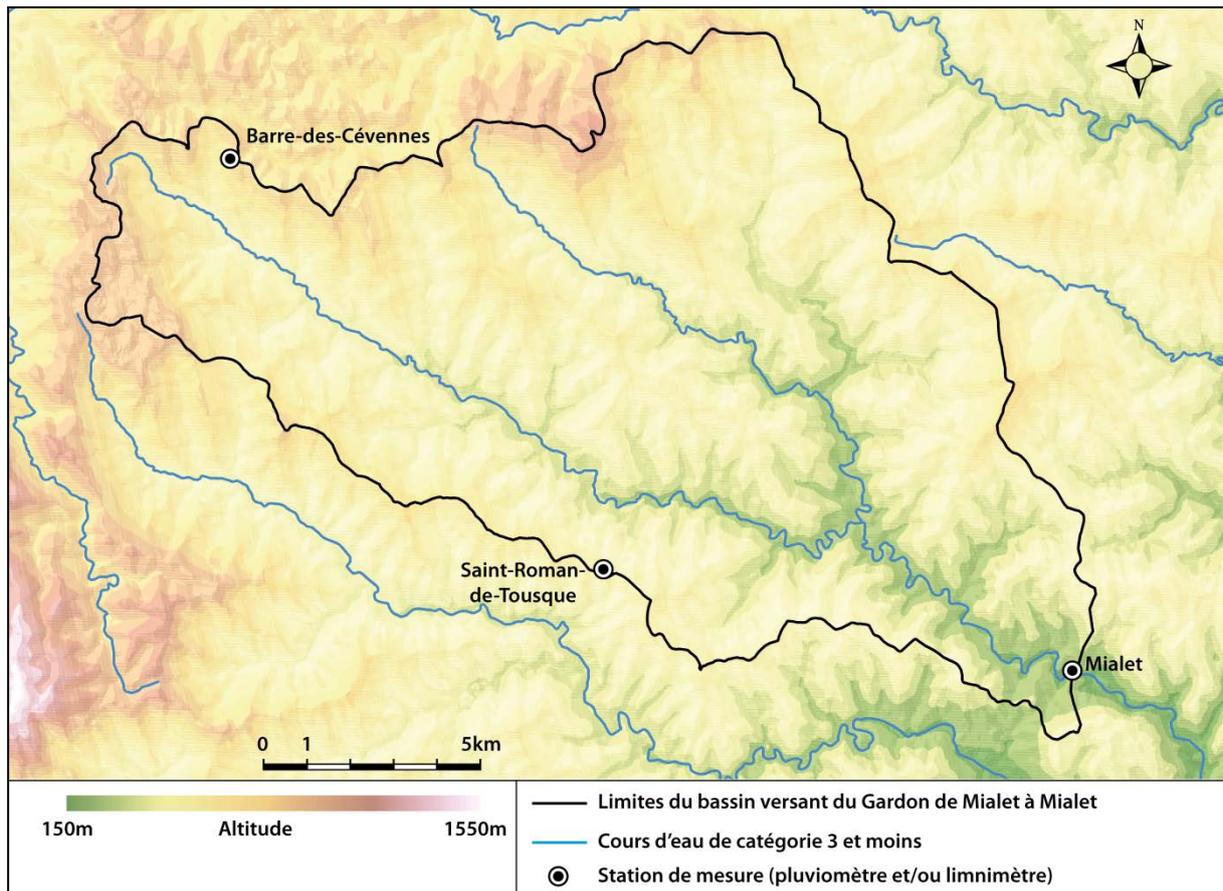


Figure 2 : Carte topographique et hydrographique simplifiée du bassin versant du Gardon de Mialet à Mialet. Altitudes d'après ASTER GDEM, un produit de METII et NASA ; Hydrographie d'après IGN BD CARTHAGE (Artigue, 2012)

La méthodologie de conception du modèle est celle présentée dans le rapport de la tâche T<sub>2</sub> (Johannet, 2013). Elle utilise la validation croisée partielle pour sélectionner les variables et la complexité. Trois ensembles sont définis : les ensembles d'arrêt, de test et d'apprentissage. La base de données extraite par Guillaume Artigue comporte 58 événements, elle est amplement décrite et discutée dans son mémoire de thèse (Artigue, 2012).

L'évènement n°222 de janvier 1998 est utilisé comme ensemble d'arrêt.

Les évènements choisis pour effectuer le test pourraient être n'importe lesquels dans la base mais nous avons choisi de tester les modèles obtenus dans quatre configurations hydrologiques différentes :

- l'évènement majeur de la base (n°19 de septembre 2002) afin de prouver que le modèle neuronal est capable de généraliser à des évènements plus intenses que ceux de sa base d'apprentissage,
- un évènement simple et intense (le n°13 de septembre 2000),
- un évènement complexe, à double pic (le n°8 de novembre 1996),
- et enfin, un évènement plus faible (le n°302 de novembre 2003).

Leurs caractéristiques sont synthétisées dans le (Tableau 2).

Tous les autres évènements constituent l'ensemble d'apprentissage.

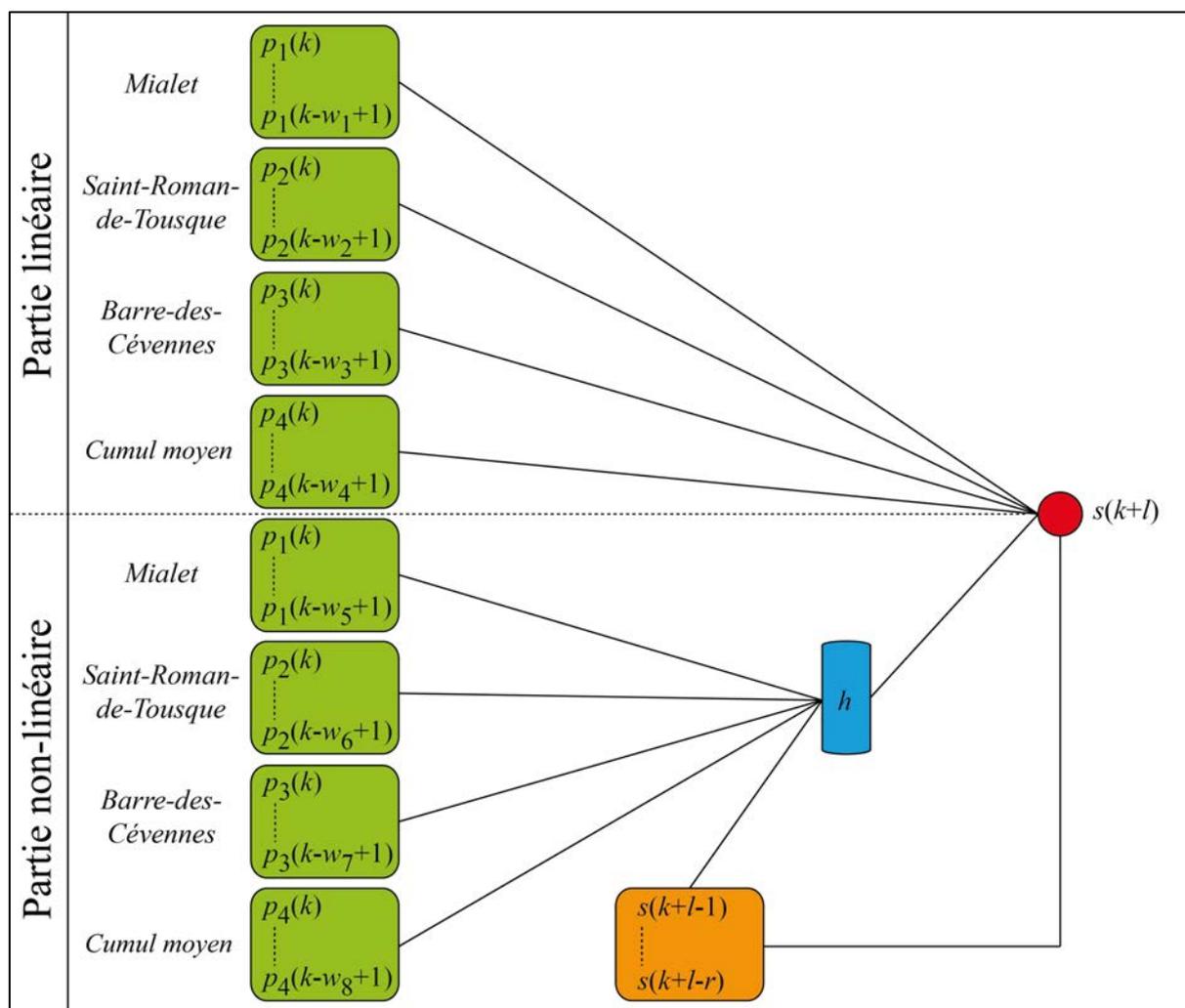


Figure 3 : Modèle récurrent utilisé dans la prévision des crues du Gardon de Mialet (Artigue, 2012)

Elément du modèle	$k+1$	$k+2$	$k+3$	$k+4$
Mialet / $w_1$	2	4	4	5
Saint-Roman-de-Tousque / $w_2$	9	8	8	8
Barre-des-Cévennes / $w_3$	9	9	9	8
Cumul moyen / $w_4$	5	3	3	3
Mialet / $w_5$	6	4	5	4
Saint-Roman-de-Tousque $w_6$	7	7	6	7
Barre-des-Cévennes / $w_7$	10	10	10	9
Cumul moyen / $w_8$	1	1	1	1
Couche cachée / $h$	3	2	2	2
Débits précédents ou ordre / $w_9$ ou $r$	1	2	2	2

Tableau 1 : Architectures optimales pour les deux types de modèles utilisés. Les valeurs sont indiquées en nombre de pas de temps (soit en 0,5h).

N°	Date	Pics	$Q_x (m^3 \cdot s^{-1})$	$Q_m (m^3 \cdot s^{-1})$	RR (mm)	$I_x (mm/h)$
13	Sep. 2000	1	454	70	230	40
19	Sep. 2002	1	819	116	235	82
8	Nov. 1996	2	269	47	222	26
302	Nov. 2003	1	200	67	155	16

**Tableau 2 : Quelques caractéristiques des événements choisis pour le test. Pics : nombre de pics ;  $Q_x$  : débit de pointe ;  $Q_m$  : débit moyen ; RR : cumul total de pluie ;  $I_x$  : intensité pluvieuse maximale (Artigue, 2012)**

La qualité des prévisions peut être évaluée grâce aux critères définis dans le rapport de la tâche 2 (Johannet, 2013). Ils sont synthétisés dans le Tableau 3.

Evt.	$l (0.5 h)$	Nash	PPD	SPPD	$V_p$	$V_o$	$D(0.5 h)$
13	1	0.94	84%	81%	O	O	1
13	2	0.95	83%	79%	O	O	1
13	3	0.93	79%	79%	J	O	1
13	4	0.93	77%	75%	J	O	1
19	1	0.85	67%	63%	O	R	-1
19	2	0.92	72%	69%	O	R	-1
19	3	0.91	74%	73%	R	R	-1
19	4	0.91	83%	83%	R	R	0
8	1	0.8	70%	70%	J	J	0
8	2	0.78	62%	62%	J	J	0
8	3	0.77	67%	63%	J	J	1
8	4	0.81	65%	59%	J	J	2
302	1	0.63	64%	64%	J	J	0
302	2	0.64	70%	70%	J	J	1
302	3	0.57	72%	72%	J	J	0
302	4	0.57	65%	63%	J	J	1
<b>Moyenne</b>		0.81	72%	70%	-	-	11min
<b>Ecart-type</b>		0.14	7%	8%	-	-	24min

**Tableau 3 : Récapitulatif des résultats obtenus avec le modèle récurrent (Evt. correspond au numéro de l'évènement,  $l$  à l'horizon de prévision,  $V_p$  à la vigilance prévue par le modèle,  $V_o$  à la vigilance correspondant au débit observé à  $k+l$  et  $D$  au décalage temporel) (Artigue, 2012)**

D'une manière générale, 70% du pic est prévu par le modèle à plus ou moins une demi-heure de décalage. La sous-estimation varie de 41% à 17% pour le *SPPD* avec une moyenne de 30% et un écart-type de moins de 8%. Malgré cette imprécision, la relative constance dans l'erreur peut permettre au prévisionniste d'adapter sa décision en conditions opérationnelles. De plus, le modèle présente un décalage temporel avec l'observation faible : le modèle récurrent représente mieux la dynamique du bassin que le modèle non-récurrent conçu dans la tâche 2. Les hydrogrammes associés sont présentés en annexe 2 (Artigue, 2011).

Partant de ce modèle il paraît possible d'envisager une généralisation à d'autres bassins de la bordure cévenole présentés dans le (Tableau 4), auparavant les changements d'échelles spatiales et temporelles doivent être étudiées.

Pour ce faire, le modèle récurrent du bassin de Mialet a été utilisé « tel quel » pour prévoir les crues des 14 bassins cibles et de leurs 35 événements de  $k+1$  à  $k+4$  (0,5h à 2h). Loin de penser que ce premier essai peut déboucher sur de bons

résultats avec des bassins si différents, tant du bassin du Gardon à Mialet qu'entre eux, nous cherchons surtout à nous renseigner sur la nature et l'ampleur des erreurs générées lors du test sur des bassins et des épisodes parfois très différents. Il s'agit donc d'identifier les problèmes induits par une généralisation « brute » dans le but de trouver un moyen de les résoudre.

On s'attend à deux réactions de la part du modèle :

- D'abord, on peut attendre un décalage temporel, très probablement en relation avec les caractéristiques topographiques des bassins déterminées et exposées dans (Tableau 4). Etant donné que les variables du modèle ont été sélectionnées en fonction des corrélations croisées sur trois pluviomètres disposés d'une manière spécifique sur un bassin différent, il semble au premier abord improbable que le modèle présente une dynamique cohérente avec celle du bassin testé.
- On peut aussi attendre des sur-, ou sous-estimations des débits. En effet, même si la superficie du bassin a été prise en compte au travers de la sortie qui est le débit spécifique, il est également connu que le débit spécifique n'évolue pas de manière proportionnelle à la surface du bassin. Il est donc probable qu'il faille pallier cette dernière source d'erreur.

## **2.2 CHANGEMENT D'ECHELLE TEMPORELLE**

Concernant le changement d'échelle temporelle, l'application de précipitations mesurées sur les bassins cibles a conduit à des prévisions dont l'avance ou le retard par rapport à la valeur observée a pu être quantifié.

Malheureusement aucune relation n'a pu être mise en évidence de manière à corriger ce décalage temporel qui paraît ne pouvoir être relié : ni à la pente du drain, ni à la surface des bassins, ni à la longueur du drain, ni à la dénivelée du bassin, ni à la pente des versants du bassin cible. Aucune compensation temporelle n'a donc pu être proposée.

## **2.3 CHANGEMENT D'ECHELLE SPATIALE**

Concernant le changement d'échelles spatiales, il est proposé d'étudier la relation entre le débit de pointe observé sur un bassin et sa superficie. Cette relation a déjà été étudiée en hydrologie. Elle est connue comme spécifique du bassin étudié. Il a donc fallu refaire cette étude pour notre base de données. La première étape de cette recherche de relation concernant les débits et les surfaces est de travailler en « débit spécifique », c'est à dire de réduire le débit du cours d'eau par la superficie de son bassin.

On peut alors représenter sur un même diagramme l'ensemble des débits spécifiques de pointe des événements disponibles sur les bassins cibles (Figure 4).

Malgré un nombre limité d'évènements, on constate qu'une tendance nette se dessine. L'incertitude porte surtout sur des débits spécifiques extrêmes dans les bassins de petite taille mais en première approximation la régression paraît assez efficace.

L'équation de la courbe de régression de la figure 66 est :

$$q = 49,09A^{-0,386}$$

Le coefficient de corrélation de la régression avec les valeurs observées est de 0,61.

Grâce à cette tendance, on peut évaluer le débit spécifique maximal théorique du Gardon à Mialet. D'après la relation établie, il serait d'environ  $6,3 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ , ce qui, au regard des valeurs rencontrées sur les bassins voisins, dont certains bassins cibles, semble être une valeur cohérente.

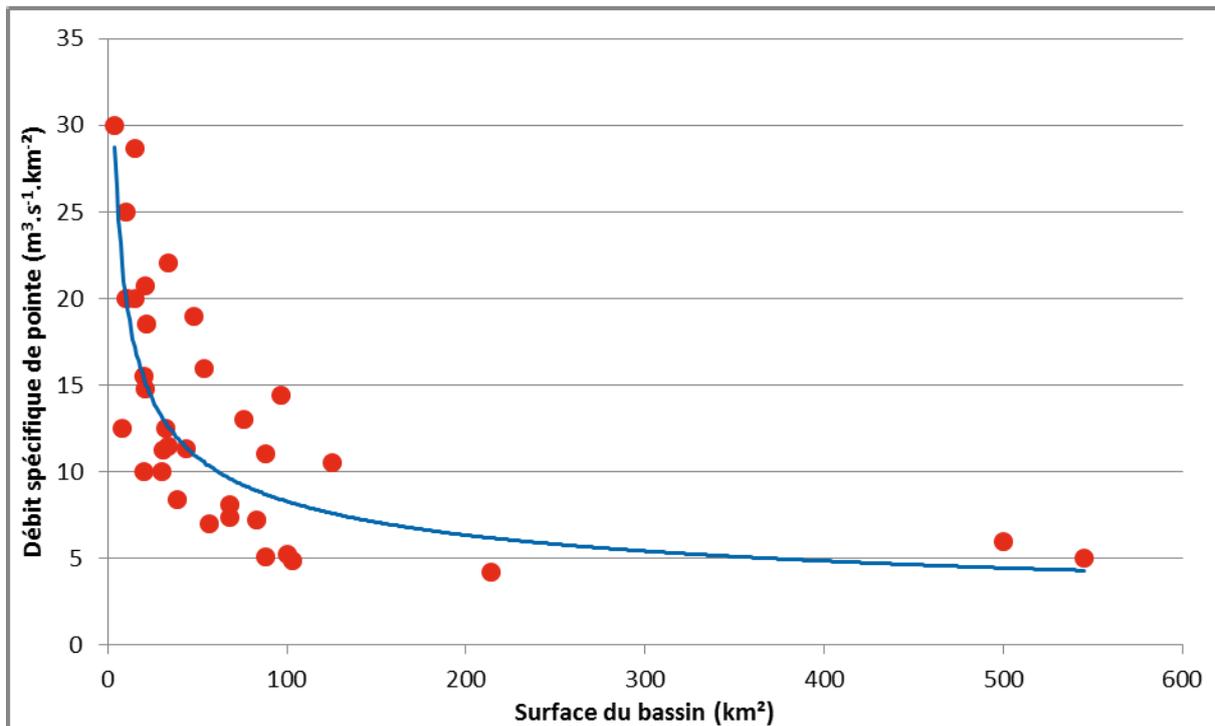


Figure 4 : Représentation du débit spécifique de pointe en fonction de la surface des bassins sur lesquels il a été observé (points rouges) et courbe de régression de type puissance (courbe bleue).

Ainsi, à l'aide de cette formule empirique et de la nouvelle valeur établie, nous allons pouvoir déterminer ce coefficient. Pour ce faire, nous devons rendre adimensionnelle la relation établie précédemment. Ainsi, nous divisons les deux variables de la relation par les valeurs associées au bassin pilote :

- Le débit spécifique théorique du bassin cible sera divisé par 6,3, la valeur théorique du débit spécifique de pointe du Gardon à Mialet.
- La surface du bassin cible sera divisée par 220, la valeur en  $\text{km}^2$  du bassin du Gardon à Mialet. Cette manipulation conduit à une nouvelle formule de régression.

C'est à partir de cette nouvelle formule empirique que nous sommes en mesure, théoriquement, pour tout bassin cible de surface et de localisation cohérentes avec celles utilisées pour son établissement, de fournir un coefficient correctif du débit sortant du modèle ayant réalisé son apprentissage sur le bassin pilote.

## 2.4 LES BASSINS VERSANT CIBLES

Les bassins versants cibles ont été choisis en fonction de leur comportement hydrologique et des mesures disponibles. Leurs caractéristiques sont présentées dans le Tableau 4.

Numéro	Bassin	Surface (km <sup>2</sup> )	Pente du drain	Longueur du drain (km)	Dénivelée du drain (m)	Pente des versants	Dénivelée du bassin (m)
1	Ruisseau du Lyonnais	4	2,80%	2	56	18%	203
2	Alzon	15	3,20%	8	144	24%	412
3	Salandre	21	6,30%	10.2	644	33%	694
4	Amous	21	3,80%	10	375	24%	503
5	Auzonnet	39	2,40%	19	450	25%	560
6	Galeizon amont	39	3,20%	18.8	597	29%	697
7	Ganière	55	3,70%	18.5	687	30%	807
8	Avène	57	2,30%	18	413	12%	550
9	Galeizon	88	2,30%	28.8	654	30%	765
10	Gardon à Saumane	100	3,70%	19.8	100	3,70%	19.8
11	Baume	210	3,80%	30	1132	35%	1352
<b>12</b>	<b>Gardon à Mialet</b>	<b>220</b>	<b>3,00%</b>	<b>29</b>	<b>864</b>	<b>33%</b>	<b>993</b>
13	Ardèche	280	4,10%	27.6	1129	35%	1222
14	Chassezac	500	2,30%	54.7	1263	28%	1505
15	Gardon à Anduze	545	2,30%	41	545	2,30%	41

**Tableau 4 : Sélection et description des bassins retenus**

Les coefficients correctifs pour effectuer le changement d'échelle entre le bassin de Mialet et les bassins cibles sont donnés dans le Tableau 5

<i>Bassin</i>	<i>Surface (km<sup>2</sup>)</i>	<i>b</i>
<i>Ruisseau du Lyonnais</i>	4	7,67
<i>Alzon</i>	15	4,60
<i>Salandre</i>	21	4,04
<i>Amous</i>	21	4,04
<i>Galeizon amont</i>	39	3,18
<i>Auzonnet</i>	49	2,92
<i>Ganière</i>	55	2,79
<i>Avène</i>	57	2,75
<i>Galeizon</i>	88	2,33
<i>Gardon à Saumane</i>	100	2,21
<i>Baume</i>	210	1,66

<i>Ardèche</i>	280	1,49
<i>Chassezac</i>	500	1,19
<i>Gardon à Anduze</i>	545	1,15

Tableau 5 : Coefficient correctif correspondant à chaque bassin visé par l'étude

Pour prendre en compte la correction envisagée, il suffit d'appliquer ces coefficients correctifs aux débits spécifiques prévus par le modèle sans autre modification.

## 2.5 PERFORMANCE DES GENERALISATIONS

Pour chaque bassin, tous les évènements ont été testés avec le modèle récurrent calé sur le bassin pilote et la correction des débits sortants a été effectuée à l'aide des coefficients déterminés dans la section précédente. Les résultats sont évalués, comme défini dans le rapport 2, sur le critère de Nash, de persistance, le *SPPD* et le *PPD* ainsi que sur le décalage temporel. L'ensemble des hydrogrammes de cette phase de test est présenté dans le mémoire de thèse de G. Artigue (Artigue, 2012b). A chaque fois, la prévision brute, sans coefficient correctif, est présentée en plus de la prévision traitée par le coefficient correctif et de l'observation ou de l'estimation du débit réalisée *a posteriori* par un retour d'expérience. Lorsque cela est nécessaire, des indications sur la substitution des pluviomètres du bassin pilote par ceux du bassin cible sont fournies. Quand ce n'est pas le cas, c'est qu'un seul vecteur de pluie est disponible : il est donc appliqué aux trois.

De manière synthétique on peut présenter le tableau récapitulatif de ces résultats (Tableau 6).

A la lumière des validations réalisées dans ce chapitre, nous pouvons définir trois types d'évènements :

- ceux se produisant sur des bassins de moins de 30km<sup>2</sup>,
- ceux où les données de pluie ne sont manifestement pas représentatives des valeurs réellement tombées,
- ceux où les données de pluie sont représentatives.

Si l'on s'en tient à une analyse qualitative des résultats, on peut juger de l'exploitabilité opérationnelle des résultats obtenus en fonction de l'apport de l'anticipation du débit par rapport à l'observation.

Ainsi, si l'on prend en compte l'ensemble des évènements et horizons de prévision, soit 136 prévisions, l'on peut considérer que les prévisions sont exploitables pour 52% d'entre eux. Si l'on ne prend en compte que les 112 prévisions hors Galeizon et bassins de moins de 30km<sup>2</sup>, l'on peut considérer que 63% des évènements ont donné lieu à des prévisions exploitables. Enfin, si l'on ne prend en compte que les 96 évènements et horizons se produisant sur des bassins de plus 30km<sup>2</sup> et disposant de pluies représentatives, ce que nous recommandons pour l'utilisation du modèle, 74% des prévisions fournies sont exploitables.

De façon plus quantitative, comme dans le chapitre 4, nous privilégions le *SPPD* pour fournir une estimation de la qualité des résultats. Le Tableau 6 présente d'une part les *SPPD* moyens et médians et d'autre part le taux d'évènements atteignant entre 70 et 130% pour les différentes catégories de bassins.

Comme on le constate aisément, sur les bassins de plus de 30km<sup>2</sup> avec des pluies représentatives, les *SPPD* sont la plupart du temps compris entre 85 et 90% avec

plus de 80% des événements dont la prévision atteint entre 70 et 130%. D'une manière générale, les *SPPD* moyens et médians sont proches de ou supérieurs à 80% alors qu'au pire, 60% des événements donnent lieu à des prévisions de 70 à 130% du pic au moment où il survient. Les prévisions obtenues sont donc très satisfaisantes compte-tenu de l'incertitude qui est couramment admise sur les événements de crue rapides (30% sur le pic de crue).

La qualité de ces prévisions permet d'envisager d'effectuer des prévisions sur les bassins versants non jaugés, c'est à dire non équipés de station de mesure. Ces bassins existent, ils concernent surtout les bassins de petite taille.

<b>SPPD sur tous les événements</b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
<i>SPPD</i> moyen	81%	76%	77%	76%	78%
<i>SPPD</i> médian	79%	80%	80%	80%	80%
Taux d'évènements entre 70 et 130%	59%	59%	65%	59%	60%
<b>SPPD sur les bassins de plus de 30km<sup>2</sup></b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
<i>SPPD</i> moyen	86%	80%	82%	82%	83%
<i>SPPD</i> médian	84%	83%	84%	81%	83%
Taux d'évènements entre 70 et 130%	67%	64%	70%	64%	66%
<b>SPPD sur tous les événements avec données de pluie représentatives</b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
<i>SPPD</i> moyen	88%	81%	82%	80%	83%
<i>SPPD</i> médian	85%	84%	85%	82%	84%
Taux d'évènements entre 70 et 130%	71%	71%	76%	71%	72%
<b>SPPD sur les bassins de plus de 30km<sup>2</sup> avec données de pluie représentatives</b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
<i>SPPD</i> moyen	94%	87%	87%	87%	89%
<i>SPPD</i> médian	87%	84%	86%	86%	86%
Taux d'évènements entre 70 et 130%	80%	80%	88%	80%	82%

Tableau 6 : *SPPD* moyens et médians et taux d'atteinte de *SPPD* compris entre 70 et 130% pour différents types d'évènements et de bassins, décrits dans le tableau

## 2.6 CONCLUSION

Les principales limites de la méthode appliquée pour la généralisation sont les suivantes :

- Bien qu'important, le nombre d'évènements utilisés pour élaborer le coefficient correctif ne pallie pas les incertitudes fortes, en particulier sur les petits bassins où la densité de valeurs incertaines est plus élevée.

- Pour différentes raisons, les bassins de moins de 30km<sup>2</sup> environ ne donnent pas des résultats fiables, même s'ils apportent un début d'information sur la crue générée ; ce défaut provient vraisemblablement du type de données utilisée : la mesure pluviométrique issue des RADARs météorologiques ; or nous avons vu dans le rapport de la tâche T<sub>2</sub> (Johannet, 2013) que les estimations de pluies sont systématiquement sous estimées par les RADARs (parfois d'un facteur 2). Ceci peut donc expliquer les sous-estimations des crues sur les petits bassins.

- Les bassins ayant des particularités ne peuvent que difficilement être appréhendés par le modèle. C'est le cas des bassins avec barrages écrêteurs, puissants karsts, vastes zones urbaines ou encore régimes hydrologiques trop différents du bassin cible. Cet élément n'est pas surprenant mais il mérite d'être souligné pour délimiter au mieux le domaine de validité de la méthode ici développée.

### 3. GENERALISATION AU BASSIN VERSANT DE LA SOMME

#### 3.1 PRESENTATION DU BASSIN VERSANT DE LA SOMME

##### 3.1.1 Géologie

Le bassin versant du fleuve Somme est assez proche par ses limites du département de la Somme. Une fraction du département de la Somme est drainée par d'autres fleuves (l'Authie et la Bresle). Inversement, la Somme a, au préalable, traversé une partie du département de l'Aisne et certains de ses affluents viennent des départements de l'Aisne ou de l'Oise (notamment la Noye et la Selle). La surface du bassin versant est de 5560 km<sup>2</sup> tandis que la surface du département est de 6170 km<sup>2</sup>. Le bassin versant de la Somme est donc 10 fois plus vaste que celui d'Anduze.



Figure 5 : Bassin versant de la Somme (AMEVA, 2013)

La caractéristique principale de cette région hydrogéologique est d'avoir une structure géologique simple : un niveau d'argile (faciès Gault) daté de l'albien fournit

une base imperméable à la masse de craies datées à la base du turonien et en partie supérieure du sénonien.

### 3.1.2 Hydrologie

La masse crayeuse est perméable et elle est le siège d'une infiltration générale et d'un écoulement régulier en direction des quelques rivières qui sont les drains naturels de la région. En tout point et en tout lieu, la nappe est en charge par rapport à la Somme, ce qui signifie que l'eau de la nappe va toujours de la base des plateaux vers la Somme (Figure 6). Il en est de même pour ses affluents. L'hydrologie de la Somme est donc principalement influencée par la nappe d'eau souterraine contenue dans les craies. Les précipitations rechargent cette nappe au travers d'un réseau de fissures. La dynamique de cette recharge peut varier considérablement selon que ces fissures sont saturées ou pas d'eau (infiltration diphasique ; transition *zone saturée/zone non saturée*). Ainsi, de même que pour les crues rapides, le niveau d'humidité du bassin est prépondérant dans la dynamique hydrologique, mais insuffisamment estimé par les mesures de débit ou de piézométrie : il ne suffit de connaître le niveau de la nappe, il faudrait également connaître le quotient de saturation en eau du réseau fissural.

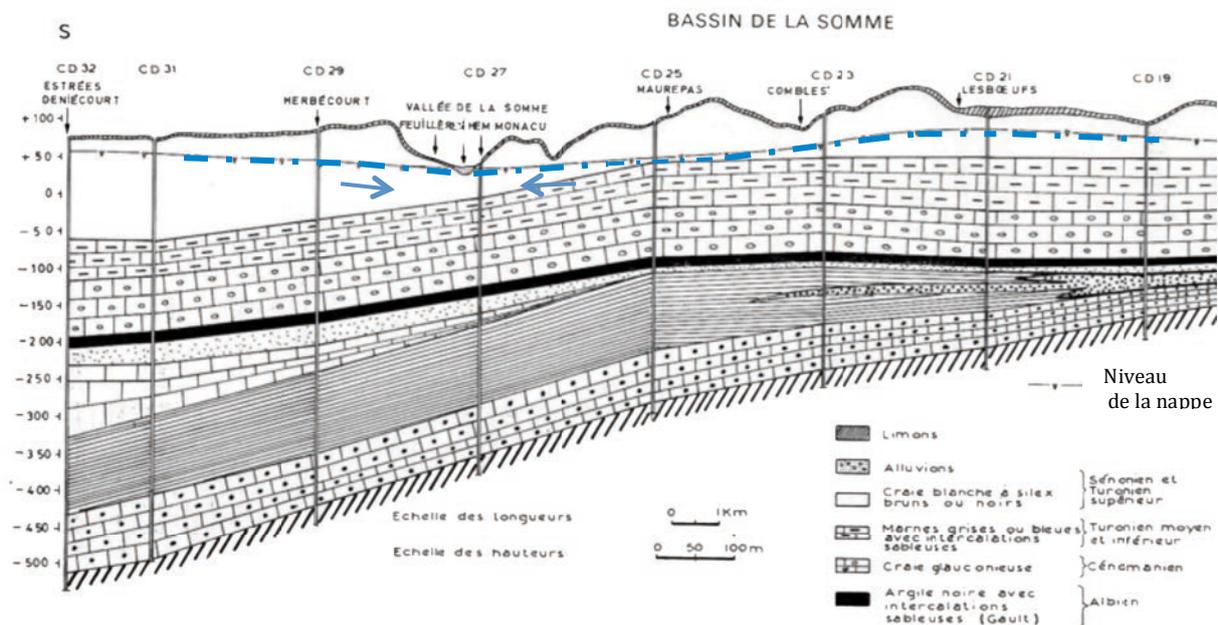


Figure 6 : Coupe Nord Sud de la vallée de la Somme (extrait de (Roux, 1978), modifié)

La partie aval de la Somme a un fonctionnement très anthropisé puisque son cours est complètement canalisé et comporte près de 25 ouvrages artificiels rendant son comportement difficile à modéliser (Figure 7).

Enfin, le fleuve Somme et de sa nappe sont en connexion hydraulique avec la Manche. Le déversement de la Somme dans la Manche est également influencé par des écluses qui empêchent la mer de remonter le cours d'eau en marée haute. En dépit des écluses, l'influence de la marée est perceptible dans le débit de la Somme jusqu'à Abbeville.



Figure 7 : Le canal de la Somme comprend 25 écluses (AMEVA, 2013)

Tous ces éléments, qu'ils soient partiellement connus comme les écoulements diphasiques et l'influence maritime ou que ce soit l'artificialisation du cours d'eau, rendent les modélisations de la Somme très difficiles à réaliser. Comme pour les crues éclair cévenoles, il n'existe pas de modèle physique satisfaisant permettant la modélisation hydrodynamique de la Somme ; les modèles sans a priori comme les modèles statistiques ont donc toute leur justification.

### 3.2 DESCRIPTION D'UNE CRUE DE NAPPE

La Somme est souvent décrite comme un fleuve régulier et les chiffres le confirment : à Abbeville, la crue décennale se situe à  $68 \text{ m}^3/\text{s}$  tandis que l'étiage quinquennal est à  $16 \text{ m}^3/\text{s}$ , soit 4,2 fois moins que la crue, pour la période 1963-1993. A titre de comparaison, ce rapport est d'environ 1000 pour la Cèze à La Roque sur Cèze (Collectif, 1994).

Néanmoins, la Somme est sujette à des crues, par exemple celle de 2001. Le débit de la Somme provient en grande partie de la nappe (Mardhel, 2001). S'il advient une succession d'années exceptionnellement pluvieuses, la nappe voit son niveau monter et le débit de la Somme croît jusqu'à être un débit de crue. La crue de 2001 en est un exemple : la pluviométrie annuelle moyenne à Abbeville est de 747 mm ; elle a été de 800 mm en 1998, 942 mm en 1999 et 1146 mm en 2000 (Vinet, 2010). Ceci a entraîné des remontées de nappe et, par voie de conséquence l'apparition d'une crue que l'on peut qualifier d'exceptionnelle : le débit de la Somme a alors atteint  $104 \text{ m}^3/\text{s}$ .

L'inertie du système crayeux vis-à-vis des écoulements est grande : on voit qu'il a fallu la succession de trois années humides pour avoir ces crues. C'est un premier point qu'il faut avoir présent à l'esprit. Un second point à souligner est que

l'information de crue est à chercher dans les eaux souterraines puisque c'est bien leur niveau qui est cause du débordement.

### **3.3 DYNAMIQUE D'UNE CRUE DE NAPPE**

#### **3.3.1 Comparaison crue éclair/ crue de nappe**

Une crue éclair se déroule dans une courte période de temps (en général inférieure à la journée). Cette période montre que, vu les vitesses des eaux souterraines, celles-ci ne participent que pour une faible part dans l'augmentation du débit des rivières. Pour une crue éclair, l'information ne provient donc pas du sous-sol. Au contraire, pour une crue de nappe, le système souterrain restitue des eaux infiltrées depuis au moins plusieurs mois en même temps que des pluies superficielles se concentrent. Une partie de l'information, en l'absence de modèle physique validé est donc à chercher dans les piézomètres qui mesurent les niveaux d'eau souterrains.

### **3.4 CHOIX DES VARIABLES POUR LE MODELE**

La tâche 6 du projet FLASH vise à évaluer la possibilité de transposer les méthodes développées sur les crues rapides aux crues lentes de nappe. Ainsi, l'information serait fournie par des piézomètres (pour les eaux souterraines provenant des pluies des mois ou des années passées) et des pluviomètres fourniraient les informations récentes (ruissellement et infiltration à faible profondeur suivie d'une exsurgence).

#### **3.4.1 Sélection des variables**

Comme pour tous les modèles à apprentissage statistique, la sélection des variables est cruciale. Celle-ci s'opère en deux étapes : d'une part le choix des variables physiques (stations de mesures) de manière experte, et d'autre part le dimensionnement fin des fenêtres temporelles appliquées au modèle qui sera déterminé comme décrit dans le rapport de la tâche 2 (Johannet, 2013) par la validation croisée partielle.

#### **Choix des piézomètres :**

Le département de la Somme dispose de 68 piézomètres dans la banque HYDRO qui peuvent être mis à profit pour anticiper sur l'occurrence de crues (Figure 8). Le travail présenté dans la suite est issu du stage de Master 2 de Délizar Ben Nessib effectué au LGEI (partenaire 1).



Figure 8 : Localisation des piézomètres dans le département de la Somme (ADES IGN 2005, BD CARTO)

La sélection de variables a donc consisté en un premier tri pour identifier les piézomètres associés à une série de données suffisamment longue. Un second choix a pris en compte une répartition assez régulière dans l'espace pour bien représenter les variations spatiales de la nappe. De plus un site particulier (vallée de l'Hallue) a été identifié par le BRGM comme représentatif de la transition *zone saturée/zone non saturée* ; le piézomètre de Senlis-le-Sec, proche de l'Hallue, a donc été également sélectionné. En plus de ce choix, les piézomètres d'Autheux, Cardonnette, Essertaux, Fricourt, Hancourt, Lignières-Châtelain, Mareuil-Caubert, Moreuil, Neuville, Nurlu, Omiécourt et Tilloloy, ainsi que Bonneuil les Eaux (dans l'Oise) ont été étudiés. L'étude corrélatoire a montré que certains piézomètres doivent être écartés.

### **Choix des pluviomètres :**

Grâce à l'appui de la DREAL Picardie nous avons pu disposer des données de précipitation de Météo-France pour les sites d'Amiens Glisy, d'Epehy, de Meaulte et de Rouvroy.

### **Choix des débits ou des hauteurs d'eau**

Les réseaux de neurones sont des outils statistiques ; ils peuvent être appliqués à des relations pluie/hauteur d'eau tout autant que pluie/débit. Les données fournies par la DREAL sont les débits ou les hauteurs d'eau. Il s'agit des mesures sur la Somme et, d'amont en aval, à Ham (débits et hauteurs), à Eclusiers-Vaux (débits), à Bray (débits), Lamotte-Brebière (débit) et Abbeville (débit) ; sur l'Hallue à Bavelincourt (débit et hauteur) et Querrieu (débit et hauteur), sur l'Avre à Moreuil (débit et hauteur).

Dans un premier temps le DREAL Picardie nous a indiqué sa priorité de travail sur le site d'Abbeville sur lequel se concentrent les enjeux, pour une modélisation en débit.

### 3.4.2 Démarche suivie

#### Validation des mesures

La démarche a consisté dans un premier temps à valider les données. Les mesures issues de certains appareils présentaient des lacunes nombreuses et ont été abandonnées. Le pas de temps des différentes mesures a nécessité un traitement préalable : les piézomètres avaient des mesures mensuelles dans les années 80 et 90, hebdomadaires entre 1998 et 2003, puis journalières entre 2006 et 2008. Les débits sont mesurés quotidiennement ainsi que les pluies. Le travail a consisté à faire une étude à pas de temps journalier, ce qui a limité la période d'utilisation des piézomètres à la période récente. Les valeurs manquantes des niveaux piézométriques ont été calculées par interpolation linéaire entre mesures connues. Enfin, pour que l'étude puisse être orientée pour les crues, elle a pris en compte la période antérieure à 2001 (pas de crues dans les périodes récentes).

#### Analyse corrélatoire simple

Le calcul des corrélations simples a été appliqué aux débits et aux mesures de piézomètres. Pour le débit mesuré à Abbeville, nous obtenons la courbe de la Figure 9 qui montre clairement la variation annuelle (décalage de 365 jours). Sur ce diagramme on note que le débit mesuré à Abbeville est corrélé avec lui-même sur une durée de 120 jours.

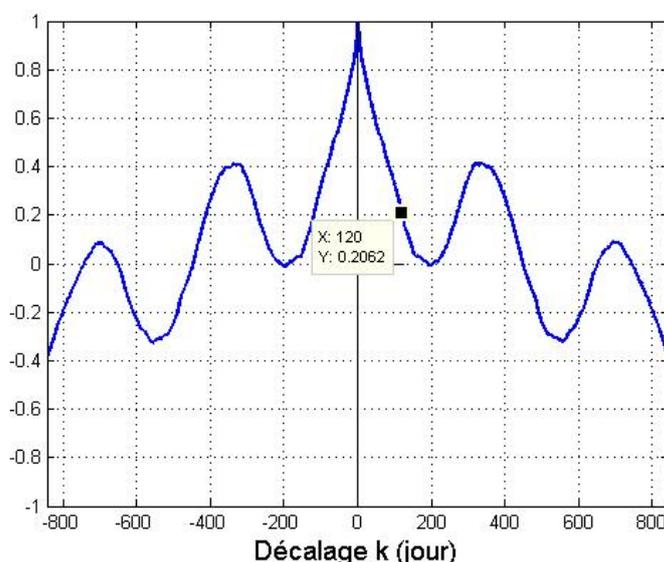


Figure 9 : Corrélogramme simple du débit journalier à Abbeville (Ben Nessib, 2013)

Ce même type de calcul a été appliqué aux piézomètres. A titre d'exemple, la Figure 10 présente l'auto-corrélogramme du piézomètre de Lignièrès-Châtelain. Là aussi, la variation annuelle est clairement visible.

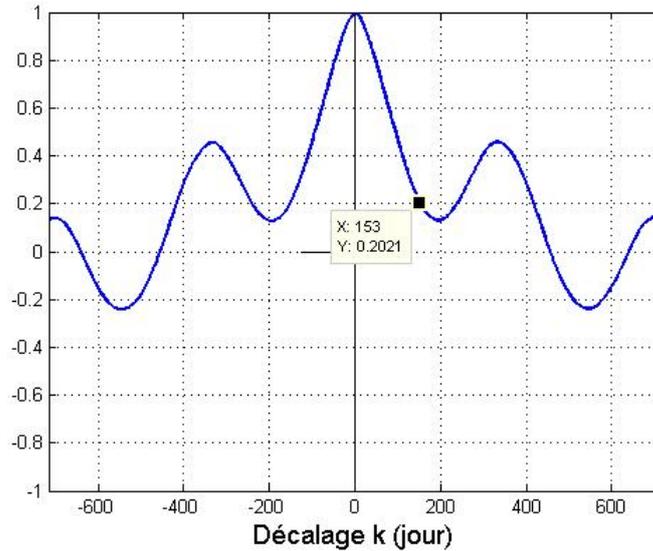


Figure 10 : Corrélogramme simple de la piézométrie journalière à Lignières-Châtelain

### **Analyse corrélation croisée**

Des corrélations croisées ont ensuite été réalisées entre niveaux piézométriques et débits. Des résultats très variés ont été obtenus ; de très bonnes corrélations croisées sont observées entre le piézomètre de Senlis-le-Sec et les débits mesurés sur l'Hallue à Bavelincourt (corrélation maximale de 0,97) et à Querrieu (valeur maximale de 0,94). Ceci montre que le niveau piézométrique de Senlis-le-Sec est fortement lié au débit de la rivière proche. Ce même résultat est confirmé dans (Amraoui, 2012). La corrélation croisée Senlis le Sec/Abbeville montre une valeur maximale de 0,83. Les corrélations croisées montrent de très faibles valeurs de corrélations pour les mesures de débits à Eclusiers-Vaux quel que soit le piézomètre utilisé pour le calcul.

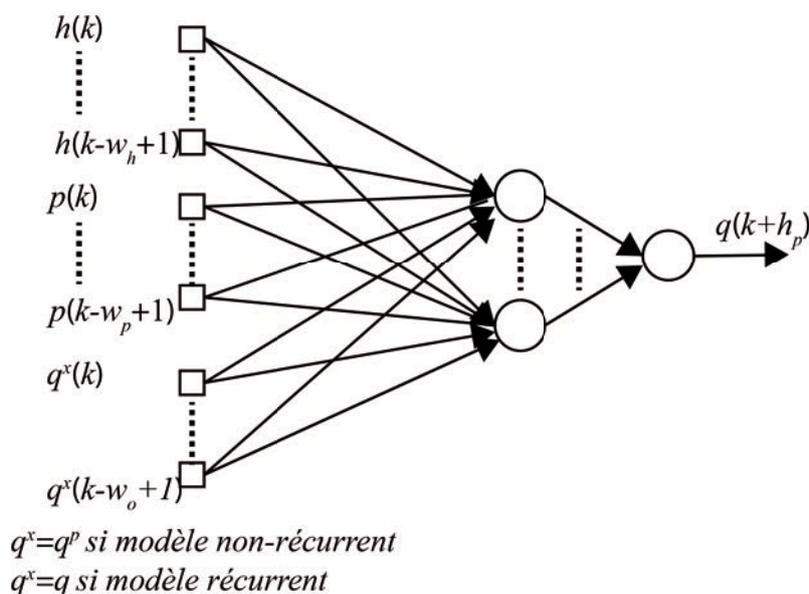
	Bavelincourt	Bray-Sur-Somme	Eclusier-Vaux	Ham
Senlis-Le-Sec	0,98	0,72	0,01	0,63
Autheux (Carouge)	0,94	0,49	0,05	0,41
Fricamps	0,60	0,18		0,23
Hancourt	0,50	0,19	0,08	0,13
Lignière	0,84	0,40	0,11	0,40
Neuville	0,94	0,29	0,29	0,31
Nurlu	0,53	0,26	0,09	0,29
Omiécourt	0,66	0,30	0,33	0,38
Bonneuil-Les-Eaux	0,77	0,26	0,08	0,24
Tilloloy	0,12	0,07	0,25	0,55

	Lamotte-Brebière	Moreuil	Querrieu	Bonnay	Abbeville
Senlis-Le-Sec	0,75	-0,60	0,94	0,85	0,83
Autheux (Carouge)	0,63	-0,50	0,92	0,83	0,73
Fricamps	0,33	-0,23	0,60	0,50	0,31
Hancourt	0,32	-0,12	0,51	0,49	0,39
Lignière	0,61	-0,46	0,93	0,84	0,71
Neuvillette	0,53	-0,52	0,86	0,78	0,62
Nurlu	0,53	-0,35	0,54	0,59	0,57
Omiécourt	0,58	-0,52	0,72	0,75	0,73
Bonneuil-Les-Eaux	0,39	-0,40	0,70	0,64	0,45
Tilloloy	0,20	-0,29	0,15	0,07	0,28

Tableau 7 : Maximum des corrélations croisées entre niveaux piézométriques (en lignes) et débits (en colonnes) à différentes stations limnimétriques

### 3.5 PREVISION PAR RESEAU DE NEURONES

Partant de ces prétraitements et analyses corrélatoires préliminaires, des modélisations ont été effectuées selon la méthodologie décrite dans le rapport de la tâche 2 (Johannet, 2013). Les performances des deux types de modèles : récurrent et non récurrent ont été évaluées ; Ils ont comme variables d'entrées : le niveau piézométrique journalier à Senlis-le-Sec  $h(k)$ , la pluie journalière  $p(k)$  et le débit journalier à Abbeville  $q(k)$ , et comme variable de sortie le débit à Abbeville *prévu* à l'horizon  $h_p$  :  $q(k+h_p)$ . A également été évaluée la variable décrivant la valeur journalière de la hauteur de la marée.



Les deux modèles : récurrents et non-récurrent ont été conçus, ils implémentent la fonction suivante.

$$q(k+h_p) = \varphi_{RN} \left( q^x(k), \dots, q^x(k-r+1), h(k), \dots, h(k-w_h+1), p(k), \dots, p(k-w_p+1), \mathbf{w} \right)$$

On a :  $w_h$  la fenêtre temporelle des hauteurs piézométriques,  $w_p$  la fenêtre temporelle des précipitations à Abbeville,  $r$  la fenêtre temporelle des débits à Abbeville qui représente l'ordre du modèle lorsque ce dernier est récurrent,  $\mathbf{w}$  est le vecteur des paramètres du modèle.

Plusieurs horizons de prévision ont été investigués : 1 jour, 4 jours et 7 jours en utilisant les pluies observées jusqu'à 3 jours dans le futur (on considère a posteriori que l'on peut alimenter le modèle avec des précipitations égales aux valeurs effectivement mesurées).

La sélection de la complexité a conduit aux modèles suivants :

	Nombre de neurones cachés	Fenêtre temporelle de pluies à Abbeville	Fenêtre temporelle de pluies futures à Abbeville	Fenêtre temporelle de piézométrie à Senlis-le-Sec	Fenêtre temporelle de débit précédent à Abbeville	Coefficient de marée
Modèle récurrent	4	5	3	3	3	
Modèle non récurrent 1	3	5	3	1	6	
Modèle non récurrent 2	3	3	2	1	1	2

**Tableau 8 : Dimensionnement des modèles de prédiction**

Les performances de ces modèles sont présentées dans le tableau suivant grâce aux critères de qualité présentés dans le rapport T2 (Johannet, 2013). On peut noter que si les critères de Nash sont excellents (supérieurs à 0,90, les critères de persistance sont en revanche très mauvais, souvent négatifs ce qui signifie que les modèles ne peuvent effectuer une meilleure prévision que la prévision naïve. En l'état actuel des travaux ils sont donc inutiles en prévision.

	$h_p$	Test	Apprentissage
Critère de Nash	1	0,993	0,992
	4	0,974	0,973
	7	0,910	0,957
Critère de persistance	1	-0,094	0,189
	4	-0,033	0,309
	7	-0,067	0,293

**Tableau 9 : Critères de performance du meilleur modèle prédictif : le modèle non-récurrent 1**

On remarque, sur le Tableau 9, que ni l'apprentissage, ni le test ne sont satisfaisants ce qui signifie que les bonnes variables n'ont pas été appliquées au réseau. En particulier il y a une contradiction entre la taille de la fenêtre temporelle sélectionnée pour l'information de piézométrie (seulement 1 ou 3 jours) et la dynamique de cette dernière qui est très lente comme indiqué sur la Figure 9. Ceci est dû au fait que le réseau ne peut être alimenté par des fenêtres de trop grande longueur pour ne pas

augmenter considérablement le nombre de variables. Un pré-traitement doit donc être recherché spécifiquement pour introduire ce type d'information au modèle. Des premiers essais ont été effectués par une décomposition multirésolution.

### 3.6 CONCLUSION

Suivant la méthodologie proposée pour réaliser des modèles de prévision hydrologique de crues rapides il a été possible de prévoir des débits à partir de données piézométriques, de mesures de pluie et des mesures antérieures (quelques jours) de débit. Néanmoins compte-tenu de la dynamique très lente du processus il est apparu que le type de variables utilisées ne devait pas être satisfaisant car les résultats de prévision n'apportent aucune anticipation par rapport à la prévision naïve même si les résultats en modélisation sont en revanche bon (critères de Nash excellents). Une étude complémentaire doit donc être réalisée dans ce sens avant de renouveler les efforts de modélisation.

## 4. GENERALISATION AU BASSIN VERSANT DE LA LIANE

### 4.1 PRESENTATION DU BASSIN VERSANT

La Liane est un fleuve côtier de 244 km<sup>2</sup> de bassin versant, orienté dans le sens Est-Ouest. Elle s'écoule sur un substrat peu perméable bordé d'une ceinture calcaire. Elle prend naissance à 101 mètres d'altitude à Quesques et est entourée par un escarpement abrupt dépassant 200 mètres d'altitude. Sa pente moyenne est de 2,8 ‰ atteignant 6 ‰ pour l'amont et est l'une des plus importantes des cours d'eau de la région. Le réseau hydrographique est relativement dense (**Error! Reference source not found.**).

De point de vue de l'occupation des sols, l'amont est essentiellement couvert par des activités rurales (zones agricoles et de forêts). A l'aval, la Liane traverse des communes plus urbanisées (Saint-Etienne-au-Mont, Saint-Léonard) et termine son cours en traversant Boulogne-Sur-Mer. C'est à son embouchure que les enjeux en terme de risque d'inondation sont les plus forts.

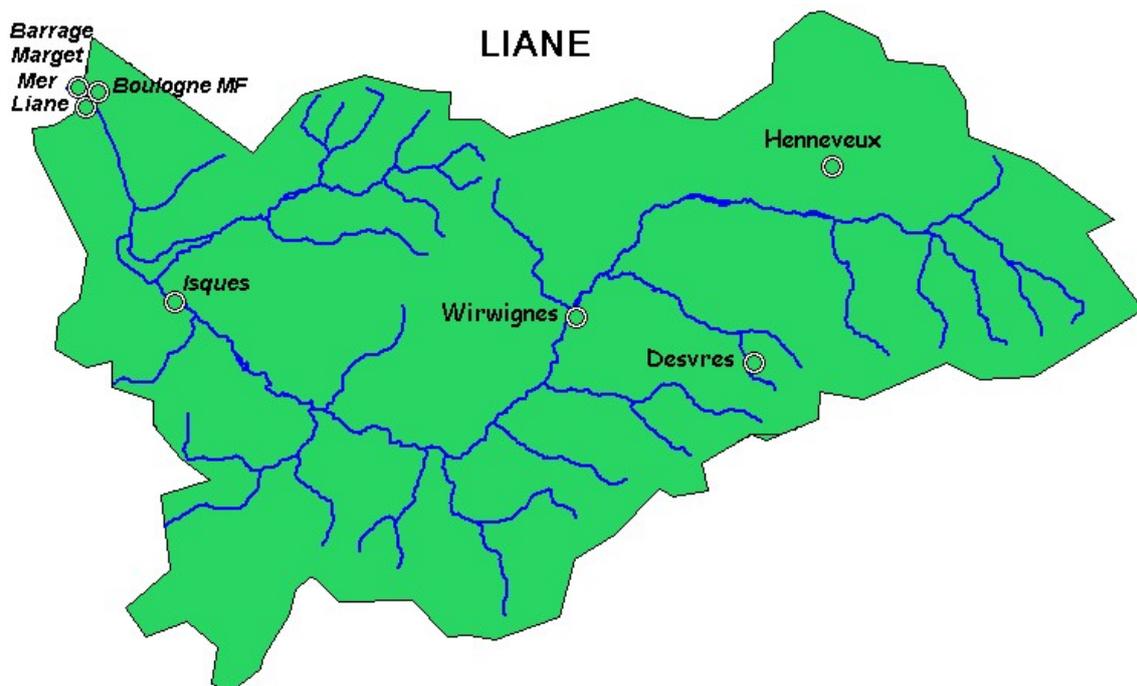


Figure 11 : Bassin versant de la Liane (source Service de Prévision des crues Artois-Picardie)

## 4.2 BASE DE DONNEES HYDROMETEOROLOGIQUE

Les crues de la Liane sont déjà citées dans l'ouvrage de Maurice CHAMPION publié en 1863 « Les inondations en France du VI<sup>ème</sup> siècle à nos jours », et ce malgré la petite taille du fleuve. On se réfère généralement à deux crues majeures : celle de 1839 et celle de 1857, cette dernière intervenant après une semaine de « pluie torrentielle ». Plus récemment, les crues du 29 octobre 1981, 1<sup>er</sup> novembre 1998, 21 novembre 2000 ont été remarquée, la dernière crue majeure est celle de novembre 2012.

L'historique des cotes les plus importantes observées à la station de Wirwignes de 1971 à nos jours est reporté dans le Tableau 10.

<i>Date de l'évènement</i>	<i>Cote à Wirwignes</i>	<i>Période de retour associée</i>
28/10/1981	4,18 m	environ 10 ans
01/11/1998	4,32 m	environ 12 ans
21/11/2000	4,16 m	environ 10 ans
02/11/2012	4,36 m	environ 12 ans

Tableau 10 : Historique des crues majeures de la Liane de ces dernières années

## 4.3 SYNTHESE DU PREDICTEUR

Initialement le bassin-versant de la Liane n'était pas compris dans les bassins versants à étudier au moyen de l'apprentissage statistique. Néanmoins, fait intéressant, c'est un bassin pour lequel les prévisions opérationnelles de crues sont déjà effectuées par un modèle à réseaux de neurones. A cette fin, le Service de Prévision des Crues Artois-Picardie utilise un fichier excel dans lequel les formules du modèle ont été introduites et qui calcule ses sorties par ce moyen très simple. Les utilisateurs peuvent introduire leurs scénarios de prévision de pluies « à la main » dans ce fichier afin de réaliser leurs prévisions des crues. Lors de la réunion du projet FLASH de mai 2013 qui se déroulait à Amiens, il est donc apparu intéressant de déployer le démonstrateur déjà réalisé sur les Gardon à ce bassin versant, de

fonctionnement plus proche des crues rapides que celui de la Somme, même si ces dernières sont beaucoup moins rapides et beaucoup moins intenses que celles des Gardons. On note un facteur de 8 sur les temps de montée (3h/24h) et un facteur 60 sur les débit de pointe ( $60\text{m}^3/\text{s}$  /  $3500\text{ m}^3/\text{s}$ ).

On note également une autre spécificité sur les pratiques du SPC Artois Picardie : les prévisions effectuées concernent le niveau maximal sur un certain horizon de prévision, ce niveau d'eau est noté  $q_{max}^p(k+h_p)$  pour la variable observée sur le processus et  $q_{max}(k+h_p)$  pour la variable estimée. Quatre horizons de prévision sont recherchés : 3h, 6h, 12h et 24h, autant de modèles seront réalisés. Les prévisions sont effectuées en supposant que l'on dispose de prévisions de pluies dites « parfaites » c'est à dire égales aux pluies mesurées.

L'architecture générique des modèles utilisés mixe les modèles récurrents et non récurrents ; de plus comme une simulation sur l'année entière est réalisée, il est nécessaire d'introduire une variable représentant l'évapotranspiration, c'est à dire la quantité d'eau évaporée, consommée et transpirée par la végétation. Cette grandeur peut être appliquée au modèle neuronal par une simple courbe en cloche qui représente les variations de cette grandeur au cours de l'année (Johannet, 2010). Les mesures de trois pluviomètres sont disponibles sur le bassin : Desvres, Henneveux et Wirwignes qui est l'exutoire considéré ici.

La base de donnée comporte les mesures journalières de 1998 à mi 2013. L'ensemble d'arrêt est constitué de l'année 2011, l'ensemble de test de l'année 2012 qui comporte la crue la plus intense de ces 30 dernières années, et l'ensemble d'apprentissage des autres années. L'échantillonnage est journalier.

Comme à chaque conception de modèles les données ont dû être validées, les lacunes complétées et la sélection des variables réalisée comme expliqué dans le rapport de la tâche  $T_2$  (Johannet, 2013).

Les prévisions parfaites de pluies sont également intégrées aux variables pluviométriques pour les trois pluviomètres ; se rajoutent donc autant de variables d'entrées pour chacun d'eux (2 variables pour l'horizon 3h, 5 pour l'horizon 6h, etc ...).

Le type de modèle utilisé est donc représenté en Figure 12. On note qu'une architecture particulière est adoptée pour parvenir à prendre en compte des fenêtres temporelles importantes sur les variables pluviométriques (pluies futures et passée) tout en diminuant le nombre de paramètres libres du réseau : une couche de neurones linéaires n'effectuant que la somme pondérée des entrées d'un même pluviomètre est ajoutée.

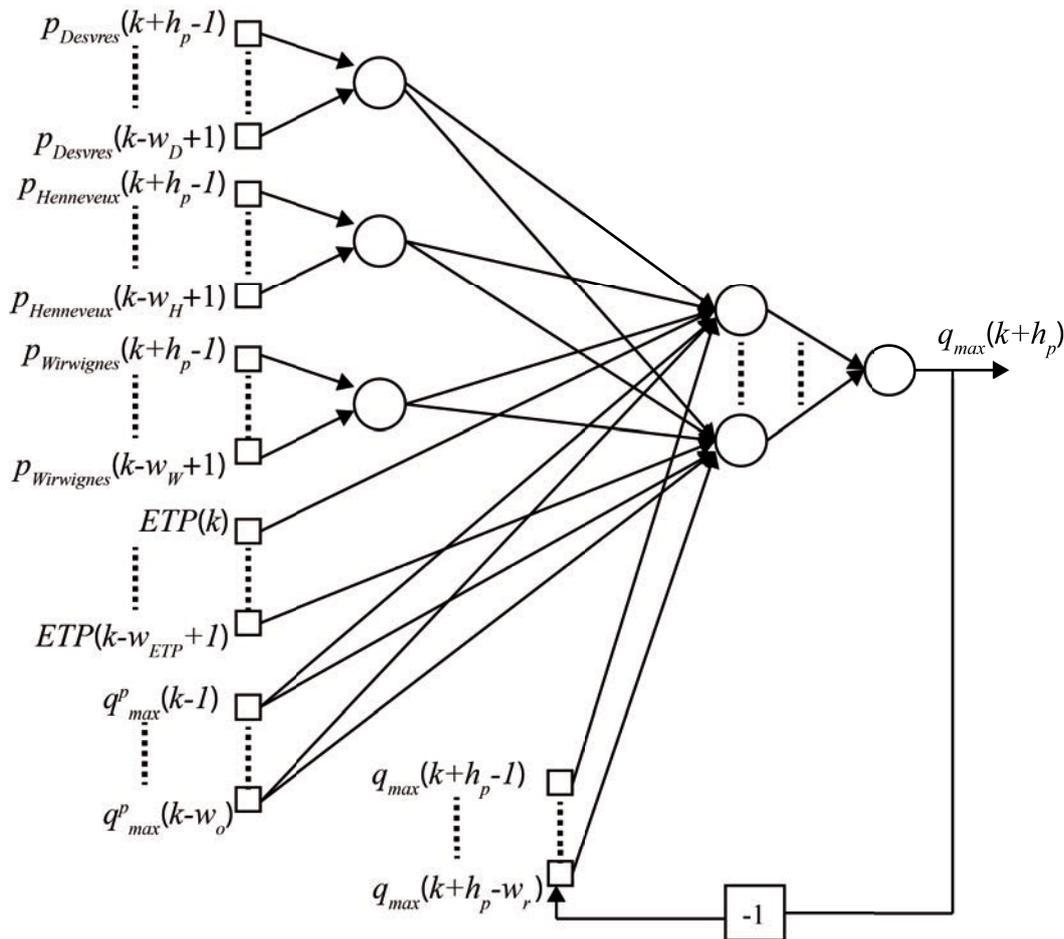


Figure 12 : Modèle générique pour la prévision des crues sur la Liane

Le Tableau 11 regroupe les définitions des structures précises des modèles.

	$w_D$	$w_H$	$w_W$	$w_{ETP}$	$w_o$	$w_r$	$N_c$
$H_p=3h$	11	8	11	2	1	6	4
$H_p=6h$	2	12	12	2	1	7	4
$H_p=12h$	4	4	3	1	1	4	6
$H_p=24h$	5	5	1	2	1	2	7

Tableau 11 : Définitions des structure précises des modèles pour la prévision des crues sur la Liane

On peut noter sur le Tableau 11 qu'il n'y a pas une grande logique dans la définition des fenêtres temporelles, en revanche l'ordre de la part récurrente du modèle tend à diminuer avec l'horizon de prévision tandis que le nombre de neurones cachés diminue avec ce dernier.

#### 4.4 PERFORMANCES SUR L'ANNEE 2012

Rappelons que l'année 2012 est l'année de test du modèle et qu'elle comporte l'événement le plus intense des 30 dernières années. Si l'on trace l'hydrogramme annuel pour la prévision à 24h, les deux courbes, mesurée et estimée sont très proches (Figure 13). Les critères de Nash (Tableau 12) sont excellents autant pour l'année complète que pour l'événement très intense de novembre 2012. La persistance ne peut être définie pour ce type de prévision.

	3h	6h	12h	24h
Nash annuel	0,99	0,98	0,96	0,94
Nash événement	0,98	0,97	0,91	0,89

Tableau 12 : Performance des prédicteurs sur la Liane

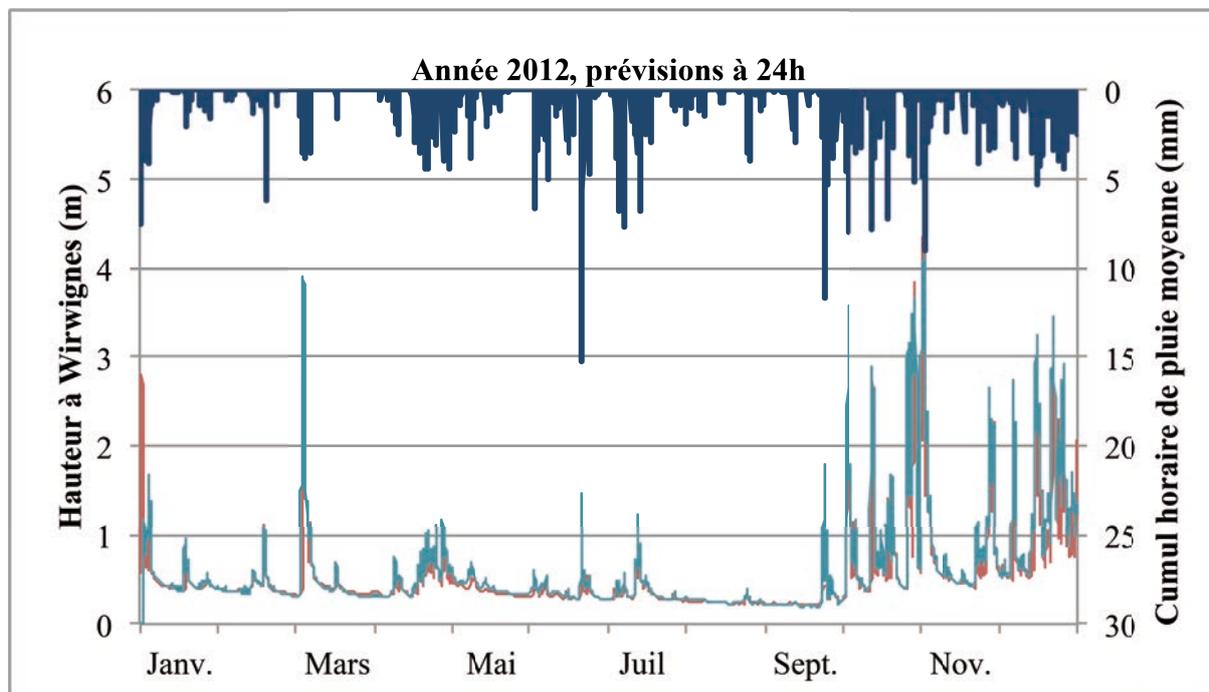
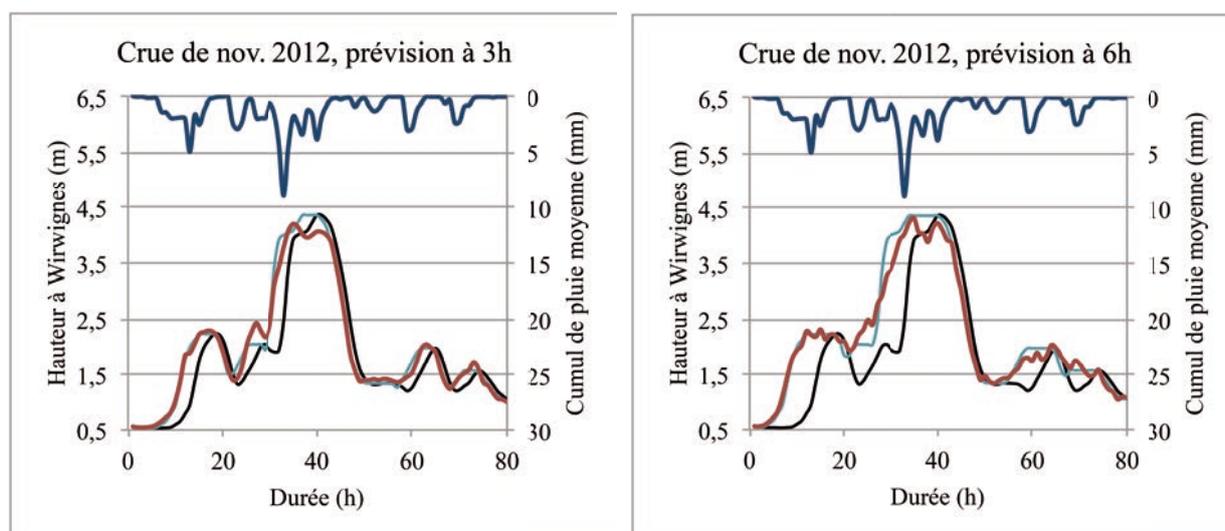


Figure 13 : Prévion à 24 du niveau maximal à Wirwignes 24h à l'avance. En bleu clair la grandeur estimée, en rouge la grandeur à prédire.



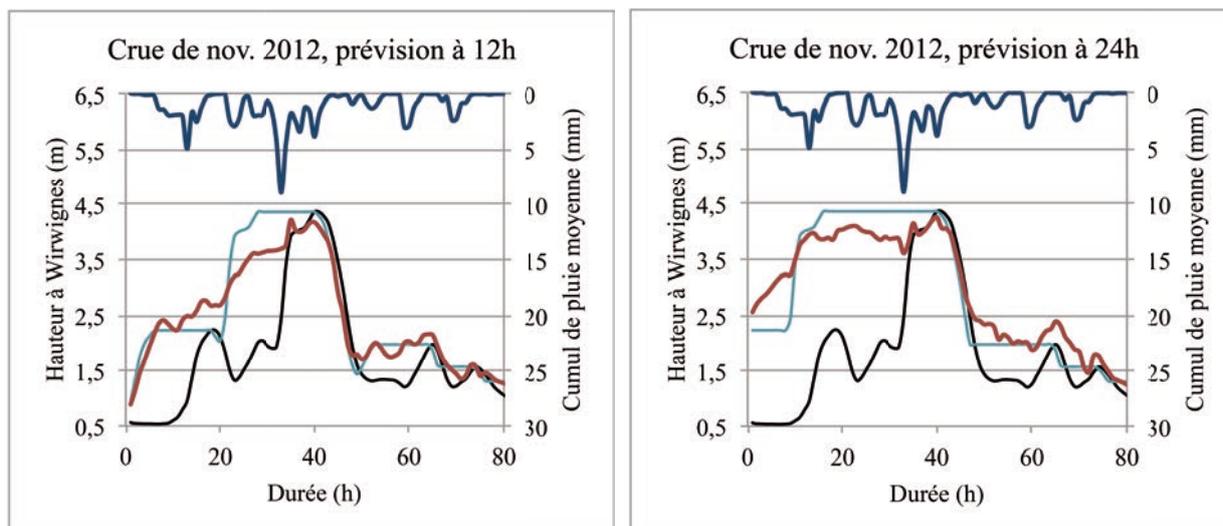


Figure 14 : Le tracé des hydrogrammes pour les quatre prévisions de l'événement de 2012 confirme les bonnes valeurs obtenues pour les critères de performance, la crue peut être anticipée, grâce aux prévisions de pluies, sur un horizon de 24. En bleu clair la grandeur à prédire, en rouge la grandeur estimée, en noir la hauteur observée.

#### 4.5 CONCLUSION

En conclusion, la généralisation des méthodes définies sur les crues cévenoles aux crues de la Liane a abouti à de très bons modèles de prévision. Il doit être noté que ces modèles ont permis d'investiguer de nouvelles pistes d'amélioration des modèles prédictifs : (i) la définition du débit maximum sur l'horizon d'anticipation souhaité, (ii) l'intégration des mesures de pluies futures qui permet d'utiliser le modèle avec plusieurs scénarios de pluies futures, (iii) l'utilisation conjointe des modèles récurrents et non récurrents, (iv) l'introduction d'une couche cachée de neurones linéaires pour diminuer la complexité du modèle tout en conservant des fenêtres temporelles de précipitation assez grandes.

### 5. CONCLUSION GENERALE

La tâche 6 du projet FLASH : Généralisations spatiales et temporelles, avait pour objectif d'évaluer le caractère générique des modèles de prévision développés dans le projet, en les déployant sur des bassins versants différents, soit par leur taille, soit par leur fonctionnement hydrologique, soit par les deux aspects. Une partie de cette généralisation a été abordée dans la tâche 2 lorsque la méthodologie de conception de modèle de prévision par réseaux de neurones, développée sur le bassin d'Anduze, a été transposée au bassin de Remoulins, englobant le précédent. Mais si utiliser une même méthodologie de conception est certes un gain dans la conception de modèles, ceci n'a pas la même simplicité et universalité que d'appliquer le même modèle à plusieurs bassins versants de comportements de même catégorie (crues cévenoles), mais cependant très différents. Cette ambition était celle de la tâche 6 qui a été relevée avec succès par le travail de thèse de Guillaume Artigue. Ceci est un des grands succès de ce projet car il n'était pas acquis, voire peu envisageable, de parvenir à réaliser un modèle statistique de fonctionnement de bassin versant, sans base de données relative à ce bassin.

Ce défi a été relevé avec succès et validé sur 15 bassins versants (de 4 à 545 km<sup>2</sup>) et 112 événements de la bordure cévenole ; mais la méthode doit gagner en fiabilité

pour les très petits bassins pour lesquels l'information pluviométrique est fournie par des RADARS. Or les enseignements de la tâche 2 qui visait à étudier les différences entre les estimations de pluie issues des RADARS ou issues des pluviomètres a mis en évidence une différence d'un facteur 2 entre les deux sources de données (Visserot, 2012). Il conviendrait donc de corriger les précipitations par ce facteur ou de réaliser un autre modèle spécifique pour les données issues des RADARS. C'est ce que nous allons tenter de réaliser comme développement futur.

Au delà de la généralisation spatiale et temporelle nous avons proposé d'effectuer une transposition des méthodologies de conception de modèle de prévision à un bassin versant de comportement très différent : la Somme, connue pour ses crues de nappe. Si l'étude a été riche d'enseignements, elle n'a en revanche pas débouché sur un modèle exploitable. Les raisons sont doubles et sont liées à la complexité même de ce bassin cumulant un comportement de la nappe de constante de temps de plusieurs années, avec une influence maritime et une influence plus rapide des précipitations. Le modèle neuronal choisi ne peut prendre en compte ces différentes composantes temporelles sans prétraitements préliminaires, un peu comme cela est réalisé dans la fonction de classification : il faut extraire des caractéristiques préalables. Cette recherche a déjà commencé avec la prise en compte d'une décomposition multi-résolution par ondelettes discrètes.

Dans cet état d'esprit, un filtrage très simple a été mis en œuvre sur les variables d'un troisième type de bassin, celui de la Liane, afin de prendre en compte les précipitations sur une fenêtre temporelle de plus grand historique. Rajouté au projet afin de développer un nouveau prototype opérationnel de prévision de crue, le bassin de la Liane est intermédiaire entre les précédents. Les prévisions effectuées en suivant la méthodologie de conception ont été excellentes.

La tâche 6 est donc une des tâches les plus fructueuses du projet FLASH tant en termes de débouchés opérationnels qu'en termes de démonstration de la capacité des modèles neuronaux à effectuer des modélisations simplifiées de systèmes complexes, ce qui était un des objets de l'appel à projet SYSCOMM (Systèmes complexes et modélisations mathématiques).

## 6. BIBLIOGRAPHIE

(AMEVA, 2013) site internet de l'AMEVA, <http://www.ameva.org/?q=content/le-bassin-versant/eaux-superficielles-et-souterraines>, *dernier accès le 20 décembre 2013*.

(Amraoui , 2012) Amraoui N., Thiéry D., Wuilleumier A., 2012, *Appui au SCHAPI-module 2 – Suivi du site expérimental de Warloy-Baillon et prévision des crues de la Somme*. Rapport final. BRGM/RP 62 062-FR, 86 p., 71 Fig., 1 tabl., 1 ann.

(Artigue, 2011) Artigue, G., Johannet, A., Borrell, V., Pistre, S., 2011, *Flash floods forecasting without rainfalls forecasts by recurrent neural networks. Case study on the Mialet basin (Southern France)*, Proceedings of the third world congress on nature and biologically inspired computing, NABIC 2011, Salamanca, Spain, pp. 303-310.

(Artigue, 2012a) Artigue, G., Johannet, A., Borrell, V., Pistre, S., 2012, *Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon de Mialet basin (southern France)*, Natural Hazards and Earth System Sciences 12, pp. 3307-3324, doi : 10.5194/nhess-12-3307-2012.

(Artigue, 2012b) Artigue G., *Prévision des crues éclair par réseaux de neurones : généralisation aux bassins non jaugés*, thèse de doctorat, Université Montpellier II, Ecole doctorale SIBAGHE, 2012.

(Ben Nessib, 2013) Ben Nessib D., *Prévision des crues de nappes à Abbeville par réseaux de neurones*, mémoire de PFE de l'Ecole Supérieure de la Statistique et de l'Analyse de l'Information de Tunis.

(Bornancin-Plantier, 2013) Bornancin-Plantier A., thèse de doctorat, *Conception de modèles de prévision des crues éclair par apprentissage artificiel*, Université Pierre et Marie Curie, Ecole Doctorale EDITE, 2013.

(Collectif, 1994), *Débits des cours d'eau : 10 ans d'observations : synthèse sur 200 stations de mesures en France : 1984-1993*, Paris, 186 p.

(Johannet, 2010) Johannet, A., *Artificial Neural Networks*. In *Environmental hydraulics 2: Mathematical Models*, J.M.Tanguy, ed.

(Johannet, 2013) Johannet A. *et al*, *Rapport du projet ANR FLASH Tâche 02 : Modélisation de la chaîne hydrométéorologique*.

(Mardhel, 2001) Mardhel V., Negrel Ph., Pointet T., *Première analyse des composantes des écoulements souterrains du bassin versant de la Somme en période de crues*. BRGM/ RP 51030 FR, 26 p., 26 Fig., 1 tabl.

(Roux, 1978) *Atlas hydrogéologique de la Somme*. BRGM, 98 p.

(Vinet, 2010) Vinet F., *Le risque inondation ; diagnostic et gestion*, Edition Tec et Doc, Collection SRD, Paris, 318 p.

## LISTE DES FIGURES

<i>Figure 1 : Méthode de généralisation (Artigue, 2012)</i>	3
<i>Figure 2 : Carte topographique et hydrographique simplifiée du bassin versant du Gardon de Mialet à Mialet. Altitudes d'après ASTER GDEM, un produit de METII et NASA ; Hydrographie d'après IGN BD CARTHAGE (Artigue, 2012)</i>	4
<i>Figure 3 : Modèle récurrent utilisé dans la prévision des crues du Gardon de Mialet (Artigue, 2012)</i>	5
<i>Figure 4 : Représentation du débit spécifique de pointe en fonction de la surface des bassins sur lesquels il a été observé (points rouges) et courbe de régression de type puissance (courbe bleue).</i>	8
<i>Figure 5 : Bassin versant de la Somme (AMEVA, 2013)</i>	12
<i>Figure 6 : Coupe Nord Sud de la vallée de la Somme (extrait de (Roux, 1978), modifié)</i>	13
<i>Figure 7 : Le canal de la Somme comprend 25 écluses (AMEVA, 2013)</i>	14
<i>Figure 8 : Localisation des piézomètres dans le département de la Somme (ADES IGN 2005, BD CARTO)</i>	16
<i>Figure 9 : Corrélogramme simple du débit journalier à Abbeville (Ben Nessib, 2013)</i>	17
<i>Figure 10 : Corrélogramme simple de la piézométrie journalière à Lignières-Châtelain</i>	18
<i>Figure 11 : Bassin versant de la Liane (source Service de Prévision des crues Artois-Picardie)</i>	22
<i>Figure 12 : Modèle générique pour la prévision des crues sur la Liane</i>	24
<i>Figure 13 : Prévision à 24 du niveau maximal à Wirwignes 24h à l'avance. En bleu clair la grandeur estimée, en rouge la grandeur à prédire.</i>	25
<i>Figure 14 : Le tracé des hydrogrammes pour les quatre prévisions de l'événement de 2012 confirme les bonnes valeurs obtenues pour les critères de performance, la crue peut être anticipée, grâce aux prévisions de pluies, sur un horizon de 24. En bleu clair la grandeur à prédire, en rouge la grandeur estimée, en noir la hauteur observée.</i>	26

## LISTE DES TABLEAUX

<i>Tableau 1 : Architectures optimales pour les deux types de modèles utilisés. Les valeurs sont indiquées en nombre de pas de temps (soit en 0,5h).</i>	5
<i>Tableau 2 : Quelques caractéristiques des évènements choisis pour le test. Pics : nombre de pics ; Qx : débit de pointe ; Qm : débit moyen ; RR : cumul total de pluie ; Ix : intensité pluvieuse maximale (Artigue, 2012)</i>	6
<i>Tableau 3 : Récapitulatif des résultats obtenus avec le modèle récurrent (Evt. correspond au numéro de l'évènement, l à l'horizon de prévision, Vp à la vigilance prévue par le modèle, Vo à la vigilance correspondant au débit observé à k+l et D au décalage temporel) (Artigue, 2012)</i>	6
<i>Tableau 4 : Sélection et description des bassins retenus</i>	9
<i>Tableau 5 : Coefficient correctif correspondant à chaque bassin visé par l'étude</i>	10
<i>Tableau 6 : SPPD moyens et médians et taux d'atteinte de SPPD compris entre 70 et 130% pour différents types d'évènements et de bassins, décrits dans le tableau</i>	11
<i>Tableau 7 : Maximum des corrélations croisées entre niveaux piézométriques (en lignes) et débits (en colonnes) à différentes stations limnimétriques</i>	19
<i>Tableau 8 : Dimensionnement des modèles de prédiction</i>	20
<i>Tableau 9 : Critères de performance du meilleur modèle prédictif : le modèle non-récurrent 1</i>	20
<i>Tableau 10 : Historique des crues majeures de la Liane de ces dernières années</i>	22
<i>Tableau 11 : Définitions des structure précises des modèles pour la prévision des crues sur la Liane</i>	24
<i>Tableau 12 : Performance des prédicteurs sur la Liane</i>	25



<http://blog.espci.fr/flash/>

